SEARCH FOR PAIR-PRODUCED HEAVY FOURTH-GENERATION BOTTOM-LIKE QUARKS DECAYING TO BZ AND TW IN 8 TEV PROTON-PROTON COLLISIONS WITH MULTILEPTON FINAL STATES

BY SANJAY R. ARORA

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Physics and Astronomy

Written under the direction of

Amitabh Lath

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

May, 2013

ABSTRACT OF THE DISSERTATION

Search for pair-produced heavy fourth-generation bottom-like quarks decaying to bZ and tW in 8 TeV Proton-Proton Collisions with multilepton final states

By SANJAY R. ARORA

Dissertation Director: Amitabh Lath

We present a search for anomalous production of events with three or more isolated leptons produced in pp collisions at $\sqrt{s} = 8\,\text{TeV}$ collected by the CMS experiment at the LHC. We analyze 9.2 fb^{-1} of data collected by the CMS experiment during the 2012 LHC run. We categorize observed multilepton events into exclusive search channels based on various quantities based on the identity and kinematics of the objects in the events. The search channels are ordered by the amount of expected Standard Model background. Explicit use of requirements such as missing transverse energy or total hadronic energy is avoided. We emphasize data-based estimation of the Standard Model backgrounds, but also use simulation to estimate some of the backgrounds when appropriate. We interpret search results in the context of a model involving the exotic bottom-like quark b' decaying to two different modes ($b' \rightarrow bZ$ and $b' \rightarrow tW$) with varying

branching ratios. We derive exclusion limits as a function of the b' mass as well as the branching ratios.

Acknowledgements

I would like to start my thanking my parents. They have supported and encouraged me through the years and given me the freedom to choose my own path. This would simply not have been possible without them.

Ashlee and Manjul have been best friends and siblings through the roller-coaster ride of a PhD. I would have given up a long time ago without their support. They also put up with obnoxious behavior that often accompanies the author (remember Jaipur?!). I look forward to adding countless hours of fun to the ones we have already had. Meekah and Sephora contributed immensely by keeping my survival instincts sharp.

The Geesey family: Mr. and Mrs. John and Donna Geesey, Grandpa Geesey, John, Ryan, Emi as well as Lucy have always been welcoming. They made New Jersey home and keep me sane whenever things get too hot at work. A week without them invariably wears me down and I hope to spend many more years bothering them.

On the physics front, my advisor Amit has been invaluable. He gave me a chance at a crucial stage in my graduate career. He has always been supportive of my wild ideas and even wilder work hours (I never did try getting wine at 8 AM). He not only taught me how an experimental physicist thinks but also how one can be interested in several things without compromising physics. I'll always be grateful to him for his support.

While not officially my advisor, Sunil has played that role too. Initially intimidating, especially when he was staring at a plot I made, I soon came to deeply value his insight into physics. Sunil also gradually takes a graduate student from

being just that to being a confident, professional scientist and for that I'll always be thankful.

Richard has been a role-model through the entire process. I have spent more time on the details of particle physics with him than anyone else. People tend to stop surprising you once you get to know them well but Richard still pulls out awe-inspiring tricks that leave me stunned. Hopefully, we'll get to work together again in the future.

Scott has been a constant mentor and friend. As for physics, I have learnt more physics from him per second than anyone else. At the risk of embarrassing him, I think he's easily the most outstanding physicist I know. His intuition as well as knowledge are at just another level and it's been an honor to work with him. I am also very thankful for all his support and entertaining conversations about space and technology.

The group would not be the same without all the postdocs, graduate and undergraduate students. Matt has always pushed us by doing everything more efficienctly and smartly and this has only made us better at what we do. Emmanuel, Patrick, Peter, and Shruti are fellow multileptoners and my days would be very boring without our sessions at the white-board (how much of that stuff have we inhaled by now!). Anthony, Christian, Claudia and Rishi are great company and always brighten things up when they come to Rutgers. David, Eric, Julia, Kelvin, Kin, and Mohamed, undergraduates only in name, have been colleagues and friends. Each of them has a very bright future given what we have seen from them and I wish them the best of luck. I would also like to thank Eva, Bob, John, JP, Pieter, and Steve for their support.

Rutgers has been wonderful over the past six years and both the university and the department have made me feel very welcome. I made great friends here and found a relaxing but intense physics environment. If I had to do it all again, I would be back in a heartbeat.

Last but definitely not the least, I thank all the engineers, technicians, post-docs and graduate students who keep CERN and the LHC running. Data analysis would be impossible without their efforts which often go unacknowledged. It's been a privilege to be part of this giant experiment especially at a time of discovery and hopefully we'll have many new discoveries in the coming years.

Dedication

For friends and family

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Introduction

The Standard Model (SM) of particle physics describes the known fundamental particles that make up matter and interactions between them. The particles fall into distinct classes called **quarks** (down, up, strange, charm, bottom and top), **leptons** (electrons, muons, taus, electron neutrinos, muon neutrinos, tau neutrinos), **gauge bosons** (photons, Z-boson, $W^{\pm}$ bosons, gluons) and lastly, the **Higgs boson**. The leptons and quarks are split into three generations. Each generation contains two quarks and two leptons. A possible Higgs candidate was discovered in the summer of 2012. All other particles in the standard model have been observed.

There are four fundamental interactions in nature. The simplest and most familiar one is **quantum electrodynamics**. It posits that charged particles interact with each other via electromagnetic potentials which are described by the exchange of photons. The next set called the **weak interactions** describe interactions between leptons (including the neutrinos) and quarks and is mediated by the exchange of W and Z bosons. The last component in the SM is the **strong interaction** which describes interactions between quarks mediated by the exchange of massless gluons. **Gravity** is not part of the SM and is negligible at high energies probed by today's accelerators.

A major research effort is directed towards possible modifications or extensions of the standard model. In general, every physics theory is applicable over some

energy (or distance) scale. When one probes interactions beyond that scale, one expects to see new phenomena or deviations from predictions of the theory. The Large Hadron Collider (LHC) described in the next section is designed to probe high energies where one might observe deviations from the SM. The analysis described in this thesis searches for a certain class of speculative models that would yield events with 3 or more leptons more frequently than the standard model alone. In particular, we look at the possibility of adding a fourth-generation of quarks with various decays to SM particles. The physics model is described in more detail later in this chapter.

An overview of the process involved in searching for new physics is as follows: Pick a physical quantity, say lepton p_T or missing transverse energy($E_{\mathrm{T}}^{\mathrm{miss}}$) and look at its distribution in a certain class of events (think of events with exactly one e^+e^- pair as an example). We can measure the observed 1-dimensional distribution in data of this quantity at the LHC. Now we want to test whether all the events in this distribution can be accounted for by the standard model. This process involves going over all possible standard model processes that can yield this class of events and adding up their respective contributions. Any discrepancy between the total standard model contribution and the measured distribution could potentially point to signs of new physics. In practice, all SM processes are fundamentally stochastic and are further smeared by detector effects. This demands a careful estimation of the uncertainties on the various standard model contributions. If no discrepancies are observed, this allows us to set bounds on underlying parameters of our physics model such as masses, branching ratios and cross-sections.

1.2   Model: Heavy Fourth-generation Exotic Quark b'

We look for exotic quarks in this analysis. In particular, we look for the heavier cousin of the bottom-quark denoted by b'. The SM has successfully described most physics phenomena probed by experiments but is still unsatisfactory in several ways. This has prompted speculations on how it could possibly be extended. Since the number of generations of quarks and leptons is not fixed by the theory, a popular and straighforward extension is the addition of a fourth generation (1; 2; 3; 4; 7; 6). While there are several constraints on the number of generations, set in particular by a limit on the number of light ($m_\nu << m_Z/2$) neutrinos, there could be additional heavier neutrinos, which would get around this obstacle (5).

b' quarks in this analysis are **exotic** in the sense that they can decay dominantly through flavor-changing neutral currents (FCNC). If the b' is heavier than the top quark t, then it could decay both to a top and a W, i.e. $b' \to tW$ and also to a b-quark and a Z, i.e. $b' \to bZ$ where the second would constitute a FCNC. Another possibility is $b' \to b+$Higgs which also constitutes a FCNC and could potentially dominate the $b' \to bZ$ mode (31). In this analysis, we only consider the first two decay modes. We look at the Higgs decay mode in an extension of this analysis which is described in 6.

We look for events with pair-produced b's where each b' can decay either to $b + Z$ or $t + W$ with varying branching ratios. This leads to three distinct event types:

- $bZbZ$

- $tWtW$

- $bZtW$

The top decays to $b + W$ so we end up getting the following decays:

- $(bb) + ZZ$

- $(bb) + WWWW$

- $(bb) + WWZ$

For the extension with the Higgs decay mode, we consider the Higgs mass to be 120 GeV. For a Higgs of this mass, the dominant decays modes are: bb (64.8%), WW (14.8%), gg (8.8%), $\tau\tau$ (7.1%), cc (3.0%), ZZ (1.6%). Since we are looking for 3 or more leptons, each event requires any H present to decay to leptons and so the dominant mode is $H \to WW$. In this case, the decay topology is exactly the same as $b' \to tW \to bWW$ since $b' \to bH \to bWW$. The second most-dominant mode is $H \to \tau\tau$ since we also look at channels with at most one hadronic-τ lepton. The last mode of interest is $H \to ZZ$ but it gets severely suppressed by the small branching ratio.

There are current bounds on masses of fourth-generation quarks. The width of the Z-boson depends on the number of generation. In particular, if there was an additional generation of quarks and leptons, the Z would decay to the fourth neutrino which would increase the Z-width. Current bounds from LEP indicate that this is possible only if the fourth-generation neutrino were heavy (with mass greater than $m_Z/2$). Additionally, heavy quarks lead to an enhancement of Higgs production through gluon-gluon fusion. The presence of a heavy fourth-generation quark would lead to an increase in the Higgs cross-section by a factor of around 9. This has not been observed and one could get around this by postulating a vector-like b' that doesn't get its mass from the Higgs (and hence doesn't couple to the Higgs).

1.3 The LHC and the CMS Detector
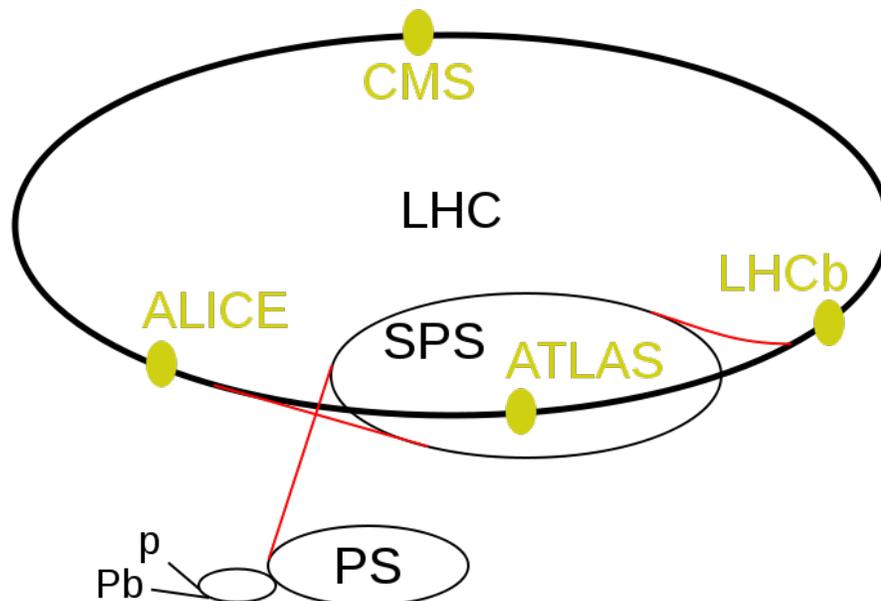
Figure 1.1: An overview of the LHC and the four detectors. Taken from (34)

The LHC (see figure 1.1) is the world's largest particle accelerator built into a tunnel 27 km in circumference straddling the French-Swiss border outside Geneva. It currently collides two beams of protons at a center-of-mass energy $(\sqrt{s})$ of 8 TeV. The beams are contained in separate beam pipes and brought together at four points along the LHC circumference. The four points contain sophisticated detectors called CMS, ATLAS, LHCb and ALICE. CMS and ATLAS are general-purpose detectors designed to search for signs of new physics. The work described in this thesis was done using the CMS detector which is described below.

CMS DETECTOR
Total weight   : 14,000 tonnes
Overall diameter : 15.0 m
Overall length   : 28.7 m
Magnetic field   : 3.8 T

STEEL RETURN YOKE
12,500 tonnes

SILICON TRACKERS
Pixel (100x150 μm) ~16m² ~66M channels
Microstrips (80x180 μm) ~200m² ~9.6M channels

SUPERCONDUCTING SOLENOID
Niobium titanium coil carrying ~18,000A

MUON CHAMBERS
Barrel: 250 Drift Tube, 480 Resistive Plate Chambers
Endcaps: 468 Cathode Strip, 432 Resistive Plate Chambers

PRESHOWER
Silicon strips ~16m² ~137,000 channels

FORWARD CALORIMETER
Steel + Quartz fibres ~2,000 Channels

CRYSTAL
ELECTROMAGNETIC
CALORIMETER (ECAL)
~76,000 scintillating PbWO₄ crystals

HADRON CALORIMETER (HCAL)
Brass + Plastic scintillator ~7,000 channels

Figure 1.2: An overview of the CMS detector. Taken from (35)

The CMS detector (see figure 1.2) has a cylindrical geometry. The coordinate system used is centered at the expected collision point at the center of the detector. The central axis of the cylinder is called the z-axis and is along the proton beam direction. The **tranverse** x-y plane has the y-axis pointing vertically up and the x-axis pointing towards the center of the LHC ring. Due to unknown initial parton momenta in the z-direction, all physics analysis is carried out in the transverse plane where momentum is conserved and known to be zero. Points are labeled by three coordinates: distance from the origin in the x-y plane (called p_T for momenta), angle ϕ with respect to the +x axis in the x-y plane. ϕ ranges from $-\pi$ to π and lastly, angle θ of the vector with respect to the z-axis. Another quantity called pseudo-rapidity η is used more often to describe θ and is defined as $\eta = -\ln\tan\theta/2$.

Figure 1.3 shows a transverse slice of the CMS detector. A tracker system consisting of silicon pixels and silicon strips forms the core of the detector. The

Figure 1.3: A transverse slice of the CMS detector showing the important components. Taken from (36)

tracker is surrounded by an electromagnetic calorimeter which in turn, is surrounded by a hadronic calorimeter. A magnetic field of 3.8 T permeates through

the tracker and the two calorimeters. Lastly, there are chambers for reconstructing muons. Details of detector components are taken from Dmitry Hits' Ph.D. thesis (37) and Frank Golf's Ph.D. thesis (38).

1.3.1 Tracking System

The tracker is responsible for reconstructing the tracks or paths of charged particles as they move out from the interaction point. It consists of layers of silicon pixels and silicon strips. As charged particles travel through the silicon components, they ionize and deposit energy in the material. This creates a small ionization current which is read by the electronics and translated into a trajectory for the particles. The trajectory is then used to measure the momentum and charge of the particle.

1.3.2 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) measures the energy of electrons and photons. Electrons undergo bremsstrahlung (i.e. radiate photons as they slow down in the material) and photons convert to e^+e^- pairs. As these processes occur, an incoming electron/photon develops into an electromagnetic shower and the resulting light is read out by electronics and converted into a measurement of energy.

The ECAL is made of lead-Tungstate crystals ($PbWO_4$) that is made up of high atomic number (high-Z) elements and completely surrounds the inner tracking system. The crystals are wedge-shaped and several radiation lengths deep so as to capture most of the particle energies. In addition, the ECAL is divided into two components: the barrel and the endcaps. The barrel covers the pseudo-rapidity range: $|\eta| < 1.479$ while the endcaps go on both ends of the barrel and cover the ranges: $1.479 < |\eta| < 3.0$. Each barrel crystal measures 22

mm x 22 mm on its front face and 26 mm x 26 mm on its back face with a depth of 230 mm or 26 radiation lengths. The endcap crystals measure 28.6 mm x 28.6 mm on the front face and 30 mm x 30 mm on the back face and are 220 mm or 25 radiation lengths deep. The ECAL is completely hermetic.

1.3.3 Hadronic Calorimeter

The hadronic calorimeter (HCAL) is designed to measure the energies of hadrons, both electrically charged and neutral. Hadrons are bound-states of colored (i.e. strongly interacting) particles. The HCal is made of alternating layers of brass plates and scintillating material. As the hadrons scatter through the brass material, they shower into additional charged and neutral hadrons. As these hadrons pass through the scintillating material, they radiate light which is then converted by electronics into measurements of energy. The hadron showers then pass subsequent brass and scintillator layers till they lose all their energy.

The HCAL hermetically surrounds the ECal. In the high pseudo-rapidity region, $3 < |\eta| < 5$ there is a forward hadron calorimeter (HF) which acts both as a hadronic and electromagnetic calorimeter. Since the HF is close to the beam-pipe and suffers from very high radiation levels, it is constructed from steel with embedded quartz fibers. Particles emit light due to Cerenkov radiation as light travels faster through quartz than the speed of light in that medium. The fibers come in two varieties: short and long. The long fiber is closer to the surface of the HF components and is used to reconstruct the energy of electromagnetic particles since electromagnetically-interacting particles dissipate their energy quicker and don't travel deep into the components. The short fibers are used to reconstruct energy of hadrons.

1.3.4 Muon System

Muons are charged particles but owing to their much heavier mass compared to electrons, they lose far less energy by radiating. So they don't deposit much energy in the calorimeters and instead pass through these components. The muon system is designed to detect muons and filter out any charged hadrons that get through the HCal. It consists of three components: drift tubes (DTs), cathode strip chambers (CSCs) and resistive plate chambers (RPCs). In addition, there's a magnetic field of 2T opposite to the magnetic field in the tracker. This creates distinctive S-shaped tracks for muons as they travel outwards from the interaction point.

1.4 Outline of the Thesis

This section gives an overview of the thesis.

Chapter 2 **Analysis Strategy and Selection Criteria** starts by describing the signal (b') model we look at. It then describes the triggers used, the exact definitions of leptons(electrons, muons, taus), jets, $E_{\mathrm{T}}^{\mathrm{miss}}$, and b-tagging recommended by CMS and used in the analysis. The general philosophy behind searching in multiple exclusive channels is described followed by a short section listing the signal cross-sections and the effect of higher b' masses on acceptances. Lastly, all the data sets and Monte-Carlo samples used are listed.

Chapter 3 **Background Estimation Techniques and Controls** describes how standard model backgrounds are estimated for all the multilepton channels. We start by describing Monte-Carlo backgrounds and show control region plots to verify that they behave as expected. This is followed by data-driven methods, i.e. using parameters measured in one dataset to estimate the background contributions in another dataset. We use data-driven methods to estimate the number of electrons and muons coming from b-jets, taus being faked by jets as

well as photons asymmetrically converting to one lepton (the other lepton is not detected because of very low momenta). Lastly, we give an example of how these data-driven methods are applied to estimate backgrounds.

Chapter 4 **Trigger Efficiencies, Selection Efficiencies and Other Corrections to Monte-Carlo Samples** describes all the corrections applied to the various Monte-Carlo samples used for estimating standard model backgrounds. Even though Monte-Carlo samples have most of the ingredients to describe certain physics processes correctly, there are still residual differences in the physics environments between Monte-Carlo and data. These include differences in trigger efficiencies, object (leptons, b-jets) selection efficiencies, as well as distributions like pileup, number of jets (N_{jet}) and $E_{\mathrm{T}}^{\mathrm{miss}}$.

Chapter 5 **Results: Statistical Techniques, Sources of Systematics and Exclusions** starts by describing the statistical techniques used to reject signal models as well as methods used to extract an upper-limit on the cross-section of multilepton processes that could come from the signal. We then describe the various sources of systematic uncertainties for backgrounds in different channels. Lastly, we describe the parameter space (the masses and branching ratios) excluded by this analysis at the 95% confidence level.

Chapter 6 **Conclusions** summarizes the results and exclusions as well as further extensions of the analysis to include the decay $b' \rightarrow bH$.

Appendix A $E_{\mathrm{T}}^{\mathbf{miss}}$ **Resolution Dependence on Pileup and Jet Activity** describes a method to characterize the resolution/width of $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in data and Monte-Carlo samples and to match them. The method parameterizes the resolution as a function of pileup (number of simultaneous collisions in an event) and total jet activity in events. It also describes a procedure to estimate systematic uncertainties due to resolution effects in different search channels.

Appendix B shows the number of observed events, background breakdowns and overlaid signal in various channels.

Appendix C lists all the triggers used for the data in this analysis.

Most of the material in this thesis is based on the following documents listed in the bibliography (30), (33), (28) and (32) which are internal CMS documents where this work was first described.

Chapter 2

Analysis Strategy and Selection Criteria

This chapter gives an overview of the analysis. It starts by introducing the signal model and the **final-states** (events with 3 or more leptons). The **triggers** applied to the data are discussed. This is followed by listing the **definitions of the objects** (electrons, muons, taus, jets etc.) used in this analysis. Lastly, **event-level cuts** and the **general strategy** used to search for new signs of physics is described. Details of background estimation are described in chapter 3 and efficiencies and other Monte-Carlo corrections are described in chapter 4. We also list the datasets and Monte-Carlo samples used in this analysis.

2.1 Introduction

As mentioned before, we look for exotic fourth-generation bottom-like quarks called b's in this analysis. It is assumed that b's are pair-produced in each event and can decay to either $b + Z$ or $t + W$. This leads to three distinct event types:

- $bZbZ$

- $tWtW$

- $bZtW$

The top decays to $b + W$ which leads to the following event topologies:

- $(bb) + ZZ$

- $(bb) + WWWW$

- $(bb) + WWZ$

Z-bosons can decay to lepton pairs (e^+e^-, $\mu^+\mu^-$, $\tau^+\tau^-$) with about 3% branching ratio each while W-bosons can decay to a lepton and a neutrino ($e\nu_e$, $\mu\nu_\mu$, $\tau\nu_\tau$) with about 10% branching ratio each.

Using a $9.2fb^{-1}$ data sample of $\sqrt{s} = 8\,\text{TeV}$ pp collisions at the LHC collected by the CMS experiment, we carry out a search for pair-produced b' quarks in events with at least three leptons. The leptons we examine, in the order of increasing difficulty of reconstruction, are muons, electrons, and the taus. The SM favors the production of particle jets over that of leptons by a wide margin. While the production of three or more leptons in SM is thus rare, physics beyond SM has proven to be so elusive that we must augment such a multilepton signature with other kinematic properties of the event that are characteristic of new physics, and are therefore effective in suppressing the rarest of backgrounds from SM processes.

The analyses carried out at CMS in 2011 at 7-TeV set limits on two scenarios where b's are pair-produced: the first one requires both b's to decay to tWs (9) and the second one requires both b's to decay to bZs (8).

In this analysis (10) carried out at CMS at 8-TeV, we bin the data in multiple exclusive channels depending on lepton and jet flavors as well as kinematical cuts. This gives us sensitivity to not just $b'b' \to tWtW$ and $b'b' \to bZbZ$ but also to intermediate scenarios where the branching ratio to bZ is not 0 or 1. This lets us adiabatically continue from the $tWtW$ end to the $bZbZ$ end and we present limits in the two-dimensional plane of the branching ratio $BF(b' \to bZ)$ vs the b' mass.

2.2   Event Triggers

Due to the large volumes of data (about 400 million pp collisions per second and about 20 million events per second) at the LHC, the number of events has to be filtered to more manageable quantities (about 100 per second) that can be written to disk and saved for more detailed analysis. For this purpose, filters called **triggers** are used to look for events satisfying certain criteria. In a multilepton analysis like this one, we look at events that have at least two leptons: either two electrons (**double-electron triggers**), two muons (**double-muon triggers**), or an electron and a muon (**electron-muon or MuEG triggers**). In addition, for some cross-checks, we look at events where the sum of all jet p_Ts is greater than some threshold (H_T **triggers**). Here we just list some important points about the triggers. More details can be found in 4.

The dilepton triggers used in this analysis are listed in Appendix C. We use the OR of several triggers i.e. we require that each event fire at least one of several dilepton triggers. While we include hadronic taus in our final-states, we don't use hadronic tau triggers.

The efficiencies of the various triggers are obtained from a comparison with independent triggers, in our case the the jet energy triggers. The trigger efficiencies are measured directly in data using an independently (H_T) triggered data sample assuming no correlations between these and the signal triggers. In such samples one searches for events containing the number of tight leptons respecting the trigger threshold and determines for which fraction the signal trigger fired. This fraction represents the trigger efficiency.

The two most important double-electron, double-muon and electron-muon trigger have efficiencies of 96±3%, 88±3% and 93±4%, respectively. The uncertainty in the correction to the simulation translates into a systematic uncertainty in the irreducible backgrounds and signal efficiencies. We also use single-lepton

triggers for tag-and-probe studies and for the one-lepton $t\bar{t}$ control region.

2.3   Object Identification

Due to the complicated nature of the CMS detector and the high amount of activity in hadron collisions, it's non-trivial to identify various objects in the detector as leptons, photons, jets etc. There are dedicated groups within CMS that study algorithms to identify these objects and suggest cuts to analysts. In this section, we list the algorithms and their requirements for the various objects used in this analysis.

Leptons in this search can be either electrons, muons, or tau leptons. Electrons and muons with $p_\mathrm{T} \geq 10$ GeV and $|\eta| < 2.4$ are reconstructed from the particle-flow (PF) algorithm(24) which uses measured quantities from the tracker, calorimeter, and muon system. The matching candidate tracks must satisfy quality requirements and spatially match with the energy deposits in the ECAL and the tracks in the muon detectors, as appropriate. Jets are reconstructed using particles identified via particle flow with $|\eta| \leq 2.5$.

Taus can decay either leptonically (τ_ℓ) to electrons or muons, or hadronically (τ_h). The hadronic decays yield either a single charged track (one-prong) or three charged tracks (three-prong) with or without additional electromagnetic energy from neutral pion decays. The hadronic τ_h are reconstructed using the HPS algorithm which reconstructs the various hadronic decay modes and rejects candidates that appear to be poorly reconstructed electrons and muons. We require the visible p_T of the τ to be greater than 20 GeV and $|\eta| < 2.3$.

Sources of background leptons include genuine leptons occurring inside or near jets, hadrons faking leptons by punch-through into the muon system, hadronic showers with large electromagnetic fractions, or photon conversions. An isolation requirement strongly reduces the background from misidentified leptons, since

most of them occur inside jets. We define the relative isolation I_{rel} as the ratio of the sum of p_{T} of other particle flow candidates in the cone defined by $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} < 0.3$ around the lepton to the p_{T} of the lepton. For electrons, muons, and isolated tracks, we require $I_{\mathrm{rel}} < 0.15$. The sum of energy in the isolation cone is corrected by subtracting out the expected contributions from additional vertices in the event. For the isolation of the hadronic tau decays we require that the sum in a cone of $\Delta R < 0.5$ is less than 2 GeV after excluding the expected contribution from additional vertices.

Leptons from decays considered in this search originate from the collision point ("prompt" leptons). After the isolation selection, the most significant background sources are residual non-prompt leptons from heavy quark decays, where the lepton tends to be more isolated because of the high p_{T} with respect to the jet axis. This background is reduced by requiring that the leptons originate from within one centimeter of the primary vertex in z and that the impact parameter d_{xy} between the track and the event vertex in the plane transverse to the beam axis be small: $d_{xy} \leq 0.02$ cm. The isolation and promptness criteria would retain the signal but largely eliminate misidentified leptons.

The precise selection criteria for the different types of objects are (28):

Muons

See table 2.1.

Electrons

See table 2.2. Furthermore,

- Veto transition region: reject electrons with |electron− >superCluster()->eta()| in range 1.4442–1.566

- Electron should not be within ΔR of 0.1 of selected muon

Cut	Value
$\|\eta\|$	< 2.4
Global muon	1
PF muon	1
χ^2/d.o.f.	< 10
$\|dz\|$ from vertex	$< 0.5 cm$
$\|d0\|$ from vertex	$< 0.02 cm$
Number of valid pixel hots	> 0
Number of tracker LayersWM*	>5
Number of valid hits in muon chamber	>0
Relative isolation within $\Delta R < 0.3$, with beta corrections for PU	< 0.15

Table 2.1: Selection criteria for muons

Cut	For Barrel	For endcap
$\|\eta\|$	<2.4	<2.4
dEtaIn	<0.007	< 0.009
dPhiIn	<0.15	< 0.10
Sigmaietaieta	<0.01	< 0.03
H/E	<0.12	< 0.10
d0(vtx)	<0.02	< 0.02
dZ(vtx)	<0.1	< 0.2
$\|1/E - 1/p\|$	<0.05	< 0.05
Relative PF isolation	<0.15	< 0.15
conversion rejection cut	0	0
Number of expected inner hits	< 2	< 2

Table 2.2: Selection criteria for barrel and endcap electrons

- Effective Area corrections for electrons

    - Isolation is calculated within $\Delta R < 0.3$

    - Effective areas are taken from official CMS numbers

  d0/dz calculated w.r.t. first *good* vertex

Taus

- HPS Taus

- $p_T > 20\,\text{GeV}$, $|\eta| < 2.3$

- ByDecayModeFinding

- AgainstElectronMVA

- AgainstMuonTight

- ByLooseCombinedIsolationDBSumPtCorr

- $\Delta R > 0.1$ with selected leptons

Photons

The cone size for all isolation sums is 0.3

Cut	For Barrel	For endcap
Conversion safe electron veto	1	1
Single tower H/E	< 0.06	< 0.05
Sigmaietaieta	< 0.011	< 0.034
Rho corrected relative PF charged hadron isolation	< 0.06	< 0.05
Rho corrected relative PF neutral hadron isolation	< 0.16	< 0.10
Rho corrected PF photon isolation	< 0.08	< 0.12

Table 2.3: Selection criteria for barrel and endcap electrons

Jets

- pfjets, apply L1FastL2L3 corrections to MC, L1FastL2L3residual corrections to data

- $p_T > 30\,\text{GeV}$, $|\eta| < 2.5$

- Neutral hadron fraction of total jet energy < 0.99

- Neutral EM fraction of total jet energy < 0.99

- Number of constituents in jet > 1

- If $|\eta| < 2.4$ additional cuts are-

  - Charged hadron fraction >0

  - Number of tracks > 0

  - Charged EM fraction < 0.99

- Apply energy corrections corresponding to JetMET correction for 53X software release

B-tagging

- CSV, medium working point

MET

- "Out-of-the-box" pfmet

- Use official list of filters

  - CSC tight beam halo filter

  - HBHE noise filter with isolated noise rejection

  - Primary vertex filter

  – ECAL dead cell trigger primitive (TP) filter

  – Tracking failure filter

  – Bad EE Supercrystal filter

2.4 Search Strategy

This section describes the general philosophy behind the analysis. The basic idea involves using exclusive channels defined by various quantities like lepton flavors, charges, number of b-jets, as well as kinematical quantities like S_T. We also describe background reduction techniques to minimize the amount of standard model contributions to channels containing the signal events.

2.4.1 Multilepton channels

Candidate events in this search must have at least three leptons, where at most one of them is an hadronic τ. The thresholds on the transverse momenta of the leptons are chosen such that triggers used are maximally efficient on these events. The leading muon (electron) is required to have $p_T > 20$ GeV and the next to leading muon (electron) is required to have $p_T > 10$ GeV.

We classify multilepton events into search channels on the basis of the number of leptons, lepton flavor, and relative charges as well as charge and flavor combinations and other kinematic quantities described below. Since the hadronic tau decays bring in additional background, search channels including these are kept separate from pure electron and muon channels.

We classify each event in terms of the maximum number of opposite-sign and same-flavor (OSSF) dilepton pairs that can be made by using each lepton only once. For example, both $\mu^+\mu^-\mu^-$ and $\mu^+\mu^-e^-$ are OSSF1, $\mu^+\mu^+e^-$ is OSSF0, and $\mu^+\mu^-e^+e^-$ is OSSF2. We denote a light lepton pair of different flavors as $\ell\ell'$. In this context, leptons are electron or muon.

The level of SM background varies considerably across the channels. Channels with hadronic tau decays or containing OSSF pairs suffer from larger backgrounds than do channels with OSSF0. Hence all these charge combinations are considered as different channels.

2.4.2 Background reduction

The main SM backgrounds in multilepton plus jet events originate from Z+jets, double vector boson production (VV+jets), $t\bar{t}$ production, rare three-body processes and QCD. Leptons associated with jets can originate from heavy quark decays, or with a lower probability, from misidentified hadrons. Leptons from heavy quark decays can be suppressed by the isolation and vertex requirements. The probability that a QCD event includes three misidentified leptons is negligible. Backgrounds from cosmic rays are also found to be negligible. Backgrounds from beam-halo muons are included in the background estimate discussed below.

For multilepton searches, the SM background is small compared to fully hadronic searches and allows multileptons to probe regions of parameter space inaccessible by other searches. However, it is still important to separate multilepton signals from background. We separate signal from background by dividing the data in the following parameters: lepton flavors (light leptons ℓ, or τ), number of OSSF lepton pairs both consistent and inconsistent with Z (75–105 GeV), S_T (scalar sum of $E_{\mathrm{T}}^{\mathrm{miss}}$, H_T and lepton p_T in bins of 300–500 GeV), and number of b-jets. Using a large number of exclusive bins in these quantities increases the sensitivity to different b' decay modes. In particular, as the branching ratio for $b' \to bZ$ increases, the 3-lepton channels lose sensitivity and the signal is picked up mostly by 4-lepton channels.

The presence of hadronic activity (used for S_T calculation) in an event is characterized by the variable H_{T}, defined as the scalar sum of the transverse jet energies for all jets with $E_{\mathrm{T}} > 30\ GeV$. Jets used for the H_{T} determination must

be well separated from any identified leptons; jets are required to have no selected leptons in a cone $\Delta R < 0.3$ around the jet axis. For jet b-tagging, we use the Combined Secondary Vertex Method, medium working point.

The missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, is defined as the magnitude of the vectorial sum of the momenta of all Particle Flow (PF) candidates. Comparison between data and simulation shows good modeling of $E_{\mathrm{T}}^{\mathrm{miss}}$ (27) and is valid for our particular selection and data collecting period as well, as will be shown later.

For dilepton masses we consider all possible OSSF dilepton pairs when deciding whether or not to declare an event Z-like or J/Ψ like. Even though $\mu^+\mu^-\mu^\pm$ is considered OSSF1, we reuse the opposite-lepton in the two possible $\ell^+\ell^-$ pairings.

- $\ell^+\ell^-\ell'^\pm$: $\ell'^\pm$ is selected to be the W-like candidate.

- $\ell^+\ell^-\ell^\pm$: The W is one of the two same-sign leptons. If there is at least one $M(\ell^+\ell^-)$ candidate in the range 75 to 105 GeV, pick ℓ where the remaining two light leptons make the best Z. If none of the OSSF pairs make a Z, then pick the SS ℓ with smallest M_T.

- $\ell^+\ell^-\tau^\pm$: The τ is selected as the W-like lepton.

- no OSSF pair: From the same-sign leptons (including τ), pick the one that makes the smallest M_T.

In this search we reject events that have a $M(\ell^+\ell^-)$ pair below 12 GeV in order to reject low mass Drell Yan and low mass resonances like J/ψ(1S) and Υ. In order to remove leptons from conversions (internal and external) that arise from final state radiation from the Z daughters, we reject events in low S_T bins that do not have an $M(\ell^+\ell^-)$ Z candidate but do have a three body Z candidate $M(\ell^+\ell^-\ell'^\pm)$ or $M(\ell^+\ell^-\ell^\pm)$.

2.4.3 S_T Calculation

For discriminating between signal and background we use the variable S_T, where S_T is the scalar sum of E_T^{miss}, H_T, and L_T. E_T^{miss} is the missing transverse energy calculated by the particle flow algorithm. L_T is the $\sum p_T$ of all selected muons, electrons, and the visible p_T of τ candidates.

The variable S_T has the property that the distribution tends to peak at the invariant mass of the interaction that initiated the event, less the mass of invisible decay products. The distribution is wide ($\sim$100 GeV) but new physics created via massive particles can be very well separated from standard model processes, especially if backgrounds are already reduced by requiring ≥ 3 lepton candidates. For example, backgrounds from $Z + \text{Jets}$ have a S_T peak around 100 GeV, while $t\bar{t}$ tends to peak around 2×175 GeV $= 350$ GeV. For new physics pair producing 300 GeV particles, with a small invariant mass in invisible particles, the S_T distribution would peak around 600 GeV.

2.5 Signal

As mentioned in the introduction to this chapter, we look at a signal model consisting of fourth-generation quarks called b's and look at the decays to $b' \to bZ$ and $b' \to tW$. This section describes some details about the generation of signal Monte-Carlo samples.

We use official pair-produced b' samples. The LHE files are made using Mad-Graph 5.1.4.4 and then decayed and showered through Pythia 6 in CMSSW532patch4. Since we are interested in two decay modes, namely $b' \to bZ$ and $b' \to tW$, we can have one of three possibilities for any given event: $b'b' \to bZbZ$, $b'b' \to tWbZ$, and $b'b' \to tWtW$. We generate each of these 3 samples for b' masses ranging from 500 GeVto 900 GeVin steps of 50 GeV.

We use HATHOR (13) to calculate next-to-next-leading (NNLO) cross-sections

for our signal model. Table 2.4 shows the leading-order (LO) as well as NNLO cross-sections at both 7 TeV and 8 TeV. Figure 2.1 shows a comparison of the LO cross-section from Pythia, the LO cross-section from HATHOR and the NNLO cross-section from HATHOR (all at 8 TeV) vs the b' mass. The cross-sections decrease roughly by a factor of 2 for every 50 GeV increase in the b' mass.
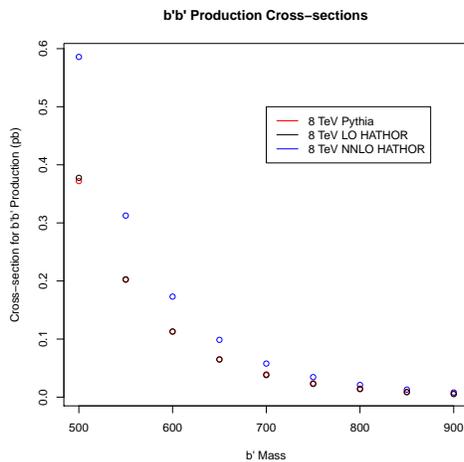
Figure 2.1: Comparison between the LO cross-section from Pythia, the LO cross-section from HATHOR and the NNLO cross-section from HATHOR at 8 TeV.

b' Mass	7 TeV LO	7 TeV NNLO	7TeV K-factor	8 TeV LO	8TeV NNLO	8TeV K-factor
500	0.2176	0.3372	1.5495	0.3776	0.5858	1.5513
550	0.1133	0.1746	1.5411	0.2029	0.3126	1.5406
600	0.0610	0.0937	1.5355	0.1129	0.1731	1.5323
650	0.0338	0.0518	1.5324	0.0647	0.0987	1.5263
700	0.0191	0.0293	1.5314	0.0379	0.0578	1.5223
750	0.0110	0.0169	1.5324	0.0227	0.0345	1.5202
800	0.0065	0.0099	1.5351	0.0138	0.0210	1.5195
850	0.0038	0.0059	1.5393	0.0085	0.0130	1.5204
900	0.0023	0.0036	1.5450	0.0053	0.0081	1.5225

Table 2.4: $b'b'$ production cross-sections from HATHOR. All numbers are in picobarns(pb).

2.5.1 Acceptances vs Mass

Since the b' masses are large enough, the decay products of the b's might be highly boosted. If there is significant jet activity, this might cause leptons and jets to merge and reduce acceptances for the signal. We show that at least for masses in the range 500-900 GeV, this is not the case.

We also account for this by assigning a systematic on the tag-and-probe efficiencies by looking at Z+Jets data and MC in bins of varying number of jets. Also, since we are looking for multilepton events (3 and 4 leptons), there are less vector bosons that can decay to jets.

Figures 2.2 shows the acceptances*branching ratios for 3 and 4 leptons channels for $b'b' \to bZbZ$, and $b'b' \to tWtW$. The behavior is quite flat within statistical uncertainties as the b' mass is increased indicating that merging of objects from boosted b's it not a problem for these masses.
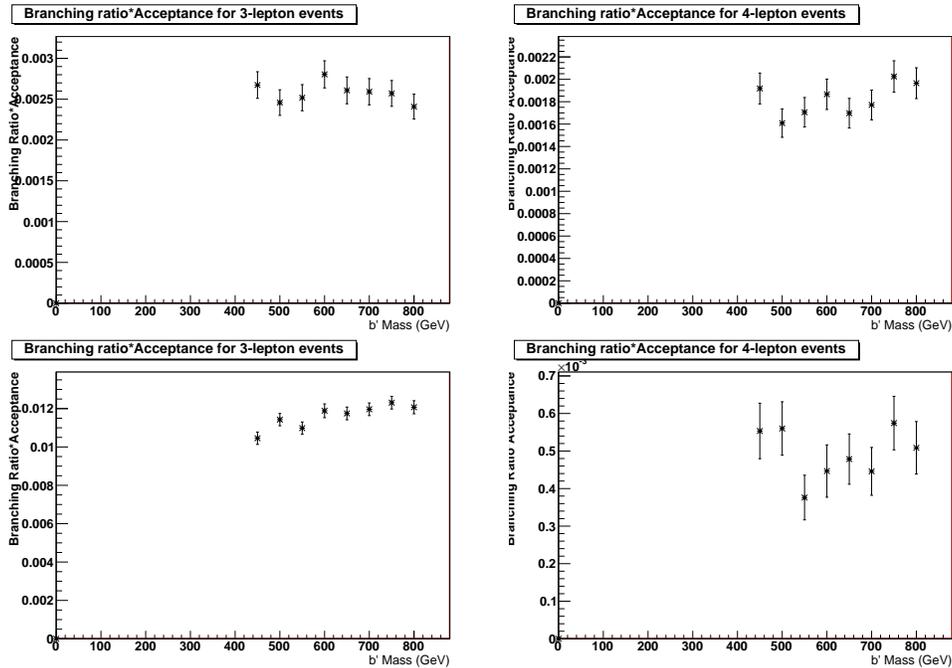
Figure 2.2: Branching Ratio*Acceptances for 3-lepton (left) and 4-lepton (right) channels vs b' mass for $b'b' \to bZbZ$ (top), $b'b' \to tWtW$ (bottom)

2.6 Datasets and Monte-Carlo Samples

Tables 2.5 and 2.6 list the datasets and simulations samples that were used in the present analysis.

