

UNIVERSITY OF CALIFORNIA, SAN DIEGO

**Search for New Physics in Proton-Proton Collisions at 8 TeV Center of Mass  
Energy with a Final State of Same-Sign Dileptons and Jets**

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

in

Physics

by

Ryan Ward Kelley

Committee in charge:

Professor Avraham Yagil, Chair  
Professor Claudio Campagnari  
Professor Robert Continetti  
Professor Aneesh Manohar  
Professor Frank Würthwein

2013

Copyright  
Ryan Ward Kelley, 2013  
All rights reserved.

The dissertation of Ryan Ward Kelley is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

---

---

---

---

---

Chair

University of California, San Diego

2013

I dedicate this thesis to Jennifer, my beautiful wife, for her endless patience, love, and support. We did it – Swipe swipe!

*perfer et obdura; dolor hic tibi proderit olim.*

*Be patient and tough; some day this pain will be useful to you.*

—old latin saying

## TABLE OF CONTENTS

Signature Page . . . . .	iii
Dedication . . . . .	iv
Epigraph . . . . .	v
Table of Contents . . . . .	vi
List of Figures . . . . .	ix
List of Tables . . . . .	xi
Acknowledgements . . . . .	xiii
Vita and Publications . . . . .	xiv
Abstract of the Dissertation . . . . .	xv
Chapter 1      Introduction . . . . .	1
1.1     Standard Model of Particle Physics . . . . .	2
1.2     Proton-Proton Collisions . . . . .	7
1.3     Motivation for Same-Sign Dilepton Signature . . . . .	8
Chapter 2      The LHC Accelerator and CMS Experiment . . . . .	13
2.1     The Large Hadron Collider . . . . .	13
2.2     The Compact Muon Solenoid Detector . . . . .	15
2.2.1     The Superconducting Solenoid . . . . .	17
2.2.2     The Silicon Detector . . . . .	18
2.2.3     The Calorimeters . . . . .	21
2.2.4     The Muon Detectors . . . . .	25
2.3     Integrated Luminosity Calculation . . . . .	29
2.4     Data Acquisition and Triggering . . . . .	30
2.5     Particle Reconstruction and Identification . . . . .	32
2.5.1     Charged Particle Reconstruction . . . . .	33
2.5.2     Vertex Reconstruction . . . . .	34
2.5.3     Electron Reconstruction . . . . .	35
2.5.4     Muon Reconstruction . . . . .	37
2.5.5     Particle Flow Reconstruction . . . . .	38
2.5.6     Calculation of Missing Transverse Energy . . . . .	40
2.5.7     Jet Reconstruction . . . . .	40
2.5.8     b-Jet Identification . . . . .	41

Chapter 3	Same-Sign Dilepton Signature . . . . .	44
	3.1 Rare Standard Model Processes . . . . .	44
	3.1.1 Top Processes . . . . .	45
	3.1.2 Two Gauge Boson Processes . . . . .	46
	3.1.3 Triple Gauge Boson Processes . . . . .	46
	3.1.4 Higgs Boson Processes . . . . .	47
	3.1.5 Same-Sign W Boson Pair Production . . . . .	48
	3.2 Other Sources of Backgrounds . . . . .	49
	3.2.1 Fake Leptons . . . . .	49
	3.2.2 Charge Flips . . . . .	50
Chapter 4	Analysis Selections . . . . .	52
	4.1 Data Samples . . . . .	53
	4.1.1 Collision Data . . . . .	53
	4.1.2 Simulated Data . . . . .	53
	4.2 Event Selections . . . . .	54
	4.2.1 Trigger Selection . . . . .	57
	4.2.2 Muon Selection . . . . .	58
	4.2.3 Electron Selection . . . . .	60
	4.2.4 Lepton Pair Disambiguation . . . . .	62
	4.2.5 Jet Selection . . . . .	62
	4.2.6 B-Tagging Selection . . . . .	63
	4.2.7 Missing Transverse Energy Selection . . . . .	63
	4.2.8 Z and $\gamma^*$ veto . . . . .	64
	4.3 Search Regions . . . . .	64
Chapter 5	Background Estimation Methods . . . . .	68
	5.1 Data-Driven method to estimate the fake rate . . . . .	70
	5.1.1 The Fake Rate Method . . . . .	71
	5.1.2 Fakeable Object Definition . . . . .	72
	5.1.3 Fake Rate Datasets . . . . .	73
	5.1.4 Fake Rate Contamination from Prompt Leptons . . . . .	74
	5.1.5 Fake Rates for Electrons and Muons . . . . .	77
	5.1.6 Study of Fake Rate Dependencies . . . . .	78
	5.2 Estimation of the Charge Mis-measurement Rate . . . . .	85
	5.3 Rare SM background estimation . . . . .	88
Chapter 6	Selection Efficiencies . . . . .	90
	6.1 Lepton Efficiencies . . . . .	91
	6.1.1 Tag-and-Probe Method . . . . .	91
	6.1.2 Lepton Composition . . . . .	96
	6.2 Trigger Efficiencies . . . . .	99
	6.3 B-tagging Efficiency . . . . .	100

Chapter 7	Results . . . . .	102
7.1	High $p_T$ Results . . . . .	103
7.1.1	High $p_T$ Control Regions . . . . .	103
7.1.2	High $p_T$ Search Regions . . . . .	110
7.2	Low $p_T$ Results . . . . .	112
7.2.1	Low $p_T$ Control Regions . . . . .	112
7.2.2	Low $p_T$ Search Regions . . . . .	119
7.3	Interpretations . . . . .	122
7.3.1	Same-Sign Top . . . . .	123
7.3.2	Standard Model Four Top . . . . .	123
7.3.3	Supersymmetry . . . . .	123
Chapter 8	Summary and Conclusions . . . . .	132
Bibliography	. . . . .	134
Appendix A	Table of Fake Rates . . . . .	142
Appendix B	Details of Simulated Samples . . . . .	144
B.1	Rare MC Samples . . . . .	144
B.2	Standard Model Background MC Samples . . . . .	144
B.3	MC Samples used for SM 4-Top Production and SS Top Pair Production . . . . .	144
B.4	MC Samples used for Fake Rate Closure Tests . . . . .	145

## LIST OF FIGURES

Figure 1.1:	The list of fundamental Standard Model particles broken down by particle type . . . . .	3
Figure 1.2:	A summary of the allowed interactions between particle types in the Standard Model . . . . .	6
Figure 1.3:	Parton distribution functions for protons at $Q^2 = 10 \text{ GeV}^2$ and $Q^2 = 10000 \text{ GeV}^2$ . . . . .	9
Figure 1.4:	Cross sections for some Standard Model process vs center-of-mass energy . . . . .	10
Figure 2.1:	The Large Hadron complex seen underground on the outskirts of Geneva, Switzerland . . . . .	14
Figure 2.2:	Overview of the Compact Muon Solenoid detector . . . . .	15
Figure 2.3:	Layout of the tracking system projected into the $rz$ plane . . . . .	19
Figure 2.4:	Radiation length and number of layers crossed in the CMS tracker as a function of pseudorapidity. . . . .	21
Figure 2.5:	Muon transverse impact parameter resolution and transverse momentum resolution as a function of pseudorapidity . . . . .	22
Figure 2.6:	One quarter of the electromagnetic calorimeter system projected into the $y-z$ plane [1]. . . . .	23
Figure 2.7:	One quarter of the hadronic calorimeter system projected into the $y-z$ plane [2]. . . . .	24
Figure 2.8:	Overview of the CMS muon subdetector system . . . . .	26
Figure 2.9:	Muon momentum resolution for two different pseudorapidity regions . . . . .	29
Figure 2.10:	Summary of the amount of data delivered by the LHC and collected by CMS in 2012 . . . . .	31
Figure 2.11:	Pictorial representation of the data acquisition system and the flow of data from the CMS detector through both trigger levels [3]. . . . .	32
Figure 2.12:	Depiction of the different portions of the electron reconstruction algorithm . . . . .	37
Figure 2.13:	Combined Secondary Vertex algorithm discriminant for data and simulation. . . . .	43
Figure 3.1:	Feynman diagram for $t\bar{t}W$ . . . . .	45
Figure 3.2:	Feynman diagram for $WZ$ . . . . .	46
Figure 3.3:	Feynman diagram for $WZ$ . . . . .	47
Figure 3.4:	Feynman diagram for $WZ$ . . . . .	48
Figure 3.5:	Feynman diagram for $qqW^\pm W^\pm$ . . . . .	49
Figure 3.6:	Feynman diagram for fakes from $t\bar{t}$ . . . . .	50
Figure 5.1:	The measured fake rate vs $p_T$ before and after the electroweak corrections . . . . .	77

Figure 5.2:	Muon Fake Rate vs $p_T$ , $ \eta $ , and number of vertices . . . . .	79
Figure 5.3:	Electron Fake Rate vs $p_T$ , $ \eta $ , and number of vertices . . . . .	80
Figure 5.4:	Muon Fake Rate vs $p_T$ , $ \eta $ , and number of vertices for different away jet $p_T$ requirements . . . . .	81
Figure 5.5:	Electron Fake Rate vs $p_T$ , $ \eta $ , and number of vertices for different away jet $p_T$ requirements (with online isolation requirement) . . . . .	82
Figure 5.6:	Electron Fake Rate vs $p_T$ , $ \eta $ , and number of vertices for different away jet $p_T$ requirements (no online isolation requirement) . . . . .	83
Figure 5.7:	The $p_T$ distribution for electrons that pass the analysis selections . . . . .	86
Figure 5.8:	The $m_{ee}$ distribution for same-sign electron pairs in data . . . . .	87
Figure 5.9:	The $m_{ee}$ distribution for opposite-sign electron pairs in data . . . . .	88
Figure 6.1:	The electron efficiency measured vs $p_T$ and $ \eta $ . . . . .	97
Figure 6.2:	The muon efficiency measured vs $p_T$ and $ \eta $ . . . . .	98
Figure 6.3:	# $b$ -tagged jets distribution for $t\bar{t}W$ MC sample to illustrate the effect of the $b$ -tag rescaling procedure . . . . .	101
Figure 7.1:	Distribution of events in the $H_T$ - $\cancel{E}_T$ plane passing the analysis selections for the high $p_T$ baseline region . . . . .	104
Figure 7.2:	Graphical representation of the event yields in the high $p_T$ baseline regions . . . . .	109
Figure 7.3:	Graphical representation of the event yields in the high $p_T$ search regions . . . . .	112
Figure 7.4:	Distribution of events in the $H_T$ - $\cancel{E}_T$ plane passing the analysis selections for the low $p_T$ baseline region . . . . .	113
Figure 7.5:	Graphical representation of the event yields in the low $p_T$ baseline region . . . . .	118
Figure 7.6:	Graphical representation of the event yields in the low $p_T$ search regions . . . . .	121
Figure 7.7:	Diagrams for the four SUSY models considered . . . . .	124
Figure 7.8:	Exclusion curves for models A1 and A2 . . . . .	126
Figure 7.9:	Exclusion curves for models B1 and B2 . . . . .	128
Figure 7.10:	Exclusion curves for model C1 . . . . .	129
Figure 7.11:	Exclusion curve for RPV model . . . . .	131

## LIST OF TABLES

Table 4.1:	Primary datasets used by the high $p_T$ analysis along with the relevant run-ranges . . . . .	53
Table 4.2:	Primary datasets used by the low $p_T$ analysis along with the relevant run-ranges . . . . .	54
Table 4.3:	Sources of true same-sign dileptons from Standard Model processes	55
Table 4.4:	Sources of “false” same-sign dileptons from Standard Model processes	56
Table 4.5:	The trigger paths used for the high $p_T$ and low $p_T$ analyses . . . . .	57
Table 4.6:	Details of the modified VBTF80 electron identification . . . . .	61
Table 4.7:	Details of the loose particle flow identification . . . . .	62
Table 4.8:	Summary of the baseline search regions considered in the high $p_T$ and low $p_T$ analysis . . . . .	65
Table 4.9:	Search regions selected for the high $p_T$ analysis . . . . .	66
Table 4.10:	Search regions selected for the low $p_T$ analysis . . . . .	67
Table 5.1:	Datasets used to measure the lepton fake rates . . . . .	74
Table 5.2:	Simulated samples used to measure the lepton fake rates . . . . .	75
Table 5.3:	The trigger paths used for the high $p_T$ and low $p_T$ analyses to calculate the Fake Rate . . . . .	76
Table 5.4:	Fake rate closure tests on $t\bar{t}$ and $W + \text{jets}$ events from simulation applied to a fake derived from QCD simulation . . . . .	84
Table 5.5:	Electron charge-flip rate as a function of $p_T$ and $ \eta $ . . . . .	86
Table 5.6:	Sources of true same-sign dileptons from Standard Model processes	89
Table 6.1:	Models used for measuring the signal or the background contribution in the tag-and-probe method . . . . .	93
Table 6.2:	Electron efficiencies measured using the tag-and-probe method . . . . .	93
Table 6.3:	Electron efficiency data-to-mc scale factors measured using the tag-and-probe method . . . . .	94
Table 6.4:	Muon efficiencies measured using the tag-and-probe method . . . . .	95
Table 6.5:	Muon efficiency data-to-mc scale factors measured using the tag-and-probe method . . . . .	96
Table 6.6:	The electron efficiency relative differences between DY and $t\bar{t}$ . . . . .	97
Table 6.7:	The muon efficiency relative differences between DY and $t\bar{t}$ . . . . .	98
Table 6.8:	The systematic uncertainties applied to the data-to-MC efficiency scale factors . . . . .	99
Table 6.9:	Scale factors from trigger inefficiencies applied to Monte Carlo predictions from the rare backgrounds . . . . .	99
Table 7.1:	Summary of the baseline search regions considered in the high $p_T$ analysis . . . . .	103

Table 7.2:	High $p_T$ baseline yields and predictions with no b-tagged jets requirement (Signal Region 0) . . . . .	105
Table 7.3:	High $p_T$ baseline yields and predictions with exactly 1 b-tagged jet (Signal Region 10) . . . . .	106
Table 7.4:	High $p_T$ baseline yields and predictions with two or more b-tagged jets (Signal Region 20) . . . . .	107
Table 7.5:	Search regions selected for the high $p_T$ analysis . . . . .	110
Table 7.6:	A summary of the results of this search for the high $p_T$ analysis . .	111
Table 7.7:	Summary of the baseline search regions considered in the low $p_T$ analysis . . . . .	113
Table 7.8:	Low $p_T$ baseline yields and predictions with no b-tagged jets requirement (Signal Region 0) . . . . .	115
Table 7.9:	Low $p_T$ baseline yields and predictions with exactly one b-tagged jet (Signal Region 10) . . . . .	116
Table 7.10:	Low $p_T$ baseline yields and predictions with two or more b-tagged jets (Signal Region 20) . . . . .	117
Table 7.11:	Search regions selected for the low $p_T$ analysis . . . . .	119
Table 7.12:	Summary of the yields for the low $p_T$ analysis . . . . .	120
Table 7.13:	Summary of the search regions used for limit setting in each SMS model considered . . . . .	125
Table A.1:	Electron fake rate measured in bins of the electron candidate $p_T$ and $ \eta $ for the high $p_T$ analysis . . . . .	142
Table A.2:	Electron fake rate measured in bins of the electron candidate $p_T$ and $ \eta $ for the low $p_T$ analysis . . . . .	142
Table A.3:	Muon fake rate measured in bins of the electron candidate $p_T$ and $ \eta $ for both the high $p_T$ and low $p_T$ analysis . . . . .	143
Table B.1:	MC datasets corresponding to contributions from rare Standard Model processes . . . . .	146
Table B.2:	MC datasets that do not contribute to rare Standard Model prediction	147
Table B.3:	MC datasets that were used for the same-sign top and 4-top signal acceptance and systematics . . . . .	147
Table B.4:	MC datasets used for the fake rate closure test . . . . .	148

## ACKNOWLEDGEMENTS

Particle physics turned out to be harder than I expected. It's a huge undertaking with many people and a huge amount of technical knowledge to absorb. This does not include actually learning particle physics itself. There were many times where I didn't think I would make it to this point and without the support of my friends, family, and colleagues, I wouldn't have. Thank you!

To make this possible, I must thank my little brother Mike. He helped me relearn and expand my programming skills to a level I didn't even know existed. His willingness to get dirty with me and talk out all sorts of design and technical issues was a huge help! Mike, if you ever need to know particle physics, I'll try to return the favor – thanks dude...

UCSD would not have been nearly as much fun without my crew: Jeff, Whit, Alex, Bourge, Frank, Toby and Diego. You guys worked and played hard and I will never forget those times. I loved the fact that we formed a winning IM soccer team with a bunch of grad students. Now if I can only stop yelling at the referee...

I'd like to thank my advisor, Professor Avi Yagil, for giving me the opportunity to return to school and finish what I started. I will never forget that.

Frank, I will not forget our intense moments in the trenches of this analysis and then forgetting about it over a nice IPA. I couldn't have done this without you dude. Furthermore, I will never forget your cross that I headed to Randy to finish. For one rare moment, we looked like Premier League quality.

To my twin brother, Randy, I appreciate your patience and encouragement. This was difficult undertaking but I couldn't have done it without the "twin" factor. Now that I'm at the end, I will actually miss our physics discussions.

Finally, to Jennifer, my wife. She has been through the whole analysis from beginning to end. She stood by me during it all. Your love and patience was the only way I really got through this. I can't wait to start the next chapter with you...

## VITA

- |      |                                                         |
|------|---------------------------------------------------------|
| 2000 | B. S. in Electrical Engineering, University of Virginia |
| 2007 | M. S. in Physics, University of California, San Diego   |
| 2013 | Ph. D. in Physics, University of California, San Diego  |

## PUBLICATIONS

Search for new physics in events with same-sign dileptons and b jets in pp collisions at  $\sqrt{s} = 8$  TeV, *CMS Collaboration*, JHEP 1303 (2013) 037 [arXiv:1212.6194 [hep-ex]]

## ABSTRACT OF THE DISSERTATION

### **Search for New Physics in Proton-Proton Collisions at 8 TeV Center of Mass Energy with a Final State of Same-Sign Dileptons and Jets**

by

Ryan Ward Kelley

Doctor of Philosophy in Physics

University of California, San Diego, 2013

Professor Avraham Yagil, Chair

A search for new physics is performed using events with isolated same-sign leptons and jets in the final state. Results are based on the full sample of proton-proton collisions collected from the Large Hadron Collider at a center-of-mass energy of 8 TeV with the CMS detector and corresponding to an integrated luminosity of  $19.5 \text{ fb}^{-1}$ . No excess above the standard model background is observed and constraints on a number of new physics models are set.

# Chapter 1

## Introduction

The goal of particle physics is to understand the basic laws governing the nature of matter and its interactions. Throughout the twentieth century, particle physics experiments and theoretical work have culminated in a model that governs the electromagnetic, weak and strong nuclear interactions, which mediate the dynamics of the known subatomic particles. Since then, discoveries of the bottom quark, the W and Z bosons, the top quark and the tau neutrino have given support to this model. Because of its success in explaining a wide variety of experimental results, this theory is regarded as the "Standard Model" (SM).

The latest particle accelerator, the Large Hadron Collider (LHC), was built to explore the predictions of different theories of particle physics, particularly to provide evidence for the existence of the theorized Higgs Boson and a large family of new particles predicted by "Supersymmetric" theories [4]. The LHC is a proton-proton collider built along the Franco-Swiss border that will eventually operate at a design energy seven times larger than the previous generation accelerator. Built to observe the high energy proton-proton collisions delivered by the LHC, the Compact Muon Solenoid (CMS) is a general-purpose particle detector used to investigate a wide range of physics, including the search for the Higgs boson, extra dimensions, Supersymmetry, and particles that make up dark matter.

The goal of this thesis is to perform one of these general searches for new physics using the final state with two leptons with the same electric charge that arise from the decay of gauge bosons. It uses the full dataset of 8 TeV collisions collected by CMS during

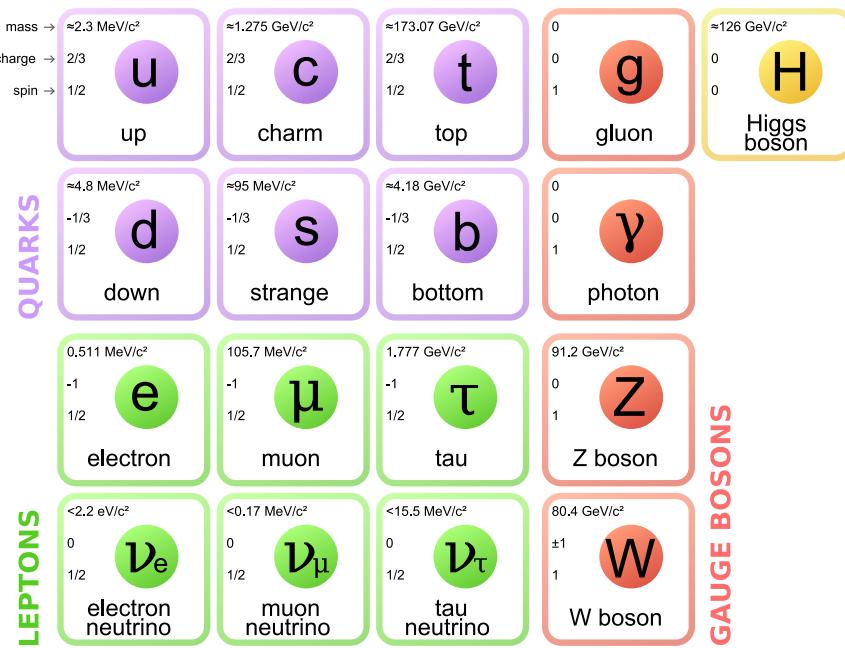
the 2012 run and corresponds to  $19.5 \text{ fb}^{-1}$  of integrated luminosity. This is a promising signature because the expected Standard Model backgrounds are relatively rare and any new physics with this final state would show up an excess of collision events. The results of this thesis expand on results previously published using  $10.5 \text{ fb}^{-1}$  of the 2012 data [5]. Several changes have been made with respect to this analysis including changes to the lepton selection and expanded search regions to achieve maximum sensitivity.

This thesis is organized into eight chapters. First, I will present a brief review of the standard model and motivate the analysis in Chapter 1. An overview of the LHC followed by a discussion of the CMS detector along with the algorithms it uses to reconstruct the decay products of the proton-proton collisions is given in Chapter 2. In Chapter 3, I present a discussion of the main sources of backgrounds in this analysis from the Standard Model and briefly discuss other sources of background due to algorithmic effects. The main analysis is shown in the next four chapters with the event selection (Chapter 4), the background estimation (Chapter 5), the efficiency measurements (Chapter 6) and finally the results (Chapter 7). The last chapter concludes the thesis.

## 1.1 Standard Model of Particle Physics

Starting with the beginning of the twentieth century, our understanding of matter has progressed from little knowledge of the structure of matter to a theoretical framework that predicts, to very high accuracy, many of the observed particle interactions from experiments performed throughout the previous century. Currently, matter and energy are best understood in terms of the kinematics and interactions of elementary particles. The laws governing the behavior and interaction of all known forms of matter and energy have been reduced to a small subset of fundamental theories. The Standard Model represents a theory that predicts the interactions of all known particles through the electromagnetic, weak, and strong nuclear forces.

The SM has 61 elementary particles which are summarized in the Figure 1.1 below and fall into two categories – fermions and bosons [6].



**Figure 1.1:** The list of fundamental Standard Model particles broken down by particle type: quarks in purple, leptons in green, gauge bosons in red, and the Higgs boson in yellow. The anti-quarks and anti-leptons are not shown [7].

## Fermions

The SM includes twelve spin- $\frac{1}{2}$  elementary particles known as fermions that make up all known matter in the universe. These fermions respect the Pauli exclusion principle and each has a corresponding anti-particle. The SM classifies the fermions according to how they interact through the quantum numbers they carry. There are six quarks each assigned a different flavor quantum number: up, down, charm, strange, top, and bottom. The remaining six fermions are the leptons which are also identified by their flavor: electron, electron neutrino, muon, muon neutrino, tau, and tau neutrino. Pairs from each are classified and grouped together to form generations, which correspond to particles exhibiting similar physical behavior.

The major property of the quarks is that they carry the color charge and interact via the strong interaction. A phenomenon called color confinement results in quarks being bound to one another, forming color neutral composite particles called hadrons. Hadrons fall into two main categories: mesons which are quark and antiquark bound states and baryons which are bound states of three quarks (or three antiquarks). The familiar proton and the neutron are the two baryons with the lowest mass. Quarks also carry electric and weak charges and hence, they interact with all fermions via both the electromagnetic and weak interactions.

The other main category of fermions, the leptons, do not carry color charge and do not participate in the strong interaction. The three neutrinos do not carry electric charge either, so they are only influenced by the weak nuclear force. This property leads neutrinos to be notoriously difficult to detect experimentally. However, the remaining three leptons do carry electric charge: the electron, muon and tau all interact electromagnetically.

All 12 fermions are grouped into three generations where each member of a generation has a greater mass than the corresponding particle in lower generations. The first generation of quarks and leptons do not decay; hence all ordinary matter is made up of such particles. Specifically, all atoms consisting of electrons orbiting the nucleus which is made up of up and down quarks. Second and third generations particles are very short lived and are only produced in very high-energy environments (such as the LHC). Neutrinos of all generations also do not decay and very rarely interact with matter.

## Bosons

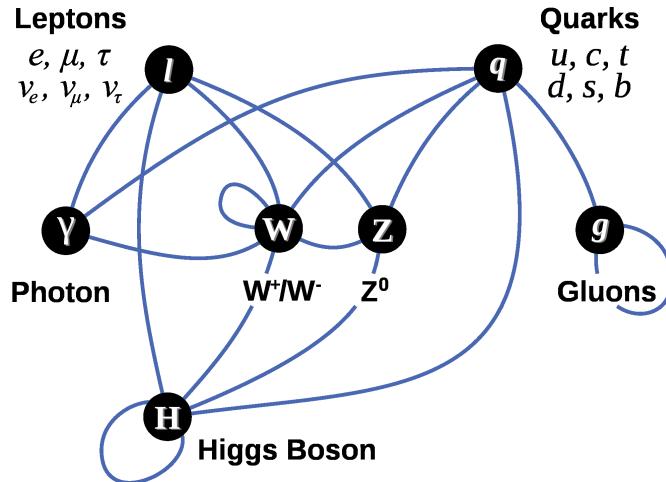
In the SM, gauge bosons play the roll as the “force carriers” and are the mediators of the strong, weak and electromagnetic fundamental forces or interactions. The SM explains such forces as resulting from matter particles exchanging these bosons, known as force mediators. At a macroscopic level, the effect is equivalent to a force field influencing both of them; however, when looking at the microscopic level, a gauge boson has been exchanged. The gauge bosons of the SM all have spin-1, and as a result, they do not follow the Pauli exclusion principle that constrains fermions and thus have no limit on their spatial density. The different types of gauge bosons are described below:

- Photons ( $\gamma$ ) mediate the electromagnetic interaction between charged particles. The photon is massless and is described by the subset of the SM known as quantum electrodynamics.
- The  $W^+$ ,  $W^-$ , and  $Z$  gauge bosons mediate the weak interactions between particles of different flavors (all quarks and leptons). They are massive with the  $W^\pm$  and  $Z$  bosons being  $80.4 \text{ GeV}/c^2$  and  $90.2 \text{ GeV}/c^2$ , respectively [8]. The  $W^+$  and  $W^-$  bosons carry an electric charge (indicated by the superscript) and the  $Z$  boson is electrically neutral. The final characteristic is that the  $W$  bosons decay weakly into final states of either two quarks or two leptons. In the case of the two quark final state, one quark must be a up-type quark (up, charm) and the other a down type anti-quark (down, strange, or bottom). The  $W$  does not decay via a top quark due to kinematic constraints imposed by the much more massive top quark. In the case where the  $W$  boson decays to a lepton pair, one lepton must be a charged lepton (electron, muon, or tau) and the other must be a neutrino (electron neutrino, muon neutrino, or tau neutrino). Finally, the  $Z$  boson must decay to a fermion and its anti-particle. For example  $Z \rightarrow \ell^+ \ell^-$  or  $Z \rightarrow q\bar{q}$  where  $q$  represent a quark,  $\bar{q}$  represents an antiquark, and  $\ell$  represents a charged lepton.
- The eight gluons ( $g$ ) mediate the strong interactions between color charged particles (the quarks). Gluons are massless. The eightfold multiplicity of gluons is labeled by a combination of color and anticolor charge, and because gluons them-

selves carry a color charge, they can interact among themselves. The gluons and their interactions are described by the theory of chromodynamics.

The final boson in the SM is the Higgs boson and is a key building block in the underlying structure of the theory. It has no intrinsic spin (scalar), is massive at  $125.7 \text{ GeV}/c^2$  [9], and plays a unique role by explaining why other elementary particles, except the photon and gluon, are massive [10]. In particular, they explain why the photon has no mass, while the W and Z bosons do. Also, in electroweak theory, the Higgs boson generates the masses of the leptons and quarks. Finally, as the Higgs is massive, it must interact with itself.

The following Figure 1.2 shows a summary of all the allowed interaction between particle types in the Standard Model.



**Figure 1.2:** A summary of the allowed interactions between particle types in the Standard Model. A line connecting two particles indicates coupling occurs between those particles. A line from one particle to itself indicates that this particular particle self-couples [11].

### Known deficiencies of the SM

The Standard Model is not considered a complete theory. Some of the major issues are outlined below:

- The SM does not provide a mechanism to explain gravitation. Attempts have been made to find a quantum theory of gravity that is also consistent with general relativity; however, these theories break down before reaching the Planck scale and therefore are not satisfactory.
- The SM is considered *ad hoc* and inelegant since it requires many numerical constants that seem arbitrary and unrelated.
- The Higgs mechanism gives rise to the hierarchy problem if any new symmetries or particle interactions are present at a high energy scale. In order for the weak scale to be much smaller than the Planck scale, severe fine tuning of the SM parameters is required.

Also there are many questions in physics that are unanswered by the SM. Some of these issues include:

- Why do particle masses and coupling constants have the values that we measure?
- Why are there three generations of particles?
- Why is there more matter than antimatter observed in the universe?
- Where does dark matter fit into the model?
- Where does dark energy fit into the model?

Clearly the SM does not provide the final explanation of all the observed interactions in nature.

## 1.2 Proton-Proton Collisions

To probe some of the unanswered question from the SM posed in the previous Section (Section 1.1), protons are collided at high energy and their debris is studied in detail. In particle physics, the probability for an interaction to occur is given by the cross section ( $\sigma$ ), which represents the effective area presented by the target to an

incoming particle. The units for cross section are given in terms of a barn (b) where  $1 \text{ b} = 10^{-24} \text{ cm}^2$ . The total number of interactions is given by the relation

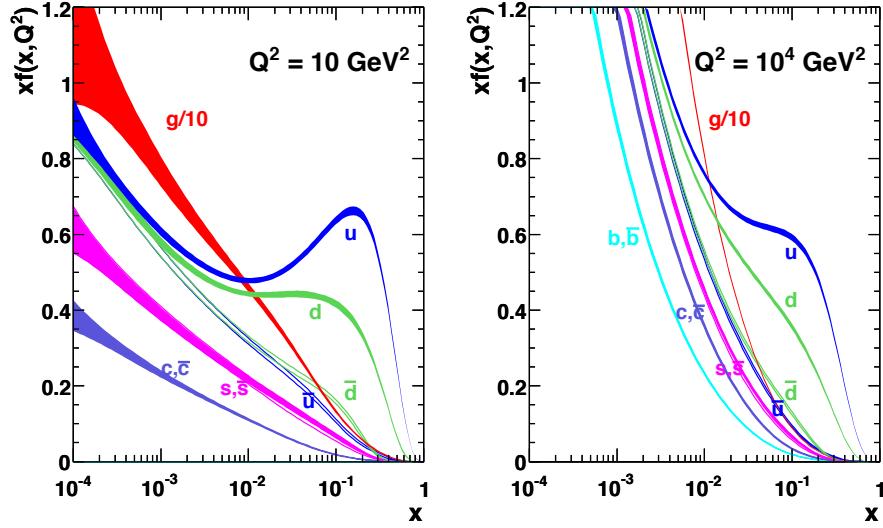
$$N = \mathcal{L} \cdot \sigma, \quad (1.1)$$

where  $N$  is the number of interactions occurring in an instantaneous luminosity of  $\mathcal{L}$  for a process with an interaction cross section  $\sigma$ . The total amount of data collected is traditionally reported as the total integrated luminosity reported by the above relation. The integrated luminosity has units of inverse barns ( $\text{b}^{-1}$ ); however, due to the inconvenience of the scale of this unit, it is customary to report the units in either inverse picobarns ( $\text{pb}^{-1} = 10^{12} \text{ b}^{-1}$ ) or inverse femtobarns ( $\text{fb}^{-1} = 10^{15} \text{ b}^{-1}$ ).

The difficulty of producing and storing antiprotons led to the decision to use proton-proton (pp) collisions at the LHC [12]. The cross section for a typical pp collision  $\approx 100 \text{ mb}$  [13]. This however, combines both elastic collisions, where the two colliding protons remain whole, and inelastic collisions, where the proton breaks apart into its constituent quarks or gluons (partons). For inelastic collisions, QCD does not allow for a quantitative calculation of the exact kinematics of the partons that make up the proton. Instead, one has to appeal to a series of experiments designed to create an empirical model of the kinematic distributions of these partons. This is done with deep inelastic scattering experiments where a high energy electron is collided with the proton [6]. By measuring the resulting momentum and angular distributions of the proton's debris, one can determine of the fraction of energy carried by the partons which make up the proton. As an example, Figure 1.3 shows the fraction of energy carried by gluons and quarks for two different energy protons. Notice that while the up and down quarks and the gluons are the dominant high momentum carriers, there is still significant momentum from the other partons (sea quarks).

### 1.3 Motivation for Same-Sign Dilepton Signature

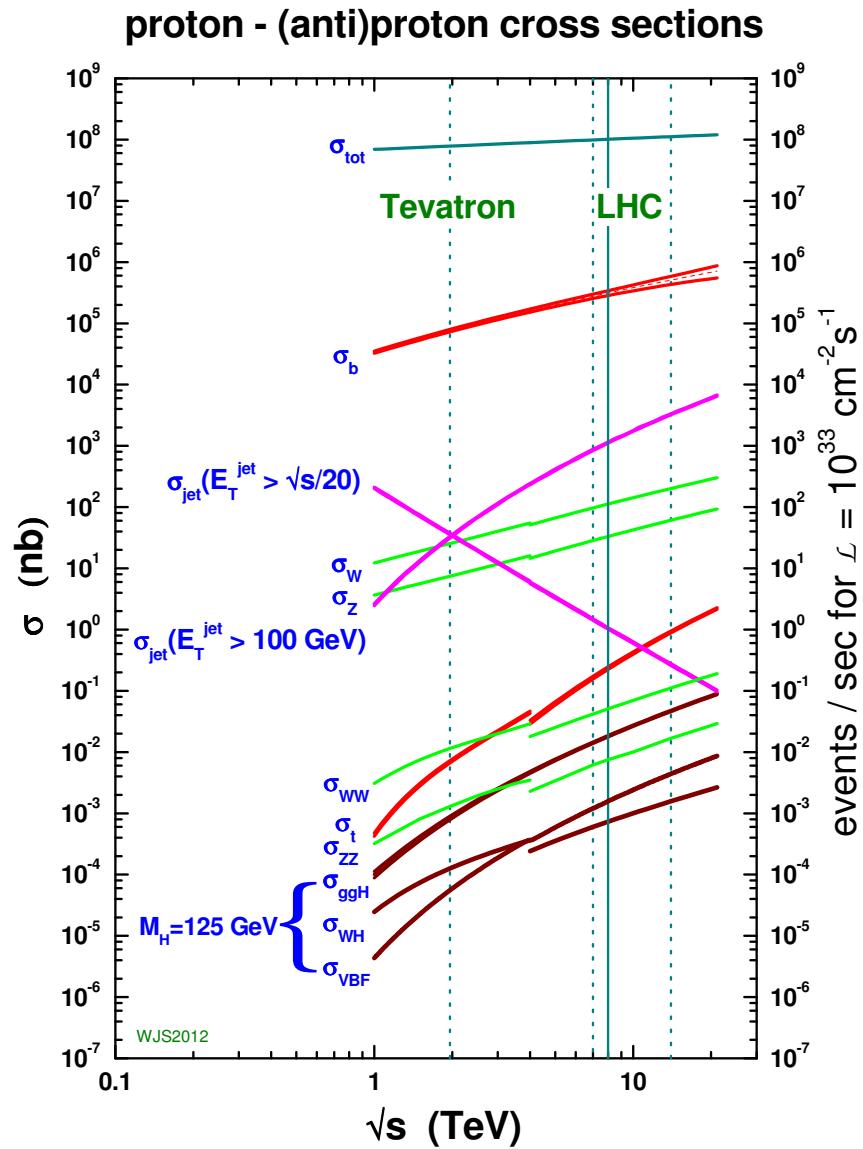
One of the purposes of the LHC and CMS is to search for physics beyond the Standard Model (BSM). As mentioned in Section 1.1, astrophysical evidence for dark matter suggests that the SM is incomplete. Although direct detection experiments have yet to make an observation, indirect measurements are pointing to dark matter being



**Figure 1.3:** The parton distribution functions and their associated error for a proton at a momentum transfer scale of  $Q^2 = 10 \text{ GeV}^2$  (left) and  $Q^2 = 10000 \text{ GeV}^2$  (right) as a function of  $x$ , the fraction of the proton’s momentum carried by the parton [14].

related to the weak interaction [15, 16]. It is clear that the SM is not the full story describing particle interactions. The physics analysis discussed in this thesis attempts to take advantage of the strengths in lepton reconstruction to search for previously unobserved physics.

Figure 1.4 shows the production cross section for a number of SM processes as a function of the center-of-mass energy of the colliding beams. Comparing the relative size of the total inelastic collisions cross section, contaminated by QCD processes, with that for W boson production, it is apparent that these electroweak processes are rare in the SM. Subsequent leptonic decays of the W boson occurs at the rate of  $\approx 30\%$ , making leptons originating from a W boson even more rare. Sources of dilepton final states are significantly more rare with the dominant source being  $Z \rightarrow \ell^+ \ell^-$ . Additional sources of dileptons include  $t\bar{t}$ ,  $WW$ ,  $WZ$ , and  $ZZ$  production, all with smaller production cross sections. Most of the dilepton pairs from these have opposite electric charge (+/-). Since these leptons are both prompt and relatively isolated, they provide an extremely clean experimental signature in the detector since they are simpler to reconstruct with smaller rates of mis-measurement. The combined effect is much smaller backgrounds for searches in dilepton final states, particularly same-sign dilepton final



**Figure 1.4:** Cross sections for some Standard Model processes vs center-of-mass energy. The cross sections of SM process giving final state leptons are many orders of magnitude smaller than QCD cross sections. From [17].

states (++) or (--), compared to single lepton and fully hadronic analyses.

From the above discussion, it is clear that a signature with two prompt and isolated leptons with the same electric charge (same-signed dileptons) is a rare occurrence in the SM relative to other processes. As a result, searches for anomalous production of same-sign dileptons can be very sensitive to new physics contributions. These include

- supersymmetry (SUSY) [18, 19, 20],
- universal extra dimensions [21]
- pair production of  $T_{5/3}$  particles (fermionic partners of the top quark) [22],
- heavy Majorana neutrinos [23],
- and same-sign top pair production [24].

New physics signatures with large cross sections are likely to be produced by strong interactions, and we thus expect significant hadronic activity in conjunction with the two leptons. Additionally, SUSY models with R-parity conservation and astrophysical evidence for dark matter suggest considering final states with undetectable particles that leads to a significant amount of missing transverse energy ( $\cancel{E}_T$ ) [25, 26]. With the above considerations, in this analysis, we search for new physics with events containing:

- same-sign dileptons (electrons and muons),
- hadronic jets (with and without b-tagging),
- and accompanying missing transverse energy.

The exact event selection will be discussed in Chapter 4.

The basic idea is to count the number of observed collision events that have the above signature and compare this to the number of events that were expected assuming the SM only. This comparison is done using statistical techniques to be discussed in Section 7.3. If there is a significant excess of events above your expected SM background, then this gives evidence that new physics exists beyond the SM. It is important to get the most accurate background prediction possible to ensure maximum sensitivity to an anomalous excess of events. The specific details of the background estimation

techniques can be found in the following chapters. Finally, in Chapter 7, we show the results of this search along with some interpretations with respect to some selected new physics models.

# **Chapter 2**

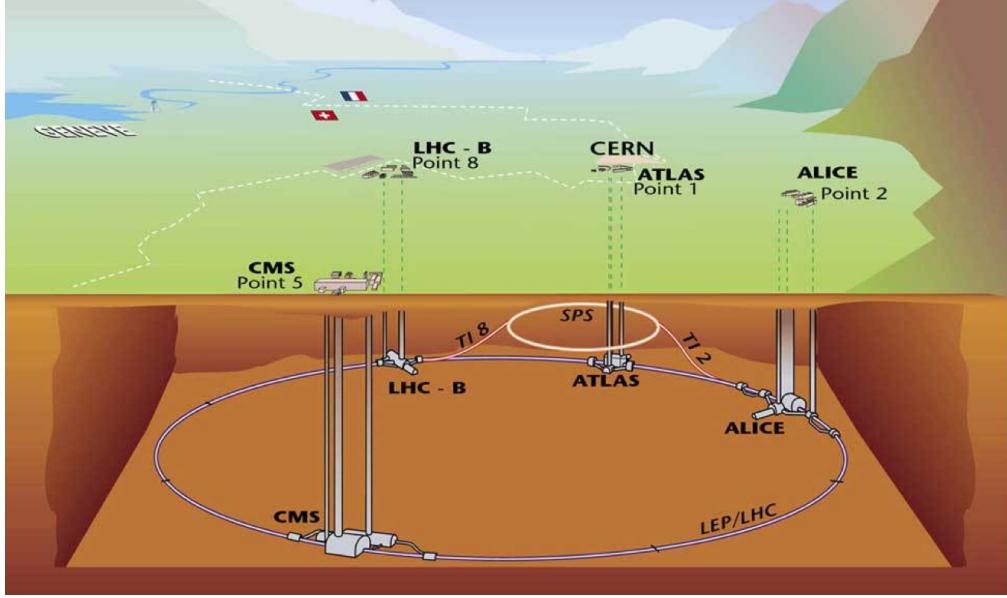
## **The LHC Accelerator and CMS Experiment**

The Compact Muon Solenoid (CMS) experiment is a collaboration of scientists working together to investigate the data streaming from the Large Hadron Collider (LHC). This chapter will give an overview of the experimental setup used by the nearly 3,000 CMS collaborators to analyze this data. First we give a short discussion of the Large Hadron Collider followed by the Compact Muon Solenoid detector itself. Finally, we will give a brief description of the algorithms used to measure the particles traversing the CMS detector used by the various analyses.

### **2.1 The Large Hadron Collider**

As previously discussed in Chapter 1, the Standard Model of particle physics (SM) is the de facto theory that currently describes all known particle interactions save gravity. The aim of the Large Hadron Collider (LHC) is to investigate the possibility of physics beyond the Standard Model. Completed in 2008, it is approximately 27 km in circumference, sits 170 meters underground and was designed to accelerate protons to 7 TeV each and collide them at four interaction points. Each interaction point has a detector positioned there to record the outcome of these proton-proton collisions: the two general purpose detectors CMS and ATLAS, and the other two specialized detectors ALICE and LHC-b. Figure 2.1 shows a cartoon of the LHC and the four particle

detectors which is located on the outskirts of Geneva, Switzerland.



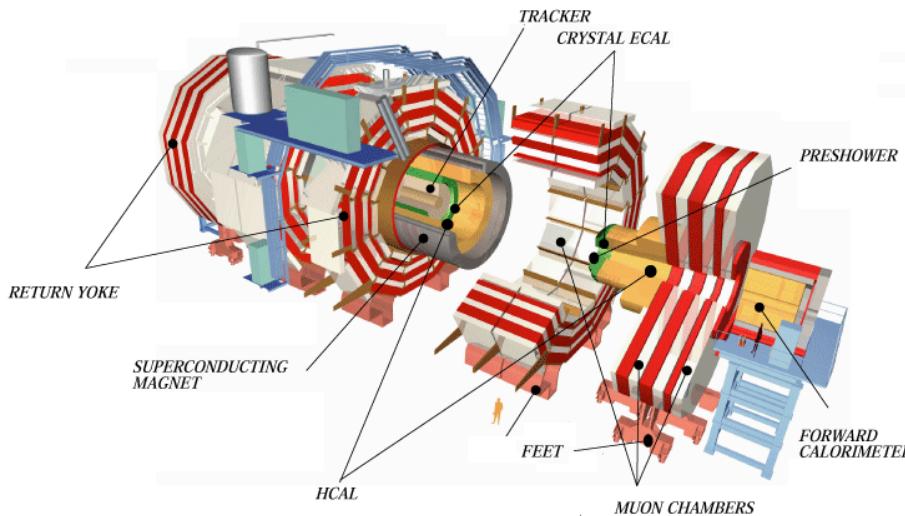
**Figure 2.1:** The Large Hadron complex seen underground on the outskirts of Geneva, Switzerland. The four detectors main detectors are also shown – LHC-B, ATLAS, ALICE, and, most relevant to this analysis, CMS at Point 5 [12].

The search for the Higgs boson carried a large influence on the design of the Large Hadron Collider (LHC) and the CMS detector. The previously unknown mass of the Higgs boson combined with its small production cross section required building a machine capable of high energy collisions at a high rate. This design leads to approximately 1 billion proton-proton collisions per second [12]. Rather than continuous beams, the protons will be “bunched” together, into 2,808 bunches, 115 billion protons in each bunch so that interactions between the two beams will take place at discrete intervals never shorter than 25 nanoseconds (ns) apart. The design luminosity of the LHC is  $10^{34} \text{ cm}^2 \text{s}^{-1}$ , providing a bunch collision rate of 40 MHz.

Although the design center-of-mass energy ( $\sqrt{s}$ ) is 14 TeV, the data collected from 2011 and 2012 was at  $\sqrt{s} = 7$  TeV and 8 TeV, respectively. The data used for this thesis is the full dataset recorded from 2012 at  $\sqrt{s} = 8$  TeV. The LHC plans to continue collisions in 2015 at  $\sqrt{s} = 13$  TeV.

## 2.2 The Compact Muon Solenoid Detector

The Compact Muon Solenoid (CMS) is one of the two general purpose detectors situated along the beam line of the LHC. The name “compact” comes from the fact that the detector is relatively small when compared to its sister experiment known as A Toroidal LHC Apparatus (ATLAS); however, CMS is by no means small. Weighing in at over 12,000 tons, cylindrical in shape, it is over 21 meters in length and 14 meters in diameter. It sits  $\sim$ 100 meters underground at LHC Point 5 near Cessy, France (see Figure 2.1). An overview of the CMS detector can be seen in Figure 2.2. CMS was de-



**Figure 2.2:** Overview of the Compact Muon Solenoid detector [27].

signed to allow for the study of many different unanswered questions in particle physics. To achieve this broad objective, CMS is capable of measuring with high precision, the trajectory, momentum and energy of many different types of particles. In addition, it is able to handle the challenging conditions presented by the high luminosity and high energy collisions from the LHC. To deal with the approximately one billion interactions per second expected from the LHC at its design luminosity, CMS is a very high granular detector, and hence has many electronic channels, each with excellent time resolution and is designed to be radiation hard to deal with the multitude of particles produced during each bunch crossing.

CMS contains a strong magnetic field produced by a superconducting solenoid.

An inner field strength of 3.8 Tesla ensures that even the highest momentum particles propagating through the detector have enough curvature to ensure a precise momentum measurement. Inside the solenoid sit both the silicon and pixel tracker systems and both the electromagnetic (ECAL) and hadronic calorimeters (HCAL). The larger number of channels in the pixel tracker ensure a precise measurement of the origin of particles both produced near the beamline and those with longer lifetimes decaying at larger diameters. The silicon strip detector has at least 10 layers along any particles' path allowing precise particle "tracking" and momentum measurements. The ECAL, with over 75,000 lead-tungstate ( $\text{PbWO}_4$ ) scintillating crystals, gives excellent energy measurements of the particles which interact primarily electromagnetically, namely electrons and photons. The HCAL, a sampling calorimeter made of brass, completes the sub-detectors existing inside of the solenoid and gives respectable energy resolution for particles that interact primarily hadronically (e.g. pions and kaons).

