

**Search for High Mass States Decaying to Tau Pairs with
the CMS Experiment at CERN**

by

A. A. Johnson

B.S., Carnegie Mellon University, 2011

M.S., University of Colorado, 2013

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Physics
2016

This thesis entitled:
Search for High Mass States Decaying to Tau Pairs with the CMS Experiment at CERN
written by A. A. Johnson
has been approved for the Department of Physics

Prof. John P. Cumalat

Prof. Kevin Stenson

Prof. Stephen Wagner

Date _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Johnson, A. A. (Ph.D., Physics)

Search for High Mass States Decaying to Tau Pairs with the CMS Experiment at CERN

Thesis directed by Prof. John P. Cumalat

Two searches for heavy neutral resonances, known as Z' 's, decaying to back-to-back tau pairs are completed using the Compact Muon Solenoid detector at the Large Hadron Collider. The first, based on data accumulated during the 2012 data-taking run at $\sqrt{s} = 8$ TeV, focuses on the fully-leptonic $\tau_e \tau_\mu$ channel, in which one tau decays to an electron and the other to a muon. The second, performed during the 2015 run at $\sqrt{s} = 13$ TeV, includes four major decay channels: the $\tau_e \tau_\mu$ channel, the $\tau_e \tau_h$ channel (in which one tau decays to an electron and the other decays hadronically into a jet), the $\tau_\mu \tau_h$ channel (in which one tau decays to a muon and the other decays hadronically), and the $\tau_h \tau_h$ channel (in which both taus decay into jets). Each search employs a hybrid of background estimation strategies using both Monte Carlo simulation and data-driven methods. The presence of neutrinos in the decays necessitates the use of missing energy in the final mass calculation. In both searches, observed data are found to be in agreement with Standard Model expectation and no excesses are observed. Limits are placed on the Z' mass at 95% confidence level such that, if the Z' exists, its mass must be greater than 2.1 TeV.

Dedication

To my parents, Carl and Sharon Johnson. Without your constant love, support, and encouragement, none of this would have been remotely possible.

Acknowledgements

I have often remarked that the fact that anything ever gets done at CERN is a miracle. This does not come from a place of cynicism, but rather from my profound sense of amazement at the sheer number of moving parts that must come together in order to operate the collider, run the detectors, analyze the data, and produce trustworthy results. Particle physics at CERN is indeed a global undertaking, and it requires the focused collaboration of thousands of staff scientists, faculty, PhD students, undergraduates, postdocs, engineers, support staff, and administrators in order to move the wheels of science along. Without them, forward progress would be impossible. In a similar sense, there are many, many people in my life without whom completing a PhD would have been impossible (or at the very least substantially more unpleasant).

First and foremost, I am forever indebted to my parents, Carl and Sharon Johnson. From an extremely early age, they both placed the utmost importance on academic achievement. My father, paragon of patience, happily spent countless hours (usually after a long day at the office) helping me through my math homework, and consistently challenged me to think analytically and to value a technical education. My mother, in addition to making me a much, much better writer, impressed upon me the interpersonal skills and social intelligence that allowed me to flourish in a multinational environment like CERN. I also owe a great deal to my delightful little sister, Madeline. Without her constant, whimsical-yet-bizarre influence growing up, I would be a much more boring person. Finally, thank you to my wonderfully-lovely girlfriend, Michelle, who made writing so much more bearable.