Primary Dataset	Reco details	Luminosity (fb^{-1})
MuEG	Run2012A-13Jul2012-v1	0.809
MuEG	Run2012A-recover-06Aug2012-v1	0.082
MuEG	Run2012B-13Jul2012-v1	4.403
MuEG	Run2012C-PromptReco-v2	4.244
DoubleMuon	Run2012A-13Jul2012-v1	0.809
DoubleMuon	Run2012A-recover-06Aug2012-v1	0.082
DoubleMuon	Run2012B-13Jul2012-v1	4.403
DoubleMuon	Run2012C-PromptReco-v2	4.218
DoubleElectron	Run2012A-13Jul2012-v1	0.809
DoubleElectron	Run2012A-recover-06Aug2012-v1	0.082
DoubleElectron	Run2012B-13Jul2012-v1	4.431
DoubleElectron	Run2012C-PromptReco-v2	4.243

Table 2.5: Data samples

Simulation sample	N events	cross section (pb)
/DYJetsToLL_M-10To50filter_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	7,131,530	11050.0
/DYJetsToLL_M-50_TuneZ2Star_8TeV-madgraph-tarball/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	30,459,503	3532.8
/TTJets_MassiveBinDECAY_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v3/AODSIM	6,915,118	225
/TTGJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	71,598	2.166
/TTWJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	196,046	0.232
/TTZJets_8TeV-madgraph_v2/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	209,677	0.208
/TTWWJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	217,213	0.002
/ZZZNoGstarJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	224,902	0.0192
/WWWJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	220,170	0.08217
/ZZJetsTo4L_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	4,804,781	0.1769
/WZJetsTo3LNu_TuneZ2_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	2,016,678	1.0575
/WWJetsTo2L2Nu_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	1,932,249	5.8123
/WJetsToLNu_TuneZ2Star_8TeV-madgraph-tarball/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	18,393,090	37509
/WWGJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	215,121	1.44
/WWZNoGstarJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1/AODSIM	222,234	0.0633

Table 2.6: Simulation samples

Chapter 3

Background Estimation Techniques and Controls

The central part of any analysis is estimating standard model background contributions to the channels of interest. As an example, if one was looking for events with exactly two opposite-sign electrons, the dominant contribution from the standard model would be $Z \to e^+e^-$ and any discrepancy between the observed number of events and the $Z \to e^+e^-$ contributions would caused by:

- some other standard model background that didn't get accounted for (say $t\bar{t}$), or

- a statistical fluctuation, or

- a source outside the standard model i.e. a new phenomenon.

In order to have a robust search for new physics, it is important to ensure that all known standard model backgrounds are accounted for accurately.

In this analysis, we look for events with 3 or more leptons. The standard model backgrounds for such processes are small but nonetheless need to be accounted for. We account for backgrounds in two distinct ways. The first is to use **Monte-Carlo** (MC) samples where all the correct relativistic kinematics, the physics of the interactions as well as detector effects are accounted for. In principle, MC samples could be generated for all standard model processes and used for all the backgrounds. In practice though, there are short-comings in MC.

These could involve approximations made while programming, inaccurate modeling of the detector, or lack of enough detailed physics knowledge about the interactions (for example, parton distribution functions). Another problem could be in modeling the tails of distributions which is often where one first searches for new physics. In these cases, the MC sample sizes would have to be large enough to ensure that enough events get generated in the tails (otherwise, the statistical uncertainties in the tails would be large).

So, for certain backgrounds, instead of using MC, we estimate the contributions from data itself. Certain classes of observed events are scaled by various factors (also measured in data) to get background contributions from specific standard model processes. These background contributions are known as **data-driven backgrounds**. We use data-driven techniques to get background contributions from electrons/muons coming from jets, jets that pass cuts for taus as well as photons asymmetrically converting to leptons. These are described in more detail below.

3.1   Monte-Carlo Backgrounds

For events with 3 or more leptons, the major sources of backgrounds are $t\bar{t}$ (with one fake lepton coming from a jet), WZ, and ZZ. There are smaller contributions from $t\bar{t} + W$ and $t\bar{t} + Z$. All these backgrounds are accounted for by using MC samples which go through a detailed detector simulator to account for noise effects.

For every major MC source, it is important to do sanity checks to make sure various distributions are modeled correctly. These sanity checks are used to assign systematic uncertainties to these sources. In addition, there are certain corrections applied to each sample which are described in more detail in Chapter 4. In this section, we show plots that demonstrate that the MC samples are

accurately modeling backrounds.

3.1.1 Controls on $t\bar{t}$

As mentioned above, we use MC for $t\bar{t}$ predictions. The reason for this is that $t\bar{t}$ is a unique background in that it is fairly well defined. Unlike fakes from Z+Jets where we cannot trust that the MC reproduces the jet composition correctly, jets in $t\bar{t}$ have a well defined composition and spectra. The kinematics of this source of background are also fairly well known. Furthermore, the fake leptons are primarily from the semileptonic decays of heavy mesons, which have been extensively studied at the b-factories (CLEO, BaBar, BELL) and should be accurately simulated in the MC.

However, it is important to confirm that the MC is accurately predicting the $t\bar{t}$ background. To do this we define control distributions that are both dominated by $t\bar{t}$ in the data and are related to the number of fake leptons that should be produced by $t\bar{t}$. We look at distributions for two data sets. The first is a single-lepton data set with one isolated muon with $p_t > 30\,\text{GeV}$, three or more jets with $p_T > 40\,\text{GeV}$, at least one b-tag, and $S_T > 200\,\text{GeV}$. The second is a dilepton data set with an isolated muon and an isolated electron where the muon and electron have opposite charges. Both of these data sets are dominated by $t\bar{t}$ for large S_T.

The S_T distributions of the two dilepton control region is shown in figure 3.1. Figure 3.1 shows that the overall number of $t\bar{t}$ is consistent with the MC prediction. Figures 3.2 and 3.3 show the comparison between the E_T^{miss} and H_T distributions respectively. We get an overall scale factor from the S_T distribution for the $t\bar{t}$ cross-section.

The above plots show that the shapes of kinematic variables is being modeled correctly by the $t\bar{t}$ samples. $t\bar{t}$ gives only two leptons though. So any background
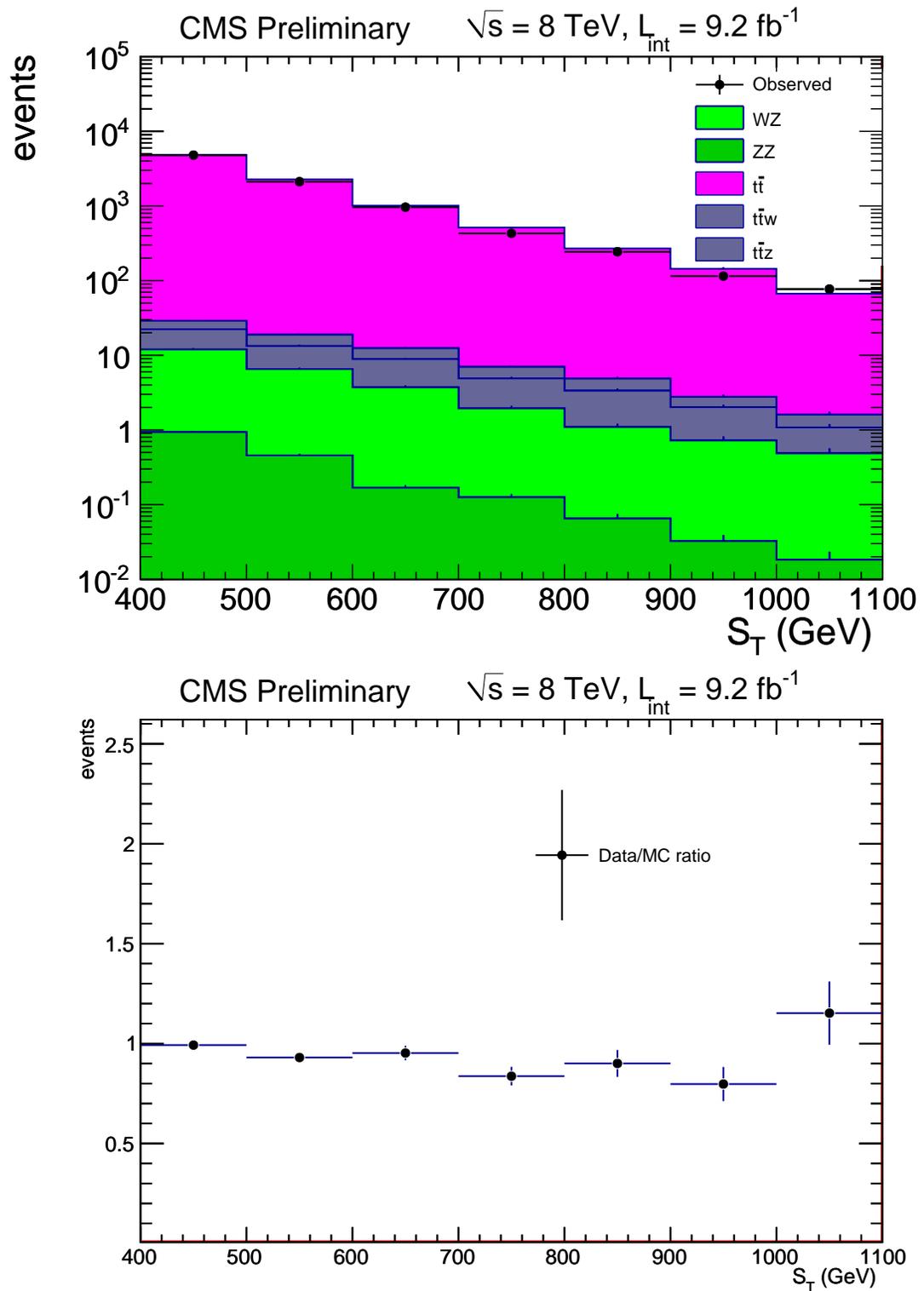
Figure 3.1: The S_T distribution of datasets dominated by $t\bar{t}$ (top). Ratio of Data/MC for the S_T distributions in the dilepton control region for $t\bar{t}$ (bottom).
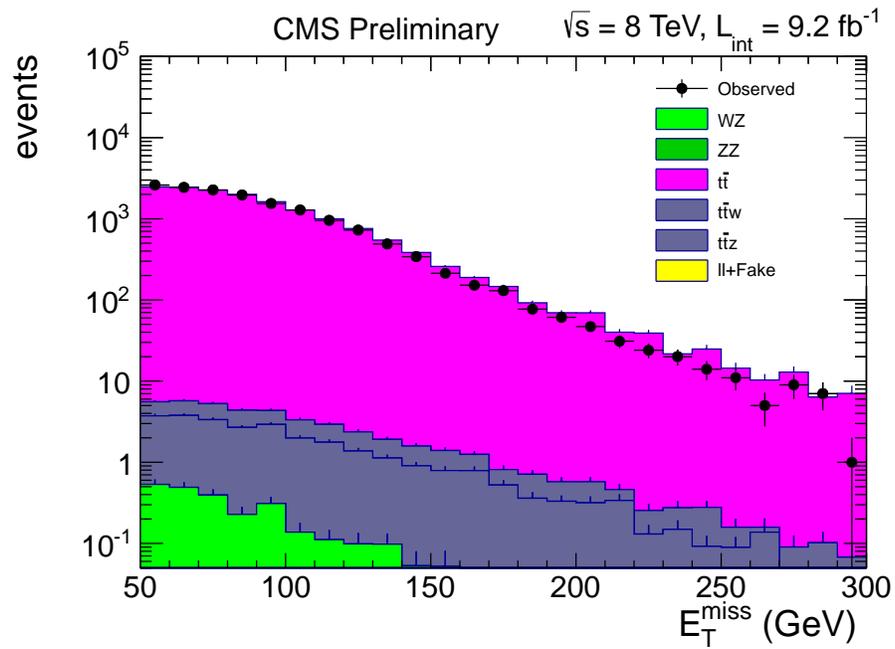
Figure 3.2: $E_\mathrm{T}^\mathrm{miss}$ distribution for the $t\bar{t}$ dilepton control region.
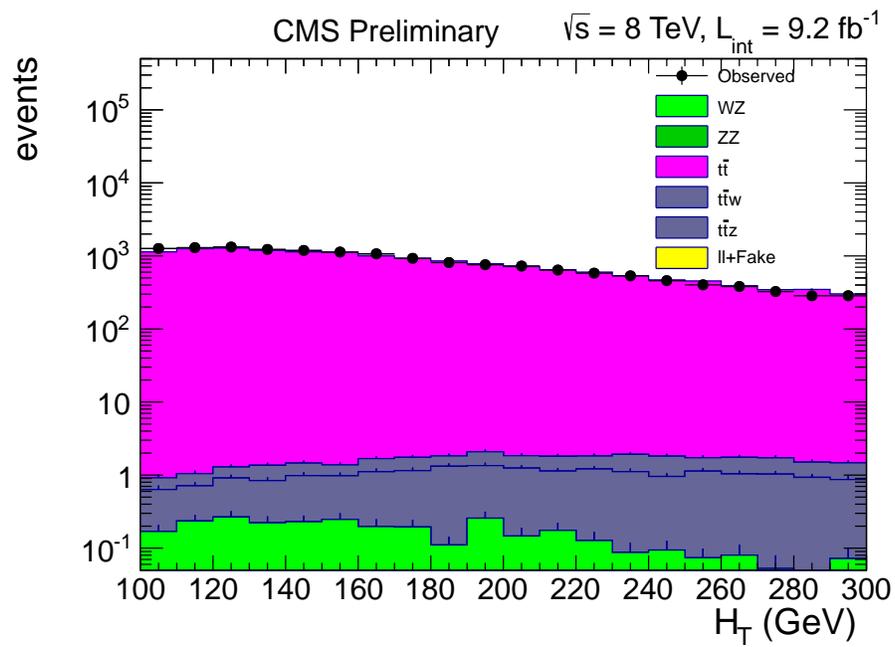
Figure 3.3: H_T distribution for the $t\bar{t}$ dilepton control region.

contributions from $t\bar{t}$ to 3 or 4 lepton channels involves at least one fake-lepton coming from a jet. To ensure that the fakes are being modeled properly, one needs to look at the isolation distribution of lepton candidates in the $t\bar{t}$ MC sample and compare it to data. Figure 3.4 shows the relative isolation distribution of muons. The MC agrees with data indicating that the fake-rate for $t\bar{t}$ jets to give a fake-lepton is the same in data and MC.
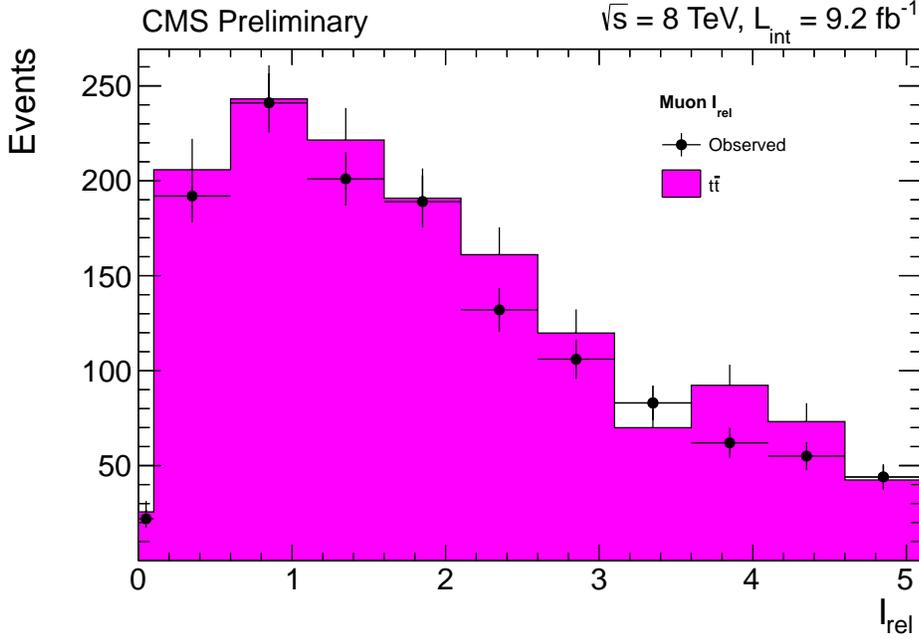
Figure 3.4: Relative isolation distribution for non-prompt muons in the $t\bar{t}$ single muon control region.

3.1.2 Controls on WZ

WZ can produce 3 leptons where $Z \rightarrow l^+l^-$ and $W \rightarrow l\nu$. We define a control region by requiring an opposite-sign same-flavor (OSSF) lepton pair with invariant mass in the Z-window (75-105 GeV). In addition, while looking at the E_T^{miss} distribution, we require the third lepton and the E_T^{miss} to form an invariant mass (transverse mass) in the W-window (60-100 GeV). On the other hand, while looking at the distribution of the transverse mass (M_T) of the third lepton and the E_T^{miss}, we require $E_T^{miss} > 50$ GeV.

In either of the control regions, if all three leptons have the same flavor, say there are three muons with two having charge +1 and one having charge -1, we choose the muon pair with mass closest to the Z-mass as the pair coming from the Z and choose the third lepton as the one coming from the W.

Figures 3.5 and 3.6 show the E_T^{miss} and transverse mass distributions respectively and show good agreement between the MC and data.
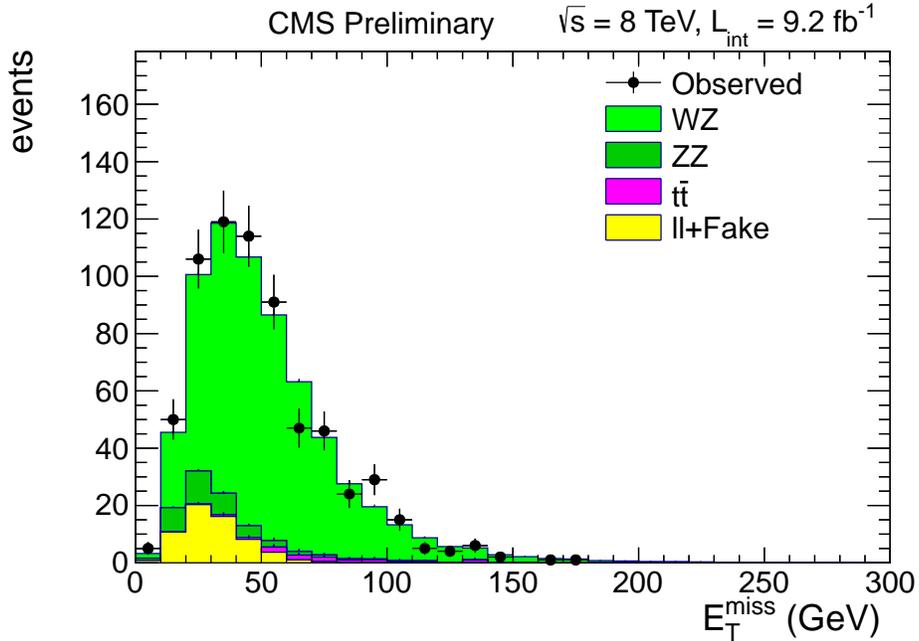
Figure 3.5: *MET* distribution for WZ control region with 3 leptons.
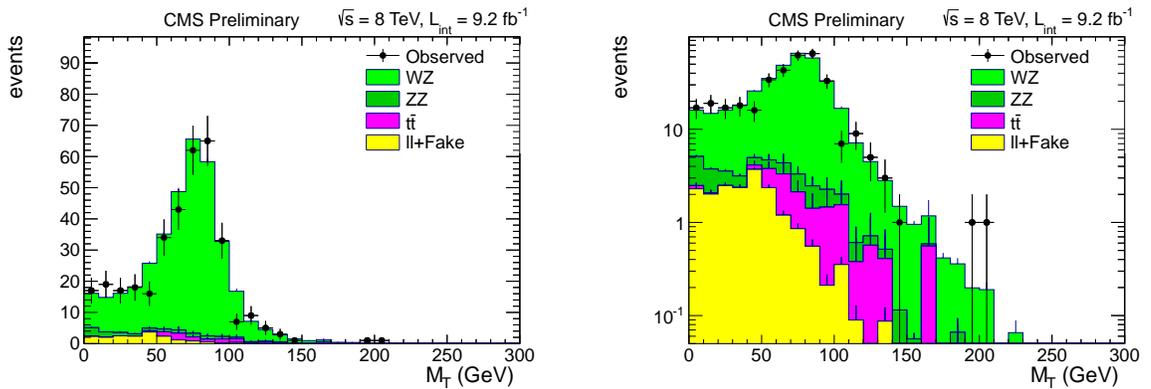
Figure 3.6: M_T distribution for WZ control region with 3 leptons on linear (left) scale and log (right) scale.

3.1.3 Controls on ZZ

ZZ can give 4 leptons with two OSSF pairs where each $Z \to l^+l^-$. The ZZ control region is defined with four leptons, with atleast one OSSF pair in the Z-window . The invariant mass distribution of all four leptons for this control region is shown in Figure 3.7 and agrees well with data.
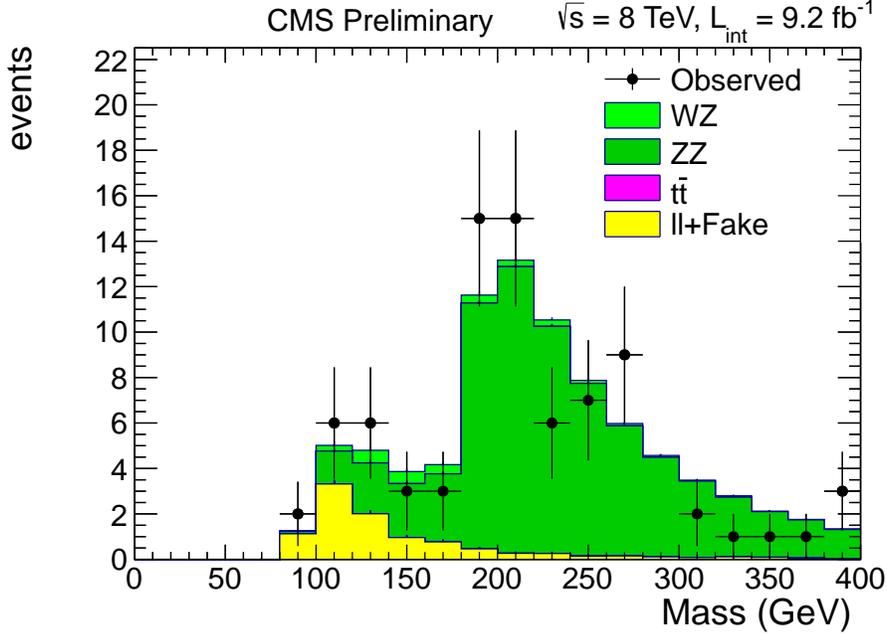
Figure 3.7: Invariant mass distribution for four-lepton events with atleast one OSSF pair with invariant mass on Z. This is the ZZ control region.

3.2 Data-driven Backgrounds

In addition to the MC backgrounds mentioned above, there are background contributions from Z+Jets (where $Z \to ll$ and there's a fake-lepton from jets), WW+Jets (where each $W \to l\nu$ and there's a fake-lepton from jets), QCD (jet-rich events), and Z+γ. We use data itself and conversion factors measured in data to estimate these backgrounds. The reason we don't use MC for these backgrounds is because they all involve fake leptons that depend on parameters (like

jet composition i.e. relative ratios of b-jets, c-jets, light jets and gluon jets) that we don't trust in MC. Also, as conditions at the LHC change, these underlying parameters might change and one would need to keep regenerating MC samples.

Before going on to details of various data-driven techniques, we describe the general philosophy behind such methods. The basic quantity we want to estimate is the rate at which an object which is not a lepton coming from the primary interaction (i.e. the objects could be jets, photons, leptons radiated by mesons in jets) passes all the cuts for a lepton coming from the primary interaction. These objects will be called "fakes". We first pick an object that is almost exactly the same as a "fake" but occurs far more frequently and is produced by similar physics processes. This more frequently occuring object is called a "proxy". Then, the ratio of the number of "fakes" to the number of "proxies" gives us what is called a **fake-rate** or a **conversion factor**. We could use this conversion factor to estimate the number of "fakes" in data except that the physics environment (jet activity etc.) usually is different in datasets where the conversion factor is measured and where it is applied. The different physics environments might affect the conversion factors and one can no longer be sure that they match in the two datasets. So one identifies a set of parameters that parameterize the physics environment (total jet p_T etc.) and measures the conversion factor as a function of this parameter. After measuring this parameter in the dataset where the conversion factor is to be applied, one can tune the conversion factor since we know it as a function of this parameter and this way we can be confident that the conversion factor being applied is appropriate for the physics environment. This abstract discussion will get clearer in the examples below.

The background estimation methods for background leptons originating from jets are based on those developed for (25). Data driven background estimations for leptons from photon conversions were added in 2011.

3.2.1   Data Driven Fake Prompt lepton Estimation

The rate for jets to produce light fake leptons depends on many factors that may not be properly simulated. These factors include, but are not limited to, jet shape, jet spectra, spectra of particles within a jet, and the probability for a jet to contain heavy flavor particles.

We use data-driven methods to estimate the probability of jets to produce light lepton candidates that appear to be prompt and isolated.

Details of measuring lepton fake-rates

As described above, for this analysis a "fake prompt" lepton is an isolated lepton candidate that appears to have come directly from the primary vertex, but actually originated from a jet or other process. These "fake prompt"s include real leptons from semileptonic decays of heavy flavor mesons. In order to estimate the the abundance of "fake prompt" lepton candidates produced via jets we determine a conversion factor to relate the number of fake prompt leptons to the number of some other type of object in the data. This is sometimes called a "fake rate" method and is similar to a "tight-loose" method. In a "fake rate" method one chooses an object in the data to act as a proxy for the fake prompt lepton candidates. A conversion factor relating the number of proxy objects to the number of fake prompt objects is measured in a control sample. Then, proxy objects can be treated as a lepton in the analysis, and each event with a proxy is scaled once by the conversion factor for each proxy in the event. This conversion factor is sometimes referred to as a "fake rate", even though the leptons being studied may be real leptons.

For this analysis we use isolated tracks (pions) as a proxy object for prompt electron and muon candidates that originate from jets. The relative abundance of pion candidates to electron or muon candidates depends on the composition

of jets, jet spectra, as well as heavy meson semileptonic branching fractions and form factors.

It is necessary for proxy objects to occur more frequently than the fake prompt lepton in order to have a good statistical measure of the background. Isolated tracks are produced by generic QCD jets 30-50 × more often than lepton candidates.The fact that both the tracks and the electron/muon candidates are isolated reduces systematic uncertainties in the conversion factor related to the jet spectra. This is because the isolation distributions of different types of objects originating from jets scale similarly with the jet energy

The relative number of leptons to tracks also depend on the types of jets in the sample, or the composition of jets. Jets initiated by b-quarks or c-quarks are much more likely to contain heavy flavor mesons, and real leptons, than jets initiated by uds quarks.

We relate the number of isolated leptons from jets to the number of isolated tracks from jets by the conversion factors, f_μ and f_e, where

$$N_\mu^{\text{Iso}} = f_\mu \times N_{\text{Track}}^{\text{Iso}} \qquad \text{and} \qquad N_e^{\text{Iso}} = f_e \times N_{\text{Track}}^{\text{Iso}}. \qquad (3.1)$$

In the above, N_μ^{Iso} is the number of isolated muons, $N_{\text{Track}}^{\text{Iso}}$ is the number of isolated tracks. The conversion factors f_μ and f_e are potentially functions of p_t, η, jet spectra, and jet composition. For this analysis, fake prompt e's and μ's predominantly have p_t less than $\sim$24 GeV. Therefore, since this is a counting experiment, we determine the conversion factors for a single bin in p_t and η (p_t from 8 to 24 GeV/c and $|\eta| < 2.4$).

$$f_{\mu/e} = \frac{N_{\mu/e}}{N_T} \times \frac{\epsilon_{\mu/e}^{Iso}}{\epsilon_T^{Iso}}, \qquad (3.2)$$

Above, N_μ is the number of non-isolated muon candidates in the data set, N_T is the number of non-isolated tracks, and $\epsilon_\mu^{Iso}/\epsilon_T^{Iso}$ is the ratio of muon and track

isolation efficiencies. We can measure N_μ and N_T directly from the sample where we apply the conversion factor (if there is enough statistics) and we parameterize the ratio of isolation efficiencies from the control sample.

We expect our conversion factor to depend on the composition of jets, or how often the jets are from heavy flavor processes. This is because for $p_t <$ 40 GeV fakes prompt leptons are mostly from heavy flavor. The parameter that we use to parameterize the isolation efficiency is R_{dxy} which is defined to be the ratio of the number of tracks with $|d_{xy}(BS)| > 0.02$ cm to the number of tracks with $|d_{xy}(BS)| < 0.02$ cm. Note that the 0.02 cm cut is based on the width of impact parameter for prompt objects. Other requirements that the tracks going in the calculation of R_{dxy} are- $p_t > 8 GeV$, track should not be within dR of 0.1 of selected e, mu and 0.3 of selected tau. Tracks coming from lead jet are vetoed. Tracks associated with jets originating from b and c quarks have larger impact parameter relative to the beam spot ($d_{xy}(BS)$) than jets form light quarks and gluons and will, therefore, have correspondingly larger values of R_{dxy}.

We expect R_{dxy} to have a minimum value of few % in a sample with no heavy flavor jets, and R_{dxy} will have a maximum value of 20%-30% in a dilepton $t\bar{t}$ sample. If fake and non-isolated fake lepton candidates are primarily from heavy flavor we expect that the efficiency ratio should be roughly linear in R_{dxy} for small values of R_{dxy}. However, for larger values of R_{dxy} the efficiency ratio should asymptotically approach the ratio in $t\bar{t} \to \ell^+\ell^-\nu\nu b\bar{b}$ as R_{dxy} approaches 20%-30%.

To measure the relationship between the efficiency ratio and R_{dxy}, we look in a sample with no bjets (nominal), that is events with a Z and a non-prompt lepton . Then we calculate efficiency ratio and R_{dxy} using these non prompt leptons and non-prompt tracks.There after we look in to a sample of pure b jets. For this we look at events with an isolated muon, three or more jets with $p_t > 40$ GeV , at least one b tagged jet, $S_T > 400 GeV$ and the selected muon should not come from the b-jet and an additional non prompt lepton. This non-prompt lepton is

then used to calculate efficiency ratio and R_{dxy} values for ttbar-like sample.

There after we parametrize Efficiency ratio and R_{dxy} in terms of a free parameter α which gives the contribution from nominal sample and pure b jet sample . Using α and the above measured quanities , we establish the following relation-

$$Efficiency\ ratio(\alpha) \ = \ \frac{[effrat\_mutr(a) \ + \ rat\_mutr \times \alpha \times effrat\_mutr(b) \times effrat\_ba\_track] \times [1 + \alpha]}{[1 + \alpha \times effrat\_ba\_track][1 + rat\_mutr \times \alpha]} \qquad (3.3)$$

$$R_{dxy}(\alpha) \ = \ [Rdxy(a) + \alpha \times Rdxy(b)]/[1 + \alpha] \qquad (3.4)$$

where,

Rdxy(a) = Rdxy of nominal sample (no b jets)

Rdxy(b) = Rdxy of a sample of pure b jets

rat_mutr = Ratio of ratio of non-isolated muons (leptons) to non isolated tracks in pure b sample to ratio of non-isolated muons (leptons) to non isolated tracks in nominal sample

effrat_mutr(a) = isolation efficiency ratio of muons to tracks in nominal sample (no b jets)

effrat_mutr(b) = isolation efficiency ratio of muons to tracks in pure b jets

effrat_ba_track = isolation efficiency ratio of tracks in sample (b) to tracks in sample (a)

Expressing Efficiency ratio in terms of R_{dxy} -

$$Efficiency\ ratio \ = \ \frac{[Rdxy(a) - Rdxy(b)] \times [effrat\_mutr(b) * effrat\_ba\_track * rat\_mutr(Rdxy(a) - Rdxy) + effrat\_mutr(a)(Rdxy - Rdxy(b))]}{[Rdxy(b) - Rdxy(a) * effrat\_ba\_track + (effrat\_ba\_track - 1)Rdxy][Rdxy(b) - rat\_mutr * Rdxy(a)] + (rat\_mutr - 1)Rdxy}$$
$$(3.5)$$

Varying α gives new values of R_{dxy} and for each value of R_{dxy} we get an efficiency ratio

Figure 3.8 shows plots for efficiency ratio versus R_{dxy} for muons and electrons.

Efficiency ratio for muon(electron) at low values of Rdxy is 1.4(1.6) and for higher values of Rdxy it is 4.8(3.6) .
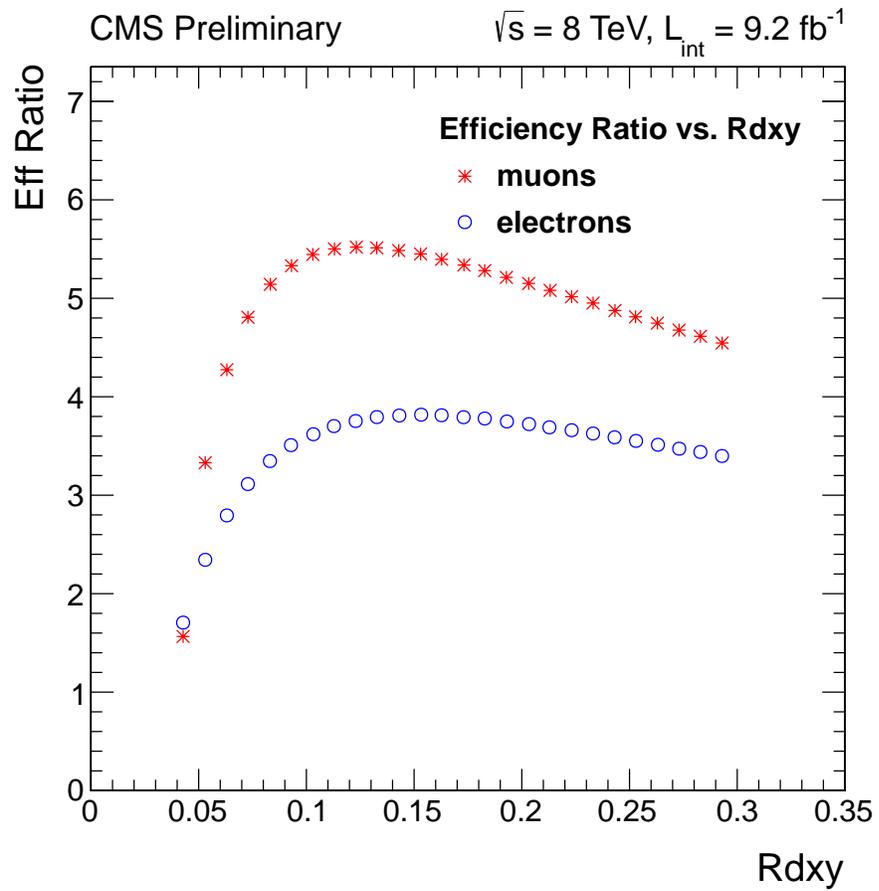
Figure 3.8: Efficiency ratio of leptons to tracks versus fraction of non-isolated tracks with a large impact parameter.

We estimate the background with "fake prompt" leptons by selecting events with two isolated leptons as a "seed" data set. We measure the number of isolated tracks, non-isolated tracks, and non-isolated leptons, and subtract out contributions from backgrounds determined via MC. We measure R_{dxy} for the remaining data, and determine the conversion factors for the data set. The conversion factors are $\sim 0.714\%$ for muons and $\sim 0.91\%$ for electrons. We apply a systematic to cover conversion factor differences on and off the Z peak, and 10% systematic from the assumption that $< (N_\ell/N_T) \times (\epsilon_\ell/\epsilon_T) > = < (N_\ell/N_T) > \times < (\epsilon_\ell/\epsilon_T) >$.

To predict our 3 lepton backgrounds, where all three leptons are electrons or muons ,we scale the number of events with isolated tracks. To estimate backgrounds with two fake prompt leptons, we select events with two isolated tracks, and apply the conversion factor twice. For events with same sign dileptons plus tau we use tight-loose method described in section For estimating the background to events with three isolated leptons with one additional isolated track, we select events with two isolated tracks and apply a scale factor, f_μ and/or f_e, to just one of the isolated tracks.

Lepton Fake Rate Given b-Tagged Jet(s)

Leptons from the semileptonic decay of B mesons are a source of fake leptons. Due to lepton isolation criteria, these leptons can fake genuine leptons only if they take up a large portion of the energy of the parent B meson. As a result, the jet that is produced from the B meson would be less likely to be reconstructed as a jet and to be tagged as a b jet. Consequently, if a b-tagged jet is found in an event, the leptons in the same event would be less likely to be fake.

We select events in a $t\bar{t}$ Monte Carlo sample where a lepton coming from the semileptonic decay of a B meson passes the lepton selection cuts. The B meson that produces such lepton is then tracked to see if it produces a reconstructed jet. In this case, we note whether the jet is tagged as a b-jet. We further check

the separation between the lepton and the reconstructed jet (dR) formed by its mother B meson.

The transverse momentum distribution of fake muons that are produced from the semileptonic decay of B mesons which also form a jet are shown from figure 3.9 to figure 3.11 for ranges of mother B mesons transverse momentum. The transverse momentum distribution shows that the fake muons indeed take up a large fraction of momentum from its mother B mesons, effectively reducing the accompanied jet multiplicity and hence reconstruction efficiency.

For B mesons that produce a fake lepton, the efficiency for which the B mesons are successfully reconstructed as a jet can be found in Table 3.1. The b-tagging efficiency of various b-taggers for such reconstructed jet can be found in Table 3.3. The rather low jet reconstruction efficiency and b-tagging efficiency confirm our expectation where a fake lepton producing B meson would be too soft to be successfully reconstructed and b-tagged. The separation between the fake lepton and the jet resulting from the mother B meson is shown in figure 3.12; note that most of the fake leptons lie within a dR cone of 0.3 in the jet produced from the mother B meson. correspondingly, the jet reconstruction efficiency with this additional dR requirement is shown in Table 3.2.

The results we found in this study conclude that we can safely ignore tracks around a b-tagged jet in the calculation of leptons fake rate, owing to the fact that a B meson that produces fake leptons can seldom be reconstructed as a jet and even more rare to be tagged as a b jet.

Lepton flavor	Jet reconstruction efficiency
Electron	11.47%±3.23%
Muon	13.00%±1.80%

Table 3.1: Jet reconstruction efficiency for B mesons that produces fake leptons
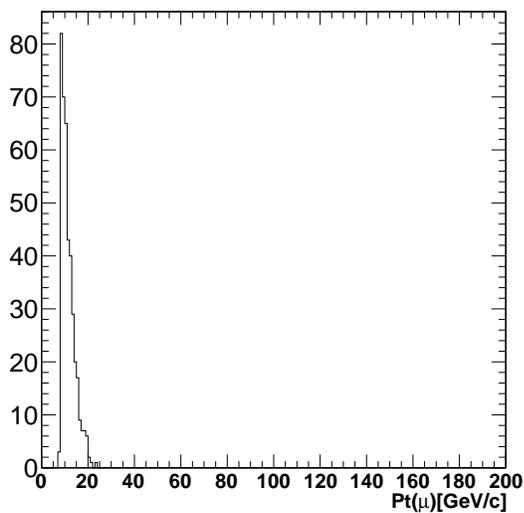
Figure 3.9: Pt distribution of muons for mother B mesons Pt range from 10 - 30 GeV/c^2
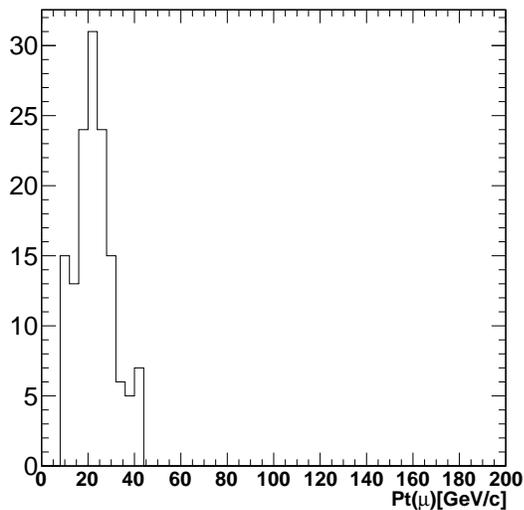
Figure 3.10: Pt distribution of muons for mother B mesons Pt range from 30 - 60 GeV/c^2

Lepton flavor	Jet reconstruction efficiency (dR > 0.3)
Electron	1.91%±1.15%
Muon	2.60%±0.70%

Table 3.2: Jet reconstruction efficiency for B mesons that produces fake leptons with additional requirment where dR > 0.3.
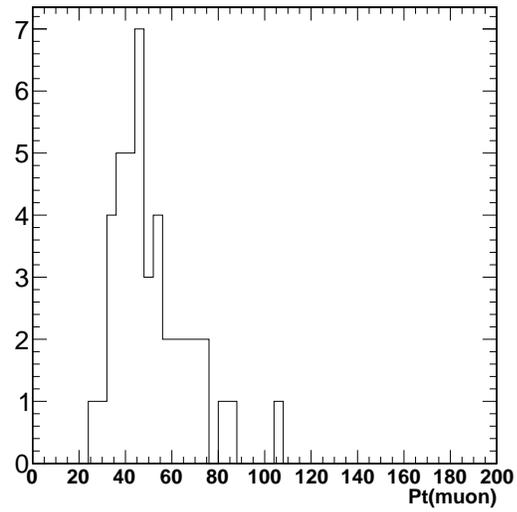
Figure 3.11: Pt distribution of muons for mother B mesons Pt range from 60 GeV/c^2 or above
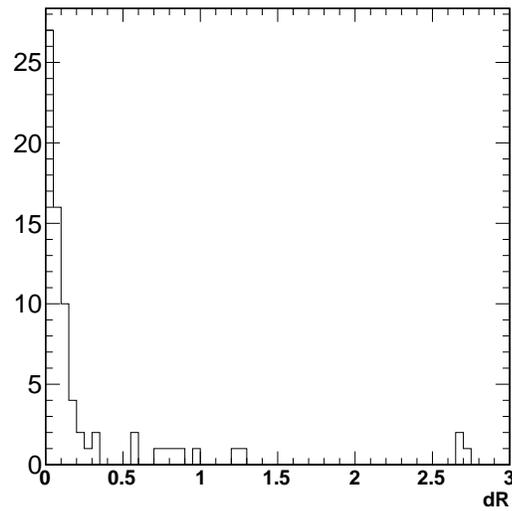
Figure 3.12: Distribution of the separation between the fake muons and the mother B mesons

Tagger	Electron efficiency	Muon efficiency
TCHPT	0.64%±0.65%	1.91%±0.59%
CSVL	8.28%±2.64%	10.05%±1.53%
CSVM	5.10%±1.98%	4.51%±0.95%
CSVT	3.18%±1.52%	3.81%±0.87%
JPL	7.00%±2.38%	7.63%±1.29%
JPM	5.10%±1.98%	4.16%±0.91%
JPT	3.18%±1.52%	1.91%±0.59%

Table 3.3: b-tagging efficiency given physics muon or electron that come from decay of the B meson that is tagged as a b.