Outside of the solenoid rests the final major sub-detector, the muon detector. Made up of three separate components, the muon chambers are placed between the large iron frame (yoke) that serves two purposes. First, they prevent particles other than muons from reaching the outermost muon chambers and hence allow for excellent muon identification. Secondly, the yoke allows for a strong return magnetic field to fill the entirety of the muon detector system which allows for better momentum resolution.

The coordinate system of CMS is centered at the nominal interaction point at the center of CMS's cylindrical shape. We define the y-axis to point upwards (away from the center of the earth), the x-axis points towards the center of the LHC ring and hence, the z-axis points along the beamline, counter clockwise around the LHC. The azimuthal angle,  $\phi$ , is measured from the zero position of the x-axis in the x-y plane while the polar angle,  $\theta$ , is measured from the positive z-axis.

Finally, we should introduce now two important quantities used in hadronic collider physics. While we know precisely the center-of-momentum frame and individual momenta of the two colliding protons, we do not know the center-of-momentum frame for the two colliding partons. For this reason, we define two quantities that are invariant under a boost in the direction of the proton's momentum. The first quantity is the partons' momentum in the plane normal to the collision – i.e. the transverse plane, the

transverse momentum ( $p_T$ ) can be defined and is, to first order, conserved. The second quantity, pseudorapidity ( $\eta$ ) is a transformation from the polar angle into an invariant quantity with respect to the frame of the collision:

$$\eta = -\ln \left[ \tan \left( \frac{\theta}{2} \right) \right], \quad (2.1)$$

where  $\theta$  is the angle between the incoming proton and the outgoing particle. These two quantities,  $p_T$  and  $\eta$ , are used extensively at the LHC to describe the kinematics of the out going particles from a proton-proton collision and will be used throughout this analysis as well.

A complete description of the CMS detector can be found in Reference [28], from which all information in this section is derived, unless explicitly stated. The following subsections give a more detailed description of the various pieces that make up the CMS detector. This is not meant to be a full detector description; however, we do highlight details that are important to the analysis in this thesis.

### 2.2.1 The Superconducting Solenoid

Charged particles in an electromagnetic field experience a force given by

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}),$$

where  $q$  is the charge of the particle,  $\mathbf{E}$  is the electric field,  $\mathbf{v}$  is the velocity of the charged particle and  $\mathbf{B}$  is the magnetic field. Therefore, for a uniform magnetic field, the radius of curvature,  $r$ , of a charge particles is directly proportional to the particle's momentum and inversely proportional to the magnetic field strength,

$$r = \frac{p_T}{q|\mathbf{B}|},$$

where  $p_T$  is the particles transverse momentum with respect to the direction of the magnetic field. With  $\sim 3$  meters from the nominal interaction point to the inside edge of the solenoid, CMS must have a strong magnetic field to achieve the maximum possible curvature for the charged particles. This ensures the most precise measurement of the momentum and this is achieved with CMS's superconducting solenoid.

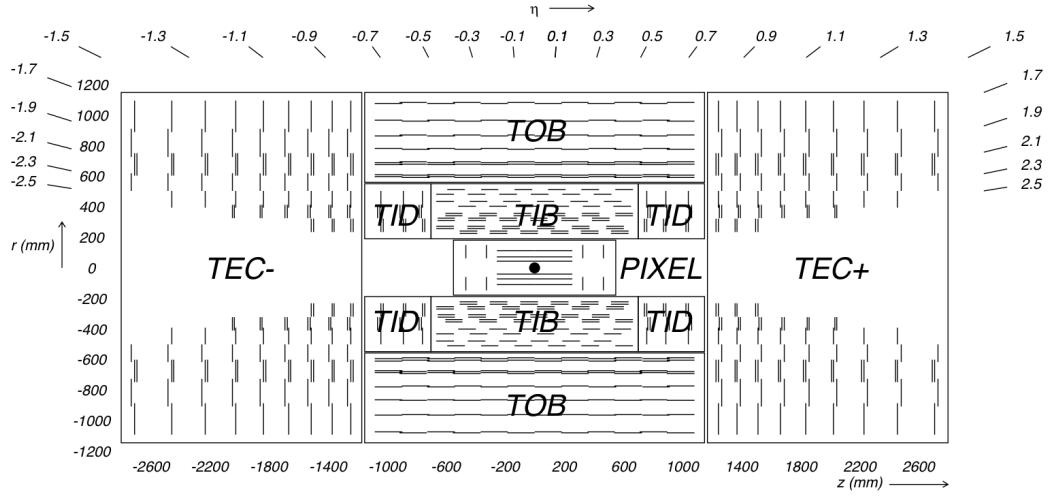
Made up of Niobium-Titanium (NbTi), the solenoid was designed to operate with a magnetic field strength of 4 Tesla, yet due to safety considerations, is operated at 3.8 Tesla. At maximum strength, the solenoid stores over 2 Gigajoules of energy and far exceeds the design strength of all other particle detectors.

To allow for this stronger magnetic field, and hence better momentum measurement, for the muon chambers, the outside of the solenoid has been designed with iron yokes which offer integral support, provide additional stopping power for non-muons and propagates the magnetic field. The yokes can be seen in red in Figure 2.2.

### 2.2.2 The Silicon Detector

The innermost portion of the CMS detector is comprised of a silicon based particle tracking system (tracker). Doped and negatively biased, when charged particles traverse the active region of the silicon, electric current is induced and measured. This gives a positional measurement of the charged particles and is recorded for later reconstruction (hits). Combining the multiple positional measurements, the curvature of a charged particle is measured, and thus, using the known magnetic field strength, the particle's momentum. The tracker was designed to have a high efficiency to properly reconstruct the trajectory, good momentum resolution and be able to identify secondary vertices from longer lived particles such as b-flavored hadrons.

Many aspects of the tracker were specifically tailored to achieve the design considerations discussed above and to physically fit within the CMS solenoid. First, in order to deal with large particle multiplicities, the detector needed to be radiation hard, have a high granularity and a fast electronic response time. The high granularity was achieved by using a pixel silicon detector in the inner most region. Outside of the pixel detector, silicon strips are used as the particle multiplicity density decreases as the square of the radius from the interaction point. This design achieves the desired position, momentum and vertex resolution while keeping the occupancy down to a manageable 1% throughout. An overview of the layout of the silicon tracker can be see in Figure 2.3 with the following subsections discussing the pixel and strip detectors, respectively.



**Figure 2.3:** Layout of the tracking system projected into the  $r_z$  plane. Modules are shown as lines and modules with back-to-back labels appear as double parallel lines [29].

### The Pixel Detector

In order to achieve the 1% occupancy nearest the beam line, pixel silicon detectors are used within a radius of  $\sim 10$  cm. The CMS pixel detector consists of three concentric cylindrical layers at 4.4, 7.4 and 10.2 cm in the radial direction with two disks at 34.5 and 45.6 cm on each side along the z-coordinate. Together, the “barrel” and “endcap” pixel detectors cover a pseudorapidity range up to  $|\eta| < 2.5$ . The pixels themselves are approximately square with a size of 100 to 150  $\mu\text{m}$  per side and have a spatial resolution of 10  $\mu\text{m}$  in the  $r - \phi$  direction and 200  $\mu\text{m}$  in the z direction. Besides allowing for low occupancy, the pixel detectors are also made radiation hard and are designed to give precise vertex position measurements.

### The Strip Detector

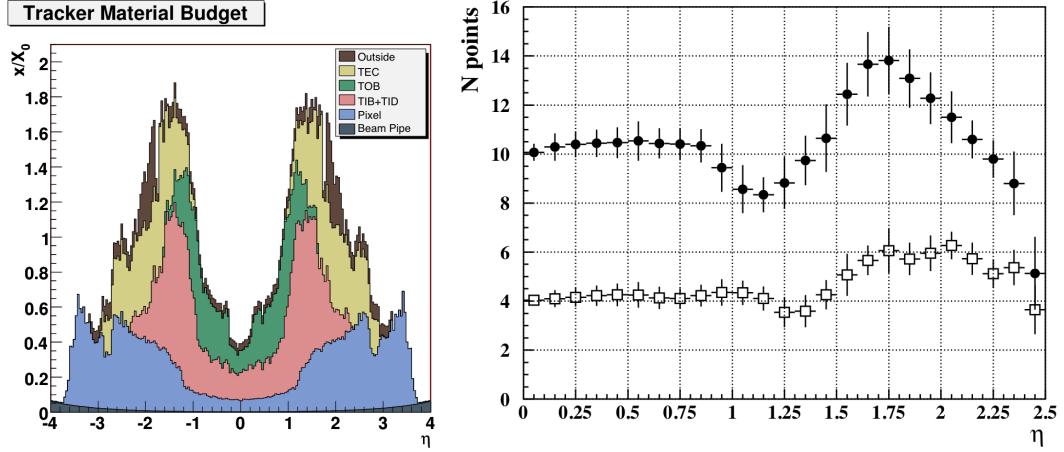
Outside of the pixel detector sits the silicon strip detector. At these distances from the interaction point, a lower density of sensors can be used to obtain the same occupancy. Strip detectors (long rectangular active regions) are significantly cheaper to manufacture than pixel detectors while still providing the desired momentum resolution

and channel occupancy. The strip detector can be divided into four different regions: the tracker inner barrel (TIB), the tracker inner disks (TID), the tracker outer barrel (TOB) and the tracker endcaps (TEC). Each region can be seen in Figure 2.3. The TIB consists of four cylindrical layers in the barrel and is surrounded on each side by three disks from the TID. The TIB and TID extend out to a radius of approximately 55 cm and provide position measurement resolution of between 23 and 35  $\mu\text{m}$ . The TOB surrounds the TIB and TID and provides another 6 cylindrical layers and provides slightly less resolution of around 35 to 53  $\mu\text{m}$ . Finally, the TEC extends out into the z-direction by providing an additional 9 disk layers with resolution between 97 and 184  $\mu\text{m}$ . The silicon strip detector also provides a pseudorapidity coverage inside  $|\eta| < 2.5$ . Some of the layers in the strip detector, namely the first two layers in each TIB, TID, TOB and TEC as well as layer five in the TEC, carry an additional offset detector mounted at a slight angle. This allows not just for additional positional measurement but an additional measurement in the “non-strip” direction (z in the barrel and r in the endcaps).

## The Tracker

With over 200 square meters of silicon, the CMS Tracker is the largest silicon detector in the world. Unfortunately, material comes at a performance cost. The silicon must all be powered, and in turn, the powered units must all be cooled (operating temperature is  $-10^\circ\text{C}$ ). The amount of material in the CMS tracker is shown in units of radiation length in Figure 2.4. A radiation length is a characteristic of a material and describes the distance over which electrons and photons lose energy and is the length at which an electron, on average, will lose all but  $1/e$  of its initial energy via bremsstrahlung. All of this material inside CMS causes a few complications. First, particle momentum resolution is degraded as each particle has a higher probability for nuclear interaction and multiple scattering as it traverses the detector. Second, electrons and photons have higher probabilities to bremsstrahlung and pair create, respectively, causing issues measuring their energy in the electromagnetic calorimeter (see next section).

The tracker provides at least ten particle position measurement for any possible particle trajectory within the tracker’s acceptance region which is about  $|\eta| < 2.5$ . Fig-

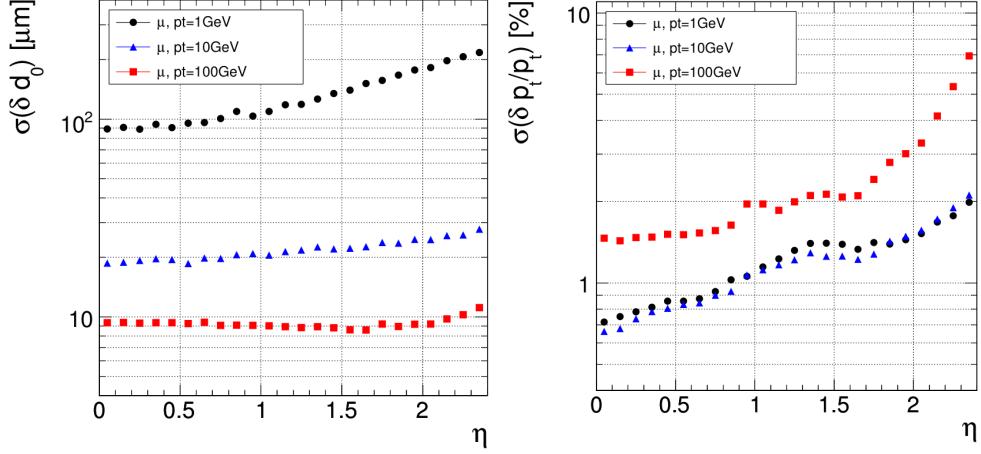


**Figure 2.4:** Radiation length of the CMS tracker as a function of pseudorapidity (left). Number of layers passed as a function of pseudorapidity (right). Black circles represent total layers while hollow squares represent the number of layers with a stereo measurement [29].

Figure 2.4 shows the number of measurements as a function of  $|\eta|$ . Some of the tracker's silicon layers have dual layers with a slight offset to facilitate a more accurate position measured. The number of these stereo layers are also shown in Figure 2.4. The overall initial transverse position (traverse impact parameter,  $d_0$ ) and the momentum resolution is shown in Figure 2.5. The overall charged particle reconstruction efficiency, calculated using standard tag and probe techniques (see Section 6.1.1), is on average greater than 99% [29].

### 2.2.3 The Calorimeters

CMS has two calorimetry detectors, both situated inside the solenoid. The first, directly behind the silicon tracker, is the electromagnetic calorimeter (ECAL), designed to measure the energy of particles which interact primarily electromagnetically, i.e. electrons and photons. The second, directly behind the ECAL, is the hadronic calorimeter (HCAL), designed to measure particles interacting predominately by the strong interaction. Both sub-detectors were designed to be fully hermetic, have good energy resolution over a large area of the detector, and hence be able to accurately reconstruct the total



**Figure 2.5:** Muon transverse impact parameter resolution (left) and transverse momentum resolution as a function of pseudorapidity for various muon momenta: 1 GeV as circles, 10 GeV as triangles and 100 GeV as squares [29].

amount of energy released for each proton-proton interaction. Combined, the calorimeters allow for the detection of neutrinos escaping from the CMS detector by measuring an energy imbalance in the transverse plane. In addition, the calorimeters are designed to be thick enough to prevent most particles from pushing through the solenoid and into the muon chambers. To achieve these features, both detectors had to be composed of dense and highly segmented material.

### Electromagnetic Calorimeter

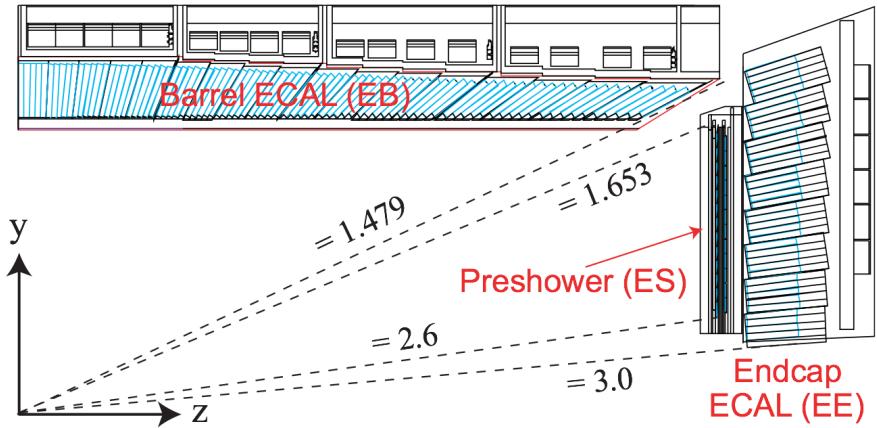
Lead Tungstate ( $\text{PbWO}_4$ ) crystals were chosen to comprise the fully hermetic and homogeneous ECAL.  $\text{PbWO}_4$  satisfies the needs imposed by the LHC environment: the radiation resistance allows for long lifetime and the high density satisfies the space constraints as well as allows for fine granularity and quick read out times. The homogeneity of the detector allows for good overall energy resolution.  $\text{PbWO}_4$  has a radiation length ( $X_0$ ) of 0.89 cm and a Molière radius of 2.2 cm. The Molière radius is a measure of the spread of energy from showering electrons and photons and is defined as the radius which contains 90% of the showering energy. Having a small Molière radius

also allows for the use of smaller crystals and hence gives higher segmentation. Finally, PbWO<sub>4</sub> has a quick scintillation decay time which allows for read out times fast enough to deal with the signed 25 ns interaction spacing of the LHC.

The ECAL barrel (EB) contains 61,200 PbWO<sub>4</sub> crystals arranged in a cylindrical formation around the silicon tracker. Each crystal is rectangular in shape, 240 mm long ( $\sim 26X_0$ ),  $22 \times 22$  mm<sup>2</sup> on the front surface and  $26 \times 26$  mm<sup>2</sup> on the near surface. The EB cover a pseudorapidity range  $|\eta| < 1.479$  with each crystal having a size in  $\eta - \phi$  space of  $0.0174 \times 0.0174$ .

The ECAL endcaps (EE+ and EE-) each contain 7,324 PbWO<sub>4</sub> crystals and cover pseudorapidity range of  $1.479 < |\eta| < 3.0$ . Each EE is split into an upper and lower disk, “dees” which arrange the crystals in an  $x - y$  pattern while situating the crystal face to point in the direction of the nominal interaction point. The crystals are approximately  $30 \times 30$  mm on their face with a length of 220 mm<sup>2</sup> ( $\sim 25X_0$ ).

In addition to the EB and EE, there is an additional electromagnetic calorimeter stationed on each side directly in front of the EE. Called the preshower (ES), this sub-detector gives additional radiation lengths and helps differentiate electrons and photons from pions. A full schematic of the ECAL system can be seen in Figure 2.6.



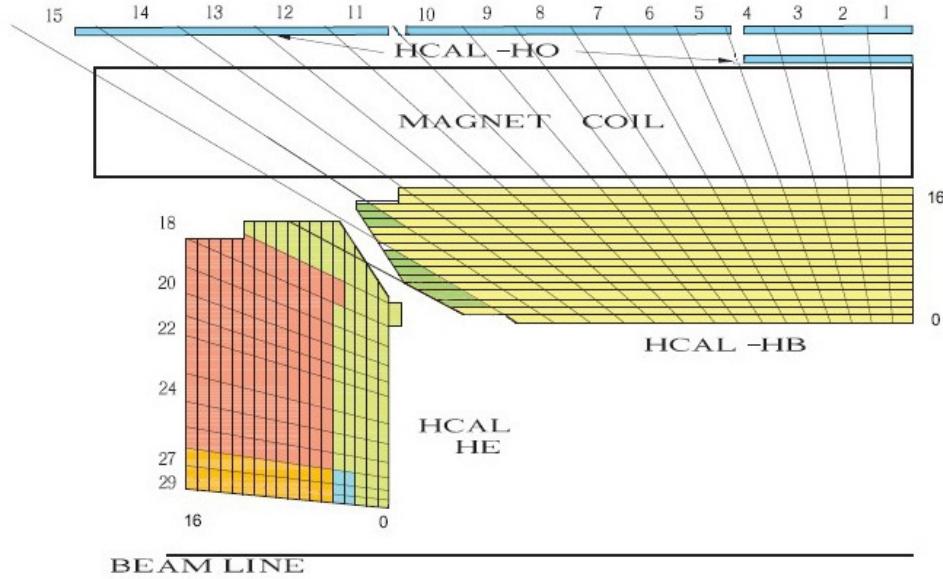
**Figure 2.6:** One quarter of the electromagnetic calorimeter system projected into the  $y - z$  plane [1].

To convert the light created in the crystals to an electrical signal, silicon avalanche diodes (APDs) and vacuum phototriodes (VPDs) are used in the barrel and endcap, re-

spectively. Both photo-detectors were designed to be fast, have a high radiation tolerance and be able to operate in the presence of CMS's strong magnetic field. In addition, the photo-detectors have to be able to withstand the bombardment of hadronic particles passing through their active area without too much disruption and to be able to amplify the low light output of the  $\text{PbWO}_4$  crystals.

### Hadronic Calorimeter

The HCAL sits between the ECAL and the solenoid and is important for the identification and reconstruction of hadronic energy [2]. The HCAL is comprised of four different sets of sub-detectors, the HCAL barrel (HB), the HCAL endcaps (HE), the outer HCAL (HO), and the forward HCAL (HF). A full schematic of the HCAL system can be seen in Figure 2.7.



**Figure 2.7:** One quarter of the hadronic calorimeter system projected into the  $y - z$  plane [2].

The HB and HE both use brass alloy (70% Cu, 30% Zn), which has an interaction length ( $\lambda_I$ ) of 15 cm. This ensures as much hadronic showering as possible while still falling within the spatial and monetary constraints. The HB covers the pseudorapidity

range of  $|\eta| < 1.3$  and contains 36 wedge shaped blocks of material (18 around the cylinder and each half of the HB). Each wedge consists of a flat brass absorber plate that runs parallel to the beam axis with plastic scintillators interspersed throughout, which convert the hadronizing energy into light for detection. The scintillators are distributed to give a spacial segmentation of  $0.087 \times 0.087$  in  $\eta - \phi$  space. The interaction length of the HB varies from 5.92 to  $10.6 \lambda_L$ , depending on the pseudorapidity. The HE takes over where the HB leaves off and covers a pseudorapidity range of  $1.3 < |\eta| < 3.0$ . The granularity in this region allows for positional measurements with a size of  $0.087 \times 0.087$  up to  $\eta < 1.6$  and  $0.17 \times 0.17$  beyond  $|\eta| > 1.6$ . The total depth of the HE is nine interaction lengths. Both the HB and HE utilize multipixel hybrid photodiodes (HPD), which operate well in strong magnetic fields, to convert the scintillator light to an electric signal.

Between  $3.0 < |\eta| < 5.0$  sits the HF which, instead of brass, uses steel to absorb the hadronic particles. Due to the very large particle fluxes in this region, scintillators were replaced by more radiation hard quartz fibers. These quartz fibers emit Cherenkov radiation which is then channeled to the readout photodiodes. The segmentation of the HF gives the approximate size of each readout channel as  $0.18 \times 0.18$  in  $\eta - \phi$  space.

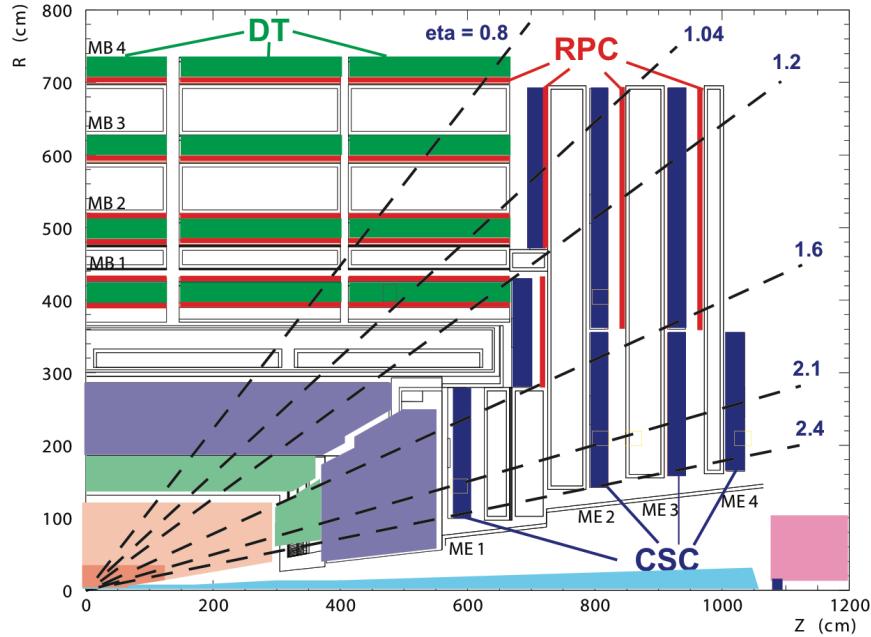
Finally, to supplement the energy measurement in the barrel, the region with the lowest interaction length, the HO sits outside the solenoid. In fact, it uses the solenoid as an extra absorber material, which gives at least 1.5 interaction lengths, depending on the pseudorapidity. The HO utilizes scintillator material directly behind the solenoid to measure the energy of any particle which hadronizes inside the solenoid. With the ECAL giving approximately one interaction length, the total interaction length of CMS is always greater than 11.8, giving  $< 0.1\%$  chance of a hadron punching through to the muon detectors.

## 2.2.4 The Muon Detectors

The Compact Muon Solenoid was designed to efficiently and precisely detect muons. Muons, unlike the much lighter and radiative electrons, interact very minimally with matter, traversing even the densest parts of CMS (the ECAL and HCAL) with relative ease while at the same time depositing very little of its energy. Muons are also

relatively long lived and their lifetime allows them to be measured by the CMS detector before it decays. It is for these reasons that the muon detectors are situated outside the solenoid, far from the interaction point.

The muon system employs three different gas-based particle detectors which allow for excellent muon measuring capabilities. In addition to positional measurements, the muon detectors also provide muon identification, momentum and triggering (see Section 2.4). The chambers sit inside CMS's iron return yoke which acts as an absorbing material for hadrons and other non-muon particles which try to penetrate through the solenoid. In the barrel, the drift tube (DT) sub-detector sits cylindrically concentric. The endcap houses cathode strip chambers (CSC) which consist of four layers on each side of the barrel. Finally, interspersed through the barrel and endcap, is the resistive plate chambers (RPC) used to enhance the timing resolution of the detection of muons. A graphical overview of the CMS muon system is shown in Figure 2.8.



**Figure 2.8:** Overview of the CMS muon sub-detector system. The drift tubes, situated in the barrel region, are shown in green, the cathode strip chambers, situated in the endcap region, are shown in blue and the resistive plate chambers, situated throughout, are shown in red [30].

## Drift Tubes

Drift tubes are long cylindrical tubes filled with gas and threaded with an equally long conductive wire (the anode). The inside of the cylinder is also lined with a conductive material (the cathode). The gas, a mixture of noble gas with full set of valence electrons, relinquishes electrons as charged particles traverse through it. A bias voltage is applied across the wire and the cylinder causing the electrons to flow one way and particle detection is achieved through measuring the amount of induced current.

The DT system sits in the barrel region outside the solenoid and covers a pseudorapidity range  $|\eta| < 1.2$ . Each set of drift tubes are arranged as a rectangular chamber and staggered throughout the return yoke. There are 250 chambers arranged in four concentric cylinders of chambers (stations) divided by five rings (wheels) of the return yoke. The chambers have a slight overlap to ensure the hermeticity in 12 different sectors of each ring and wheel, each covering  $30^\circ$  in azimuthal angle. The four stations and two and a half of the wheels can be seen in green in Figure 2.8.

Each chamber except the last has three sets (superlayers) of four planes of drift tubes. Two of the superlayers in each chamber run parallel to the beam direction to measure the  $r - \phi$  position. The third superlayer in the first three chambers have drift tubes that run perpendicular to the beam line allowing an additional measurement in the z-direction.

In all, there are over 172,000 drift tubes in the DT system each with an approximate length of 2.5 m. This length, coupled with the low background and muon rates and the uniform magnetic field in this region of the detector allow for the one-dimensional position measurements of the DT system. The DT's give a time resolution of  $\sim 25$  ns, well within the timing constraints of the LHC, and a global resolution in the  $r - \phi$  direction of  $100\ \mu\text{m}$ .

## Cathode Strip Chambers

A cathode strip chamber is similar to a drift tube chamber in that it is a gaseous detector filled with anode wires; however, instead of a wire tube per cathode tube, the anode wires are stretched in parallel and are surrounded on either side by a plan of cathode material. One plane of the cathode is segmented into strips which run perpendicular

to the anode wires. The induced charge distribution on the wires and strips gives a fully two-dimensional measurement of a passing particle (in addition to the third known position given by the position of the chamber). In addition to giving a three-dimensional positional measurements, the CSCs were also designed with a fast response time, fine segmentation and to be radiation hard in order to deal with increased background and muon rates and the non-uniform magnetic field in the endcap region (as opposed to the lower rates in the barrel where positional ambiguity can be tolerated).

The CSCs operate in the pseudorapidity region  $0.9 < |\eta| < 2.4$ . Similar to the barrel, there are four stations in each endcap region, arranged in radial disks perpendicular to the beam line. There are 468 chambers in total, each trapezoidal in shape and arranged side-by-side circularly around the beam line in two concentric circles. The anode wires run along the azimuthal direction and provide the radial measurement while the strips run lengthwise in a radial direction and give a measurement in  $\phi$ . The overall positional resolution of the CSC is between 75 and 150  $\mu\text{m}$ . The CSC system is shown in blue in Figure 2.8.

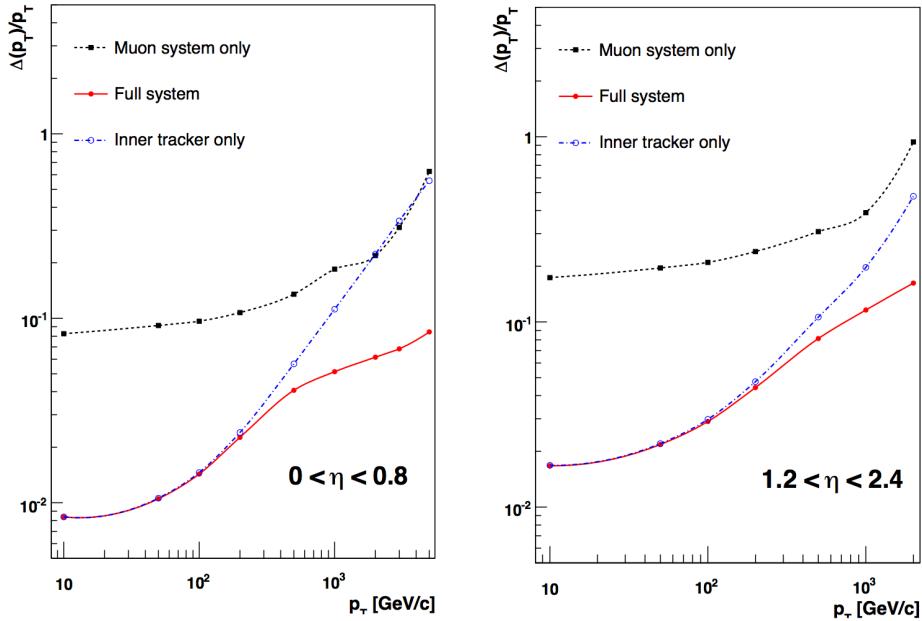
## Resistive Plate Chambers

The goal of the RPC muon detector system is to compliment the CSCs and DTs by adding faster readout times at the expense of position resolution. The RPC chambers consist of two gaseous regions separated by a common strip readout anode plane which gives adequate positional resolution. Operating with a time resolution well below the minimum of 25 ns, the RPCs are ideal detectors to be used to “trigger”, the CMS detector (see Section 2.4). The RPC system covers a pseudorapidity range of  $|\eta| < 1.6$ . Originally designed to cover the full range, it will be completed during the next upgrade. The RPC chambers can be seen in red in Figure 2.8. Running along the beam direction (and hence measuring the azimuthal angle), six (three) layers of chambers exist in the barrel (endcap) with over 480 chambers in total each approximately 2.5 meters in length.

## Summary

Overall, the muon system was designed to have an efficiency of over 95% for detecting muons with a resolution of less than 10% for muons up to  $p_T = 200\text{GeV}$ ,

increasing with resolution between 15% and 40% for 200 GeV to 1 TeV muons (depending on  $|\eta|$ ). Using the tracker as an additional set of measurements on the muon, the expected resolution improves to around 1% for lower  $p_T$  muons and down to approximately 5% for muon in the TeV range. The expected resolution on the transverse momentum measurement can be seen for two different pseudorapidity ranges in Figure 2.9.



**Figure 2.9:** Muon momentum resolution for two different pseudorapidity regions,  $0 < |\eta| < 0.8$  (left)  $1.2 < |\eta| < 2.4$  and (right), as a function of transverse momentum using only the muon system (black), only the inner tracker (blue) and the combined muon and tracker detectors (red) [1].

## 2.3 Integrated Luminosity Calculation

CMS uses two different techniques to measure the instantaneous luminosity, and subsequently the total integrated luminosity [31, 32, 33]. Both methods employ the use of Van der Meer scans [34], modulation of the two beam positions until maximum overlap is achieved to determine the maximum instantaneous luminosity.

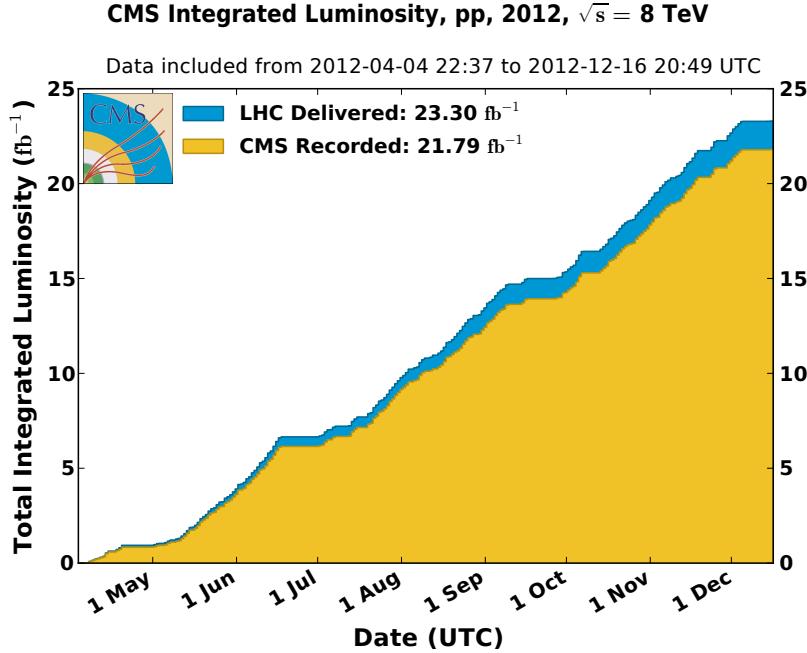
The first method uses HF calorimeters which are forward at  $|\eta| > 3.0$  and measures the average fraction of empty calorimeter readouts when triggering on zero-bias events defined by random triggers which are completely agnostic of activities in the detector. This fraction is converted into a cross section measurement which can be used to measure the instantaneous luminosity [31].

The non-linear HF response as a function of the instantaneous luminosity, among other problems, led to the creation and use of the second method [32, 33]. This method relies on the fine granularity of the pixel detector. With a very small fraction ( $\sim 0.1\%$ ) of particles leaving deposits in the same pixel, the number of pixel clusters during a bunch crossing is found to be directly related to the proton-proton cross section, which can be related to the luminosity.

The data used in this analysis was taken over the course of 2012, during which time the LHC delivered approximately  $23.3 \text{ fb}^{-1}$ . Between April and December of that year, CMS was able to record  $21.8 \text{ fb}^{-1}$  of this data of which  $19.5 \text{ fb}^{-1}$  was certified by the CMS collaboration as usable for this analysis. The uncertainty from the luminosity measurement used for all analyses using 2012 data in CMS is 4.4%. This error must be taken into account for any process which is estimated from simulation. Figure 2.10 shows the amount of data delivered by the LHC and recorded by the CMS detector as a function of time in 2012.

## 2.4 Data Acquisition and Triggering

With nearly 80 million channels, and a bunch crossing every 25 ns having up to 40 proton-proton interactions per crossing, a huge amount of data is produced by the CMS detector when the LHC is running at or near design luminosity. At approximately 100 kilobytes of data per interaction, computing design constraints limit the number of interactions that can be written to disk to  $\sim 1000$  per second. The 25 ns between bunches give an interaction rate of 40 MHz resulting in a data reduction requirement of  $\sim 10^5$ . To accomplish this, CMS has designed a data acquisition (DAQ) and trigger system which utilizes a massive computing infrastructure and custom hardware. It is designed to keep the most interesting interactions across a broad range of physics goals. If a

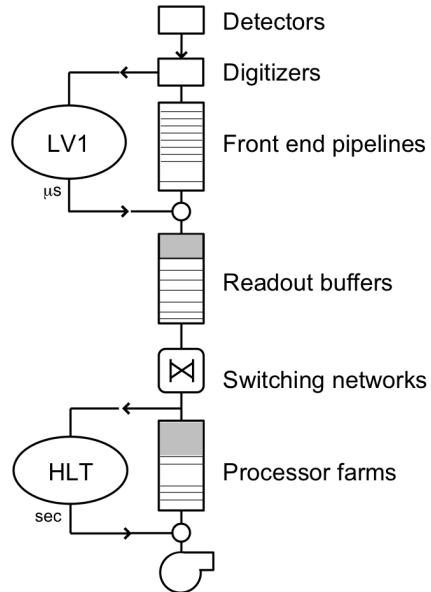


**Figure 2.10:** The amount of data delivered by the LHC (red) and the amount of data collected by the CMS detector (blue) as a function of date in the year 2012 [35].

proton-proton interaction (an event), doesn't pass the trigger system, it is lost forever.

Triggering occurs in two stages, the first of which is called Level 1 (L1) and has the task of reducing the load from 40 MHz down to 100 kHz [36]. To do so, it employs a farm of custom electronics using only coarse and lower resolution data to make decisions while storing the rest of the event data in pipelines awaiting processing. Here, only calorimetric and muon chamber information is available to make decisions based on the energy and quantity of depositions in the detector. The data from the events passing the L1 trigger are processed, compressed and zero-suppressed and sent for handling by the DAQ. The DAQ then stores and processes the information and sends it along to the second trigger phase. The High Level Trigger (HLT) is a farm of standard CPUs running parallel that can run a more complete event reconstruction to facilitate more informed decisions on whether to keep the event. The object reconstruction algorithms used are simplified version of the algorithms described in the following Section. A cartoon representation of the trigger and DAQ chain can be seen in Figure 2.11. Between

these two trigger systems, the desired data reduction is achieved.



**Figure 2.11:** Pictorial representation of the data acquisition system and the flow of data from the CMS detector through both trigger levels [3].

## 2.5 Particle Reconstruction and Identification

The digitization of the events selected by the trigger only constitute raw ingredients that need to be further processed in order to give a complete description of the particles present as the result of a proton-proton collision. Many algorithms have been developed by the CMS collaboration to reconstruct these particles; however, only the reconstruction algorithms used in this thesis will be described here with a full description given in the references. In this section we discuss charged particle, electron, muon, jet and missing transverse energy reconstruction; which are the important objects used in this analysis.

### 2.5.1 Charged Particle Reconstruction

Charged particles traversing the CMS detector first deposit energy in the CMS tracker, leaving thousands of positional measurements to interpret per bunch crossing. Determining the trajectory of these charged particles amongst this collection of “hits” is an exercise in pattern recognition. Each charged particle can be described using five parameters that model its trajectory as it is bent through the magnetic field. “Tracking” is the process of finding these “tracks” through the CMS tracking system [37, 38, 29].

Individual silicon strip and pixel channels with a threshold above the signal-to-noise ratio are clustered together based on proximity producing a collection of positional measurements (along with the associated uncertainty). These “hits” in the inner portion of the detector are used to seed multiple tracking steps, each designed to find tracks from charged particles with different properties. Each step takes an initial trajectory measurement from the seed and propagates the trajectory from the inside to the outside of the tracker using a combinatorial Kalman filter (CKF) [39]. After propagation ends, the tracks are measured for quality and hits on high quality tracks are removed from consideration before creating the seeds for the next tracking step. The tracks produced with each step are then merged into a single collection, eliminating duplicate tracks by comparing shared hits and keeping tracks with higher quality.

The seeding step searches for combinations of two or three hits near the beam line compatible with a particle trajectory with an energy above a given threshold coming from the interaction point. The initial particle trajectory is estimated using either a seed triplet or a seed pair and the beam spot. There are seven tracking steps. The thresholds in each step is lowered so that the initial steps have less background of seed candidates that lead to tracks reconstructed for non-existent particles (fakes). The first two steps are seeded using only triplets from the pixel detector, the second step dropping the  $p_T$  threshold after the earlier higher  $p_T$  step. The third step uses pairs of pixel hits to gain additional efficiency. The fourth step again uses pixel triplets, but with a looser requirement on the compatibility with the interaction point to search for displaced tracks and particles with longer lifetimes. The fifth step again uses triplets, but allows for combinations of hits in both the pixel and strip detectors. This allows for seeding of particles which have slightly longer decay lengths or have a missed pixel hit (due to

detector inefficiency). The final two steps use pairs and triplets of hits at further radii in the strip detector to search for particles which have a large displacement.

In each step, after the seeds have been created, the initial trajectory estimate is propagated inside-out, layer-by-layer through the detector using the CKM algorithm. The algorithm accounts for losses of energy as the particle traverse the material in the detector as well as the possibility of multiple scatter. At each layer, the position of the trajectory is estimated and any hits compatible with the position are added to the track. If more than one compatible hit is found, multiple trajectories are kept and further propagated. In addition, for each layer, a new trajectory is created assuming there is no hit in the current layer to allow for detector inefficiency. The process continues until either too many layers have been crossed without a hit or the trajectory reaches the last layer of the tracker.

After the trajectories have been propagated, a final trajectory fit to the collected hits is performed. Any outlying hits that fall too far from the overall fit are removed. This gives a final accurate measurement of the tracks' parameters. Vertex compatibility, number of hits in the track and the  $\chi^2$  of the fit are used as quality control selections. Hits from tracks passing tight quality selections are removed from consideration in subsequent tracking steps.

After all seven steps have been performed, the resulting tracks from each step are merged into a single collection. To ensure that no charged particles have been reconstructed twice, any tracks that share a large fraction of its hits are compared and only the best quality track is kept.

### 2.5.2 Vertex Reconstruction

In the high multiplicity environment of the LHC, it is essential to measure the number of proton-proton interactions for each bunch crossing. Typically, only one of the these interactions is responsible for the high  $p_T$  process that triggered the event, and measuring the lepton impact parameters with respect to the correct vertex is essential for the rejection of muons from semi-leptonic flavor decays and electrons from photon conversions. Thus, a high vertex reconstruction efficiency, a low fake vertex rate and the ability to differentiate nearby interactions are all requirements needed for CMS's vertex

reconstruction.

Vertex reconstruction is performed in two steps. First, tracks are clustered together using the deterministic annealing algorithm [40, 41]. The algorithm iteratively clusters tracks with nearby impact parameters first with large windows and allowing a single track to be clustered multiple times. Each iteration tightens the clustering window until a stopping condition is met. Tracks are not allowed to cluster to vertices if their transverse impact parameter is larger than 3 cm or the longitudinal impact parameter is larger than 4 cm. The deterministic annealing algorithm was specifically chosen due to its high clustering efficiency even when faced with the noisy LHC environment. Studies on the vertex efficiency show that the vertex reconstruction response as a function of the number of interactions is linear.

The second step takes each cluster of tracks and fits the vertex position using the adaptive vertex fitting algorithm [42, 43]. This algorithm weighs tracks in the cluster based on compatibility with the vertex position (and un-weights the tracks that are more incompatible) giving a good vertex resolution of less than 50  $\mu\text{m}$  depending on the number of tracks present in the cluster. More vertex reconstruction performance results can be found in Reference [44].

### 2.5.3 Electron Reconstruction

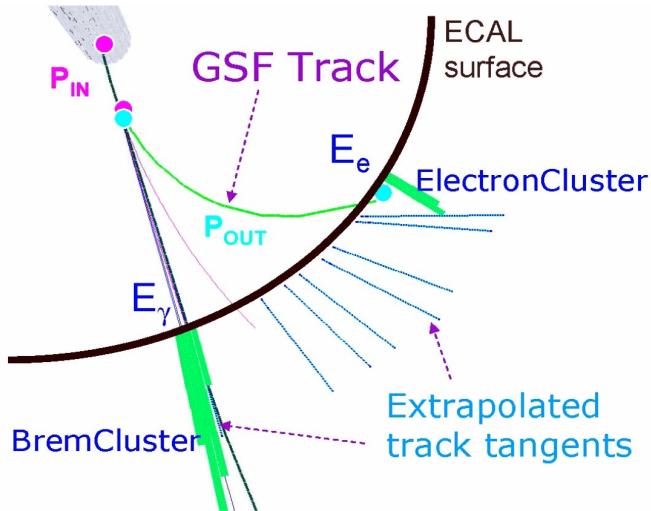
It is important to reconstruct and identify electrons efficiently and to know their energy to a high precision. Electrons interact primarily via the electromagnetic force and hence deposit a majority of their energy in the ECAL. However, due to the large amount of material the electron must traverse in the silicon tracker, the electron radiates a significant amount of energy via bremsstrahlung and hence the “footprint” of the electron’s energy distribution is very complex. In addition, other particles also leave a significant amount of energy in the ECAL, namely hadrons, in particular the  $\pi^0$ , which decays predominantly to two photons . Therefore, strong rejection of reconstructed electrons originating from sources other than true electrons from the hard interaction (fakes) is needed in the form of electron identification. The identification of electrons and rejection of fake electrons are left to be presented in the context of this analysis in Section 4.2.3 while the reconstruction algorithm is summarized below. More informa-

tion about electron reconstruction can be found in References [45, 46, 47].

The first step of electron reconstruction is to identify clusters of energy in the ECAL, i.e. areas where many crystals in close proximity have large energy deposits. As the electron travels through the silicon tracker, bremsstrahlung causes losses of energy through radiated photons which then also deposit energy in the ECAL. The strong magnetic field bends the electron in the azimuthal direction and it will radiate energy, in the form of photons, which will spread energy along  $\phi$  in the ECAL. Therefore the footprint of the electron is such that it is long in  $\phi$  and narrow in  $\eta$  and multiple deposits of energy need to be clustered together to form a “Super Cluster (SC)” containing the full amount of energy originally possessed by the electron.

After all SCs have been assembled, electron reconstruction continues by searching for the footprint of the electron in the tracker. Hits in the pixel detector compatible with the energy-weighted center of the SC are identified allowing for both positive and negative charge hypotheses. If two or three compatible pixel hits are discovered, they are used to seed a track building algorithm similar to that described in Section 2.5.1. The sole difference in the case of electrons is that track fitting is done with a Gaussian sum filter (GSF) [48] algorithm which can handle the large changes in electron trajectory due to bremsstrahlung.

In addition to this approach seeded by SCs, a separate algorithm is run in parallel. This is a “particle-flow” based approach which will be described in full in Section 2.5.5. Instead of first building a SC, the algorithm searches for pixels seeds for all clustered energy. For each found seed, the algorithm completes the trajectory building and then searches for clusters to add to the electron. At each layer in the tracker along the electron track, the position in the ECAL pointed to by the electron’s current momentum is used to search for additional compatible clusters that originate from a radiated photon. After both algorithms are complete, the two collections of electrons are merged. The SC seeded algorithm dominates at higher energies ( $> 20\text{ GeV}$ ) while the track-seeded algorithm adds significant efficiency at lower energies as well as in crowded environments such as near or inside a jet. A graphic showing the different pieces of the two algorithms can be seen in Figure 2.12.



**Figure 2.12:** Depiction of the different portions of the electron reconstruction algorithm. Here the electron radiated a photon with a significant amount of energy early in its traversal of the tracker and therefore there are two large clusters (depicted in green) of energy present in the ECAL (depicted in black) [3].

### 2.5.4 Muon Reconstruction

Muons, in contrast to electrons, penetrate through to the outer regions of the CMS detector due to their much higher mass and hence have a smaller background contamination. Objects mis-reconstructed as muons in this analysis come mostly from high energy hadrons which “punch-through” the dense calorimeters and leave deposits in the muon detector. Another source of background muons come from semi-leptonic decays of pions and kaons to real muons. Reconstruction efficiency for muons is quite high and we can therefore apply tight identification requirements to reduce these backgrounds for this analysis. This is discussed in Section 4.2.2. The muon reconstruction algorithm is discussed below, with more detailed information available in References [30, 49].