I cannot begin to express my gratitude towards the academic role models who helped mold

me into a physicist. Thank you to Mike Hazeltine for showing me, with that first AP physics exam, that physics is a challenge worth taking on. Thanks to Prof. Kunal Ghosh, who first lured me into the Carnegie Mellon physics department with brownies and a ridiculous laugh, and who cultivated my leadership skills through several student organizations. Thanks also to Prof. Jim Russ, my undergraduate advisor, who helped me through the grad school process and who fostered my love of high energy physics. I'm tremendously grateful to my supervisors at Argonne Lab: Sasha Paramonov, for showing me those first ropes, and Prof. Jimmy Proudfoot, whose letter of recommendation is without a doubt what got me into grad school. Sincere thanks to Eduardo Luiggi, my postdoc for the first few years of grad school. He was extremely patient with me, and guided me through literally every step of our 8 TeV analysis. This was absolutely critical in preparing me to do research on my own, and I am so, so grateful for the help. Eduardo, someday I will find you and someday I will get you to smile. I'm also very grateful to my BSM3G group members, who did an extraordinary job with the analyses presented here. In particular, thanks to Francesco Romeo, who helped me a great deal with tau ID & reconstruction, to Andres Florez, without whom I would never have been able to figure out the limit tools, and especially to Prof. Alfredo Gurrola, who developed a large amount of the software framework used in our analyses and who was an endless source of friendly advice and encouragement. Thanks also to Annapaola de Cosa and Christian Veelken for their invaluable guidance throughout all my service work. Thanks to Prof. Kevin Stenson, for teaching me that first particle physics class way back in 2011 and for letting bounce all those CMSSW questions off him, and thanks to Prof. Oliver DeWolfe for his friendship and for his willingness to assist me with all things high energy theory. Finally, thanks also to Doug Johnson for his tireless upkeep of the CU-HEP computing system, the backbone upon which all of this work rests.

There exists an ingrained stereotype of the miserable graduate student. I have never related to this, and I think the primary reason for that is my group of friends here in Colorado. I could not have asked for a more entertaining, supportive, and hilarious cohort of fellow grad students and I feel exceptionally lucky to have gone through this experience with each of them by my side. In no

particular order, they are: Jamey Szalay, Yomay Shyur, Andy Missert, Liz Shanblatt, Cathy Klauss, Jack Houlton, Andrew Koller, Victor Colussi, Scott Johnson, Frank Jensen, Oscar Henriksson, Brian O'Callahan, Sara Campbell, Bob Peterson, Ben Pollard, Brad Dinardo, Teesa Christian, and Casey Gardiner. Thanks for the ski days, wine clubs, trips to the Hill, and innumerable nights of laughter at Holmes Place.

Special thanks is also due to the CERN/Geneva contingent - a group of PhD students, postdocs, and undergrads from all over the world who made my time there unforgettable. They are: Laser Kaplan, Prof. Rachel Yohay, Prof. Ted Kolberg, Charlie Mueller, Sarah Buss, Grace Brown, Matthias Weber, Maria Hoffman, Vytautas Vislavicius, Shannon Ruse, Sarah Charley, Evan Wolfe, Stany Sevova, Andrew Hard, Christine McLean, Kira Burt, Siyuan Sun, Nate Woods, and Justin Pilot. Monday night trivia with Colonel Moutarde and Friday night shenanigans at Champs Frechet were an absolute blast, and I am going to sincerely miss the unwavering tradition that was 3:00 PM coffee in R1. Apologies are also due to Lindt & Sprüngli AG, who I am sure noticed a significant dip in revenues after my departure from Switzerland.

Finally, I am so unbelievably lucky to have had Prof. John Cumalat as an advisor. Research in experimental particle physics can be extremely frustrating, especially when the majority of one's collaborators are an ocean and eight time zones away. John was incredibly understanding of this, and worked tirelessly to get me the help I needed, to the point of sending me to CERN twice as well as to multiple schools and workshops. Indeed, John has been phenomenally supportive throughout my graduate research career, and his guidance and insights have truly taught me to think like a particle physicist. John, you are an outstanding advisor and friend. Thank you for everything.