3.2.2 Tau Backgrounds from Jets

Taus are much heavier than either electrons or muons and are unstable with short lifetimes. They decay within the detector to either electrons/muons or to pions. In the latter case, they are called **hadronically-decaying taus**. Unlike electrons and muons, hadronically decaying τ leptons cannot be easily identified without isolation (in other words, a non-isolated hadronic tau has almost no distinguishing features that separate it from jets). Therefore, the dominant background to τ leptons are fakes from jets. To determine τ fakes in a data set, we use an isolation side band, and determine a conversion factor for the number of τ in the sideband to the number of fake τ that pass our isolation cut. We call this conversion factor f_t. Unfortunately since isolation distributions vary dramatically with jet multiplicity and jet spectra, f_t can differ greatly between data sets. Therefore, we define a second parameter f_{SB}, the ratio of non-isolated τ candidates in the isolation side band divided by the total number of non-isolated τ candidates in a given data set. f_{SB} parameterizes the amount of jet activity in the dataset.

To determine the dependence between f_t and f_{SB}, we divide dilepton data with one OSSF pair on-Z into bins of the $\sum p_T$ of all tracks in the event associated with the primary vertex. This gives us different data sets with different jet spectra and multiplicity. For each such bin, we calculate f_t and f_{SB} from the isolation distribution of tau candidates and determine their functional relationship. We

find that the dependence of f_t on f_{SB} is the same for on-Z dilepton data and off-Z dilepton data within a systematic uncertainty of 25%.

The value of f_{SB} is large (close to 1) if the data set contains very soft jets, and f_t is also large and fake τs can easily pass the isolation criteria. For a data set with a large multiplicity of high p_T jets, the value of f_{SB} and f_t is small. Figure 3.13 shows f_t versus f_{SB} for two different p_T ranges.

As the total sum p_t increases, taus tend to find it harder to stay isolated and the isolation distribution moves to the right. This results in decreasing f_{SB}, which acts as a free parameter. We plot f_t vs f_{SB} for the on-Z dilepton dataset. We do a fit to get an analytic expression for $f_t(f_{SB})$ and assign an error that is either the difference between lead jet and sum track p_T binned curves or 25% of the f_t, whichever is larger.

To predict the $N\ell$+fake τ background we bin the $N\ell$ data in bins of $\sum p_T$ and for each bin calculate f_{SB} and use it to determine f_t using the correlation described above. Candidate τ in the isolation side band can then be used as actual τ candidates, and their contributions to the background are scaled by f_t (10%-20%).
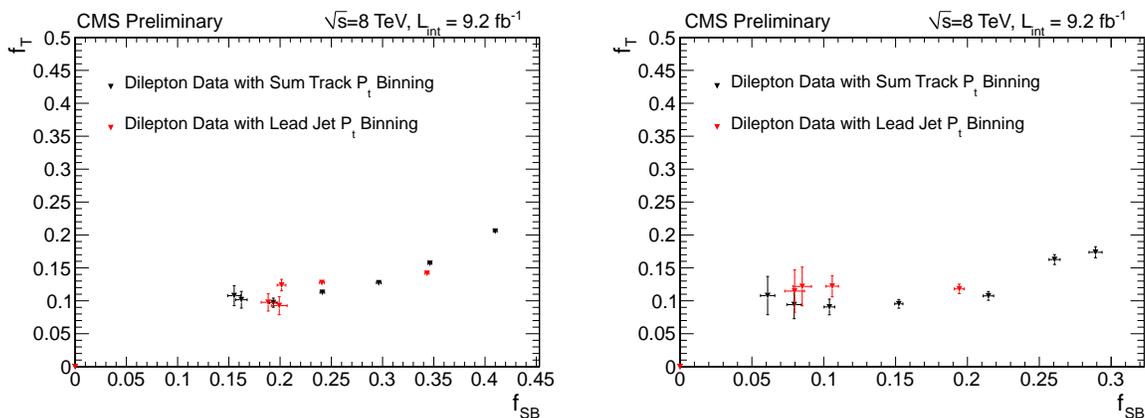
Figure 3.13: f_t vs f_{SB} for taus with visible tau P_t between 20-40 GeV (left) and 40-60 GeV (right) in on-Z diLepton data.

3.2.3 Backgrounds From Asymmetric Photon Conversions

Photons converting to $\ell^+\ell^-$ are a source of backgrounds to analyses with leptons. If both leptons are reconstructed the invariant mass will be small and a cut of $M(\ell^+\ell^-) > 12$ GeV should remove them (along with upsilon, J/Ψ, and low mass Drell Yan). Therefore the mode by which conversions give backgrounds to multiplepton analyses is if the conversion is asymmetric, where one lepton takes most of the photon energy and the second lepton is very soft and not measured.

There are two different types of photon conversions that can give rise to backgrounds in multilepton analyses. The first type is an "external conversion". External conversions occur when a photon radiated by the collision interacts with the material in the detector and generates an $\ell^+\ell^-$ pair. External conversions primarily produce e^+e^- pairs as opposed to $\mu^+\mu^-$ pairs since muons are much heavier than electrons. The ratio of the rate of external conversions to e^+e^- and the rate of external conversions to $\mu^+\mu^-$ is between 6.0×10^4 and 2.0×10^5. Therefore, conversions occurring within the detector material rarely produce muons. The electron identification requirements include cuts designed to remove most of the external conversions from electrons.

The second type of photon conversions are "internal conversion" where the photon is virtual and never interacts with the material in the detector. Internal photon conversions can produce muons almost as often as electrons, and can occur in any process with virtual photons (see figure 3.14).

Internal conversions are an important background to any analysis with leptons and may be underestimated by various MC generators. While MC generators do simulate internal conversion of emitted photons, by necessity they have a cutoff on the conversion lepton momentum. Therefore highly asymmetric conversions, with one of the leptons at ≤ 1 GeV/c and the other one carrying all the momentum of emission, is not properly accounted for.
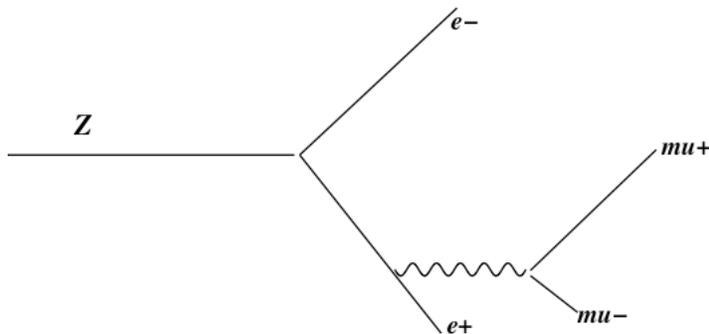
Figure 3.14: A Feynman diagram showing a Z decay to electrons, and an asymmetric FSR decay to muons (indicated by length of the muon legs).

We note that for the multilepton analysis, the most important source of this background involves Z bosons decaying to leptons, and an asymmetric internal conversion of an FSR γ^* from one of the Z's decay leptons. The FSR would cause the invariant mass of the leptons from the Z decay to not reconstruct at the Z-pole, and the asymmetric internal conversion would add one more lepton to the event. Figure 3.14 shows a diagram of this process. The Z is shown decaying to electrons, and the e^+ emitting an FSR γ^* which produces a $\mu^+\mu^-$ pair. If one of the muons carries most of the γ^* momentum, this could appear as a 3-lepton event.

Since internal conversions may not be properly simulated by the MC it is important to have a data driven method for estimating the background. To do this data driven background estimation we assume that the rate for standard model to produce on-shell photons is proportional to the rate for producing virtual photons that yield asymmetric conversions. We use isolated photon candidates as our "fakeable object". The conversion factor (fake rate) is a ratio of the probability for a photon to produce a valid lepton candidate via asymmetric conversion divided by the probability for the photon to be on-shell and pass all photon selection criteria.

We use the final state radiation (FSR) of the Z to measure the conversion

factor for both muons and electrons. Both $Z + \gamma$ (ISR) and $Z \rightarrow \ell^+\ell^-\gamma$ (FSR) can clearly be seen in the data by plotting the $M(\ell^+\ell^-\gamma)$ mass versus $M(\ell^+\ell^-)$ for low MET, HT, ST events 3.15. The FSR band is seen spread along the x-axis at 90 GeV on the y-axis. The ISR band is seen spread along the y-axis at 90 GeV on the x-axis. The 3-body Z-peak is a clear indication that the photon (or third lepton) involved is produced by FSR.
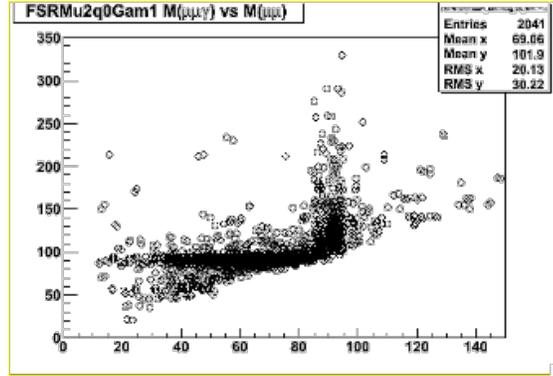
Figure 3.15: $M(\ell^+\ell^-\gamma)$ versus $M(\ell^+\ell^-)$. The FSR band is seen spread along at the x-axis at 90 GeV on the y-axis. The ISR band is seen spread along the y-axis at 90 GeV on the x-axis.

In a control region devoid of new physics (low S_T or low MET and low H_T) we find clean FSR events by searching for a 3-body Z peak. As mentioned earlier, if on-resonance $Z \rightarrow \ell^+\ell^-$ has FSR on one of the legs, the dilepton mass will be pushed off the Z peak. However, the 3-body mass of the $\ell^+\ell^-\gamma(*)$ will still be on the Z peak.

To get the conversion factor we divide the number of $\ell^+\ell^-\ell^\pm$ on the Z peak by the number of $\ell^+\ell^-\gamma$ on the Z peak. Figure 3.16 shows the 3-body $\ell^+\ell^-\gamma$ mass in data where the $\ell^+\ell^-$ is not on the Z peak (< 75 GeV and > 105 GeV). Figure 3.17 shows the 3-body $\ell^+\ell^-e^\pm$ peak where no opposite sign same flavor (OSSF) pair makes a Z candidate. Figure 3.18 shows the 3-body $\mu^+\mu^-\mu^\pm$ peak where no opposite sign same flavor (OSSF) pair makes a Z candidate. The 3μ plot is especially interesting because there is essentially no contributions from external conversions, and the peak that we see is entirely from internal conversions.
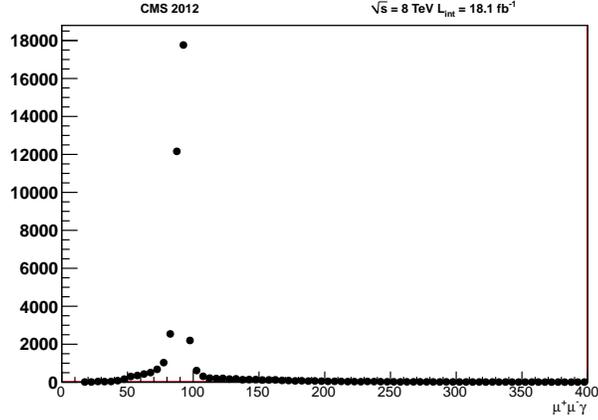
Figure 3.16: $M(\ell^+\ell^-\gamma)$ where $M(\ell^+\ell^-)$ is either < 75 GeV or > 105 GeV.
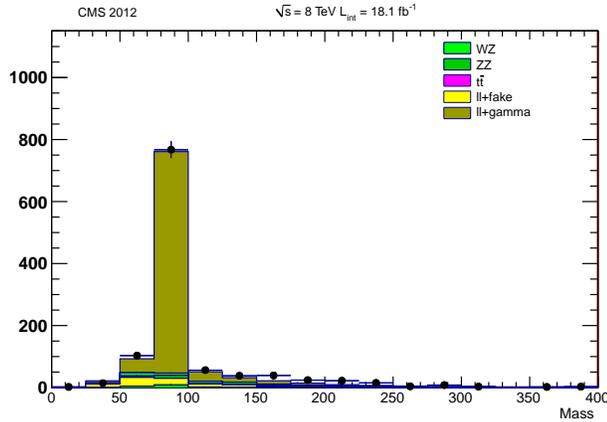
Figure 3.17: $M(\ell^+\ell^- e^\pm)$ where $M(\ell^+\ell^-)$ is either < 75 GeV or > 105 GeV. Note that most external conversions to electrons have already been removed in the electron identification requirements.

We use the 3-body mass peak in the range 80-100 GeV and find that the FSR conversion factor for muons (C_μ) is $0.8\% \pm 0.1\%$ and the conversion factor for electrons (C_e) is $1.8\% \pm 0.3\%$ where uncertainties are statistical only. We assign theoretical systematic uncertainties of 100% to these conversion factors from our underlying assumption that the number of isolated photons is proportional to the number of leptons from asymmetric internal and external conversions from ISR or FSR photons. Keep in mind that the relative acceptance of observed photons and lepton candidates is folded in the fake rate measurment, which is therefore very specific to the selections of this analysis.
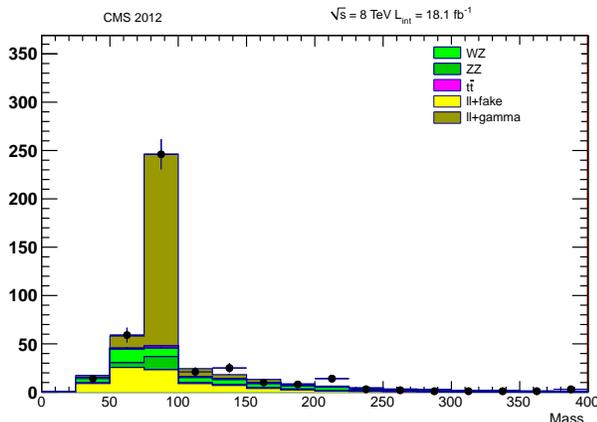
Figure 3.18: $M(\mu^+\mu^-\mu^\pm)$ where both $M(\mu^+\mu^-)$ are either < 75 GeV or > 105 GeV.

3.3 A walk through the background prediction methods

The previous sections described the measurement of several data-driven fake-rates. Our eventual goal though is to calculate the total background contribution for each channel. For MC backgrounds, this is straightforward. As an example, suppose the channel consists of two opposite-charge electrons and a muon. In this case, one would take an MC simulation sample (say $t\bar{t}$ for concreteness), count the fraction of total events (also called **acceptance**) that pass the cuts for the $e^+e^-\mu^\pm$ channel and scale the fraction by $\sigma_{MC}\mathcal{L}$ where σ_{MC} is the cross-section for the MC source and $\mathcal{L}$ is the total luminosity of the data collected. The quantity acceptance $*\sigma_{MC} * \mathcal{L}$ is the contribution from $t\bar{t}$ to this channel.

Even after accounting for all the MC backgrounds, there might be backgrounds not accounted for which require the data-driven techniques described above. This section gives an example of how all the above fake-rates/conversion factors are applied to data.

In the case of e and μ, the source of backgrounds is largely heavy flavor production in jets, with associated decay to leptons which can pass isolation

requirements. For τ, the main source of backgrounds are jets which satisfy τ-isolation requirements, which can be affected by the amount of jet activity in the event. Therefore we parametrize the probability for an e or μ arising from a jet using a variable (R_{dxy}) that correlates with the heavy flavor content of jets. We parametrize the probability for a jet faking a τ in terms of another variable (f_{SB}) that corresponds to the amount of jet activity in the event.

3.3.1 Electron and Muon backgrounds from jets

Jets produce isolated e and μ largely through heavy flavor. We need a parameter that correlates with the amount of heavy flavor in an event. We can then parametrize the probability of a jet to produce an isolated lepton in terms of this parameter.

We define R_{dxy} as the ratio of the number of tracks in a given data sample with impact parameter (d_{xy}) greater than 200 microns to that less than 200 microns. Events with heavy flavor jets will have many tracks with large impact parameter, leading to larger R_{dxy}. Events with light flavor jets will have R_{dxy} of 0.02 - 0.03, while $t\bar{t}$ events will have R_{dxy} around 0.2 - 0.3.

$$R_{dxy} \equiv \frac{\text{Ntrack high } d_{xy}}{\text{Ntrack low } d_{xy}}$$

In practice, since we use MC to predict Weak decay backgrounds ($t\bar{t}$, diboson) we subtract the number of tracks expected from these when calculating R_{dxy} and the background prediction from jets. Figure 3.8 shows the variation of $\epsilon_{e,\mu}/\epsilon_t$ with R_{dxy}.

In order to parametrize the correlation of R_{dxy} with heavy flavor content, we divide the jet-enriched dataset into 12 bins of b-tag discriminant (-5 to 7 in steps of 1). We further divide these into 3 bins of leading jet pt (0-60 GeV, 60-150 GeV, and >150 GeV) for a total of 36 bins.

For each of these 36 bins, we also calculate ϵ_ℓ and ϵ_{track}. These are simply the efficiencies of finding *isolated* leptons or tracks, given a sample of leptons or tracks. They are defined as follows:

$$\epsilon_{\ell(\text{track})} \equiv \frac{\text{Num iso } \ell(\text{track})}{\text{Num total } \ell(\text{track})}$$

Note that some of the 36 bins have little to no statistics, and are ignored. The efficiency ratio increases monotonically with R_{dxy} as we expected, since heavy flavor jets (large R_{dxy}) can produce real isolated leptons. We can now determine the efficiency ratio $(\frac{\epsilon_\ell}{\epsilon_{\text{track}}})$ for a given dataset by calculating its R_{dxy}.

In order to predict the number of backgrounds in the 3ℓ dataset, we look at the 2ℓ dataset, and calculate the R_{dxy}. From this, and the correlation taken from the jet-enriched dataset above, we extract $\frac{\epsilon_\ell}{\epsilon_{\text{track}}}$

Now we know the ratio of efficiencies for turning a non-isolated object into an isolated one, for the 2ℓ dataset. We can now predict the number of isolated leptons from background in the 3ℓ sample by counting the number of non-isolated leptons and tracks, as well as the number of isolated tracks in the 2ℓ dataset.

$$\frac{\epsilon_\ell}{\epsilon_{\text{track}}} \times \frac{\text{Num non} - \text{isolated } \ell}{\text{Num non} - \text{isolated track}} \times \text{Num isolated track}$$

$$= \text{Predicted number of isolated leptons in } 3\ell \text{ sample}$$

Illustration of $\mu\mu e$ Background from $\mu\mu$

We illustrate the method by calculating the number of $\mu\mu e$ events expected from the $\mu\mu$ sample plus a background e. This is just one component of the background in this sample, but the method is the same. For brevity's sake we neglect error propagation in this illustration.

We determine the following numbers from the data ($\mu\mu$ opposite charge, 190/pb sample): These numbers are for the background e in the p_t range of

8 GeV/c to 24 GeV/c. The first number is the actual count, while the second number (bold, and in brackets) is the count with MC components subtracted.

- Num non-isolated electrons in 2μ dataset: 100, [**89.1**]

- Num non-isolated tracks in 2μ dataset: 14329, [**13714.9**]

    - with $d_{xy} > 200$ microns: 680, [**506.5**]

- Num isolated tracks in 2μ dataset: 488, [**485.2**]

    - with $d_{xy} > 200$ microns: 20, [**19.1**]

From these we can determine that for this $\mu\mu$ dataset, $R_{dxy} = 0.038$. We use the fit to R_{dxy} vs. $\frac{\epsilon_e}{\epsilon_{\text{track}}}$ to extract $\frac{\epsilon_e}{\epsilon_{\text{track}}} = 0.91$.

We now predict the number of background e in the $\mu\mu e$ sample to be

$$0.91 \times \frac{89.1}{13714.9} \times 485.2 = 2.9 \text{events}$$

Of course, to determine the full background for this sample we have to include irreducible MC, backgrounds from photon conversions, etc. These calculations are carried out separately for the various S_T and on-Z/off-Z bins.

3.3.2 Background for Taus from Jets

The background for τ's actually consists of "fake" τ's from jets. Tau-identification depends heavily on isolation, which in turn is greatly affected by the total jet activity in the event. Therefore we parametrize the probability for jets faking τ's (f_t) in terms of a variable that characterizes the activity in the event (f_{SB}). We create this parametrization with an on-Z dilepton dataset, and cross check it with an off-Z dilepton dataset.

We define the following regions of relIso:

Isolated: RelIso from 0.0 - 0.15

Side Band: RelIso from 0.15 - 1.0

Other: RelIso greater than 1.0

We can now take ratios of the number of tracks in a given dataset in the various relIso regions. We use the off-Z dilepton dataset with sum-p_T binning as nominal and compare off-Z dilepton dataset with lead jet p_T binning and on-Z dilepton datasets to it later.

- $f_t \equiv \frac{\text{Ntrack Isolated}}{\text{Ntrack Side Band}}$

- $f_{SB} \equiv \frac{\text{Ntrack Side Band}}{\text{Ntrack Side Band + Other}}$

The variable f_{SB} is anti-correlated with activity. In a dataset with higher activity, there will be more tracks with larger relIso, moving the distribution to the right. As the number of tracks in the "Other" region increases, the ratio f_{SB} decreases.

In order to visualize the variation of f_{SB} with event activity, we divide the off-Z dilepton data in bins of total track sumPT in the event. For sumPt up to 100 GeV, the binsize is 10 GeV, and above 100 GeV, the binsize is 20 GeV.

For each bin of sumPt, we calculate the f_t and f_{SB}. Note that points with low f_{SB} correspond to datasets with high activity, and correspondingly low f_t, while high f_{SB} (low activity) yields a higher f_t. In other words, events with less activity are more likely to produce fake τ's, while τ candidates in events with high activity are more likely to fail the isolation requirement in τ-ID.

This exercise is repeated with the off-Z dilepton dataset with lead jet p_T binning and the on-Z dilepton dataset. Although the statistics are poor, the f_t vs. f_{SB} is consistent. This gives us confidence that given any dataset, we can calculate f_{SB}, and then extract f_t from the plot.

In order to predict the number of fake τ's in the 2ℓ+track dataset, we look at the relIso for the 2ℓ sample. Note that at very low values of relIso, tracks in this sample are dominated by real leptons. However, we can calculate f_{SB} from this sample, and use the f_t vs. f_{SB} curve obtained from the off-Z dilepton sample above to get f_t. Then we multiply f_t by the number of tracks in the side band region of the 2ℓ dataset to determine the number of background events in the 2ℓ+track dataset.

Illustration of $\mu\mu + \tau$ Background from $\mu\mu$

We can use these values to predict the background due to jets for tau candidates. Examining the same 190/pb dataset as before, we look a the $\mu\mu$ sample and find that between p_t of 8 GeV/c and 24 GeV/c and SumPt of tracks in the event between 0-60 GeV/c we get

- 890 non-isolated tracks in the Side Band.

- 4961 non-isolated tracks in the Side Band + Other regions

We thus determine $f_{SB} = 0.179$ and the look at the curves to extract $f_t = 0.11$. Multiplying by 890 tracks in the side band region gives a background expectation of 97.9 events in the $\mu\mu+$ 1 τ sample. This is repeated for the various lepton flavors, and track pt and SumPt ranges to arrive at the background prediction.

3.3.3   Conclusion

The backgrounds for the lepton and lepton plus tau samples are calculated from data, parametrized by variables most correlated with the the probability to create backgrounds. The background e, μ are parametrized in terms of heavy flavor content, while the background τ are parametrized by the relative isolation of isolated tracks. We have illustrated the method for a specific subchannel and for one bin in the p_t spectrum.

Chapter 4

Trigger Efficiencies, Selection Efficiencies and Other Corrections to Monte-Carlo Samples

As explained in Chapter 3, many of the standard model backgrounds are taken from Monte-Carlo (MC) samples. There are checks done on these MC samples (shown in the control region plots) that confirm that MC gets kinematical distributions correct. At the same though, there are underlying parameters that might be different in data and the detector simulator that the MC samples go through. In these cases, one needs to make certain corrections to MC to match these parameters in data.

In this chapter, we describe the corrections done to MC. We start out with **trigger efficiencies**. To keep the amount of data gathered at reasonable amounts, the LHC applies filters called **triggers**. The triggers are conditions that an event is required to satisfy before it can be stored on disk. As an example, a trigger could simply require an electron with $p_T > 10$ GeV. Any event with no electrons satisfying this condition would be discarded. In principle, any event with at least one electron satisfying this condition would be accepted. In practice though, the probability of accepting an event satisfying the conditions is less than 100% and is called the **trigger efficiency**. On the other hand, MC samples have no triggers and so contains 100% of the events satisfying the trigger that the data went through. If the trigger efficiency (always measured in data) was 95% and we just blindly used the MC, we would over-estimate the backgrounds by 5%. To correct for this, we scale the MC contributions by 95% in this case.

This discussion is then followed by **selection efficiencies**. This refers to the leptons in our case. Objects in the detector are classified as leptons if they pass two kinds of cuts: identification (ID) cuts that distinguish between electrons, muons and τs, and isolation cuts that ensure that the leptons are well-separated from other objects in the event and ensure that the leptons didn't come from jets. The probability that a real lepton passes the cuts is not 100% and is called the **selection efficiency** for that lepton type. It is a combination of both the **ID efficiency** and the **isolation efficiency**. As with the trigger efficiency, the selection efficiencies could be different in data and MC samples, and would lead to a systematic bias in the background estimation from MC. This is also corrected in MC.

We then discuss **pileup-reweighting**. At the LHC, proton bunches are steered into each other and can lead to more than one simultaneous collision. This is known as **pileup** and the number of simultaneous interactions/collisions in a given event is denoted by N_{vertex} (number of vertices). Pileup affects selection efficiencies as well as $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution and we do corrections to MC to match the pileup distribution with data.

As explained in more detail in Appendix A, we correct MC samples to match $E_{\mathrm{T}}^{\mathbf{miss}}$ **resolution** in data. This analysis also uses **b-jets** (jets coming from b-quarks). The important parameters here are the efficiency: probability that a b-jet passes the cuts for a b-jet and the fake-rate: rate at which non b-jets (either gluon jets or light quark jets) pass the cuts identifying a b-jet. These parameters are different in MC and data and we correct the MC.

4.1   Trigger Efficiencies

In order to estimate the trigger efficiencies, we use the HT primary data set and select events that have a single isolated lepton that passes our selection criteria

and that fires a HT trigger. If we assume that the efficiencies of trigger i, ϵ_i, and trigger j, ϵ_j, are uncorrelated, then the efficiency for an event to satisfy both triggers , ϵ_{ij}, is given by $\epsilon_{ij} = \epsilon_i \times \epsilon_j$. We estimate the trigger efficiency by taking the ratio of the number of events that satisfied both the HT and lepton trigger to the number of events that satisfied the HT trigger. This ratio will be equal to $\epsilon_{ij}/\epsilon_i = \epsilon_j$ if the HT and lepton triggers are uncorrelated. Additionally we require $pfMET > 180$ or $HT > 550$ or $HT > 300$ and $pfMET > 70$ to try and remove trigger correlations or biases between HT and lepton triggers, especially electron triggers.

In data, we determine the efficiency of the "OR" of all selected single electron and single muon triggers in the run range 190456 to 207518 for the p_t range 20 GeV/c to 140 GeV/c. We get the isolated electron trigger efficiency to be 89.9%, and the muon trigger efficiency to be 86.8%. For dielectrons it is, 95.5% and that for dimuons is 89%. Using a single non-lepton trigger, e.g. HLT_HT750_v7, as 'tag' trigger, the electron trigger efficiency is 90%±0.8%and the muon trigger efficiency is 85.8%±0.5%. We have also used electron-muon cross triggers. These cross triggers are important since we may not have plain single lepton triggers for upcoming runs, and even if we do, the thresholds will be extremely high. In keeping with the method mentioned above, for electron-muon cross trigger, we get 89%±1.2% efficiency.

The cuts on the leptons for them to fire the dilepton triggers are looser than the ones for single lepton triggers, including the L1 seeds for the dilepton and single lepton triggers. This is one of the primary reasons why the dilepton trigger efficiencies are larger than, or comparable to single lepton trigger efficiencies . This true especially for electrons. The single electron triggers rely on either L1_SingleEG15 and L1_SingleEG20 where as the dielectron triggers rely on L1_SingleEG12. The single and dilepton trigger efficiencies for all of the HT triggers used in this study are given in table 4.1 and 4.2.
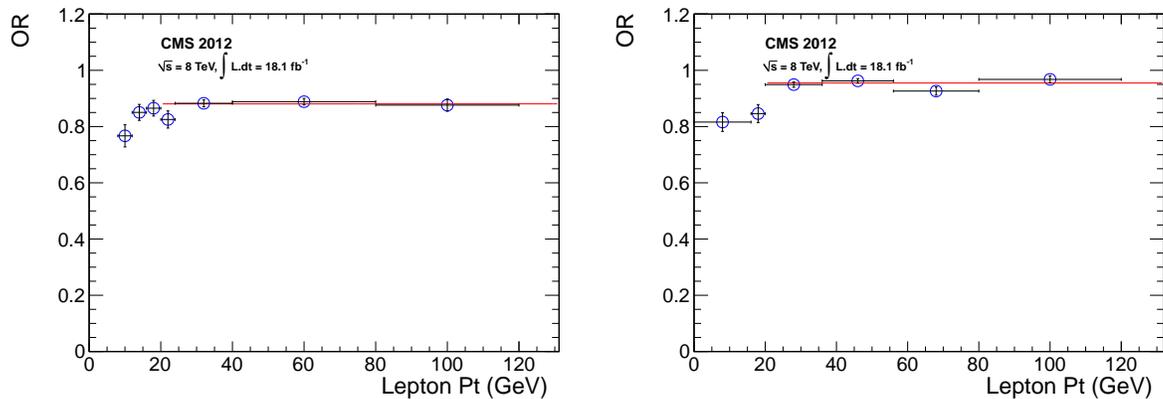
Figure 4.1: , Dimuon "OR" Efficiency (left) and Dielectron "OR" efficiency (right) by the method described in this section.

Tag Trigger	Electron "OR" Efficiency	Muon "OR" Efficiency
HLT_HT650_v3	86.6%±5.7%	86%±2.7%
HLT_HT750_v3	88%±1.5%	86%±0.8%
HLT_HT750_v4	88.4%±3.0%	87%±1.5%
HLT_HT550_v7	88%±4.7%	87%±1.8%
HLT_HT750_v7	90%±0.8%	85.8%±0.5%

Table 4.1: Efficiency for the "OR" of all single electron (muon) triggers determined using different HT triggers to tag events.

Tag Trigger	Dilectron "OR" Efficiency	Dimuon "OR" Efficiency
HLT_HT750_v3	93.6%±1.5%	94%±2%
HLT_HT750_v4	96.7%±2.4%	91%±6.4%
HLT_HT650_v7	95.9%± %	93%±6%
HLT_HT750_v7	96%±0.7%	87%±1.8%

Table 4.2: Efficiency for the "OR" of all dielectron (dimuon) triggers determined using different HT triggers to tag events.

4.2 Lepton Identification and Isolation Efficiencies

We use a **tag and probe** method on $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ events to measure the efficiency of the identification and isolation requirements for electrons and muons. This involves selecting an event with a tight-lepton ("tag") that passes all the lepton cuts for the event and a loose-lepton ("probe") that has only a subset of the lepton cuts applied. We require the tag and probe to have invariant mass in the Z-window to be sure that both are leptons coming from a Z. Then we count the fraction of events where the probe also passes the full lepton cuts. This ratio (parameterized as a function of the p_T and η of the probe) gives the selection efficiency.

For both the electron and muon cases, we require that the tag muon have the same selection cuts as for isolated muons used in our analysis, except that we raise the p_t requirement to $p_t > 20$ GeV/c. For a probe muon, we require only that it be a global muon with $p_t > 5$ GeV and $|\eta| < 2.1$. For a probe electron, we loosen the cut values on $\sigma_{I\eta I\eta}$, $\Delta\phi$, $\Delta\eta$, and HoverE to the WP95 selection values and then further doubled the $\Delta\eta$ cut. We know that the probe requirements are very loose by looking at their distributions for candidate Z's.

The invariant mass distributions of di-muons and di-electrons for three probe p_t bins are shown in figure 4.2 and figure 4.3, respectively, where the Monte Carlo has been normalized to have the same number of events under the Z peak, 80 to 100 GeV/c^2, as the data. For each bin in p_t, we find the number of events on the Z peak before and after applying cuts. In order to remove background under the Z peak, we fit the mass distribution in the range of 55 GeV/c^2 to 125 GeV/c^2 to a linear background distribution plus a variable width Lorentzian line-shape centered at the mass of the Z. We then remove the background in the Z peak range by using the linear component of the fit. We find the lepton identification and isolation efficiencies separately by first determining the probe

selection efficiency applying the identification requirements and then determine the isolation efficiency by applying the isolation requirement to probes that pass the identification selection.

The Monte Carlo models the identification efficiencies to within a few percent throughout the whole p_t range. Figures 4.4 and 4.5 show the identification efficiency and corresponding data to Monte Carlo ratio as a function of probe p_t for muons and electrons, respectively. Additional figures are given in the Appendix showing the muon and electron identification efficiencies as a function of p_t for $|\eta| < 2.1$, $|\eta| < 1.5$ (barrel), and $1.5 < |\eta| < 2.1$ (endcap). For muon ID efficiency $= 0.9925 \pm 0(stat) \pm 0.0021(syst_{BE}) \pm 0.0018(syst_{jet}) \pm 0.0017(syst_{vert})$

For Electron ID efficiency-

- $\sigma_e = 10.6258 \pm 0.3236(stat) \pm 9.582(syst_{BE}) \pm 0.3659(syst_{jet}) \pm 0.429(syst_{vert})$

- $(\epsilon\_inf)_e = 1.012 \pm 0.0002(stat) \pm 0.0116(syst_{BE}) \pm 0.0008(syst_{jet}) \pm 0.0013(syst_{vert})$

- $(\epsilon\_const)_e = 0.7513 \pm 0.0151(stat) \pm 0.1283(syst_{BE}) \pm 0.054(syst_{jet}) \pm 0.0113(sys_{vert})$

The meaning of $\sigma_e, (\epsilon\_inf)_e, (\epsilon\_const)_e$ is explained in the next paragraph.

For the isolation efficiency, the measured efficiency in the range of $p_t < 25$ GeV/c is significantly smaller than in Monte Carlo. Figures 4.6 and 4.7 show the isolation efficiency and corresponding data to Monte Carlo ratio as a function of probe p_t for muons and electrons, respectively. Additional figures are given in the Appendix showing the muon and electron isolation efficiencies as a function of p_t for $|\eta| < 2.1$, $|\eta| < 1.5$ (barrel), and $1.5 < |\eta| < 2.1$ (endcap). In order to model this, we fit the ratio of data and MC isolation efficiencies to the equation below which was proposed in the same sign dilepton analysis for lepton efficiencies, (**?**

).

$$\text{Efficiency}(p_T) = \epsilon_\infty \times \text{Erf}\left(\frac{p_t - C}{\sigma}\right) + \epsilon_C \times \left(1 - \text{Erf}\left(\frac{p_t - C}{\sigma}\right)\right) \qquad (4.1)$$

- ϵ_∞ = Value in plateau region at high momenta

- $C = p_T$ cut for leptons

- ϵ_C = value at $p_T = C$

- σ = Describes the rate of change in value as p_T drops.

Tables 4.4 and 4.6 gives the fit parameter for different regions in η. Although the isolation efficiency may depend on the number of jets(vertices) in the event, we find that variations in the parameters with the number of jets (veritces) is consistent with the statistical uncertainties. We assign systematic errors equal to the difference between the barrel and endcap values, syst_{BE}, systematic equal to difference between the values determined for events with one jet and those with three jets(syst_{Jet}) and another systematic equal to difference between the values determined for events with six to ten vertices and those with eleven to fifteen vertices(syst_{vert})

One more possible concern with the ever growing luminosity is pileup. For lepton $p_t > 25$ GeV/c we need not worry about ID or Isolation efficiencies as the MC agrees well with data. This can be seen from the plots provided in the Appendix, where the lepton efficiency is binned according to the number of good vertices in the event. The resulting parameters are:

- $\sigma_\mu = 11.6361 \pm 0.3416(stat) \pm 2.3697(syst_{BE}) \pm 1.8662(syst_{jet}) \pm 1.7979(syst_{vert})$

- $(\epsilon\_inf)_\mu = 0.9985 \pm 0(stat) \pm 0.002(syst_{BE}) \pm 0.0009(syst_{jet}) \pm 0.0002(syst_{vert})$

- $(\epsilon\_const)_\mu = 0.9324 \pm 0.0039(stat) \pm 0.0371(syst_{BE}) \pm 0.1041(syst_{jet}) \pm 0.0166(syst_{vert})$

- $\sigma_e = 16.4017 \pm 0.5597(stat) \pm 0.5075(syst_{BE}) \pm 1.9723(syst_{jet}) \pm 2.839(syst_{vert})$

- $(\epsilon\_inf)_e = 0.9982 \pm 0.0001(stat) \pm 0.001(syst_{BE}) \pm 0.0004(syst_{jet}) \pm 0.0001(syst_{vert})$

- $(\epsilon\_const)_e = 0.9316 \pm 0.0052(stat) \pm 0.0054(syst_{BE}) \pm 0.015(syst_{jet}) \pm 0.005(sys_{vert})$

Detector	muon ID correction
All	0.9925 ± 0
barrel	0.9922 ± 0.0001
endcap	0.9943 ± 0.0002
0Jet	0.9935 ± 0.0001
1Jet	0.9925 ± 0.0001
2Jet	0.9912 ± 0.0002
3Jet	0.9907 ± 0.0005
1to5Vert	0.9961 ± 0.0006
6to10Vert	0.995 ± 0.0001
11to15Vert	0.9933 ± 0.0001
gt15Vert	0.9906 ± 0.0001
barrel_0Jet	0.9933 ± 0.0001
barrel_1Jet	0.9921 ± 0.0001
barrel_2Jet	0.9907 ± 0.0002
barrel_3Jet	0.9903 ± 0.0005
endcap_0Jet	0.9949 ± 0.0003
endcap_1Jet	0.9943 ± 0.0003
endcap_2Jet	0.9932 ± 0.0005
endcap_3Jet	0.993 ± 0.0012
barrel_1to5Vert	0.9968 ± 0.0006
barrel_6to10Vert	0.9951 ± 0.0002
barrel_11to15Vert	0.993 ± 0.0001
barrel_gt15Vert	0.9899 ± 0.0002
endcap_1to5Vert	0.9933 ± 0.0015
endcap_6to10Vert	0.9953 ± 0.0004
endcap_11to15Vert	0.9948 ± 0.0003
endcap_gt15Vert	0.9935 ± 0.0004

Table 4.3: Fit parameters for ratio of data and MC , for muon identification efficiencies.

Detector	σ_μ	$(\epsilon_\infty)_\mu$	$(\epsilon_C)_\mu$
All	11.6361 ± 0.3416	0.9985 ± 0	0.9324 ± 0.0039
barrel	12.2389 ± 0.3754	0.998 ± 0	0.9157 ± 0.0055
endcap	9.8692 ± 0.7535	1 ± 0.0001	0.9528 ± 0.0063
0Jet	55.4351 ± 26.4702	1.0003 ± 0.0012	0.9957 ± 0.001
1Jet	9.6283 ± 0.771	0.9988 ± 0.0001	0.9561 ± 0.0071
2Jet	7.4864 ± 0.7597	0.9995 ± 0.0002	0.8819 ± 0.0159
3Jet	7.7621 ± 1.7972	0.9997 ± 0.0005	0.852 ± 0.0377
1to5Vert	12.2914 ± 2.6407	0.9993 ± 0.0004	0.9581 ± 0.0215
6to10Vert	12.5855 ± 0.7197	0.9993 ± 0.0001	0.9432 ± 0.0068
11to15Vert	10.7876 ± 0.4942	0.9991 ± 0.0001	0.9266 ± 0.0066
gt15Vert	10.3852 ± 0.7404	0.9987 ± 0.0001	0.94 ± 0.0086
barrel_0Jet	33.5737 ± 9.0701	0.9995 ± 0.0005	0.9917 ± 0.0024
barrel_1Jet	11.0107 ± 0.7944	0.9983 ± 0.0001	0.9378 ± 0.0089
barrel_2Jet	8.3722 ± 0.8184	0.999 ± 0.0002	0.8751 ± 0.0193
barrel_3Jet	9.27 ± 1.5185	0.9991 ± 0.0006	0.8224 ± 0.0379
endcap_0Jet	20.8295 ± 19.3673	0.9993 ± 0.0001	0.9999 ± 0.0025
endcap_1Jet	99.9999 ± 81.0492	0.9999 ± 0.0004	0.9999 ± 0
endcap_2Jet	6.6619 ± 1.2311	1.0012 ± 0.0004	0.8786 ± 0.0252
endcap_3Jet	3.4584 ± 0.8873	1.0005 ± 0.0011	0.8 ± 0.1447
barrel_1to5Vert	12.0254 ± 2.3029	0.9992 ± 0.0004	0.9282 ± 0.0319
barrel_6to10Vert	13.3633 ± 0.9299	0.999 ± 0.0001	0.9411 ± 0.0093
barrel_11to15Vert	11.2392 ± 0.538	0.9987 ± 0.0001	0.9076 ± 0.0093
barrel_gt15Vert	11.1216 ± 0.6682	0.998 ± 0.0001	0.9109 ± 0.0115
endcap_1to5Vert	23.2338 ± 60.4503	0.9992 ± 0.0008	0.9999 ± 0.1686
endcap_6to10Vert	11.3227 ± 1.1681	1.0002 ± 0.0002	0.9468 ± 0.0101
endcap_11to15Vert	9.343 ± 1.1165	1.0003 ± 0.0002	0.9489 ± 0.0105
endcap_gt15Vert	5.1567 ± 2.9421	1.0005 ± 0.0002	0.9605 ± 0.0261

Table 4.4: Fit parameters for ratio of data and MC , for muon isolation efficiencies.