Muon reconstruction begins by building segments in the muon sub-detector. Positions of hits in the DT and CSCs are matched together to form small segments compatible with a single particle passing through each chamber. Using these segments (as well as hits in the RPCs) as positional measurements, tracking is performed starting from the inside of the muon chamber and working out to the outer edge of the muon system.

Here, the Kalman Filter is used to propagate and fit the muon tracks accounting for the loss of energy and multiple scattering as the muon passes through the iron yoke. These tracks built in the muon system are called “Stand-Alone” muons (SAM).

Since muons first passed through the CMS tracker, the SAMs are used to look for tracks built in the silicon tracker matching the expected trajectory and energy. Matched tracks are combined with the muon-only tracks and a global refit is performed (again with the Kalman Filter algorithm). The resulting muons are called “global muons.”

In addition, another muon reconstruction algorithm is used to build muons with information from both the tracker and the muon system. “Tracker” muons are seeded from the tracks built in the silicon tracker. Each track is extrapolated through both the ECAL and HCAL taking into account the expected trajectory and uncertainty based on the magnetic field and the amount of material it traverses. The energy deposited in the ECAL and HCAL at the expected position of the muon’s trajectory is checked to ensure compatibility with the expectation from the minimum ionizing particle. The expected muon trajectory is then extended into the muon system checking for matching segments. This algorithm gives higher efficiency for lower  $p_T$  muons at the expense of a larger background; however, muon identification selections can be applied for further suppression.

### 2.5.5 Particle Flow Reconstruction

The algorithms discussed up to this point have been the so-called “detector-based” algorithms. That is, reconstruction in each of the local sub-detectors drives the higher level reconstruction of particles that leave their signature in the respective sub-detectors. The “particle-flow” (PF) algorithm is a paradigm shift from this type of algorithm in that it uses the very fine segmentation of the CMS detector to search for an individual particle across all the sub-detectors. For example, a charged hadron will leave a track in the silicon tracker, a small amount of energy in the ECAL and will be deposited in the HCAL, dissipating most of its energy. PF is used to identify particles in this manner. When complete, the algorithm aims to have produced a list of all the particles produced during the collision.

One large disadvantage of detector-based methods is the fact that energy could

be double counted, while with the PF algorithm, a global reconstruction of the event is done on a particle-by-particle basis ensuring that double counted energy is minimized. The resulting list of particles produced by this algorithm is used extensively across many analyses at CMS. The calculation of missing transverse energy ( $\cancel{E}_T$ ), jet reconstruction, and  $\tau$  reconstruction are among the most important. The following briefly describes the particle flow algorithm; however, a complete description can be found in References [50, 51].

Before the PF reconstruction begins, local reconstruction in each of the sub-detectors is completed and provides tracks from the silicon detector (Section 2.5.1), clustered energy in both the ECAL and HCAL (similar to those discussed in Section 2.5.3) and the local reconstruction product in the muon system (Section 2.5.4). The general strategy is then to “link” these different inputs together to form the complete picture of particles traversing through the detector. Objects from each of the sub-detectors in close proximity are first grouped together in blocks. Each block usually contains a few different inputs from each of the detectors, i.e. a track pointing to an ECAL cluster, or an ECAL and HCAL energy cluster adjacent to each other. Any blocks linked to signatures from isolated muons or electrons are removed, as these have a very clean signature with very little background. The blocks belonging to these particles (including the clusters from radiated products as discussed in Section 2.5.3) are then removed from the list of unidentified blocks and added the list of PF muons and electrons. Next, the momentum of any track pointing to calorimetric clusters is compared to the energy contained in the ECAL or HCAL clusters. If the energy and momentum of the two are compatible, the object is labeled as a charged hadron, its energy is estimated as a weighted sum from both objects, and its constituents are removed from the list and added to the list of PF charged hadrons. If there is significantly more energy in the track than is deposited in the calorimeter, a secondary muon identification is performed searching for non-isolated lower  $p_T$  muons. If identified, again, these constituents are removed and added to the PF muon list. If no muon is found, tighter track requirements are applied to reject mis-reconstructed tracks. If there is significantly more energy in the calorimeters, neutral hadrons and photons are created comprising the energy unaccounted for by the track. At this point, only unlinked clusters remain and all ECAL clusters are hypothesized to

be photons with all HCAL clusters hypothesized to be neutral hadrons and added to their respective lists.

### 2.5.6 Calculation of Missing Transverse Energy

Neutrinos and other non-interacting particles from BSM theories, such as neutralinos from supersymmetry, leave no signature as they traverse the detector. Yet, we can infer their existence, transverse direction and energy by looking at the sum of all of the transverse energy in the event. As the two protons collide, all of their momentum is in the  $z$ -direction (and negative- $z$ ), with no momentum transverse to the beam line. Momentum conservation then requires that the transverse vectorial sum of all energy in the event should sum to zero. If this calculation is performed and an imbalance (missing transverse energy,  $\cancel{E}_T$ ) is detected, a non-interacting particle is hypothesized to have escaped from the detector.

The  $\cancel{E}_T$  reconstruction algorithm used by this analysis takes as input the full collection of particle flow candidates and produces the negative vectorial sum. Particle flow  $\cancel{E}_T$  is defined as

$$\cancel{E}_T = - \sum_i \mathbf{p}_T^i,$$

where the summation is over all particles reconstructed with the particle flow algorithm described in the previous section. Further information about  $\cancel{E}_T$  may be found in Reference [52].

### 2.5.7 Jet Reconstruction

Jets are the experimental signatures of quarks and gluons produced in high-energy processes such as proton-proton collisions. As quarks and gluons have a net color charge and cannot exist freely due to color-confinement, they are not directly observed in Nature. Instead, they come together to form color-neutral hadrons, a process called hadronization that leads to a collimated spray of hadrons called a jet.

Jet reconstruction is done by clustering nearby reconstructed PF candidates. Starting with the highest  $p_T$  candidates as seeds, clustering is done using the anti- $k_T$  algorithm [53] with a distance parameter of  $\Delta R = 0.5$ , defined in the  $\eta - \phi$  plan. This

clustering algorithm uses a distance measurement inversely proportional to the square of each particle’s  $p_T$  and thus gives stability in the cases of infrared or collinear radiation.

The use of PF candidates gives good jet energy response due to the high reconstruction efficiencies of charged hadrons and photons which make up about 90% of the jet’s energy. However, some corrections are still applied to the jet energy with these energy scale corrections factorized into three steps. First, an offset correction is applied to reduce the energy of the jet based on the amount of pile-up in the event. A pile-up density is calculated using the FastJet method [54] to compute mean energy of jets in the event and the ghost particle method for calculating jet areas [55].

Next, corrections for the detector response are made as a function of jet  $p_T$  and  $\eta$ . The first portion of this correction is applied based on measurements of simulated jet energies. The last correction is made to account for the differences in the true CMS response between the physical and simulated detector. The differences are measured by looking at jet energies in di-jet,  $\gamma + \text{jets}$  and  $Z + \text{jets}$  events where the object recoiling against the jet can be used to measure the true jet energy. The uncertainty on this method is less than 5%, smaller than the overall jet energy resolution which ranges from 8% to 15%, depending on the jet  $p_T$  and  $\eta$ . The remaining differences are explicitly corrected for in data in the residual correction step, which completes the jet energy correction chain. A full picture of jet reconstruction and energy corrections can be found in Reference [56].

### 2.5.8 b-Jet Identification

The identification of b-jets will be a major handle to select possible signal event. One of the design requirements of the tracker was to find displaced vertices of the long lived hadrons produced by b-quarks. Many algorithms were designed to identify displaced vertices present in b flavored jets (b-jets); however, we only discuss the one algorithm that is used in this analysis, the Combined Secondary Vertex method (CSV) [57].

The presence of a secondary vertex and the kinematic variables associated with this vertex can be used to discriminate between a b and non-b jet. Two of these variables are the distance and direction between the primary and secondary vertices (the flight distance and direction). The significance of the flight distance (the ratio of the flight

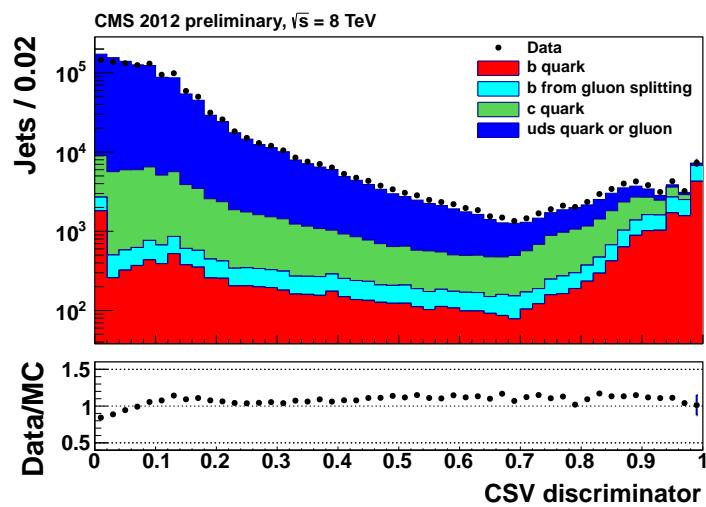
distance to its estimated uncertainty) can be used as a discriminating variable. In the CSV algorithm, the flight distance information is combined with track-based lifetime information. By using this additional information, the algorithm provides discrimination also in cases where no secondary vertices are found, increasing the efficiency with respect to using the secondary vertex information alone. The vertex category is defined as

- real: There exists a secondary vertex.
- pseudo: When no real vertex is found, tracks with a high impact parameter significance ( $S_{IP}$ , the impact parameter divided by its uncertainty) are used to create a “pseudo” vertex.
- no vertex: When no real or pseudo vertex is found.

In addition to the vertex category, the following variables are used to form a discriminating variable (in the “no vertex” category, only the last two are used):

- the flight distance significance in the transverse plane (“2D”);
- the vertex mass;
- the number of tracks at the vertex;
- the ratio of the energy carried by the tracks at the vertex with respect to all tracks in the jet;
- the pseudorapidity of the tracks at the vertex with respect to the jet axis;
- the 2D  $S_{IP}$  of the first track that raises the invariant mass above the charm threshold of  $1.5 \text{ GeV}/c^2$ ;
- the number of tracks in the jet;
- the 3D  $S_{IP}$  for each track in the jet.

A likelihood ratio is built from these variables. It is used to discriminate between b and c jets and between b and light-parton jets. The distribution of the CSV discriminator is shown in Figure 2.13. In this analysis, we use the “medium” working point, which is any jet with a discriminator value greater than 0.679 is “tagged” as a b-jet.



**Figure 2.13:** Combined Secondary Vertex algorithm discriminant for data (black points) and compared to simulation of lighter flavor jets (green and blue) as well as b-jets (cyan and red). Larger values of the discriminant are more indicative of heavy flavor jets. The ratio between data and simulation is shown at bottom [57].

# Chapter 3

## Same-Sign Dilepton Signature

As discussed in Section 1.3, there are many models that predict same-sign dileptons, hadronic jets, and  $\cancel{E}_T$ . In order to determine if there is an excess of events, an accurate prediction of the expected background to this signature must be performed. In this chapter, a brief description will be given for the SM background processes that give genuine same-sign dileptons that are both prompt and isolated (rare SM). These constitute an irreducible background to the analysis that must be accounted for. These processes are discussed in Section 3.1.

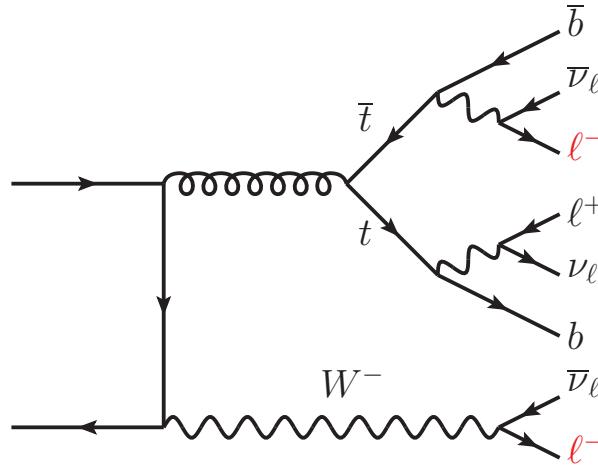
The other two major sources of background include objects mis-identified as selected leptons (fakes) and leptons with a mis-identified charge and from genuine opposite-sign pairs (charge flips). These backgrounds sources are a result of the failure of the selection to properly classify these events and these sources are discussed in Section 3.2.

### 3.1 Rare Standard Model Processes

As discussed previously, the main source of irreducible background is processes that produced genuine prompt and isolated same-sign leptons in the final state. In this section, we provide a brief description of the main processes that the SM predicts and are included in this analysis.

### 3.1.1 Top Processes

Processes involving top production with an associated gauge boson are a significant source of same-sign dileptons. This can be seen by the example Feynman diagram shown in Figure 3.1. Top processes that are considered in this analysis are listed



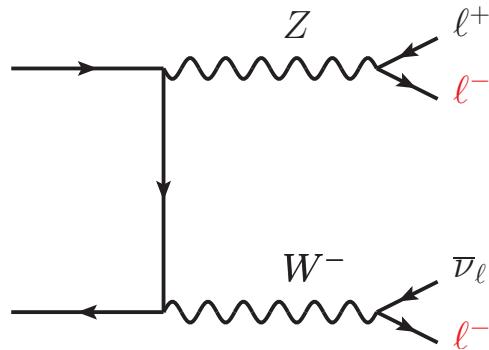
**Figure 3.1:** Example leading order Feynman diagram for  $pp \rightarrow t\bar{t}W \rightarrow 3\ell\nu + X$ .

below:

- $t\bar{t}W$ : The  $W$  decays leptonically and the top quark with the same charge as the  $W$  also decays leptonically ( $t \rightarrow \ell^\pm \nu b$  and  $W^\pm \rightarrow \ell^\pm \nu$ ).
- $t\bar{t}Z$ : The  $Z$  decays leptonically and at least one of the top quarks also decays leptonically ( $t \rightarrow \ell^\pm \nu b$  and  $Z \rightarrow \ell^+ \ell^-$ ).
- $t\bar{t}\gamma$ : The  $\gamma$  decays leptonically and at least one of the top quarks also decays leptonically ( $t \rightarrow \ell^\pm \nu b$  and  $\gamma \rightarrow \ell^+ \ell^-$ ).
- $t\bar{t}Z$ : The  $Z$  decays leptonically and the top quarks also decays leptonically ( $t\bar{b} \rightarrow \ell^- \nu \bar{b}$  and  $Z \rightarrow \ell^+ \ell^-$ ).
- $t\bar{t}WW$ : At least one of the  $W$ 's decay leptonically and one of the top quarks with the same sign also decays leptonically ( $t \rightarrow \ell^\pm \nu b$  and  $W^\pm \rightarrow \ell^\pm \nu$ ).

### 3.1.2 Two Gauge Boson Processes

Processes involving the production of two electroweak bosons in the same event can easily produce same-sign leptons. What makes them rare is their actual production is quite low due to the mass of these bosons. A typical example can be seen from the Feynman diagram shown in Figure 3.2 below: Processes than involve diboson produc-



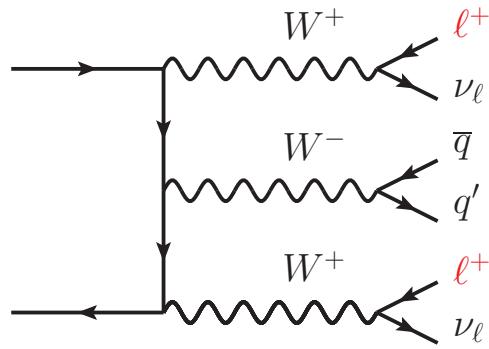
**Figure 3.2:** Example leading order Feynman diagram for  $pp \rightarrow WZ \rightarrow 3\ell\nu + X$ .

tion considered in this analysis are listed below:

- $W\gamma$ : The  $W$  and  $\gamma$  decay leptonically ( $W^\pm \rightarrow \ell^\pm \nu$  and  $\gamma \rightarrow \ell^+ \ell^-$ ).
- $WZ$ : The  $W$  and  $Z$  decay leptonically ( $W^\pm \rightarrow \ell^\pm \nu$  and  $Z \rightarrow \ell^+ \ell^-$ ).
- $ZZ$ : Both  $Z$ 's decay leptonically ( $2 \times Z \rightarrow \ell^+ \ell^-$ ).

### 3.1.3 Triple Gauge Boson Processes

Processes involving the production of three electroweak bosons in the same event can easily produce same-sign leptons. Producing three gauge boson is considerably suppressed with respect to processes with two bosons since this requires even more energy to produce due to the additional mass of the extra boson. A typical example can be seen from the Feynman diagram shown in Figure 3.3. Processes that involve triboson production considered in this analysis are listed below:

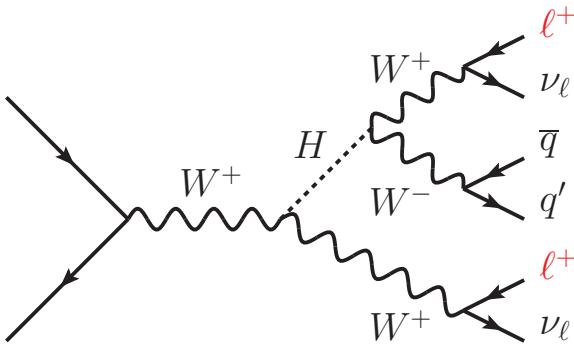


**Figure 3.3:** Example leading order Feynman diagram for  $pp \rightarrow WWW$ .

- $WW\gamma$ : One or both  $W$  bosons decay leptonically and the  $\gamma$  decays leptonically ( $W^\pm \rightarrow \ell^\pm \bar{\nu}_\ell$  and  $\gamma \rightarrow \ell^+ \ell^-$ ).
- $WWW$ : Two of the  $W$  bosons with the same sign decay leptonically ( $2 \times W^+ \rightarrow \ell^+ \bar{\nu}_\ell$  or the charge conjugate).
- $WWZ$ : One of the  $W$  bosons and the  $Z$  boson decay leptonically ( $W^\pm \rightarrow \ell^\pm \bar{\nu}_\ell$  and  $Z \rightarrow \ell^+ \ell^-$ ).
- $WZZ$ : Either the  $W$  boson and at least one of the  $Z$  bosons decay leptonically or both  $Z$  bosons decay leptonically ( $W^\pm \rightarrow \ell^\pm \bar{\nu}_\ell$  and at least one  $Z \rightarrow \ell^+ \ell^-$  or  $2 \times Z \rightarrow \ell^+ \ell^-$ ).
- $ZZZ$ : At least two of the  $Z$  bosons decay leptonically ( $2$  or  $3 \times Z \rightarrow \ell^+ \ell^-$ ).

### 3.1.4 Higgs Boson Processes

Processes involving the Higgs boson production in conjunction with a  $W$  or  $Z$  boson or a top quark pair. The Higgs can decay to  $W$  or  $Z$  bosons or tau leptons which can lead to same sign dileptons. A typical example can be seen from the Feynman diagram shown in Figure 3.4 below: Processes than involve associated Higgs Boson production considered in this analysis are listed below:



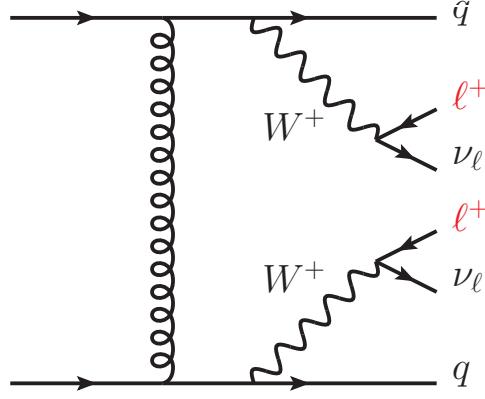
**Figure 3.4:** Example leading order Feynman diagram for  $pp \rightarrow WH; H \rightarrow W^+W^-$ .

- $WH, ZH, t\bar{t}H; H \rightarrow WW$ : One lepton comes from the decay of the associated  $W$ ,  $Z$  or  $t\bar{t}$  pair. The other lepton comes from a leptonically decaying  $W$  boson.
- $WH, ZH, t\bar{t}H; H \rightarrow ZZ$ : One lepton comes from the decay of the associated  $W$ ,  $Z$  or  $t\bar{t}$  pair. The other lepton comes from a leptonically decaying  $Z$  boson.
- $WH, ZH, t\bar{t}H; H \rightarrow \tau\tau$ : One lepton comes from the decay of the associated  $W$ ,  $Z$  or  $t\bar{t}$  pair. The other lepton comes from a leptonically decaying  $\tau$  lepton.

### 3.1.5 Same-Sign W Boson Pair Production

Processes involving the production of two bosons with the same electric charge where both  $W$ 's decay leptonically will also produce same-sign dileptons. The two processes we consider for this situation include double parton scattering (DPS) and  $qqW^\pm W^\pm$ . Processes than involve same-sign  $W$  pair production considered in this analysis are listed below:

- $qqW^\pm W^\pm$ : Two same signed quarks are produced and each radiates a  $W$  boson which both decay leptonically ( $pp \rightarrow qqWW$ ).
- $W^\pm W^\pm(DPS)$ : Two independent hard parton collisions that both produce  $W$  bosons which both decay leptonically ( $pp \rightarrow 2 \times W \rightarrow \ell \bar{\nu}_\ell$ ).



**Figure 3.5:** Example leading order Feynman diagram for  $qqW^\pm W^\mp \rightarrow \ell^\pm \nu_\ell \ell^\pm \nu_\ell + qq$ .

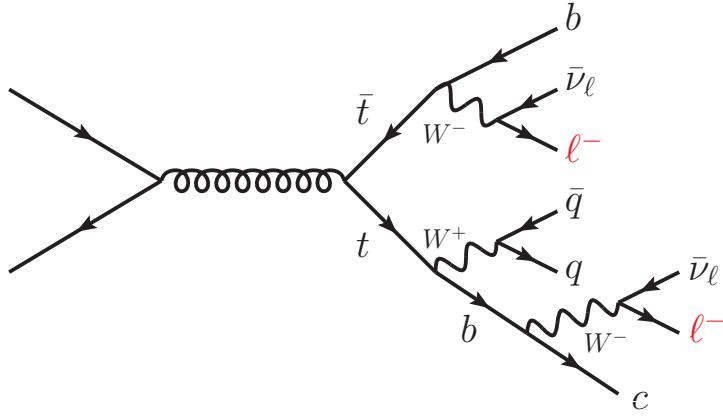
## 3.2 Other Sources of Backgrounds

In the previous subsection, we discussed SM process that give rise to events that possess genuine same-sign dileptons in the final state. In an ideal scenario, this is all we would need to consider; however, there are two other major sources of background that are caused by our inability to select the correct events with 100% certainty. The first of these is the non-negligible situation when an electron or muon is reconstructed and mis-identified as a prompt and isolated lepton. The second situation is less significant than the first but nevertheless is still important and is accounted for. This is the case where a real electron is reconstructed and selected; however, its electric charge is mis-reconstructed. The following subsections will give a brief discussion of these two cases, respectively.

### 3.2.1 Fake Leptons

The dominant source of same-sign dileptons are events where one lepton is from  $W \rightarrow \ell \bar{\nu}_\ell$  and the other originates from a semi-leptonic b-quark decay. We refer to leptons that originate from a electroweak boson or something with similar properties (such as a super-symmetric object) as “real leptons” since these give rise to true prompt and isolated leptons. For leptons that originate from other sources, we refer to this as a “fake lepton” since this is not the lepton we are interested in.

By far, the dominant source of fake leptons is  $t\bar{t}$  events. A representative diagram of this sort of event can be seen in the figure below (Figure 3.6). Here we get a same-sign



**Figure 3.6:** Diagram for  $t\bar{t}$  decays giving rise to same-sign dileptons final state.

dilepton final state by having one of the leptons from a top quark decaying leptonically from  $W \rightarrow \ell\bar{\nu}_\ell$  and the other from a “fake” lepton from the semi-leptonic b-quark decay from the other top [58]. Another important but sub-dominant source of fake lepton background is from  $pp \rightarrow W \rightarrow \ell\bar{\nu}_\ell + \text{jets}$  events where the jet decays semi-leptonically.

The actual rate that the fake lepton will pass your analysis selection is dependent on the details of the analysis. In Section 5.1, we give a prediction on the number of same-sign dilepton event containing these fake leptons.

### 3.2.2 Charge Flips

The measurement of a lepton’s charge is determined from the curvature of the reconstructed track as it travels through the magnetic field. The rate at which the charge is mis-measured is negligibly small for muons at the W/Z momentum scale; however, for electrons it is significant. Bremsstrahlung is emitted as electrons are deflected by the silicon-based material of the track. A radiated photon has a significant probability to convert into an  $e^+e^-$  pair due to the relatively large amount of material in the tracker [1]. The combination of the radiation of a hard photon and a subsequent asymmetric conversion (i.e. when one of the two resulting electrons carries a significant fraction of the

momentum of the original photon) can result in the reconstruction of a track that combines hits from the original electron with those from the converted photon. As a result, the reconstructed track may have a curvature opposite to that of the prompt electron, resulting in an incorrect charge assignment.

We estimate the rate that an electron that passes the analysis selection but has an incorrect charge measurement to be  $\sim 0.01\%$  [59]. The actual rate has been measured and will be fully discussed in Section 5.2.

# Chapter 4

## Analysis Selections

In Chapter 3, we discussed the motivation for selecting same-sign dilepton pairs. In this Chapter, we discuss the specific selections used to carry out the analysis. We consider a number of final states characterized by the number of jets, number of b-tagged jets, scalar sum  $p_T$  of selected jets ( $H_T$ ) and the missing transverse energy ( $\cancel{E}_T$ ). To provide coverage for a wide-range of generic signatures, we provide separate results for different lepton  $p_T$  requirements. Specifically, we perform a high  $p_T$  analysis where we select leptons with a  $p_T$  requirement of 20 GeV (high  $p_T$  analysis) as is common for leptons from  $W/Z$  and in many scenarios of stop/sbottom production. We separately consider dilepton pairs where the  $p_T$  requirement is lowered to 10 GeV (low  $p_T$  analysis). This slightly looser selection provides sensitivity to scenarios with a (partially) compressed spectrum, where at least one lepton is expected to come from the decay of an off-shell  $W/Z$ . Finally, we define search region in bins of # jets, # b-tagged jets ,  $H_T$  and  $\cancel{E}_T$  in order to improve statistical sensitivity.

This chapter begins with a discussion on the collision and simulated data used for this thesis. This is followed by the definitions of the selections for all the physics objects used; and finally, this chapter concludes with the discussion on the various search regions used in the analysis.

## 4.1 Data Samples

### 4.1.1 Collision Data

This analysis uses data collected during the 2012 run at  $\sqrt{s} = 8$  TeV. Dilepton events are selected from primary datasets collected using electron and muon trigger paths. Only runs and luminosity sections from good data taking periods are used, where such periods are defined using the flags delivered by the CMS data quality monitoring, detector, and physics validation teams. The datasets used in this analysis are listed in Tables 4.1, and 4.2 with the integrated luminosity corresponding to  $19.5\text{ fb}^{-1}$ .

**Table 4.1:** Primary datasets used by the high  $p_T$  analysis along with the relevant run-ranges.

Name	Run Range
/DoubleMu/Run2012A-13Jul2012-v1/AOD	190456-193621
/DoubleMu/Run2012B-13Jul2012-v1/AOD	193834-196531
/DoubleMu/Run2012A-recover-06Aug2012-v1/AOD	190949 190945 190906 190895 190782
/DoubleMu/Run2012C-24Aug2012-v1/AOD	198022-198523
/DoubleMu/Run2012C-PromptReco-v2/AOD	198934-203755
/DoubleMu/Run2012D-PromptReco-v1/AOD	203773-208913
/DoubleElectron/Run2012A-13Jul2012-v1/AOD	190456-193621
/DoubleElectron/Run2012B-13Jul2012-v1/AOD	193834-196531
/DoubleElectron/Run2012A-recover-06Aug2012-v1/AOD	190949 190945 190906 190895 190782
/DoubleElectron/Run2012C-24Aug2012-v1/AOD	198022-198523
/DoubleElectron/Run2012C-PromptReco-v2/AOD	198934-203755
/DoubleElectron/Run2012D-PromptReco-v1/AOD	203773-208913
/MuEG/Run2012A-13Jul2012-v1/AOD	190456-193621
/MuEG/Run2012B-13Jul2012-v1/AOD	193834-196531
/MuEG/Run2012A-recover-06Aug2012-v1/AOD	190949 190945 190906 190895 190782
/MuEG/Run2012C-24Aug2012-v1/AOD	198022-198523
/MuEG/Run2012C-PromptReco-v2/AOD	198934-203755
/MuEG/Run2012D-PromptReco-v1/AOD	203773-208913

### 4.1.2 Simulated Data

The analysis uses simulated data samples for candidate signal models and the relevant backgrounds. The contributions for real same-sign dilepton events are estimated directly from simulated data. These “rare” processes were discussed in Chapter 3.1 and constitute an irreducible background for this analysis. Simulated samples of  $Z/\gamma^*$ ,

**Table 4.2:** Primary datasets used by the low  $p_T$  analysis along with the relevant run-ranges.

Name	Run Range
/MuHad/Run2012A-13Jul2012-v1/AOD	190456-193621
/MuHad/Run2012B-13Jul2012-v1/AOD	193834-196531
/MuHad/Run2012A-recover-06Aug2012-v1/AOD	190949 190945 190906 190895 190782
/MuHad/Run2012C-24Aug2012-v1/AOD	198022-198523
/MuHad/Run2012C-PromptReco-v2/AOD	198934-203755
/MuHad/Run2012D-PromptReco-v1/AOD	203773-208913
/ElectronHad/Run2012A-13Jul2012-v1/AOD	190456-193621
/ElectronHad/Run2012B-13Jul2012-v1/AOD	193834-196531
/ElectronHad/Run2012A-recover-06Aug2012-v1/AOD	190949 190945 190906 190895 190782
/ElectronHad/Run2012C-24Aug2012-v1/AOD	198022-198523
/ElectronHad/Run2012C-PromptReco-v2/AOD	198934-203755
/ElectronHad/Run2012D-PromptReco-v1/AOD	203773-208913

$W + \text{jets}$ , and  $t\bar{t}$  processes are used to cross-check the partially data-driven estimate of the background contribution from charge mis-reconstruction, as well as systematic uncertainties of the lepton selection. The data-driven technique developed to estimate the background from fake leptons is applied to simulated  $t\bar{t}$  and  $W + \text{jets}$  samples as part of the systematic uncertainty studies on these methods. Simulated SUSY samples are used for the study of the systematic uncertainties as well as to develop a model of the selection efficiencies. These samples are also used in the extraction of upper limits on the observed and expected cross section for these models. The contribution from double fakes from QCD processes is determined using a data-driven method, but the available simulated samples are used to study some dependencies of the fake lepton prediction method. Simulated data samples are normalized to an integrated luminosity of  $19.5 \text{ fb}^{-1}$ , unless otherwise stated. Details of the simulated samples can be found in Tables 4.3 and 4.4.

## 4.2 Event Selections

The event selection consists of the following general requirements with amplifying details in the following sub-sections:

- we require a dilepton trigger;

**Table 4.3:** Sources of true same-sign dileptons from Standard Model processes. Cross sections are next-to-leading order. The equivalent integrated luminosity of these simulated events is listed in the column on the far right. The contributions from these process to the background is taken directly from simulation.

Sample	Cross Section (pb)	Equivalent Luminosity ( $\text{fb}^{-1}$ )
$t\bar{t}Z$	0.208	1021
$t\bar{t}W$	0.232	845
$t\bar{t}WW$	0.00204	106931
$t\bar{t}\gamma$	2.17	33
$t\bar{b}Z$	0.0114	13026
$ZZZ$	0.00554	40692
$WWW$	0.0822	2737
$WW\gamma$	0.528	407
$WZZ$	0.0192	12946
$WWZ$	0.0580	3832
$ZZ$	0.177	27177
$WZ$	1.058	1908
$qqW^-W^-$	0.0889	1084
$qqW^+W^+$	0.248	402
$W^\pm W^\pm(\text{DPS})$	0.588	1418
$W\gamma^* \rightarrow \ell\nu\mu\mu$	1.91	156
$W\gamma^* \rightarrow \ell\nu\tau\tau$	0.336	148
$WH, ZH, t\bar{t}H; H \rightarrow WW$	0.260	769
$WH, ZH, t\bar{t}H; H \rightarrow ZZ$	0.0320	15652
$WH, ZH, t\bar{t}H; H \rightarrow \tau\tau$	0.0177	5478

- we select events with two good same sign isolated leptons; the high  $p_T$  analysis requires both leptons to have  $p_T > 20$  GeV and the low  $p_T$  analysis requires the leptons to have  $p_T > 10$  GeV.
- $m_{\ell\ell} > 8$  GeV to reject low mass resonances.
- For most search regions, a significant amount of  $\cancel{E}_T$  to reduce SM backgrounds, particularly that from charge mis-measurement in  $Z \rightarrow \ell\ell$  events.
- We require a significant amount of hadronic energy to reduce standard model backgrounds, particularly mis-identified leptons in  $W + \text{jets}$  events.

**Table 4.4:** Sources of “fake” same-sign dileptons from Standard Model processes. Cross sections are next-to-leading order. The equivalent integrated luminosity of these simulated events is listed in the column on the far right. The contributions from these process to the background is taken directly from simulation.

Sample	Cross Section (pb)	Equivalent Luminosity ( $\text{fb}^{-1}$ )
$t\bar{t}$	225	30
$t\bar{t} \rightarrow \ell\ell X$	25	493
$t\bar{t} \rightarrow \ell(q \rightarrow \ell)X$	103	248
$t\bar{t} \rightarrow \text{hadronic}$	107	292
$t, s\text{-channel}$	3.89	69
$\bar{t}, s\text{-channel}$	1.76	80
$t, s\text{-channel}$	55.5	66
$\bar{t}, s\text{-channel}$	30.0	63
$tW$	11.2	45
$overline{t}W$	11.2	44
$Z/\gamma^* \rightarrow \ell\ell$	3533	8.62
$W + \text{jets}$	37509	varies
$WW$	5.81	332

- We veto events if the invariant mass of either hypothesis lepton and a third good lepton in the event is consistent with the Z mass. This requirement is chosen to suppress the irreducible background from  $WZ$  and  $ZZ$  events.
- We veto events if the invariant mass of either hypothesis lepton and a third good lepton in the event is  $< 12$  GeV. This requirement is chosen to suppress the irreducible background from  $\gamma^*$  and low mass resonances.

To reject known machine background resulting in high activity in the pixel layers, we require that no fewer than 25% of the tracks in event are high purity in events having 10 or more tracks. Also, to ensure that there was a reconstructed collision, we require that at least one good primary vertex (PV). A good PV is selected by requiring:

- the PV is not considered a fake,
- the number of degrees of freedom  $> 4$ ,

- $|\rho| < 2$  cm,
- $|z| < 24$  cm.

### 4.2.1 Trigger Selection

We use a handful of double lepton trigger paths to select events in data. An event in the  $ee$  final state is required to pass the relevant double electron trigger, a  $\mu\mu$  event must pass the double muon trigger and an  $e\mu$  event is required to pass one of the electron-muon cross triggers. Because of the rapidly changing trigger menu, many trigger paths were not implemented in the simulation and thus no trigger requirement is made here. Instead, as discussed in Section 6.2, a weight is applied to each event, based on the trigger efficiencies measured from data. A list of trigger paths used to select signal-like events can be found in the following table:

**Table 4.5:** The trigger paths used for the high  $p_T$  and low  $p_T$  analyses. Note that for the high  $p_T$  triggers, XXLstr stands for CaloIdT\_CaloIsoVL\_TrkIdVL\_TrkIsoVL which is explained in the text below. For the low  $p_T$  triggers, PFHT175 refers to particle flow based  $H_T > 175$  GeV.

Analysis Type	Channel	Trigger
high $p_T$	$ee$	HLT_Ele17_XXLstr_Ele8_XXLstr
	$e\mu$	HLT_Mu17_Ele8_XXLstr or HLT_Mu8_Ele17_XXLstr
	$\mu\mu$	HLT_Mu17_Mu8
low $p_T$	$ee$	HLT_DoubleEle8_CaloIdT_TrkIdVL_Mass8_PFHT175
	$e\mu$	HLT_Mu8_Ele8_CaloIdT_TrkIdVL_Mass8_PFHT175
	$\mu\mu$	HLT_DoubleMu8_Mass8_PFHT175

The high  $p_T$  analysis selects events with two leptons having  $p_T > 20$  GeV. Dielectron events in the high  $p_T$  triggers are selected using triggers with  $E_T$  requirements of 17 and 8 GeV on the two legs. Triggers with electron legs impose online calorimeter and tracker identification and isolating requirements. The online selections are sufficiently loose compared to the analysis isolation selection that little inefficiency is introduced. However, care must be taken in the data-driven background estimate so that a bias is not introduced. This issue was studied in detail elsewhere and found to not be a significant

concern [60]. The electron-muon cross triggers have muon  $p_T$  requirements of 17 GeV (8 GeV) and corresponding  $E_T$  requirements of 8 GeV (17 GeV). The dimuon triggers have only a  $p_T$  requirement of 17 and 8 GeV on the two legs.

The trigger paths used for the low  $p_T$  analyses have lower  $p_T$  requirements due to the lower threshold on the lepton (i.e.  $p_T > 10$  GeV); however, they impose an additional  $H_T$  requirement to keep the rate within the allotted bandwidth. The online  $H_T$  requirement was 175 GeV using particle flow based jets. Later runs also imposed a pile-up correcting to the  $H_T$  calculation. All three channels use triggers that also impose a dilepton invariant mass of 8 GeV. The dielectron triggers have an  $E_T$  requirement of 8 GeV on both legs. Unlike the high  $p_T$  triggers, the electron legs impose only calorimeter and tracker identification requirements – no isolation requirement. The electron-muon cross triggers require the same identification requirements on the electron leg and select electrons with  $E_T > 8$  GeV and muons with  $p_T > 8$  GeV. The  $p_T$  requirement in the dimuon trigger is 8 GeV on both legs.

### 4.2.2 Muon Selection

Muon reconstruction was discussed previously in Section 2.5.4. To reject fakes with relatively little loss in efficiency, we require muons to be reconstructed as both global muons and particle flow muons. To further reject fakes and poorly reconstructed muons we apply the following quality requirements [59, 61]:

- We require the muon to have  $p_T > 20$  GeV and  $p_T > 10$  GeV for the high  $p_T$  and low  $p_T$  analysis, respectively.
- We require the muon to have  $|\eta| < 2.4$  which is the edge of the muon detector coverage. Due to the finite luminous region, some muons can have a track with  $|\eta| > 2.4$ ; however, we do not consider these as the inefficiency is relatively large.
- We require that the global fit of the muon have  $\chi^2/\text{ndof} < 10$  to reject poorly reconstructed muons.
- We require that at least one muon sub-detector hit is used in the global fit.
- We require that at least one hit is in the pixel layers.

- We require that the muon have at least 6 layers in the silicon tracker, to reject poorly reconstructed muons as well as muons originating late in the tracker from decays in flight.
- We require that the energy deposits in the veto cones are less than 4 GeV in the ECAL and 6 GeV in the HCAL. Veto deposits are calculated in cones of size  $\Delta R = 0.07, 0.1$  for the ECAL and HCAL, respectively. Vetoing on the energy in the cones rejects fake muons produced when hadrons punch through the calorimeter into the muon system. The requirement can introduce an inefficiency for muons with high  $p_T$  or in a busy event environment such as a  $t\bar{t}$ -like event [62].
- We require the muon to have at least two segments in the muon chambers, to reject reconstructed objects that are mis-identified as muons because they punched through the calorimeter and interacted in the first layer of the muons system (e.g. from kaons).
- We require that the muon have transverse impact parameter ( $d_0 < 50 \mu\text{m}$ ), where the  $d_0$  has been calculated with respect to the primary vertex. A tight selection reduces background muons from the decay of b-hadrons, as well as decays in flight.
- We require that the inner track  $z$  be less than 1 mm from the first good vertex, which we take to be the event vertex. This requirement helps reject mis-reconstructed muons as well as those originating from pile-up interactions.

The isolation follows the POG recommendation, using particle-flow based isolation with a  $\Delta\beta$  correction for PU. However, a smaller cone size of  $\Delta R = 0.3$  is adopted due to the high hadronic activity expected in signal-like events. The isolation is calculated using

$$Iso = [\Sigma_{\text{ch}} + \max(0, \Sigma_{\text{nh}} + \Sigma_{\text{ph}} - 0.5\Delta\beta)]/p_T,$$

where  $\Sigma_{\text{ch,nh,ph}}$  are the sums of the  $p_T$  of the charged hadron, neutral hadron, and photon particle flow candidates, respectively. Here the charged hadrons are matched to the PV and a 0.5 GeV threshold is applied on neutral hadrons and photons. The  $\Delta\beta$  correction is determined from the sum  $p_T$  of charged hadrons not matched to the PV with a threshold

of 0.5 GeV in a cone of the same size as the isolation. The  $I_{so}$  is required to be less than 0.1 [63].

### 4.2.3 Electron Selection

Electrons in this analysis are reconstructed using the Gaussian Sum Filter (GSF) algorithm as was discussed in Section 2.5.3. For this analysis we require the GSF electrons to pass the additional requirements:

- We require the electron to have  $p_T > 20$  GeV and  $p_T > 10$  GeV for the high  $p_T$  and low  $p_T$  analyses, respectively.
- We require the electron to have  $|\eta| < 2.4$ .
- We require the number of missing expected inner hits to be zero.
- We require the  $H/E < 0.1$  (0.075) in the barrel (endcap) to match the requirements on the trigger.
- We require that the electron have a  $|d_0| < 100$   $\mu\text{m}$ , where the  $d_0$  has been calculated with respect to the primary vertex. A tight selection on the impact parameter reduces background from fake electrons and electrons from photon conversions in the tracker.
- We require that there is no muon within a cone of  $\Delta R = 0.1$  about the electron. Only muons passing the selection in Section 4.2.2 are considered. This veto rejects the situation where a muon is mis-reconstructed as an electron.
- We require the electron to pass the VBF80 identification (see Table 4.6).
- We reject electrons with a supercluster  $|\eta|$  in the transition region between the barrel and endcap of the ECAL ( $1.4442 < |\eta_{SC}| < 1.566$ ). Electrons in this region are poorly reconstructed.
- We require that the GSF track  $z$  be less than 1 mm from the first good vertex, which we take to be the event vertex. This requirement helps reject mis-reconstructed electrons as well as those originating from pile-up interactions.

- We require that all three charge measurements for an electron agree. One charge comes from the curvature of the GSF track. A second comes from the curvature of the associated CTF track. We require all electrons to have an associated CTF track. The last charge, the so-called supercluster charge, is determined using the relative position of the supercluster with respect to the projected track from the pixel seed.
- To reject conversions, we apply a veto of a good reconstructed conversion vertex. A conversion vertex is considered good if it has no tracker hits towards the beam, has a fit probability above  $10^{-6}$ , has a displacement of more than 2 cm, and the CTF track matching to the electron should be a part of the conversion vertex. No requirement is made on the vertex quality flag corresponding to merging and arbitration.

The isolation follows the POG recommendation, using particle-flow based isolation with a cone size of  $\Delta R = 0.3$  [64]. In the endcap, an inner veto of  $\Delta R = 0.015$  (0.08) is imposed for charged hadrons (photons). The isolation is corrected for PU by subtracting from the neutral isolation components a term defined by the product of the average event energy density and the effective area of the isolation cone [65]. The neutral component after correction is required to be non-negative. The isolation relative to the electron  $p_T$  is required to be less than 0.09.

The VBTF80 identification uses the selection from Table 4.6. CMS defines the barrel/endcap transition to be at  $|\eta_{SC}| = 1.479$ . The selection values are chosen to have approximately 80% efficiency for the electrons from W and Z decays. The values reported are from CMS recommended values [66].

**Table 4.6:** Details of the modified VBTF80 electron identification.

Selection Variable	Selection Value (barrel)	Selection Value (endcap)
$\sigma_{in\eta}$	< 0.01	< 0.03
$\Delta\phi_{in}$	< 0.06	< 0.03
$\eta_{in}$	< 0.004	< 0.007
$H/E$	< 0.1	< 0.075
$(1/E - 1/P)$	> 0.05	

The  $\sigma_{i\eta i\eta}$  variable cuts on the energy deposits in the ECAL, electrons have a characteristic shower shape that is narrow in  $\eta$  but wide in  $\phi$ , the latter due to the bremsstrahlung in the tracker of the electrons. The  $\Delta\phi_{in}$  and  $\Delta\eta_{int}$  variables are cut on the track-to-cluster matching. The final variable,  $H/E$ , is calculated as a ratio of energies in the ECAL and HCAL veto cones. An electron is expected to deposit nearly all of its energy in the ECAL, while a hadron usually deposits significant energy in the HCAL, as well.

#### 4.2.4 Lepton Pair Disambiguation

In events with multiple same-sign lepton pairs passing the selection above, only one pair is selected according to the following prescription:

- We give preference to  $\mu\mu$  pairs over  $e\mu$ , which are chosen over  $ee$  pairs.
- If multiple candidates remain, the pair with the highest scalar sum  $p_T$  is chosen.

#### 4.2.5 Jet Selection

We require the presence of energetic hadronic activity in the event as new physics with a large cross section is expected to be produced via the strong interaction. We choose to use jets built from particle-flow candidates, as they provide the best scale and resolution for jets typical of those produced in SM process such as  $t\bar{t}$  events. Particle-flow candidate reconstruction was discussed in section Section 2.5.5 and jet reconstruction was discussed in Section 2.5.7.

**Table 4.7:** Details of the loose particle flow identification. From [67].

Selection Variable	Selection Value	Comment
fraction of energy from neutral hadrons	< 0.99	
fraction of energy from neutral EM particles	< 0.99	
number of particle flow candidates	> 1	
fraction of energy from charged hadrons	> 0	$ \eta  < 2.4$
fraction of energy from charged EM particles	< 0.99	$ \eta  < 2.4$
number of charged particle flow candidates	> 0	$ \eta  < 2.4$

Jets are reconstructed with the anti- $k_T$  algorithm with parameter  $R = 0.5$ . Jets in simulation have L1FastJetL2L3 corrections applied, as well as jets in data have L2L3 residual corrections applied [68]. Selected jets are required to pass the loose particle flow jet ID described in Table 4.7. These selections are very loose and are primarily intended to reject jets that are unambiguously due to detector noise. The inefficiency of this selection on simulation is less than 1% [67]. Additionally, jets are required to be separated by  $\Delta R > 0.4$  from any hypothesis leptons and other leptons passing the selection above.

In the event, there must be at least two particle-flow jets with  $p_T > 40$  GeV and  $|\eta| < 2.4$ . The fiducial cut coincides with the extent of the tracker. In addition to counting jets, the analysis also makes a selection based on the total hadronic activity in the event,  $H_T$ . The  $H_T$  is calculated as the scalar sum of the  $p_T$  of all jets passing the jet counting criteria discussed above. Even a modest  $H_T$  requirement suppresses the background from failures of lepton identification in  $W + \text{jets}$  events. Having two jets in the event imposes a minimal  $H_T$  requirement of 80 GeV.

#### 4.2.6 B-Tagging Selection

Jets selected based on the requirements of Section 4.2.5 are b-tagged using the Combined Secondary Vertex method with the medium working point (CSVM) tagger. This tagger identifies jets with discriminant larger than 0.679 as b-tagged. The details of b-jet identification was discussed in Section 2.5.8.