Contents

Chapter

1	Introduction	1
1.1	Introduction	1
2	The Standard Model and Beyond	3
2.1	Particles of the Standard Model	3
2.2	Quantum Electrodynamics	6
2.3	Electroweak Theory	7
2.4	Quantum Chromodynamics	9
2.5	Spontaneous Symmetry Breaking and the Higgs Mechanism	10
2.6	Beyond the Standard Model	11
2.6.1	Models Predicting New Neutral Bosons	14
3	The Large Hadron Collider (LHC)	16
4	The Compact Muon Solenoid	20
4.1	The Silicon Tracker	22
4.1.1	The Pixel Detector	22
4.1.2	The Strip Tracker	25
4.2	The Electromagnetic Calorimeter	27
4.3	The Hadronic Calorimeter	28

4.4	The CMS Magnet	30
4.5	The Muon System	32
4.5.1	Drift Tubes (DTs)	32
4.5.2	Cathode Strip Chambers (CSCs)	34
4.5.3	Resistive Plate Chambers (RPCs)	34
4.6	Particle Reconstruction	37
4.6.1	Particle Flow	38
4.7	The Trigger System	45
4.7.1	The Level-1 (L1) Trigger)	45
4.7.2	The High level Trigger (HLT)	46
5	Search for $Z' \rightarrow \tau\tau$ events at $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV	47
5.1	Strategy	47
5.2	Data and Monte Carlo Samples	49
5.3	Physics Object Reconstruction	51
5.3.1	Jet Reconstruction	51
5.3.2	Electron Reconstruction and Identification	53
5.3.3	Muon Reconstruction and Identification	54
5.3.4	Tau Reconstruction and Identification	55
5.3.5	Missing Transverse Energy (\cancel{E}_T)	59
5.3.6	Corrections for Pile-Up	61
5.4	Triggers	61
5.4.1	Single Lepton Trigger Efficiency	63
5.4.2	Single Electron Trigger Efficiency	64
5.4.3	Single Muon Trigger Efficiency	64
5.4.4	Di-Tau Trigger Efficiency	67
5.5	Muon + Hadronic Tau Channel	70

5.5.1	W + Jets Background Estimation	81
5.5.2	DY + Jets Background Estimation	89
5.5.3	Other Small Backgrounds	91
5.6	The Fully-Hadronic Channel	93
5.6.1	QCD Background Estimation & Validation Strategy	94
5.6.2	Background Estimation for $Z(\rightarrow \tau\tau) + \text{Jets}$	100
5.7	Electron + Hadronic Tau Channel	101
5.7.1	Event selection	101
5.7.2	Genuine dilepton events	108
5.7.3	QCD multijet background	110
5.7.4	W+Jets background	112
5.7.5	Overlays of observations and SM predictions	115
5.8	Electron + Muon Channel	115
5.8.1	Event selection	115
5.8.2	Genuine dilepton events	117
5.8.3	Drell-Yan process	122
5.8.4	$t\bar{t}$ and single top processes	122
5.8.5	Diboson process	123
5.8.6	QCD multijet background	123
5.8.7	W+Jets background	123
5.8.8	Overlays of observations and SM predictions	123
5.9	Systematic Uncertainty	123
5.10	Results	129
5.10.1	13 TeV Results	129
5.10.2	8 TeV Results	132

6	Tau Lifetime Studies	137
6.1	Motivation	137
6.2	Methods	139
6.3	Impact on Limit	142
6.3.1	8 TeV Results in the $e\mu$ Channel	142
6.3.2	13 TeV Results in the $\tau_h\tau_h$ Channel	143
7	Summary, Conclusion, and What's Next for $Z' \rightarrow \tau\tau$	148
Bibliography		152