Detector	σ_e	$(\epsilon_\infty)_e$	$(\epsilon_C)_e$
All	10.6258 ± 0.3236	1.012 ± 0.0002	0.7513 ± 0.0151
barrel	0.4687 ± 74.277	1.0083 ± 0.0002	0.8283 ± 0.1572
endcap	10.0507 ± 0.3626	1.0199 ± 0	0.7 ± 0.1306
0Jet	0.0135 ± 71.671	1.0106 ± 0.0004	0.8992 ± 0.0364
1Jet	10.4727 ± 0.5153	1.013 ± 0.0004	0.754 ± 0.0224
2Jet	12.3987 ± 0.9065	1.0126 ± 0.0006	0.8145 ± 0.0281
3Jet	10.8386 ± 0.6977	1.0122 ± 0.0013	0.7 ± 0.0172
1to5Vert	11.1416 ± 1.5819	1.0127 ± 0.0016	0.7533 ± 0.0649
6to10Vert	9.8051 ± 0.7264	1.0116 ± 0.0004	0.777 ± 0.0313
11to15Vert	10.2341 ± 0.5026	1.0129 ± 0.0004	0.7657 ± 0.0234
gt15Vert	10.0286 ± 0.6917	1.0134 ± 0.0005	0.7328 ± 0.0346
barrel_0Jet	9.6754 ± 0.7383	1.0084 ± 0.0004	0.7777 ± 0.039
barrel_1Jet	0.1238 ± 72.7863	1.009 ± 0.0004	0.8284 ± 0.0094
barrel_2Jet	14.1301 ± 1.1131	1.01 ± 0.0007	0.8613 ± 0.0243
barrel_3Jet	12.4544 ± 1.807	1.0091 ± 0.0014	0.811 ± 0.0566
endcap_0Jet	10.9178 ± 0.9642	1.0199 ± 0	0.7552 ± 0.0574
endcap_1Jet	10.1282 ± 0.7978	1.0199 ± 0	0.7146 ± 0.0394
endcap_2Jet	8.5768 ± 1.2695	1.0199 ± 0.0002	0.718 ± 0.1101
endcap_3Jet	11.1761 ± 1.1711	1.0199 ± 0.0007	0.7 ± 0.0085
barrel_1to5Vert	12.1944 ± 2.6278	1.0108 ± 0.0017	0.812 ± 0.0698
barrel_6to10Vert	10.9329 ± 0.79	1.0089 ± 0.0004	0.8186 ± 0.028
barrel_11to15Vert	0.1171 ± 72.7414	1.0091 ± 0.0003	0.8902 ± 0.0104
barrel_gt15Vert	10.4565 ± 0.9144	1.0109 ± 0.0005	0.8013 ± 0.036
endcap_1to5Vert	11.4995 ± 1.5606	1.0199 ± 0.0019	0.7 ± 0.1468
endcap_6to10Vert	7.9054 ± 0.8618	1.0199 ± 0.0001	0.7 ± 0.1488
endcap_11to15Vert	10.2641 ± 0.7102	1.0199 ± 0	0.7207 ± 0.0406
endcap_gt15Vert	8.9149 ± 1.7388	1.0199 ± 0.0001	0.7135 ± 0.1762

Table 4.5: Fit parameters for ratio of data and MC , for electron identification efficiencies.

Detector	σ_e	$(\epsilon_\infty)_e$	$(\epsilon_C)_e$
All	16.4017 ± 0.5597	0.9982 ± 0.0001	0.9316 ± 0.0052
barrel	16.3637 ± 0.6326	0.998 ± 0.0001	0.9286 ± 0.0064
endcap	16.8712 ± 1.096	0.999 ± 0.0002	0.934 ± 0.0087
0Jet	19.4253 ± 1.8304	0.9987 ± 0.0001	0.9723 ± 0.0057
1Jet	15.1912 ± 1.1269	0.9985 ± 0.0001	0.944 ± 0.0086
2Jet	14.1848 ± 1.1502	0.9984 ± 0.0002	0.9174 ± 0.0152
3Jet	17.1635 ± 3.8043	0.9989 ± 0.0009	0.929 ± 0.0322
1to5Vert	16.0531 ± 3.9225	0.9991 ± 0.0007	0.9388 ± 0.032
6to10Vert	13.7765 ± 1.0033	0.9989 ± 0.0001	0.9363 ± 0.0113
11to15Vert	16.6155 ± 0.8405	0.9988 ± 0.0001	0.9313 ± 0.0078
gt15Vert	16.0444 ± 1.0988	0.9981 ± 0.0002	0.9272 ± 0.0112
barrel_0Jet	19.4655 ± 2.345	0.9986 ± 0.0001	0.9745 ± 0.0069
barrel_1Jet	15.3074 ± 1.3	0.9983 ± 0.0001	0.9456 ± 0.01
barrel_2Jet	13.9547 ± 1.2622	0.9983 ± 0.0003	0.9112 ± 0.0185
barrel_3Jet	19.3132 ± 3.7455	0.9991 ± 0.001	0.9283 ± 0.0291
endcap_0Jet	19.5334 ± 2.8515	0.9991 ± 0.0004	0.9656 ± 0.0102
endcap_1Jet	15.1518 ± 2.0229	0.9992 ± 0.0004	0.9298 ± 0.0169
endcap_2Jet	15.2169 ± 2.2253	0.9988 ± 0.0007	0.9203 ± 0.0244
endcap_3Jet	1.7864 ± 94.3294	0.9983 ± 0.0017	0.8124 ± 0.1049
barrel_1to5Vert	16.3334 ± 4.4642	0.9992 ± 0.0007	0.942 ± 0.036
barrel_6to10Vert	13.943 ± 1.3374	0.9989 ± 0.0001	0.9478 ± 0.0131
barrel_11to15Vert	16.4398 ± 0.9971	0.9985 ± 0.0001	0.932 ± 0.0096
barrel_gt15Vert	15.5966 ± 0.9932	0.9978 ± 0.0002	0.9046 ± 0.0139
endcap_1to5Vert	14.1943 ± 7.1898	0.9971 ± 0.0014	0.9366 ± 0.0713
endcap_6to10Vert	13.3371 ± 1.446	0.9989 ± 0.0004	0.8987 ± 0.0217
endcap_11to15Vert	17.7963 ± 1.5166	0.9997 ± 0.0004	0.9272 ± 0.0124
endcap_gt15Vert	3.9169 ± 0.8946	0.9988 ± 0.0004	0.8 ± 0.1276

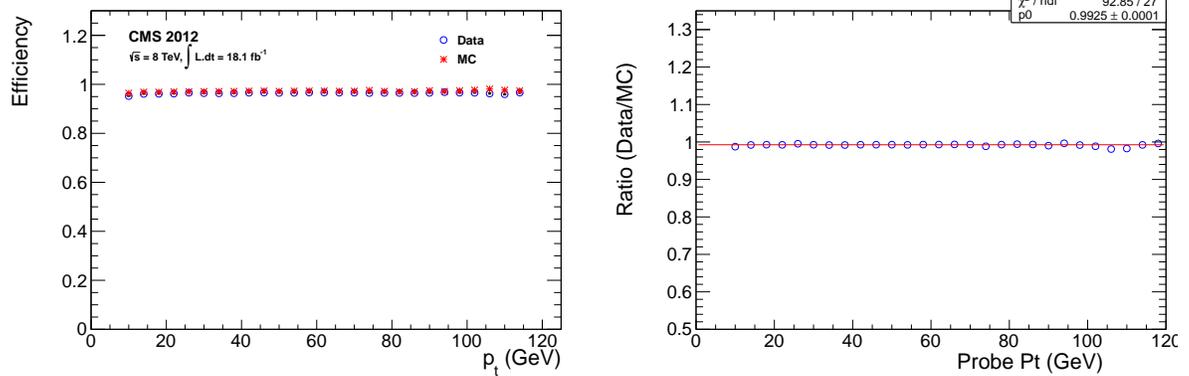Table 4.6: Fit parameters for ratio of data and MC , for electron isolation efficiencies.

Figure 4.4: Muon identification efficiency as a function of probe p_t (left) and ratio of data and MC (right).
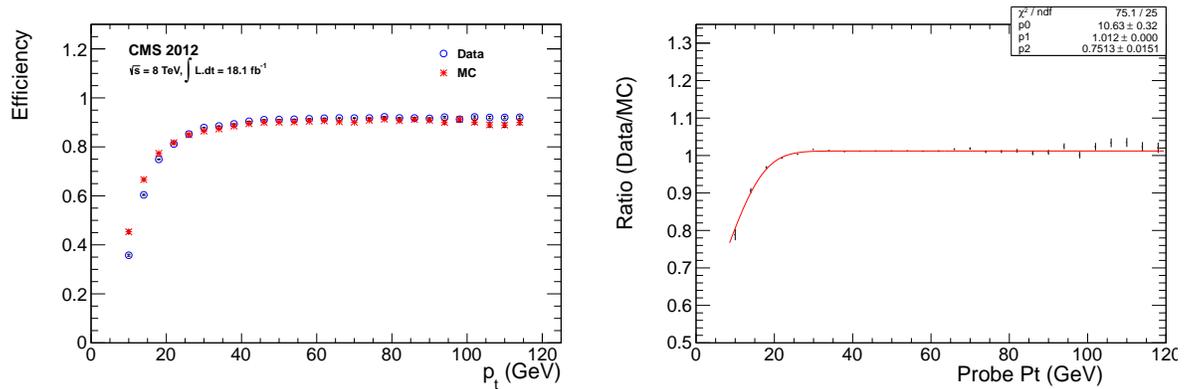

Figure 4.5: Electron identification efficiency as a function of probe p_t (left) and ratio of data and MC (right).
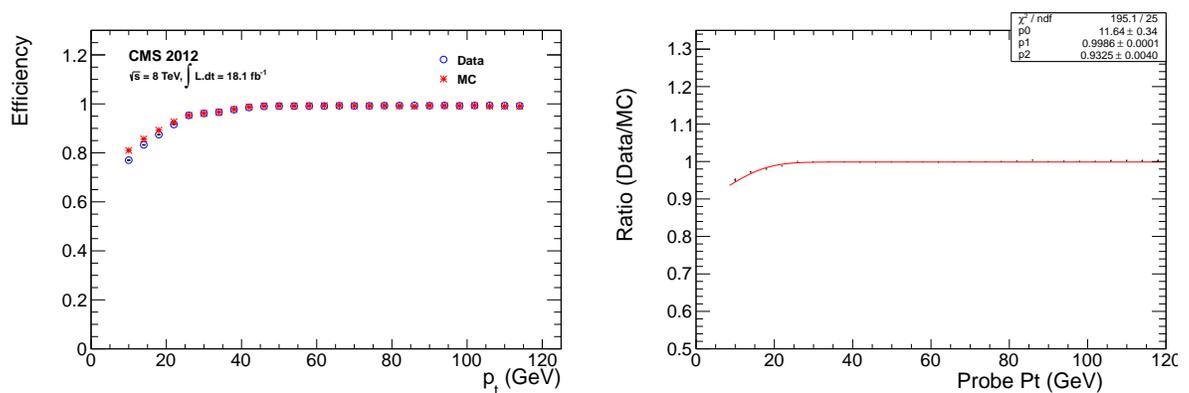

Figure 4.6: Muon isolation efficiency as a function of probe p_t (left) and ratio of data and MC (right).
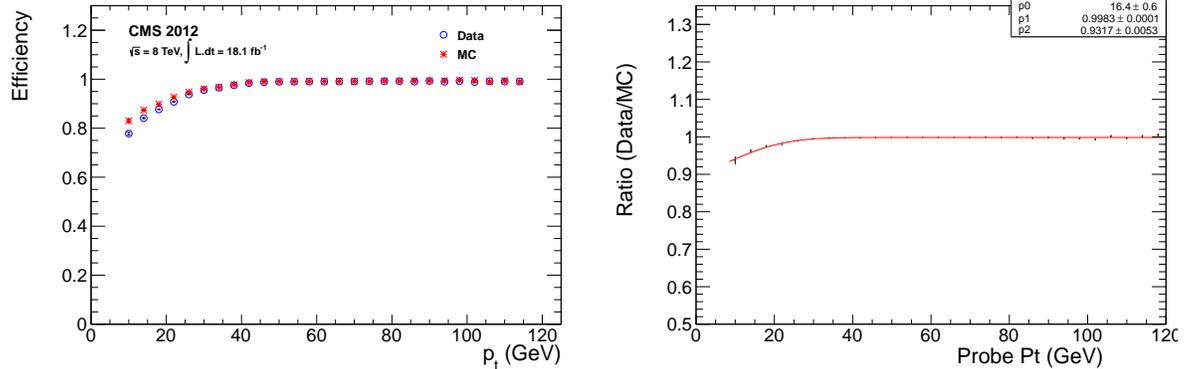
Figure 4.7: Electron isolation efficiency as a function of probe p_t (left) and ratio of data and MC (right).

Obtaining efficiencies for taus has well known difficulties. One of the main obstacles is the missing energy that accompanies tau decay, making it unreliable for Z boson mass reconstruction. Additionally, to see how tau isolation cuts vary withnumber of jets vertices, using $Z \to \tau^+\tau^-$ decays is a poor method because one cuts on the number of jets and vertices in identifying taus. Instead of using the tag and probe method on $Z \to \tau^+\tau^-$ decays, we propose using muons as the tag and probe objects where the probes match tau candidates in p_t, η, and ϕ to within .0001 respectively. We then pass these probe muons through the various tau discriminants (loose,medium,tight) to get our typical efficiency numerator. The results of using muons to get the tau isolation efficiencies is shown in the figures below. More figures showing the dependence of tau isolation on number of jets and vertices are located in the appendix.

For Tau isolation efficiency-

- $\sigma_{ltau} = 12.5016 \pm 0.79475(stat) \pm 2.742(syst_{BE}) \pm 1.15442(syst_{jet}) \pm 1.1089(syst_{vert})$

- $(\epsilon\_inf)_{ltau} = 0.999458 \pm 0.000229813(stat) \pm 0.010013(syst_{BE}) \pm 0.00126(syst_{jet}) \pm 0.000301(syst_{vert})$

- $(\epsilon\_const)_{ltau} = 0.957157 \pm 0.00517409(stat) \pm 0.023456(syst_{BE}) \pm 0.132649(syst_{jet}) \pm 0.007182(syst_{vert})$

- $\sigma_{mtau} = 13.9118 \pm 1.02537(stat) \pm 5.1059(syst_{BE}) \pm 1.68068(syst_{jet}) \pm 0.6844(syst_{vert})$

- $(\epsilon\_inf)_{mtau} = 0.997172 \pm 0.000381543(stat) \pm 0.020346(syst_{BE}) \pm 0.00342(syst_{jet}) \pm 0.002406(syst_{vert})$

- $(\epsilon\_const)_{mtau} = 0.952994 \pm 0.00626046(stat) \pm 0.033473(syst_{BE}) \pm 0.102442(syst_{jet}) \pm 0.002982(sys_{vert})$

- $\sigma_{ttau} = 11.2522 \pm 0.974878(stat) \pm 5.07203(syst_{BE}) \pm 0.35385(syst_{jet}) \pm 26.5816(syst_{vert})$

- $(\epsilon\_inf)_{ttau} = 0.992449 \pm 2.11862e-05(stat) \pm 0.006507(syst_{BE}) \pm 0.002351(syst_{jet}) \pm 0.001252(syst_{vert})$

- $(\epsilon\_const)_{ttau} = 0.951698 \pm 0.0080301(stat) \pm 0.036837(syst_{BE}) \pm 0.159393(syst_{jet}) \pm 0.02909(sys_{vert})$
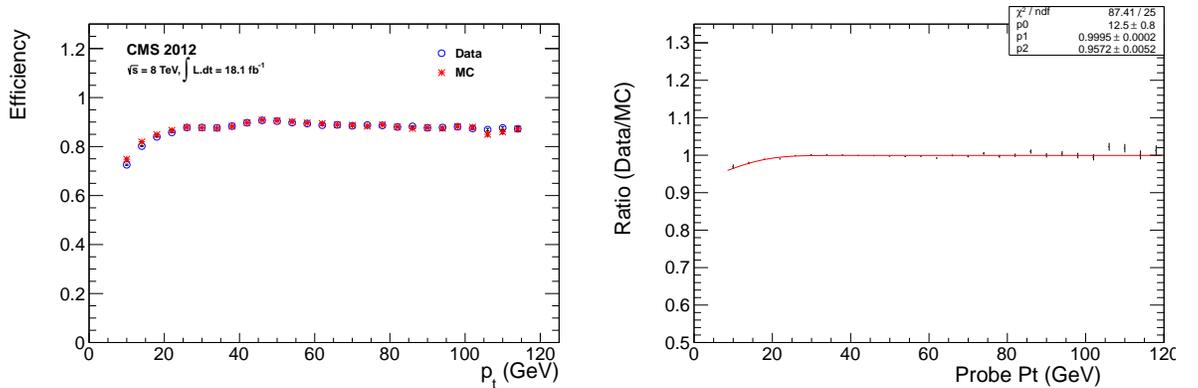

Figure 4.8: Loose Tau isolation efficiency as a function of probe p_t (left) and ratio of data and MC (right).
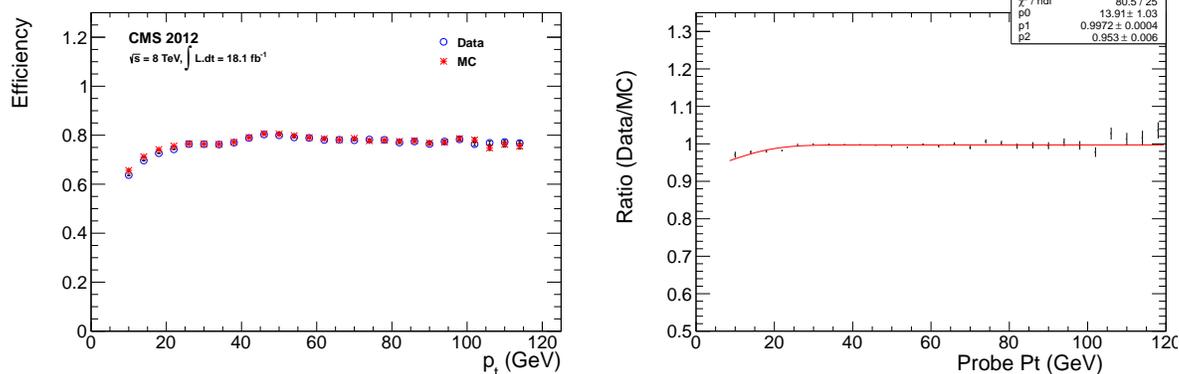
Figure 4.9: Medium Tau isolation efficiency as a function of probe p_t (left) and ratio of data and MC (right).
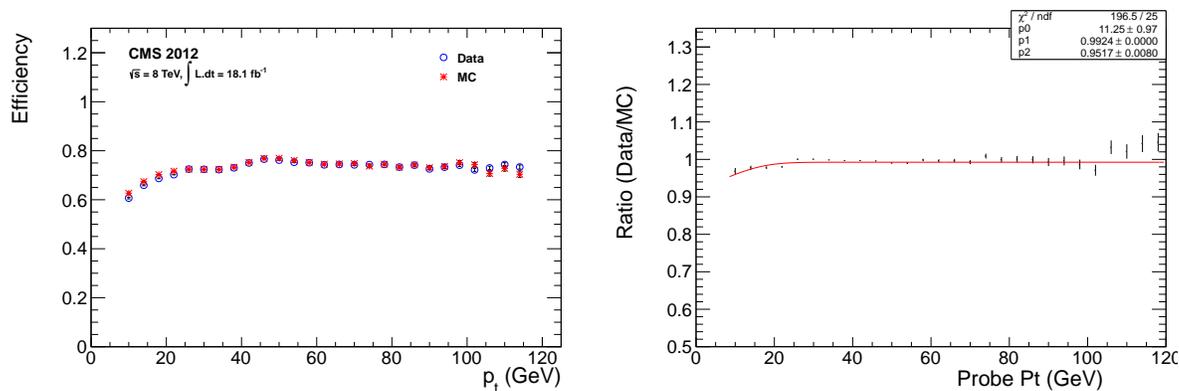

Figure 4.10: Tight Tau isolation efficiency as a function of probe p_t (left) and ratio of data and MC (right).

4.3 Other Corrections

4.3.1 Pileup Reweighting

The number of simultaneous collisions or N_{vertex} is a quantity which depends only on the conditions in the proton beams and should be identical in data and all the Monte-Carlo samples. The Monte-Carlo samples though tend to have different N_{vertex} distributions depending on the generator. Figure 4.11 shows the N_{vertex}

distributions in data and some Monte-Carlo samples.

N_{vertex} affects quantities like lepton efficiencies, $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution, jet activity and total isolation of objects. So it is important to reweigh the N_{vertex} distributions in MC to match data. This is done by re-scaling MC events in each N_{vertex} bin by the fraction of events in that particular bin in data divided by the fraction of events in that bin in MC.

For example, if a $t\bar{t}$ MC event has $N_{vertex} = 20$ and in $t\bar{t}$ MC, 70% events have $N_{vertex} = 20$ while in data 60% events have $N_{vertex} = 20$, we would scale this 1 event by $\frac{0.60}{0.70}$ to get 0.86 events.

One has to be careful though. If the N_{vertex} distributions are very different in data and MC, then these reweight factors can be very large (more than 10). In this case, one might get very events contributing a lot (with large statistical errors) to backgrounds in certain channels. To protect against this, we check to make sure that the N_{vertex} distributions are largely similar in all our samples and that the reweights are never more than 10.

4.3.2 $E_{\mathrm{T}}^{\mathbf{miss}}$ Resolution

Details on $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution modeling and $E_{\mathrm{T}}^{\mathrm{miss}}$ systematics can be found in Chapter A. We give a brief summary of the method, and how it is used to determine a systematic here.

We parameterize the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a function of the number of vertices and the H_T or number of jets in the event. We model the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a sum of Rayleigh distributions given by

$$p(x) = \sum_{ij} W_{ij} \frac{x}{\sigma_{ij}^2} e^{-x^2/2\sigma_{ij}^2}, \qquad (4.2)$$

where "i" represents the number of vertices and "j" is for the number of jets or H_T in bins of 40 GeV. The weight W_{ij} represents the faction of events that have

i vertices and $j \times 40 \ GeV$ of H_T.

We parameterize σ_{ij} in both data and MC. Additional smearing is added to the MC to make it match data, where the amount of smearing is decided on an event by event basis based on the number of vertices and H_T. In order to determine a systematic, we calculate how much the number of events in each bin would change if we added even more smearing to the MC. It should be noted that the resolution systematic is either correlated or anti-correlated between different $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_T bins since there is a conservation of the number of events. If events enter the higher $E_{\mathrm{T}}^{\mathrm{miss}}$ regions they do so by leaving the lower $E_{\mathrm{T}}^{\mathrm{miss}}$ regions, causing an anti-correlation in the systematic uncertainty between the low and high $E_{\mathrm{T}}^{\mathrm{miss}}$ regions.

4.3.3   b-Jet Efficiencies and Fake-rates

The b-tagging efficiency for jets can be different in data and Monte-Carlo (MC) samples. This requires that we correct the MC b-tagging efficiency to make it match data. The b-tagging POG provides official numbers and a recipe for this correction (https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagPOG). In particular, there are fit functions provided for the following:

- efficiencies for generator-level b-jets, c-jets and light jets to get tagged as b-jets. These efficiencies are parameterized either as a function of the jet p_t or are constant for a given tagger.

- scale-factors defined as the ratio of the efficiency in data divided by the efficiency in MC parameterized as a function of the jet p_t.

- Uncertainties are provided for the scale factors which are the same as 2012 but inflated by a constant factor. For certain scale factors, an additional multiplicative correction function is provided which we include.

We apply the corrections in the following way:

For each event in MC, we have the following information (we have access to generator-level information):

- Flavor-composition of all jets that pass our cuts.

- Whether each jet got b-tagged or not.

So, one can build a weight for each event by multiplying weights for each jet:

$$w = \Pi_{i=1}^{N_{jet}} w_i$$

where the weight for each jet w_i is defined as follows:

Efficiency in MC (ϵ_{MC}) is found by the functions described above (and provided by the b-tag POG). The efficiency will generally depend on the p_t of the jet as well as the generator-level type of the jet (b/c/light). Efficiency in data is: $\epsilon_{Data} = SF * \epsilon_{MC}$ where SF is the scale-factor described by functions above and the relationship is just the definition of the scale-factor.

Now, there are two possibilities:

- the jet got tagged as a b-jet (the generator-level information goes into choosing the different functions for efficiencies and scale factors):

$$w_i = \frac{Probability\,to\,get\,tagged\,in\,data}{Probability\,to\,get\,tagged\,in\,MC} = \frac{\epsilon_{Data}}{\epsilon_{MC}} = SF_i$$

- the jet didn't get tagged as a b-jet:

$$w_i = \frac{Probability\,to\,not\,get\,tagged\,in\,data}{Probability\,to\,not\,get\,tagged\,in\,MC} = \frac{(1 - \epsilon_{Data})}{(1 - \epsilon_{MC})} = \frac{(1 - SF_i\epsilon_{MC})}{(1 - \epsilon_{MC})}$$

The product of weights for all jets then gives a cumulative scale factor for the event.

To find the uncertainty on this weight, we vary all the light jet scale-factors up and down while keeping the b/c scale-factors fixed (to account for correlations). Similarly, we vary all the b/c-jet scale-factors up and down while keeping the light jet scale-factors fixed. We then add the deviations from the total weight w in quadrature.
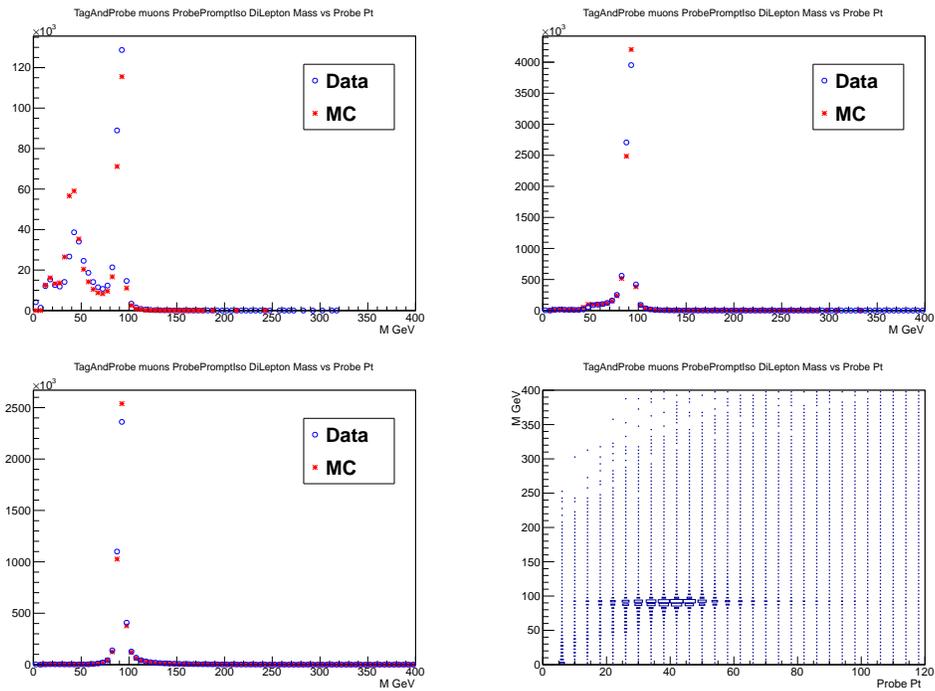
Figure 4.2: Di-muon invariant mass of tag muon and probe muon. Shown is mass for probe p_t from 12-24 GeV (top left), 24-48 GeV (top right), and > 48 GeV (bottom left). The mass versus probe p_t is shown bottom right.
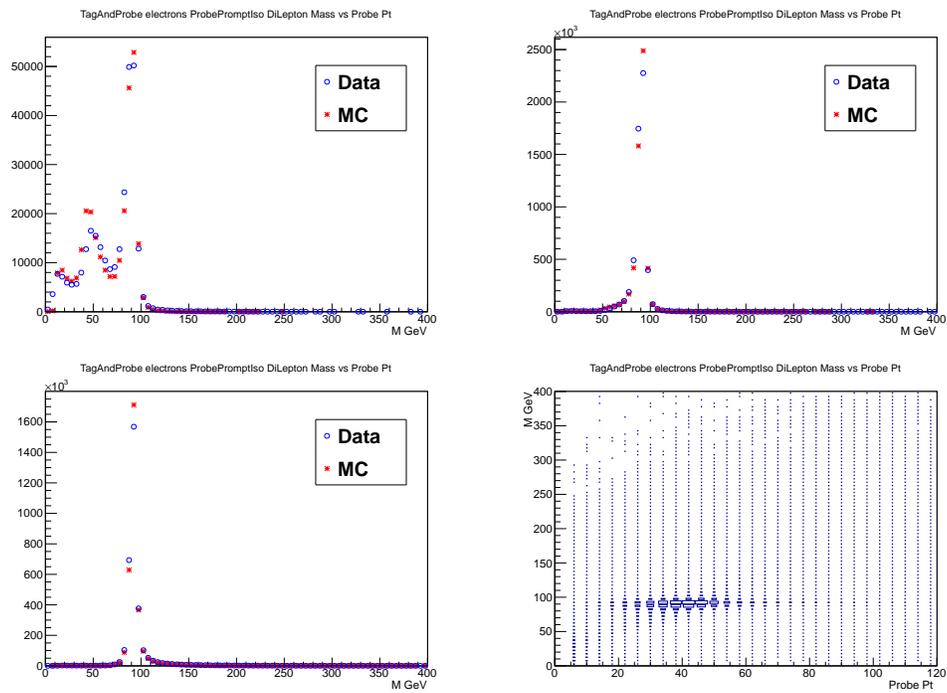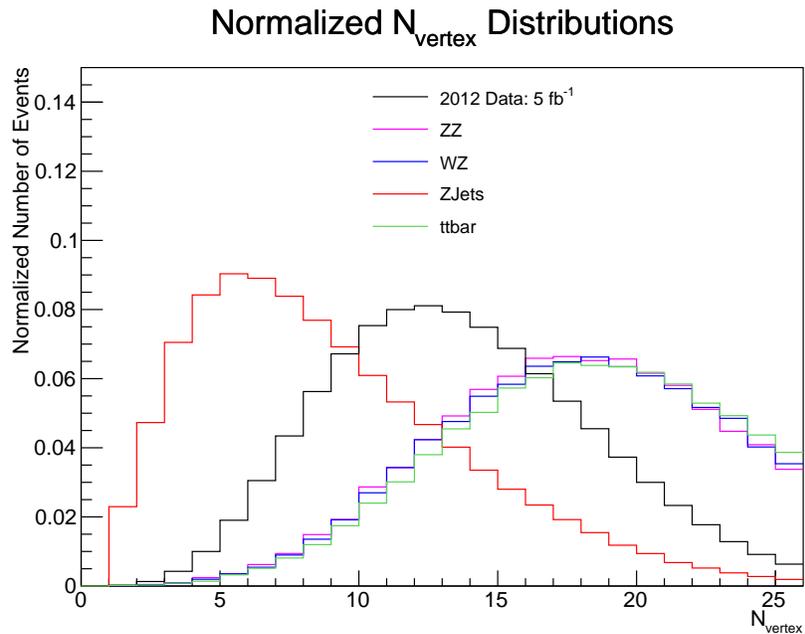
Figure 4.3: Di-electron invariant mass of tag electron and probe electron. Shown is the mass for probe p_t from 12-24 GeV (top left), 24-48 GeV (top right), and $>$ 48 GeV (bottom left). The mass versus probe p_t is shown bottom right.

Figure 4.11: N_{vertex} Distributions of data and MC samples.

Chapter 5

Results: Statistical Techniques, Sources of Systematics and Exclusions

This chapter starts by describing the results i.e. the number of observed events in various exclusive channels and the standard model background in each channel. There is an overall good statistical agreement between observations and backgrounds implying no deviations from the standard model. Then, we describe the statistical procedure used to quantify the likelihood of our signal model ($b'b' \rightarrow$ Multileptons) being consistent with the data. This is followed by a back-of-the-envelope calculation of what b' masses we expect to exclude and the shape of the exclusion contour in the two-dimensional plane of branching ratio of $b' \rightarrow bZ$ vs b' mass. Lastly, we look at the exclusion contour that displays the b' masses and $BF(b' \rightarrow bZ)$ values that are excluded and those that are not.

5.1   Results

Tables 5.1 and 5.2 show the number of observed events and the expected total standard model backgrounds for various exclusive 4-lepton and 3-lepton channels respectively. All the results in this chapter at for 9.2 fb^{-1} of CMS data at $\sqrt{s}$=8 TeV.

4 Lepton Results

N_{OSSF}	onZ	S_T(GeV)	0-τ, 0-b		1-τ, 0-b		0-τ, 1+b		1-τ, 1+b	
			obs	expect	obs	expect	obs	expect	obs	expect
0	-	$S_T > 2000$ GeV	0	0 ± 0.009	0	0 ± 0.009	0	0 ± 0.009	0	0 ± 0.009
0	-	$1500 < S_T < 2000$ GeV	0	0 ± 0.009	0	0 ± 0.009	0	0 ± 0.009	0	0 ± 0.009
0	-	$1000 < S_T < 1500$ GeV	0	0 ± 0.009	0	0 ± 0.009	0	0 ± 0.009	0	0 ± 0.009
0	-	$600 < S_T < 1000$ GeV	0	0 ± 0.009	0	0.01 ± 0.01	0	0.01 ± 0.02	0	0 ± 0.009
0	-	$300 < S_T < 600$ GeV	0	0.009 ± 0.01	0	0.6 ± 0.5	0	0.0007 ± 0.009	0	0.11 ± 0.07
0	-	$0 < S_T < 300$ GeV	0	0.004 ± 0.009	2	0.16 ± 0.08	0	0.0002 ± 0.009	0	0.14 ± 0.09
1	offZ	$S_T > 2000$ GeV	0	0 ± 0.009	0	0 ± 0.009	0	0 ± 0.009	0	0 ± 0.009
1	onZ	$S_T > 2000$ GeV	0	0 ± 0.009	0	0 ± 0.009	0	0 ± 0.009	0	0 ± 0.009
1	offZ	$1500 < S_T < 2000$ GeV	0	0 ± 0.009	0	0.007 ± 0.01	0	0.009 ± 0.01	0	0 ± 0.009
1	onZ	$1500 < S_T < 2000$ GeV	0	0 ± 0.009	0	0.01 ± 0.01	0	0.01 ± 0.01	0	0.001 ± 0.009
1	offZ	$1000 < S_T < 1500$ GeV	0	0.001 ± 0.009	0	0.06 ± 0.03	0	0.06 ± 0.04	0	0.02 ± 0.02
1	onZ	$1000 < S_T < 1500$ GeV	0	0.03 ± 0.02	0	0.05 ± 0.03	0	0.03 ± 0.02	0	0.09 ± 0.05
1	offZ	$600 < S_T < 1000$ GeV	0	0.02 ± 0.02	2	0.15 ± 0.05	0	0.22 ± 0.13	0	0.32 ± 0.14
1	onZ	$600 < S_T < 1000$ GeV	0	0.18 ± 0.06	0	0.7 ± 0.13	0	0.1 ± 0.06	0	0.47 ± 0.21
1	offZ	$300 < S_T < 600$ GeV	0	0.07 ± 0.02	1	0.7 ± 0.15	0	0.47 ± 0.25	1	0.7 ± 0.23
1	onZ	$300 < S_T < 600$ GeV	2	0.6 ± 0.17	5	4.7 ± 0.7	0	0.009 ± 0.01	0	0.19 ± 0.11
1	offZ	$0 < S_T < 300$ GeV	1	0.17 ± 0.05	9	4 ± 1.2	0	0.02 ± 0.02	0	0.37 ± 0.17
1	onZ	$0 < S_T < 300$ GeV	0	1.2 ± 0.38	18	18 ± 5.2	2	0 ± 0.009	2	0 ± 0
2	offZ	$S_T > 2000$ GeV	0	0 ± 0.009	0	0 ± 0	0	0.01 ± 0.01	0	0 ± 0
2	onZ	$S_T > 2000$ GeV	0	0.001 ± 0.009	0	0 ± 0	0	0 ± 0.009	0	0 ± 0
2	offZ	$1500 < S_T < 2000$ GeV	0	0 ± 0.009	0	0 ± 0	0	0.002 ± 0.009	0	0 ± 0
2	onZ	$1500 < S_T < 2000$ GeV	0	0.02 ± 0.01	0	0 ± 0	0	0 ± 0.009	0	0 ± 0
2	offZ	$1000 < S_T < 1500$ GeV	0	0.004 ± 0.01	0	0 ± 0	0	0.04 ± 0.02	0	0 ± 0
2	onZ	$1000 < S_T < 1500$ GeV	0	0.27 ± 0.06	0	0 ± 0	0	0.04 ± 0.02	0	0 ± 0
2	offZ	$600 < S_T < 1000$ GeV	0	0.04 ± 0.01	0	0 ± 0	0	0.45 ± 0.14	0	0 ± 0
2	onZ	$600 < S_T < 1000$ GeV	1	2.6 ± 0.5	0	0 ± 0	1	0.1 ± 0.06	0	0 ± 0
2	offZ	$300 < S_T < 600$ GeV	1	0.46 ± 0.1	0	0 ± 0	1	1.4 ± 0.39	0	0 ± 0
2	onZ	$300 < S_T < 600$ GeV	10	19 ± 3.8	0	0 ± 0	2	0.07 ± 0.03	0	0 ± 0
2	offZ	$0 < S_T < 300$ GeV	4	3.4 ± 0.9	0	0 ± 0	0	0.07 ± 0.03	0	0 ± 0
2	onZ	$0 < S_T < 300$ GeV	68	56 ± 13	0	0 ± 0	1	0.44 ± 0.12	0	0 ± 0
Total4	All	All	87	84 ± 19	37	29 ± 6.9	7	3.6 ± 1.1	3	2.5 ± 0.7

Table 5.1: Observed yields for four lepton events from $9.2 fb^{-1}$ recorded in 2012. The channels are broken down by the number of and mass of any opposite-sign, same-flavor pairs (whether on or off Z), whether the leptons include taus, whether there are any b-jets present and the S_T. Expected yields are the sum of simulation and data-driven estimates of backgrounds in each channel. The channels are exclusive.

N_{OSSF}	onZ	S_T(GeV)	0-τ, 0-b obs	0-τ, 0-b expect	1-τ, 0-b obs	1-τ, 0-b expect	0-τ, 1+b obs	0-τ, 1+b expect	1-τ, 1+b obs	1-τ, 1+b expect
3 Lepton Results										
0	-	$S_T > 2000$ GeV	0	0 ± 0.009	0	0 ± 0.2	0	0 ± 0.01	0	0 ± 0.2
0	-	$1500 < S_T < 2000$ GeV	0	0.01 ± 0.01	0	0.003 ± 0.2	0	0 ± 0.01	0	0.5 ± 0.48
0	-	$1000 < S_T < 1500$ GeV	0	0.07 ± 0.03	0	0.4 ± 0.22	0	0.6 ± 0.5	2	1.3 ± 0.9
0	-	$600 < S_T < 1000$ GeV	2	2.1 ± 1.2	17	9 ± 3.5	1	3.3 ± 1.6	23	20 ± 10
0	-	$300 < S_T < 600$ GeV	14	13 ± 5.7	129	134 ± 53	20	16 ± 6.5	206	186 ± 98
0	-	$0 < S_T < 300$ GeV	30	37 ± 10	555	581 ± 130	22	13 ± 5.9	150	150 ± 72
1	$m_{\ell+\ell-} > 105$ GeV	$S_T > 2000$ GeV	0	0.0005 ± 0.01	0	0 ± 0.2	0	0 ± 0.03	0	0 ± 0.2
1	$m_{\ell+\ell-} < 75$ GeV	$S_T > 2000$ GeV	0	0.002 ± 0.01	0	0 ± 0.2	0	0 ± 0.03	0	0 ± 0.2
1	onZ	$S_T > 2000$ GeV	0	0.12 ± 0.04	0	0.005 ± 0.2	0	0.01 ± 0.04	0	0 ± 0.2
1	$m_{\ell+\ell-} > 105$ GeV	$1500 < S_T < 2000$ GeV	0	0.08 ± 0.04	0	0.2 ± 0.2	0	0.06 ± 0.04	0	0.05 ± 0.05
1	$m_{\ell+\ell-} < 75$ GeV	$1500 < S_T < 2000$ GeV	1	0.02 ± 0.03	0	0 ± 0.2	0	0.06 ± 0.04	0	0 ± 0.2
1	onZ	$1500 < S_T < 2000$ GeV	2	0.5 ± 0.28	0	0.12 ± 0.08	0	0.11 ± 0.07	0	0.07 ± 0.05
1	$m_{\ell+\ell-} > 105$ GeV	$1000 < S_T < 1500$ GeV	0	0.46 ± 0.11	0	0.6 ± 0.28	0	0.15 ± 0.07	1	0.9 ± 0.6
1	$m_{\ell+\ell-} < 75$ GeV	$1000 < S_T < 1500$ GeV	0	0.41 ± 0.08	0	0.2 ± 0.12	0	0.16 ± 0.08	0	0.6 ± 0.6
1	onZ	$1000 < S_T < 1500$ GeV	6	7.6 ± 1.3	3	2.4 ± 0.5	1	1.6 ± 0.43	1	0.8 ± 0.6
1	$m_{\ell+\ell-} > 105$ GeV	$600 < S_T < 1000$ GeV	6	5.2 ± 1.2	12	8.5 ± 2.6	3	3.9 ± 1.5	13	9.8 ± 5.4
1	$m_{\ell+\ell-} < 75$ GeV	$600 < S_T < 1000$ GeV	2	4.7 ± 0.9	11	6.8 ± 2.5	0	3.3 ± 1.1	5	5.1 ± 2.8
1	onZ	$600 < S_T < 1000$ GeV	42	56 ± 7.6	48	35 ± 7.2	7	10 ± 2.7	10	6.5 ± 1.9
1	$m_{\ell+\ell-} > 105$ GeV	$300 < S_T < 600$ GeV	34	31 ± 5.3	149	170 ± 39	12	17 ± 6.1	80	73 ± 35
1	$m_{\ell+\ell-} < 75$ GeV	$300 < S_T < 600$ GeV	34	38 ± 6	139	128 ± 29	26	23 ± 9	87	81 ± 35
1	onZ	$300 < S_T < 600$ GeV	314	356 ± 45	1023	1219 ± 290	63	44 ± 8.1	131	132 ± 31
1	$m_{\ell+\ell-} > 105$ GeV	$0 < S_T < 300$ GeV	81	97 ± 9.5	799	761 ± 182	11	11 ± 4.6	50	41 ± 17
1	$m_{\ell+\ell-} < 75$ GeV	$0 < S_T < 300$ GeV	308	325 ± 36	4933	4208 ± 1033	31	35 ± 13	146	129 ± 38
1	onZ	$0 < S_T < 300$ GeV	2054	2260 ± 213	24078	22191 ± 5517	57	67 ± 9.3	391	369 ± 87
Total3	All	All	2930	3239 ± 308	31896	29460 ± 7204	254	252 ± 59	1296	1211 ± 351

Table 5.2: Observed yields for three lepton events from $9.2 fb^{-1}$ recorded in 2012. The channels are broken down the number of and mass of any opposite-sign, same-flavor pairs (whether on or off Z), whether the leptons include taus, whether there are any b-jets present and the S_T. Expected yields are the sum of simulation and data-driven estimates of backgrounds in each channel. The channels are exclusive.