#### 4.2.7 Missing Transverse Energy Selection

Missing transverse energy is a natural requirement for new physics searches. Many models contain weakly interacting particles. For example, some of the SUSY models considered produce leptons via decay chains ending in a lightest, non-interacting particle (LSP). CMS uses three different  $\cancel{E}_T$  reconstruction algorithms; however, for this analysis, we use  $\cancel{E}_T$  reconstructed from a vector sum of the particle flow candidates which was discussed in detail in Section 2.5.6. For almost all of the search scenarios, we impose a minimum  $\cancel{E}_T$  requirement of at least 30 GeV. Event a modest  $\cancel{E}_T$  requirement

is effective at suppressing background from Drell-Yan process in the  $ee$  and  $e\mu$  final states where the charge on one of the final state electrons is mis-reconstructed.

#### 4.2.8 Z and $\gamma^*$ veto

One of the primary irreducible backgrounds to the same-sign dilepton search comes from  $WZ$  and  $ZZ$  production, where the bosons both decay to leptons. A natural same sign hypothesis is formed using a lepton from each of the two bosons. In the case of the  $WZ$ , the lepton from the  $W$  comes together with a neutrino of the same flavor providing a natural source of  $\cancel{E}_T$ . To reduce this background, we reject events for which one of the hypothesis leptons and a third lepton in the event have an invariant mass consistent with the  $Z$ , defined to be between 76 and 106 GeV. Along the same lines, to reject events with a virtual photon, we also require the invariant mass to be less than 12 GeV. We require the third lepton to be the same flavor and opposite sign and to pass the identification and isolation criteria described in Sections 4.2.3 and 4.2.2.

### 4.3 Search Regions

As in previous versions of this analysis [58, 69, 70], there is a minimal baseline selection asking for a same-sign lepton pair and at least two jets (or b-tagged jets). Here, we define separately a control region (baseline) selection for each of the two lepton selections (high  $p_T$  and low  $p_T$ ) in the following b-tagged categories:

- no  $b$ -tagged jet requirement,
- require exactly one  $b$ -tagged jet,
- require two or more  $b$ -tagged jets.

In the baseline regions, a relatively low  $\cancel{E}_T$  requirement is made in an effort to be inclusive to a wide variety of signatures such as same-sign top productions or models with R-parity violation. Specifically, for R-parity violating models we expect events that have many high  $p_T$  jets. Thus a relatively high  $H_T > 500$  GeV requirement is required when

searching for this signature and any  $\cancel{E}_T$  requirement is relaxed. Many of the viable models, however, involve leptons arising from  $W$  decays and thus naturally have  $\cancel{E}_T$  from the accompanying neutrino. To improve sensitivity to these signatures, a region with  $\cancel{E}_T > 30 \text{ GeV}$  is introduced when  $H_T < 500 \text{ GeV}$ . Table 4.8 summarizes the baseline search regions used in this analysis where we define baselines for each of the analysis types. There are a total of six baseline regions. For convenience, each search region is assigned a number, given in the far right column.

**Table 4.8:** Summary of the baseline search regions considered in the high  $p_T$  and low  $p_T$  analysis.

Analysis	min lepton $p_T$ ( $\mu, e$ ) (GeV)	$H_T$ (GeV)	$\cancel{E}_T$ (GeV)	# jets	# $b$ -tagged jets	Search Region #
high $p_T$	20, 20	80	30 if $H_T < 500$ else 0	2	$\geq 0$	SR0
					= 1	SR10
					$\geq 2$	SR20
low $p_T$	10, 10	250	30 if $H_T < 500$ else 0	2	$\geq 0$	SR0
					= 1	SR10
					$\geq 2$	SR20

In contrast from the fall analysis, this analysis is performed as a simultaneous counting experiment in multiple independent bins or search regions. The search regions are defined by event selections on the lepton  $p_T$ ,  $H_T$ ,  $\cancel{E}_T$ , # jets, and #  $b$ -tagged jets . To improve sensitivity to SUSY scenarios involving top and bottom squark production, we define regions with tighter  $H_T$  and  $\cancel{E}_T$  requirements. The  $\cancel{E}_T$ - $H_T$  plane is divided into 4 major regions:

- Low- $H_T$  Low- $\cancel{E}_T$  region:  $200 \text{ GeV} < H_T < 400 \text{ GeV}$ ,  $50 \text{ GeV} < \cancel{E}_T < 120 \text{ GeV}$ .
- Low- $H_T$  high- $\cancel{E}_T$  region:  $200 \text{ GeV} < H_T < 400 \text{ GeV}$ ,  $\cancel{E}_T > 120 \text{ GeV}$ .
- high- $H_T$  Low- $\cancel{E}_T$  region:  $H_T > 400 \text{ GeV}$ ,  $50 \text{ GeV} < \cancel{E}_T < 120 \text{ GeV}$ .
- High- $H_T$  high- $\cancel{E}_T$  region:  $H_T > 400 \text{ GeV}$ ,  $\cancel{E}_T > 120 \text{ GeV}$ .

As will be discussed in Section 7.3 all of the SUSY scenarios that we consider explicitly have at least two quarks, zero to four b-quarks, up to two hadronically decaying W bosons, and at least two neutrinos and two LSPs. Thus, in all search regions, we require at least 2 jets but the # jets requirement is broken into two categories:

- two to three jets
- four or more jets

Additionally, the SUSY inspired search regions also have exclusively defined values on the #  $b$ -tagged jets with exactly zero, exactly one and two or more  $b$ -tagged jets. The search regions for the high  $p_T$  analysis are summarized in Table 4.9.

**Table 4.9:** Search regions selected for the high  $p_T$  search where we require  $H_T > 200$  GeV.

# $b$ -tagged jets	$\cancel{E}_T$	# jets	$H_T$ [200-400]	$H_T$ [ $> 400$ ]
$= 0$	50-120	2-3	SR1	SR2
		$\geq 4$	SR3	SR4
	$> 120$	2-3	SR5	SR6
		$\geq 4$	SR7	SR8
$= 1$	50-120	2-3	SR11	SR12
		$\geq 4$	SR13	SR14
	$> 120$	2-3	SR15	SR16
		$\geq 4$	SR17	SR18
$\geq 2$	50-120	2-3	SR21	SR22
		$\geq 4$	SR23	SR24
	$> 120$	2-3	SR25	SR26
		$\geq 4$	SR27	SR28

The search regions for the low  $p_T$  analysis are essentially the same as the high  $p_T$  analysis with one change. In order to avoid the threshold effect from the  $H_T > 175$  GeV requirement from the triggers used in low  $p_T$  analysis, the minimum  $H_T$  requirement is raised to 250 GeV. The search regions for the low  $p_T$  analysis are summarized in Table 4.10.

**Table 4.10:** Search regions selected for the low  $p_T$  search where we require  $H_T > 250$  GeV.

# $b$ -tagged jets	$\cancel{E}_T$	# jets	$H_T$ [250-400]	$H_T$ [ $> 400$ ]
$= 0$	50-120	2-3	SR1	SR2
		$\geq 4$	SR3	SR4
	$> 120$	2-3	SR5	SR6
		$\geq 4$	SR7	SR8
$= 1$	50-120	2-3	SR11	SR12
		$\geq 4$	SR13	SR14
	$> 120$	2-3	SR15	SR16
		$\geq 4$	SR17	SR18
$\geq 2$	50-120	2-3	SR21	SR22
		$\geq 4$	SR23	SR24
	$> 120$	2-3	SR25	SR26
		$\geq 4$	SR27	SR28

# Chapter 5

## Background Estimation Methods

In Chapter 3, we described the different sources of background events in this analysis. Chapter 4 described the actual selections used to differentiate interesting events from background. Even after these selections, which were designed to isolate the "real" leptons, a non-trivial amount of background leptons still remain. Same-sign dilepton events formed from non genuine same-sign processes come from two main sources: failures of the identification (fakes) and failure to properly reconstruct the electron's charge (charge flips).

Fake leptons can be the dominant background to this analysis depending on the search region. The term fake leptons here will refer to leptons from heavy flavor decays and decays in flight as well as jets that leave a lepton-like signature in the detector (see Chapter 3). In either case, fake leptons originate from jets. For the fake lepton to pass identification and isolation selections, the underlying parton must have showered, or fragmented, in such an unlikely way as to leave a lepton-like signature. This scenario may not be well modeled by the simulation. Thus, data-driven methods to estimate these backgrounds are needed.

A typical fake muon comes from the decay of a heavy flavor parton. For example, a bound state containing a b-quark decays semi-leptonically ( $b \rightarrow W + c; W \rightarrow \ell \bar{\nu}_\ell$ ). A bit more than 20% of the time, the  $W$  decays to an electron or muon. As this is a real lepton, the muon will pass the identification requirement most of the time – as expected. However, because of the boost of the b hadron system, the lepton from the virtual  $W$  will frequently have a small separation from the original hadron's (and remaining frag-

ments') trajectory. Isolation was designed to be a handle against leptons from this type of scenario since prompt leptons tend to be relatively well separated spatially from other objects in the collision.

Fake electrons come from heavy flavor decays, in a manner similar to muons as well as from light flavor jets. Details of the electron reconstruction algorithm make it less likely to reconstruct an electron embedded in a high  $p_T$  jet as compared to a muon. Despite this, fake electrons are relatively common. A representative example is a light flavor jet that fragments to a  $\pi^0$  and a charged pion. The  $\pi^0$  will deposit its energy in the ECAL, and if the track from the charged pion is matched to the supercluster reconstructed from the  $\pi^0$ , an electron object can be constructed. Electron identification described in Section 4.2.3 is a powerful tool for rejecting this scenario; however, it is not perfect and fake electrons remain. If, however, an electron comes from a heavy flavor decay, isolation remains the best handle for rejection.

Leptons that originate from heavy flavor decays (bottom or charm hadrons), for which isolation is the best handle to differential these from leptons from electroweak processes, pass the full selection with a probability that depends on the momentum of the parent parton. If the lepton carries a large fraction of the momentum of the original parton or is emitted at a large relative angle, the lepton has a non-negligible chance of being isolated. The larger the difference between the trajectory of the originating parton and the daughter lepton, the more ambient energy there is potentially surrounding the lepton and the less likely the lepton is to pass the isolation requirement. Thus, the probability for a lepton of a given momentum to pass the analysis selection is dependent upon the underlying parton kinematics.

We attempt to estimate the probability for a reconstructed lepton passing some loose requirement to also pass the full analysis selections by measuring a ratio of these two in a statistically independent sample. We assume that a jet is sufficiently similar in all samples and the likelihood for a jet to fake a lepton is largely independent of the underlying process and kinematics – that is, the fake rate is universal. The validity of this assumption is tested in simulation and the measured dependence on the sample composition and the underlying kinematics is taken as a systematic uncertainty on the method.

Lepton charge is primarily determined from the curvature of the reconstructed trajectory in the form of a track. The rate to mis-measure the charge for a muon is negligibly small at the W/Z momentum scale. Yet, the charge mis-reconstruction rate for electrons is significant enough to warrant attention. Bremsstrahlung is emitted as electrons interact with the silicon atoms from the layers of the tracker. A radiated photon has a significant probability to convert into an  $e^+e^-$  pair due the large amount of material in the tracking system. The combination of the radiation of a hard photon and a subsequent asymmetric conversion (specifically when one of the two resulting electrons carries a significant fraction of the momentum of the original photon) can result in the reconstruction of a track that combines hits from the original electron and those from the converted photon. As a result, the reconstructed track may have a curvature opposite to that of the prompt electron, resulting in an incorrect charge assignment. We estimate the background from this source by determining the electron charge mis-reconstruction rate using a simulated sample of Drell-Yan and  $t\bar{t}$  events. We apply that rate to data to events passing the full analysis selection described in Chapter 4, but with the same sign electric charge requirement placed by an opposite sign requirement.

Standard model source of prompt, same-sign dileptons are rare and contribute between 25-60%, depending on the search region. The remaining contributions from the two data-driven estimates are significant at the 40-75% level, also depending on the search region. This chapter is broken into three sections. The first describes the details of the data-driven method to estimate the number of events with “fake” leptons, still one of the dominant sources of backgrounds for this analysis. The second describes the method that uses information from both data and simulation to estimate the number of prompt electrons with a mis-reconstructed charge. Finally, the chapter concludes with a brief discussion of the irreducible estimates for the rare SM processes.

## 5.1 Data-Driven method to estimate the fake rate

In this section, we discuss the method used to estimate the background from fake leptons. An overview of the method is provided, followed by a determination of the fake rate in an independent sample. The fake rates themselves are given, followed by a study

of the systematic uncertainties associated with this method using simulation to perform a closure test.

### 5.1.1 The Fake Rate Method

The concept of the fake rate (FR) method is that multijet data is used to measure a lepton FR as a function of the lepton’s  $p_T$  and  $|\eta|$ . The multijet data will mostly be from QCD processes. The fake rate is defined as the probability for a lepton candidate passing a loose lepton selection to also pass the tight, full analysis selection. Leptons passing the loose but not the tight selection are called “fakeable objects” (FO). The fake rate is used to estimate the background from fake leptons as follows:

- Events are selected using all analysis cuts, except the lepton selection. Events with one lepton passing the tight selection and one failing it, but passing the looser selection, are used to estimate the background to dilepton final states from one fake lepton (single fakes). Backgrounds with two fake leptons are estimated by requiring both leptons to pass the loose selection and fail the full analysis selection (double fakes).
- The background from a single fake lepton is estimated by weighing each tight-FO pair by factor of  $\epsilon_{FR}/(1 - \epsilon_{FR})$ , where  $\epsilon_{FR}$  is the value of the fake rate for the chosen FO’s  $p_T$  and  $|\eta|$ .
- The background from two fake leptons is estimated by weighting each FO-FO pair by the product of two factors of  $\epsilon_{FR}/(1 - \epsilon_{FR})$ , where the value of  $\epsilon_{FR}$  is determined separately for each of the two FOs.
- The combination of the two weights over the selected events is the background estimate.

Note that the estimate of single fakes for a pure sample of two fake leptons overestimates the background by a factor of two<sup>1</sup>. Implicit in the above description is two assumptions

---

<sup>1</sup> The reason for the over prediction is due to fact that a single fake event can have contributions from real single fake events (e.g.  $W + \text{jets}$ ) and real double fake events where one fake leptons fluctuates to passes the full selection and the other is still a FO. See [71] for more details.

upon which the method relies:

1. The lepton fake rate measured in a inclusive QCD sample represents the lepton fake rate in the signal sample (universality).
2. The lepton fake rate is independent for the two leptons. Specifically, to predict the contribution to an  $e\mu$  event, we consider the electron and muon fakes separately assuming no correlations between the two estimates (locality).

The validity of these two assumptions is tested as part of the study of the systematic uncertainty of the method.

### 5.1.2 Fakeable Object Definition

The fake rate method uses an extrapolation from the analysis lepton requirements to a looser lepton selection. A lepton passing the full analysis selection (or tight selection) is referred to as a “numerator” lepton. Similarly, a lepton passing the loose selection is defined as a “denominator” object. The fake rate itself is determined by the ratio of the number to denominator lepton counts in bins of the  $p_T$  and  $|\eta|$  of the leptons. As described in the Section 5.1.1, the estimation of the background counts leptons that pass the denominator selection but fails the numerator. We refer to these as non-numerator leptons.

The numerator selections were described in Chapter 4. The denominator selections are described below specifying only the looser selections. The muon denominator definition is to relax the following muon requirements from Section 4.2.2:

- global fit  $\chi^2/\text{ndof} < 50$  (numerator cut is  $< 10$ );
- transverse impact parameter with respect to the selected vertex is  $< 2 \text{ mm}$  (numerator cut  $< 50 \mu\text{m}$ );
- the MIP-like requirement on deposits in the ECAL and HCAL is removed [numerator cut was  $\text{ECAL(HCAL)} < 4(6) \text{ GeV}$ ];
- $I_{\text{so}}$  is set to  $I_{\text{so}} < 0.4$  (numerator cut is  $< 0.10$ ).

The electron denominator definition is to relax the following requirements from Section 4.2.3:

- transverse impact parameter cut is removed (numerator cut  $< 100 \mu\text{m}$ );
- $I_{\text{so}}$  is set to  $I_{\text{so}} < 0.6$  (numerator cut is  $< 0.09$ ).

We thus use an extrapolation in isolation and impact parameter to estimate the fake lepton backgrounds for electrons and muons. Based on simulation, we expect that the background is comprised of by  $t\bar{t}$  events, where the fake leptons is predominantly a real lepton from the leptonic decay of a b-quark. In this scenario, relaxing the isolation and impact parameter requirements, while keeping the identification tight, provides sufficient lever arm while keeping roughly the same composition in events with denominator leptons.

### 5.1.3 Fake Rate Datasets

The fake rate is measured in a multi-jet sample. Since the single jet triggers are heavily pre-scaled, we select a sample using single lepton triggers. The triggers are chosen so that the online lepton requirements coincide as closely as possible with the possible with those used in the dileptons triggers. A list of primary datasets is given in Table 5.1 and for the simulated datasets in Table 5.2.

Events are required to have a lepton passing the denominator requirement discussed in Section 5.1.2. To enrich the sample in fakes, we require an away jet separated from the lepton by  $\Delta R > 1$ . Here a jet is a particle-flow jet that is required to pass the loose particle-flow jet ID listed in Table 4.7. The jet is further required to have a  $p_{\text{T}} > 40 \text{ GeV}$  and  $|\eta| < 2.4$ . The jet  $p_{\text{T}}$  is corrected using the methods described in Section 2.5.7. To reject a known machine background resulting in high activity in the pixel layers, we require that no fewer than 25% of track are high purity in events having 10 or more tracks. Also, to ensure that there was a reconstructed collision, we require at least one good vertex. Only runs and luminosity sections certified as good are used in the fake rate determination.

The muon fake rate was measured using a combination of the single muon triggers that had various  $p_{\text{T}}$  and jet requirements. These triggers have various pre-scales but

**Table 5.1:** Datasets used to measure the lepton fake rates.

Name	Run Range
/DoubleMu/Run2012A-13Jul2012-v1/AOD	190456-193621
/DoubleMu/Run2012B-13Jul2012-v1/AOD	193834-196531
/DoubleMu/Run2012A-recover-06Aug2012-v1/AOD	190949 190945 190906 190895 190782
/DoubleMu/Run2012C-24Aug2012-v1/AOD	198022-198523
/DoubleMu/Run2012C-PromptReco-v2/AOD	198934-203755
/DoubleMu/Run2012D-PromptReco-v1/AOD	203773-208913
/DoubleElectron/Run2012A-13Jul2012-v1/AOD	190456-193621
/DoubleElectron/Run2012B-13Jul2012-v1/AOD	193834-196531
/DoubleElectron/Run2012A-recover-06Aug2012-v1/AOD	190949 190945 190906 190895 190782
/DoubleElectron/Run2012C-24Aug2012-v1/AOD	198022-198523
/DoubleElectron/Run2012C-PromptReco-v2/AOD	198934-203755
/DoubleElectron/Run2012D-PromptReco-v1/AOD	203773-208913
/SingleMu/Run2012A-13Jul2012-v1/AOD	190456-193621
/SingleMu/Run2012B-13Jul2012-v1/AOD	193834-196531
/SingleMu/Run2012A-recover-06Aug2012-v1/AOD	190949 190945 190906 190895 190782
/SingleMu/Run2012C-24Aug2012-v1/AOD	198022-198523
/SingleMu/Run2012C-PromptReco-v2/AOD	198934-203755
/SingleMu/Run2012D-PromptReco-v1/AOD	203773-208913

there was sufficient statistics to calculate the fake rate. The high  $p_T$  and low  $p_T$  analysis uses the same triggers and they are listed in Table 5.3.

The high  $p_T$  and low  $p_T$  analyses each require a different electron fake rates. The high  $p_T$  analysis uses dilepton triggers with an online isolation requirement, as listed in Table 4.2.1. For consistency, we measure the fake rate using the single lepton triggers with the same isolation and identification requirements. In order to increase the statistics, we use the trigger with the same  $E_T$ , isolation and identification requirements but have an additional jet requirement of  $> 30$  GeV. For the low  $p_T$  analysis, we require the electrons to be  $> 10$  GeV. Since these dilepton triggers don't have isolation, we use the corresponding single leptons triggers without the isolation requirement. The list of electron triggers used for the fake rate measurements is shown in Table 5.3.

### 5.1.4 Fake Rate Contamination from Prompt Leptons

Even after the away jet requirement on the event, there is a contamination from leptons from  $W$  and  $Z$  boson decays. These events are further pruned of the contamination from the electroweak processes with  $W$  or  $Z$  production. The  $W$  events

**Table 5.2:** Simulated samples used to measure the lepton fake rates. The common part of each dataset name Summer12\_DR53X-PU\_S10\_START53\_V7A is replaced with a shorthand Su12. All datasets are in the AODSIM data tier.

Name	Cross section, pb
TT_CT10_TuneZ2star_8TeV-powheg-tauola/Su12-v2	334
TT_CT10_TuneZ2star_8TeV-powheg-tauola/Su12-v1	334
TTJets_SemiLeptMGDecays_8TeV-madgraph/Su12-v2	102.50
W1JetsToLNu_TuneZ2Star_8TeV-madgraph/Su12-v1	6663
W2JetsToLNu_TuneZ2Star_8TeV-madgraph/Su12-v1	2159
W3JetsToLNu_TuneZ2Star_8TeV-madgraph/Su12-v1	640
W4JetsToLNu_TuneZ2Star_8TeV-madgraph/Su12-v1	264
QCD_Pt_5to15_TuneZ2star_8TeV_pythia6/Su12-v1	4.2639499e10
QCD_Pt_15to30_TuneZ2star_8TeV_pythia6/Su12-v2	9.8828742e8
QCD_Pt_30to50_TuneZ2star_8TeV_pythia6/Su12-v2	6.6285328e7
QCD_Pt_50to80_TuneZ2star_8TeV_pythia6/Su12-v2	8148778.0
QCD_Pt_80to120_TuneZ2star_8TeV_pythia6/Su12-v3	1033680.0
QCD_Pt_120to170_TuneZ2star_8TeV_pythia6/Su12-v3	156293.3
QCD_Pt_170to300_TuneZ2star_8TeV_pythia6/Su12-v2	34138.15
QCD_Pt_15to20_MuEnrichedPt5_TuneZ2star_8TeV_pythia6/Su12-v2	7.022e8
QCD_Pt_20to30_MuEnrichedPt5_TuneZ2star_8TeV_pythia6/Su12-v1	2.87e8
QCD_Pt_30to50_MuEnrichedPt5_TuneZ2star_8TeV_pythia6/Su12-v1	6.609e7
QCD_Pt_50to80_MuEnrichedPt5_TuneZ2star_8TeV_pythia6/Su12-v1	8082000.0
QCD_Pt_80to120_MuEnrichedPt5_TuneZ2star_8TeV_pythia6/Su12-v1	1024000.0
QCD_Pt_120to170_MuEnrichedPt5_TuneZ2star_8TeV_pythia6/Su12-v1	157800.0
QCD_Pt_20_MuEnrichedPt_15_TuneZ2star_8TeV_pythia6/Su12-v3	3.64e8

are suppressed by a requirement that  $\cancel{E}_T$  is below 20 GeV and the transverse mass  $M_T < 25$  GeV. The Z events are initially suppressed by removing dielectron and dimuon events with another lepton passing the denominator requirement and forming a pair with an invariant mass between 71 to 111 GeV (events are removed only with dileptons with both  $p_T > 20$  GeV and the other lepton passing a looser ID and isolation selection. Further stringent suppression of remaining source of genuine, prompt leptons is achieved by vetoing events satisfying the following conditions,

- for electrons:
  - other fakeable objects with  $p_T > 10$  GeV are present;
  - there is a GSF track making an opposite-sign pair with the fakeable object and an invariant mass of 76 – 106 GeV;
  - the away jet has an EM-fraction higher than 0.8.

**Table 5.3:** The trigger paths used for the high  $p_T$  and low  $p_T$  analyses to calculate the Fake Rate. Note that for the high  $p_T$  triggers, XXLstr stands for CaloIdT\_CaloIsoVL\_TrkIdVL\_TrkIsoVL which is explained in the text below. For the low  $p_T$  triggers, PFHT175 refers to particle flow based  $H_T > 175 \text{ GeV}$ .

Lepton Type	Analysis Type	Trigger
muons	high and low $p_T$	HLT_Mu5 HLT_Mu8 HLT_Mu12 HLT_Mu17 HLT_Mu24_eta2p1 HLT_Mu30_eta2p1
electrons	high $p_T$	HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_Jet30 HLT_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL
electrons	high $p_T$	HLT_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_Jet30 HLT_Ele8_CaloIdT_CaloIsoVL_TrkIdVL HLT_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_Jet30

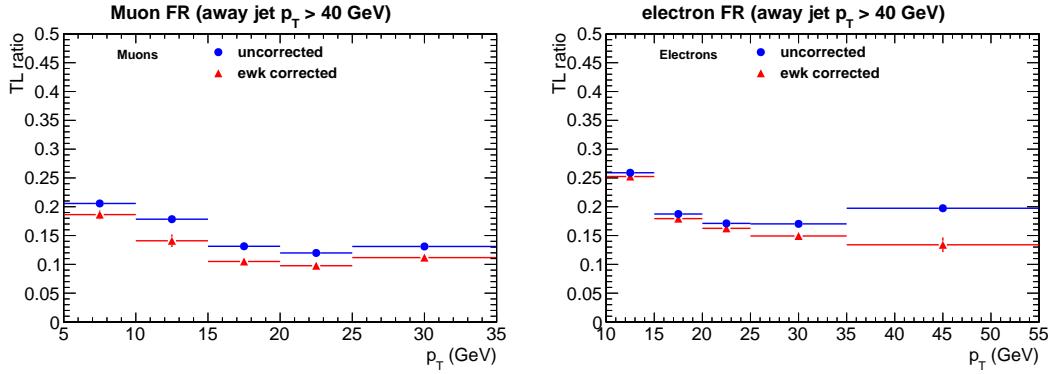
- for muons:
  - other fakeable objects with  $p_T > 10 \text{ GeV}$  are present;
  - there is another muon (no ID requirement) with  $p_T > 10 \text{ GeV}$  making a opposite-sign pair with the fakeable object with an invariant mass of 76 – 106 GeV;
  - there is another muon (no ID requirement) with  $p_T > 10 \text{ GeV}$  making a opposite-sign pair with the fakeable object with an invariant mass between 8 and 12 GeV to additionally suppress Upsilon production contribution.

Despite the above requirements, there still is some contamination from the  $W/Z$  events that contribute to the fake rate. In order to subtract this effect, we use simulation to estimate and subtract this contamination in the data derived fake rate. The procedure is as follows:

- Normalize  $Z/\gamma^* \rightarrow \ell\ell$  and  $W \rightarrow \ell\bar{\nu}_\ell$  simulated events to the effective luminosity from each pre-scaled trigger from 5.3,
- apply all FO selections except the  $\cancel{E}_T$  and  $M_T$  selections;

- select a region enriched in prompt leptons from data and simulation:
  - apply a  $\cancel{E}_T > 30 \text{ GeV}$  to selected event (apply truth matching for simulation),
  - require the  $M_T$  to be inside a window of 60-100 GeV,
- Extract a data-to-MC scale factor from the lepton selections per  $p_T$  and  $|\eta|$  bin,
- apply the data-to-MC scale factor as a function of  $p_T$  and  $|\eta|$  to the MC count of selected leptons in this  $M_T$  window. This represents an estimation of the true prompt leptons that are contaminating the fake rate measurement,
- subtract this binned count from the measured fake rate from data. This is the corrected fake rate that is used in the analysis.

The following plots in Figure 5.1 show that the difference in the  $p_T$  projection of the fake rate is fairly small for muons and increases with  $p_T$  for electrons.



**Figure 5.1:** Projection of the muon (left) and electron (right) fake rate in data vs  $p_T$ . The blue plot is the fake rate before the EWK correction and the red plot is after.

### 5.1.5 Fake Rates for Electrons and Muons

The fake rate is measured using the selections described in Section 5.1.1. As discussed in Section 5.1.3, the electron fake rates are measured separately for the triggers with and without an online isolation requirement. The results of the measurement for electrons are summarized in Tables A.1 and A.2. The muon fake rates are measured

using the single-muon triggers described in Section 5.1.3. The results are summarized in Table A.3.

Figure 5.4 shows the  $p_T$ ,  $|\eta|$ , and  $N_{\text{vtx}}$  projections of the fake rate as measured in data. The fake rate is approximately flat in  $|\eta|$ . The FR decreases with increasing  $p_T$ , from a value of approximately 0.2 at 5 GeV to about 0.13 at high  $p_T$ . The highest  $p_T$  bin shows a slight increase in the fake rate that may be an indication that  $W$  contamination is present, but the increase is not significant enough to warrant considerable attentions at this point. We assume that the fake rate is mostly constant above a  $p_T > 35$  GeV so we assume this value in the last bin for all leptons greater than 35 GeV.

Figure 5.5 shows the  $p_T$ ,  $|\eta|$ , and  $N_{\text{vtx}}$  projections for the electron fake rate measured with and without an online isolation requirement. The fake rate is relatively stable, although it does show a slight upward trend at high  $p_T$ . This is more pronounced in the fake rate with the online isolation. Overall, the fake rate without the isolation requirement is higher than with it, which is expected since this allows more non-prompt leptons to be considered (due to higher denominator lepton counts). We assume that the fake rate is mostly constant above a  $p_T > 55$  GeV so we assume this value in the last bin for all leptons greater than 55 GeV.

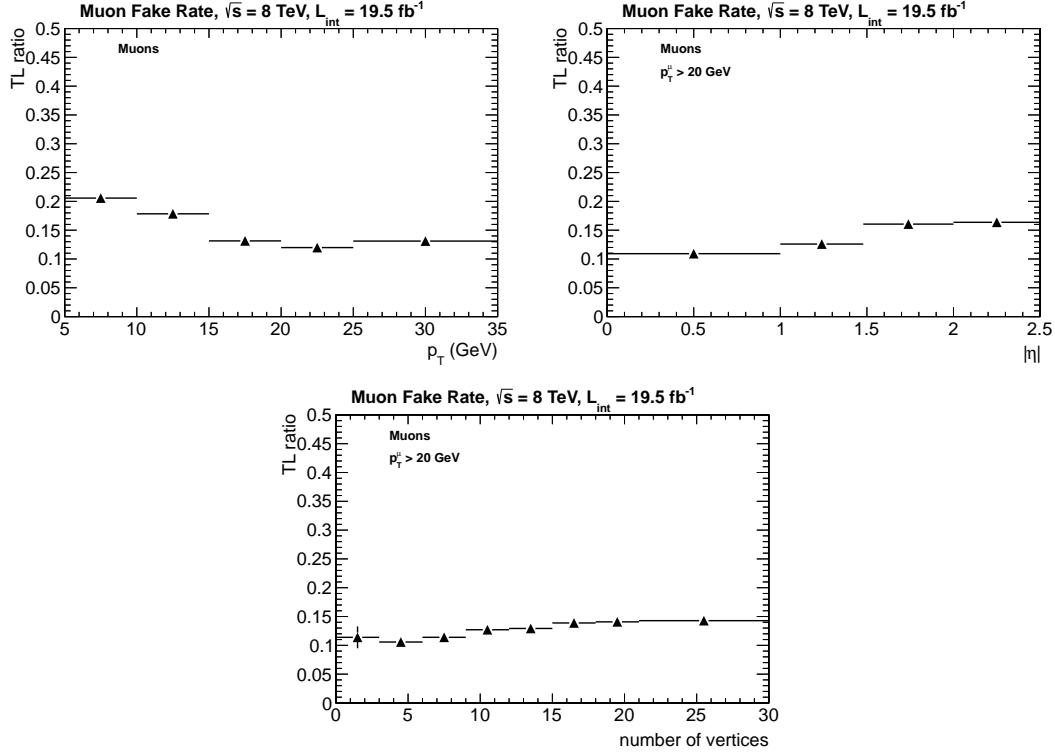
### 5.1.6 Study of Fake Rate Dependencies

Several studies are performed in data and simulation to assess the dependence of the fake rate on the kinematics. Studies of the contaminations of the fake rate in data due to prompt leptons was already considered in Section 5.1.4. Here we perform the following two studies:

- Measurement of the fake rate dependence on the away-side jet  $p_T$ , as a measure of the dependence of the parent parton momentum.
- Closure test are performed using simulated QCD,  $W + \text{jets}$ , and  $t\bar{t}$  samples.

#### Fake Rate Dependence on Away Jet $p_T$

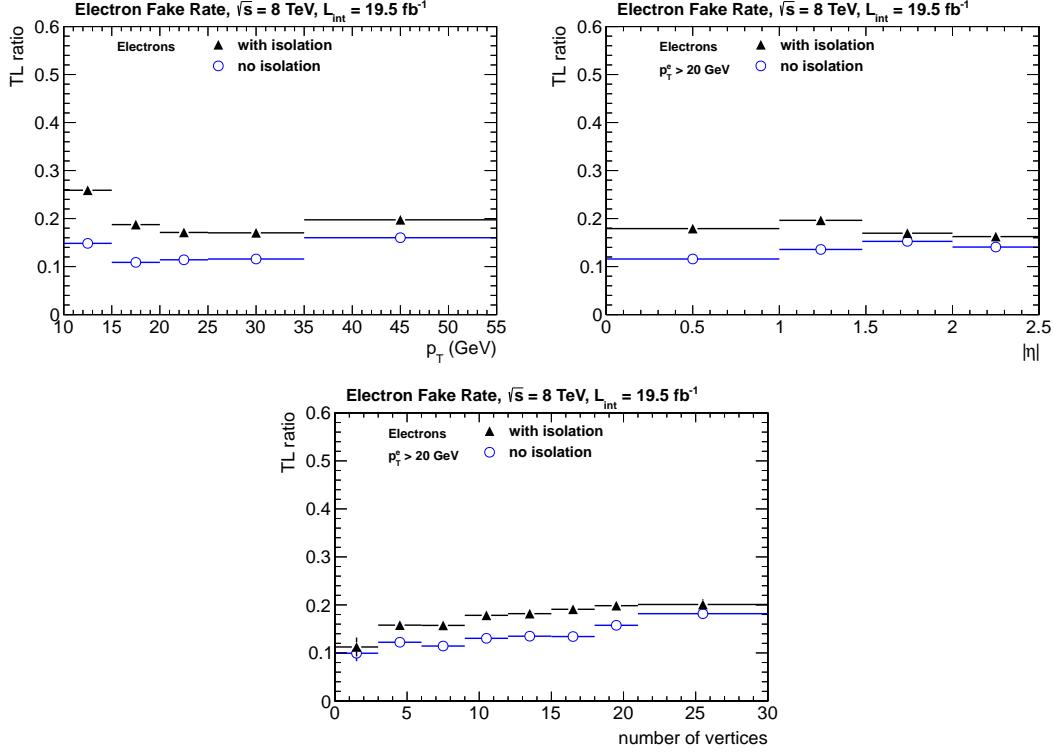
The lepton fake rates are dependent upon the away-jet requirement because of the extrapolation in isolation. Lepton candidates in high  $p_T$  jets have a smaller proba-



**Figure 5.2:** Projection of the muon fake rate in data vs  $|\eta|$ ,  $p_T$ , and the number of vertices. This fake rate is used for both the high  $p_T$  and low  $p_T$  analysis.

bility to pass an isolation requirement than a similar  $p_T$  lepton candidate in a low  $p_T$  jet. The difference can be understood by considering the relative size of the lepton  $p_T$  and the difference between the parent parton's  $p_T$  with the lepton's  $p_T$  ( $|p_T^{\text{parton}} - p_T^{\text{lepton}}|$ ). The latter is a measure of the maximum transverse energy that can be deposited in the isolation cone. Thus, for a given lepton  $p_T$ , the higher the  $p_T$  of the originating parton, the less likely it is for this lepton to pass the numerator isolation criteria. The probability levels off when the momentum of the parton becomes much larger than that of the lepton candidate.

Figures 5.4-5.6 show the  $|\eta|$  and  $p_T$  projections for the muon and electron fake rates measured in the data with  $p_T > 20, 40$ , and  $60$  GeV away-jet requirements. The range of jet  $p_T$  considered was determined by looking at the typical  $p_T$  spectrum of  $b$ -quarks with FO daughters in simulated  $t\bar{t}$  events. The dependence of the muon fake rate on the away-jet  $p_T$  requirement is  $\sim 30\%$ . The dependence is flat in  $|\eta|$ , but in-

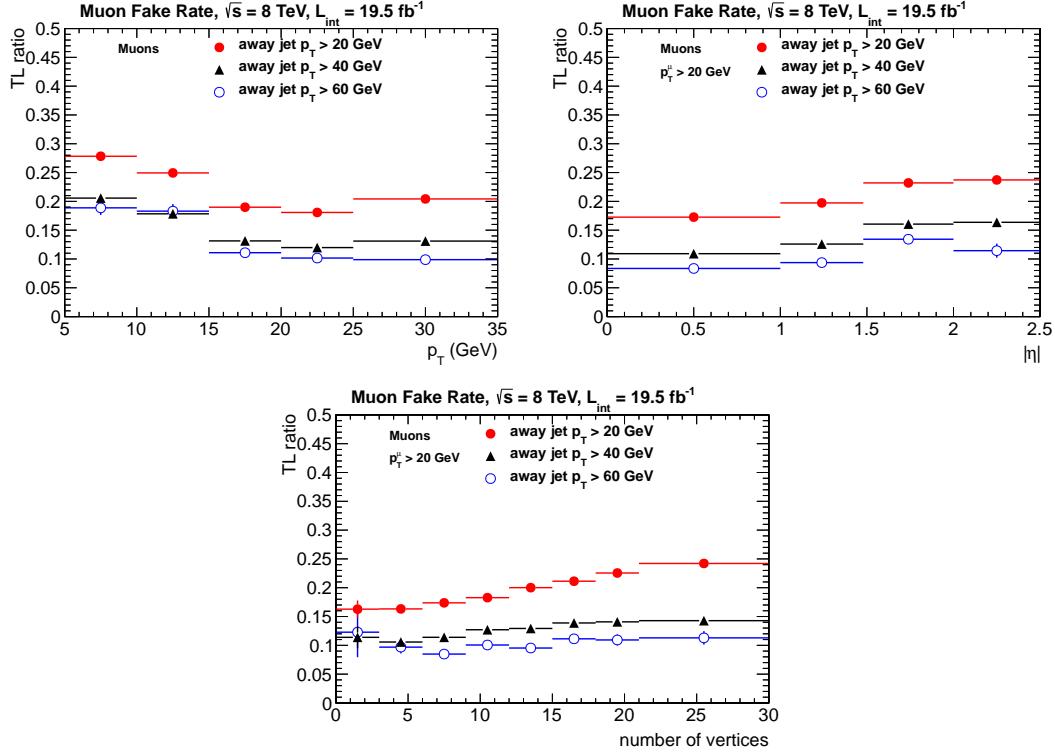


**Figure 5.3:** Projection of the electron fake rate in data vs  $|\eta|$ ,  $p_T$ , and the number of vertices with (triangles) and without (open circles) an online isolation requirement. The fake rate measured with an online isolation requirement is used for the high  $p_T$  analysis. The triggers without the online isolation requirement are used in the fake rate for the low  $p_T$  analysis.

creases with  $p_T$  for the muon candidate. The dependence of the electron fake rates on the away-jet  $p_T$  is similarly  $\sim 30\%$ . There is not a strong trend in either  $|\eta|$  or  $p_T$ . The dependence appears similar for electron fake rates measured with and without an isolation requirement.

### Fake Rate Closure Tests

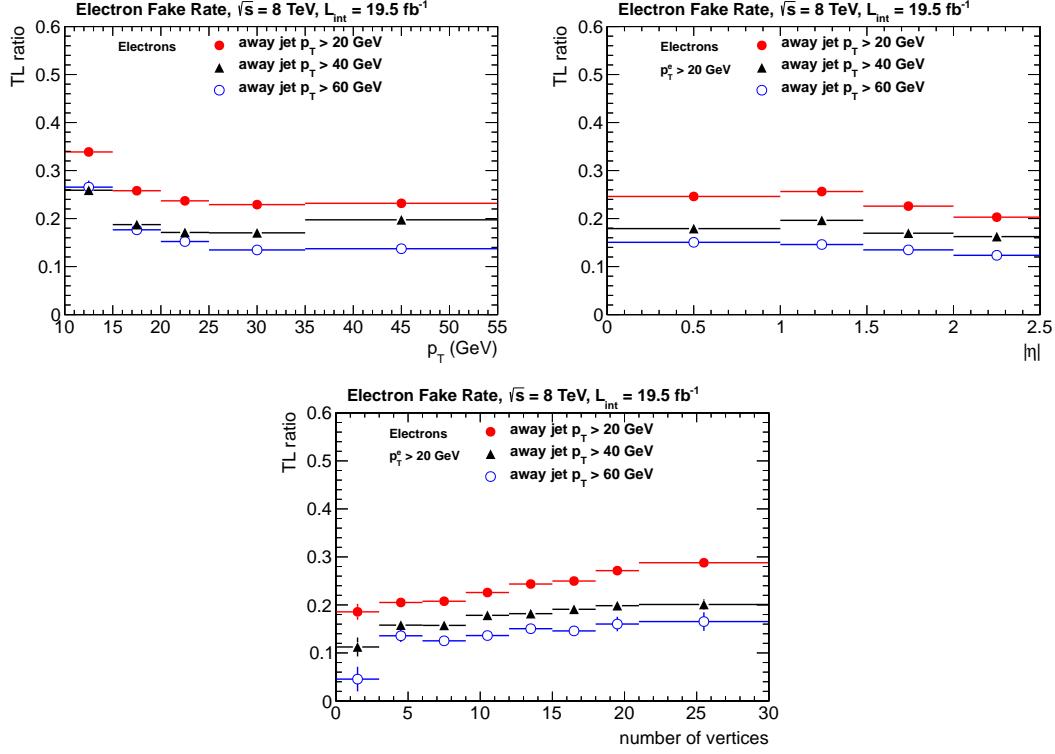
We test that the fake rates measured in QCD are applicable to the dilepton samples by performing a closure test on a simulated  $t\bar{t}$  and  $W + \text{jets}$  sample. The closure test uses QCD MC to derive the fake rate itself and applied events selected from  $t\bar{t}$  and



**Figure 5.4:** Projection of the muon fake rate in data vs  $|\eta|$ ,  $p_T$ , and  $N_{\text{vtx}}$  for different away jet  $p_T$  requirements. The away jet is a corrected particle flow jet and is required to be  $\Delta R > 0.1$  away from the muon candidate.

$W + \text{jets}$  MC sample. See Table 5.2 for the MC samples used for this test. Other dilepton processes with fakes, such as QCD with two fakes, are expected to contribute negligibly to the events selected in this analysis. First we describe the results of the closure test using  $t\bar{t}$  and  $W + \text{jets}$  events with the QCD derived fake rate. The following selections are applied to events entering the  $t\bar{t}$  and  $W + \text{jets}$  closure tests:

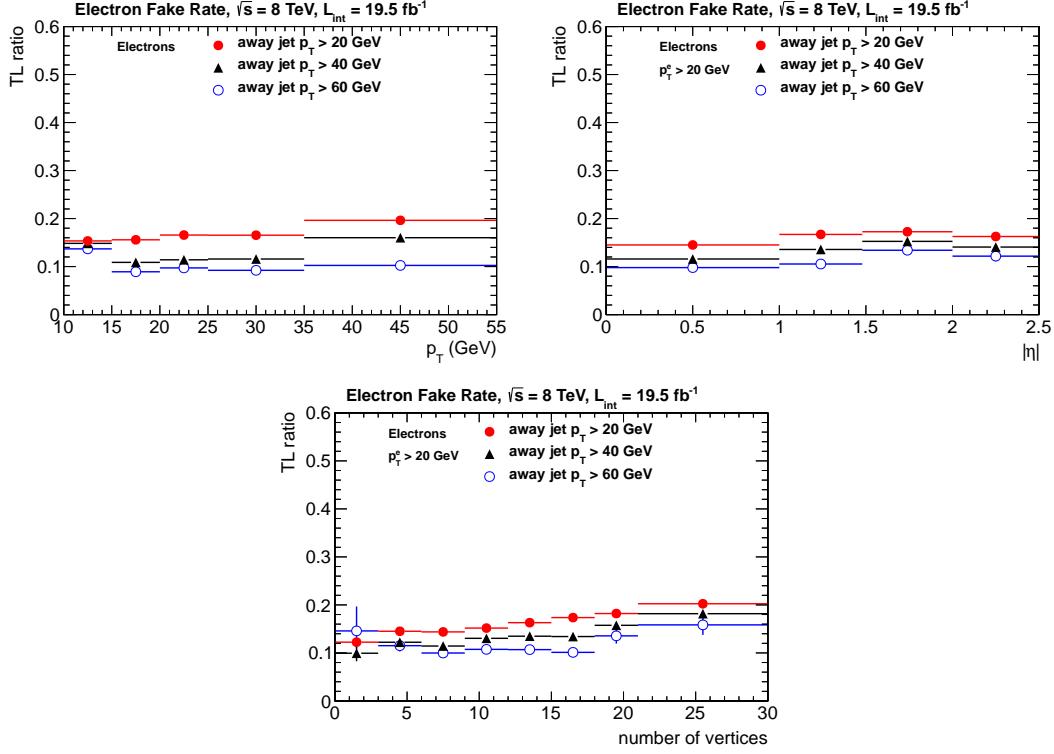
1. Select events that pass the baseline selections (SR0, SR10, SR20). For  $W + \text{jets}$ , the statistics was too low so only the SR0 is considered.
2. Require at truth level that for  $t\bar{t}$  events, the top-quark pairs decay semi-leptonically (i.e. one  $W \rightarrow \ell \bar{\nu}_\ell$  and the other  $W \rightarrow q\bar{q}$ ).
3. Require that one lepton (real) is matched to the leptonic  $W$  decay and the other



**Figure 5.5:** Projection of the electron fake rate in data vs  $|\eta|$ ,  $p_T$ , and  $N_{\text{vtx}}$  for different away jet  $p_T$  requirements. The away jet is a corrected particle flow jet and is required to be  $\Delta R > 0.1$  away from the electron candidate. The electron fake rate is measured using triggers with an online isolation requirement.

(fake) lepton is not matched to the leptonic W decay and passes the denominator but **not** the full lepton selections. This is the FO count.

4. Scale the FO count by  $\epsilon_{FR}/(1 - \epsilon_{FR})$  where  $\epsilon_{FR}$  is the QCD MC derived fake rate as a function of the fake lepton  $p_T$  and  $|\eta|$  – this is the prediction of the number of fakes passing the full lepton selections.
5. Require that one lepton (real) is matched to the leptonic W decay and the other (fake) lepton is not matched to the leptonic W decay and passes the full lepton selections – this is the observed number of fake events.
6. Compare the predicted and observed number of fake events.



**Figure 5.6:** Projection of the electron fake rate in data vs  $|\eta|$ ,  $p_T$ , and  $N_{\text{vtx}}$  for different away jet  $p_T$  requirements. The away jet is a corrected particle flow jet and is required to be  $\Delta R > 0.1$  away from the electron candidate. The electron fake rate is measured using triggers without an online isolation requirement.

Table 5.4 shows the results of the closure test. As in past iterations of this analysis, there is an over-prediction in the  $t\bar{t}$  MC [72, 73]. The systematic uncertainty of  $\pm 50\%$  per fake lepton is estimated for the fake rate method. Our understanding of the these results is that the main underlying cause is the dependence of the fake rate on the parent parton momentum. The momentum spectrum of partons from ISR/FSR differs from that of the b-jets or light-flavor-quark jets ( $W \rightarrow q\bar{q}$ ) arising from  $t \rightarrow Wb$  decays. The mix of the spectra varies, but the range of the fake rate variation can be tested in data QCD used to measure the fake rate by applying varying threshold to the away jet as illustrated in Fig. 5.4 and 5.6.

We compute the contributions from double-fake and single-fake events separately and based on the results of the closure test, we assign a 50% systematic uncer-

tainty on the combined estimate.

### Contamination from Signal Events

Signal contamination enters when there is a significant source of two “real” leptons, with one or both failing the numerator selections, but passing the denominator cuts and comprising a significant fraction of the total number of single fake or double fake counts. An additional simulation based correction is applied where simulated truth information is used to select events with two real same-sign dileptons failing the numerator selections but passing the denominator selection. These datasets correspond to the rare SM samples discussed in Section 3.1. These events are weighted using the same fake rate as we apply to data and then subtracted from the data driven prediction for fake events. This corrected fake prediction is then used in the analysis.