Tables

Table

5.1	Collision Data Samples	49
5.2	MC Samples used in 13 TeV analysis. The multiplicative terms take the cross section from leading order to NLO.	50
5.3	Loose Jet-ID Selections.	52
5.4	Electron ID selection criteria	54
5.5	μ Identification requirements	55
5.6	Reconstructed Tau Decay Modes	56
5.7	The trigger paths used to collected the data. Emulated trigger paths, in particular those most similar to the paths used to collect the data, are applied to the simulated samples.	63
5.8	The number of events in the mu-tau control region after requiring the denominator level selection criteria	70
5.9	Event summary table after signal region selection	79
5.10	Signal and background yields after various stages of the $\mu\tau_h$ selection.	80
5.11	Background and data yields in W + Jets control regions B, C, D , under nominal \cancel{E}_T conditions ($\cancel{E}_T > 30$ GeV).	85
5.12	Background and data yields in W + Jets control regions B, C, D obtained without a \cancel{E}_T requirement.	87
5.13	Events in the $Z \rightarrow \tau\tau$ control region for data and MC.	90

5.14	Yields in the $\mu\tau_h$ QCD control region for data and MC.	92
5.15	Yields in the control regions B and D used for calculation of OS/LS ratio.	97
5.16	Background and data yields in QCD control regions A and C under nominal isolation and \cancel{E}_T conditions (i.e. isolated + $\cancel{E}_T > 30$ GeV).	100
5.17	QCD yields in the isolated regions A (signal region), B, C, and D.	100
5.18	Summary of systematic uncertainties. Values are given in percent. “s” indicates template variations (“shape” uncertainties).	129
5.19	Number of observed events in data and estimated background events for the entire mass range. The uncertainties quoted on the number of background events represent the combined statistical and systematic uncertainty.	130
5.20	Number of observed events in data and estimated background events in the region $M(\tau_1, \tau_2, \cancel{E}_T) > 300$ GeV. The uncertainties quoted on the number of background events represent the combined statistical and systematic uncertainty.	131
5.21	Event summary table after signal region selection	132
5.22	Expected background rates compared to observed rates in bins of $m(e, \tau_h, \cancel{E}_T)$. Un- certainties are statistical in each bin.	135
5.23	Expected signal rates and associated significances in bins of $m(e, \tau_h, \cancel{E}_T)$	136

Figures

Figure

2.1	Table showing the quarks, leptons, gauge bosons (force carriers), and Higgs boson in the SM[82].	4
2.2	Typical QED interaction[51].	7
2.3	$Z' \rightarrow \tau\tau$ event at the LHC, wherein two quarks from colliding protons form a Z' which then decays to an oppositely-charged, back-to-back pair of τ leptons.	13
3.1	The Large Hadron Collider located underneath the Franco-Swiss border at CERN near Geneva, Switzerland. The CMS Experiment is located on the French side in Cessy, France	17
3.2	Total integrated luminosity delivered by the LHC (blue), and total integrated luminosity recorded by CMS (yellow) during the 2016 run. The difference is a measure of the experiment's efficiency.	19
4.1	The CMS Experiment (open for maintenance)	21
4.2	Schematic of the CMS Experiment showing the silicon trackers, electromagnetic calorimeter (ECAL), hadronic calorimeter (HCAL), forward calorimeters, superconducting solenoid, steel magnetic flux return yoke, and muon system.	23
4.3	Overview of the silicon tracker layout (projected onto the Rz plane, including the pixel layers and disks (red), the TIB and TOB layers, and the TID and TEC disks. .	24