5.2 Statistical Procedure

This section describes methods for setting statistical limits for various models with multiple channels. We use the LandS tool to compute limits with LHC-type CL_s with 3000 toys as recommended for CMS analyses. This computation yields the observed limit as well as the expected limit with one- and two-sigma uncertainty bands. We thank both the LandS developers for providing this very useful tool.

The input datacards for LandS describe the channels used for limit setting along with the number of observed events, signal yield and background events in those channels. For each channel, nuisance parameters are defined to describe the systematic signal uncertainty, the systematic background uncertainty as well as the statistical uncertainties in background and signal. While systematic uncertainties in many cases are correlated across channels, statistical uncertainties are not. The datacard is produced in a way such that these correlations are taken into account properly.

In order to speed up the limit computation, channels with no expected signal are removed. The channels are then ordered according to the expected single-channel r-value (low to high) where the r-value is defined to be $\frac{\sigma(95\% excluded)}{\sigma_{theory}}$, and only the first 40 channels in this list are used to make the datacards.

When a two-dimensional parameter range is probed (such as a region in the $b' \to bZ$ branching ratio and the b' mass plane), the $r = 1$ contour in the parameter plane is computed. To do so, we use a standard algorithm called Delaunay triangulation between the grid points to obtain a continuous r-value landscape without gaps, in which we then determine the $r = 1$ contour with high resolution. The resulting curve is then smoothed such that misleading edges that are only due to the parameter bin size are removed.

5.3 Sources of Systematic Uncertainties

This section gives more detail on how the systematic uncertainties are calculated. In general systematic uncertainties are found by weighing events up or down or smearing them, then propagating those changes into the various bins of the analysis. The change in the expected backgrounds or yields in each bin corresponds to a systematic uncertainty, where we keep track of the relative sign of changes between different bins in order to keep track of correlations and anti-correlations.

$E_{\mathrm{T}}^{\mathrm{miss}}$ Resolution Systematic

Details on $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution modeling and $E_{\mathrm{T}}^{\mathrm{miss}}$ systematics can be found in Appendix A. We give a brief summary of the method, and how it is used to determine a systematic here.

We parameterize the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a function of the number of vertices and the H_T or number of jets in the event. We model the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a sum of Rayleigh distributions given by

$$p(x) = \sum_{ij} W_{ij} \frac{x}{\sigma_{ij}^2} e^{-x^2/2\sigma_{ij}^2}, \tag{5.1}$$

where "i" represents the number of vertices and "j" is for the number of jets or H_T in bins of 40 GeV. The weight W_{ij} represents the faction of events that have i vertices and $j \times 40\ GeV$ of H_T.

We parameterize σ_{ij} in both data and MC. Additional smearing is added to the MC to make it match data, where the amount of smearing is decided on an event by event basis based on the number of vertices and H_T. In order to determine a systematic, we calculate how much the number of events in each bin would change if we added even more smearing to the MC. It should be noted that the resolution systematic is either correlated or anti-correlated between different $E_{\mathrm{T}}^{\mathrm{miss}}$ bins since there is a conservation of the number of events. If events enter

the higher E_T^miss regions they do so by leaving the lower E_T^miss regions, causing an anti-correlation in the systematic uncertainty between the low and high E_T^miss regions.

Jet Energy Scale Systematic

In order to determine a systematic on the jet energy scale we simultaneously move up or down each jet in the event by the uncertainty on the jet p_T. Doing this, it is possible for jets with p_T above the cutoff (30 GeV) to propagate below the cutoff, taking them out of the jet count, the H_T calculation, or out of the b-tag count. Also, jets with p_T below 30 GeV are allowed to propagate up above the cutoff, which discretely adds 30 GeV to the H_T and an additional jet for the b-tag veto to consider. We additionally propagate variations in the jet energy scale to the E_T^miss and check how that affects the efficiency for passing a E_T^miss cut. For each bin, we take the change that caused the largest change in the expected number of events in each bin.

Lepton Efficiency Systematic

Please see Chapter 4 for the results of tag and probe studies of the lepton identification and isolation efficiency. The isolation efficiency does not perfectly match between the data and the MC. The disagreement between data and MC is larger at low p_T values. We parameterize the ratio between data and MC. We determine systematic uncertainties to these parameters by seeing how they vary between barrel and endcap and between the different numbers of jets in the event. The MC is corrected by weighing events by this ratio for each reconstructed lepton. We determine a systematic by varying the parameters and adding the changes in the weights in quadrature. Both the systematics and the importance of the correction increase with decreasing p_T.

B-jet weights and Systematics

The b-tagging efficiency for jets can be different in data and MC samples. This requires that we correct the MC b-tagging efficiency to make it match data. The b-tagging POG provides official numbers and a recipe for this correction (https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagPOG). In particular, there are fit functions provided for the following:

- efficiencies for generator-level b-jets, c-jets and light jets to get tagged as b-jets. These efficiencies are parameterized either as a function of the jet p_t or are constant for a given tagger.

- scale-factors defined as the ratio of the efficiency in data divided by the efficiency in MC parameterized as a function of the jet p_t.

- Uncertainties are provided for the scale factors which are the same as 2012 but inflated by a constant factor. For certain scale factors, an additional multiplicative correction function is provided which we include.

We apply the corrections in the following way:

For each event in MC, we have the following information (we have access to generator-level information):

- Flavor-composition of all jets that pass our cuts.

- Whether each jet got b-tagged or not.

So, one can build a weight for each event by multiplying weights for each jet:

$$w = \Pi_{i=1}^{N_{jet}} w_i$$

where the weight for each jet w_i is defined as follows:

Efficiency in MC (ϵ_{MC}) is found by the functions described above (and provided by the b-tag POG). The efficiency will generally depend on the p_t of the jet as well as the generator-level type of the jet (b/c/light). Efficiency in data is: $\epsilon_{Data} = SF * \epsilon_{MC}$ where SF is the scale-factor described by functions above and the relationship is just the definition of the scale-factor.

Now, there are two possibilities:

- the jet got tagged as a b-jet (the generator-level information goes into choosing the different functions for efficiencies and scale factors):

$$w_i = \frac{Probability\,to\,get\,tagged\,in\,data}{Probability\,to\,get\,tagged\,in\,MC} = \frac{\epsilon_{Data}}{\epsilon_{MC}} = SF_i$$

- the jet didn't get tagged as a b-jet:

$$w_i = \frac{Probability\,to\,not\,get\,tagged\,in\,data}{Probability\,to\,not\,get\,tagged\,in\,MC} = \frac{(1 - \epsilon_{Data})}{(1 - \epsilon_{MC})} = \frac{(1 - SF_i\epsilon_{MC})}{(1 - \epsilon_{MC})}$$

The product of weights for all jets then gives a cumulative scale factor for the event.

To find the uncertainty on this weight, we vary all the light jet scale-factors up and down while keeping the b/c scale-factors fixed (to account for correlations). Similarly, we vary all the b/c-jet scale-factors up and down while keeping the light jet scale-factors fixed. We then add the deviations from the total weight w in quadrature.

5.3.1 Summary of Systematics

Table 5.3 summarizes uncertainties on background estimates due to various systematics.

Source of Uncertainty	Uncertainty
Luminosity	4.5%
PDF	14%
$E_{\mathrm{T}}^{\mathrm{miss}}$ Res ($E_{\mathrm{T}}^{\mathrm{miss}}$): 0-50 GeV, 50-100 GeV, > 100 GeV	(-3%, +4%, +4%)
Jet Energy Scale $W^{\pm}Z$	0.5% (WZ)
B-Tag Veto (CSVM)	0.1% (WZ), 6% ($t\bar{t}$)
Muon ID/Isolation at 10 (100) GeV	11% (0.2%)
Electron ID/Isolation at 10 (100) GeV	14 % (0.6%)
$t\bar{t}$ xsec/fake rate	50%
WZ xsec	6%
ZZ xsec	12%

Table 5.3: The systematic uncertainties associated with this analysis. The $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution systematic is given for WZ background on Z for different cuts on $E_{\mathrm{T}}^{\mathrm{miss}}$ and for different cuts on M_T given a cut of $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV.

5.4 Exclusion Limits

We apply our analysis to get exclusion limits for pair-produced b's that decay to 3 and 4-leptons as a function of the branching fraction $BF(b' \to bZ) = 1 - BF(b' \to tW)$ and the b' mass.

5.4.1 Back-of-the-envelope Calculation

To get a rough idea of what the exclusion curve should look like, we look at the branching ratios into 3 and 4 lepton channels as a function of $\alpha \equiv BF(b' \to bZ)$. Table 5.4 shows the branching ratios to 3 and 4 lepton events for $b'b' \to bZbZ$, $b'b' \to bZtW$ and $b'b' \to tWtW$. Furthermore, the branching ratio to $bZbZ$ is α^2, to $bZtW$ is $2\alpha(1-\alpha)$ and to $tWtW$ is $(1-\alpha)^2$. So we can get a parametric equation for the total branching ratio to 3/4 lepton channels when $BF(b' \to bZ) = \alpha$ given by:

$$f(\alpha) = (0.36\%)\alpha^2 + (2.65\%)2\alpha(1 - \alpha) + (5.1\%)(1 - \alpha)^2$$

$\frac{f(0)}{f(1)}$ gives the ratio between the $tWtW$ and the $bZbZ$ mode and comes out to approximately 14. If the cross-sections for $b'b'$ production drop by a factor of 2 for

Event Type	4-lepton BF	3-lepton BF	Total
bZbZ	$(6\%)(6\%) = 0.36\%$	0%	0.36%
bZtW	$(25\%)(25\%)(6\%) = 0.375\%$	$2(25\%)(75\%)(6\%) = 2.25\%$	2.65%
tWtW	$(25\%)^4 = 0.4\%$	$4(25\%)^3(75\%) = 4.7\%$	5.1%

Table 5.4: Branching ratios to 3 and 4 lepton channels for different $b'b'$ decays.

every 50 GeV increase in the b' mass, we expect a difference in the limits between $tWtW$ and $bZbZ$ of $(50GeV) * \log_2 14 = 190$ GeV. Please note that this is a rough back-of-the-envelope calculation and gives only an estimate of how different the two limits should be. One can also get a shape of the contour (quadratic in α) and it is shown in figure 5.1. The mass of 625 GeV on the $bZbZ$ side is put in by hand after looking at the actual exclusion but the shape and width of the curve come from the above argument. In short, as the bZ mode becomes dominant, the 3-lepton channel gets choked off and we get worse limits. This calculation is the most optimistic one in that it assumes that all efficiencies are 100%.

In figure 5.4, we present the exclusion plot with the y-axis being the branching fraction $b' \to bZ$ and the x-axis being the b' mass. The exclusion has the same general shape as figure 5.1 and a difference in the two end-point limits of 100 GeV.

5.4.2 Sensitivity of Channels to Different Branching Ratios

Unlike an analysis looking at just one channel for the decay modes $b'b' \to bZbZ$ or $b'b' \to tWtW$, we look at many exclusive channels. As the branching ratio for $b' \to bZ$ varies adiabatically, the signal yield shifts to different channels. Figure 5.2 shows the number of observed events, background breakdown and signal yields in S_T bins for events with 4 electrons and muons, with two opposite-sign same-flavor (OSSF) pairs where at least one pair is on-Z. There are no taus

Rough Estimate of the Shape of Exclusion Curve
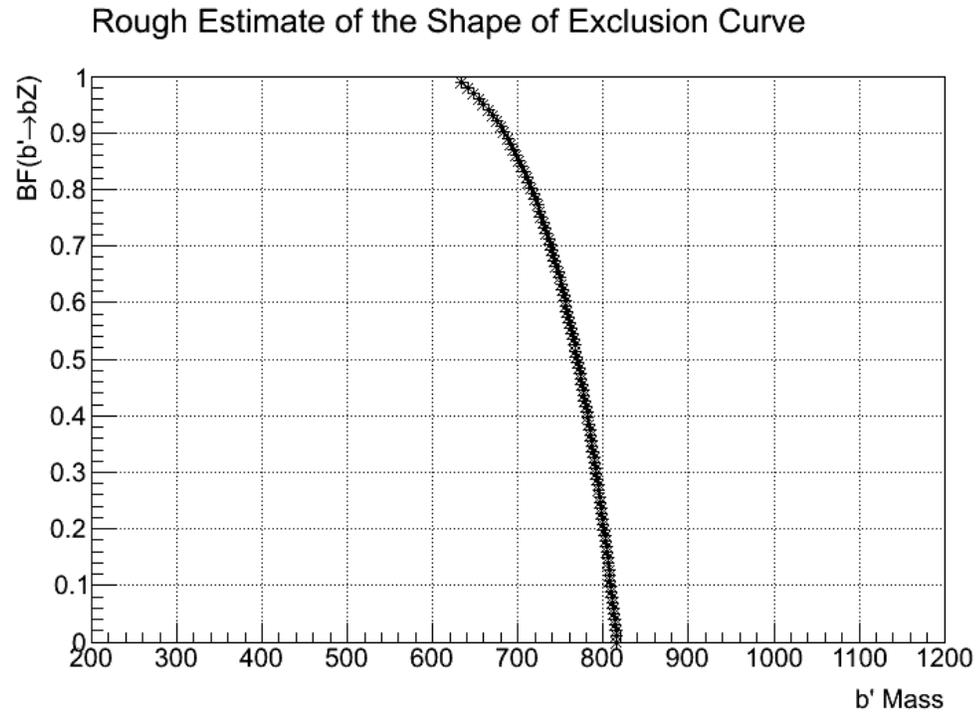
Figure 5.1: A curve based on the branching ratio argument in the text. Note that the 625 GeVlimit on the $bZbZ$ axis is put in by hand after looking at the actual exclusion but the arguments gets the shape and width of the curve in the right ballpark. As the $bZbZ$ mode becomes more dominant, the 3-lepton channel gets choked off and the limits get worse.
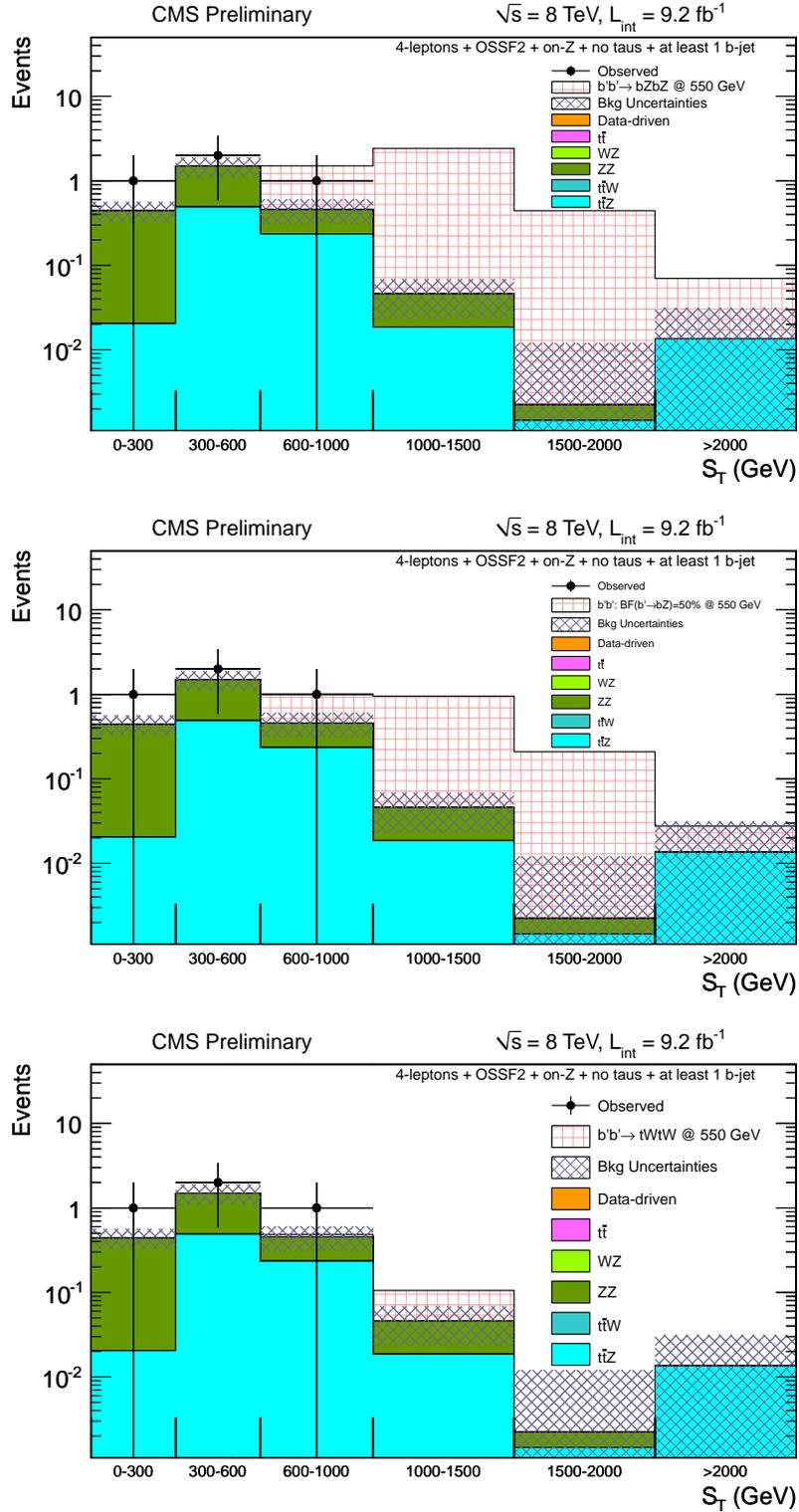
Figure 5.2: S_T distributions in events with 2 opposite-sign same-flavor pairs of leptons with at least one on-Z, no taus and at least 1 b-jet showing expected signal yield in the case where $BR(b' \to bZ) = 100\%$ (top), $BR(b' \to bZ) = 50\%$ (middle), and $BR(b' \to bZ) = 0\%$ (bottom).
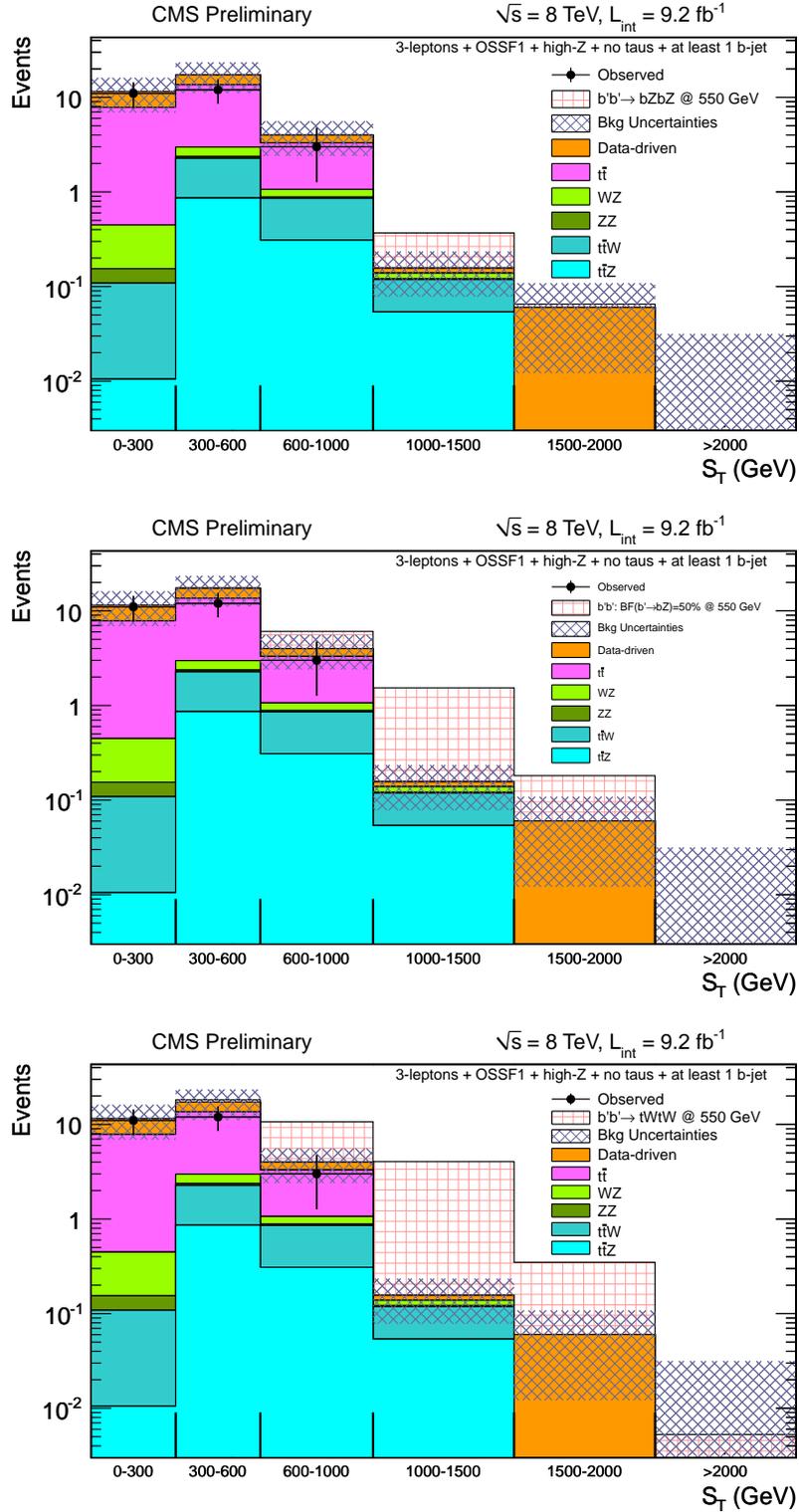
Figure 5.3: S_T distributions in events with 3 leptons with one opposite-sign same-flavor pair that is above the Z-mass, no taus and at least 1 b-jet showing expected signal yield in the case where $BR(b' \to bZ) = 100\%$ (top), $BR(b' \to bZ) = 50\%$ (middle), and $BR(b' \to bZ) = 0\%$ (bottom).

but at least 1 b-jet in these events. In other words, these channels are sensitive to the presence of b-jets and two Zs, i.e. $bZbZ$. So, if the branching ratio for $b' \rightarrow$ is 100%, we expect most of the signal to lie in these channels. As the branching ratio decreases though, the middle plot shows decreasing signal yield till the branching ratio is completely to $b' \rightarrow tW$ in which case almost no signal lies in these channels (bottom plot).

Figure 5.3 shows channels that are sensitive to the other end of the spectrum i.e. $b'b \rightarrow tWtW$. In this case, we expect to see 3-lepton events with one OSSF pair that is off-Z. We show events where the OSSF pair is above the Z-mass, with no taus and at least 1 b-jet. The top figure shows signal where $BR(b' \rightarrow bZ) = 100\%$ and we see almost no signal but as the branching ratio decreases, we see even more signal in the middle and bottom plots.

These two sets of plots demonstrate the merit of doing an analysis with several exclusive channels. As the branching ratio to $bZbZ$ decreases, while the 4-lepton channels lose sensitivity, the 3-lepton channels pick up the signal. This enables one to look at more realistic scenarios/models rather than the artificial case of pure $bZbZ$ or pure $tWtW$ decays.

5.4.3 Exclusion Contour

Figures 5.5, 5.6, and 5.7 show the exclusion plots overlaid on the expected cross-section, the observed cross-section and the theory cross-sections respectively.

Figure 5.8 shows the 1-dimensional exclusion curve for $BR(b' \rightarrow bZ) = 0$ or in other words when both b's decay to tW. Figure 5.9 shows the corresponding curve for $BR(b' \rightarrow bZ) = 50\%$ and figure 5.10 for $BR(b' \rightarrow bZ) = 100\%$.

Figure 5.8 has the theory cross-section line (blue) crossing the observed limit line at 760 GeV. Figure 5.10 has the theory cross-section line crossing the observed limit line at 660 GeV. Lastly, figure 5.9 has the lines crossing at 720 GeV. For each of these branching ratios scenarios, masses below the quoted values are
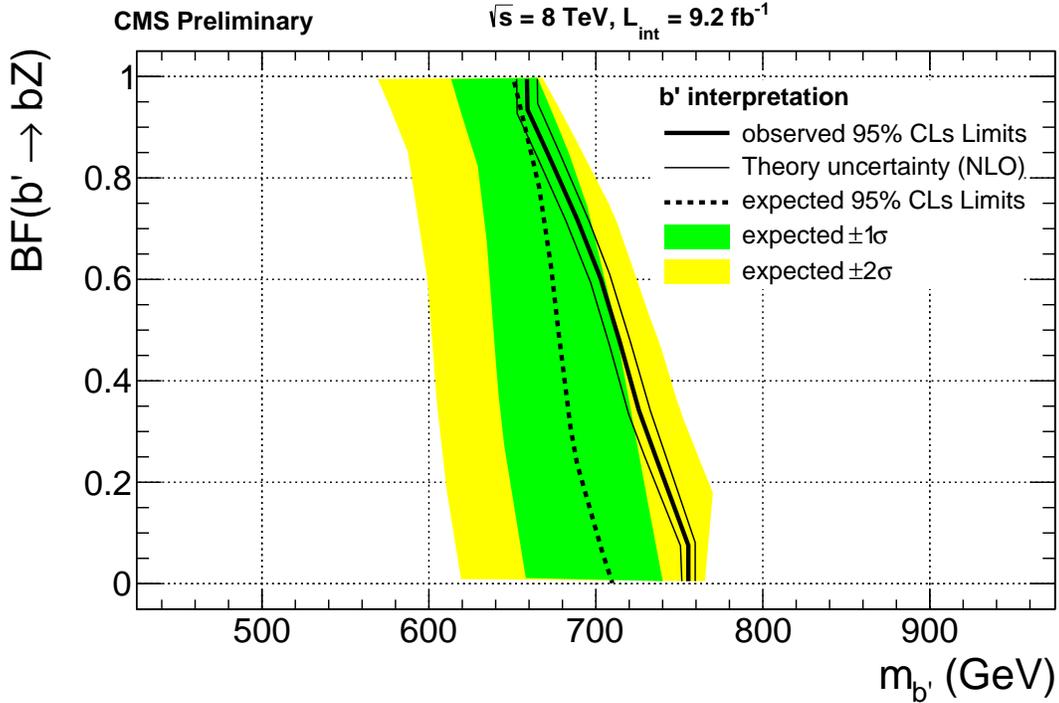
Figure 5.4: Exclusion limits for pair-produced b's going to multileptons in the two-dimensional plane of branching fraction of b' to bZ vs. b' mass. Signal points to the left of the curve are excluded. The y = 0 axis corresponds to $b'b' \to tWtW$ and the y=1 axis to $b'b' \to bZbZ$. The limits vary from 625 GeVto 730 GeVas the branching ratio is varied.

excluded while masses above that could possibly exist. Several such 1D curves are converted into a 2D plane shown in 5.4. In the 2D plot, all masses and branching ratio combinations to the left of the observed curve are excluded while the ones to the right might possibly exist.
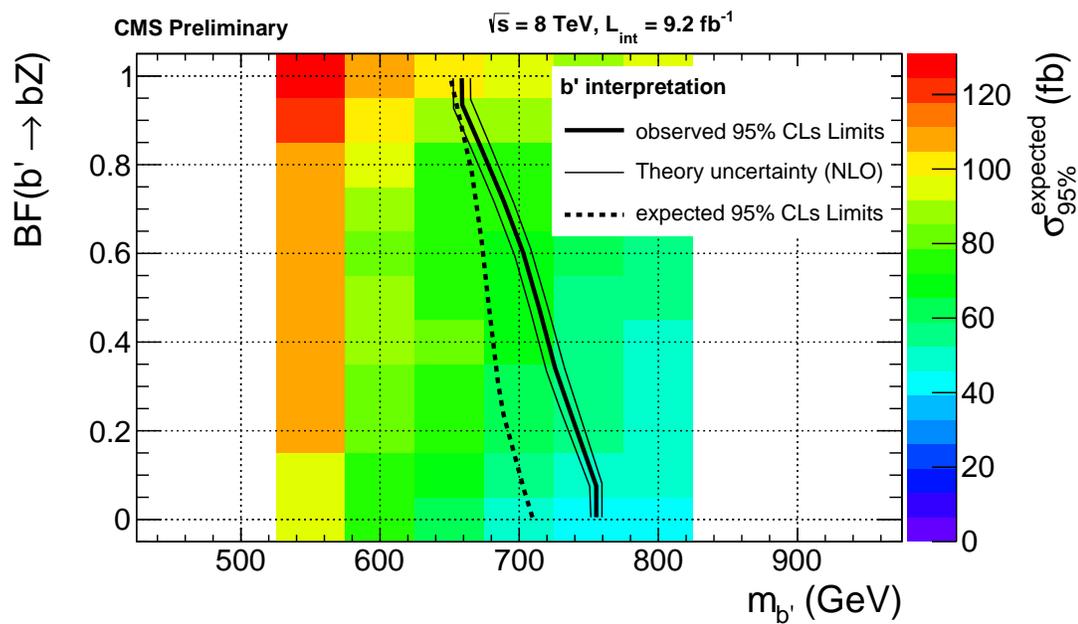
Figure 5.5: Exclusion curve for Bprime with expected 95% UL on cross section overlay.
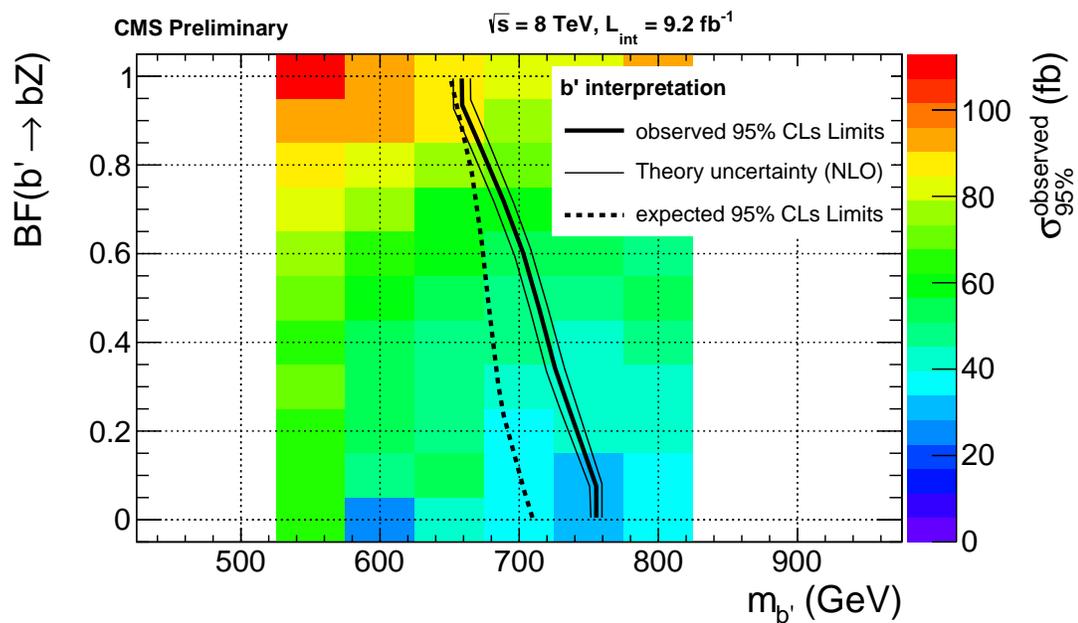
Figure 5.6: Exclusion curve for Bprime with observed 95% UL on cross section overlay.
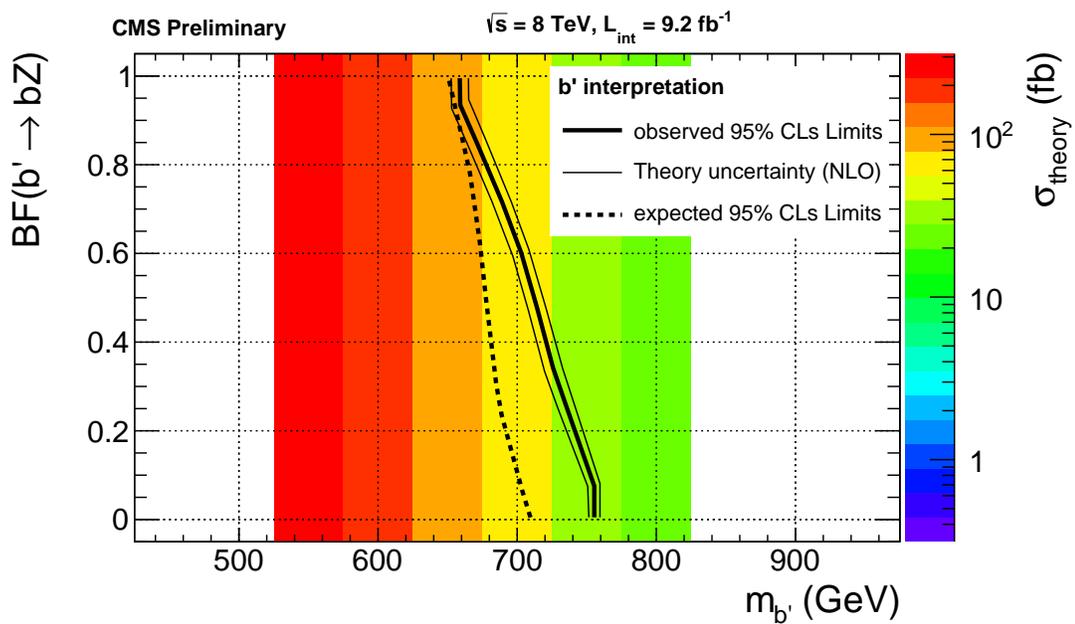
Figure 5.7: Exclusion curve for Bprime with theory cross section overlay.
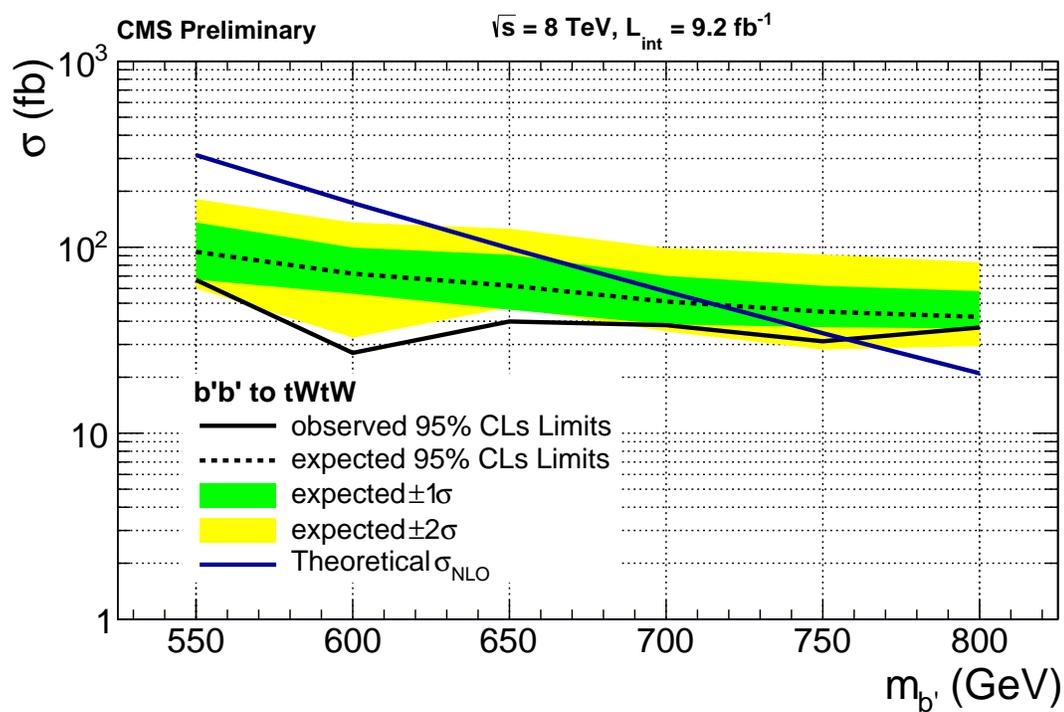
Figure 5.8: Exclusion curve for Bprime vs b' mass for $b'b' \to tWtW$.
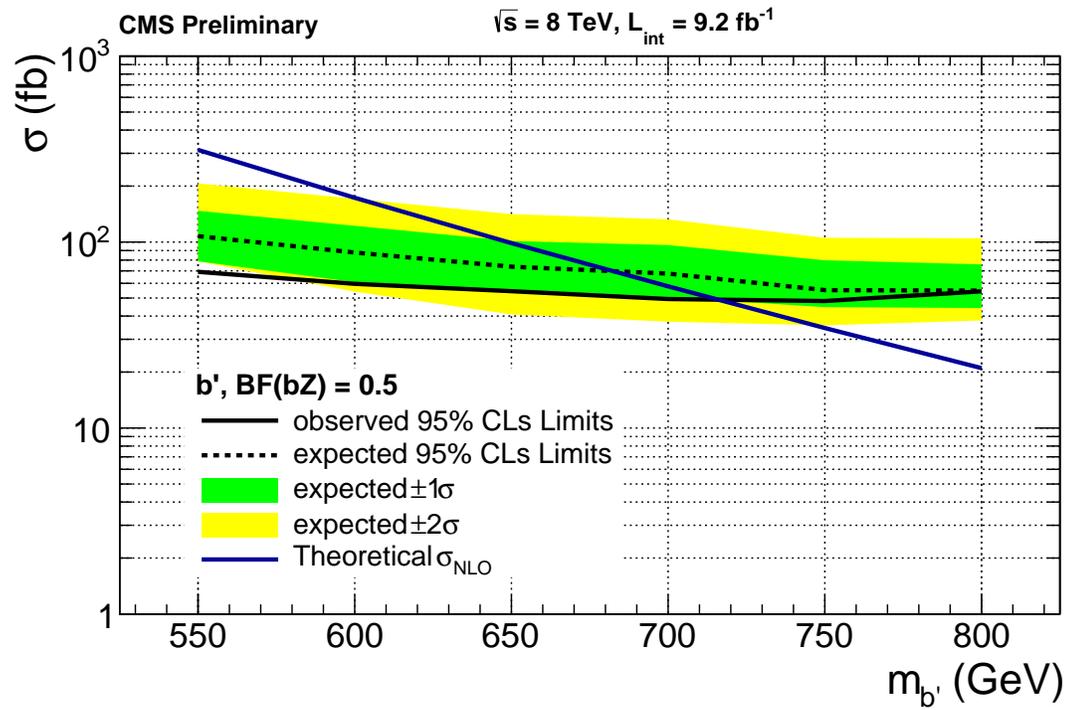
Figure 5.9: Exclusion curve for Bprime with $BR(b' \to bZ) = 50\%$.
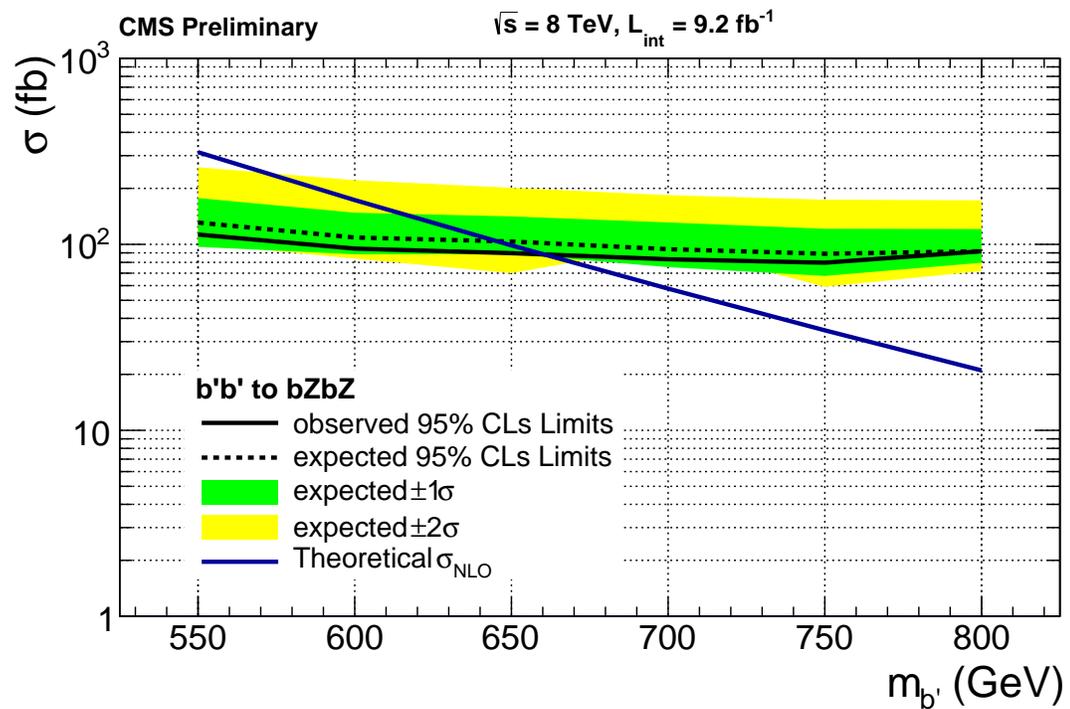
Figure 5.10: Exclusion curve for Bprime vs b' mass for $b'b' \to bZbZ$.