**Table 5.4:** Fake rate closure tests on  $t\bar{t}$  and  $W + \text{jets}$  events from simulation applied to a fake derived from QCD simulation. The  $t\bar{t}$  closure test shows results for the baseline search regions with exactly zero, exactly one and at least two  $b$ -tagged jets (SR0, SR10, SR20). Due to limited statistics, the  $W + \text{jets}$  closure test is only performed on the baseline search region with a  $b$ -tagged jet veto applied (SR0). “p” and “o” refer to the predicted and the observed counts, respectively.

Sample	search region	result	$ee$	$\mu\mu$	$e\mu$	$\ell\ell$
$t\bar{t}$	SR0	pred	$756.8 \pm 194.8$	$830.9 \pm 13.0$	$1592.7 \pm 230.0$	$3180.5 \pm 301.6$
		obs	339	318	656	1313
		pred/obs	$2.23 \pm 0.59$	$2.61 \pm 0.15$	$2.43 \pm 0.36$	$2.42 \pm 0.24$
		(p-o)/p	$0.55 \pm 0.12$	$0.62 \pm 0.02$	$0.59 \pm 0.06$	$0.59 \pm 0.04$
	SR10	pred	$381.3 \pm 99.1$	$420.3 \pm 7.9$	$804.4 \pm 116.3$	$1606.0 \pm 153.0$
		obs	167	177	354	698
		pred/obs	$2.28 \pm 0.62$	$2.37 \pm 0.18$	$2.27 \pm 0.35$	$2.30 \pm 0.24$
		(p-o)/p	$0.56 \pm 0.12$	$0.58 \pm 0.03$	$0.56 \pm 0.07$	$0.57 \pm 0.04$
	SR20	pred	$59.0 \pm 14.9$	$45.3 \pm 2.1$	$99.0 \pm 13.7$	$203.4 \pm 20.3$
		obs	34	23	53	110
		pred/obs	$1.74 \pm 0.53$	$1.97 \pm 0.42$	$1.87 \pm 0.36$	$1.85 \pm 0.26$
		(p-o)/p	$0.42 \pm 0.18$	$0.49 \pm 0.11$	$0.46 \pm 0.10$	$0.46 \pm 0.07$
$W + \text{jets}$	SR0	pred	$23.7 \pm 5.9$	$4.5 \pm 0.7$	$33.7 \pm 7.3$	$61.9 \pm 9.4$
		obs	29	4	21	54
		pred/obs	$0.82 \pm 0.25$	$1.13 \pm 0.59$	$1.60 \pm 0.49$	$1.15 \pm 0.23$
		(p-o)/p	$-0.22 \pm 0.38$	$0.12 \pm 0.46$	$0.38 \pm 0.19$	$0.13 \pm 0.18$

## 5.2 Estimation of the Charge Mis-measurement Rate

As discussed in the Section 3.2.2 and the beginning of this chapter, a possible source of background is lepton charge mis-measurement. As with the fake leptons from the previous section, this may not be well modeled by simulation, so we use a data-driven method to estimate this background.

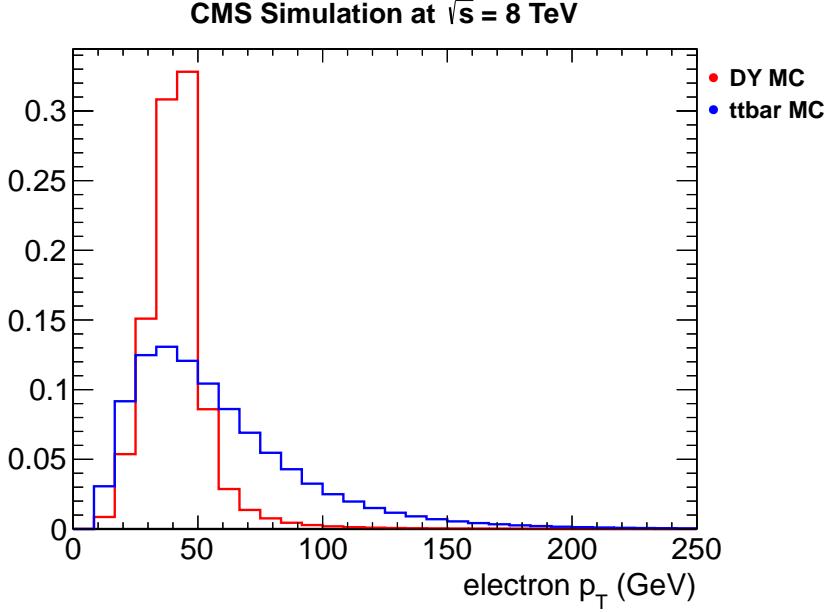
We define a charge flip as an electron whose charge is mis-reconstructed. We do not consider muons, as we have found previously that muon charge flips are several orders of magnitude smaller than electron charge flips [74]. Further, it is difficult to correctly reconstruct the muon  $p_T$  while incorrectly reconstructing the charge, as both come from the track (this effect is not present in electrons since the electron  $p_T$  is determined by the ECAL).

We expect the charge flip rate to depend on  $p_T$  and  $|\eta|$ . Events with a higher  $p_T$  will have a straighter track, and with less curvature, the charge will be more difficult to measure. Similarly, bremsstrahlung photons in events with higher  $|\eta|$  will be more likely to convert, since there is more material at higher  $|\eta|$ .

It is difficult to find a charge flip enriched data sample with kinematics similar to our search regions. Z events give the most charge flips, but have a very limited  $p_T$  distribution. We therefore choose to measure the flip rate in simulation, and to test the flip rate on data. We choose a combination of Drell-Yan and  $t\bar{t}$  simulated events because these samples have good statistics and a reasonable range of  $p_T$ , see Figure 5.7.

We calculate the flip rate as a fraction. The denominator is the number of electrons that pass the full selection requirements (see Section 4.2.3). The numerator is the number of denominator electrons with an incorrectly reconstructed charge sign according to truth information from the simulation. We further calculate this in bins of  $p_T$  and  $|\eta|$ . The result is shown in table 5.5. Notice that the flip rate has a very significant dependence on both  $p_T$  and  $|\eta|$ .

We validate by making a prediction and an observation using the flip rate applied to data. We start by selecting electron pairs that pass the analysis  $p_T$ ,  $|\eta|$ , and electron ID and isolation requirements which are listed in Section 4.2.3. Also, these events must pass the dilepton trigger (for data) from Section 4.2.1, and the dilepton invariant mass should be between 76 GeV and 106 GeV for exactly one pair of electrons. The observation



**Figure 5.7:** The  $p_T$  distribution for electrons that pass the selections from Section 4.2.3 for Drell-Yan and  $t\bar{t}$  simulated events.

**Table 5.5:** Electron charge-flip rate as a function of  $p_T$  and  $|\eta|$ .

$ \eta  \diagdown p_T$	0-20 GeV	20-40 GeV	40-60 GeV	60-80 GeV	80-200 GeV
0.0-1.0	$2.52 \times 10^{-05} \pm 2.52 \times 10^{-05}$	$3.44 \times 10^{-05} \pm 6.18 \times 10^{-06}$	$2.06 \times 10^{-05} \pm 4.38 \times 10^{-06}$	$1.03 \times 10^{-04} \pm 3.27 \times 10^{-05}$	$9.16 \times 10^{-05} \pm 4.58 \times 10^{-05}$
1.0-2.0	$1.67 \times 10^{-04} \pm 6.29 \times 10^{-05}$	$1.81 \times 10^{-04} \pm 1.82 \times 10^{-05}$	$2.16 \times 10^{-04} \pm 1.94 \times 10^{-05}$	$5.31 \times 10^{-04} \pm 1.00 \times 10^{-04}$	$9.87 \times 10^{-04} \pm 2.06 \times 10^{-04}$
2.0-2.4	$0.00 \times 10^{+00} \pm 4.52 \times 10^{-05}$	$2.96 \times 10^{-04} \pm 4.52 \times 10^{-05}$	$3.96 \times 10^{-04} \pm 5.49 \times 10^{-05}$	$1.34 \times 10^{-03} \pm 3.34 \times 10^{-04}$	$8.79 \times 10^{-04} \pm 4.39 \times 10^{-04}$

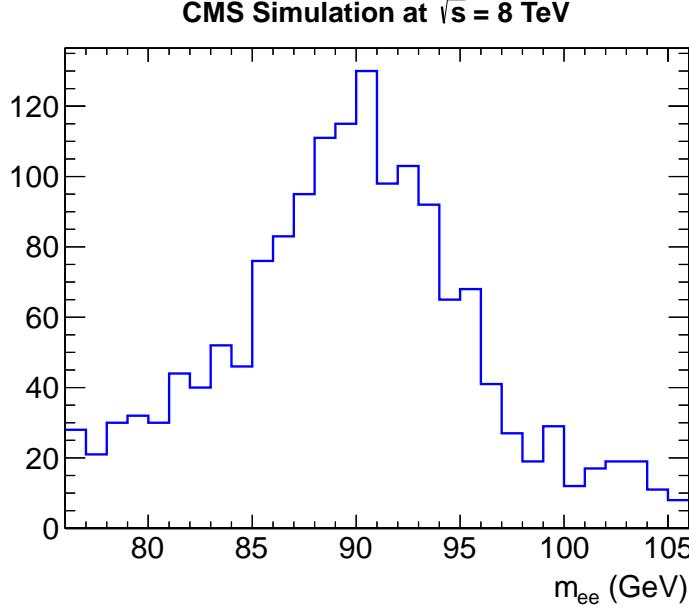
is then simply the number of same sign dilepton events occurring in the above dilepton selection. For the prediction, the flip rate (which was derived in simulation) is applied to this selection: we select the opposite-sign dielectron events in the sample, weighting each one with a factor of

$$\frac{\varepsilon_{fl,1}}{1 - \varepsilon_{fl,1}} + \frac{\varepsilon_{fl,2}}{1 - \varepsilon_{fl,2}}$$

where the flip rate for the  $n$ th electron,  $\varepsilon_{fl,n}$ , is taken from table 5.5. The denominator is necessary because we are not summing over the same-sign dielectron pairs.

In data, we predict 944 same-sign events and observe 1561 events. The invariant mass distribution for these 1561 events is shown in figure 5.8. However, we must investigate the possibility of non-Z events occurring in the data sample, causing our

observation to be artificially high.



**Figure 5.8:** The  $m_{ee}$  distribution for same-sign electron pairs in data.

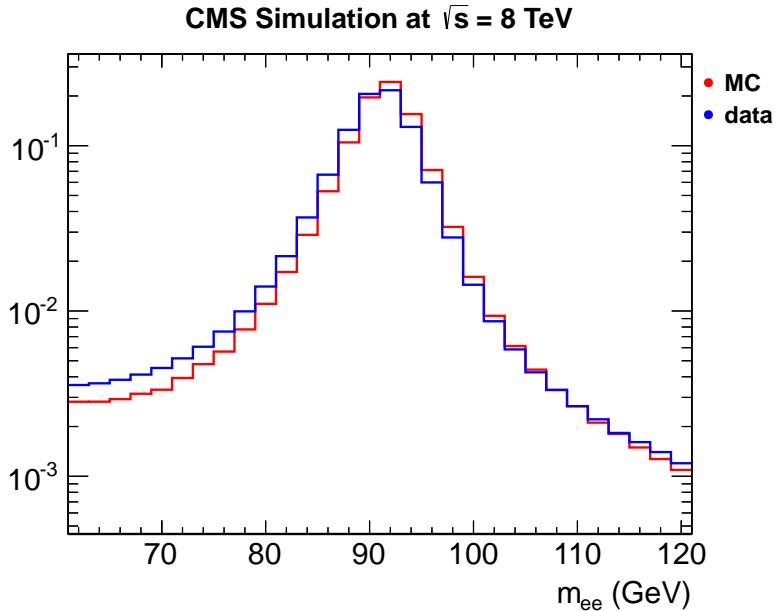
We consider the upper sidebands in the dilepton mass plot of the data charge flips ( $106 \text{ GeV} < m_{ee} < 121 \text{ GeV}$ ). Figure 5.9 shows that events in this region are unlikely to be flips from high mass Zs, and so we may assume as an approximation that everything in this region is a contamination. Figure 5.8 shows that a significant number of electrons are indeed predicted in this region. In order to compare with the prediction from simulation, we subtract these events from the full data prediction. We extrapolate from this figure that about 16% of the observed flips are actually contamination rather than charge flips. We use this factor to adjust our number of observed charge flips.

The result is that we expect 944 charge flips, and we observe 1311. We reconcile these two differences with a scale factor:

$$SF = \text{observed/predicted} = 1.39$$

So the flip rates in Table 5.5, multiplied by the scale factor, should give a good approximation of the overall flip rate.

This remaining discrepancy, coupled with the kinematic limitation of the control



**Figure 5.9:** The  $m_{ee}$  distribution for opposite-sign electron pairs in data.

sample, allow for considerable error in our estimation of the charge flip background. We therefore assign an overall systematic of  $\pm 30\%$  to allow for these.

### 5.3 Rare SM background estimation

Backgrounds arising from failures of lepton identification and electron charge mis-reconstruction together account for 40-70% of the estimated background in a given search region. The remaining 30-60% of the background is irreducible and is taken directly from simulation. A list of the process and associated cross sections is provided in Table 4.3 and repeated here for convenience. Note that we neglect contributions from  $W\gamma^* \rightarrow \ell\nu ee$  because the simulated dataset was known to be faulty and this process contributes less than a percent to the irreducible background total.

Most of the background involving rare SM process have never been measured before. The cross sections used for these processes are next-to-leading (NLO) order and are used to normalize the rare predictions. The cross section values used for the most relevant processes,  $t\bar{t}W$  and  $t\bar{t}Z$ , are 232 fb and 208 fb, respectively [75, 76].

**Table 5.6:** Sources of true same-sign dileptons from Standard Model processes. Cross sections are next-to-leading order. The equivalent integrated luminosity of these simulated events is listed in the column on the far right. The contributions from these process to the background is taken directly from simulation.

Sample	Cross Section (pb)	Equivalent Luminosity ( $\text{fb}^{-1}$ )
$t\bar{t}Z$	0.208	1021
$t\bar{t}W$	0.232	845
$t\bar{t}WW$	0.00204	106931
$t\bar{t}\gamma$	2.17	33
$t\bar{b}Z$	0.0114	13026
$ZZZ$	0.00554	40692
$WWW$	0.0822	2737
$WW\gamma$	0.528	407
$WZZ$	0.0192	12946
$WWZ$	0.0580	3832
$ZZ$	0.177	27177
$WZ$	1.058	1908
$qqW^-W^-$	0.0889	1084
$qqW^+W^+$	0.248	402
$W^\pm W^\pm(\text{DPS})$	0.588	1418
$W\gamma^* \rightarrow \ell\nu\mu\mu$	1.91	156
$W\gamma^* \rightarrow \ell\nu\tau\tau$	0.336	148
$WH, ZH, t\bar{t}H; H \rightarrow WW$	0.260	769
$WH, ZH, t\bar{t}H; H \rightarrow ZZ$	0.0320	15652
$WH, ZH, t\bar{t}H; H \rightarrow \tau\tau$	0.0177	5478

The systematic uncertainty for the rare SM backgrounds accounts for the theoretical uncertainty on the cross sections as well as for the ratio between LO (as in simulation) and NLO cross-sections being non-uniform as a function of jet multiplicity,  $H_T$  and  $\cancel{E}_T$ . Due to this, a 50% systematic uncertainty for the prediction from these rare SM processes is assigned.

# Chapter 6

## Selection Efficiencies

Selection efficiencies is one of the ingredients needed to measure a cross section or set an exclusion limit. The combined selection efficiency represents the fraction of events within the acceptance that are also reconstructed and pass the analysis selections. Acceptances here is defined as the probability for a generated object from simulation to actually traverse through the detector. It is customary to measure the efficiency for the various aspects of the selections and then combine them as a measure of the total efficiency of the analysis.

To determine where a specific model is consistent with the data, the same analysis selections are applied to a simulated sample of this model. Here, the simulation has been generated with a Monte Carlo generator that has been passed through the detector simulation and then the full off-line reconstruction. No trigger requirement is made in the simulation. Instead, we scale the simulated events by the trigger efficiency measured for the relevant signature. Similarly, differences exist in the efficiency of the lepton identification and isolation requirements, as well as the b-tagging efficiency between data and simulation. Scale factors are measured and then applied to simulation to account for these differences.

In this chapter we describe the techniques used to measure the selection efficiencies and acceptances. The first section describes the measurement of the lepton efficiencies using the tag-and-probe method and its associated uncertainties. The next section discusses the efficiencies for the various trigger requirements. Finally, the data-to-simulation scale factor for b-tagged jets is shown and the procedure used to apply it

is discussed.

## 6.1 Lepton Efficiencies

The efficiencies of the lepton isolation and identification requirements (including all quality requirements) are measured with the tag-and-probe method in dilepton Z events using the full 2012 dataset. The efficiency of the identification requirements is a property of the lepton itself and is directly applicable to the leptons in signal events. The efficiency of the isolation requirement, however, is a strong function of all other (mainly hadronic) activity in the event.

In this section, the lepton efficiencies are measured and their data-to-simulation scale factors are computed. The systematic uncertainty for this method is assigned for both the tag-and-probe method itself and due to the fact that the efficiency is measured in Z events; however, our signal events will have more hadronic activity and are more like  $t\bar{t}$  events. To account for this, a short study is also done to show the validity of measuring the efficiency in Z events and applying them to signal-like events.

### 6.1.1 Tag-and-Probe Method

The basic tag-and-probe method essentially takes advantage of the fact that we can reconstructed  $Z \rightarrow \ell\ell$  events with high accuracy. We select  $Z \rightarrow \ell^+\ell^-$  events from a lepton triggered sample by requiring opposite-sign, same-flavor leptons in a window around the nominal Z mass,  $|m_Z - m_{\ell\ell}| < 15$  GeV. Out of window events are not used to extract efficiencies but are retained as sidebands to monitor the contaminations from backgrounds. Contributions from  $W$  and QCD processes are found to be small. An event is “tagged” by requiring one lepton to pass the full lepton selection described in Chapter 4 and to be matched to an online trigger object by requiring  $\Delta R < 0.1$  to that object. Tagged events are classified by the other “probe” leptons as either

- Pass: The probe passes the full selection.
- Fail: The probe fails the full selection but passes a looser criteria chosen to not bias the selection under study.

The ratio of the count of the passing probe leptons to the total number of tagged events (both passing and failing probes) is the measured efficiency. The efficiency is extracted in bins of  $p_T$  and  $|\eta|$  for both electrons and muons separately.

### Tag-and-Probe Results for Electrons

The electron selection efficiencies are measured in events passing the trigger, which require one well-identified electron passing the WP80 electron ID and with  $p_T > 27$  GeV (see Table 4.6). The tag electron is required to match the well-identified electron from the HLT\_Ele27\_WP80 trigger and also to pass all the electron requirements described in Section 4.2.3. Additionally, the tag electron is chosen to have a  $p_T > 32$  GeV to eliminate turn-on effects from the trigger efficiency. The probe electron is required to pass a simple loose selection as follows

- reconstructed electron
- $p_T > 10$  GeV,
- $|\eta| < 2.4$ ,
- excluding electrons with a supercluster between  $1.4442 < |\eta_{SC}| < 1.566$ .

The isolation efficiency is measured with the probes passing all electron selections described in Section 4.2.3, except for the isolation itself. The identification efficiency is measured with probes passing the isolation requirement. The overall efficiency is measured with probes passing only the nominal requirements above. Results of the measurement are summarized in Table 6.2. The contribution from the Z events in data is based on fitting the mass range 60-120  $\text{GeV}/c^2$  using a simultaneous fit of various models for the signal and backgrounds. The models used are summarized in the following Table 6.1. Since the kinematics of the background shape varies in the different  $p_T - |\eta|$  bins, different background shapes were chosen on a bin-by-bin basis in order to get the best fits. The results of the tag-and-probe show a strong dependence on the lepton  $p_T$  and a weaker but non-negligible dependence on  $|\eta|$ .

The following sources of systematic uncertainty are attributed to this measurement: variation in the models used to fit the signal and background distributions and

**Table 6.1:** Models used for measuring the signal or the background contribution in the tag-and-probe method.

Model	Usage
Breit-Wigner function * Crystal-Ball function	Signal
MC-based template function	Signal
Exponential	Background
Exponential * error function	Background
Polynomial	Background
Polynomial $\times$ exponential function	Background
Chebyshev Polynomial	Background

**Table 6.2:** Electron efficiencies measured using the tag-and-probe method. The uncertainties are statistical only.

$ \eta  \setminus p_T$		10 – 15 GeV	15 – 20 GeV	20 – 30 GeV	30 – 40 GeV	40 – 50 GeV	50 – 200 GeV
0.0 – 0.8	MC	0.363 $\pm$ 0.004	0.503 $\pm$ 0.002	0.646 $\pm$ 0.001	0.764 $\pm$ 0.000	0.819 $\pm$ 0.000	0.841 $\pm$ 0.001
	Data	0.303 $\pm$ 0.003	0.462 $\pm$ 0.002	0.617 $\pm$ 0.001	0.733 $\pm$ 0.000	0.796 $\pm$ 0.000	0.815 $\pm$ 0.001
	Data/MC	0.834 $\pm$ 0.012	0.918 $\pm$ 0.006	0.954 $\pm$ 0.002	0.960 $\pm$ 0.001	0.972 $\pm$ 0.001	0.969 $\pm$ 0.001
0.8 – 1.4442	MC	0.379 $\pm$ 0.004	0.480 $\pm$ 0.003	0.600 $\pm$ 0.001	0.736 $\pm$ 0.001	0.809 $\pm$ 0.000	0.830 $\pm$ 0.001
	Data	0.369 $\pm$ 0.008	0.435 $\pm$ 0.003	0.554 $\pm$ 0.001	0.688 $\pm$ 0.001	0.772 $\pm$ 0.000	0.794 $\pm$ 0.001
	Data/MC	0.973 $\pm$ 0.023	0.906 $\pm$ 0.009	0.923 $\pm$ 0.003	0.935 $\pm$ 0.001	0.955 $\pm$ 0.001	0.956 $\pm$ 0.001
1.566 – 2.0	MC	0.206 $\pm$ 0.004	0.344 $\pm$ 0.003	0.482 $\pm$ 0.002	0.615 $\pm$ 0.001	0.681 $\pm$ 0.001	0.697 $\pm$ 0.001
	Data	0.197 $\pm$ 0.004	0.313 $\pm$ 0.003	0.444 $\pm$ 0.002	0.569 $\pm$ 0.001	0.647 $\pm$ 0.001	0.693 $\pm$ 0.001
	Data/MC	0.954 $\pm$ 0.028	0.909 $\pm$ 0.012	0.921 $\pm$ 0.005	0.924 $\pm$ 0.002	0.950 $\pm$ 0.001	0.995 $\pm$ 0.002
2.0 – 2.4	MC	0.199 $\pm$ 0.004	0.321 $\pm$ 0.004	0.419 $\pm$ 0.002	0.515 $\pm$ 0.001	0.576 $\pm$ 0.001	0.593 $\pm$ 0.002
	Data	0.223 $\pm$ 0.005	0.303 $\pm$ 0.004	0.416 $\pm$ 0.000	0.494 $\pm$ 0.001	0.558 $\pm$ 0.001	0.575 $\pm$ 0.002
	Data/MC	1.119 $\pm$ 0.036	0.944 $\pm$ 0.015	0.993 $\pm$ 0.004	0.959 $\pm$ 0.003	0.968 $\pm$ 0.002	0.969 $\pm$ 0.004

measuring the isolation (Iso) and identification (ID) components simultaneously (ID+Iso) as opposed to a factorized measurement (ID\*Iso). Varying the models for signal and backgrounds seems to make the most significant difference to the data-to-MC scale factors in the lowest  $p_T$  bins where the backgrounds are the largest and least understood. The effect was sub-dominant to the difference in factorization scheme (using ID+Iso vs ID\*Iso). To show the difference between the data-to-MC scale factors using different factorization schemes, the relative difference of the data-to-MC scale factors when measured using ID+Iso versus ID\*Iso is shown in the following Table 6.3. As you move to lower  $p_T$  and higher  $|\eta|$ , we see relatively larger differences between the two factorization schemes with some differences as high as  $\sim 25\%$ . The two methods offer a trade

off with the ID+Iso method accounting for the correlation between the ID and isolation; whereas, the ID\*Iso does not account for the correlation but has a smaller background. This leads to differences in the efficiencies between the two methods. At higher  $p_T$  and lower  $|\eta|$ , we see this effect is mitigated with relative differences on the order of a few percent. To account for these differences, we apply at 10% systematic uncertainty for electrons with  $p_T < 15$  GeV where we saw larger differences and 5% above where the differences settle out.

**Table 6.3:** Electron efficiency data-to-MC scale factors measured using the tag-and-probe method in different factorization schemes. “rel diff” refers to the relative difference between the two factorization schemes. The uncertainties are statistical only.

$ \eta  \backslash p_T$		10 – 15 GeV	15 – 20 GeV	20 – 30 GeV	30 – 40 GeV	40 – 50 GeV	50 – 200 GeV
0.0 – 0.8	ID*Iso	0.787 ± 0.022	0.943 ± 0.010	0.963 ± 0.002	0.969 ± 0.001	0.975 ± 0.000	0.973 ± 0.001
	ID+Iso	0.834 ± 0.012	0.918 ± 0.006	0.954 ± 0.002	0.960 ± 0.001	0.972 ± 0.001	0.969 ± 0.001
	rel diff	0.055 ± 0.030	-0.028 ± 0.013	-0.009 ± 0.003	-0.010 ± 0.001	-0.004 ± 0.001	-0.004 ± 0.002
0.8 – 1.4442	ID*Iso	0.861 ± 0.015	0.910 ± 0.011	0.921 ± 0.002	0.943 ± 0.001	0.956 ± 0.001	0.950 ± 0.002
	ID+Iso	0.973 ± 0.023	0.906 ± 0.009	0.923 ± 0.003	0.935 ± 0.001	0.955 ± 0.001	0.956 ± 0.001
	rel diff	0.115 ± 0.029	-0.005 ± 0.015	0.003 ± 0.004	-0.009 ± 0.001	-0.001 ± 0.001	0.006 ± 0.003
1.566 – 2.0	ID*Iso	0.798 ± 0.028	0.886 ± 0.026	0.910 ± 0.005	0.928 ± 0.002	0.948 ± 0.001	0.964 ± 0.002
	ID+Iso	0.954 ± 0.028	0.909 ± 0.012	0.921 ± 0.005	0.924 ± 0.002	0.950 ± 0.001	0.995 ± 0.002
	rel diff	0.164 ± 0.042	0.025 ± 0.031	0.012 ± 0.008	-0.004 ± 0.003	0.002 ± 0.002	0.031 ± 0.003
2.0 – 2.4	ID*Iso	0.866 ± 0.027	0.929 ± 0.015	0.935 ± 0.005	0.954 ± 0.002	0.962 ± 0.002	0.955 ± 0.004
	ID+Iso	1.119 ± 0.036	0.944 ± 0.015	0.993 ± 0.004	0.959 ± 0.003	0.968 ± 0.002	0.969 ± 0.004
	rel diff	0.226 ± 0.041	0.016 ± 0.023	0.058 ± 0.007	0.006 ± 0.004	0.006 ± 0.003	0.015 ± 0.006

### Tag-and-Probe Results for Muons

For muons, the technique for electrons is essentially repeated with appropriate modifications for muon specific selections. The muon selection efficiencies are measured using events that pass the HLT\_IsoMu24\_eta2p1 trigger which require one isolated muon with a  $p_T > 24$  GeV. In the tag-and-probe analysis, the tag muon is required to match the well-identified muon from the trigger and also to pass all the muon requirements described in Section 4.2.2. Additionally, the tag muon is chosen to have a  $p_T > 30$  GeV to eliminate turn-on effects from the trigger efficiency. The probe muon is required to have

- $p_T > 10$  GeV

- $|\eta| < 2.4$
- have both the global and the particle-flow muon types.

The isolation efficiency is measured with the probes passing all muon selections described in Section 4.2.2, except for the isolation itself. The identification efficiency is measured with probes passing the isolation requirement. The overall efficiency is measured with probes passing only the nominal requirements above. Results of the measurement are summarized in Table 6.4. The contribution from the Z events in data is based on fitting the mass range of 60-120 GeV/ $c^2$  using a simultaneous fit of various models for the signal and backgrounds. The models used are the same as those used by the electrons and are summarized in the same Table 6.1. Since the kinematics of the background shape varied in the different  $p_T - |\eta|$  bins, different models were chosen on a bin-by-bin basis in order to get the best fits.

**Table 6.4:** Muon efficiencies measured using the tag-and-probe method. The uncertainties are statistical only.

$ \eta  \diagdown p_T$		10 – 15 GeV	15 – 20 GeV	20 – 30 GeV	30 – 40 GeV	40 – 50 GeV	50 – 200 GeV
0.0 – 1.2	MC	0.582 ± 0.003	0.644 ± 0.002	0.754 ± 0.001	0.857 ± 0.000	0.901 ± 0.000	0.900 ± 0.000
	Data	0.556 ± 0.005	0.617 ± 0.002	0.727 ± 0.001	0.832 ± 0.000	0.881 ± 0.000	0.877 ± 0.001
	Data/MC	0.956 ± 0.010	0.957 ± 0.004	0.964 ± 0.001	0.971 ± 0.000	0.978 ± 0.000	0.974 ± 0.001
1.2 – 2.4	MC	0.646 ± 0.003	0.682 ± 0.002	0.763 ± 0.001	0.861 ± 0.000	0.907 ± 0.000	0.895 ± 0.001
	Data	0.620 ± 0.010	0.662 ± 0.003	0.731 ± 0.001	0.843 ± 0.000	0.892 ± 0.000	0.875 ± 0.001
	Data/MC	0.960 ± 0.016	0.971 ± 0.005	0.959 ± 0.001	0.978 ± 0.001	0.984 ± 0.000	0.977 ± 0.001

Since the technique for the muons is the same as the electrons, we consider the same sources of systematic uncertainties. As with the electrons, varying the fitting models showed the most difference in the lower  $p_T$  bins for the same reason. Also, the factorization scheme was varied with the results shown in Table 6.5. As you move to lower  $p_T$  and higher  $|\eta|$ , we see differences between the two factorization schemes of  $\sim 5\%$ . At higher  $p_T$  and lower  $|\eta|$ , we see this effect is much smaller being a few percent or less. To account for these differences, we apply at 5% systematic uncertainty for muons with  $p_T < 15$  GeV where we saw larger differences and 3% above where the differences settle out.

### 6.1.2 Lepton Composition

In the previous sub-sections, we saw there there is an uncertainty associated with the data-to-simulation scale factors on the lepton efficiency. Another source of uncertainty in the efficiencies measured above stems from the fact that the efficiency is extracted from Drell-Yan (DY) events; however, the individual efficiencies for ID and isolation for non-DY events may differ. A Monte Carlo study was done to measure the ID and isolation efficiencies in  $Z \rightarrow \ell\ell$  with respect to  $t\bar{t} \rightarrow \ell\nu q\ell\nu q$  events. The exact datasets used are listed in Appendix B.2. The denominator selection for these efficiencies was any generator-level  $e/\mu/\tau$  ( $\tau \rightarrow e/\mu$ ) which is truth matched to a W/Z boson decay at the parton level (Pythia6 status 3) within the acceptance and  $p_T > 10$  GeV. The numerator selection requires that this generated level lepton is truth matched to a reconstructed lepton of the appropriate flavor which passes the full analysis selection outlined in Chapter 4. The results from this study are shown in Figures 6.1, 6.2 and in Tables 6.6, 6.7 for electrons and muons, respectively.

Here we see that the majority of the differences come at low  $p_T$  where we see relative differences on the order of 10% for muons and 5% for the electrons. As we move to higher  $p_T$  the differences are only on the order of a few percent. To account for these differences we assign a systematic uncertainty which is roughly half the relative difference stated in tables 6.6, 6.7:

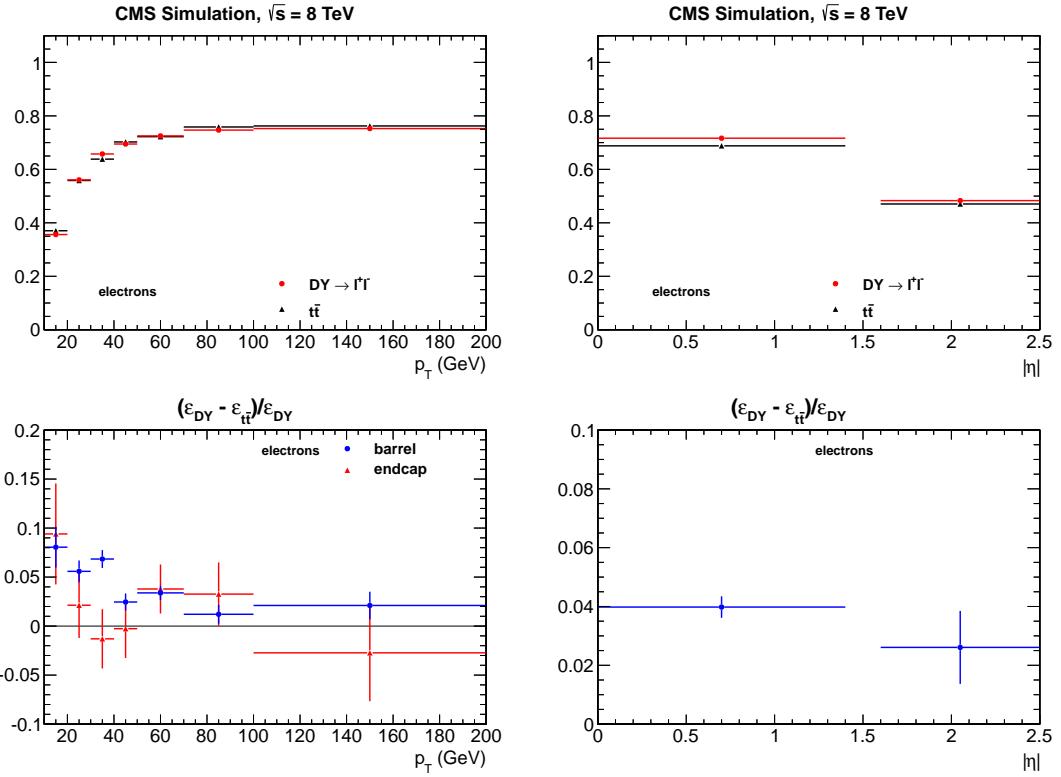
- muons: 5% for  $p_T < 30$  GeV and 3% for  $p_T > 30$  GeV
- electrons: 3%

**Table 6.5:** Muon efficiency data-to-MC scale factors measured using the tag-and-probe method in different factorization schemes. “rel diff” refers to the relative difference between the two factorization schemes. The uncertainties are statistical only.

$ \eta $	$p_T$	10 – 15 GeV	15 – 20 GeV	20 – 30 GeV	30 – 40 GeV	40 – 50 GeV	50 – 200 GeV
0.0 – 1.2	ID*Iso	$0.918 \pm 0.017$	$0.955 \pm 0.008$	$0.965 \pm 0.001$	$0.972 \pm 0.000$	$0.979 \pm 0.000$	$0.974 \pm 0.001$
	ID+Iso	$0.956 \pm 0.010$	$0.957 \pm 0.004$	$0.964 \pm 0.001$	$0.971 \pm 0.000$	$0.978 \pm 0.000$	$0.974 \pm 0.001$
	rel diff	$0.040 \pm 0.021$	$0.003 \pm 0.009$	$-0.001 \pm 0.002$	$-0.001 \pm 0.001$	$-0.001 \pm 0.000$	$-0.000 \pm 0.001$
1.2 – 2.5	ID*Iso	$0.962 \pm 0.011$	$1.034 \pm 0.006$	$0.972 \pm 0.001$	$0.980 \pm 0.001$	$0.985 \pm 0.000$	$0.980 \pm 0.001$
	ID+Iso	$0.960 \pm 0.016$	$0.971 \pm 0.005$	$0.959 \pm 0.001$	$0.978 \pm 0.001$	$0.984 \pm 0.000$	$0.977 \pm 0.001$
	rel diff	$-0.003 \pm 0.020$	$-0.065 \pm 0.008$	$-0.013 \pm 0.002$	$-0.002 \pm 0.001$	$-0.001 \pm 0.001$	$-0.003 \pm 0.001$

**Table 6.6:** The electron efficiency relative differences between DY and  $t\bar{t}$  ( $(\varepsilon_{DY} - \varepsilon_{t\bar{t}})/\varepsilon_{DY}$ ). The uncertainties are statistical only.

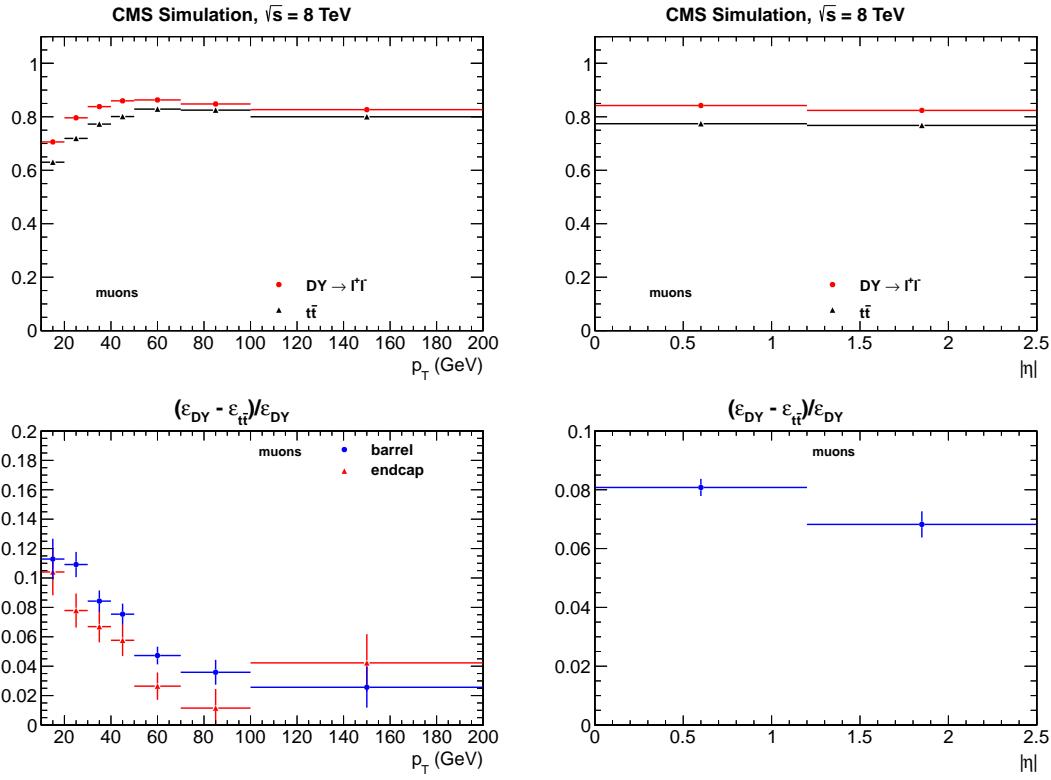
$p_T \backslash  \eta $	0.00 - 1.44	1.57 - 2.50
10 - 20 GeV	$0.080 \pm 0.021$	$0.094 \pm 0.051$
20 - 30 GeV	$0.056 \pm 0.011$	$0.021 \pm 0.033$
30 - 40 GeV	$0.068 \pm 0.009$	$-0.013 \pm 0.030$
40 - 50 GeV	$0.025 \pm 0.009$	$-0.003 \pm 0.030$
50 - 70 GeV	$0.034 \pm 0.007$	$0.038 \pm 0.025$
70 - 100 GeV	$0.012 \pm 0.010$	$0.033 \pm 0.032$
100- 200 GeV	$0.021 \pm 0.014$	$-0.027 \pm 0.049$



**Figure 6.1:** The electron efficiency measured using DY (red) and  $t\bar{t}$  MC (black). The top two plots are the efficiencies projected vs  $p_T$  and  $|\eta|$ , respectively. The bottom two plots are the relative differences between DY and  $t\bar{t}$  also projected vs  $p_T$  and  $|\eta|$ .

**Table 6.7:** The muon efficiency relative differences between DY and  $t\bar{t}$  ( $(\varepsilon_{DY} - \varepsilon_{t\bar{t}})/\varepsilon_{DY}$ ). The uncertainties are statistical only.

$p_T$	$ \eta $	0.00 - 1.20	1.20 - 2.50
10 - 20 GeV		$0.113 \pm 0.014$	$0.104 \pm 0.016$
20 - 30 GeV		$0.109 \pm 0.009$	$0.078 \pm 0.012$
30 - 40 GeV		$0.084 \pm 0.007$	$0.067 \pm 0.011$
40 - 50 GeV		$0.075 \pm 0.007$	$0.058 \pm 0.011$
50 - 70 GeV		$0.047 \pm 0.006$	$0.026 \pm 0.009$
70 - 100 GeV		$0.036 \pm 0.009$	$0.012 \pm 0.013$
100- 200 GeV		$0.026 \pm 0.014$	$0.042 \pm 0.020$



**Figure 6.2:** The muon efficiency measured using DY (red) and  $t\bar{t}$  MC (black). The top two plots are the efficiencies projected vs  $p_T$  and  $|\eta|$ , respectively. The bottom two plots are the relative differences between DY and  $t\bar{t}$  also projected vs  $p_T$  and  $|\eta|$ .

To account for the differences in the efficiency between data and simulation, the scale factors reported in Tables 6.6 and 6.7 are applied to each MC sample used in the prediction with Table 6.8 summarizing the various systematic uncertainties taken.

**Table 6.8:** The systematic uncertainties applied to the data-to-MC efficiency scale factors.

	lepton flavor	$p_T < 15 \text{ GeV}$	$p_T > 15 \text{ GeV}$
tag-and-probe	electron	10%	5%
	muon	5%	3%
composition		$p_T < 30 \text{ GeV}$	$p_T > 30 \text{ GeV}$
	electron	3%	3%
	muon	5%	3%

## 6.2 Trigger Efficiencies

The trigger efficiencies and data-to-MC scale factor is calculated for every trigger and is applied  $p_T$  and  $|\eta|$  dependent for each lepton. Details on how the scale factors have been studied and calculated can be found in [60]. The value of those scale factors is shown in Table 6.9.

**Table 6.9:** Scale factors from trigger inefficiencies applied to Monte Carlo predictions from the  $e\mu$  irreducible backgrounds. For some channels, scale factors are parametrized by the trailing lepton  $p_T$  or  $|\eta|$ .

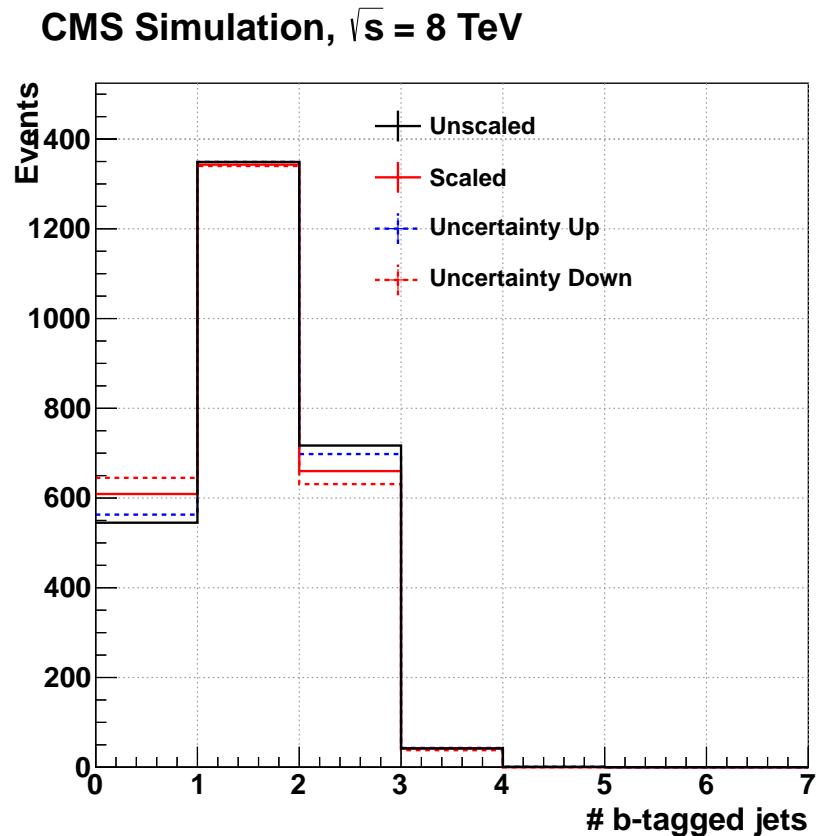
Low- $p_T$	Scale Factor	High- $p_T$	Scale Factor
$\mu\mu,  \eta  < 1$	0.94	$\mu\mu,  \eta  < 1$	0.90
$\mu\mu,  \eta  > 1$	0.90	$\mu\mu,  \eta  > 1$	0.81
$e\mu$	0.93	$e\mu$	0.93
$ee$	0.93	$ee, p_T < 30$	0.92
		$ee, p_T > 30$	0.96

## 6.3 B-tagging Efficiency

Since the b-tagging efficiencies measured in data are somewhat different than those measured by simulation, a scale factor is applied to simulated events to take this difference into account. The CMS B-tagging and Vertex Group (BTV) has measured the b-tagging scale factors between data and MC. The scale factors in general depend on the jet flavor,  $p_T$ , and  $|\eta|$  and the details are described by the BTV in [77, 78]. These scale factors were measured on muon-jet and  $t\bar{t}$  data and are documented in [78]. In contrast to the previous iterations of this analysis where the scale factor was applied as an additional event weight, the scale factor will now be used to update the b-tagging status on a jet-by-jet basis. A full description is provided in the stated reference; however, we provide a short summary of the method here. In this method, scale factors are used to update the b-tagging status on each jet individually. In order to upgrade or downgrade the b-tagging status of each individual jet, a random number generator is used. This gives the advantage that the yields from MC can be treated the same way as data. The method in our case is relatively simple since we only have one operating point. If  $SF < 1$ , then the fraction of  $f = 1 - SF$  b-tagged jets from the “tagged” collection are to be downgraded to the “non-tagged” status and in this case, it is not necessary to know the MC b-tagging efficiency. The situation gets more complicated when  $SF > 1$ . It is necessary to upgrade the b-tagging status of some of the untagged jets and the fraction of such jets that needs to be upgraded is

$$f = \frac{1 - SF}{1 - 1/\varepsilon_{MC}}.$$

Using this relationship and a random number, we recount the number of b-tagged jets on all events and use the count of this new collection in all selections regarding the #  $b$ -tagged jets . To estimate the systematic uncertainty on the b-tagging efficiency scale factor, the uncertainty is scaled up or down and then the above procedure is repeated. The difference between the unscaled and the scaled acceptance efficiency is used in determining the systematic uncertainty. These are considered fully correlated between the search regions. Figure 6.3 shows the effect of the this rescaling procedure on the #  $b$ -tagged jets distribution on a  $t\bar{t}W$  MC sample which is required to pass the baseline selections (SR0).



**Figure 6.3:**  $\# b$ -tagged jets distribution for  $t\bar{t}W$  MC sample to illustrate the effect of the b-tag rescaling procedure on the number of b-tagged jets.

# Chapter 7

## Results

We have conducted a search for new physics in a final state with same-sign leptons, missing transverse energy, and significant hadronic activity. The selection for this analysis was discussed in Chapter 4, which described search regions sensitive to new physics, while rejecting a large amount of the SM background. Observed events yields in the baseline control regions are compared with the background estimates whose methods were described in Chapter 5. The event yields in the search regions defined in Section 4.3 are reported with their associated estimated background yield prediction. No new physics is observed.