4.4	Diagram indicating the layout of the barrel pixel (BPIX) sensors (in green) and endcap pixel (FPIX) sensors (in pink)	24
4.5	Hit efficiency vs. the average number of pp collisions per bunch crossing for pixel barrel layers and forward disks.	26
4.6	Distribution of hit residuals on the second layer of the pixel barrel in the transverse direction to the beam. The distribution is fitted with a student's t-function for which sigma is shown on the plot.	26
4.7	Layout of the electromagnetic calorimeter (ECAL) indicating the configuration of the barrel (EB) and endcap (EE) crystals. Numbers on the layout are in terms of η	29
4.8	Relative energy resolution vs $ \eta $ in EB (a) and EE (b) for both Monte Carlo (MC) and data. The vertical dashed lines indicate module boundaries in EB. The mismatch between data and MC is corrected by applying η -dependent correcting factors to subsequent models.	29
4.9	Layout of the hadronic calorimeter (HCAL) indicating the configuration of the barrel (HB) and endcap (HE) components as well as the outer HCAL (HO) lining the solenoid.	31
4.10	Layout of the muon system highlighting the layered structure of the drift tubes (DTs, green), cathode strip chambers (CSCs, blue), and resistive plate chambers (RPCs, red).	33
4.11	The measured DT efficiency as a function of the muon transverse momentum p_T , η , and ϕ . Results for the 4 stations are superimposed.	35
4.12	The measured CSC efficiency as a function of η , ϕ , and the muon transverse momentum p_T for ME1 (left) and ME2 (right) stations. The vertical lines on the p_T distributions divide the ranges between values covered by the J/ψ and Z probes.	36
4.13	Residuals (distance between “good” track and RPC hit) distribution in RB3 (the RPC associated with MB3 in the barrel of the muon system), fitted to a Gaussian.	37

5.1	Tau Decay Mode Finding efficiency as a function of p_T , η , and number of reconstructed vertices	58
5.2	Efficiency of tau anti-muon discriminators (“Loose” working point in red and “Tight” in blue) as a function of p_T , η , and number of reconstructed vertices	58
5.3	Efficiency of MVA-based anti-electron discriminators for taus (“Loose” working point in red and “Medium” in blue) as a function of p_T , η , and number of reconstructed vertices	59
5.4	Efficiency of isolation discriminator “CombinedIsoDB3Hits” for taus (“Loose” working point in red, “Medium” in blue and “Tight” in purple) as a function of p_T , η , and number of reconstructed vertices	60
5.5	Energy response of reconstructed taus, raw and vs. generated tau p_T	61
5.6	Distributions of the number of reconstructed primary vertices. Top: $e\mu$ channel. Bottom: $e\tau_h$ channel. Left: before reweighting. Right: after reweighting. The apparent substantial differences in the QCD “integrals” are artifacts of the plotting procedure, in which bins with negative contributions are zero-ed.	62
5.7	The efficiency vs. p_T curves of the single-electron triggers used. Left column: barrel. Right column: endcap. The offline electron ID requirements used are HEEP (top row), MVANonTrig80 (middle row), MVATrigWP80 (bottom row).	65
5.8	The efficiency vs. p_T curves of the single-electron triggers used. Left column: barrel. Right column: endcap. The offline electron ID requirements used are HEEP (top row), MVANonTrig80 (middle row), MVATrigWP80 (bottom row). The two bins each contain half of the events with $p_T > 35$ GeV.	66
5.9	The efficiency vs. p_T curves of the single-muon triggers used. Left column: $ \eta < 1.2$. Right column: $ \eta > 1.2$. Bottom row: the two bins each contain half of the events with $p_T > 25$ GeV.	68
5.10	The per-leg τ_h trigger efficiency as a function of $p_T(\tau_h)$ for simulated Z' signal sample (left) and for data and all backgrounds obtained by requiring mu-tau region (right).	70