Chapter 6

Conclusions and Extensions

We carried out a search for exotic fourth-generation b' quarks in $\sqrt{s}$=8 TeV CMS 2012 data where the b' could decay to one of two final-states: b+Z and t+W. The search was carried out in multilepton ($\geq$3 leptons) events and was parameterized as a function of the b' mass and the branching ratio of $b' \to bZ$. No evidence for new physics was discovered and we set 95% upper-limits on the model.

The search was carried out in multiple exclusive channels and was the first one to look at the case where $BR(b' \to bZ)$ was varied adiabatically. As the branching ratio changes, different channels become sensitive to the signal as shown in the previous chapter (Chapter 5).

Backgrounds in this analysis were estimated using either Monte-Carlo (MC) samples corrected for pileup, trigger efficiencies, lepton efficiencies, b-jet efficiencies and E_T^{miss} resolution or using data-driven methods for electrons/muons coming from b-jets, jets faking taus, and asymmetric conversion of photon into a hard and soft lepton where the hard lepton gets lost. A new method was introduced for estimating the amount of E_T^{miss} smearing (or resolution) due to pileup and jet activity. This method will become more important in 2015 when the LHC restarts at $\sqrt{s}$=13 TeV with much higher instantaneous luminosities and thus higher pileup. We also observed contributions from rare standard model processes like $t\bar{t}$ W and $t\bar{t}$ Z.

6.1 Extensions of this Analysis

There are two important extensions of this analysis:

- Add the decay mode: $b' \to bH$

- Add dilepton channels with same-sign and opposite-sign leptons.

6.1.1 Addition of decay mode: $b' \to bH$

With the discovery of a Higgs-like scalar particle in the summer of 2012, it is important to search for new processes that could potentially decay to a Higgs. In addition to the decay modes, $b' \to bZ$ and $b' \to tW$, we now consider $b' \to bH$. This leads to six distinct event types: $bZbZ$, $tWtW$, $bHbH$, $bZtW$, $bZbH$, and $tWbH$. The top decays to $b + W$ so we end up getting the following decays: $(bb) + ZZ$, $(bb) + WWWW$, $(bb) + HH$, $(bb) + WWZ$, $(bb) + ZH$, $(bb) + WWH$.

For a Higgs at 120 GeV, the dominant decays modes are: bb (64.8%), WW (14.8%), gg (8.8%), $\tau\tau$ (7.1%), cc (3.0%), ZZ (1.6%). Since we are looking for 3 or more leptons, each event requires any H present to decay to leptons and so the dominant mode is $H \to WW$. In this case, the decay topology is exactly the same as $b' \to tW \to bWW$ since $b' \to bH \to bWW$. The second most-dominant mode is $H \to \tau\tau$ since we also look at channels with at most one hadronic-τ lepton. The last mode of interest is $H \to ZZ$ but it gets severely suppressed by the small branching ratio.

We carried out an analysis similar to the one describes in this thesis where we bin the data in multiple exclusive channels and present results in several slices of the 3-dimensional branching ratio parameter space including the cases where $BR(bZ) = 1$, $BR(tW) = 1$, $BR(bH) = 1$, $BR(bZ) = 0$ (tW-bH plane), $BR(tW) = 0$ (bZ-bH plane), $BR(bH) = 0$ (bZ-tW plane) and lastly, various plots where $BR(bH)$ is fixed a some value and the exclusion is presented in the $bZ - tW$

plane.

When $BF(b' \to bH) \neq 0$, we could use the fact that the event topology for $b'b' \to bHbH \to bbWWWW$ is exactly the same as $b'b' \to tWtW \to bbWWWW$ when $H \to WW$ which is the dominant relevant decay with a branching fraction of 15%. This results in an effective cross-section $\sigma(b'b' \to bHbH \to bbWWWW) = 0.15^2\sigma(b'b' \to tWtW \to bbWWWW)$. Assuming that b' cross-sections drop by a factor of 2 for every 50 GeVincrease in mass, this would cause the limit for $bHbH$ to be about 100 GeVworse than the $tWtW$ limit.

Figures 6.4, 6.5, 6.6 show one-dimensional exclusions plots for $b'b' \to bZbZ$, $b'b' \to tWtW$, and $b'b' \to bHbH$ respectively.

Figures 6.1 and 6.2 summarize the expected and observed exclusion contours for all possible branching ratio combinations.

Figure 6.3 shows the exclusion contours with $BR(b' \to bH) = 0$ and also shows 1-sigma and 2-sigma bands on the expected limits.

Differences between m_{Higgs}=120 GeVand m_{Higgs}=125 GeV

The Higgs mass used in the analysis presnted here is 120 GeVwhile the measured Higgs mass is around 125 GeV. Note that the extension to the bH mode presented here is preliminary and the Higgs mass will be updated to 125 GeVin the final result. We argue below that having a mass of 120 GeVresults is more conservative exclusions.

The main Higgs decay mode contributing to multileptons is $H \to WW$. The branching ratio to this mode is 14% for a mass of 120 GeVand 21% for a mass of 125 GeV. For the decay $b'b' \to bHbH$, this will result in a signal efficiency lower by a factor of about 50% ($\frac{0.14*0.14}{0.21*0.21}$). Since the $b'b'$ production cross-sections
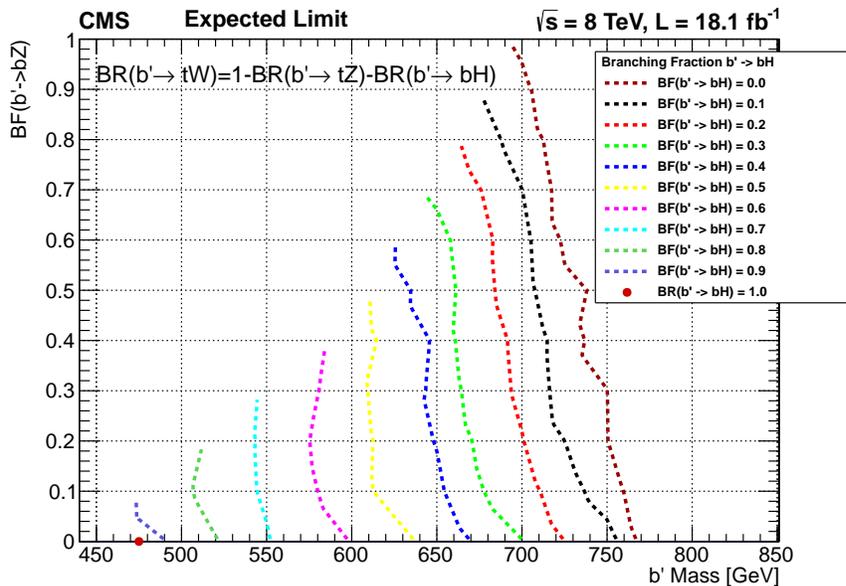
Figure 6.1: Expected exclusion curves as a function of branching ratios. The x-axis is the b' mass, the y-axis is $BR(b' \to bZ)$ and the various curves represent fixed $BR(b' \to bH)$. As $BR(b' \to bH)$ goes up, the acceptance to 3 and 4-lepton events goes down because of the $H \to WW$ branching ratio of 15% which makes the limits worse.

fall by a factor of 2 for every 50 GeVincrease in b' mass, our exclusions are more conservative (compared to a mass of 125 GeV) by around 50 GeV.

6.1.2 Addition of dilepton channels

This is on-going work with CMS collaborators from Taiwan and Korea. Adding in both same-sign dilepton and opposite-sign dilepton channels will increase sensitivity since there is significant branching ratio to dileptons in all the cases: $b'b' \to bZbZ$, $b'b' \to tWtW$, and $b'b' \to bHbH$.
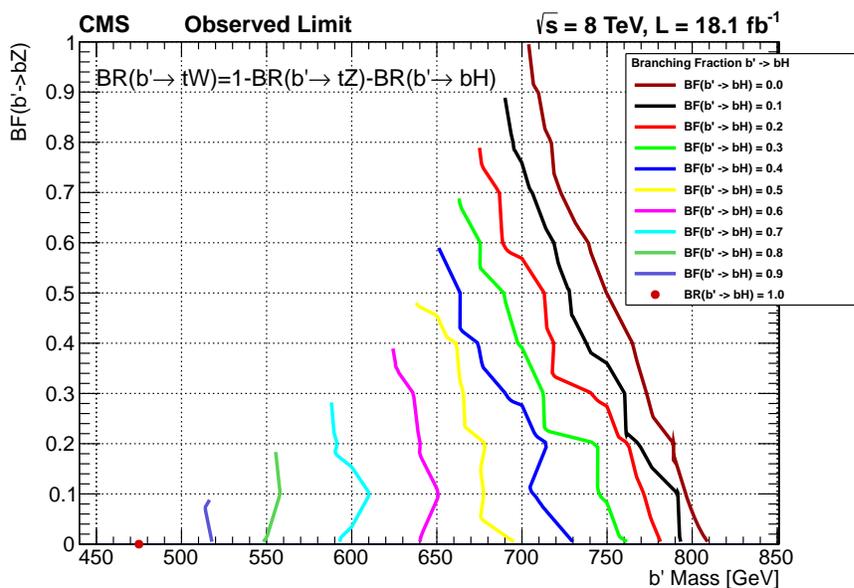
Figure 6.2: Observed exclusion curves as a function of branching ratios. The x-axis is the b' mass, the y-axis is $BR(b' \to bZ)$ and the various curves represent fixed $BR(b' \to bH)$. As $BR(b' \to bH)$ goes up, the acceptance to 3 and 4-lepton events goes down because of the $H \to WW$ branching ratio of 15% which makes the limits worse.
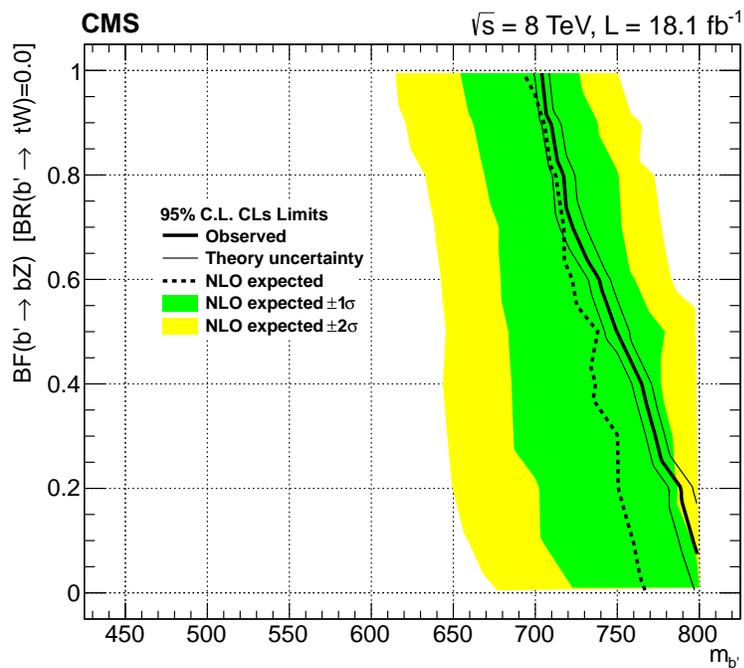
Figure 6.3: Exclusion limits for pair-produced b's going to multileptons in the two-dimensional plane of branching fraction of b' to bZ vs. b' mass. Signal points to the left of the curve are excluded. The y = 0 axis corresponds to $b'b' \to tWtW$ and the y=1 axis to $b'b' \to bZbZ$. The branching ratio for $b' \to bH$ is set to zero.
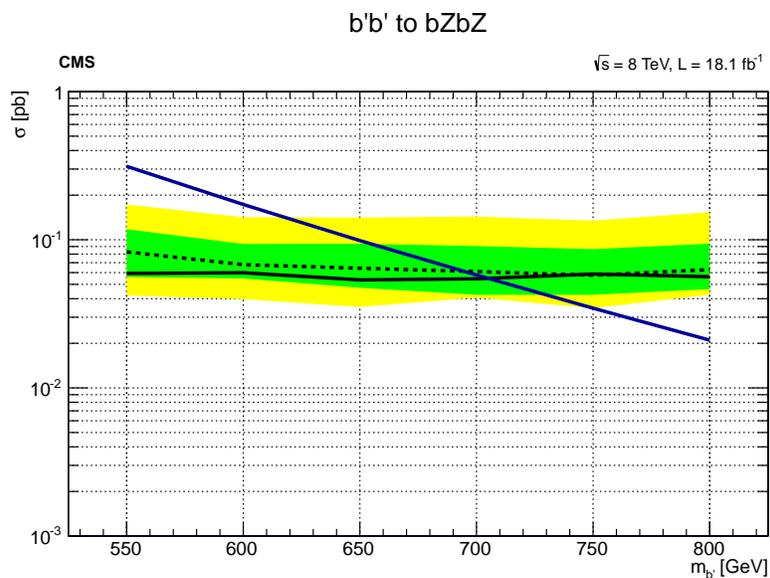
Figure 6.4: Exclusion curve for Bprime vs b' mass for $b'b' \to bZbZ$.
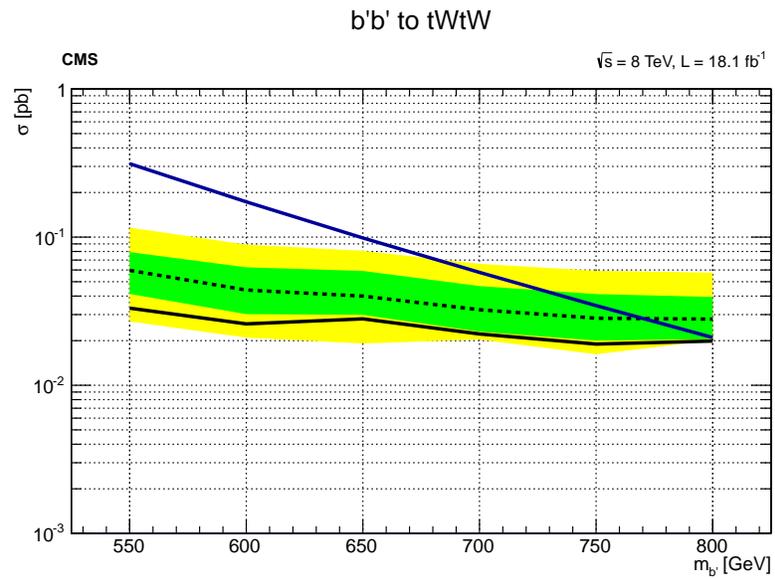
Figure 6.5: Exclusion curve for Bprime vs b' mass for $b'b' \to tWtW$.
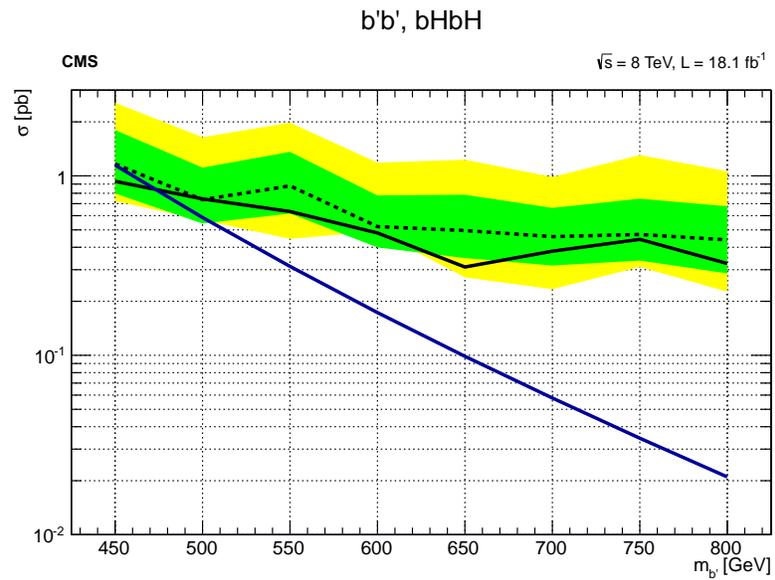
Figure 6.6: Exclusion curve for Bprime vs b' mass for $b'b' \to bHbH$.

Appendix A
MET Resolution Dependence on Pileup and Jet Activity

Many models of new physics hypothesize particles that don't interact via electromagnetic or strong interactions (heavier neutrinos, lightest supersymmetric particles etc.). These would interact very weakly with detectors and would thus not be identifiable using the various detector sub-systems at CMS. Their presence can be inferred though through momentum conservation in the two-dimensional transverse plane. Since the initial protons and thus the quarks travel along the beam-axis, the net momentum in the transverse direction is zero. After adding all the decay products from an event, any remnant momentum in the transverse plane is balanced by a vector-quantity called missing transverse momentum. The magnitude of this two-dimensional vector is called missing transverse energy or MET for short.

Since MET is often used to distinguish standard model processes from signals of new physics, it is crucial to get an accurate measurement of this quantity. At the same time, since MET is inferred from the sum of momenta of all visible decay products in an event, it is very sensitive to mismeasurements in these quantities. To make matters worse, as the instantaneous luminosity rises at the LHC, pileup i.e. more than one simultaneous interaction between proton pairs becomes more probable and leads to more objects being produced in each event. Each of these objects can be mismeasured leading to greater uncertainty in MET. The goal in this chapter is to understand the effect of pileup as well as mismeasurements of jet momenta on MET measurements.

In this chapter, we measure MET in events which are expected to have no real MET ($Z \rightarrow ee$ or $Z \rightarrow \mu\mu$ which result in no standard model neutrinos) and study the width of the MET distribution as a function of pileup and jet activity. Since any MET seen should be due to detector effects, this gives us a probe to characterize the behavior of this fake- or noisy-MET.

A.1 Introduction

The method presented in this chapter is used to match MET resolution between data and Monte Carlo. The method deals with fake MET, i.e. MET arising from detector mismeasurements, pileup, jet activity etc. The method is applicable to generic Monte Carlo samples.

Typically, the tails of MET are parametrized with distributions such as double-gaussians etc. We show that the need for the "second gaussian" is greatly reduced when MET is binned in variables indicative of hadronic activity and pileup. We choose HT and Nvertex for this purpose.

It must be noted that the purpose of this method is to correctly characterize the MET resolution which dominantly comes from the central part of the MET distribution. Binning in Nvertex and HT allows us to discern the true central MET region and then match the data with Monte Carlo. **We do not consider the long MET tails which are best dealt with data-driven methods such as using Z+jets dilepton data with fakeable objects as in the CFO method described in chapter 3**. Our analysis is protected against additional deviation in MET tails between data and simulation by using our data-driven fake rate methods such as the CFO method applied to Z+Jets, which accounts for contributions of high MET tails.

A.2 Basic Idea

With the LHC running at 8 TeV and higher instantaneous luminosities, pileup (number of simultaneous collisions) is a bigger problem. More interaction vertices result in more physics objects (jets, tracks, photons etc.). Each object has an uncertainty in its momentum measurement which results in a higher uncertainty in the MET measurement as the number of objects increase. We present a method to quantify the effect of pileup and jet activity on MET measurement. We do this by studying the MET distribution in events which should yield no real MET, i.e. events with exactly two leptons coming from a Z. Any MET measured here should come predominantly from resolution effects.

Consider a particle that carries all the MET. If the MET is completely fake and stochastic, then we can model the momenta components of this particle by Gaussians centered at 0 with some width σ, i.e.:

$$MET_x \sim N(0, \sigma^2), \ MET_y \sim N(0, \sigma^2), \ MET_z \sim N(0, \sigma^2)$$

We measure the transverse missing energy,

$$MET = \sqrt{MET_x^2 + MET_y^2}$$

which will follow the distribution:

$$f(a) = Pr[MET = a] = \int dx \, dy \, g(x)g(y)\delta(\sqrt{x^2 + y^2} - a)$$

where $g(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{x^2}{2\sigma^2})$ for MET_x and MET_y, and we are integrating over the full range, $x, y \in \mathbb{R}$.

Switching to polar coordinates, we get

$$f(a) = \int r dr d\theta \frac{1}{2\pi\sigma^2} \exp(-\frac{r^2}{2\sigma^2})\delta(r - a) = \frac{1}{\sigma^2}a \exp(-\frac{a^2}{2\sigma^2}) \text{ where } a \geq 0$$

So, the distribution of MET is given by:

$$f_{MET}(x) = \frac{1}{\sigma^2} x \exp(-\frac{x^2}{2\sigma^2}); \; x \geq 0 \tag{A.1}$$

Figure A.1 shows some plots of the MET distribution on linear and log scales for different values of the resolution or width parameter σ.
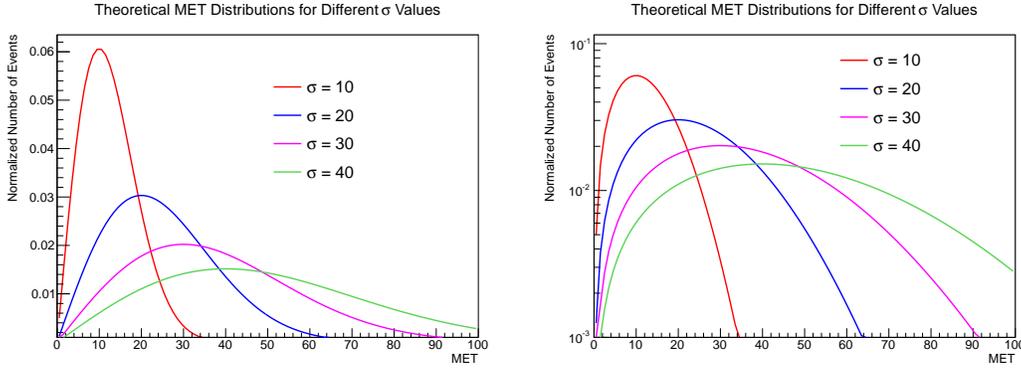
Figure A.1: Expected MET distribution from equation A.1 for different values of σ in linear (left) and log (right) scales.

A.2.1 MET Tail:

In any analysis where MET is being cut upon, we are interested in the background that comes from low-MET events that leak into the high-MET region because of MET resolution/smearing effects. In particular, we are interested in the fraction of events that lie above some MET cut d. This is given by:

$$\alpha = \frac{\frac{1}{\sigma^2} \int_d^\infty r dr \, \exp(-\frac{r^2}{2\sigma^2})}{\frac{1}{\sigma^2} \int_0^\infty r dr \, \exp(-\frac{r^2}{2\sigma^2})} = \frac{1}{\sigma^2} \int_d^\infty r dr \, \exp(-\frac{r^2}{2\sigma^2}) = 1 - \frac{1}{\sigma^2} \int_0^d r dr \, \exp(-\frac{r^2}{2\sigma^2})$$

Define $\beta \equiv \frac{r}{\sigma}$ to get

$$\alpha = 1 - \int_0^{\frac{d}{\sigma}} \beta d\beta \, \exp(-\frac{\beta^2}{2}) = 1 + \int_0^{\frac{d}{\sigma}} \frac{d}{d\beta}(\exp(-\frac{\beta^2}{2})) = 1 + \exp(-\frac{d^2}{2\sigma^2}) - 1 = \exp(-\frac{d^2}{2\sigma^2})$$

So, fraction of events in the tail for $MET > d$ is given by:

$$\alpha(\frac{d}{\sigma}) = \exp(-\frac{d^2}{2\sigma^2}) \tag{A.2}$$
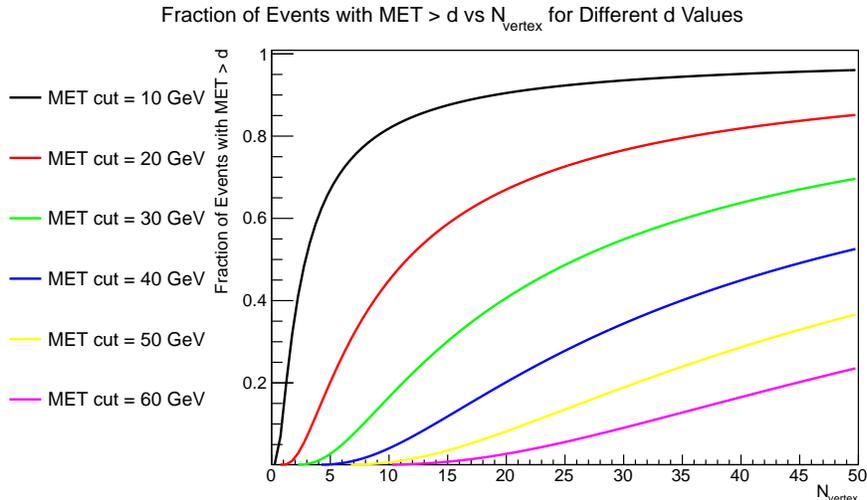
Figure A.2: Fraction of events with MET > d vs N_{vertex} for various MET cuts (d). The functional form is: $\exp\left(-\frac{d^2}{2\sigma^2 N_{\text{vertex}}}\right)$ which is a variant of equation A.2.

As expected, when $\frac{d}{\sigma} \to 0$, $\alpha \to 1$ and when $\frac{d}{\sigma} \to \infty$, $\alpha \to 0$.

Figure A.2 plots the fraction of events with MET>d for various values of MET cut d vs N_{vertex} if each collision vertex is treated independently and the total width is $\sigma\sqrt{N_{\text{vertex}}}$.

A.2.2 Modeling Variance:

Now, both in the MET distribution (Eq. A.1) and the fraction of events in the MET tail (Eq. A.2), we have a parameter σ that describes the mean as well as the width of the MET distribution. The more MET sources we have, the higher σ will be. In our case, we expect σ to describe all the underlying contributions to fake MET. In particular, we expect it to depend on pile-up i.e. number of collision vertices N_{vert} as well as some parameterization of jet activity like number of jets N_{jets} or sum p_t of all the jets H_T(we chose H_T). We can write:

$$\sigma^2 = \sigma_0^2 + \sigma_{\text{vert}}^2 N_{\text{vert}} + \sigma_{HT}^2 N_{HT} \tag{A.3}$$

where N_{HT} is the H_T-bin with some conventional discretization i.e.

$$\sigma^2 = \sigma_0^2 + \sigma_{\text{vert}}^2 N_{\text{vert}} + \sigma_{HT}^2 \lfloor \frac{H_T}{30} \rfloor \quad\quad\quad (A.4)$$

In this study, we binned H_T in 30-GeV wide bins and the jet pt cut was 30 GeV. σ_0^2 is the part of variance that is not described by either pileup or jet activity.

A.2.3   Justification of Linear Dependence of σ^2 on N_{vert} and H_T

We claim in equation A.3 that σ^2 is expected to be linear in N_{vertex} and H_T. Figures A.3 and A.4 demonstrate this by plotting measured σ^2 vs N_{vert} and H_T respectively.
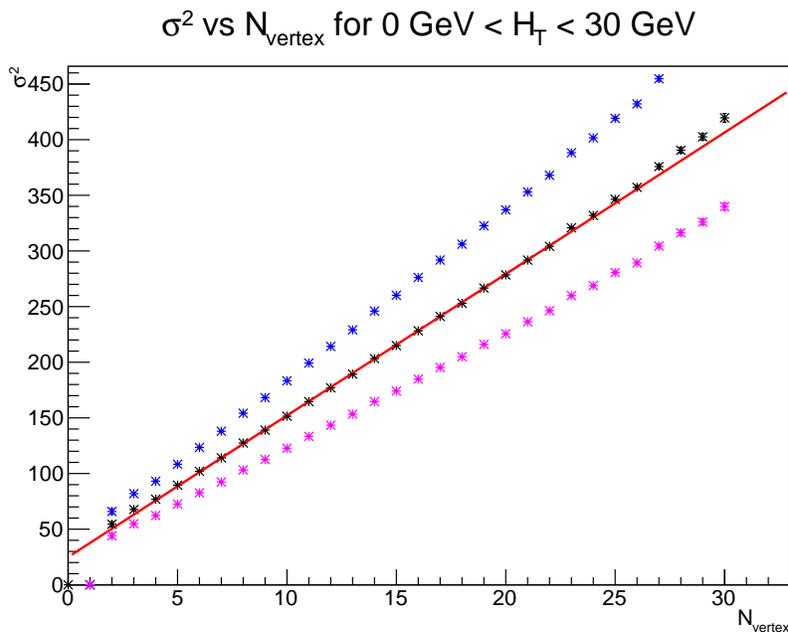
Figure A.3: Dependence of σ^2 vs N_{vert} for events with $0 \text{ GeV} < H_T < 30 \text{ GeV}$. The black dots are the nominal value. The blue (purple) dots are the measured σs shifted up (down) by the systematic (10%). The assumption that σ^2 is linear as a function of N_{vertex} is justified within the systematic.
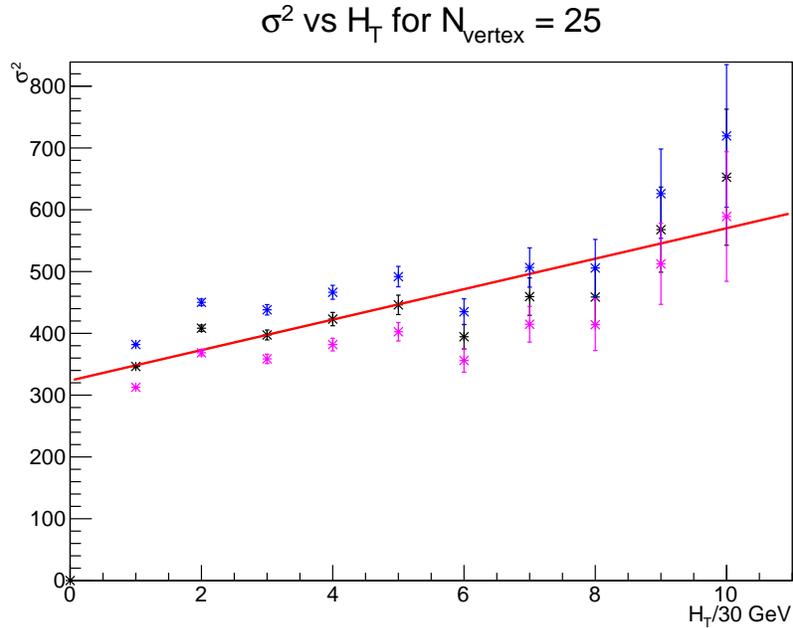
Figure A.4: Dependence of σ^2 vs $H_T/30GeV$ for events with N_{vert}=25. The black dots are the nominal value. The blue (purple) dots are the measured σs shifted up (down) by the systematic (5%). The assumption that σ^2 is linear as a function of H_T is justified within the systematic.

A.3 Removing MET Backgrounds

We are looking for events in data with exactly two opposite-sign same-flavor leptons $(e^+e^-/\mu^+\mu^-)$ with the presumption that they come from $Z \to ll$ and thus have no real MET. There are still some backgrounds that can give ll and real MET. These are:

$Z \to ll + X$ Sources:

1. $Z(ll) + W(\tau_h) + MET$

2. $Z(ll) + Z(\nu\nu)$

3. $Z \to \tau_l\tau_l \to ll + MET$

The cross-section for $WZ \to l^+l^- + \tau(W)$ is roughly $(40\,pb) * (7\%\ for\ Z(ll)) * (10\%\ for\ W \to \tau * 65\%\ for\ \tau_h) = 0.182\,pb$ as compared to $\sigma_{Z \to ll} \sim 1600\,pb$. So

we ignore $Z(ll) + W(\tau_h) + MET$. Similarly, the cross-section for $Z(ll) + Z(\nu\nu)$ is small enough that we can ignore it. For $Z \to \tau_l\tau_l$, see below.

WW Sources:

1. $WW \to ll + MET$

2. $t\bar{t} \to bbWW \to ll + MET$

We can remove a large part of WW contributions (including $Z \to \tau_l\tau_l$) by requiring that the two leptons have an invariant mass within the Z-window, $[m_Z - \Delta, m_Z + \Delta]$. In addition, for any source with WW (including $Z \to \tau_l\tau_l$), we expect to see both same-flavor as well as mixed-flavor lepton pairs with equal branching ratios:

$$BR(WW \to ee) = BR(WW \to \mu\mu) = \frac{1}{2}BR(WW \to e\mu)$$

So, we remove these backgrounds by adding the ee and $\mu\mu$ channels together and subtracting off the $e\mu$ contribution.

$W(l\nu) + Fake$:

Here the assumption is that the lepton coming from a W (with real MET) combines with a fake lepton from a jet. This contribution is made small by requiring that the two leptons lie in the Z-window. It is further suppressed by the small fake rate. In addition, fakes coming from a jet can give e^+, e^-, μ^+, μ^- democratically. So, the $e\mu$ subtraction procedure described above removes this background.

A.3.1 Gaussian Assumption

One of the key assumptions of the method is that the x and y components of fake MET (after real MET from neutrinos is subtracted) are Gaussians with mean

0 and some width σ. This assumption is definitely not true for the METx and METy distributions over all Nvertex and HT bins added together. The overall MET distribution has non-Gaussian tails (see figure A.5).
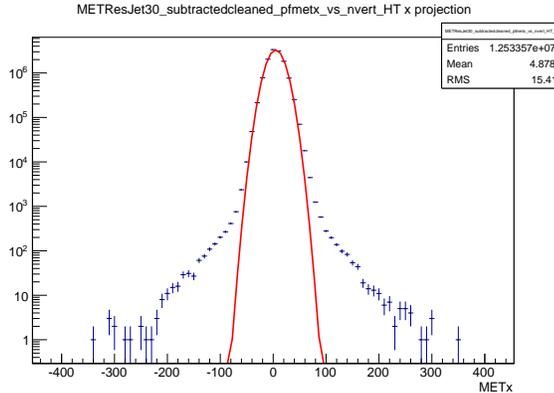
Figure A.5: MET x-component distribution in 18.1 fb^{-1} of dilepton data with $e\mu$ subtraction. The fit is a Gaussian.

In particular, note that the METx distribution peaks at around $\sim 10^7$ events and starts deviating from the Gaussian fit at around 10^3 events (after 3 orders of magnitude). The width of the shape is ~ 15 GeV (either RMS or the width of the Gaussian).

We hypothesize that that MET resolution is mainly driven by pileup (Nvertex) and jet activity (parameterized by H_T in our case). If we bin the data only in Nvertex and look at the METx distribution, we get the following plots for Nvertex = 10 (see figure A.6) and Nvertex = 25 (see figure A.7) respectively.

In both the above plots (Nvertex = 10 in figure A.6 and Nvertex = 25 in figure A.7), the METx distribution starts deviating after a 3 orders of magnitude drop just like the baseline overall METx distribution (figure A.5). The major difference is in the width of the Gaussian (or correspondingly the RMS). When Nvertex = 10, the width is ~ 12.6 GeV and when Nvertex = 25, the width is ~ 19 GeV. This suggests that binning in Nvertex alone has almost no effect on the tails but controls the width of the Gaussian, i.e. each Nvertex is very close to a Gaussian source.
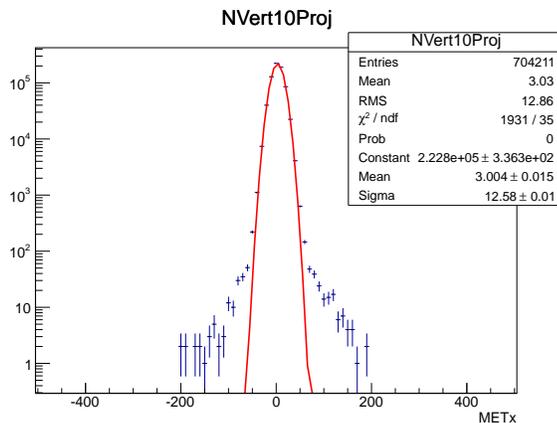
Figure A.6: MET x-component distribution in events with $N_{\text{vert}} = 10$ but summed over all H_T bins.
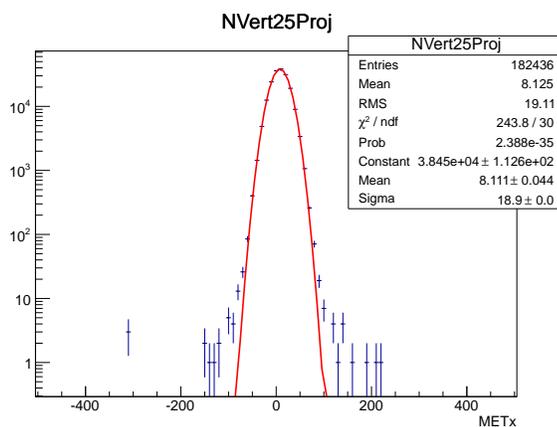
Figure A.7: MET x-component distribution in events with $N_{\text{vert}} = 25$ but summed over all H_T bins.

On the other hand, if we bin the data only in HT, then we get the following METx distributions for HT = 0-30 GeV (see figure A.8) and HT = 90-120 GeV (see figure A.9) respectively.

When HT = 0-30 GeV, the width of the Gaussian is ~15 GeV while the tail starts developing much later: after a 4 orders of magnitude drop (at 10^2 events out of a peak of 10^6 events). On the other hand, when HT = 90-120 GeV, the width of the Gaussian is ~17.5 GeV while the tail starts much earlier after only a 2 orders of magnitude drop (10^2 events out of a peak of 10^4 events). This suggests that the effect of each HT bin is two-fold. First, each HT bin contributes a Gaussian

HT1Proj

HT1Proj	
Entries	1.087781e+07
Mean	4.829
RMS	14.98
χ^2 / ndf	6.33e+04 / 39
Prob	0
Constant	2.933e+06 ± 1.147e+03
Mean	4.706 ± 0.005
Sigma	14.71 ± 0.00
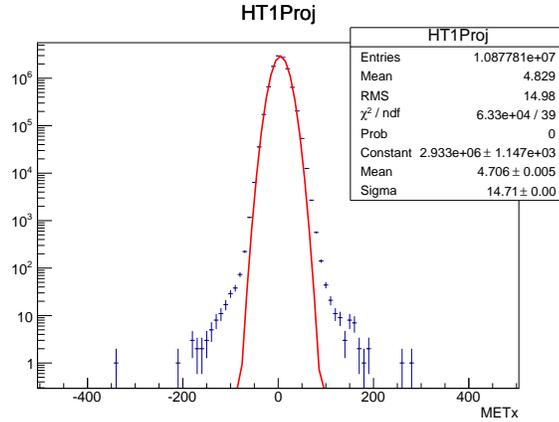
Figure A.8: MET x-component distribution in events with $H_T = 0 - 30$ GeV but summed over all N_{vert} bins.

HT4Proj

HT4Proj	
Entries	198247
Mean	5.239
RMS	18.1
χ^2 / ndf	1318 / 36
Prob	0
Constant	4.535e+04 ± 1.315e+02
Mean	5.132 ± 0.039
Sigma	17.32 ± 0.03
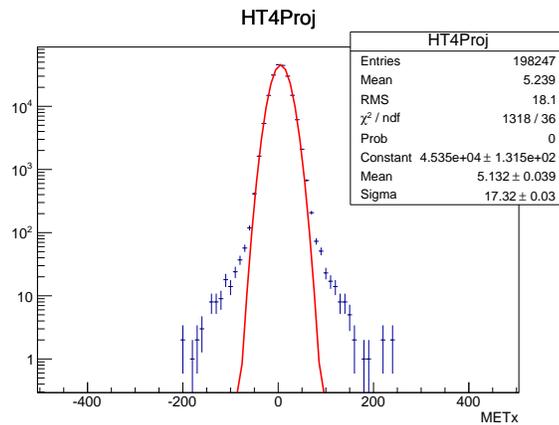
Figure A.9: MET x-component distribution in events with $H_T = 90 - 120$ GeV but summed over all N_{vert} bins.

component that increases the width of METx. Secondly, each HT bin controls a non-Gaussian tail. The primary advantage of binning in HT then is to decrease the non-Gaussian tails.

Lastly, we bin the data in both Nvertex and HT bins. The following four (see figures A.10, A.11, A.12, A.13) plots show the METx distribution in various Nvertex and HT bins and the non-Gaussian tails while present have been reduced significantly. Also note that since the aim of the method is to provide a correction to the Monte-Carlo samples to match data, the central part of the METx distribution (and its width) is the real driving factor. **We don't attempt to**

model the tails very accurately and they are accounted for by data-driven methods like using Z+jets dilepton data with fakeable objects as in the CFO method.
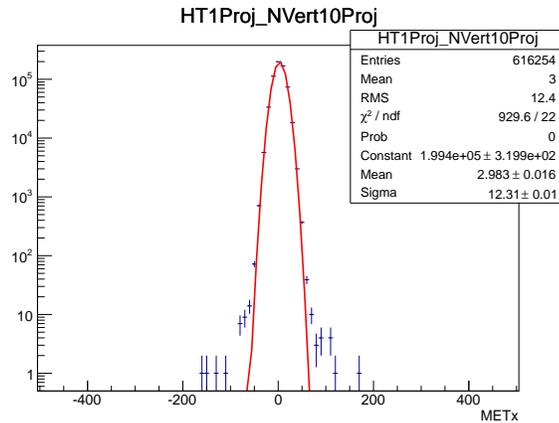
Figure A.10: MET x-component distribution in dilepton events with $N_{\text{vertex}} = 10$ and $H_T = 0 - 30$ GeV.
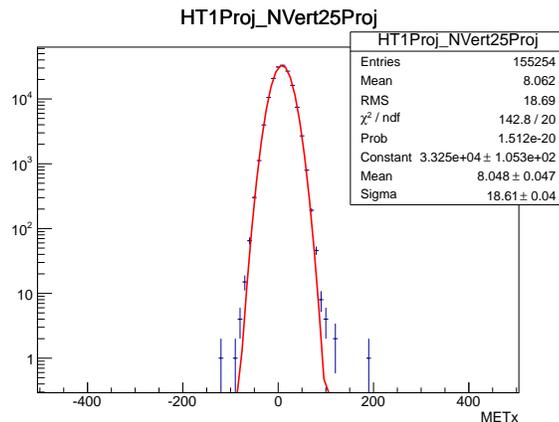
Figure A.11: MET x-component distribution in dilepton events with $N_{\text{vertex}} = 25$ and $H_T = 0 - 30$ GeV.