Given the lack of a significant excess over the predicted backgrounds, the results of this search are used to place limits on the cross section of rare SM processes and on the parameters for various model of new physics. The 95% confidence level (CL) upper limit on the signal  $\sigma_{prod} \times BR$  is calculated. For each model considered, limits are obtained by performing a statistical combination of the most sensitive, exclusive, search regions.

The selections chosen for this analysis were motivated by considering generic new physics signatures and their Standard Model backgrounds. Although opposite-sign dileptons are significantly more prevalent in the SM than same-sign, any new physics model with pairs of Majorana particles will produce the two in equal abundance. Astrophysical evidence for dark matter suggests the existence of a weakly interacting, massive particle; thus search regions with significant  $E_T$  are motivated. Finally, particles produced via strong interactions tend to have larger cross sections than those produced

via electroweak interactions; therefore, final states with significant hadronic activity are thus likely to require a smaller integrated luminosity for observation. These motivating factors are seen in a variety of models such as SUSY, models with extra dimensions, Majorana neutrinos, and same-sign top resonances (See discussion in Chapters 1 and 3).

This chapter consists of three sections. The first section discusses the results of the high  $p_T$  analysis starting with a discussion of the yields for the three baseline regions. Results from each search region follow where we discuss the variation in the background composition with changed # jets, #  $b$ -tagged jets,  $H_T$  and  $\cancel{E}_T$  requirements. The following section provides an analogous discussion of the low  $p_T$  analysis. The final section discusses limits set on various models of new physics.

## 7.1 High $p_T$ Results

### 7.1.1 High $p_T$ Control Regions

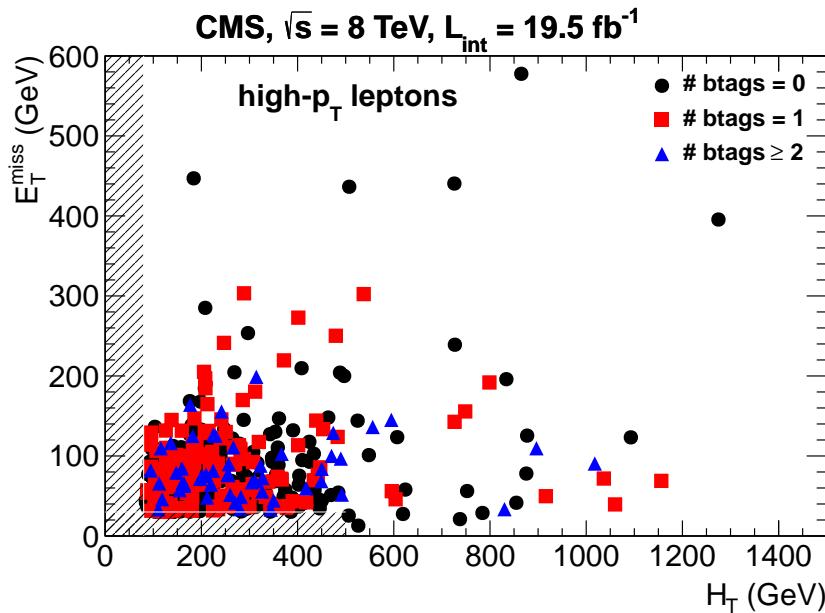
In Section 4.3, we defined a control region enriched in background by requiring a minimal selection. These baseline selections was listed in Table 4.8 but are repeated here for convenience in Table 7.1. SR0 is designed to be inclusive to all the search regions

**Table 7.1:** Summary of the baseline search regions considered in the high  $p_T$  and low  $p_T$  analysis.

Search Region #	min lepton $p_T$ ( $\mu, e$ ) (GeV)	$H_T$ (GeV)	$\cancel{E}_T$ (GeV)	# jets	# $b$ -tagged jets
SR0					$\geq 0$
SR10	20, 20	80	30 if $H_T < 500$ else 0	2	= 1
SR20					$\geq 2$

and is mostly sensitive to fake leptons in SM multi-jet processes: predominately  $t\bar{t}$  and  $W + \text{jets}$ . The contribution from backgrounds without intrinsic  $b$ -quarks is suppressed in SR10 and SR20 since we add a #  $b$ -tagged jets requirement of exactly one and two or more, respectively. Figure 7.1 shows the distribution of events from SR0 in the  $H_T$ - $\cancel{E}_T$  plane. The events cluster in the lower left corner as expected for background events.

The observed event yields and estimated backgrounds in the high  $p_T$  control regions can be seen in Tables 7.2, 7.3, and 7.4. The yield reported as *MC Pred* includes



**Figure 7.1:** Distribution of events in the  $H_T$ - $E_T$  plane passing the analysis selections for the high  $p_T$  baseline region (SR0).

contributions from genuine same-sign lepton pairs (sum of the rows from the  $W\gamma$  down to the Higgs samples). The statistical uncertainty on the simulated contributions is calculated using the Clopper-Pearson method [79]. Systematic uncertainties (the second uncertainty if present) are displayed only for the final combined background component and no uncertainty is added for estimates with zero entries. Systematic uncertainties are 100% correlated among the channels. The upper portion of the table (from the top down to  $ZZ$ ) is based on simulation and is used as a reference only. The total background when taken from simulation only is reported as *Total MC* and again is given as a reference only. The SF (DF) contributions are for events with one (two) fake leptons and the SC refers to the contamination to the fake prediction from rare SM events.

**Table 7.2: Baseline with 0 btags (Signal Region 0):** Observed event yields in the high  $p_T$  baseline (lepton  $p_T > 20/20$  GeV,  $H_T > 80$  GeV,  $\cancel{E}_T > 30$  GeV if  $H_T < 500$  GeV (otherwise no  $\cancel{E}_T$  cut), at least 2 jets with  $p_T > 40$  GeV) compared to expectations from simulation alone, and from the data-driven methods. The upper part of the table is based on simulation only and is used only as a reference. The lower part is the main result of the analysis. The SF (DF) contributions are for events with one (two) fake leptons. The  $e\mu$  MC Pred contribution includes contributions from genuine same-sign lepton pairs (a sum of the rows from  $W\gamma$  down to the Higgs samples). The statistical uncertainties on the MC contribution are calculated using the Clopper-Pearson method. Systematic uncertainties (the second uncertainty if present) are displayed only for the final combined type of background, no systematic uncertainty is added for estimates with zero entries. Systematic uncertainties are 100% correlated among the channels.

source	$ee$	$\mu\mu$	$e\mu$	$\ell\ell$
$t\bar{t} \rightarrow \ell\ell X$	$6.40 \pm 2.90$	$0.00 \pm 1.21$	$5.81 \pm 2.80$	$12.21 \pm 3.71$
$t\bar{t} \rightarrow \ell(b \rightarrow \ell)X$	$13.86 \pm 3.98$	$10.25 \pm 3.56$	$22.25 \pm 4.78$	$46.36 \pm 6.65$
$t\bar{t} \rightarrow \ell(\not{b} \rightarrow \ell)X$	$3.47 \pm 2.35$	$0.51 \pm 1.51$	$6.95 \pm 2.99$	$10.93 \pm 3.56$
$t\bar{t}$ other	$0.00 \pm 1.21$	$0.00 \pm 1.21$	$0.00 \pm 1.21$	$0.00 \pm 1.21$
t, s-channel	$0.00 \pm 0.52$	$0.00 \pm 0.52$	$0.00 \pm 0.52$	$0.00 \pm 0.52$
t, t-channel	$0.27 \pm 0.68$	$0.75 \pm 0.86$	$1.55 \pm 1.05$	$2.56 \pm 1.25$
tW	$0.76 \pm 1.14$	$0.71 \pm 1.14$	$2.26 \pm 1.55$	$3.72 \pm 1.85$
$DY \rightarrow \ell\ell$	$18.03 \pm 9.25$	$0.00 \pm 4.14$	$5.81 \pm 6.57$	$23.85 \pm 10.26$
$W + jets \rightarrow \ell\nu$	$0.00 \pm 73.20$	$0.00 \pm 73.20$	$0.00 \pm 73.20$	$0.00 \pm 73.20$
WW	$0.00 \pm 0.11$	$0.00 \pm 0.11$	$0.00 \pm 0.11$	$0.00 \pm 0.11$
$W\gamma^* \rightarrow \ell\nu\mu\mu$	$0.00 \pm 0.23$	$0.20 \pm 0.33$	$0.22 \pm 0.33$	$0.42 \pm 0.39$
$W\gamma^* \rightarrow \ell\nu\tau\tau$	$0.11 \pm 0.30$	$0.22 \pm 0.34$	$0.00 \pm 0.24$	$0.33 \pm 0.38$
WZ	$18.53 \pm 0.47$	$17.10 \pm 0.47$	$36.72 \pm 0.66$	$72.35 \pm 0.93$
ZZ	$1.19 \pm 0.03$	$0.90 \pm 0.03$	$2.22 \pm 0.04$	$4.31 \pm 0.06$
$t\bar{t}\gamma$	$0.53 \pm 1.35$	$0.00 \pm 1.08$	$3.07 \pm 2.10$	$3.60 \pm 2.21$
$t\bar{t}W$	$10.61 \pm 0.55$	$14.72 \pm 0.66$	$26.55 \pm 0.85$	$51.88 \pm 1.19$
$t\bar{t}Z$	$2.80 \pm 0.26$	$3.16 \pm 0.29$	$6.25 \pm 0.39$	$12.20 \pm 0.54$
$t\bar{t}Z(Z \rightarrow \ell\ell)$	$0.44 \pm 0.03$	$0.34 \pm 0.03$	$0.82 \pm 0.04$	$1.59 \pm 0.05$
$t\bar{t}WW$	$0.22 \pm 0.01$	$0.29 \pm 0.01$	$0.51 \pm 0.01$	$1.01 \pm 0.01$
$WW\gamma$	$0.00 \pm 0.09$	$0.00 \pm 0.09$	$0.00 \pm 0.09$	$0.00 \pm 0.09$
WWW	$1.84 \pm 0.13$	$2.44 \pm 0.15$	$4.17 \pm 0.19$	$8.45 \pm 0.27$
WWZ	$0.40 \pm 0.05$	$0.32 \pm 0.05$	$0.80 \pm 0.07$	$1.52 \pm 0.10$
WZZ	$0.08 \pm 0.01$	$0.08 \pm 0.01$	$0.19 \pm 0.02$	$0.36 \pm 0.03$
ZZZ	$0.01 \pm 0.00$	$0.00 \pm 0.00$	$0.01 \pm 0.00$	$0.02 \pm 0.00$
$qqW^\pm W^\pm$	$8.07 \pm 0.66$	$11.04 \pm 0.79$	$20.11 \pm 1.01$	$39.22 \pm 1.40$
WW(DPS)	$0.10 \pm 0.06$	$0.19 \pm 0.07$	$0.31 \pm 0.08$	$0.60 \pm 0.11$
WH, ZH, $t\bar{t}H$ ; $H \rightarrow WW$	$2.72 \pm 0.31$	$2.97 \pm 0.33$	$5.94 \pm 0.44$	$11.63 \pm 0.61$
WH, ZH, $t\bar{t}H$ ; $H \rightarrow ZZ$	$0.12 \pm 0.01$	$0.15 \pm 0.02$	$0.29 \pm 0.02$	$0.55 \pm 0.03$
WH, ZH, $t\bar{t}H$ ; $H \rightarrow \tau\tau$	$0.40 \pm 0.04$	$0.44 \pm 0.05$	$0.74 \pm 0.06$	$1.58 \pm 0.08$
Total MC	$90.94 \pm 74.03$	$66.77 \pm 73.48$	$153.54 \pm 73.85$	$311.24 \pm 74.51$
SF	$90.86 \pm 8.69$	$92.64 \pm 5.45$	$200.91 \pm 11.00$	$384.41 \pm 15.04$
DF	$5.89 \pm 0.48$	$4.09 \pm 0.29$	$9.72 \pm 0.55$	$19.69 \pm 0.79$
SC	$2.30 \pm 0.53$	$2.24 \pm 0.50$	$4.59 \pm 1.04$	$9.14 \pm 2.00$
SF + DF	$96.75 \pm 8.65$	$96.73 \pm 5.43$	$210.62 \pm 10.96$	$404.10 \pm 14.98$
SF + DF - SC	$94.44 \pm 8.66 \pm 47.22$	$94.49 \pm 5.45 \pm 47.24$	$206.03 \pm 11.01 \pm 103.02$	$394.96 \pm 15.11 \pm 197.48$
Charge Flips	$23.44 \pm 1.16 \pm 7.03$	$0.00 \pm 0.00 \pm 0.00$	$5.22 \pm 0.45 \pm 1.57$	$28.66 \pm 1.24 \pm 8.60$
MC Pred	$48.15 \pm 1.77 \pm 24.08$	$54.56 \pm 1.71 \pm 27.28$	$108.91 \pm 2.68 \pm 54.46$	$211.62 \pm 3.20 \pm 105.81$
Total Pred	$166.04 \pm 8.92 \pm 53.47$	$149.04 \pm 5.71 \pm 54.55$	$320.17 \pm 11.34 \pm 116.53$	$635.24 \pm 15.50 \pm 224.21$
Data	146	111	220	477

**Table 7.3: Baseline with 1 btag (Signal Region 10):** Observed event yields in the high  $p_T$  baseline (lepton  $p_T > 20/20$  GeV,  $H_T > 80$  GeV,  $\cancel{E}_T > 30$  GeV if  $H_T < 500$  GeV (otherwise no  $\cancel{E}_T$  cut), at least 2 jets with  $p_T > 40$  GeV and exactly one b-tagged jet using CSV) compared to expectations from simulation alone, and from the data-driven methods. The upper part of the table is based on simulation only and is used only as a reference. The lower part is the main result of the analysis. The SF (DF) contributions are for events with one (two) fake leptons. The  $e\mu$  MC Pred contribution includes contributions from genuine same-sign lepton pairs (a sum of the rows from  $W\gamma$  down to the Higgs samples). The statistical uncertainties on the MC contribution are calculated using the Clopper-Pearson method. Systematic uncertainties (the second uncertainty if present) are displayed only for the final combined type of background, no systematic uncertainty is added for estimates with zero entries. Systematic uncertainties are 100% correlated among the channels.

source	$ee$	$\mu\mu$	$e\mu$	$\ell\ell$
$t\bar{t} \rightarrow \ell\ell X$	$3.51 \pm 2.35$	$0.00 \pm 1.21$	$3.47 \pm 2.35$	$6.99 \pm 2.99$
$t\bar{t} \rightarrow \ell(b \rightarrow \ell)X$	$6.13 \pm 2.90$	$5.39 \pm 2.80$	$11.45 \pm 3.64$	$22.96 \pm 4.89$
$t\bar{t} \rightarrow \ell(\not{b} \rightarrow \ell)X$	$1.70 \pm 1.91$	$0.51 \pm 1.51$	$4.06 \pm 2.47$	$6.27 \pm 2.90$
$t\bar{t}$ other	$0.00 \pm 1.21$	$0.00 \pm 1.21$	$0.00 \pm 1.21$	$0.00 \pm 1.21$
t, s-channel	$0.00 \pm 0.52$	$0.00 \pm 0.52$	$0.00 \pm 0.52$	$0.00 \pm 0.52$
t, t-channel	$0.27 \pm 0.68$	$0.26 \pm 0.68$	$0.53 \pm 0.77$	$1.06 \pm 0.93$
tW	$0.37 \pm 1.00$	$0.00 \pm 0.80$	$1.87 \pm 1.46$	$2.24 \pm 1.55$
$DY \rightarrow \ell\ell$	$8.11 \pm 7.12$	$0.00 \pm 4.14$	$1.97 \pm 5.18$	$10.08 \pm 7.61$
$W + jets \rightarrow \ell\nu$	$0.00 \pm 73.20$	$0.00 \pm 73.20$	$0.00 \pm 73.20$	$0.00 \pm 73.20$
WW	$0.00 \pm 0.11$	$0.00 \pm 0.11$	$0.00 \pm 0.11$	$0.00 \pm 0.11$
$W\gamma^* \rightarrow \ell\nu\mu\mu$	$0.00 \pm 0.23$	$0.00 \pm 0.23$	$0.00 \pm 0.23$	$0.00 \pm 0.23$
$W\gamma^* \rightarrow \ell\nu\tau\tau$	$0.00 \pm 0.24$	$0.00 \pm 0.24$	$0.00 \pm 0.24$	$0.00 \pm 0.24$
WZ	$1.41 \pm 0.14$	$1.21 \pm 0.13$	$2.63 \pm 0.19$	$5.26 \pm 0.26$
ZZ	$0.09 \pm 0.01$	$0.08 \pm 0.01$	$0.18 \pm 0.01$	$0.35 \pm 0.02$
$t\bar{t}\gamma$	$0.26 \pm 0.67$	$0.00 \pm 1.08$	$1.52 \pm 1.04$	$1.78 \pm 1.10$
$t\bar{t}W$	$5.63 \pm 0.41$	$7.36 \pm 0.48$	$13.35 \pm 0.61$	$26.33 \pm 0.86$
$t\bar{t}Z$	$1.53 \pm 0.20$	$1.69 \pm 0.22$	$2.96 \pm 0.27$	$6.17 \pm 0.39$
$t\bar{t}B(Z \rightarrow \ell\ell)$	$0.22 \pm 0.02$	$0.17 \pm 0.02$	$0.43 \pm 0.03$	$0.82 \pm 0.04$
$t\bar{t}WW$	$0.10 \pm 0.00$	$0.15 \pm 0.01$	$0.25 \pm 0.01$	$0.49 \pm 0.01$
$WW\gamma$	$0.00 \pm 0.09$	$0.00 \pm 0.09$	$0.00 \pm 0.09$	$0.00 \pm 0.09$
WWW	$0.15 \pm 0.04$	$0.22 \pm 0.05$	$0.37 \pm 0.06$	$0.74 \pm 0.09$
WWZ	$0.05 \pm 0.02$	$0.08 \pm 0.03$	$0.12 \pm 0.03$	$0.24 \pm 0.04$
WZZ	$0.01 \pm 0.01$	$0.01 \pm 0.01$	$0.03 \pm 0.01$	$0.05 \pm 0.01$
ZZZ	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$qqW^\pm W^\pm$	$0.44 \pm 0.20$	$1.04 \pm 0.28$	$1.48 \pm 0.31$	$2.96 \pm 0.42$
WW(DPS)	$0.01 \pm 0.03$	$0.00 \pm 0.03$	$0.04 \pm 0.04$	$0.05 \pm 0.04$
WH, ZH, $t\bar{t}H$ ; $H \rightarrow WW$	$1.34 \pm 0.22$	$1.12 \pm 0.21$	$2.89 \pm 0.31$	$5.35 \pm 0.42$
WH, ZH, $t\bar{t}H$ ; $H \rightarrow ZZ$	$0.05 \pm 0.01$	$0.05 \pm 0.01$	$0.09 \pm 0.01$	$0.18 \pm 0.02$
WH, ZH, $t\bar{t}H$ ; $H \rightarrow \tau\tau$	$0.11 \pm 0.02$	$0.13 \pm 0.03$	$0.24 \pm 0.03$	$0.48 \pm 0.05$
Total MC	$31.49 \pm 73.70$	$19.46 \pm 73.43$	$49.92 \pm 73.60$	$100.86 \pm 73.93$
SF	$41.00 \pm 4.36$	$45.33 \pm 3.19$	$91.84 \pm 5.69$	$178.17 \pm 7.84$
DF	$1.53 \pm 0.23$	$1.57 \pm 0.17$	$2.97 \pm 0.28$	$6.07 \pm 0.40$
SC	$0.65 \pm 0.19$	$0.56 \pm 0.16$	$1.14 \pm 0.29$	$2.35 \pm 0.59$
SF + DF	$42.53 \pm 4.34$	$46.90 \pm 3.17$	$94.81 \pm 5.67$	$184.24 \pm 7.81$
SF + DF - SC	$41.88 \pm 4.35 \pm 20.94$	$46.34 \pm 3.18 \pm 23.17$	$93.66 \pm 5.67 \pm 46.83$	$181.89 \pm 7.83 \pm 90.94$
Charge Flips	$4.00 \pm 0.22 \pm 1.20$	$0.00 \pm 0.00 \pm 0.00$	$2.52 \pm 0.22 \pm 0.76$	$6.53 \pm 0.31 \pm 1.96$
MC Pred	$11.40 \pm 0.94 \pm 5.70$	$13.30 \pm 1.31 \pm 6.65$	$26.57 \pm 1.38 \pm 13.29$	$51.27 \pm 1.63 \pm 25.63$
Total Pred	$57.28 \pm 4.45 \pm 21.74$	$59.64 \pm 3.43 \pm 24.11$	$122.76 \pm 5.84 \pm 48.69$	$239.68 \pm 8.01 \pm 94.51$
Data	42	35	75	152

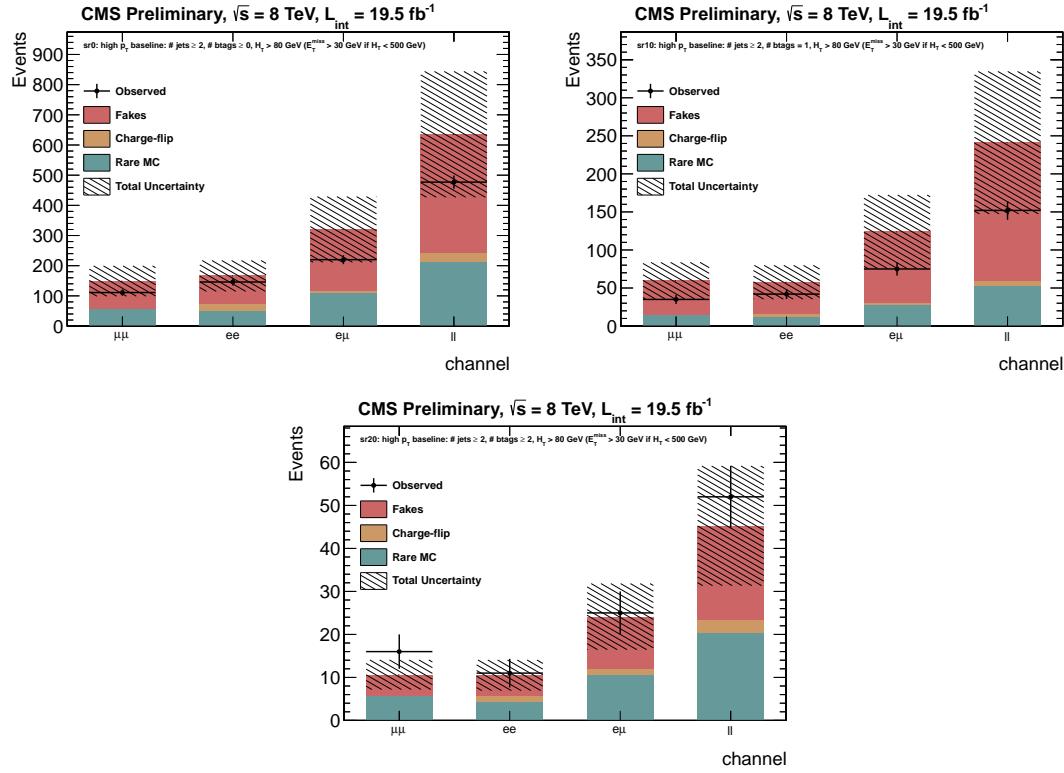
**Table 7.4: Baseline with 2 btag (Signal Region 20):** Observed event yields in the high  $p_T$  baseline (lepton  $p_T > 20/20$  GeV,  $H_T > 80$  GeV,  $\cancel{E}_T > 30$  GeV if  $H_T < 500$  GeV (otherwise no  $\cancel{E}_T$  cut), at least 2 jets with  $p_T > 40$  GeV and at least two b-tagged jets using CSVM) compared to expectations from simulation alone, and from the data-driven methods. The upper part of the table is based on simulation only and is used only as a reference. The lower part is the main result of the analysis. The SF (DF) contributions are for events with one (two) fake leptons. The  $e\mu$  MC Pred contribution includes contributions from genuine same-sign lepton pairs (a sum of the rows from  $W\gamma$  down to the Higgs samples). The statistical uncertainties on the MC contribution are calculated using the Clopper-Pearson method. Systematic uncertainties (the second uncertainty if present) are displayed only for the final combined type of background, no systematic uncertainty is added for estimates with zero entries. Systematic uncertainties are 100% correlated among the channels.

source	$ee$	$\mu\mu$	$e\mu$	$\ell\ell$
$t\bar{t} \rightarrow \ell\ell X$	$1.16 \pm 1.73$	$0.00 \pm 1.21$	$1.74 \pm 1.91$	$2.89 \pm 2.22$
$t\bar{t} \rightarrow \ell(b \rightarrow \ell)X$	$0.00 \pm 1.21$	$0.56 \pm 1.51$	$2.26 \pm 2.07$	$2.81 \pm 2.22$
$t\bar{t} \rightarrow \ell(\not{b} \rightarrow \ell)X$	$1.16 \pm 1.73$	$0.00 \pm 1.21$	$0.00 \pm 1.21$	$1.16 \pm 1.73$
$t\bar{t}$ other	$0.00 \pm 1.21$	$0.00 \pm 1.21$	$0.00 \pm 1.21$	$0.00 \pm 1.21$
t, s-channel	$0.00 \pm 0.52$	$0.00 \pm 0.52$	$0.00 \pm 0.52$	$0.00 \pm 0.52$
t, t-channel	$0.00 \pm 0.54$	$0.00 \pm 0.54$	$0.77 \pm 0.86$	$0.77 \pm 0.86$
tW	$0.00 \pm 0.80$	$0.00 \pm 0.80$	$0.00 \pm 0.80$	$0.00 \pm 0.80$
$DY \rightarrow \ell\ell$	$0.00 \pm 4.14$	$0.00 \pm 4.14$	$0.00 \pm 4.14$	$0.00 \pm 4.14$
$W + jets \rightarrow \ell\nu$	$0.00 \pm 73.20$	$0.00 \pm 73.20$	$0.00 \pm 73.20$	$0.00 \pm 73.20$
WW	$0.00 \pm 0.11$	$0.00 \pm 0.11$	$0.00 \pm 0.11$	$0.00 \pm 0.11$
$W\gamma^* \rightarrow \ell\nu\mu\mu$	$0.00 \pm 0.23$	$0.00 \pm 0.23$	$0.00 \pm 0.23$	$0.00 \pm 0.23$
$W\gamma^* \rightarrow \ell\nu\tau\tau$	$0.00 \pm 0.24$	$0.00 \pm 0.24$	$0.00 \pm 0.24$	$0.00 \pm 0.24$
WZ	$0.04 \pm 0.03$	$0.08 \pm 0.04$	$0.15 \pm 0.05$	$0.26 \pm 0.07$
ZZ	$0.01 \pm 0.00$	$0.01 \pm 0.00$	$0.01 \pm 0.00$	$0.02 \pm 0.01$
$t\bar{t}\gamma$	$0.15 \pm 0.37$	$0.00 \pm 1.08$	$0.85 \pm 0.58$	$0.99 \pm 0.61$
$t\bar{t}W$	$2.73 \pm 0.29$	$3.80 \pm 0.35$	$7.20 \pm 0.46$	$13.73 \pm 0.63$
$t\bar{t}Z$	$0.70 \pm 0.14$	$0.69 \pm 0.15$	$1.78 \pm 0.22$	$3.17 \pm 0.28$
$t\bar{t}Z(Z \rightarrow \ell\ell)$	$0.05 \pm 0.01$	$0.05 \pm 0.01$	$0.10 \pm 0.01$	$0.19 \pm 0.02$
$t\bar{t}WW$	$0.06 \pm 0.00$	$0.07 \pm 0.00$	$0.14 \pm 0.01$	$0.27 \pm 0.01$
$WW\gamma$	$0.00 \pm 0.09$	$0.00 \pm 0.09$	$0.00 \pm 0.09$	$0.00 \pm 0.09$
WWW	$0.01 \pm 0.02$	$0.02 \pm 0.02$	$0.04 \pm 0.03$	$0.07 \pm 0.03$
WWZ	$0.00 \pm 0.01$	$0.00 \pm 0.01$	$0.01 \pm 0.01$	$0.02 \pm 0.02$
WZZ	$0.01 \pm 0.00$	$0.00 \pm 0.00$	$0.01 \pm 0.00$	$0.01 \pm 0.01$
ZZZ	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$qqW^\pm W^\pm$	$0.00 \pm 0.09$	$0.03 \pm 0.05$	$0.07 \pm 0.12$	$0.10 \pm 0.12$
WW(DPS)	$0.00 \pm 0.03$	$0.00 \pm 0.03$	$0.00 \pm 0.03$	$0.00 \pm 0.03$
WH, ZH, $t\bar{t}H; H \rightarrow WW$	$0.55 \pm 0.15$	$0.70 \pm 0.17$	$0.95 \pm 0.19$	$2.20 \pm 0.28$
WH, ZH, $t\bar{t}H; H \rightarrow ZZ$	$0.02 \pm 0.01$	$0.03 \pm 0.01$	$0.06 \pm 0.01$	$0.12 \pm 0.01$
WH, ZH, $t\bar{t}H; H \rightarrow \tau\tau$	$0.05 \pm 0.02$	$0.06 \pm 0.02$	$0.10 \pm 0.02$	$0.21 \pm 0.03$
Total MC	$6.69 \pm 73.39$	$6.09 \pm 73.38$	$16.23 \pm 73.41$	$29.02 \pm 73.44$
SF	$5.05 \pm 1.05$	$5.05 \pm 0.87$	$12.40 \pm 1.54$	$22.50 \pm 2.06$
DF	$0.20 \pm 0.08$	$0.28 \pm 0.07$	$0.25 \pm 0.07$	$0.73 \pm 0.13$
SC	$0.30 \pm 0.10$	$0.24 \pm 0.08$	$0.50 \pm 0.17$	$1.05 \pm 0.32$
SF + DF	$5.25 \pm 1.04$	$5.33 \pm 0.86$	$12.65 \pm 1.53$	$23.23 \pm 2.04$
SF + DF - SC	$4.95 \pm 1.04 \pm 2.48$	$5.09 \pm 0.87 \pm 2.54$	$12.15 \pm 1.54 \pm 6.07$	$22.18 \pm 2.07 \pm 11.09$
Charge Flips	$1.41 \pm 0.09 \pm 0.42$	$0.00 \pm 0.00 \pm 0.00$	$1.42 \pm 0.12 \pm 0.43$	$2.83 \pm 0.15 \pm 0.85$
MC Pred	$4.37 \pm 0.63 \pm 2.19$	$5.54 \pm 1.21 \pm 2.77$	$11.47 \pm 0.88 \pm 5.73$	$21.38 \pm 1.03 \pm 10.69$
Total Pred	$10.74 \pm 1.22 \pm 3.33$	$10.62 \pm 1.49 \pm 3.76$	$25.03 \pm 1.78 \pm 8.36$	$46.39 \pm 2.32 \pm 15.43$
Data	11	16	25	52

Consider first the inclusive baseline SR0. Simulation alone indicates an expected background contribution of 20% fake leptons, 10% leptons with incorrect charge assignment and 70% from SM source is prompt, isolated same-sign dileptons. This is in contrast to previous iterations of this analysis which saw the contribution from fake leptons dominating the background composition. This is due to the tighter impact parameter requirement on the lepton selection and increased yield of the rare process with the increase in both luminosity and center-of-mass energy of the collisions. The background from fake leptons is mostly from top events. It is difficult to estimate the contribution from  $W + \text{jets}$  due to the limited statistics in this sample with a large uncertainty of  $\pm 72$  events; but this could be a significant component of the fake lepton background. The background from electrons with a mis-reconstructed charge comes from fully leptonic  $t\bar{t}$  decays and Drell-Yan events. The remaining irreducible background comes from several different process that involve genuine same-sign dileptons, although most predominately are  $t\bar{t}W$ ,  $WZ$ , and  $qqW^\pm W^\pm$ .

The data-driven background estimation over-predicts by a factor of almost seven higher than that expected from pure simulation ( $\approx 63$  vs.  $395$ ). This is attributed to several factors. First, the fake rate method itself over-predicted by 60% in the  $t\bar{t}$  closure test shown in Section 5.1.6. This accounts for the 50% systematic uncertainty driven by this method. Secondly, the simulation also has a large uncertainty since the  $W + \text{jets}$  sample has low statistics. Finally, the simulation itself may be failing to accurately account for all the sources of fake leptons. When one accounts for the large uncertainties on both the simulated and data-driven fake lepton prediction, the methods are consistent and thus no correction is made at this time.

In baselines SR10 and SR20, a b-tagged jet requirement is made. This reduces the contributions from rare SM processes such as  $WZ$  and  $qqW^\pm W^\pm$  where there are no intrinsic b-quarks in the final state. In these regions, the contributions from fakes leptons becomes reduced due to the increased b-tagged jet requirement. Figure 7.2 shows a graphical representation of the yields and there full background predictions for SR0, SR10 and SR20. Here we see the background prediction and the data yield are in agreement within the uncertainty. This lends confidence that our data-driven methods are performing reasonably well.



**Figure 7.2:** Graphical representation of the event yields in the high  $p_T$  baseline region with no (top left), exactly one (top right) and  $\geq 2$  (bottom column) #  $b$ -tagged jets requirement (Search Region 0, 10 and 20). Also shown as a histogram is the result of the background prediction. The shading around the histogram represents the uncertainty in the background prediction.

### 7.1.2 High $p_T$ Search Regions

We now consider the search regions shown in Table 4.9 and repeated in Table 7.5 for convenience. The observed yields and estimated backgrounds for each region are shown in Table 7.6. The uncertainty is the total statistical and systematic uncertainty. A graphical representation of these results can also be seen in Figure 7.3. We see no evidence of an event yield in significant excess of the background estimations.

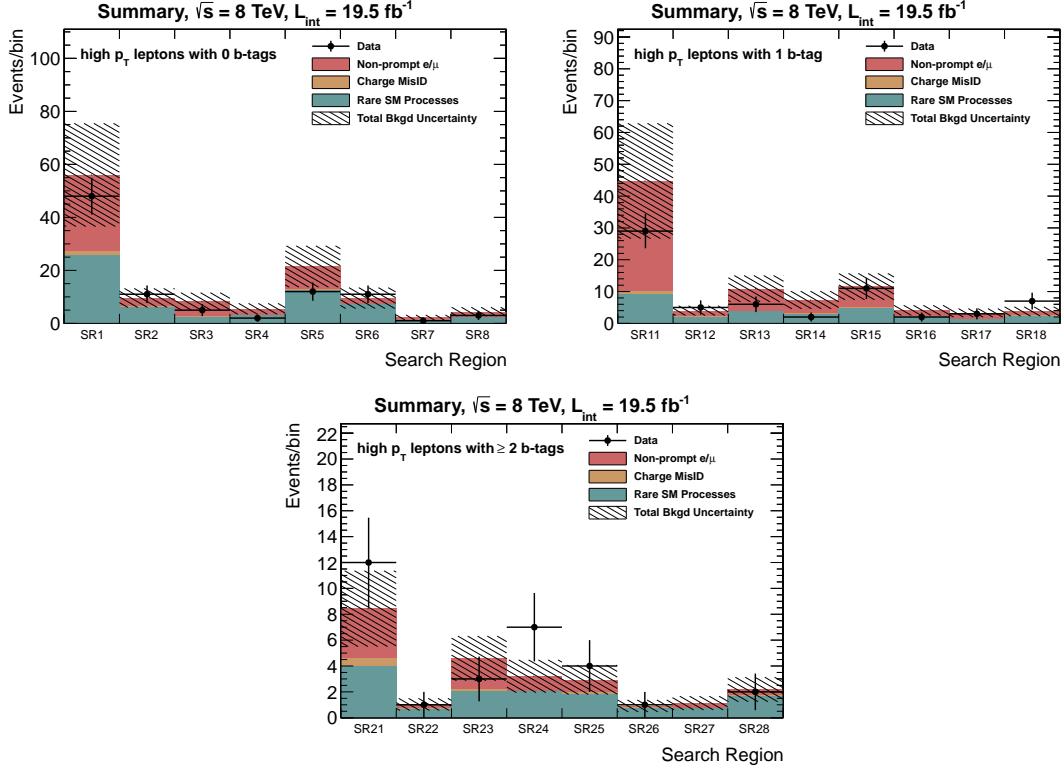
**Table 7.5:** Search regions selected for the high  $p_T$  search where we require  $H_T > 200$  GeV.

# $b$ -tagged jets	$\cancel{E}_T$	# jets	$H_T$ [200-400]	$H_T$ [ $> 400$ ]
$= 0$	50-120	2-3	SR1	SR2
		$\geq 4$	SR3	SR4
	$> 120$	2-3	SR5	SR6
		$\geq 4$	SR7	SR8
$= 1$	50-120	2-3	SR11	SR12
		$\geq 4$	SR13	SR14
	$> 120$	2-3	SR15	SR16
		$\geq 4$	SR17	SR18
$\geq 2$	50-120	2-3	SR21	SR22
		$\geq 4$	SR23	SR24
	$> 120$	2-3	SR25	SR26
		$\geq 4$	SR27	SR28

In the various search regions, irreducible backgrounds, primarily from  $t\bar{t}W$ ,  $WZ$  and  $qqW^\pm W^\pm$ , are the dominant backgrounds at approximately 25-80%, depending on the search region. Fake leptons account for approximately 15-75% of the background also depending on the search region. The background from charge mis-reconstruction is small contributing less than 5%.

**Table 7.6:** A summary of the results of this search for the high  $p_T$  analysis. For each signal region, we show its most important kinematical requirements, the prediction for the three background (BG) components as well as the total, and the event yield.

# $b$ -tagged jets	$\cancel{E}_T$	# jets	$H_T$	SR	Fake BG	Flip BG	Rare MC	Total BG	Observed	
$= 0$	$\geq 0$	30 if $H_T < 500$ else 0	2	80	0	$394.96 \pm 198.06$	$28.66 \pm 8.69$	$211.62 \pm 105.86$	$635.24 \pm 224.74$	477
	50-120	2-3	200-400	1	$28.62 \pm 14.52$	$1.56 \pm 0.48$	$25.84 \pm 12.96$	$56.02 \pm 19.47$	48	
		$> 400$	2	$2.94 \pm 1.63$	$0.37 \pm 0.12$	$6.30 \pm 3.20$	$9.61 \pm 3.60$	11		
		$\geq 4$	200-400	3	$5.64 \pm 2.98$	$0.13 \pm 0.04$	$2.54 \pm 1.36$	$8.32 \pm 3.28$	5	
		$> 400$	4	$2.33 \pm 1.33$	$0.15 \pm 0.05$	$2.99 \pm 1.58$	$5.47 \pm 2.07$	2		
	$> 120$	2-3	200-400	5	$8.53 \pm 4.45$	$0.19 \pm 0.06$	$12.74 \pm 6.41$	$21.47 \pm 7.81$	12	
		$> 400$	6	$2.11 \pm 1.23$	$0.06 \pm 0.02$	$7.38 \pm 3.75$	$9.55 \pm 3.94$	11		
		$\geq 4$	200-400	7	$1.32 \pm 0.81$	$0.02 \pm 0.01$	$0.87 \pm 0.60$	$2.21 \pm 1.01$	1	
		$> 400$	8	$1.97 \pm 1.13$	$0.02 \pm 0.01$	$2.41 \pm 1.29$	$4.40 \pm 1.72$	3		
$= 1$	$\geq 0$	30 if $H_T < 500$ else 0	2	80	10	$181.89 \pm 91.28$	$6.53 \pm 1.98$	$51.27 \pm 25.69$	$239.68 \pm 94.85$	152
	50-120	2-3	200-400	11	$34.63 \pm 17.51$	$1.00 \pm 0.31$	$9.13 \pm 4.61$	$44.76 \pm 18.10$	29	
		$> 400$	12	$1.88 \pm 1.10$	$0.12 \pm 0.04$	$2.01 \pm 1.10$	$4.01 \pm 1.56$	5		
		$\geq 4$	200-400	13	$7.00 \pm 3.67$	$0.16 \pm 0.05$	$3.73 \pm 1.93$	$10.89 \pm 4.15$	6	
		$> 400$	14	$4.27 \pm 2.31$	$0.09 \pm 0.03$	$3.00 \pm 1.57$	$7.35 \pm 2.79$	2		
	$> 120$	2-3	200-400	15	$6.31 \pm 3.31$	$0.33 \pm 0.10$	$4.91 \pm 2.51$	$11.55 \pm 4.16$	11	
		$> 400$	16	$1.70 \pm 1.02$	$0.07 \pm 0.02$	$2.33 \pm 1.25$	$4.10 \pm 1.62$	2		
		$\geq 4$	200-400	17	$2.03 \pm 1.16$	$0.04 \pm 0.01$	$1.24 \pm 0.75$	$3.31 \pm 1.38$	3	
		$> 400$	18	$1.27 \pm 0.78$	$0.05 \pm 0.02$	$2.45 \pm 1.30$	$3.76 \pm 1.52$	7		
$\geq 2$	$\geq 0$	30 if $H_T < 500$ else 0	2	80	20	$22.18 \pm 11.28$	$2.83 \pm 0.86$	$21.38 \pm 10.74$	$46.39 \pm 15.60$	52
	50-120	2-3	200-400	21	$3.87 \pm 2.10$	$0.56 \pm 0.17$	$4.00 \pm 2.06$	$8.43 \pm 2.95$	12	
		$> 400$	22	$0.28 \pm 0.26$	$0.06 \pm 0.02$	$0.69 \pm 0.52$	$1.03 \pm 0.59$	1		
		$\geq 4$	200-400	23	$2.40 \pm 1.35$	$0.11 \pm 0.04$	$2.10 \pm 1.14$	$4.61 \pm 1.77$	3	
		$> 400$	24	$1.13 \pm 0.72$	$0.06 \pm 0.02$	$2.03 \pm 1.11$	$3.23 \pm 1.32$	7		
	$> 120$	2-3	200-400	25	$0.99 \pm 0.63$	$0.14 \pm 0.04$	$1.81 \pm 1.00$	$2.94 \pm 1.19$	4	
		$> 400$	26	$0.13 \pm 0.19$	$0.03 \pm 0.01$	$0.76 \pm 0.54$	$0.91 \pm 0.58$	1		
		$\geq 4$	200-400	27	$0.38 \pm 0.34$	$0.03 \pm 0.01$	$0.73 \pm 0.54$	$1.14 \pm 0.64$	0	
		$> 400$	28	$0.45 \pm 0.34$	$0.03 \pm 0.01$	$1.67 \pm 0.94$	$2.16 \pm 1.00$	2		



**Figure 7.3:** Graphical representation of the event yields in the high  $p_T$  search regions with one (top left), exactly one (top right) and  $\geq 2$  (bottom column) #  $b$ -tagged jets requirement (Search Region 1-8, 11-18, and 21-28). Also shown as a histogram is the result of the background prediction. The shading around the histogram represents the uncertainty in the background prediction.

## 7.2 Low $p_T$ Results

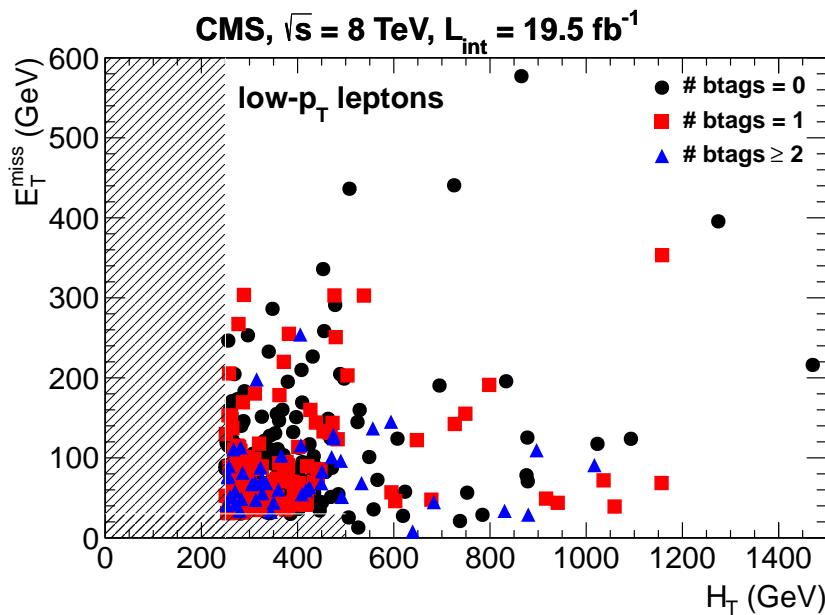
### 7.2.1 Low $p_T$ Control Regions

For the low  $p_T$  analysis, the control region (baseline) selections were listed in Table 4.8 but are repeated here for convenience in Table 7.7. The control regions in the low  $p_T$  analysis are essentially the same as the high  $p_T$  analysis with two notable exceptions: the lepton  $p_T$  threshold is lowered from 20 to 10 GeV on both leptons and the minimum  $H_T$  requirement is raised from 80 to 250 GeV. The lowered lepton  $p_T$  allows more sensitivity for regions with a compressed mass spectrum that would lead to

**Table 7.7:** Summary of the baseline search regions considered in the low  $p_T$  analysis.

Search Region #	min lepton $p_T$ ( $\mu, e$ ) (GeV)	$H_T$ (GeV)	$\cancel{E}_T$ (GeV)	# jets	# $b$ -tagged jets
SR0	10, 10	250	30 if $H_T < 500$ else 0	2	$\geq 0$
SR10					= 1
SR20					$\geq 2$

lower  $p_T$  leptons. The  $H_T$  requirement is raised since the low  $p_T$  online trigger selection has a minimum  $H_T$  requirement of 175 GeV; thus, 250 GeV was chosen to account for the trigger turn-on. As with the high  $p_T$  analysis, the control region SR0 was designed to be inclusive to all the search regions and is mostly sensitive to fake leptons in SM multi-jet processes: predominately  $t\bar{t}$  and  $W + \text{jets}$ . The contribution for backgrounds without intrinsic  $b$ -quarks is suppressed in SR10 and SR20 which add a #  $b$ -tagged jets requirement of exactly one and at least two, respectively. Figure 7.4 shows the distribution of events from SR0 in the  $H_T$ - $\cancel{E}_T$  plane. The events cluster in lower left corner as expected for background events.



**Figure 7.4:** Distribution of events in the  $H_T$ - $\cancel{E}_T$  plane passing the analysis selections for the low  $p_T$  baseline region (SR0).

The observed event yields and estimated backgrounds in the low  $p_T$  control regions can be seen in Tables 7.8, 7.9, and 7.10. The yield reported as *MC Pred* includes contributions from genuine same-sign lepton pairs (sum of the rows from the  $W\gamma$  down to the Higgs samples). The statistical uncertainty on the simulated contributions is calculated using the Clopper-Pearson method [79]. Systematic uncertainties (the second uncertainty if present) are displayed only for the final combined background component and no uncertainty is added for estimates with zero entries. Systematic uncertainties are 100% correlated among the channels. The upper portion of the table (from the top down to ZZ) is based on simulation only and is used as a reference only. The total background when taken from simulation only is reported as *Total MC* and again is given as a reference only. The SF(DF) contributions are for events with one(two) fake leptons and the SC refers to the contamination to the fake prediction from rare SM events.

The data-driven background estimation over-predicts by a factor of almost four higher than that expected from pure simulation ( $\approx 73$  vs. 240). This is attributed to several factors. First, the fake rate method itself over-predicted by 60% in the  $t\bar{t}$  closure test shown in Section 5.1.6. This accounts for the 50% systematic uncertainty driven by this method. Secondly, the simulation also has a large uncertainty since the  $W + \text{jets}$  sample has low statistics. Finally, the simulation itself may be failing to accurately account for all the sources of fake leptons. When one accounts for the large uncertainties on both the simulated and data-driven fake lepton prediction, the methods are consistent and thus no correction is made at this time.

Consider first the inclusive baseline SR0. Simulation alone indicates a that background contribution of 20% fake leptons, 10% leptons with incorrect charge assignment and 70% from SM source is prompt, isolated same-sign dileptons. The background from fake leptons is mostly from top events. Its difficult to estimate the contribution from  $W + \text{jets}$  due to the limited statistics in this sample with a large uncertainty of  $\pm 72$  events; but this could be a significant component of the fake lepton background. The background from electrons with a mis-reconstructed charge comes from fully leptonic  $t\bar{t}$  decays and Drell-Yan events. The remaining irreducible background comes from several different process that involve genuine same-sign dileptons, although most predominately are  $t\bar{t}W$ ,  $WZ$ , and  $qqW^\pm W^\pm$ .