5.11 The distributions of $p_T(\mu)$ (left) and $p_T(\tau_h)$ (right) after acceptance, μ identification, and τ_h identification cuts.	72
5.12 Left: \cancel{E}_T after acceptance, μ identification, and τ_h identification cuts. Right: $P_\zeta - 3.1 \times P_\zeta^{vis}$ after acceptance, μ identification, and τ_h identification cuts.	73
5.13 Sketch depicting the W+Jets rejection power of a $\Delta\phi$ cut.	74
5.14 Definitions for p_ζ and p_ζ^{vis}	76
5.15 p_ζ vs. p_ζ^{vis} for $Z' \rightarrow \tau\tau$ (left) and for W+Jets (right).	76
5.16 Left: $\cos\Delta\phi$ between the muon and tau/jet directions. Right: $m(\mu, \tau_h, \cancel{E}_T)$ for signal Z' and various backgrounds. Both plots normalized to unity.	79
5.17 $p_T(\tau_h)$, $\eta(\tau_h)$, \cancel{E}_T , and $m_T(\mu, \cancel{E}_T)$ control plots using events satisfying the “pre-selection” requirements. The background predictions are entirely based on MC. Agreement between data and MC is good.	82
5.18 $p_T(\tau_h)$, $\eta(\tau_h)$, \cancel{E}_T , and $m_T(\mu, \cancel{E}_T)$ control plots using events satisfying the signal region requirements. The background predictions are entirely based on MC. Agreement between data and MC is good.	83
5.19 The four regions used in the “ABCD” method of estimating W+Jets background . .	83
5.20 Left: Comparison of the ζ distribution in W + Jets MC, normalized to unity, for events with isolated and non-isolated τ_h . Right: Comparison of the $\cos\Delta\phi$ distribution in W + Jets MC, normalized to unity, for events with isolated and non-isolated τ_h	85
5.21 Top: W + Jets closure test performed with simulation (normal scale to focus on low mass) under nominal \cancel{E}_T conditions (i.e. $\cancel{E}_T > 30$ GeV). Middle: W + Jets closure test performed with simulation requiring $\cancel{E}_T > 0$ GeV. Bottom: Comparison of the reconstructed mass distributions in regions A and B. Left: normal scale to focus on low mass. Right: Log scale to focus on the high mass tails.	86
5.22 $m_T(\mu, \cancel{E}_T)$ distributions in W + Jets control region D (left) and B (right) with the nominal \cancel{E}_T cut of $\cancel{E}_T > 30$ GeV (top) and with no \cancel{E}_T cut (bottom).	88

5.23 $m(\mu, \tau_h, \cancel{E}_T)$ distribution in the $Z \rightarrow \tau\tau$ control region defined by $\cancel{E}_T < 30$ GeV and $m(\mu, \tau_h, \cancel{E}_T) < 300$ GeV, with otherwise similar selection criteria to the signal region. The $Z \rightarrow \tau\tau$ (DY+Jets) events are shown in blue.	90
5.24 QCD estimation and validation strategy for the $\tau_h \tau_h$ channel.	95
5.25 $\tau_h - \tau_h$ mass distribution in isolated OS, low- \cancel{E}_T sideband (region B). Left: linear scale. Right: log scale.	97
5.26 $m(\tau_h, \tau_h, \cancel{E}_T)$ mass distribution in anti-isolated OS, low- \cancel{E}_T sideband (region 2B) Left: linear scale. Right: log scale.	98
5.27 $m(\tau_h, \tau_h, \cancel{E}_T)$ mass distribution in anti-isolated OS region with nominal \cancel{E}_T (Left: linear scale. Right: log scale).	99
5.28 $m(\tau_h \tau_h)$ in the $Z \rightarrow \tau\tau$ control region obtained using the “pre-selection” cuts, and additionally requiring $\tau_h \tau_h$ pairs with invariant mass less than 100 GeV.	102
5.29 Distributions, after preselection, of the variables used for $e\tau_h$ signal selection: \cancel{E}_T (top left), “ p_ζ ” (top right), $\cos \Delta\phi(e, \tau_h)$ (bottom left), and n_b (bottom right).	104
5.30 Distributions, after $e\tau_h$ preselection, of electron p_T (top left), electron pseudorapidity (top right), $\tau_h p_T$ (bottom left), τ_h pseudorapidity (bottom right).	105
5.31 Distributions, after $e\tau_h$ preselection, of $m_T(e, \cancel{E}_T)$ (left), $m(e, \tau, \cancel{E}_T)$ (right).	106
5.32 Distributions of the variables used for $e\tau_h$ signal selection, after all other signal selection requirements on variables other than the one plotted: \cancel{E}_T (top left), “ ζ ” (top right), $\cos \Delta\phi(e, \tau_h)$ (bottom left), and n_b (bottom right).	107
5.33 Distributions of $m(\tau_e, \tau_h, \cancel{E}_T)$. Left: validation region with $\cancel{E}_T < 30$ GeV, $n_b = 0$ $m(\tau_e, \tau_h, \cancel{E}_T) < 200$ GeV. Right: validation region with $n_b \geq 1$	109
5.34 QCD estimation strategy for the $e\tau_h$ channel.	110
5.35 The distribution of reconstructed parent mass, $m(\tau_e, \tau_h, \cancel{E}_T)$, in the same-charge, tightly-isolated sample: $e\tau_h$ channel.	111