In the low HT (0-30 GeV) plots, for both low and high Nvertex (10 and 25 respectively), the non-Gaussian tails start developing after 4 orders of magnitude. The only difference is the width of the Gaussian, which is 12.3 GeV for Nvertex = 10 and 18.6 GeV for Nvertex = 25. In the high HT(90-120 GeV) regions, the tail starts developing after 2-3 orders of magnitude like the baseline distribution and the width goes from $\sim$15 GeV for low Nvertex to $\sim$21 GeV for high Nvertex.
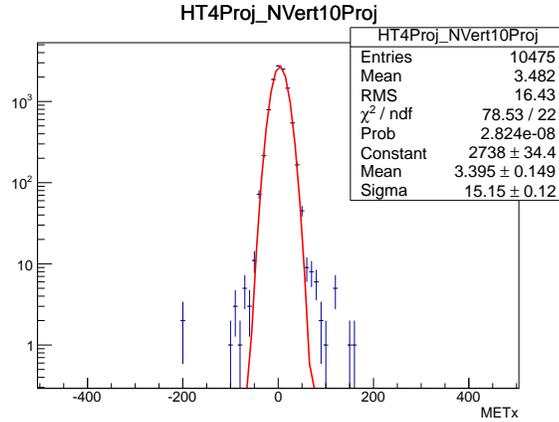
Figure A.12: MET x-component distribution in dilepton events with $N_{\text{vertex}} = 10$ and $H_T = 90 - 120$ GeV.
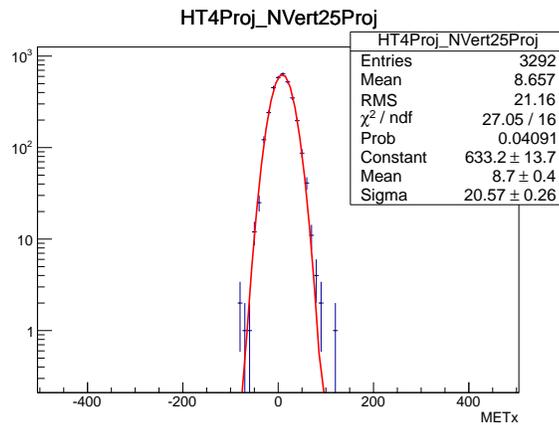
Figure A.13: MET x-component distribution in dilepton events with $N_{\text{vertex}} = 25$ and $H_T = 90 - 120$ GeV.

So, in conclusion, each Nvertex behaves almost exactly like a Gaussian source. Each HT bin contributes both a Gaussian part as well as a non-Gaussian tail. On binning in HT, we see a reduction in the tails by an order of magnitude in the low HT bins. Also, for our purposes, the width of the Gaussian part is the most important parameter and is largely independent of the tails, which are accounted for by the data-driven CFO method using Z+Jets data. This can be seen by the calculating the resolution parameters: sigmaNvert and sigmaHT.

So (see table A.1) we can fix HT in either bin (0-30 or 90-120 GeV) to get sigmaNverts of 3.6 GeV and 3.6 GeV respectively and fix Nvert in at either value

Plot	Sigma for Gaussian (GeV)
HT = 0-30 GeV, Nvert = 10	12.3
HT = 0-30 GeV, Nvert = 25	18.6
HT = 90-120 GeV, Nvert = 10	15.2
HT=90-120 GeV, Nvert = 25	20.6

Table A.1: Widths of Gaussians from fits in different H_T and Nvertex bins.

(10 or 25) to get sigmaHTs of 5.2 GeV and 5.1 GeV respectively. The central values for sigmaNvert and sigmaHT are really similar even though the tails are not modeled properly by the Gaussian because the width of the Gaussians is driven by the central part of the MET distribution and not the tails.

A.3.2 Gaussian Assumption: What if means are non-zero?

In addition to assuming that METx and METy are Gaussians, we assume that the Gaussians have means 0. As the plots in the previous section show, the means are often non-zero but small compared to the widths.

If we account for non-zero means α for METx and β for METy, then we get the following for the MET distribution for a fixed resolution σ:

$$f(a) = Pr[MET = a] = \int dx\, dy\, g_x(x)g_y(y)\delta(\sqrt{x^2 + y^2} - a)$$

where $g_x(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{(x-\alpha)^2}{2\sigma^2})$, $g_y(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{(y-\beta)^2}{2\sigma^2})$ for MET_x, MET_y and we are integrating over the full range, $x, y \in \mathbb{R}$.

So,

$$Pr[MET = a] = \frac{1}{2\pi\sigma^2}e^{-\frac{\alpha^2+\beta^2}{2\sigma^2}}\int dx\, dy\, e^{-\frac{x^2+y^2}{2\sigma^2}}e^{\frac{\alpha x+\beta y}{\sigma^2}}\delta(\sqrt{x^2+y^2} - a)$$

Do a rotation in the x-y plane to get

$$Pr[MET = a] = \frac{1}{2\pi\sigma^2}e^{-\frac{\alpha^2+\beta^2}{2\sigma^2}}\int dx\, dy\, e^{-\frac{x^2+y^2}{2\sigma^2}}e^{\frac{\gamma x}{\sigma^2}}\delta(\sqrt{x^2+y^2} - a)$$

where $\gamma = \sqrt{\alpha^2 + \beta^2}$ and so

$$Pr[MET = a] = N \int dx\,dy\, e^{-\frac{x^2+y^2}{2\sigma^2}}(1 + \frac{1}{2}\frac{\gamma^2 x^2}{\sigma^4} + \frac{1}{24}\frac{\gamma^4 x^4}{\sigma^8} + ...)\delta(\sqrt{x^2+y^2} - a)$$

where N is a normalization constant independent of MET and the terms odd in x vanish.

$$Pr[MET = a] \sim \int r\,dr\,d\theta\, e^{-\frac{r^2}{2\sigma^2}}(1 + \frac{1}{2}\frac{\gamma^2 r^2 cos^2(\theta)}{\sigma^4} + \frac{1}{24}\frac{\gamma^4 r^4 cos^4(\theta)}{\sigma^8} + ...)\delta(r - a)$$

$$Pr[MET = a] \sim a\,e^{-\frac{a^2}{2\sigma^2}}\int d\theta\,(1 + \frac{1}{2}\frac{\gamma^2 a^2 cos^2(\theta)}{\sigma^4} + \frac{1}{24}\frac{\gamma^4 a^4 cos^4(\theta)}{\sigma^8} + ...)$$

$$Pr[MET = a] = \frac{1}{2\pi\sigma^2}a\,e^{-\frac{a^2}{2\sigma^2}}(2\pi + \frac{1}{2}\frac{\gamma^2 a^2 \pi}{\sigma^4} + \frac{1}{32}\frac{\gamma^4 a^4 \pi}{\sigma^8} + ...)$$

The non-zero means adds higher order terms to the Rayleigh tail. These tails are suppressed by powers of the factor $\frac{\alpha^2+\beta^2}{\sigma^2}$.

The leading term and next-to-leading term become comparable at a MET of:

$$2\pi = \frac{1}{2}\frac{\gamma^2 a^2 \pi}{\sigma^4} \implies a = 4\frac{\sigma^2}{\gamma}$$

If $\sigma = \sigma_{Nvert}\sqrt{N_{\text{vert}}} = (4\,GeV) * \sqrt{16} = 16\,GeV$ (ignoring contributions from H_T and σ_0 which will only increase σ) and the mean is around $5\,GeV$, we get non-Rayleigh tails at $200\,GeV$. As long as one is in the regime,

$$\frac{\gamma^2 a^2}{\sigma^4} < 1$$

the higher-order terms will get suppressed even more. At higher METs (a), the non-Rayleigh tails get severely suppressed by the decay exponential. In addition, we are only interested in getting a measurement of the width/resolution and the tails don't matter. So, we still assume that the means of METx and METy are

zero in the rest of this chapter.

A.4   Procedure and Results

As described above, in 18.1 fb^{-1} of data, we combine e^+e^- and $\mu^+\mu^-$ channels and subtract off any $e^{\pm}\mu^{\mp}$ contributions where in each case the two leptons are required to form an invariant mass in the Z-window. We also look at the following Monte-Carlo samples: ZZ, WZ, $ZJets$ and $t\bar{t}$. For WZ and $t\bar{t}$, we subtract off the neutrino contributions from the MET to get just the fake MET. For ZZ, we just look at the $e^+e^-\mu^+\mu^-$ channel. For WZ, we look at the $e^+e^-\mu^{\pm}$ and $\mu^+\mu^-e^{\pm}$ channels. For $t\bar{t}$, we look at the $e^{\pm}\mu^{\mp}$, e^+e^- and $\mu^+\mu^-$ channels. In $t\bar{t}$, we require exactly two generator-level neutrinos which ensures that none of the leptons comes from leptonic taus, which would give additional real MET.

For each of these samples, we bin the events as a function of N_{vert} and H_T and for each bin look at the fraction of events with $MET > d$ where d is a variable MET cut. These distributions are then fit to equation A.2 with variance coming from equation A.3. The fits give us values for σ_{vert} and σ_{HT} and we get a systematic on these fit values by varying the MET cut d.

A.4.1   Results from Fitting MET Tail Distributions

Note: The following two subsections: "Binning in N_{vert} only" and "Binning in H_T only" are meant to show a comparison of different Monte-Carlo samples and the effect of differing underlying H_T and N_{vert} distributions. The data plots are not relevant to the discussion in these subsections and haven't been updated to the full luminosity.
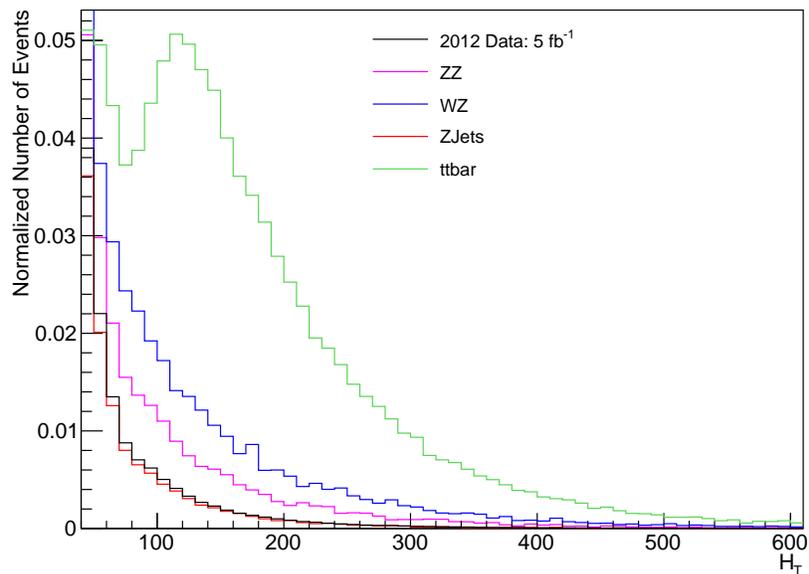
Figure A.14: H_T Distributions of data and MC samples.
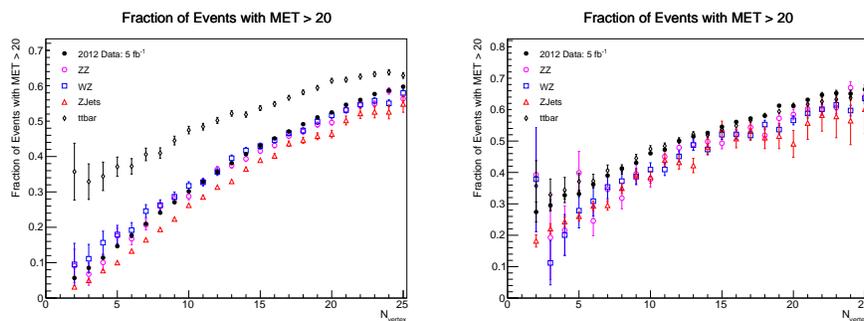
Figure A.15:   Fraction of Events vs N_{vertex} with MET>20 GeV with nominal H_T distributions (left) and with H_T distributions matched by reweighing (right)
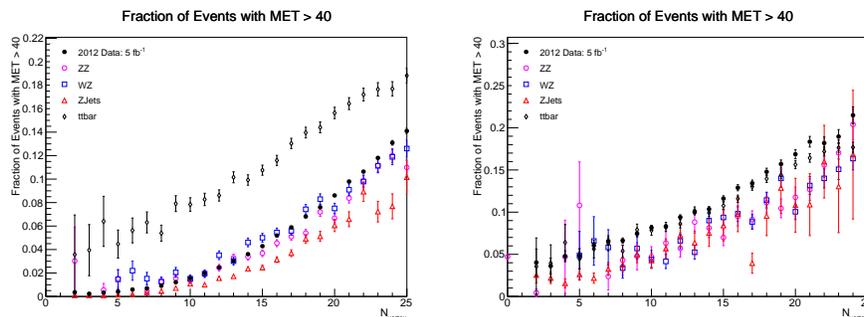
Figure A.16:   Fraction of Events vs N_{vertex} with MET>40 GeV with nominal H_T distributions (left) and with H_T distributions matched by reweighing (right)
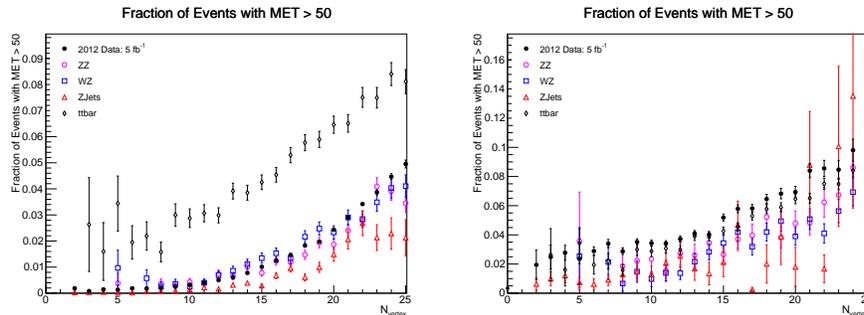
Figure A.17: Fraction of Events vs N_{vertex} with MET>50 GeV with nominal H_T distributions (left) and with H_T distributions matched by reweighing (right)

Binning in N_{vert} only:

Figures A.15, A.16, and A.17 (left plots) show the fraction of events with $MET > d$ for $d = 20, 40,$ and $50\ GeV$ for data as well as MC. In all the plots, $t\bar{t}$ has more events with $MET > d$ for different N_{vert} as compared to the other samples. This is because the $t\bar{t}$ H_T distribution has a much higher mean compared to all these other samples and if we bin in just N_{vert}, we are averaging over the H_T distribution (See figure A.14).

Figures A.15, A.16, and A.17 (right plots) show the effect of scaling all the H_T distributions to match $t\bar{t}$. Now, all the distributions collapse into roughly the same regions. The error bars are quite big because of the effect of reweighing two completely different distributions to match.

Binning in H_T only:

Figures A.19, A.20, and A.21 (left plots) show the fraction of events in the $MET > d$ tail in different H_T bins with N_{vert} averaged over. Now, ZJets is much lower than the other samples and this is because the ZJets N_{vert} distribution peaks at a much lower values (see figure A.18).

Figures A.19, A.20, and A.21 (right plots) shows the effect of reweighing the N_{vert} distributions to match. Now, the shapes collapse into the same regions.

Both the above two plots indicate the MET smearing of vastly different physics
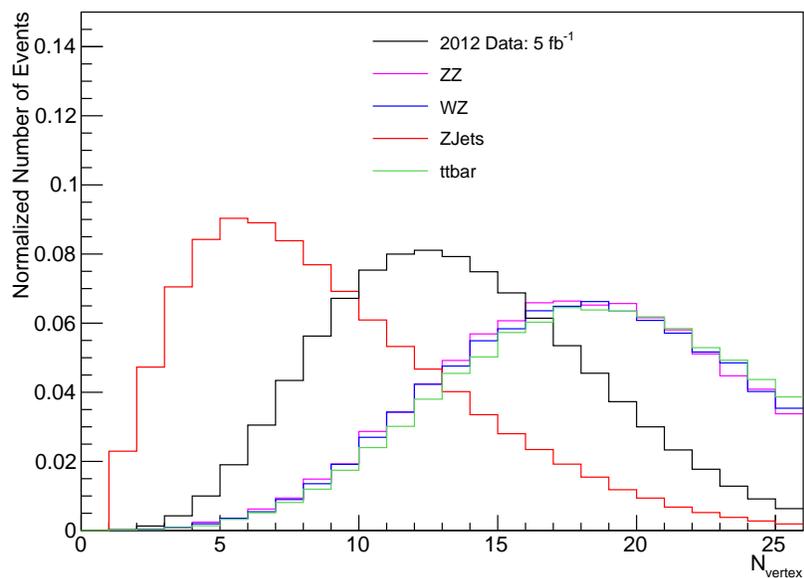
Figure A.18: N_{vertex} Distributions of data and MC samples.
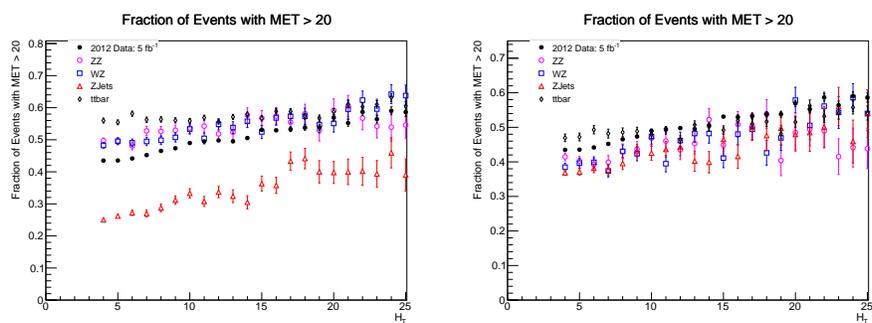
Figure A.19: Fraction of Events vs H_T with MET>20 GeV with nominal N_{vertex} distributions (top) and with N_{vertex} distributions matched by reweighing (bottom)
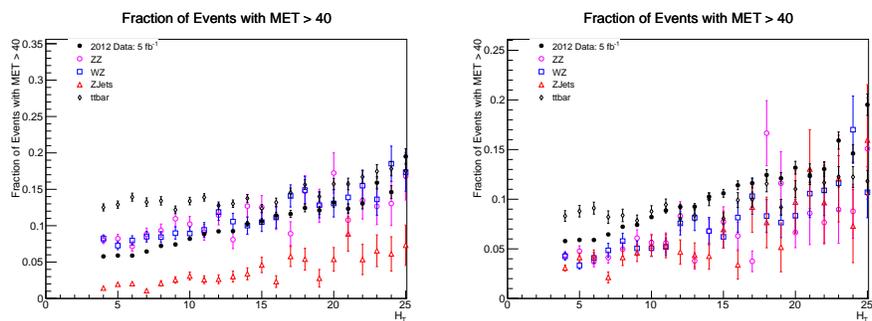
Figure A.20: Fraction of Events vs H_T with MET>40 GeV with nominal N_{vertex} distributions (top) and with N_{vertex} distributions matched by reweighing (bottom)
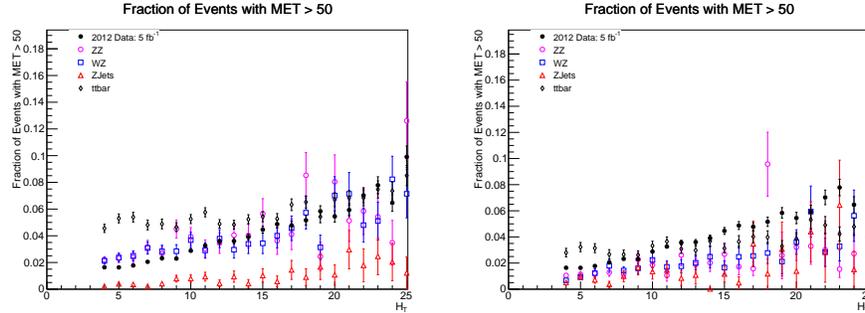
Figure A.21:   Fraction of Events vs H_T with MET>50 GeV with nominal N_{vertex} distributions (top) and with N_{vertex} distributions matched by reweighing (bottom)

samples is driven mainly by pileup (N_{vert}) and jet activity (H_T). On matching these distributions in samples as different as $t\bar{t}$ and ZJets, we get close agreement between the fake-MET distributions.

Binning in N_{vert} and H_T:

Here we give an example of how σ_0, σ_{vert} and σ_{HT} are fitted for in data. We bin the data in bins of N_{vert} and H_T. We select a MET cut d and for each N_{vert} and H_T bin, we compute the fraction of events with MET $> d$. So now we have a 2-dimensional grid of fraction of events with MET $> d$ for every d that we choose.

We expect the fraction of events to be given by the functional form:

$$\alpha(\frac{d}{\sigma}) = \exp(-\frac{d^2}{2\sigma^2})$$

where

$$\sigma^2 = \sigma_0^2 + \sigma_{\text{vert}}^2 N_{\text{vert}} + \sigma_{HT}^2 N_{HT}$$

Now, fix the H_T bin and look at the fractions vs N_{vert} (see Figure A.22).

We fit these distributions to:

$$\alpha(\frac{d}{\sigma}) = \exp(-\frac{d^2}{2\sigma^2})$$
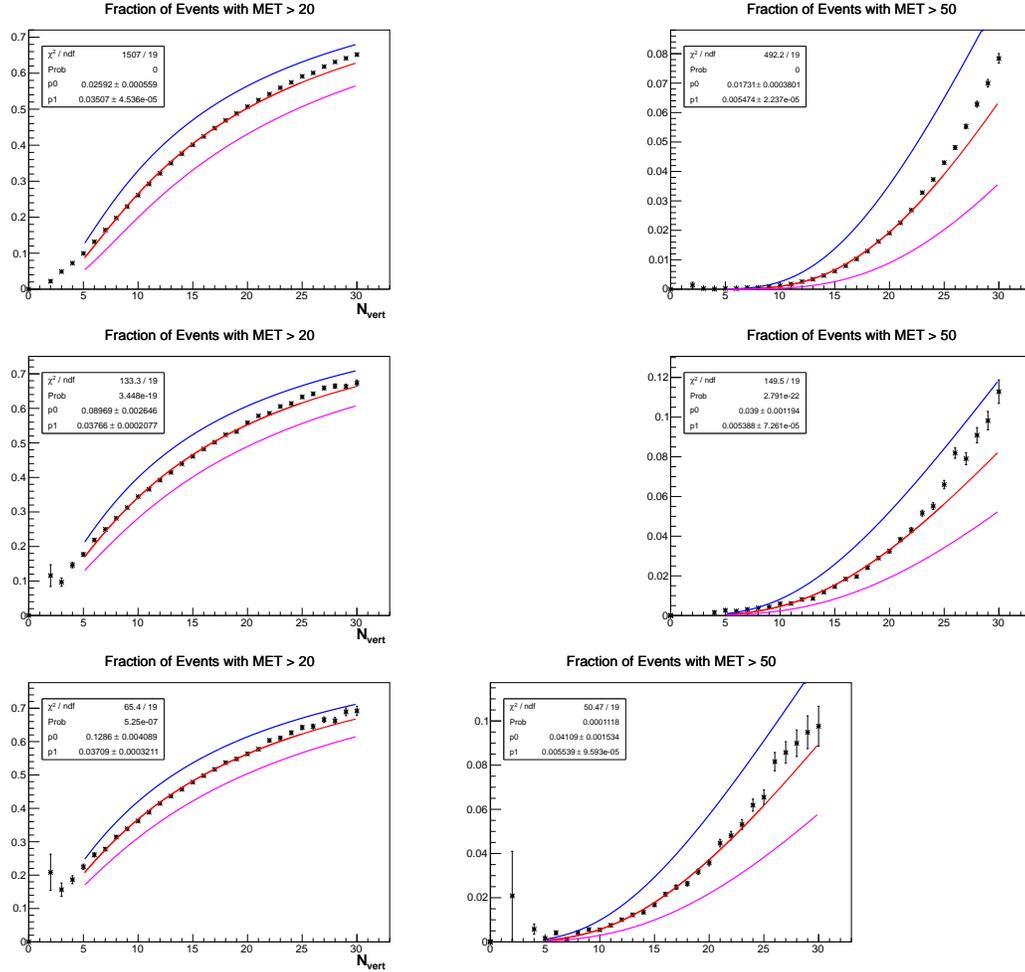
Figure A.22: Fraction of Events with MET > 20 GeV (left) and MET > 50 GeV (right) vs N_{vert} for fixed $H_T = 0\text{-}30$ GeV (top), $H_T = 30\text{-}60$ GeV (middle), $H_T = 60\text{-}90$ GeV (bottom). The blue (purple) curves has σ shifted up (down) by the systematic (10%).

where

$$\sigma^2 = \sigma_{0,HT}^2 + \sigma_{\mathrm{vert}}^2 N_{\mathrm{vert}}$$

and $\sigma_{0,HT}^2 = \sigma_0^2 + \sigma_{HT}^2 N_{HT}$ and N_{HT} is fixed as mentioned above. These fits for various values of d give us different values of σ_{vert} and differences of $\sigma_{0,HT}^2$ between different HT bins give us σ_{HT}. Also, $\sigma_{0,HT}^2$ from the lowest H_T bin is used as an estimate of σ_0. We get a systematic on these σs from looking at how much the fit values vary as d is changed from 20 GeV to 60 GeV.
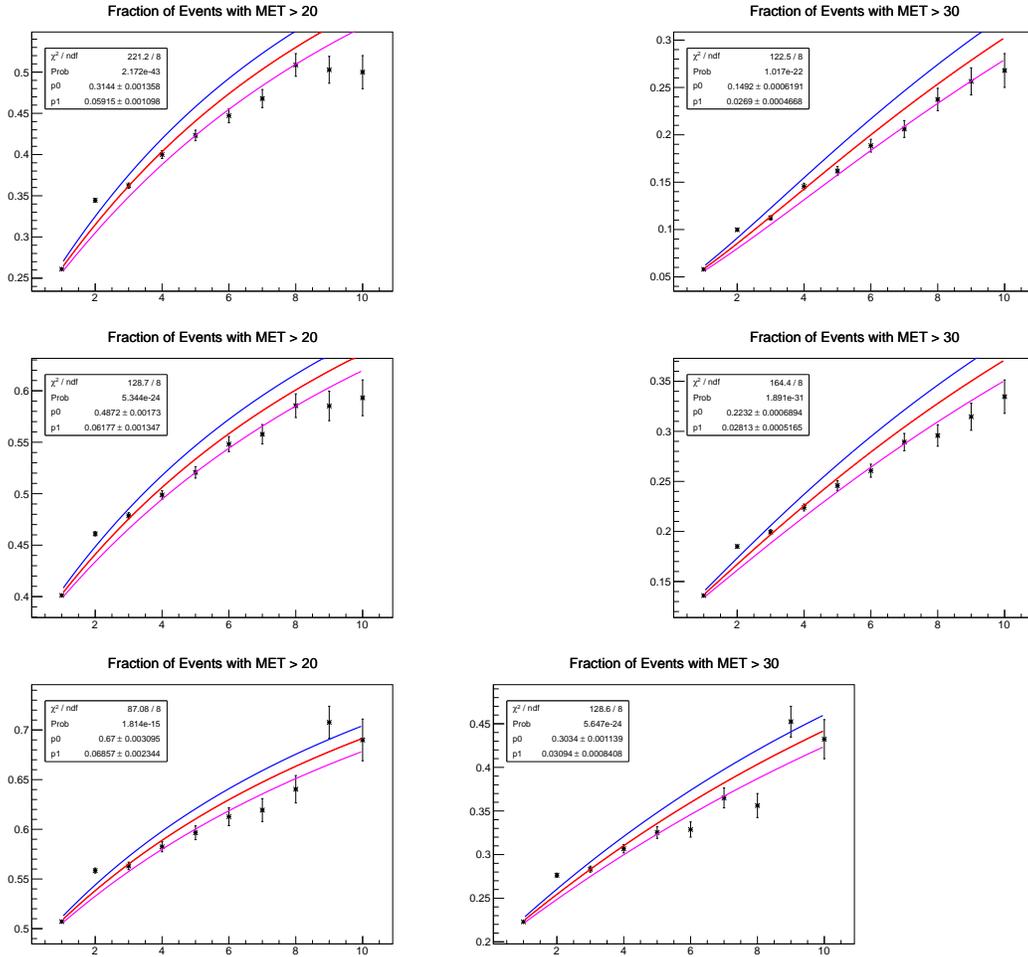
Figure A.23: Fraction of Events with MET > 20 GeV (left) and MET > 30 GeV (right) vs H_T in bins of 30 GeV for fixed $N_{\text{vert}} = 10$ (top), $N_{\text{vert}} = 15$ (middle), $N_{\text{vert}} = 20$ (bottom). The blue (purple) curve has σ shifted up (down) by the systematic (5%).

Similarly, we fix the N_{vert} bin and plot the fraction vs N_{HT}(the H_T bin) (see Figure A.23). Now, the fit function is:

$$\alpha(\frac{d}{\sigma}) = \exp(-\frac{d^2}{2\sigma^2})$$

where

$$\sigma^2 = \sigma_{0,vert}^2 + \sigma_{HT}^2 N_{HT}$$

Dataset	σ_0	σ_{vert}	σ_{HT}	$\delta\sigma_0$	$\delta\sigma_{\text{vert}}$	$\delta\sigma_{HT}$
$18.1fb^{-1}$Data	4.5	3.75	5.01	100%(4.5)	10%(0.38)	5%(0.25)
ZZ	5.41	3.46	4.07	40%(2.2)	20%(0.69)	25%(1.02)
WZ	5.5	3.52	4.00	100%(5.5)	20%(0.70)	10%(0.40)
$t\bar{t}$	11.00	3.1	3.00	50%(5.5)	30%(0.93)	60%(1.8)
ZJets	4.55	3.51	4.20	70%(3.2)	10%(0.35)	15%(0.63)

Table A.2: Resolutions from Fits for Dilepton Data and MC Samples

and $\sigma_{0,vert}^2 = \sigma_0^2 + \sigma_{\text{vert}}^2 N_{\text{vert}}$ and N_{vert} is fixed. These fits for various values of d give us a measurement of σ_{HT} and the differences of $\sigma_{0,vert}^2$ between different N_{vert} bins gives us a measurement of σ_{vert}. We get a systematic on these measurements from varying d.

In addition, we make sure that the values of the parameters from both fixed H_T slices and fixed N_{vert} slices are consistent with each other. This gives us additional confidence in our fitting procedure.

A.4.2   MET Closure

Figure A.24 shows the MET distribution in data (after $e - \mu$ subtraction) compared with a sum of Rayleigh distributions

$$p(x) = \sum_{ij} W_{ij} \frac{x}{\sigma_{ij}^2} e^{-x^2/2\sigma_{ij}^2}, \tag{A.5}$$

where "i" represents the number of vertices and "j" is for the number of jets or H_T in bins of 30 GeV. The weight W_{ij} represents the fraction of events that have i vertices and $j \times 40\ GeV$ of H_T and is gotten from data too. The plot shows the a sum of Rayleigh distributions describes the overall MET distribution in data well over a range varying by four orders of magnitude.
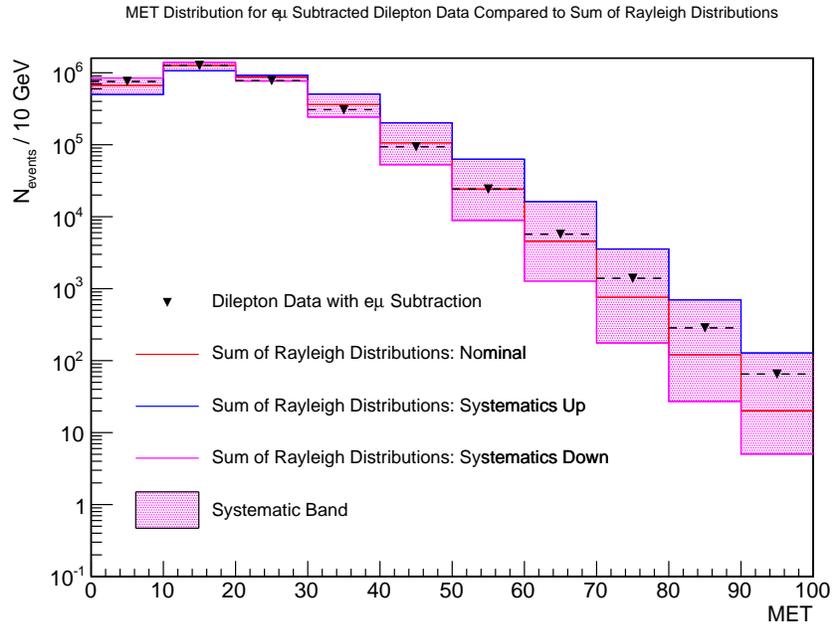
Figure A.24: MET Distribution in $e - \mu$ subtracted dilepton data compared with sum of Rayleigh distributions with widths taken from $e - \mu$ subtracted data.

A.5 Applying MET Corrections

We want an estimate of the systematic on the background estimates due to MET smearing effects. For data we have values of $\sigma_{0,Data}$, $\sigma_{vert,Data}$, and $\sigma_{HT,Data}$ and for any give MC sample, we have $\sigma_{0,MC}$, $\sigma_{vert,MC}$ and $\sigma_{HT,MC}$ from the above fitting procedure.

We first do an additional MET smearing on every event in MC depending on the H_T and N_{vert} for that event. For each event, smear the x and y components of MET, MET_x and MET_y independently (i.e. generate two different random numbers) by Gaussians $N(0, \sigma^2)$ where

$$\sigma^2 = (\sigma_{0,Data}^2 - \sigma_{0,MC}^2) + (\sigma_{vert,Data}^2 - \sigma_{vert,MC}^2)N_{\text{vert}} + (\sigma_{HT,Data}^2 - \sigma_{HT,MC}^2)N_{HT}$$

where N_{vert} and $N_{HT} = \lfloor \frac{H_T}{30} \rfloor$ are for this particular event. This corrects the MET smearing for MC so that it matches data.

The following steps give a procedure for getting the systematic on the back-ground:

1. Bin the background (MC) after the above-mentioned smearing in a two-dimensional X vs Y plane where X and Y are physical quantities of interest, say transverse mass M_{trans} and MET. This is the **nominal plane**.

2. Now, for each event, smear the x and y components of MET, MET_x and MET_y independently (i.e. generate two different random numbers) by Gaussians $N(0, \sigma^2)$ with the resolution constants moved up and down within their systematic errors where as before:

$$\sigma^2 = (\sigma_{0,Data}^2 - \sigma_{0,MC}^2) + (\sigma_{vert,Data}^2 - \sigma_{vert,MC}^2)N_{\text{vert}} + (\sigma_{HT,Data}^2 - \sigma_{HT,MC}^2)N_{HT}$$

and N_{vert} and N_{HT} are for this particular event. This will smear the MET for the MC event and will move events around in the XY plane. This is the **shifted plane**.

3. The difference between the nominal and the shifted plane gives the effect of the uncertainty in MET smearing on the background estimation in different XY regions. In particular, note that the backgrounds in different XY regions will be either correlated on anti-correlated because there's a zero-sum game between different XY regions. So, the effect of this smearing gives an estimate of the systematic on the background due to MET resolution effects.

A.6   Conclusion

We derive the expected distribution of MET if the the underlying sources of MET are completely random. This is used to parameterize MET resolution as a function of pile-up and jet activity in data and various Monte-Carlo samples like WZ+Jets,

$t\bar{t}$+Jets, Z+Jets, ZZ+Jets. Uncertainties on MET resolution parameters are used to estimate a systematic on background predictions from MET resolution effects.

The method is independent of the underlying physics as long as one subtracts off the contributions of neutrinos. It can also be generalized to include more causes of MET smearing (in addition to pile-up and jet-activity).

A.7 Modeling of MET by Rayleigh Distributions

If the x and y components of MET are exactly Gaussians, then we expect the MET to follow a Rayleigh distribution. As we showed above, METx and METy are better modeled by Gaussians as the data is binned in Nvertex and HT. We then look at MET distributions in these Nvertex and HT bins and fit them to Rayleigh distributions. The first plot (see figure A.25) shows the MET distribution for all Nvertex and HT bins combined. **Note**: In all the plots below, p0 refers to the constant for normalization in the Rayleigh distribution and p1 refers to the width of the Rayleigh
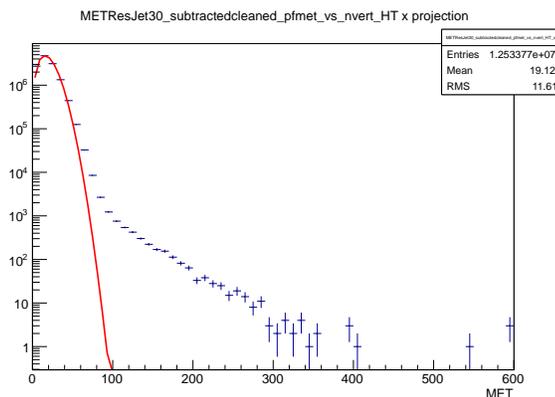
Figure A.25: MET distribution in 18.1 fb^{-1} of dilepton data with $e\mu$ subtraction. The fit is a Rayleigh distribution.

This clearly has non-Rayleigh tails which is due to the non-Gaussian tails in the overall METx distribution shown above.

If we bin the data only in Nvertex and look at the MET distributions, we

get the following plots for Nvertex = 10 (see figure A.26) and Nvertex = 25 (see figure A.27) respectively.
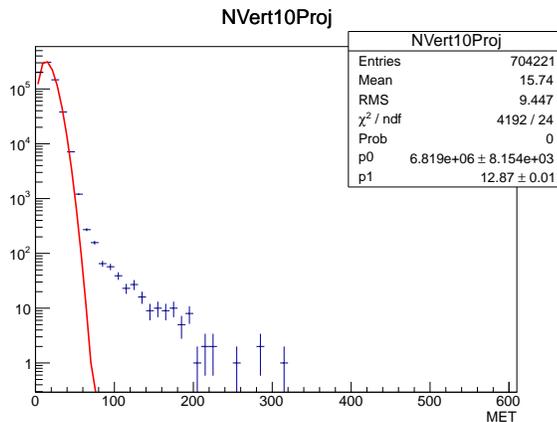
Figure A.26: MET Distribution in dilepton events with $N_{\text{vertex}} = 10$ (summed over H_T bins).
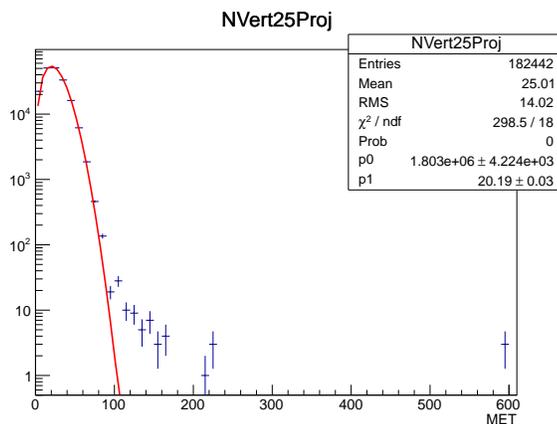
Figure A.27: MET Distribution in dilepton events with $N_{\text{vertex}} = 25$ (summed over H_T bins).

On the other hand, if we bin the data only in HT, then we get the following MET distributions for HT = 0-30 GeV (see figure A.28) and HT = 90-120 GeV (see figure A.29) respectively.

In all these 4 plots (two with just Nvertex binning and two with just HT binning), the non-Gaussian tails (reduced but clearly present) in the METx distributions create non-Rayleigh tails in the MET distributions. The tails get reduced
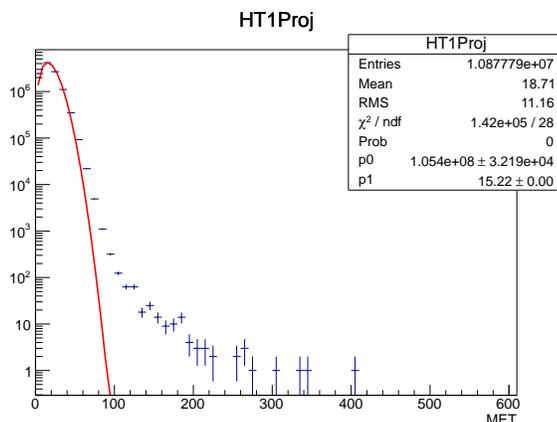
Figure A.28: MET Distribution in dilepton events with $H_T = 0 - 30$ GeV (summed over N_{vertex} bins).
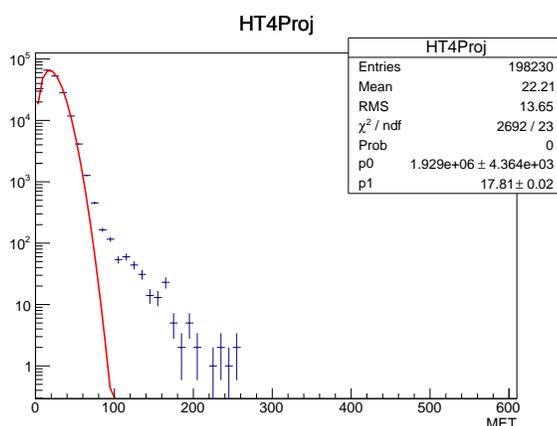
Figure A.29: MET Distribution in dilepton events with $H_T = 90 - 120$ GeV (summed over N_{vertex} bins).

on binning in Nvertex and HT compared to the overall MET distribution but are still present.