**Table 7.8: Baseline with 0 btags (Signal Region 0):** Observed event yields in the low  $p_T$  baseline (lepton  $p_T > 10/10$  GeV,  $H_T > 250$  GeV,  $\cancel{E}_T > 30$  GeV if  $H_T < 500$  GeV (otherwise no  $\cancel{E}_T$  cut), at least 2 jets with  $p_T > 40$  GeV) compared to expectations from simulation alone, and from the data-driven methods. The upper part of the table is based on simulation only and is used only as a reference. The lower part is the main result of the analysis. The SF (DF) contributions are for events with one (two) fake leptons. The  $e\mu$  MC Pred contribution includes contributions from genuine same-sign lepton pairs (a sum of the rows from  $W\gamma$  down to the Higgs samples). The statistical uncertainties on the MC contribution are calculated using the Clopper-Pearson method. Systematic uncertainties (the second uncertainty if present) are displayed only for the final combined type of background, no systematic uncertainty is added for estimates with zero entries. Systematic uncertainties are 100% correlated among the channels.

source	$ee$	$\mu\mu$	$e\mu$	$\ell\ell$
$t\bar{t} \rightarrow \ell\ell X$	$2.21 \pm 2.07$	$0.00 \pm 1.21$	$1.74 \pm 1.91$	$3.95 \pm 2.47$
$t\bar{t} \rightarrow \ell(b \rightarrow \ell)X$	$8.86 \pm 3.33$	$19.77 \pm 4.57$	$31.96 \pm 5.63$	$60.59 \pm 7.49$
$t\bar{t} \rightarrow \ell(\not{b} \rightarrow \ell)X$	$1.05 \pm 1.73$	$1.18 \pm 1.73$	$6.16 \pm 2.90$	$8.38 \pm 3.25$
$t\bar{t}$ other	$0.00 \pm 1.21$	$0.00 \pm 1.21$	$0.00 \pm 1.21$	$0.00 \pm 1.21$
t, s-channel	$0.00 \pm 0.52$	$0.00 \pm 0.52$	$0.00 \pm 0.52$	$0.00 \pm 0.52$
t, t-channel	$0.47 \pm 0.77$	$0.77 \pm 0.86$	$0.53 \pm 0.77$	$1.76 \pm 1.11$
tW	$0.36 \pm 1.00$	$0.76 \pm 1.14$	$1.50 \pm 1.37$	$2.63 \pm 1.63$
$DY \rightarrow \ell\ell$	$5.86 \pm 6.57$	$2.01 \pm 5.18$	$0.00 \pm 4.14$	$7.87 \pm 7.12$
$W + \text{jets} \rightarrow \ell\nu$	$0.00 \pm 73.20$	$0.00 \pm 73.20$	$0.00 \pm 73.20$	$0.00 \pm 73.20$
WW	$0.00 \pm 0.11$	$0.00 \pm 0.11$	$0.00 \pm 0.11$	$0.00 \pm 0.11$
$W\gamma^* \rightarrow \ell\nu\mu\mu$	$0.00 \pm 0.23$	$0.00 \pm 0.23$	$0.00 \pm 0.23$	$0.00 \pm 0.23$
$W\gamma^* \rightarrow \ell\nu\tau\tau$	$0.00 \pm 0.24$	$0.00 \pm 0.24$	$0.00 \pm 0.24$	$0.00 \pm 0.24$
WZ	$6.54 \pm 0.29$	$7.37 \pm 0.30$	$15.12 \pm 0.43$	$29.03 \pm 0.59$
ZZ	$0.40 \pm 0.02$	$0.36 \pm 0.02$	$0.82 \pm 0.03$	$1.58 \pm 0.04$
$t\bar{t}\gamma$	$1.00 \pm 1.55$	$0.00 \pm 1.08$	$1.01 \pm 1.55$	$2.01 \pm 1.86$
$t\bar{t}W$	$6.76 \pm 0.45$	$10.59 \pm 0.55$	$16.61 \pm 0.68$	$33.96 \pm 0.96$
$t\bar{t}Z$	$1.80 \pm 0.22$	$2.68 \pm 0.26$	$4.77 \pm 0.34$	$9.25 \pm 0.47$
$t\bar{t}Z(Z \rightarrow \ell\ell)$	$0.20 \pm 0.02$	$0.18 \pm 0.02$	$0.40 \pm 0.03$	$0.78 \pm 0.04$
$t\bar{t}WW$	$0.18 \pm 0.01$	$0.28 \pm 0.01$	$0.44 \pm 0.01$	$0.91 \pm 0.01$
$WW\gamma$	$0.00 \pm 0.09$	$0.00 \pm 0.09$	$0.00 \pm 0.09$	$0.00 \pm 0.09$
WWW	$1.00 \pm 0.10$	$1.39 \pm 0.11$	$2.21 \pm 0.14$	$4.60 \pm 0.20$
WWZ	$0.23 \pm 0.04$	$0.23 \pm 0.04$	$0.45 \pm 0.06$	$0.91 \pm 0.08$
WZZ	$0.04 \pm 0.01$	$0.05 \pm 0.01$	$0.12 \pm 0.02$	$0.22 \pm 0.02$
ZZZ	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.01 \pm 0.00$
$qqW^\pm W^\pm$	$4.91 \pm 0.54$	$7.74 \pm 0.66$	$13.23 \pm 0.84$	$25.88 \pm 1.15$
WW(DPS)	$0.01 \pm 0.03$	$0.01 \pm 0.03$	$0.02 \pm 0.04$	$0.05 \pm 0.04$
WH, ZH, $t\bar{t}H; H \rightarrow WW$	$1.78 \pm 0.26$	$2.37 \pm 0.29$	$4.10 \pm 0.37$	$8.24 \pm 0.51$
WH, ZH, $t\bar{t}H; H \rightarrow ZZ$	$0.09 \pm 0.01$	$0.12 \pm 0.01$	$0.22 \pm 0.02$	$0.43 \pm 0.03$
WH, ZH, $t\bar{t}H; H \rightarrow \tau\tau$	$0.15 \pm 0.03$	$0.22 \pm 0.03$	$0.29 \pm 0.04$	$0.65 \pm 0.06$
Total MC	$43.89 \pm 73.67$	$58.09 \pm 73.60$	$101.71 \pm 73.68$	$203.69 \pm 74.12$
SF	$50.69 \pm 4.86$	$69.22 \pm 5.05$	$109.64 \pm 6.66$	$229.55 \pm 9.67$
DF	$3.28 \pm 0.28$	$5.59 \pm 0.37$	$8.60 \pm 0.46$	$17.47 \pm 0.65$
SC	$1.06 \pm 0.24$	$1.83 \pm 0.41$	$2.64 \pm 0.56$	$5.54 \pm 1.15$
SF + DF	$53.97 \pm 4.83$	$74.80 \pm 5.01$	$118.25 \pm 6.61$	$247.02 \pm 9.60$
SF + DF - SC	$52.91 \pm 4.84 \pm 26.45$	$72.97 \pm 5.03 \pm 36.49$	$115.60 \pm 6.64 \pm 57.80$	$241.48 \pm 9.67 \pm 120.74$
Charge Flips	$6.75 \pm 0.47 \pm 2.03$	$0.00 \pm 0.00 \pm 0.00$	$1.68 \pm 0.17 \pm 0.50$	$8.43 \pm 0.50 \pm 2.53$
MC Pred	$25.09 \pm 1.79 \pm 12.54$	$33.60 \pm 1.51 \pm 16.80$	$59.82 \pm 2.04 \pm 29.91$	$118.51 \pm 2.59 \pm 59.25$
Total Pred	$84.75 \pm 5.18 \pm 29.35$	$106.57 \pm 5.25 \pm 40.17$	$177.10 \pm 6.94 \pm 65.08$	$368.42 \pm 10.03 \pm 134.52$
Data	75	110	164	349

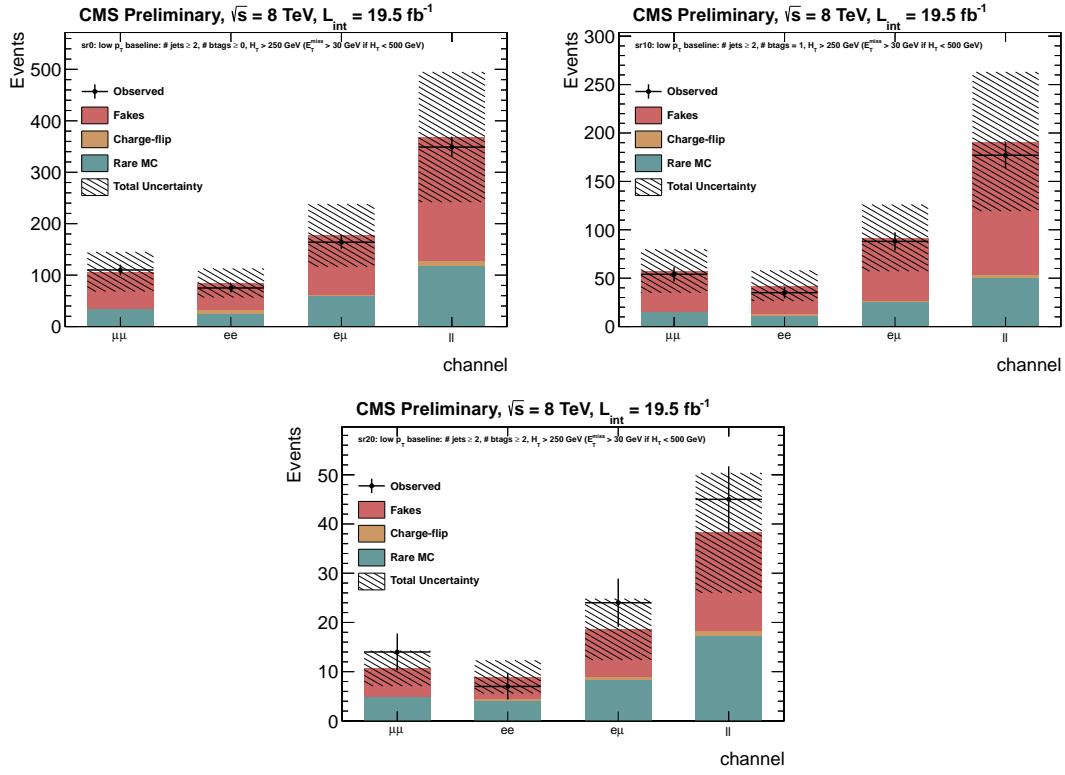
**Table 7.9: Baseline with 1 btag (Signal Region 10):** Observed event yields in the low  $p_T$  baseline (lepton  $p_T > 10/10$  GeV,  $H_T > 250$  GeV,  $\cancel{E}_T > 30$  GeV if  $H_T < 500$  GeV (otherwise no  $\cancel{E}_T$  cut), at least 2 jets with  $p_T > 40$  GeV and exactly one b-tagged jet using CSVM) compared to expectations from simulation alone, and from the data-driven methods. The upper part of the table is based on simulation only and is used only as a reference. The lower part is the main result of the analysis. The SF (DF) contributions are for events with one (two) fake leptons. The  $e\mu$  MC Pred contribution includes contributions from genuine same-sign lepton pairs (a sum of the rows from  $W\gamma$  down to the Higgs samples). The statistical uncertainties on the MC contribution are calculated using the Clopper-Pearson method. Systematic uncertainties (the second uncertainty if present) are displayed only for the final combined type of background, no systematic uncertainty is added for estimates with zero entries. Systematic uncertainties are 100% correlated among the channels.

source	$ee$	$\mu\mu$	$e\mu$	$\ell\ell$
$t\bar{t} \rightarrow \ell\ell X$	$1.11 \pm 1.73$	$0.00 \pm 1.21$	$1.14 \pm 1.73$	$2.25 \pm 2.07$
$t\bar{t} \rightarrow \ell(b \rightarrow \ell)X$	$1.67 \pm 1.91$	$10.11 \pm 3.49$	$16.77 \pm 4.29$	$28.55 \pm 5.37$
$t\bar{t} \rightarrow \ell(\not{b} \rightarrow \ell)X$	$0.00 \pm 1.21$	$0.00 \pm 1.21$	$3.40 \pm 2.35$	$3.40 \pm 2.35$
$t\bar{t}$ other	$0.00 \pm 1.21$	$0.00 \pm 1.21$	$0.00 \pm 1.21$	$0.00 \pm 1.21$
t, s-channel	$0.00 \pm 0.52$	$0.00 \pm 0.52$	$0.00 \pm 0.52$	$0.00 \pm 0.52$
t, t-channel	$0.00 \pm 0.54$	$0.24 \pm 0.68$	$0.26 \pm 0.68$	$0.50 \pm 0.77$
tW	$0.00 \pm 0.80$	$0.37 \pm 1.00$	$1.50 \pm 1.37$	$1.87 \pm 1.46$
$DY \rightarrow \ell\ell$	$5.86 \pm 6.57$	$0.00 \pm 4.14$	$0.00 \pm 4.14$	$5.86 \pm 6.57$
$W + jets \rightarrow \ell\nu$	$0.00 \pm 73.20$	$0.00 \pm 73.20$	$0.00 \pm 73.20$	$0.00 \pm 73.20$
WW	$0.00 \pm 0.11$	$0.00 \pm 0.11$	$0.00 \pm 0.11$	$0.00 \pm 0.11$
$W\gamma^* \rightarrow \ell\nu\mu\mu$	$0.00 \pm 0.23$	$0.00 \pm 0.23$	$0.00 \pm 0.23$	$0.00 \pm 0.23$
$W\gamma^* \rightarrow \ell\nu\tau\tau$	$0.00 \pm 0.24$	$0.00 \pm 0.24$	$0.00 \pm 0.24$	$0.00 \pm 0.24$
WZ	$0.57 \pm 0.09$	$0.63 \pm 0.10$	$1.35 \pm 0.14$	$2.55 \pm 0.18$
ZZ	$0.04 \pm 0.01$	$0.04 \pm 0.01$	$0.09 \pm 0.01$	$0.16 \pm 0.01$
$t\bar{t}\gamma$	$0.47 \pm 0.73$	$0.00 \pm 1.08$	$0.48 \pm 0.73$	$0.95 \pm 0.88$
$t\bar{t}W$	$3.42 \pm 0.33$	$5.37 \pm 0.40$	$8.42 \pm 0.49$	$17.21 \pm 0.69$
$t\bar{t}Z$	$0.97 \pm 0.17$	$1.25 \pm 0.18$	$2.15 \pm 0.24$	$4.38 \pm 0.33$
$t\bar{t}Z(Z \rightarrow \ell\ell)$	$0.11 \pm 0.02$	$0.10 \pm 0.01$	$0.21 \pm 0.02$	$0.41 \pm 0.03$
$t\bar{t}WW$	$0.08 \pm 0.00$	$0.14 \pm 0.01$	$0.21 \pm 0.01$	$0.44 \pm 0.01$
$WW\gamma$	$0.00 \pm 0.09$	$0.00 \pm 0.09$	$0.00 \pm 0.09$	$0.00 \pm 0.09$
WWW	$0.06 \pm 0.03$	$0.15 \pm 0.04$	$0.27 \pm 0.05$	$0.48 \pm 0.07$
WWZ	$0.04 \pm 0.02$	$0.05 \pm 0.02$	$0.07 \pm 0.03$	$0.17 \pm 0.04$
WZZ	$0.01 \pm 0.00$	$0.01 \pm 0.01$	$0.02 \pm 0.01$	$0.03 \pm 0.01$
ZZZ	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$q\bar{q}W^\pm W^\pm$	$0.30 \pm 0.18$	$0.90 \pm 0.27$	$1.09 \pm 0.27$	$2.29 \pm 0.38$
WW(DPS)	$0.00 \pm 0.03$	$0.00 \pm 0.03$	$0.00 \pm 0.03$	$0.00 \pm 0.03$
WH, ZH, $t\bar{t}H$ ; $H \rightarrow WW$	$0.80 \pm 0.18$	$1.06 \pm 0.20$	$2.09 \pm 0.27$	$3.95 \pm 0.36$
WH, ZH, $t\bar{t}H$ ; $H \rightarrow ZZ$	$0.04 \pm 0.01$	$0.04 \pm 0.01$	$0.08 \pm 0.01$	$0.17 \pm 0.02$
WH, ZH, $t\bar{t}H$ ; $H \rightarrow \tau\tau$	$0.06 \pm 0.02$	$0.08 \pm 0.02$	$0.10 \pm 0.02$	$0.25 \pm 0.04$
Total MC	$15.60 \pm 73.58$	$20.54 \pm 73.46$	$39.72 \pm 73.54$	$75.86 \pm 73.80$
SF	$24.31 \pm 2.69$	$35.48 \pm 2.93$	$53.34 \pm 3.87$	$113.13 \pm 5.55$
DF	$0.98 \pm 0.14$	$1.94 \pm 0.21$	$3.03 \pm 0.26$	$5.94 \pm 0.36$
SC	$0.32 \pm 0.09$	$0.54 \pm 0.16$	$0.75 \pm 0.20$	$1.61 \pm 0.41$
SF + DF	$25.29 \pm 2.68$	$37.42 \pm 2.90$	$56.37 \pm 3.84$	$119.07 \pm 5.51$
SF + DF - SC	$24.97 \pm 2.68 \pm 12.48$	$36.87 \pm 2.91 \pm 18.44$	$55.62 \pm 3.85 \pm 27.81$	$117.46 \pm 5.53 \pm 58.73$
Charge Flips	$1.25 \pm 0.09 \pm 0.37$	$0.00 \pm 0.00 \pm 0.00$	$0.75 \pm 0.08 \pm 0.22$	$2.00 \pm 0.12 \pm 0.60$
MC Pred	$6.97 \pm 0.93 \pm 3.49$	$9.82 \pm 1.27 \pm 4.91$	$16.64 \pm 1.06 \pm 8.32$	$33.43 \pm 1.34 \pm 16.71$
Total Pred	$33.19 \pm 2.84 \pm 12.97$	$46.69 \pm 3.17 \pm 19.08$	$73.00 \pm 3.99 \pm 29.03$	$152.88 \pm 5.69 \pm 61.06$
Data	28	40	64	132

**Table 7.10: Baseline with 2 btag (Signal Region 20):** Observed event yields in the low  $p_T$  baseline (lepton  $p_T > 10/10$  GeV,  $H_T > 250$  GeV,  $\cancel{E}_T > 30$  GeV if  $H_T < 500$  GeV (otherwise no  $\cancel{E}_T$  cut), at least 2 jets with  $p_T > 40$  GeV and at least two b-tagged jets using CSV) compared to expectations from simulation alone, and from the data-driven methods. The upper part of the table is based on simulation only and is used only as a reference. The lower part is the main result of the analysis. The SF (DF) contributions are for events with one (two) fake leptons. The  $e\mu$  MC Pred contribution includes contributions from genuine same-sign lepton pairs (a sum of the rows from  $W\gamma$  down to the Higgs samples). The statistical uncertainties on the MC contribution are calculated using the Clopper-Pearson method. Systematic uncertainties (the second uncertainty if present) are displayed only for the final combined type of background, no systematic uncertainty is added for estimates with zero entries. Systematic uncertainties are 100% correlated among the channels.

source	$ee$	$\mu\mu$	$e\mu$	$\ell\ell$
$t\bar{t} \rightarrow \ell\ell X$	$0.56 \pm 1.51$	$0.00 \pm 1.21$	$0.60 \pm 1.51$	$1.16 \pm 1.73$
$t\bar{t} \rightarrow \ell(b \rightarrow \ell)X$	$2.69 \pm 2.22$	$2.22 \pm 2.07$	$1.69 \pm 1.91$	$6.61 \pm 2.99$
$t\bar{t} \rightarrow \ell(\not{b} \rightarrow \ell)X$	$1.05 \pm 1.73$	$0.59 \pm 1.51$	$0.50 \pm 1.51$	$2.13 \pm 2.07$
$t\bar{t}$ other	$0.00 \pm 1.21$	$0.00 \pm 1.21$	$0.00 \pm 1.21$	$0.00 \pm 1.21$
t, s-channel	$0.00 \pm 0.52$	$0.00 \pm 0.52$	$0.00 \pm 0.52$	$0.00 \pm 0.52$
t, t-channel	$0.00 \pm 0.54$	$0.00 \pm 0.54$	$0.00 \pm 0.54$	$0.00 \pm 0.54$
tW	$0.00 \pm 0.80$	$0.00 \pm 0.80$	$0.00 \pm 0.80$	$0.00 \pm 0.80$
$DY \rightarrow \ell\ell$	$0.00 \pm 4.14$	$2.01 \pm 5.18$	$0.00 \pm 4.14$	$2.01 \pm 5.18$
$W + jets \rightarrow \ell\nu$	$0.00 \pm 73.20$	$0.00 \pm 73.20$	$0.00 \pm 73.20$	$0.00 \pm 73.20$
WW	$0.00 \pm 0.11$	$0.00 \pm 0.11$	$0.00 \pm 0.11$	$0.00 \pm 0.11$
$W\gamma^* \rightarrow \ell\nu\mu\mu$	$0.00 \pm 0.23$	$0.00 \pm 0.23$	$0.00 \pm 0.23$	$0.00 \pm 0.23$
$W\gamma^* \rightarrow \ell\nu\tau\tau$	$0.00 \pm 0.24$	$0.00 \pm 0.24$	$0.00 \pm 0.24$	$0.00 \pm 0.24$
WZ	$0.04 \pm 0.03$	$0.03 \pm 0.03$	$0.09 \pm 0.04$	$0.16 \pm 0.05$
ZZ	$0.01 \pm 0.00$	$0.00 \pm 0.00$	$0.01 \pm 0.00$	$0.02 \pm 0.00$
$t\bar{t}\gamma$	$0.35 \pm 0.54$	$0.00 \pm 1.08$	$0.35 \pm 0.54$	$0.70 \pm 0.65$
$t\bar{t}W$	$2.21 \pm 0.27$	$2.97 \pm 0.30$	$5.17 \pm 0.39$	$10.34 \pm 0.54$
$t\bar{t}Z$	$0.52 \pm 0.13$	$0.88 \pm 0.16$	$1.56 \pm 0.20$	$2.96 \pm 0.27$
$t\bar{t}Z(Z \rightarrow \ell\ell)$	$0.03 \pm 0.01$	$0.03 \pm 0.01$	$0.07 \pm 0.01$	$0.13 \pm 0.02$
$t\bar{t}WW$	$0.06 \pm 0.00$	$0.08 \pm 0.00$	$0.13 \pm 0.01$	$0.27 \pm 0.01$
$WW\gamma$	$0.00 \pm 0.09$	$0.00 \pm 0.09$	$0.00 \pm 0.09$	$0.00 \pm 0.09$
WWW	$0.01 \pm 0.02$	$0.01 \pm 0.02$	$0.03 \pm 0.02$	$0.05 \pm 0.03$
WWZ	$0.00 \pm 0.01$	$0.00 \pm 0.01$	$0.01 \pm 0.01$	$0.02 \pm 0.02$
WZZ	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.01 \pm 0.01$
ZZZ	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$qqW^\pm W^\pm$	$0.00 \pm 0.09$	$0.02 \pm 0.04$	$0.07 \pm 0.12$	$0.09 \pm 0.12$
WW(DPS)	$0.00 \pm 0.03$	$0.00 \pm 0.03$	$0.00 \pm 0.03$	$0.00 \pm 0.03$
WH, ZH, $t\bar{t}H; H \rightarrow WW$	$0.54 \pm 0.15$	$0.70 \pm 0.17$	$1.03 \pm 0.20$	$2.27 \pm 0.28$
WH, ZH, $t\bar{t}H; H \rightarrow ZZ$	$0.02 \pm 0.01$	$0.04 \pm 0.01$	$0.07 \pm 0.01$	$0.14 \pm 0.02$
WH, ZH, $t\bar{t}H; H \rightarrow \tau\tau$	$0.02 \pm 0.01$	$0.06 \pm 0.02$	$0.08 \pm 0.02$	$0.16 \pm 0.03$
Total MC	$8.12 \pm 73.41$	$9.66 \pm 73.47$	$11.45 \pm 73.40$	$29.23 \pm 73.52$
SF	$4.47 \pm 0.92$	$5.83 \pm 1.00$	$9.68 \pm 1.35$	$19.98 \pm 1.91$
DF	$0.28 \pm 0.08$	$0.28 \pm 0.07$	$0.49 \pm 0.10$	$1.05 \pm 0.15$
SC	$0.20 \pm 0.07$	$0.27 \pm 0.09$	$0.38 \pm 0.11$	$0.85 \pm 0.25$
SF + DF	$4.75 \pm 0.91$	$6.10 \pm 1.00$	$10.18 \pm 1.33$	$21.03 \pm 1.90$
SF + DF - SC	$4.56 \pm 0.91 \pm 2.28$	$5.83 \pm 1.00 \pm 2.91$	$9.79 \pm 1.34 \pm 4.90$	$20.18 \pm 1.91 \pm 10.09$
Charge Flips	$0.48 \pm 0.04 \pm 0.14$	$0.00 \pm 0.00 \pm 0.00$	$0.50 \pm 0.05 \pm 0.15$	$0.98 \pm 0.06 \pm 0.29$
MC Pred	$3.82 \pm 0.73 \pm 1.91$	$4.83 \pm 1.20 \pm 2.42$	$8.67 \pm 0.81 \pm 4.33$	$17.32 \pm 1.00 \pm 8.66$
Total Pred	$8.85 \pm 1.17 \pm 2.98$	$10.66 \pm 1.56 \pm 3.79$	$18.96 \pm 1.57 \pm 6.54$	$38.47 \pm 2.16 \pm 13.30$
Data	7	14	24	45

In baselines SR10 and SR20, a b-tagged jet requirement is made. This reduces the contributions from mostly rare SM process such as  $WZ$  and  $qqW^\pm W^\pm$  where there are no intrinsic b-quarks in the final state. In these regions, the contributions from fakes leptons becomes enhanced and the main rare SM background is now  $t\bar{t}W$ . Figure 7.2 shows a graphical representation of the yields and there full background predictions for SR0, SR10 and SR20. Here we see the background prediction and the data yield are in agreement within the uncertainty. This lends confidence that our data-driven methods are performing reasonably well. In all search regions, we see reasonable consistency between the data yields and the background prediction.



**Figure 7.5:** Graphical representation of the event yields in the low  $p_T$  baseline region with no (top left), exactly one (top right) and  $\geq 2$  (bottom column) # b-tagged jets requirement (Search Region 0, 10 and 20). Also shown as a histogram is the result of the background prediction. The shading around the histogram represents the uncertainty in the background prediction.

### 7.2.2 Low $p_T$ Search Regions

The general search regions for the low  $p_T$  analysis are defined in Table 4.10 and are repeated in Table 7.11. The observed yields and estimated backgrounds for each region are shown in Table 7.12. The uncertainty is the total statistical and systematic uncertainty. A graphical representation of these results can also be seen in Figure 7.6. We see no evidence of an event yield in significant excess of the background estimations.

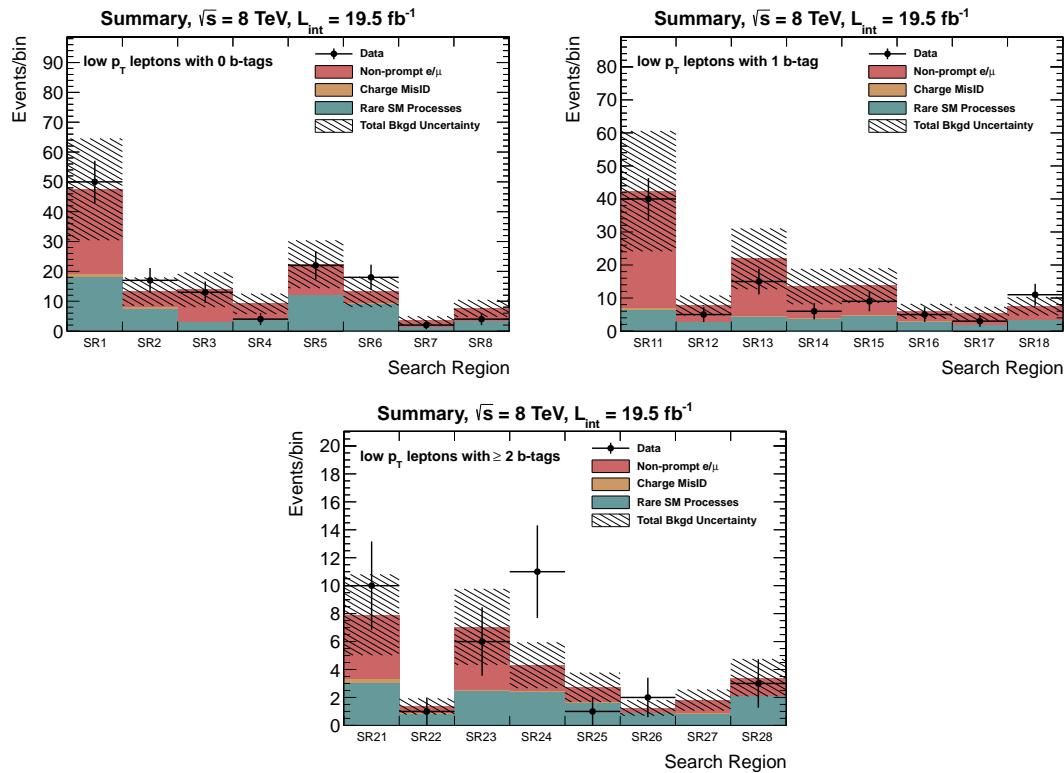
**Table 7.11:** Search regions selected for the low  $p_T$  search where we require  $H_T > 250$  GeV.

# $b$ -tagged jets	$\cancel{E}_T$	# jets	$H_T$ [250-400]	$H_T$ [ $> 400$ ]
$= 0$	50-120	2-3	SR1	SR2
		$\geq 4$	SR3	SR4
	$> 120$	2-3	SR5	SR6
		$\geq 4$	SR7	SR8
$= 1$	50-120	2-3	SR11	SR12
		$\geq 4$	SR13	SR14
	$> 120$	2-3	SR15	SR16
		$\geq 4$	SR17	SR18
$\geq 2$	50-120	2-3	SR21	SR22
		$\geq 4$	SR23	SR24
	$> 120$	2-3	SR25	SR26
		$\geq 4$	SR27	SR28

In the various search regions, irreducible backgrounds, primarily from  $t\bar{t}W$ ,  $WZ$  and  $qqW^\pm W^\pm$ , are the dominant backgrounds at approximately 15-70%, depending on the search region. Fake leptons account for approximately 25-80% of the background also depending on the search region. The background from charge mis-reconstruction is small contributing less than 5%.

**Table 7.12:** A summary of the results of this search for the low  $p_T$  analysis. For each signal region, we show its most important kinematical requirements, the prediction for the three background (BG) components as well as the total, and the event yield.

# $b$ -tagged jets	$\cancel{E}_T$	# jets	$H_T$	SR	Fake BG	Flip BG	Rare MC	Total BG	Observe	
$= 0$	$\geq 0$	30 if $H_T < 500$ else 0	2	250	0	241.48 $\pm$ 121.13	8.43 $\pm$ 2.58	118.51 $\pm$ 59.31	368.42 $\pm$ 134.89	349
	50-120	2-3	250-400	1	28.49 $\pm$ 14.44	0.95 $\pm$ 0.29	18.05 $\pm$ 9.07	47.49 $\pm$ 17.05	50	
			> 400	2	5.46 $\pm$ 2.89	0.41 $\pm$ 0.13	7.35 $\pm$ 3.73	13.22 $\pm$ 4.72	17	
		$\geq 4$	250-400	3	10.94 $\pm$ 5.63	0.12 $\pm$ 0.04	2.80 $\pm$ 1.51	13.86 $\pm$ 5.83	13	
			> 400	4	5.43 $\pm$ 2.87	0.16 $\pm$ 0.05	3.51 $\pm$ 1.84	9.10 $\pm$ 3.41	4	
	> 120	2-3	250-400	5	10.48 $\pm$ 5.41	0.13 $\pm$ 0.04	11.78 $\pm$ 5.93	22.39 $\pm$ 8.02	22	
			> 400	6	3.98 $\pm$ 2.16	0.07 $\pm$ 0.02	9.06 $\pm$ 4.58	13.11 $\pm$ 5.07	18	
		$\geq 4$	250-400	7	2.39 $\pm$ 1.35	0.02 $\pm$ 0.01	1.08 $\pm$ 0.68	3.49 $\pm$ 1.51	2	
			> 400	8	4.44 $\pm$ 2.36	0.03 $\pm$ 0.01	3.14 $\pm$ 1.65	7.60 $\pm$ 2.88	4	
$= 1$	$\geq 0$	30 if $H_T < 500$ else 0	2	250	10	117.46 $\pm$ 58.99	2.00 $\pm$ 0.61	33.43 $\pm$ 16.77	152.88 $\pm$ 61.33	132
	50-120	2-3	250-400	11	35.51 $\pm$ 17.93	0.56 $\pm$ 0.17	6.29 $\pm$ 3.20	42.36 $\pm$ 18.22	40	
			> 400	12	5.13 $\pm$ 2.73	0.14 $\pm$ 0.04	2.47 $\pm$ 1.32	7.74 $\pm$ 3.03	5	
		$\geq 4$	250-400	13	17.60 $\pm$ 8.96	0.14 $\pm$ 0.05	4.11 $\pm$ 2.12	21.85 $\pm$ 9.21	15	
			> 400	14	9.74 $\pm$ 5.02	0.10 $\pm$ 0.03	3.64 $\pm$ 1.89	13.47 $\pm$ 5.36	6	
	> 120	2-3	250-400	15	9.15 $\pm$ 4.72	0.23 $\pm$ 0.07	4.40 $\pm$ 2.26	13.78 $\pm$ 5.24	9	
			> 400	16	3.16 $\pm$ 1.73	0.07 $\pm$ 0.02	2.75 $\pm$ 1.46	5.98 $\pm$ 2.26	5	
		$\geq 4$	250-400	17	3.58 $\pm$ 1.94	0.04 $\pm$ 0.01	1.61 $\pm$ 0.92	5.23 $\pm$ 2.14	3	
			> 400	18	4.22 $\pm$ 2.25	0.05 $\pm$ 0.02	3.27 $\pm$ 1.71	7.54 $\pm$ 2.82	11	
$\geq 2$	$\geq 0$	30 if $H_T < 500$ else 0	2	250	20	20.18 $\pm$ 10.27	0.98 $\pm$ 0.30	17.32 $\pm$ 8.72	38.47 $\pm$ 13.47	45
	50-120	2-3	250-400	21	4.60 $\pm$ 2.45	0.32 $\pm$ 0.10	3.00 $\pm$ 1.58	7.92 $\pm$ 2.91	10	
			> 400	22	0.46 $\pm$ 0.36	0.06 $\pm$ 0.02	0.83 $\pm$ 0.57	1.35 $\pm$ 0.68	1	
		$\geq 4$	250-400	23	4.50 $\pm$ 2.41	0.11 $\pm$ 0.03	2.47 $\pm$ 1.32	7.07 $\pm$ 2.75	6	
			> 400	24	1.84 $\pm$ 1.07	0.07 $\pm$ 0.02	2.42 $\pm$ 1.29	4.33 $\pm$ 1.68	11	
	> 120	2-3	250-400	25	1.05 $\pm$ 0.65	0.10 $\pm$ 0.03	1.57 $\pm$ 0.89	2.72 $\pm$ 1.11	1	
			> 400	26	0.34 $\pm$ 0.29	0.03 $\pm$ 0.01	0.89 $\pm$ 0.59	1.25 $\pm$ 0.66	2	
		$\geq 4$	250-400	27	0.95 $\pm$ 0.60	0.03 $\pm$ 0.01	0.83 $\pm$ 0.58	1.81 $\pm$ 0.83	0	
			> 400	28	1.29 $\pm$ 0.79	0.04 $\pm$ 0.01	2.07 $\pm$ 1.12	3.40 $\pm$ 1.37	3	



**Figure 7.6:** Graphical representation of the event yields in the low  $p_T$  search regions with one (top left), exactly one (top right) and  $\geq 2$  (bottom column) #  $b$ -tagged jets requirement (Search Region 1-8, 11-18, and 21-28). Also shown as a histogram is the result of the background prediction. The shading around the histogram represents the uncertainty in the background prediction.

## 7.3 Interpretations

Given the lack of a significant excess over the expected SM background, the results of the search are used to place limits on the cross section of rare SM processes and on the parameters of various models of new physics. In total, we have 54 search regions (SR) for this analysis, but for the limit setting procedure for each model we take into account only the dedicated search regions as described in Section 4.3. For each model in consideration, all relevant search regions are included into a multi-bin fit in order to maximize sensitivity. The 95% confidence level (CL) upper limits on the signal yields are calculated using the modified frequentist  $CL_s$  method. Formally we construct a likelihood function, i.e.,

$$\mathcal{L}(\text{data}|\mu, \theta) = \text{Poisson}\left(\text{data}|\mu \cdot s(\theta) + b(\theta)\right) \cdot p(\hat{\theta}|\theta), \quad (7.1)$$

where “data” represents our observed yields, and  $\text{Poisson}(\text{data}|\mu, \theta)$  represents the product of probabilities to observe  $n_i$  events in the  $i$ th region, i.e.,

$$\prod_i \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-\mu s_i - b_i}. \quad (7.2)$$

Here, the parameter  $\mu$  represents our signal strength modifier and corresponds to  $\sigma \times BR$  in the context of the R-parity-conserving simplified SUSY models (SMS) [80]. The parameter  $\theta$  denotes the collection of nuisance parameters that can influence the values of  $b$  and  $s$ . Log-normal pdf’s are assumed for the signal and background nuisance parameters. These are taken to be either 0% or 100% correlated where appropriate. The profile likelihood ratio is used to test the compatibility of the observed data with the background-only and signal+background hypotheses. The full prescription of the statistical model can be found elsewhere [81, 82, 83].

In this section, we first give the upper limits on the cross sections for same-sign top production followed by the Standard Model four top production. We conclude this chapter with the limits on various Supersymmetry models with Simplified Model Spaces (SMS), which can lead to a same-sign dilepton signature.

### 7.3.1 Same-Sign Top

The results of the search regions SStop1, SStop1++, SStop2, and SStop2++ are used to set limits on the cross section for same-sign top-quark pair production,  $\sigma(pp \rightarrow tt + \bar{t}\bar{t})$  from SStop1 and SStop2, and  $\sigma(pp \rightarrow tt)$  from SStop1++ and SStop2++. Here  $\sigma(pp \rightarrow tt + \bar{t}\bar{t})$  is short hand for the sum  $\sigma(pp \rightarrow tt) + \sigma(pp \rightarrow \bar{t}\bar{t})$ . Note that in most new physics scenarios  $pp \rightarrow \bar{t}\bar{t}$  is suppressed with respect to  $pp \rightarrow tt$  because of the parton distribution functions of the proton. These limits are calculated using an acceptance obtained using simulated  $pp \rightarrow t\bar{t}$  events and an opposite-sign selection. This acceptance, including branching fractions, is 0.43% (0.26%) for SStop1 (SStop2) search region. The relative uncertainty on this acceptance is 14%. The observed upper limits are  $\sigma(pp \rightarrow tt + \bar{t}\bar{t}) < 0.71$  pb and  $\sigma(pp \rightarrow tt) < 0.33$  pb at 95% CL; the median expected limits are 0.48 and 0.30 pb, respectively. The full procedure was documented in [69].

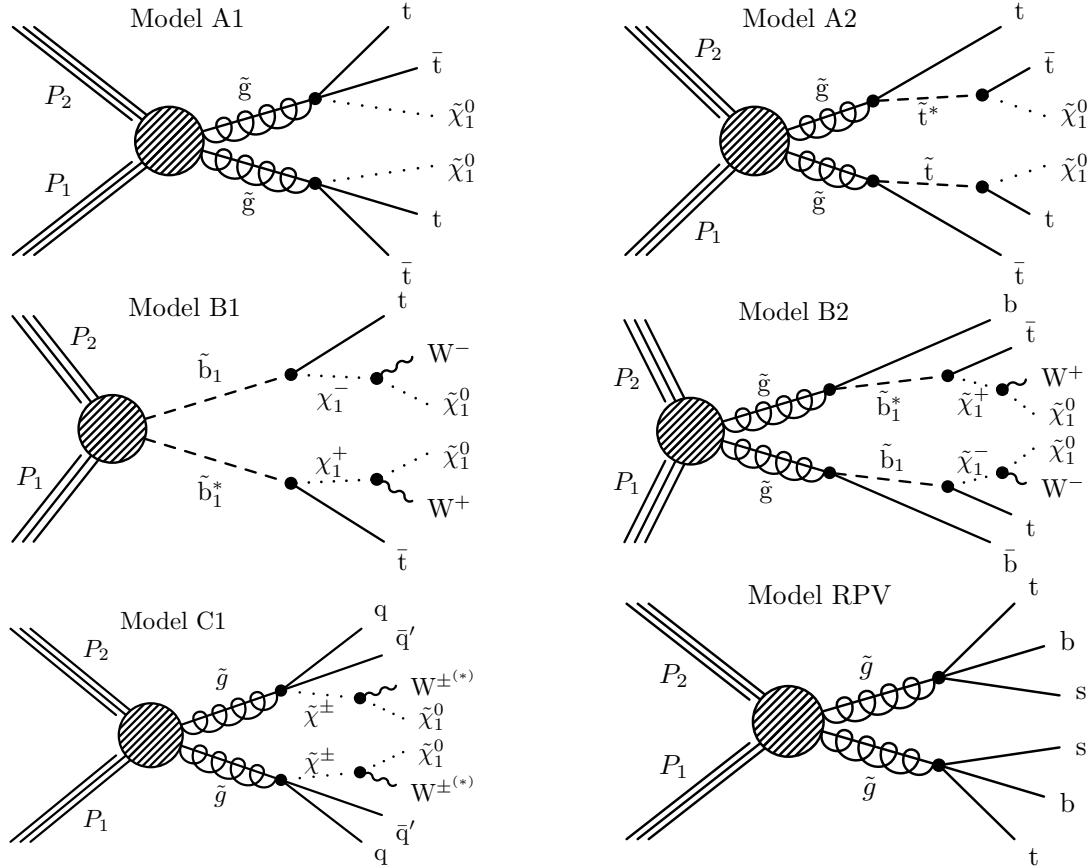
### 7.3.2 Standard Model Four Top

Similarly, the results from search regions SR21–SR28 and high- $p_T$  lepton selection are used to set limits on the SM cross section for the Standard Model four-top-quark production. The observed upper limit is  $\sigma(pp \rightarrow tt\bar{t}\bar{t}) < 47$  fb at 95% CL, compared to a median expected limit of 36 fb. The SM cross section as computed with the MC@NLO program [84] is  $\sigma_{\text{SM}} = (9.14 \pm 0.05) \times 10^{-1}$  fb. The most sensitive search regions, SR24 and SR28, have a signal acceptance of 0.52% and 0.49%, respectively, with relative uncertainties of 13% and 17%.

### 7.3.3 Supersymmetry

Next, we present limits on the parameter spaces of various R-parity-conserving simplified Supersymmetry models (SMS) [80]. The exclusion contours are obtained with the gluino (sbottom) pair production cross-sections at the NLO + NLL accuracy that are calculated in the limit where other sparticles are heavy and decoupled [85, 86, 87, 88, 89, 90]. The production of SUSY particles and the decay chains under consideration are shown schematically in Figure 7.7 and the summary of the search regions used

for each SMS model is given in Table 7.13 with the results from each model discussed below.

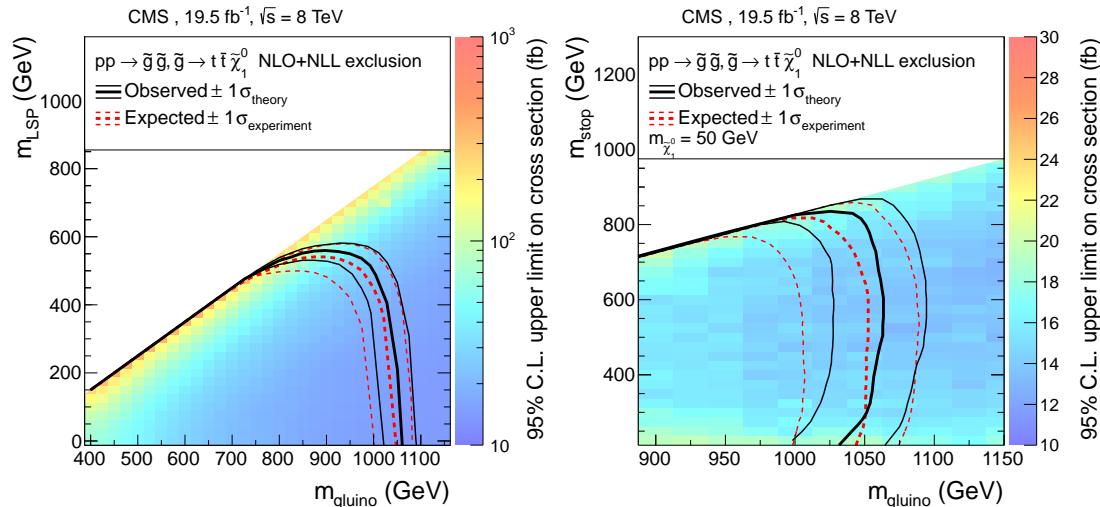


**Figure 7.7:** Feynman diagrams for the four SUSY models considered (A1, A2, B1, B2, C1 and RPV).

**Table 7.13:** Summary of the search regions used for limit setting in each SMS model considered.

Model	Model Restriction	Analysis Type	Signal Regions Used
A1		high- $p_T$	21–28
A2	$m_{\tilde{\chi}^0} = 50 \text{ GeV}$	high- $p_T$	21–28
B1	$m_{\tilde{\chi}^0} = 50 \text{ GeV}$	high- $p_T$	11–18, 21–28
B1	$x = m_{\tilde{\chi}^0}/m_{\tilde{\chi}_1^\pm} = 0.5$	high- $p_T$	11–18, 21–28
B1	$x = m_{\tilde{\chi}^0}/m_{\tilde{\chi}_1^\pm} = 0.8$	low- $p_T$	11–18, 21–28
B2	$m_{\tilde{\chi}^0} = 50 \text{ GeV}, m_{\tilde{\chi}_1^\pm} = 150 \text{ GeV}$	high- $p_T$	21–28
B2	$m_{\tilde{\chi}^0} = 50 \text{ GeV}, m_{\tilde{\chi}_1^\pm} = 300 \text{ GeV}$	high- $p_T$	21–28
C1	$x = 0.5$	high- $p_T$	01–08
C1	$x = 0.8$	low- $p_T$	01–08
RPV		high- $p_T$	RPV2

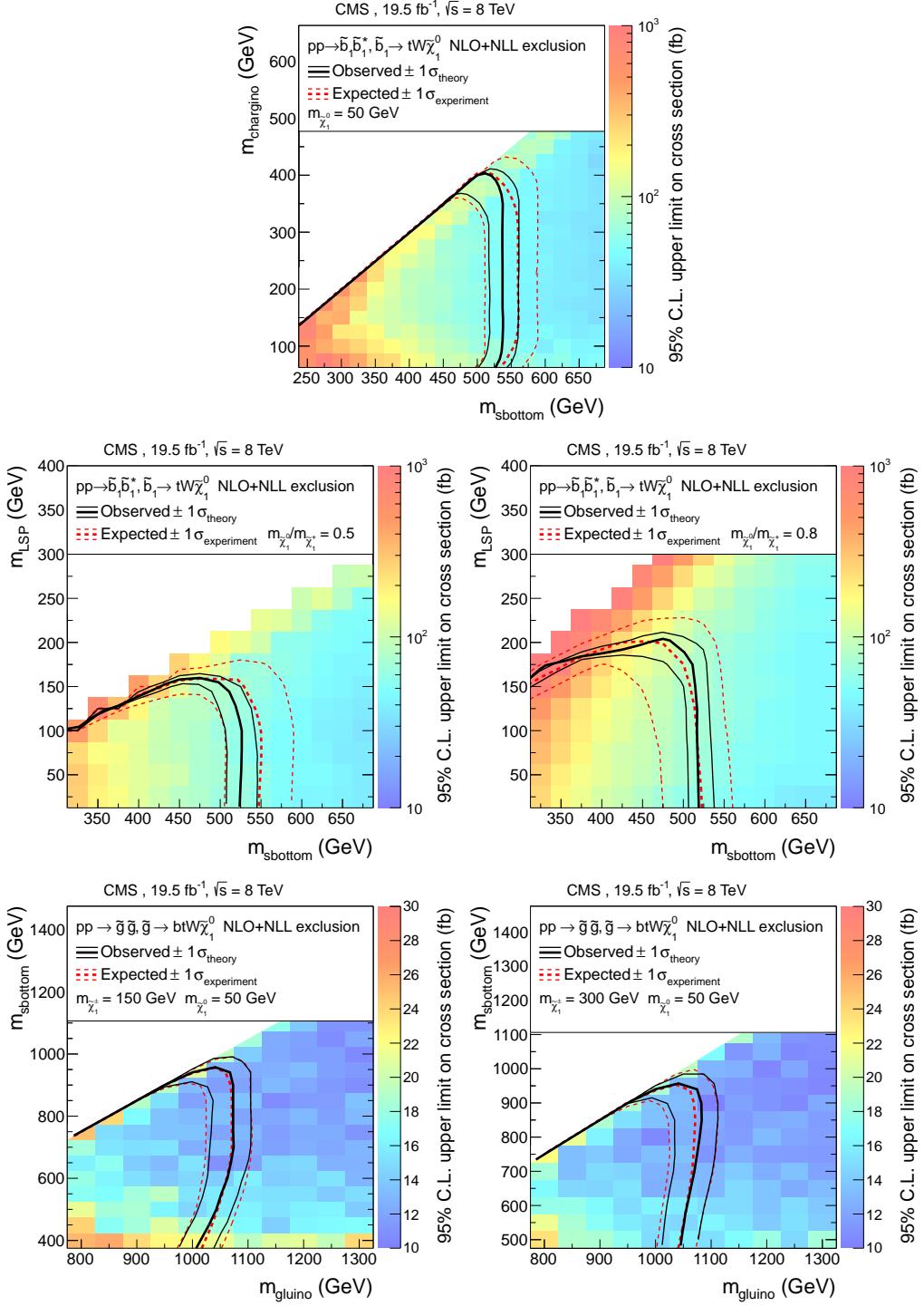
Scenarios A1 and A2 represent models of gluino pair production resulting in the  $t\bar{t}t\bar{t}\tilde{\chi}_1^0\tilde{\chi}_1^0$  final state, where  $\tilde{\chi}_1^0$  is the lightest neutralino [91, 92, 80, 93, 94]. In model A1, the gluino undergoes a three-body decay  $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$  mediated by an off-shell top squark. In model A2, the gluino decays to a top quark and an anti-top squark, with the on-shell anti-squark further decaying into an anti-top quark and a neutralino. Both of these models produce four on-shell W bosons and four bottom quarks. Therefore, these models can be probed with search regions SR21–SR28 where at least two b-tagged jets and high  $p_T$  leptons are required, the region with the best sensitivity is SR28. The 95% CL upper limits on the cross section times branching ratio, as well as the exclusion contours, are shown in Figure 7.8. For model A1 these are presented as a function of gluino mass and  $\tilde{\chi}_1^0$  mass and for model A2 as a function of gluino mass and stop mass for  $\tilde{\chi}_1^0$  mass set to 50 GeV.