Lastly, we bin both in Nvertex and HT as before with METx. The non-Gaussian tails are reduced significantly in the METx plots with fine binning and as a consequence, we see a much reduced tail in the Rayleigh distributions (see figures A.30, A.31, A.32, A.33) . Also, to reiterate, **we don't attempt to model the tails very accurately and they are accounted for by data-driven methods like using Z+jets dilepton data with fakeable objects as in the CFO method**. The aim of this method is to get a measure of the

width of the distribution so that we can correct Monte-Carlo samples to match data.
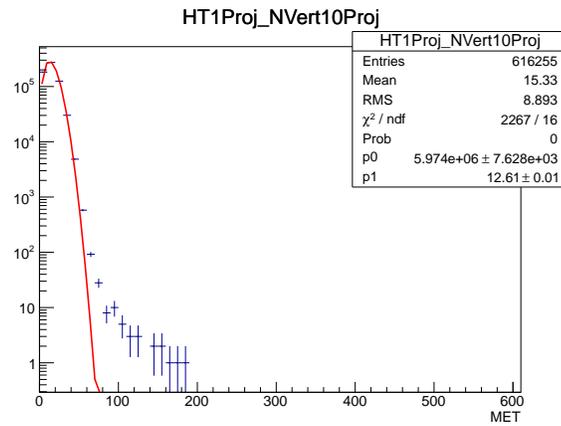
Figure A.30: MET Distribution in dilepton events with $N_{\text{vertex}} = 10$ and $H_T = 0 - 30$ GeV.
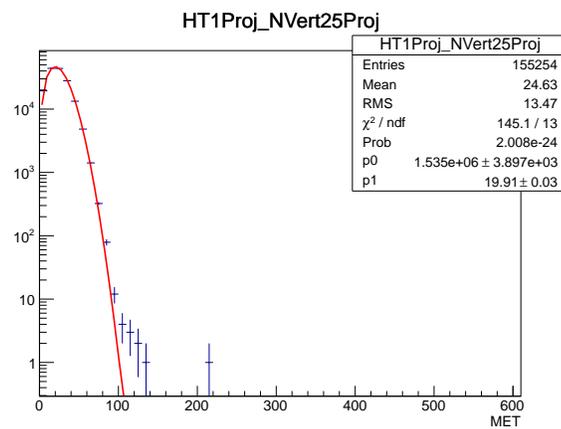
Figure A.31: MET Distribution in dilepton events with $N_{\text{vertex}} = 25$ and $H_T = 0 - 30$ GeV.
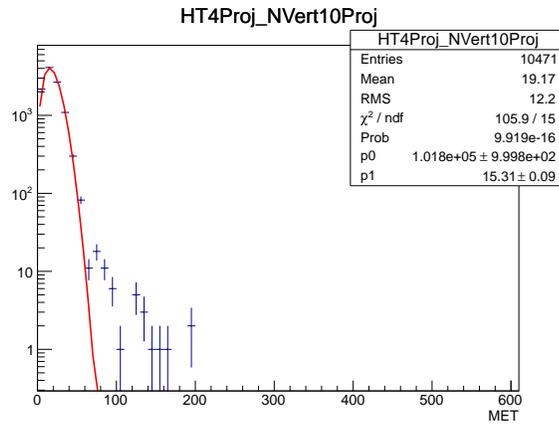
Figure A.32: MET Distribution in dilepton events with $N_{\text{vertex}} = 10$ and $H_T = 90 - 120$ GeV.
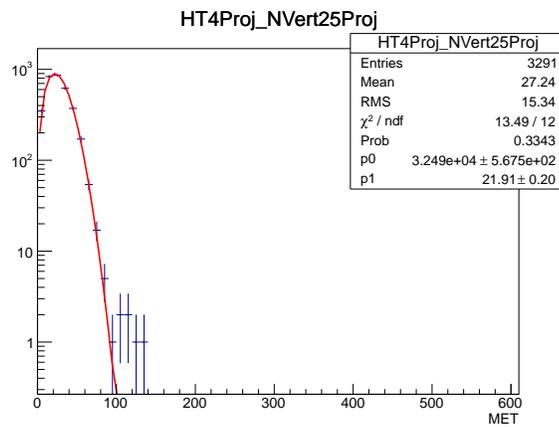
Figure A.33: MET Distribution in dilepton events with $N_{\text{vertex}} = 25$ and $H_T = 90 - 120$ GeV.

Appendix B

S_T Distribution Plots for All Channels

B.1 Plots for 3-lepton and 4-lepton Events

The following plots show observations, background estimations and signal for various 3-lepton and 4-lepton bins. The left plot always shows $b'b' \to bZbZ$ and the right always shows $b'b' \to tWtW$ where the b' mass is 550 GeVin both cases.
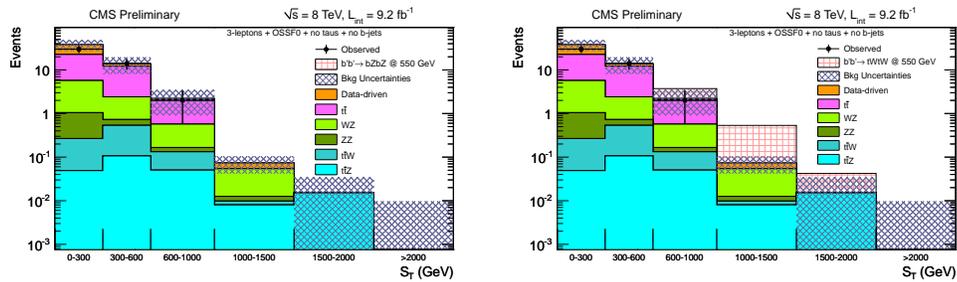
Figure B.1: 3-lepton + OSSF0 + on-Z + Tau0 + b0
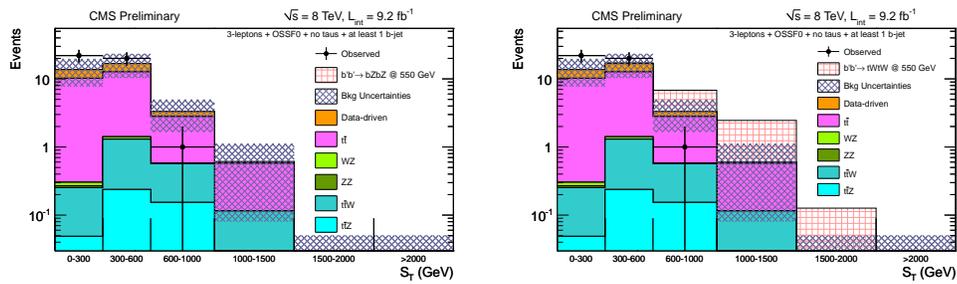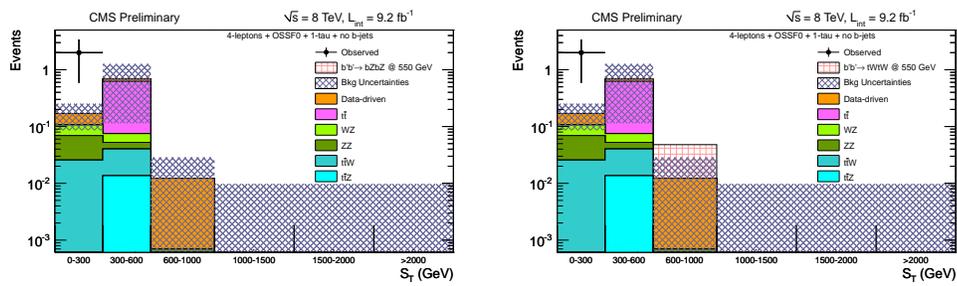

Figure B.2: 3-lepton + OSSF0 + on-Z + Tau0 + b1
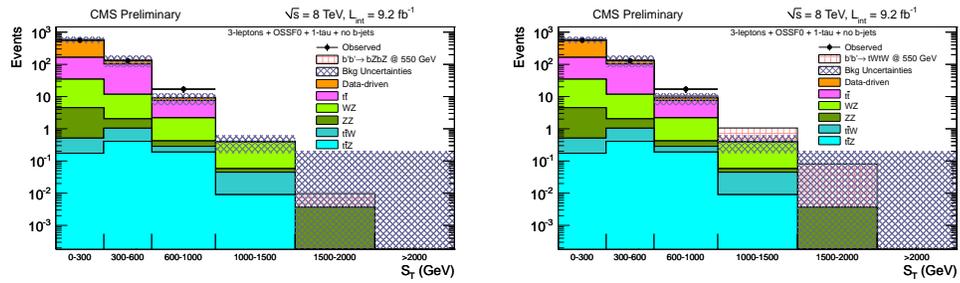
Figure B.3: 3-lepton + OSSF0 + on-Z + Tau1 + b0
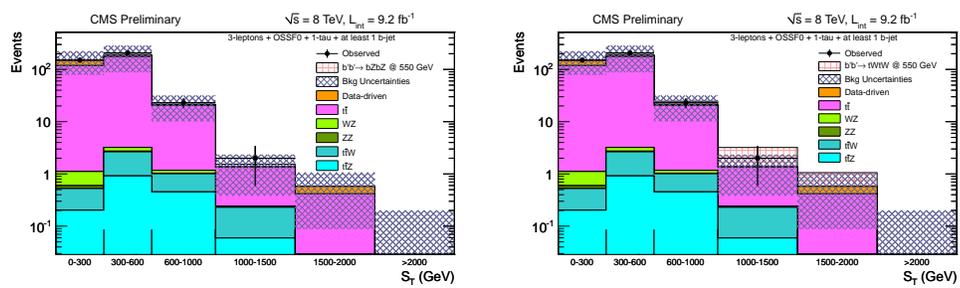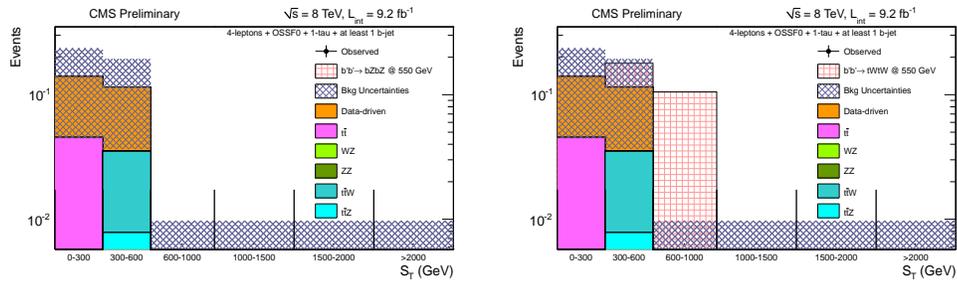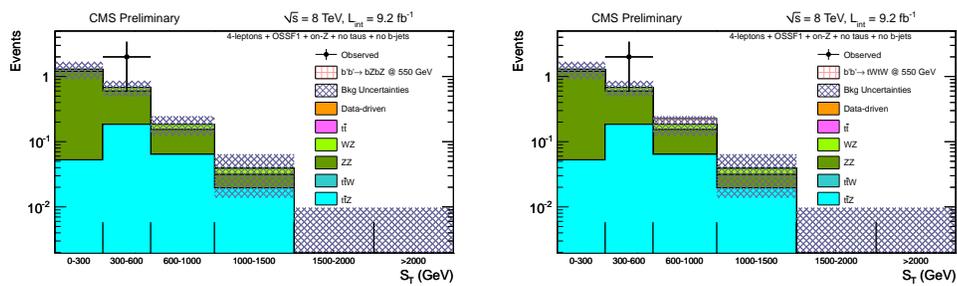

Figure B.4: 3-lepton + OSSF0 + on-Z + Tau1 + b1
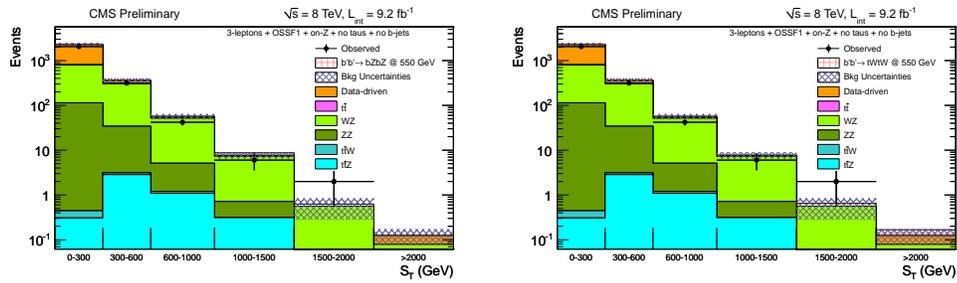
Figure B.5: 3-lepton + OSSF1 + on-Z + Tau0 + b0
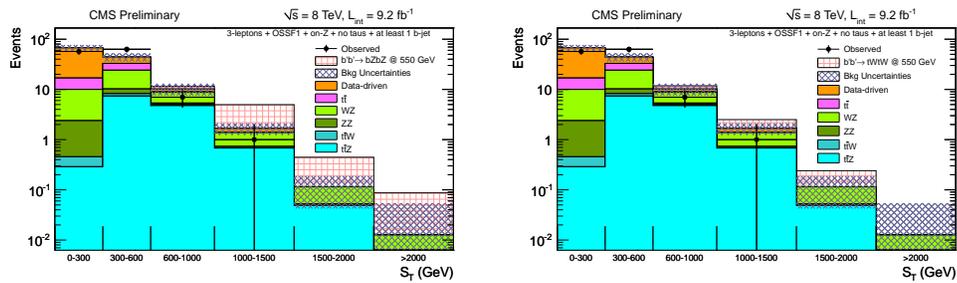
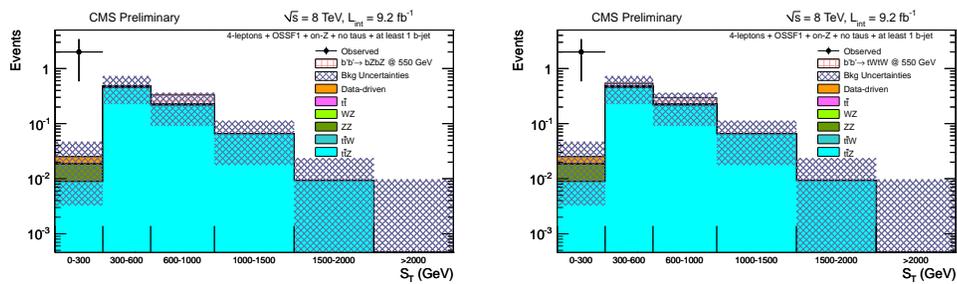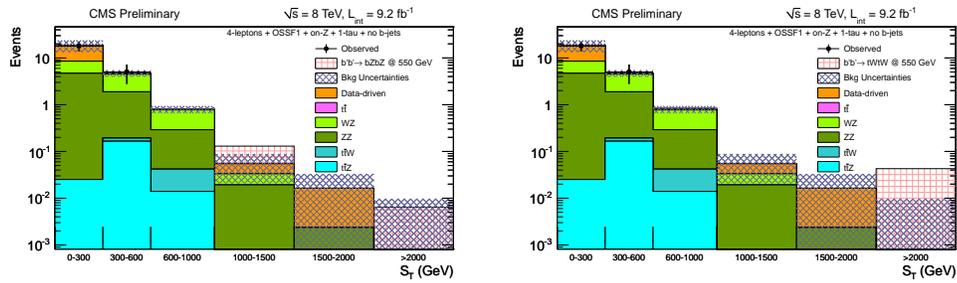Figure B.6: 3-lepton + OSSF1 + on-Z + Tau0 + b1
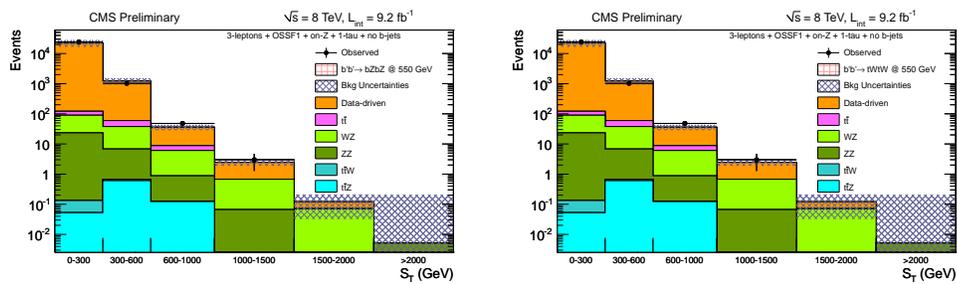
Figure B.7: 3-lepton + OSSF1 + on-Z + Tau1 + b0
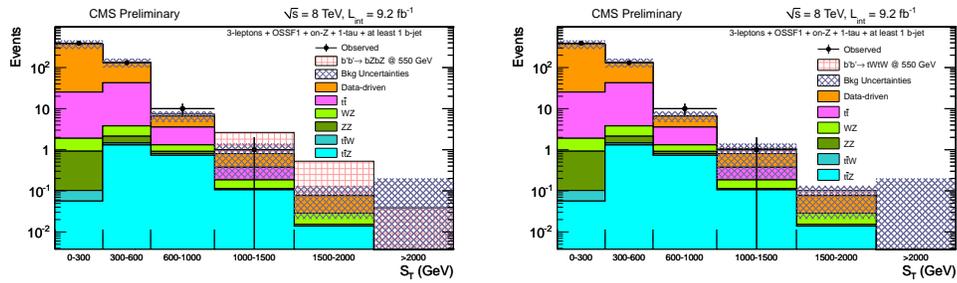
Figure B.8: 3-lepton + OSSF1 + on-Z + Tau1 + b1
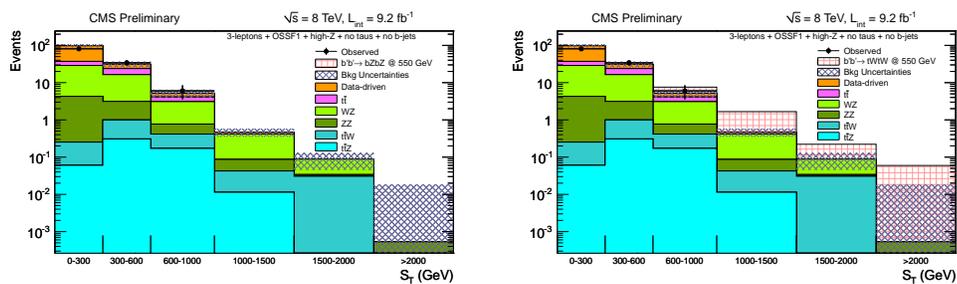
Figure B.9: 3-lepton + OSSF1 + above-Z + Tau0 + b0
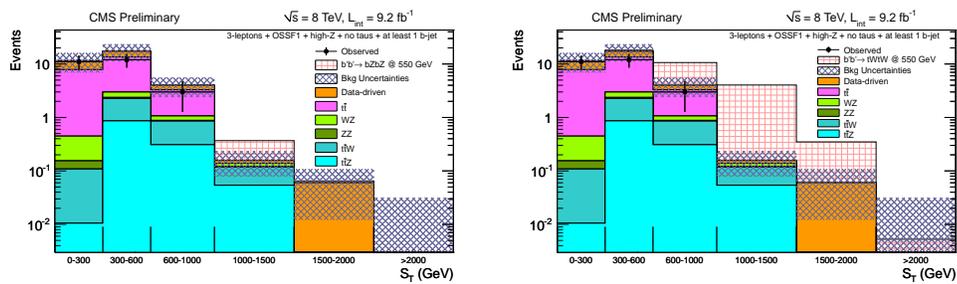
Figure B.10: 3-lepton + OSSF1 + above-Z + Tau0 + b1
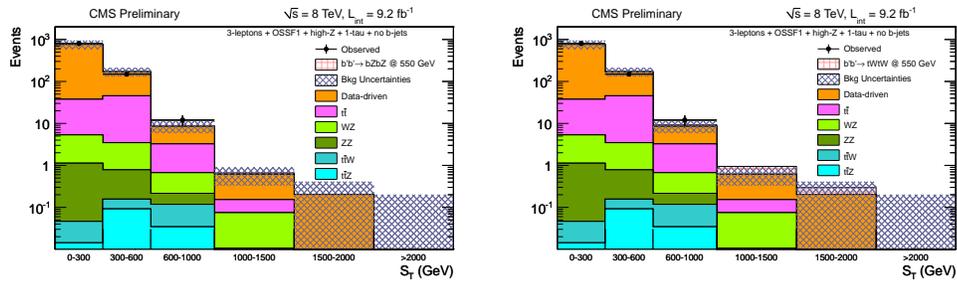
Figure B.11: 3-lepton + OSSF1 + above-Z + Tau1 + b0
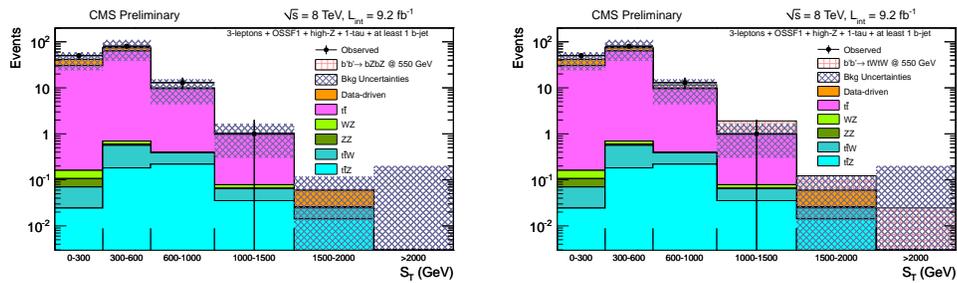
Figure B.12: 3-lepton + OSSF1 + above-Z + Tau1 + b1
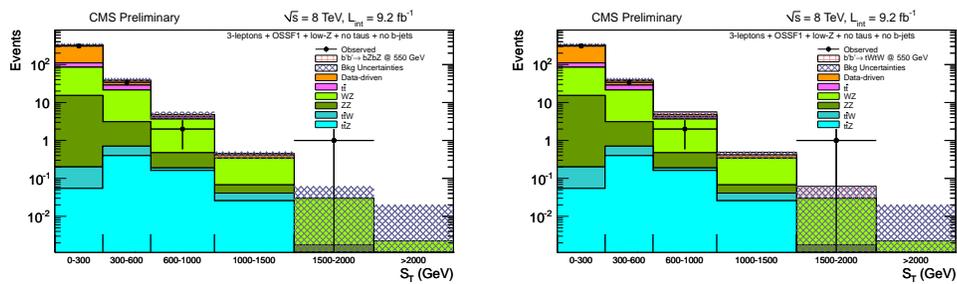
Figure B.13: 3-lepton + OSSF1 + below-Z + Tau0 + b0
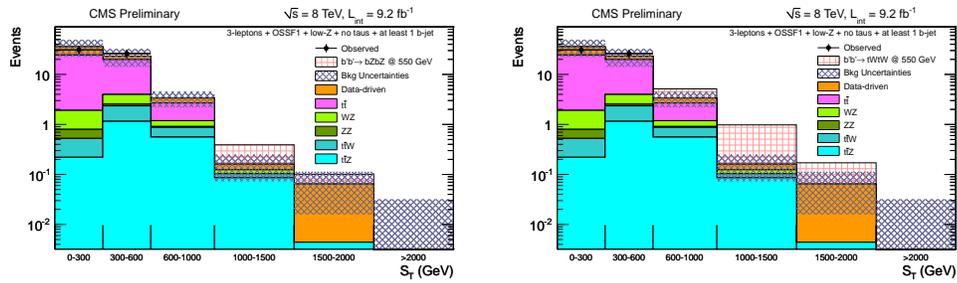
Figure B.14: 3-lepton + OSSF1 + below-Z + Tau0 + b1
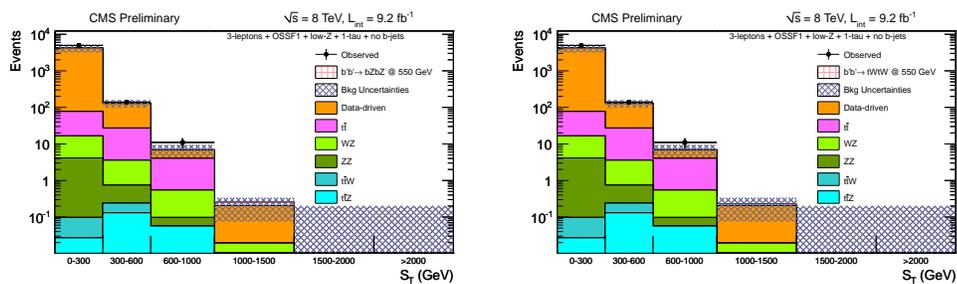
Figure B.15: 3-lepton + OSSF1 + below-Z + Tau1 + b0
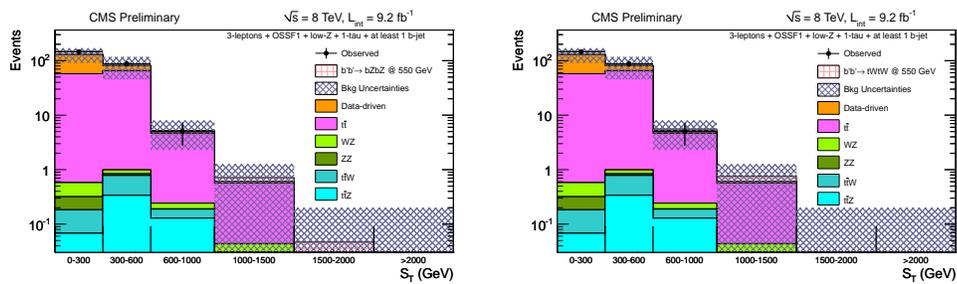
Figure B.16: 3-lepton + OSSF1 + below-Z + Tau1 + b1
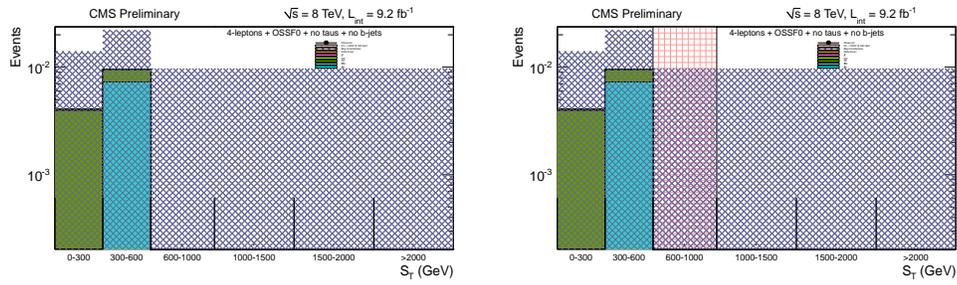
Figure B.17: 4-lepton + OSSF0 + on-Z + Tau0 + b0
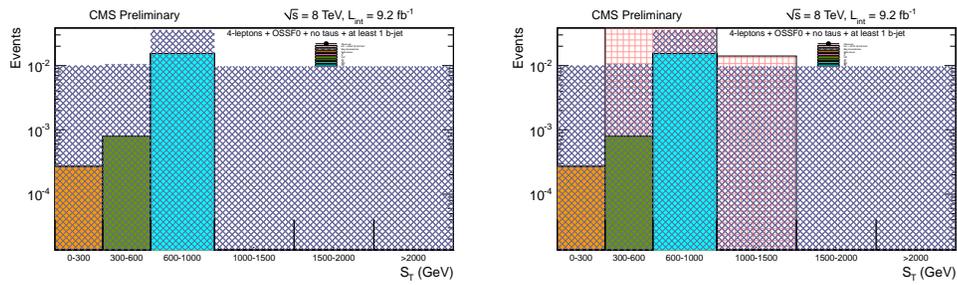
Figure B.18: 4-lepton + OSSF0 + on-Z + Tau0 + b1
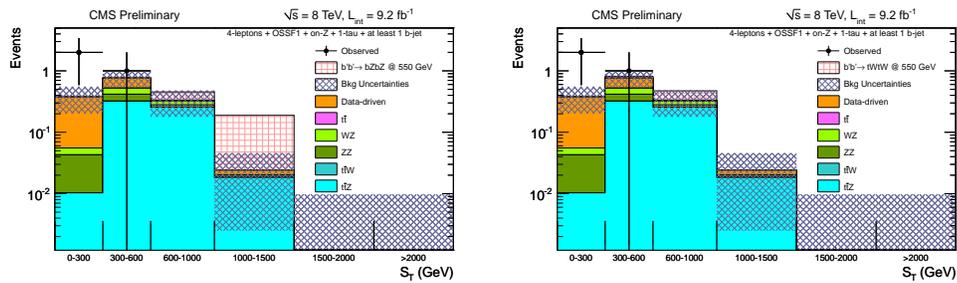

Figure B.19: 4-lepton + OSSF0 + on-Z + Tau1 + b0

Figure B.20: 4-lepton + OSSF0 + on-Z + Tau1 + b1

Figure B.21: 4-lepton + OSSF1 + on-Z + Tau0 + b0

Figure B.22: 4-lepton + OSSF1 + on-Z + Tau0 + b1

Figure B.23: 4-lepton + OSSF1 + on-Z + Tau1 + b0
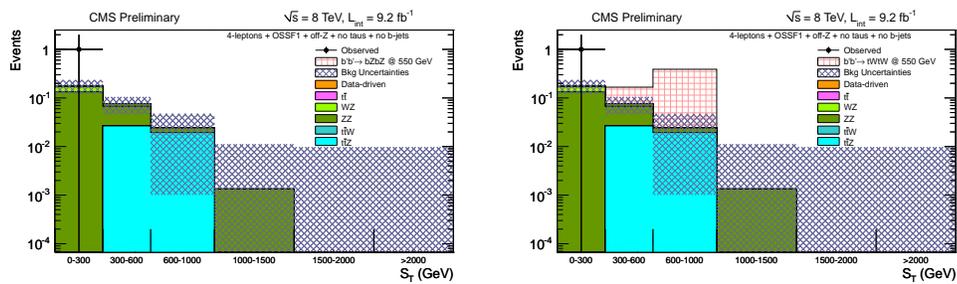

Figure B.24: 4-lepton + OSSF1 + on-Z + Tau1 + b1

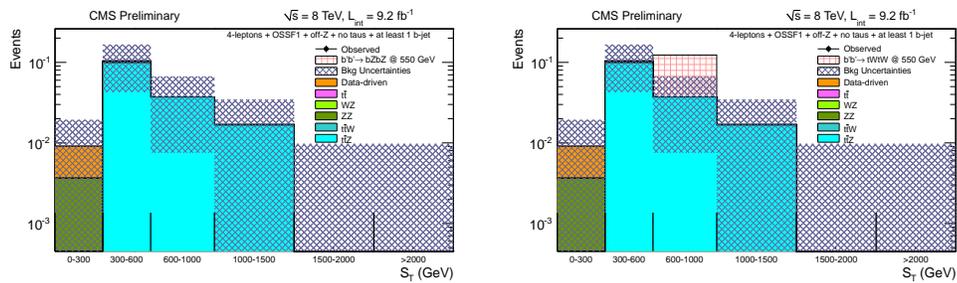Figure B.25: 4-lepton + OSSF1 + off-Z + Tau0 + b0

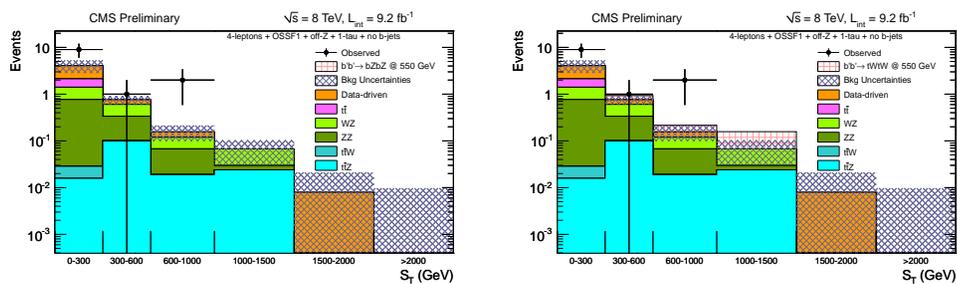Figure B.26: 4-lepton + OSSF1 + off-Z + Tau0 + b1

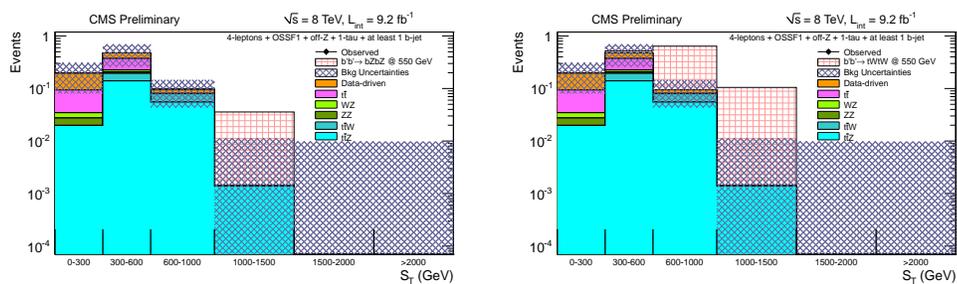

Figure B.27: 4-lepton + OSSF1 + off-Z + Tau1 + b0

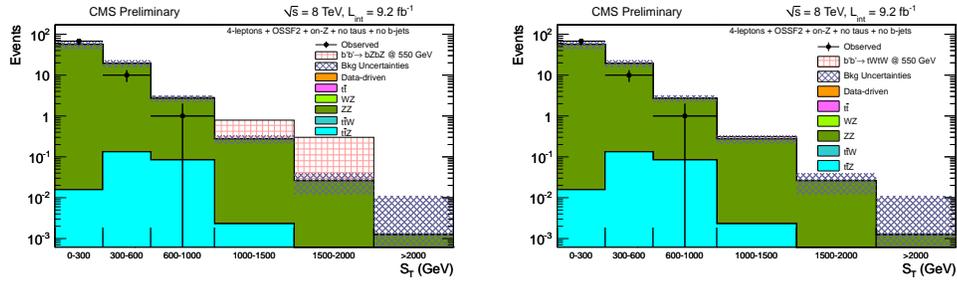Figure B.28: 4-lepton + OSSF1 + off-Z + Tau1 + b1
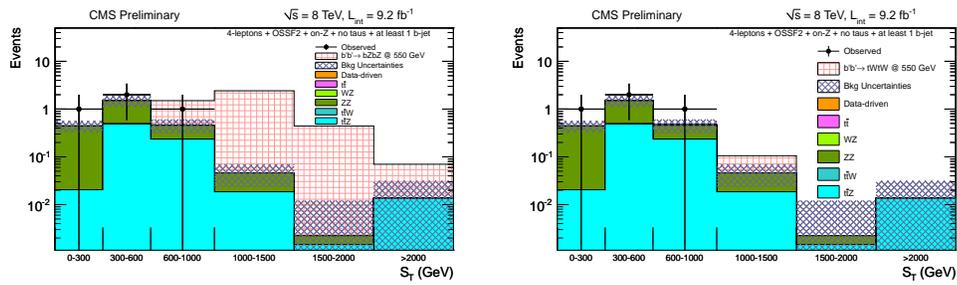
Figure B.29: 4-lepton + OSSF2 + on-Z + Tau0 + b0
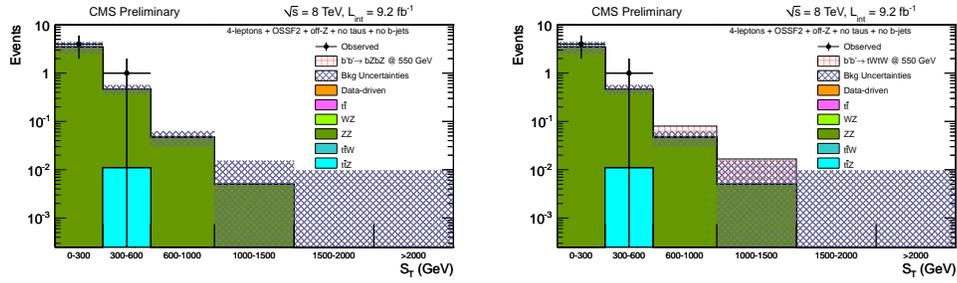

Figure B.30: 4-lepton + OSSF2 + on-Z + Tau0 + b1

Figure B.31: 4-lepton + OSSF2 + off-Z + Tau0 + b0
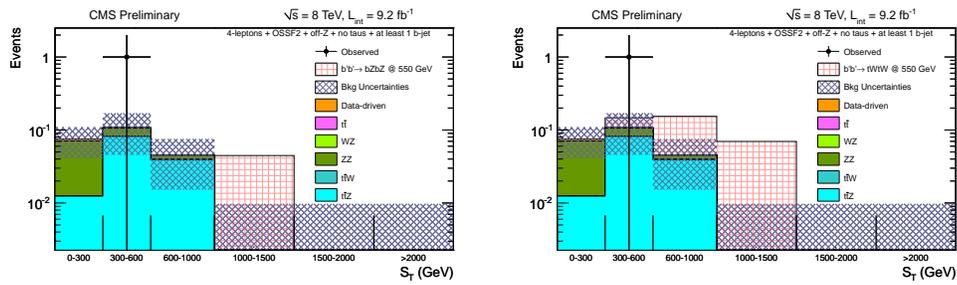
Figure B.32: 4-lepton + OSSF2 + off-Z + Tau0 + b1

Appendix C
List of Triggers Used

C.1   Trigger List 2012

DoubleMuon

- HLT_Mu17_Mu8_v*

Double Electron

- HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_Ele8_CaloIdT_CaloIsoV
  L_TrkIdVL_TrkIsoVL_v*

Muon-Electron

- HLT_Mu8_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*

- HLT_Mu17_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*

Single Electron

- HLT_Ele80_CaloIdVT_TrkIdT_v*

- HLT_Ele100_CaloIdVT_TrkIdT_v*

- HLT_Ele90_CaloIdVT_GsfTrkIdT_v*

Single Muon

- HLT_Mu40_eta2p1_v*

- HLT_Mu50_eta2p1_v*

- HLT_IsoMu24_eta2p1_v*

- HLT_IsoMu30_eta2p1_v*

- HLT_IsoMu34_eta2p1_v*

- HLT_IsoMu40_eta2p1_v*

Bibliography

[1] A. Soni, A.K. Alok, A. Giri et al.,"The fourth family: a natural explanation for the observed pattern of anomalies in B-CP asymmetries", arXiv:0807.1971

[2] R. Ciftci, Anomalous Single Production of the Fourth Generation Quarks at the LHC, arXiv:0807.4291.

[3] R. Ciftci, A. K. Ciftci, and S. Sultansoy, Possible impact of the fourth generation quarks on production of a charged Higgs boson at the LHC, arXiv:0807.0831.

[4] G. D. Kribs, T. Plehn, M. Spannowsky et al., Four generations and Higgs physics, Phys. Rev. D76 (2007) 075016, arXiv:0706.3718. doi:10.1103/PhysRevD.76.075016.

[5] B. Holdom, W.S. Hou, M. Mangano, S. Sultansoy, G. Unel, "Four Statements about the Fourth Generation", Summary of the "Beyond the 3-generation SM in the LHC Era" Workshop, CERN-PH-TH/2009-091.

[6] A. Arhrib, W.S. Hou, "Flavor Changing Neutral Currents involving Heavy Quarks with Four Generations", arXiv:hep-ph/0602035.

[7] D. Atwood, S.K. Gupta, A. Soni, "Detecting Fourth Generation Quarks at Hadron Colliders", arXiv:hep-ph:1104.3874.

[8] "Search for a vector-like quark of charge -1/3 and decaying to bZ in pp collisions at $\sqrt{(s)} = 7$ TeV, "CMS Physics Analysis Summary"

[9] "Search for heavy bottom-like quarks in 4.9 fb^{-1} of pp collisions at $\sqrt{(s)} = 7$ TeV", JHEP 1205 (2012) 123.

[10] "Search for RPV supersymmetry with three or more leptons and b-tags", "CMS Physics Analysis Summary"

[11] The CMS Collaboration, "The CMS experiment at the CERN LHC", JINST 3 (2008) S08004

[12] S. Agostinelli et al., "GEANT4: A Simulation toolkit", Nucl. Instrum. Meth. A506 250303 (2003), doi:10.1016/S0168-9002(03)01368-8

[13] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, M. Wiedermann, "HATHOR-HAdronic Top and Heavy quarks crOss section calculatoR", Comput.Phys.Commun.182:1034-1046,2011/arXiv:1007:1327.

[14] F. Maltoni et al., "MadEvent: Automatic event generation with MadGraph", JHEP 02, 027 (2003), arXiv:hep-ph/0208156v1

[15] T. Sjostrand, S. Mrenna, and P. Skands, "A Brief Introduction to PYTHIA 8.1", Comput. Phys. Commun. 178 (2008) 852, doi:10.1016/j.cpc.2008.01.036.

[16] P. M. Nadolsky et al., "Implications of CTEQ global analysis for collider observables", Phys. Rev. D 78 (2008) 013004, doi:10.1103/PhysRevD.78.013004

[17] ElectronIdentificationBasedonSimpleCuts,https://twiki.cern.ch/twiki/bin/viewauth/CMS/SimpleCutBasedEleID.

[18] N. Adam et al., Measurements of Inclusive W and Z Cross Sections in pp Collisions at sqrt(s) = 7 TeV, CMS AN-2010/116. J. Alcaraz Maestre et al., Updated Measurements of the Inclusive W and Z Cross Sections at 7 TeV, CMS AN-2010/264.

[19] Combined Fakeable Object (CFO) Method to Measure Lepton Fake Rates Using Multiple Data Objects", AN-2010/159 for SUS-10-008, "Search for new physics with multileptons at the LHC".

[20] A. Abulencia *et al.* [CDF Collaboration], "Search for neutral MSSM Higgs bosons decaying to tau pairs in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Rev. Lett. **96**, 011802 (2006)

[21] D. Jang, "Search for MSSM Higgs decaying to tau pairs in p anti-p collision at $sqrts = 1.96$ TeV at CDF,", Thesis, Rutgers U, Piscataway, FERMILAB-THESIS-2006-11.

[22] T. Aaltonen *et al.* [CDF Collaboration], "Search for Supersymmetry in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$-TeV Using the Trilepton Signature of Chargino-Neutralino Production," Phys. Rev. Lett. **101**, 251801 (2008)

[23] CMS Collaboration, "Study of tau reconstruction algorithms using pp collisions data collected at $\sqrt{s} = 7$ TeV", CMS Physics Analysis Summary, http://cms-physics.web.cern.ch/cms-physics/public/PFT-10-004-pas.pdf

[24] CMS Collaboration, "Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV","CMS Physics Analysis Summary", http://cms-physics.web.cern.ch/cms-physics/public/PFT-10-002-pas.pdf

[25] SUSY Multileptons, SUS-10-008.

[26] SUSY Multileptons, SUS-11-013.

[27] CMS Collaboration, "Missing Transverse Energy Performance in Minimum-Bias and Jet Events from Proton-Proton Collisions at $\sqrt{s} = 7\,$TeV", CMS Physics Analysis Summary, http://cdsweb.cern.ch/record/1279142/files/JME-10-004-pas.pdf

[28] "Background and Efficiency Determination Methods for Multilepton Analyses", AN-2012/257

[29] "Combined Fakeable Object (CFO) Method to Measure Lepton Fake Rates Using Multiple Data Objects", AN-2010/159 for SUS-10-008, "Search for new physics with multileptons at the LHC".

[30] "MET Resolution Dependence on Pileup", AN-2012/342

[31] M. Sher, "Fourth Generation b' decays into b + Higgs", arXiv:hep-ph/9908238v2, 1999.

[32] "A Search for Anomalous Production of Events with three or more leptons using 19.5 fb^{-1} of $\sqrt{s}$=8 TeV CMS data", CMS AN-2012/343

[33] "Search for heavy fourth-generation bottom-like quarks in s = 8 TeV pp Collisions with multilepton final states", CMS AN-2012/380

[34] Wikimedia Commons, http://commons.wikimedia.org/wiki/File:CMS_Slice.gif

[35] CERN website, http://cms.web.cern.ch/news/cms-detector-design

[36] Wikimedia Commons, http://commons.wikimedia.org/wiki/File:CMS_Slice.gif

[37] Dmitry Hits, A Multi-Lepton Search for New Physics in 35 pb^{-1} Proton-Proton Collisions at the LHC for a Center of Mass Energy $\sqrt{s} = 7$ TeV Using the CMS Detector, Ph.D. Thesis, Rutgers University, 2011.

[38] Frank Golf, Search for New Physics in a Final State with Same-Sign Dileptons, Jets, and Missing Transverse Energy at 7 TeV Center of Mass Energy, Ph.D. Thesis, University of California, San Diego, 2012.