**Figure 7.8:** Exclusion regions at 95% CL in the planes of  $m(\tilde{\chi}_1^0)$  vs.  $m(\tilde{g})$  (model A1) and  $m(\tilde{t}_1)$  vs.  $m(\tilde{g})$  (model A2). The excluded regions are those within the kinematic boundaries and to the left of the curves. The effects of the theoretical uncertainties in the next-to-leading-order plus next-to-leading-log calculations of the production cross sections [95] are indicated by the black-thin curves; the expected limits and their  $\pm 1$  standard-deviation variations are also shown in dashed red curves.

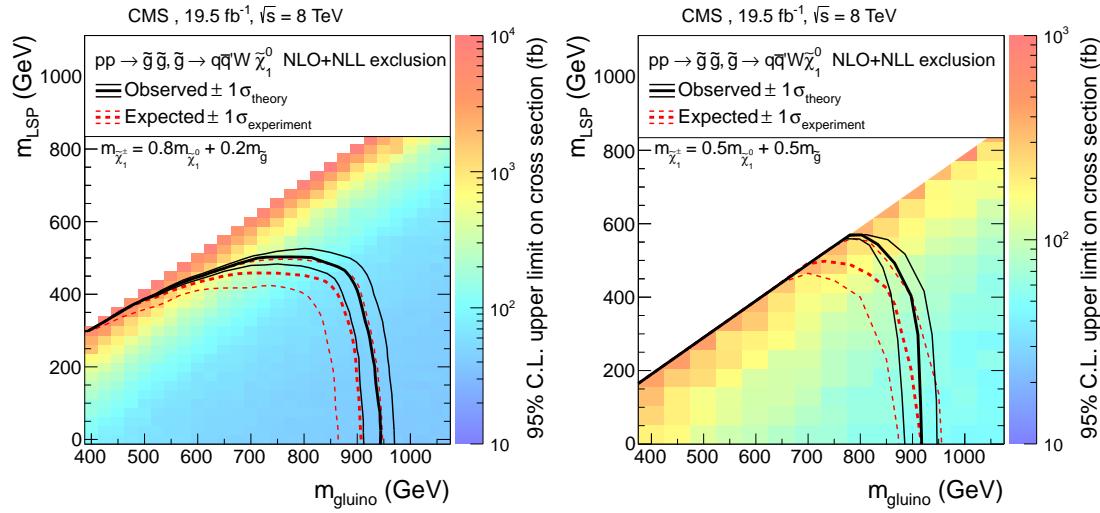
Model B1 is a model of bottom-squark pair production, followed by one of the most likely decay modes of the bottom squark,  $\tilde{b}_1 \rightarrow t\tilde{\chi}_1^-$  and  $\tilde{\chi}_1^- \rightarrow W^-\tilde{\chi}_1^0$ . We consider three scenarios in this decay mode. We either fix the  $\tilde{\chi}_1^0$  mass at 50 GeV and present the limits in the  $m_{\tilde{b}_1}$ - $m_{\tilde{\chi}_1^\pm}$  plane, or use  $m_{\tilde{b}_1}$ - $m_{\tilde{\chi}_1^0}$  plane with variable mass of  $m_{\tilde{\chi}_1^\pm}$  using the equation  $m_{\tilde{\chi}_1^\pm} = xm_{\tilde{\chi}_1^0} + (1-x)m_{\tilde{b}_1}$  with  $x = 0.5$  or  $x = 0.8$ . The value of  $x$  determines whether the top and W are produced on- or off-shell. The limits on this model, obtained using search regions SR11 to SR28, are presented in Figure 7.9. For the  $x = 0.8$  scenario the low- $p_T$  lepton selection is used, while high- $p_T$  leptons are used for the other two scenarios. SR28 is again the most sensitive search region, followed by the regions requiring one b-tagged jet: SR18, SR15, and SR13.

Model B2 consists of gluino pair production followed by  $\tilde{g} \rightarrow \tilde{b}_1\bar{b}$ . The gluino decay modes in models A1 and A2 would be dominant if the top squark is the lightest supersymmetric quark. Conversely, if the bottom squark is the lightest, the decay mode in model B2 would be the most probable. The limits on this model, calculated using search regions SR21–SR28 and the high- $p_T$  lepton selection, are presented as a function of  $m(\tilde{b}_1)$  and  $m(\tilde{g})$  for two fixed masses of  $m_{\tilde{\chi}_1^\pm}$ , 150 and 300 GeV. The region with the largest sensitivity to this model is SR28.



**Figure 7.9:** Exclusion regions at 95% CL in the planes of  $m(\tilde{\chi}_1^\pm)$  vs.  $m(\tilde{b}_1)$  and  $m(\tilde{\chi}_1^0)$  vs.  $m(\tilde{b}_1)$  (model B1) and  $m(\tilde{b}_1)$  vs.  $m(\tilde{g})$  (model B2). The convention for the exclusion curves are same as in Figure 7.9.

Model C1 is another decay chain following a gluino pair production  $pp \rightarrow \tilde{g}\tilde{g}$ , where each gluino decays to light quarks and a chargino via a heavy virtual squark:  $\tilde{g} \rightarrow q\bar{q}'\tilde{\chi}_1^\pm, \tilde{\chi}_1^\pm \rightarrow W^{(*)}\tilde{\chi}_1^0$ . The decay is charge democratic, resulting in an equal fraction of same-sign and opposite-sign W pairs in the final state. In this model there are three parameters:  $m_{\tilde{g}}$ ,  $m_{\tilde{\chi}_1^\pm}$ , and  $m_{\tilde{\chi}_1^0}$ . The scan is performed in the  $m_{\tilde{\chi}_1^0}$ ,  $m_{\tilde{g}}$  plane, while the chargino mass is defined as  $m_{\tilde{\chi}_1^\pm} = xm_{\tilde{\chi}_1^0} + (1-x)m_{\tilde{g}}$  where we examine  $x$  values of 0.5 and 0.8. Depending on the value of the  $x$ -parameter, for a given  $m_{\tilde{\chi}_1^0}$  and  $m_{\tilde{g}}$ , the W can be produced on-shell or off-shell giving rise to either high- or low- $p_T$  leptons. In this model no enrichment of heavy-flavour jets is expected. Therefore, the search regions SR01–SR08, with both low- and high- $p_T$  lepton selection, are used for cross section upper limit calculation. In Figure 7.10 we present the limit in the  $m_{\tilde{\chi}_1^0}$ – $m_{\tilde{g}}$  plane with  $x = 0.8$ . In this model, gluino masses up to 900 GeV are probed. Most of the sensitivity for this model is obtained from search region SR08.



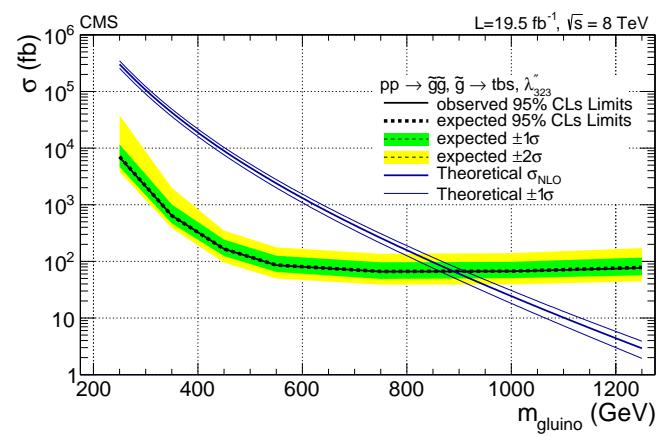
**Figure 7.10:** Exclusion regions at 95% CL in the planes of  $m(\tilde{\chi}_1^0)$  vs.  $m(\tilde{g})$  (model C1). The convention for the exclusion curves are same as in Figure 7.10.

Excluded regions in the parameter space of the five models are shown in Figures 7.8, 7.9, and 7.10. These results extend the sensitivity obtained in the previous version of this analysis [5] on gluino and sbottom masses. For the gluino-initiated models (A1, A2, B2, and C1), we probe gluinos with masses below around 1050 GeV, with

relatively small dependence on the details of the models. This is because the limits are driven by the common gluino-pair production cross section. In the case of the direct bottom-squark pair production, model B1, our search shows sensitivity for sbottom masses up to 500 GeV.

These models are also probed by other CMS new physics searches in different decay modes, although other searches have so far been interpreted in the context of model A1 but not A2, B1, or B2. For model A1 the limits given here are complementary to the limits from other searches [96, 97, 98, 99]: less stringent at low  $m(\tilde{\chi}_1^0)$  but more stringent at high  $m(\tilde{\chi}_1^0)$ . A similar conclusion applies to model A2, since the final state is the same, while in the case of bottom-squark pair production limits on  $m(\tilde{b}_1)$  are about 500(600) GeV, but assuming the decay mode  $\tilde{b}_1 \rightarrow b\tilde{\chi}_1^0$  instead of the model B1 mode  $\tilde{b}_1 \rightarrow t\tilde{\chi}_1^-$  [100]. Comparable limits for model A1, as well as for similar models with top and bottom quarks from gluino decays, have been reported by the ATLAS collaboration [101, 102, 103, 104].

One R-parity violating scenario is considered in the scope of gluino-pair production, where each gluino decays to three quarks:  $\tilde{g} \rightarrow tbs(\bar{t}\bar{s})$  (model RPV). Such decays lead to same-sign W pairs in the final state in 50% of the cases. The model is governed by one parameter ( $m_{\tilde{g}}$ ) which dictates the production cross section and the final state kinematics. The dedicated search region RPV2 with the high- $p_T$  lepton selection is used to place an upper limit on the production cross section. The signal selection efficiency is 0.1% for low gluino masses and reaches to 0.5% for the higher gluino masses considered in this analysis. The result is shown in Figure 7.11. In this scenario, the gluino mass is probed up to approximately 900 GeV.



**Figure 7.11:** Exclusion at 95% CL upper limit on the gluino production cross section for an RPV simplified model,  $p p \rightarrow \tilde{g}\tilde{g}, \tilde{g} \rightarrow tbs(\bar{t}\bar{b}\bar{s})$ .

# Chapter 8

## Summary and Conclusions

This thesis reports on a search for new physics in a final state with two same-sign leptons, missing transverse energy, and significant hadronic activity. The results use the data collected by the CMS detector from 2012 corresponding to  $19.5 \text{ fb}^{-1}$  of integrated luminosity. No significant deviation from the Standard Model expectations were observed.

The results are used to set upper limits on the same-sign top-pair production cross section  $\sigma(pp \rightarrow tt + \bar{t}\bar{t}) < 0.71 \text{ pb}$  and  $\sigma(pp \rightarrow tt) < 0.33 \text{ pb}$  at the 95% confidence level. Also the upper limit on the SM cross section for the four-top-quark production is computed to be  $\sigma(pp \rightarrow ttt\bar{t}) < 47 \text{ fb}$  at the 95% confidence level.

Upper limits on the cross section for several models of supersymmetry are also set. Specifically, sparticle production cross-sections calculated assuming that gluinos decay exclusively into top or bottom squarks. The probed models show an upper limit on gluinos masses up to  $\sim 1050 \text{ GeV}/c^2$  and on bottom squarks with masses up to  $\sim 500 \text{ GeV}/c^2$ . In models where no third generation squarks were involved in the gluino decays, a somewhat weaker limit is obtained for the probed gluino masses of up to  $\sim 900 \text{ GeV}/c^2$ .

The future of this analysis looks towards the next LHC run that will increase the center-of-mass energy to 13 or 14 TeV. This, along with improvements to the estimation for the main SM backgrounds (i.e.  $t\bar{t}W$ ,  $t\bar{t}Z$ ,  $qqW^\pm W^\pm$ , and  $WH, ZH, t\bar{t}H$ ;  $H \rightarrow \tau\tau$ ) will result in a large increase in the sensitivity of this analysis for all of the supersymmetry models probed. The CMS collaboration intends to pursue this analysis when the LHC

comes back online sometime in 2015.

# Bibliography

- [1] G. Bayatian et al., “CMS Physics Technical Design Report Volume I: Detector Performance and Software”. Technical Design Report CMS. CERN, Geneva, 2006.
- [2] CMS HCAL Collaboration, “Design, performance, and calibration of CMS hadron-barrel calorimeter wedges”, *Eur. Phys. J. C* **55** (2008) 159–171.  
*doi:10.1140/epjc/s10052-008-0573-y*.
- [3] M. LeBourgeois, “Search for the Standard Model Higgs boson in the fully leptonic WW decay channel at CMS”,.
- [4] CERN, “Unified forces”, (2013).
- [5] CMS Collaboration, “Search for new physics in events with same-sign dileptons and b jets in pp collisions at  $\sqrt{s} = 8 \text{ TeV}$ ”, *JHEP* (2013)  
*arXiv:1014.3168*. *doi:10.1007/JHEP03(2013)037*.
- [6] F. Halzen and A. D. Martin, “Quarks & Leptons: An Introductory Course in Modern Particle Physics”. John Wiley & Sons, Inc., 1984.
- [7] [http://upload.wikimedia.org/wikipedia/commons/thumb/0/00/Standard\\_Model\\_of\\_Elementary\\_Particles.svg/2000px-Standard\\_Model\\_of\\_Elementary\\_Particles.svg.png](http://upload.wikimedia.org/wikipedia/commons/thumb/0/00/Standard_Model_of_Elementary_Particles.svg/2000px-Standard_Model_of_Elementary_Particles.svg.png).
- [8] W.-M. Yao, C. Amsler, D. Asner et al., “Review of Particle Physics”, *Journal of Physics G* **33** (2006) 1+.
- [9] C. Collaboration, “CMS Higgs Physics Results”, (2013).
- [10] CMS Collaboration, “Observation of a new boson with a mass near 125 GeV”, *CMS Physics Analysis Summary CMS-PAS-HIG-12-020* (2012).
- [11] [http://upload.wikimedia.org/wikipedia/commons/thumb/4/4c/Elementary\\_particle\\_interactions.svg/2000px-Elementary\\_particle\\_interactions.svg.png](http://upload.wikimedia.org/wikipedia/commons/thumb/4/4c/Elementary_particle_interactions.svg/2000px-Elementary_particle_interactions.svg.png).

- [12] L. Evans and P. Bryant, “LHC Machine”, *Journal of Instrumentation* **3** (2008), no. 08, S08001.
- [13] J. W. H. J. M. Campbell and W. J. Stirling, “Hard Interactions of Quarks and Gluons: a Primer for LHC Physics”,.
- [14] A. Martin et al., “Parton distributions for the LHC”, *Eur.Phys.J. C***63** (2009) 189–285, arXiv:0901.0002. doi:10.1140/epjc/s10052-009-1072-5.
- [15] J. S. G. Bertone, D. Hooper, “Particle Dark Matter: Evidence, Candidates, Constraints”, *Phys.Rept.* **405** (2004) arXiv::hep-ph/0404175v2. doi:10.1016/j.physrep.2004.08.031.
- [16] H. Baer, “TASI 2008 Lectures on Collider Signals II:  $E_T^{missing}$  Signatures and the Dark Matter Connection”, arXiv::0901.4732v1.
- [17] W. Stirling, “Standard Model cross sections as a function of collider energy”, (2012).
- [18] R. M. Barnett, J. F. Gunion, and H. E. Haber, “Discovering supersymmetry with like-sign dileptons”, *Phys. Lett. B* **315** (1993) 349, arXiv:hep-ph/9306204. doi:10.1016/0370-2693(93)91623-U.
- [19] H. Baer, C.-h. Chen, F. Paige et al., “Signals for minimal supergravity at the CERN large hadron collider. II: Multi - lepton channels”, *Phys. Rev. D* **53** (1996) 6241, arXiv:hep-ph/9512383. doi:10.1103/PhysRevD.53.6241.
- [20] M. Guchait and D. P. Roy, “Like-sign dilepton signature for gluino production at CERN LHC including top quark and Higgs boson effects”, *Phys. Rev. D* **52** (1995) 133, arXiv:hep-ph/9412329. doi:10.1103/PhysRevD.52.133.
- [21] H.-C. Cheng, K. T. Matchev, and M. Schmaltz, “Bosonic supersymmetry? Getting fooled at the CERN LHC”, *Phys. Rev. C* **66** (2002) 056006, arXiv:hep-ph/0205314. doi:10.1103/PhysRevD.66.056006.
- [22] R. Contino and G. Servant, “Discovering the top partners at the LHC using same-sign dilepton final states”, *JHEP* **06** (2008) 026, arXiv:0801.1679. doi:10.1088/1126-6708/2008/06/026.
- [23] F. M. L. Almeida Jr., Y. D. A. Coutinho, J. A. Martins Simões et al., “Same-sign dileptons as a signature for heavy Majorana neutrinos in hadron-hadron collisions”, *Phys. Lett. B* **400** (1997) 331, arXiv:hep-ph/9703441. doi:10.1016/S0370-2693(97)00143-3.
- [24] Y. Bai and Z. Han, “Top-antitop and top-top resonances in the dilepton channel at the CERN LHC”, *JHEP* **04** (2009) 056, arXiv:0809.4487. doi:10.1088/1126-6708/2009/04/056.

- [25] AMS Collaboration Collaboration, “First Result from the Alpha Magnetic Spectrometer on the International Space Station: Precision Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5–350 GeV”, *Phys.Rev.Lett.* **110** (2013), no. 14, 141102. doi:10.1103/PhysRevLett.110.141102.
- [26] G. Bertone, D. Hooper, and J. Silk, “Particle dark matter: Evidence, candidates and constraints”, *Phys.Rept.* **405** (2005) 279–390, arXiv:hep-ph/0404175. doi:10.1016/j.physrep.2004.08.031.
- [27] T. Virdee, “CMS: Building on innovation”, *CERN Courier* (2009).
- [28] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3 S08004** (2008) 361. doi:10.1088/1748-0221/3/08/S08004.
- [29] CMS Collaboration, “CMS tracking performance results from early LHC operation”, *The European Physical Journal C - Particles and Fields* **70** (2010) 1165–1192. 10.1140/epjc/s10052-010-1491-3.
- [30] CMS Collaboration, “The CMS muon project: Technical Design Report”,.
- [31] CMS Collaboration, “Absolute Calibration of Luminosity Measurement at CMS: Summer 2011 Update”, *CMS Physics Analysis Summary CMS-PAS-EWK-11-001* (2011).
- [32] CMS Collaboration, “Absolute Calibration of the Luminosity Measurement at CMS: Winter 2012 Update”, *CMS Physics Analysis Summary CMS-PAS-SMP-12-008* (2012).
- [33] CMS Collaboration, “CMS Luminosity Based on Pixel Cluster Counting: Summer 2012 Update”, *CMS Physics Analysis Summary* (2012).
- [34] S. van der Meer, “Calibration of the effective beam height in the ISR. oai:cds.cern.ch:296752”, Technical Report CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, (1968).
- [35] C. Collaboration, “CMS Luminosity - Public Results”, (2013).
- [36] CMS Collaboration, “CMS TriDAS project: Technical Design Report; 1, the trigger systems”,.
- [37] S. Cucciarelli, M. Konecki, D. Kotlinski et al., “Track reconstruction, primary vertex finding and seed generation with the Pixel Detector”, Technical Report CMS-NOTE-2006-026. CERN-CMS-NOTE-2006-026, CERN, Geneva, (Jan, 2006).
- [38] W. Adam, B. Mangano, T. Speer et al., “Track Reconstruction in the CMS tracker”, Technical Report CMS-NOTE-2006-041. CERN-CMS-NOTE-2006-041, CERN, Geneva, (Dec, 2006).

- [39] R. Frühwirth, “Application of Kalman filtering to track and vertex fitting”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **262** (1987), no. 2–3, 444 – 450. doi:10.1016/0168-9002(87)90887-4.
- [40] K. Rose, “Deterministic annealing for clustering, compression, classification, regression, and related optimization problems”, *Proceedings of the IEEE* **86** (nov, 1998) 2210 –2239. doi:10.1109/5.726788.
- [41] CMS Collaboration, A. Aimar, J. Harvey, and N. Knoors, “Deterministic annealing for vertex finding at CMS”, in *Computing in High Energy Physics and Nuclear Physics 2004*, CERN. CERN, Geneva, 2005.
- [42] T. Speer, K. Prokofiev, R. Frühwirth et al., “Vertex Fitting in the CMS Tracker”, Technical Report CMS-NOTE-2006-032. CERN-CMS-NOTE-2006-032, CERN, Geneva, (Feb, 2006).
- [43] R. Frühwirth, W. Waltenberger, and P. Vanlaer, “Adaptive Vertex Fitting”, Technical Report CMS-NOTE-2007-008. CERN-CMS-NOTE-2007-008, CERN, Geneva, (Mar, 2007).
- [44] CMS Collaboration, “Tracking and Primary Vertex Results in First 7 TeV Collisions”, *CMS Physics Analysis Summary* **CMS-PAS-TRK-10-005** (2010).
- [45] S. Baffioni, C. Charlot, F. Ferri et al., “Electron reconstruction in CMS”, *The European Physical Journal C - Particles and Fields* **49** (2007) 1099–1116. 10.1140/epjc/s10052-006-0175-5.
- [46] CMS Collaboration, “Electron reconstruction and identification at  $\sqrt{s} = 7$  TeV”, *CMS Physics Analysis Summary* **CMS-PAS-EGM-10-001** (2010).
- [47] CMS Collaboration, “Electron performance with  $19.6^{-1}$  of data collected at  $\sqrt{s} = 8$  TeV with the CMS detector”, *CMS Performance Note* **DP2013\_003** (2013).
- [48] W. Adam, R. Frühwirth, A. Strandlie et al., “RESEARCH NOTE FROM COLLABORATION: Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC”, *Journal of Physics G Nuclear Physics* **31** (September, 2005) 9, arXiv:arXiv:physics/0306087. doi:10.1088/0954-3899/31/9/N01.
- [49] CMS Collaboration, “Performance of CMS muon reconstruction in cosmic-ray events”, *Journal of Instrumentation* **5** (March, 2010) 3022–+, arXiv:0911.4994. doi:10.1088/1748-0221/5/03/T03022.
- [50] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET”, *CMS Physics Analysis Summary* **CMS-PAS-PFT-09-001** (2009).

- [51] CMS Collaboration, “Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV”, *CMS Physics Analysis Summary CMS-PAS-PFT-10-002* (2010).
- [52] CMS Collaboration, “Missing transverse energy performance of the CMS detector”, *Journal of Instrumentation* **6** (2011), no. 09, P09001.
- [53] M. Cacciari, G. P. Salam, and G. Soyez, “The anti-k<sub>t</sub> jet clustering algorithm”, *Journal of High Energy Physics* **4** (April, 2008) 63–+, arXiv:0802.1189. doi:10.1088/1126-6708/2008/04/063.
- [54] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas”, *Physics Letters B* **659** (2008), no. 1–2, 119 – 126. doi:10.1016/j.physletb.2007.09.077.
- [55] M. Cacciari, G. P. Salam, and G. Soyez, “The catchment area of jets”, *Journal of High Energy Physics* **2008** (2008), no. 04, 005.
- [56] CMS Collaboration, “Determination of jet energy calibration and transverse momentum resolution in CMS”, *Journal of Instrumentation* **6** (2011), no. 11, P11002.
- [57] CMS Collaboration, “Identification of b-quark jets with the CMS Experiment”, *Journal of Instrumentation* **8** (2013), no. 13, P04013.
- [58] D. Barge et al., “Search for New Physics in the Same Sign Dilepton final state with b jets and Missing Energy at the LHC”, *CMS AN-2011/367 2011/367* (2011).
- [59] C. Campagnari 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”, *CMS AN-2013/120 2013/120* (2013).
- [60] D. Dobur et al., “Search for new physics with same-sign dilepton events in CMS using full 2012 data”, *CMS AN-2013/051 2013/051* (2013).
- [61] C. Collaboration, “Muon ID Recommendations”, (2012).
- [62] M. Mulders et al. *CMS AN-2008/098* (2008).
- [63] C. Collaboration, “Muon PF-Based Isolation Recommendations”, (2012).
- [64] C. Collaboration, “EGammaPOG PF-Based Isolation Recommendations”, (2012).
- [65] C. Collaboration, “EGammaPOG PF-Based  $\rho$  Correction”, (2012).
- [66] C. Collaboration, “EGammaPOG ID Recommendations”, (2012).

- [67] N. Saoulidou et al., “Particle Flow Jet Identification Criteria”, *CMS AN-2010/003* (2010).
- [68] C. Collaboration, “Jet Energy Corrections”, (2012).
- [69] D. Barge et al., “Search for new physics in a final state with a same-sign lepton pair, b-tagged jets, and missing transverse energy”, *CMS AN-2012/177 2012/177* (2012).
- [70] D. Barge 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”, *CMS AN-2012/364 2012/364* (2012).
- [71] D. Barge et al., “Search for New Physics with Same-Sign Dileptons using the 2011 dataset of CMS”, *CMS AN-2011/468 2011/468* (2011).
- [72] CMS Collaboration, “Search for new physics with same-sign isolated dilepton events with jets and missing transverse energy at the LHC”, *JHEP* (2011) arXiv:1014.3168. doi:10.1007/JHEP06(2011)077.
- [73] CMS Collaboration, “Search for new physics with same-sign isolated dilepton events with jets and missing transverse energy”, *Phys. Rev. Lett.* **109** (2012) arXiv:1205.6615. doi:10.1103/PhysRevLett.109.071803.
- [74] D. Barge et al., “Inclusive Search for New Physics with Same-Sign Dileptons using early LHC data”, *CMS AN-2010/247 2010/247* (2010).
- [75] J. M. Campbell and R. K. Ellis, “ $t\bar{t}W^\pm$  production and decay at NLO”, *JHEP* **07** (2012) 052, arXiv:1204.5678. doi:10.1007/JHEP07(2012)052.
- [76] A. Kardos, Z. Trocsanyi, and C. Papadopoulos, “Top quark pair production in association with a Z-boson at NLO accuracy”, *Phys. Rev. D* **85** (2012) 054015, arXiv:1111.0610. doi:10.1103/PhysRevD.85.054015.
- [77] CMS Collaboration, “Identification of b-quark jets with the CMS experiment”, *JINST* **8** (2013) P04013, arXiv:1211.4462. doi:10.1088/1748-0221/8/04/P04013.
- [78] C. Collaboration, “Methods to apply b-tagging efficiency scale factors”, (2012).
- [79] C. Clooper and E. S. Pearson, “The use of confidence or fiducial limits illustrated in the case of the binomial”,. doi:doi:10.1093/biomet/26.4.404.
- [80] LHC New Physics Working Group Collaboration, “Simplified models for LHC new physics searches”, *J. Phys. G* **39** (2012) 105005, arXiv:1105.2838. Model of Section IV.E with topology (B+B). doi:10.1088/0954-3899/39/10/105005.

- [81] A. L. Read, “Presentation of search results: The  $CL_s$  technique”, *J. Phys. G* **28** (2002) 2693. doi:10.1088/0954-3899/28/10/313.
- [82] ATLAS and CMS Collaborations, “Procedure for the LHC Higgs boson search combination in summer 2011”, ATL-PHYS-PUB-2011-011, CMS NOTE-2011/005, (2011).
- [83] T. Junk, “Confidence level computation for combining searches with small statistics”, *Nucl. Instrum. Meth. A* **434** (1999) 435, arXiv:hep-ex/9902006. doi:10.1016/S0168-9002(99)00498-2.
- [84] S. Frixione and B. R. Webber, “Matching NLO QCD computations and parton shower simulations”, *JHEP* **0206** (2002) 029, arXiv:hep-ph/0204244.
- [85] W. Beenakker, R. Höpker, M. Spira et al., “Squark and gluino production at hadron colliders”, *Nucl. Phys. B* **492** (1997) 51. doi:10.1016/S0550-3213(97)00084-9.
- [86] A. Kulesza and L. Motyka, “Threshold resummation for squark-antisquark and gluino-pair production at the LHC”, *Phys. Rev. Lett.* **102** (2009) 111802. doi:10.1103/PhysRevLett.102.111802.
- [87] A. Kulesza and L. Motyka, “Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC”, *Phys. Rev. D* **80** (2009) 095004. doi:10.1103/PhysRevD.80.095004.
- [88] W. Beenakker, S. Brensing, M. Krämer et al., “Soft-gluon resummation for squark and gluino hadroproduction”, *J. High Energy Phys.* **12** (2009) 041. doi:10.1088/1126-6708/2009/12/041.
- [89] W. Beenakker, S. Brensing, M. Krämer et al., “Squark and gluino hadroproduction”, *Int. J. Mod. Phys. A* **26** (2011) 2637. doi:10.1142/S0217751X11053560.
- [90] M. Kramer, A. Kulesza, R. van der Leeuw et al., “Supersymmetry production cross sections in  $pp$  collisions at  $\sqrt{s} = 7$  TeV”, arXiv:1206.2892.
- [91] B. S. Acharya, P. Grajek, G. L. Kane et al., “Identifying Multi-Top Events from Gluino Decay at the LHC”, (2009). arXiv:0901.3367.
- [92] G. L. Kane, E. Kuflik, R. Lu et al., “Top Channel for Early SUSY Discovery at the LHC”, *Phys. Rev. D* **84** (2011) 095004, arXiv:1101.1963. doi:10.1103/PhysRevD.84.095004.
- [93] R. Essig, E. Izaguirre, J. Kaplan et al., “Heavy Flavor Simplified Models at the LHC”, *JHEP* **01** (2012) 074, arXiv:1110.6443. doi:10.1007/JHEP01(2012)074.

- [94] M. Papucci, J. T. Ruderman, and A. Weiler, “Natural SUSY Endures”, *JHEP* **09** (2012) 035, arXiv:1110.6926. doi:10.1007/JHEP09(2012)035.
- [95] M. Kramer, A. Kulesza, R. van der Leeuw et al., “Supersymmetry production cross sections in  $pp$  collisions at  $\sqrt{s} = 7$  TeV”, arXiv:1206.2892.
- [96] CMS Collaboration, “Interpretation of searches for supersymmetry with simplified models”, (2013). arXiv:1301.2175. Accepted by *Phys. Rev. D*.
- [97] CMS Collaboration Collaboration, “Search for supersymmetry in final states with missing transverse energy and 0, 1, 2, or at least 3 b-quark jets in 7 TeV pp collisions using the variable  $\alpha_T$ ”, *JHEP* **1301** (2013) 077, arXiv:1210.8115. doi:10.1007/JHEP01(2013)077.
- [98] CMS Collaboration Collaboration, “Search for supersymmetry in hadronic final states with missing transverse energy using the variables  $\alpha_T$  and b-quark multiplicity in pp collisions at  $\sqrt{s} = 8$  TeV”, arXiv:1303.2985.
- [99] CMS Collaboration Collaboration, “Search for gluino mediated bottom- and top-squark production in multijet final states in pp collisions at 8 TeV”, arXiv:1305.2390.
- [100] ATLAS Collaboration, “Search for Scalar Bottom Quark Pair Production with the ATLAS Detector in pp Collisions at  $\sqrt{s} = 7$  TeV”, *Phys. Rev. Lett.* **108** (2012) 181802, arXiv:1112.3832. doi:10.1103/PhysRevLett.108.181802.
- [101] ATLAS Collaboration, “Search for Gluinos in Events with Two Same-Sign Leptons, Jets and Missing Transverse Momentum with the ATLAS Detector in pp Collisions at  $\sqrt{s} = 7$  TeV”, *Phys. Rev. Lett.* **108** (2012) 241802, arXiv:1203.5763. doi:10.1103/PhysRevLett.108.241802.
- [102] ATLAS Collaboration, “Search for supersymmetry in pp collisions at  $\sqrt{s} = 7$  TeV in final states with missing transverse momentum and b-jets with the ATLAS detector”, *Phys. Rev. D* **85** (2012) 112006, arXiv:1203.6193. doi:10.1103/PhysRevD.85.112006.
- [103] ATLAS Collaboration, “Search for top and bottom squarks from gluino pair production in final states with missing transverse energy and at least three b-jets with the ATLAS detector”, *Eur. Phys. J. C* **72** (2012) 2174, arXiv:1207.4686. doi:10.1140/epjc/s10052-012-2174-z.
- [104] ATLAS Collaboration Collaboration, “Multi-channel search for squarks and gluinos in  $\sqrt{s} = 7$  TeV  $pp$  collisions with the ATLAS detector”, *Eur.Phys.J.* **C73** (2013) 2362, arXiv:1212.6149. doi:10.1140/epjc/s10052-013-2362-5.

# Appendix A

## Table of Fake Rates

This appendix contains the tables of fakes rate values determined for electrons and muons from data.

**Table A.1:** Electron fake rate measured in bins of the electron candidate  $p_T$  and  $\eta$  for the high  $p_T$  analysis. The EWK correction has been applied and the uncertainties are statistical only.

$  \eta   \backslash p_T$	10.000 – 15.000	15.000 – 20.000	20.000 – 25.000	25.000 – 35.000	35.000 – 55.000
0.000 – 1.000	$0.251 \pm 0.009$	$0.171 \pm 0.007$	$0.149 \pm 0.007$	$0.151 \pm 0.011$	$0.149 \pm 0.029$
1.000 – 1.479	$0.271 \pm 0.012$	$0.199 \pm 0.010$	$0.194 \pm 0.009$	$0.160 \pm 0.014$	$0.120 \pm 0.028$
1.479 – 2.000	$0.245 \pm 0.011$	$0.184 \pm 0.008$	$0.166 \pm 0.007$	$0.145 \pm 0.008$	$0.128 \pm 0.022$
2.000 – 2.500	$0.239 \pm 0.014$	$0.172 \pm 0.010$	$0.152 \pm 0.008$	$0.142 \pm 0.008$	$0.134 \pm 0.022$

**Table A.2:** Electron fake rate measured in bins of the electron candidate  $p_T$  and  $\eta$  for the low  $p_T$  analyses. The EWK correction has been applied and the uncertainties are statistical only.

$  \eta   \backslash p_T$	10.000 – 15.000	15.000 – 20.000	20.000 – 25.000	25.000 – 35.000	35.000 – 55.000
0.000 – 1.000	$0.132 \pm 0.004$	$0.089 \pm 0.004$	$0.092 \pm 0.006$	$0.089 \pm 0.007$	$0.106 \pm 0.016$
1.000 – 1.479	$0.153 \pm 0.006$	$0.116 \pm 0.007$	$0.104 \pm 0.009$	$0.103 \pm 0.011$	$0.100 \pm 0.018$
1.479 – 2.000	$0.183 \pm 0.007$	$0.130 \pm 0.008$	$0.131 \pm 0.010$	$0.111 \pm 0.010$	$0.118 \pm 0.017$
2.000 – 2.500	$0.155 \pm 0.009$	$0.122 \pm 0.010$	$0.136 \pm 0.013$	$0.116 \pm 0.012$	$0.110 \pm 0.020$

**Table A.3:** Muon fake rate measured in bins of the muon candidate  $p_T$  and  $\eta$  for the high  $p_T$ , and low  $p_T$  analysis. The EWK correction has been applied and the uncertainties are statistical only.

$ \eta  \backslash p_T$	5.000 – 10.000	10.000 – 15.000	15.000 – 20.000	20.000 – 25.000	25.000 – 35.000
0.000 – 1.000	$0.167 \pm 0.013$	$0.128 \pm 0.018$	$0.082 \pm 0.012$	$0.080 \pm 0.009$	$0.095 \pm 0.008$
1.000 – 1.479	$0.196 \pm 0.017$	$0.134 \pm 0.020$	$0.110 \pm 0.012$	$0.093 \pm 0.011$	$0.114 \pm 0.008$
1.479 – 2.000	$0.206 \pm 0.016$	$0.161 \pm 0.021$	$0.138 \pm 0.013$	$0.139 \pm 0.011$	$0.144 \pm 0.008$
2.000 – 2.500	$0.217 \pm 0.020$	$0.169 \pm 0.025$	$0.135 \pm 0.020$	$0.105 \pm 0.020$	$0.120 \pm 0.028$

# **Appendix B**

## **Details of Simulated Samples**

In this appendix, the MC datasets used in this analysis are listed.

### **B.1 Rare MC Samples**

Table B.1 contains a list of Monte Carlo samples contributing to the SM background from rare processes that is taken from simulation. The cross section and equivalent luminosity of each sample is also provided.

### **B.2 Standard Model Background MC Samples**

Table B.2 contains a list of Monte Carlo samples contributing to the SM background from background processes. The cross section and equivalent luminosity of each sample is also provided.

### **B.3 MC Samples used for SM 4-Top Production and SS Top Pair Production**

Table B.3 contains a list of Monte Carlo samples contributing to the Same-Sign top pair and SM 4-top production. The cross section and equivalent luminosity of each sample is also provided.

## B.4 MC Samples used for Fake Rate Closure Tests

Table B.4 contains a list of Monte Carlo samples used for the fake rate closure tests described in Section 5.1.6. The cross section and equivalent luminosity of each sample is also provided.

**Table B.1:** MC datasets corresponding to contributions not covered by the data-driven methods. Predicted yields from the SM samples listed here are used directly in the analysis. The common part of each dataset name Summer12\_DR53X-PU\_S10\_START53\_V7A is replaced with a shorthand Su12. All datasets are in the AODSIM data tier.

Name	Cross section, pb	Luminosity, $\text{fb}^{-1}$
TTZJets_8TeV-madgraph/Su12-v1	0.206	1021.68
TTWWJets_8TeV-madgraph/Su12-v1	0.002037	106931.76
TTWJets_8TeV-madgraph/Su12-v1	0.232	845.03
TTGJets_8TeV-madgraph/Su12-v1	2.166	33.06
TBZToLL_4F_TuneZ2star_8TeV-madgraph/Su12-v1	0.0114	13026.67
ZZZNoGstarJets_8TeV-madgraph/Su12-v1	0.0055269	40692.61
WWWJets_8TeV-madgraph/Su12-v1	0.08217	2737.02
WWGJets_8TeV-madgraph/Su12-v1	0.528	407.43
WZZNoGstarJets_8TeV-madgraph/Su12-v1	0.01922	12946.70
WWZNoGstarJets_8TeV-madgraph/Su12-v1	0.05798	3832.94
ZZJetsTo4L_TuneZ2star_8TeV-madgraph-tauola/Su12-v1	0.1769	27177.36
WZJetsTo3LNu_TuneZ2_8TeV-madgraph-tauola/Su12-v1	1.0575	1908.25
WmWmqq_8TeV-madgraph/Su12-v1	0.08888	1084.52
WpWpqq_8TeV-madgraph/Su12-v1	0.2482	402.84
WW_DoubleScattering_8TeV-pythia8/Su12-v1	0.5879	1418.68
WGstarToLNu2E_TuneZ2star_8TeV-madgraph-tauola/Su12-v1	5.873	53.58
WGstarToLNu2Mu_TuneZ2star_7TeV-madgraph-tauola/Su12-v1	1.914	156.73
WGstarToLNu2Tau_TuneZ2star_7TeV-madgraph-tauola/Su12-v1	0.336	148.80
WH_ZH_TTH_HToWW_M-125_8TeV-pythia6/Su12-v1	0.2604	769.72
WH_ZH_TTH_HToZZ_M-125_8TeV-pythia6/Su12-v1	0.0320	15652.26
WH_ZH_TTH_HToTauTau_M-125_lepdecay_8TeV-pythia6-tauola/Su12-v1	0.0177	5478.02

**Table B.2:** MC datasets that do not contribute to MC Pred. The contribution to the background from these processes is covered by data-driven methods, but expected yields based on simulation are nevertheless provided as a reference. Predicted yields from the SM samples listed here are used directly in the analysis. The common part of each dataset name Summer12\_DR53X-PU\_S10\_START53\_V7A is replaced with a shorthand Su12. All datasets are in the AODSIM data tier.

Name	Cross section, pb	Luminosity, $\text{fb}^{-1}$
TTJets_MassiveBinDECAY_TuneZ2star_8TeV-madgraph/Su12-v1	225.21	29.59
TTJets_FullLeptMGDecays_8TeV-madgraph/Su12-v1	24.56	493.45
TTJets_SemiLeptMGDecays_8TeV-madgraph/Su12-v1	102.50	247.66
TTJets_HadronicMGDecays_8TeV-madgraph/Su12-v1	106.93	292.00
T_s-channel_Tune2star_8TeV-powheg-tauola/Su12-v1	3.89	68.59
Tbar_s-channel_Tune2star_8TeV-powheg-tauola/Su12-v1	1.76	79.53
T_t-channel_Tune2star_8TeV-powheg-tauola/Su12-v1	55.5	66.10
Tbar_t_channel_Tune2star_8TeV-powheg-tauola/Su12-v1	30.0	63.03
T_tW-channel-DR_Tune2star_8TeV-powheg-tauola/Su12-v1	11.18	44.83
Tbar_tW-channel-DR_Tune2star_8TeV-powheg-tauola/Su12-v1	11.18	44.45
DYJetsToLL_M-50_TuneZ2Star_8TeV-madgraph-tarball/Su12-v1	3532.8	8.62
WJetsToLNu_TuneZ2Star_8TeV-madgraph-tarball/Su12-v1	37509	inf
WWJetsTo2L2Nu_TuneZ2star_8TeV-madgraph-tauola/Su12-v1	5.81	332.61

**Table B.3:** MC datasets that were used for the same-sign top and 4-top signal acceptance and systematics used in Section 7.3.1. The common part of each dataset name Summer12\_DR53X-PU\_S10\_START53\_V7A is replaced with a shorthand Su12. All datasets are in the AODSIM data tier. The SMS scans cross-sections and luminosities varies on the particular point.

Name	Cross section, pb	Luminosity, $\text{fb}^{-1}$
TTJets_FullLeptMGDecays_8TeV-madgraph/Su12-v1	24.56	493.45
TTTT_TuneZ2star_8TeV-madgraph-tauola/Su12-v1	.000716	139656.43

**Table B.4:** MC datasets used for the fake rate closure test in Section 5.1. The common part of each dataset name Summer12\_DR53X-PU\_S10\_START53\_V7A is replaced with a shorthand Su12. All datasets are in the AODSIM data tier.

Name	Cross section, pb
TT_CT10_TuneZ2star_8TeV-powheg-tauola/Su12-v2	334
TT_CT10_TuneZ2star_8TeV-powheg-tauola/Su12-v1	334
TTJets_SemiLeptMGDecays_8TeV-madgraph/Su12-v2	102.50
WJetsToLNu_TuneZ2Star_8TeV-madgraph-tarball/Su12-v1	37509
WJetsToLNu_TuneZ2Star_8TeV-madgraph-tarball/Su12-v2	37509
W1JetsToLNu_TuneZ2Star_8TeV-madgraph/Su12-v1	6663
W2JetsToLNu_TuneZ2Star_8TeV-madgraph/Su12-v1	2159
W3JetsToLNu_TuneZ2Star_8TeV-madgraph/Su12-v1	640
W4JetsToLNu_TuneZ2Star_8TeV-madgraph/Su12-v1	264
QCD_Pt-5to15_TuneZ2star_8TeV_pythia6/Su12-v1	4.2639499e10
QCD_Pt-15to30_TuneZ2star_8TeV_pythia6/Su12-v2	9.8828742e8
QCD_Pt-30to50_TuneZ2star_8TeV_pythia6/Su12-v2	6.6285328e7
QCD_Pt-50to80_TuneZ2star_8TeV_pythia6/Su12-v2	8148778.0
QCD_Pt-80to120_TuneZ2star_8TeV_pythia6/Su12-v3	1033680.0
QCD_Pt-120to170_TuneZ2star_8TeV_pythia6/Su12-v3	156293.3
QCD_Pt-170to300_TuneZ2star_8TeV_pythia6/Su12-v2	34138.15
QCD_Pt-15to20_MuEnrichedPt5_TuneZ2star_8TeV_pythia6/Su12-v2	7.022e8
QCD_Pt-20to30_MuEnrichedPt5_TuneZ2star_8TeV_pythia6/Su12-v1	2.87e8
QCD_Pt-30to50_MuEnrichedPt5_TuneZ2star_8TeV_pythia6/Su12-v1	6.609e7
QCD_Pt-50to80_MuEnrichedPt5_TuneZ2star_8TeV_pythia6/Su12-v1	8082000.0
QCD_Pt-80to120_MuEnrichedPt5_TuneZ2star_8TeV_pythia6/Su12-v1	1024000.0
QCD_Pt-120to170_MuEnrichedPt5_TuneZ2star_8TeV_pythia6/Su12-v1	157800.0
QCD_Pt_20_MuEnrichedPt_15_TuneZ2star_8TeV_pythia6/Su12-v3	3.64e8