5.36 Left: comparison of the simulated distributions of $m(\tau_e, \tau_h, \cancel{E}_T)$, using tight τ_h isolation and with relaxed τ_h isolation, in the control region described in Section 5.7.4. Center: analogous comparison for the signal region. Right: comparison in the signal region of the “tighter” and “looser” halves of the sideband.	112
5.37 Distributions of $m_T(e, \cancel{E}_T)$ obtained in the W-enriched control region, when scanning the W+Jets event rate from 1.0 to 1.25 times its nominal value, in steps of 0.05. For each assumed W+Jets event rate, the QCD prediction is determined anew using the procedure described in Section 5.7.3.	113
5.38 $e\tau_h$ channel: the distribution of reconstructed parent mass, $m(\tau_e, \tau_h, \cancel{E}_T)$, in the W-validation region described in Section 5.7.4.	114
5.39 Left: predicted background yields and observed event yields in the $e\tau_h$ channel after signal selection. Right: the distribution of transverse mass.	115
5.40 Distributions, after $e\tau_h$ final selection, of electron p_T (top left), electron pseudorapidity (top right), $\tau_h p_T$ (bottom left), τ_h pseudo-rapidity (bottom right).	116
5.41 Distributions, after preselection, of the variables used for the $e\mu$ signal selection: \cancel{E}_T (top left), “ ζ ” (top right), $\cos \Delta\phi(e, \mu)$ (bottom left), and n_b (bottom right).	118
5.42 Distributions, after $e\mu$ preselection, of electron p_T (top left), electron pseudorapidity (top right), muon p_T (bottom left), muon pseudorapidity (bottom right).	119
5.43 Distributions, after $e\mu$ preselection, of $m(e, \mu, \cancel{E}_T)$ (top), $m_T(e, \cancel{E}_T)$ (bottom left), and $m_T(\mu, \cancel{E}_T)$ (bottom right).	120
5.44 Distributions of the variables used for $e\mu$ signal selection, after all other signal selection requirements on variables other than the one plotted: \cancel{E}_T (top left), “ ζ ” (top right), $\cos \Delta\phi(e, \mu)$ (bottom left), and n_b (bottom right).	121
5.45 Distributions of $m(\tau_e, \tau_\mu, \cancel{E}_T)$. Left: validation region with $\cancel{E}_T < 30$ GeV, $n_b = 0$ and $m(\tau_e, \tau_\mu, \cancel{E}_T) < 125$ GeV. Right: validation region with $\cancel{E}_T > 30$ GeV, $n_b \geq 1$ and $P_\zeta - 3.1 \times P_\zeta^{vis} > -50$ GeV.	124

5.46 The distribution of reconstructed parent mass, $m(\tau_e, \tau_\mu, \cancel{E}_T)$, in the same-sign, tightly-isolated sample: $e\mu$ channel.	125
5.47 Predicted background yields and observed event yields in the $e\mu$ channel after signal selection.	125
5.48 Distributions, after $e\mu$ final selection, of electron p_T (top left), electron pseudorapidity (top right), muon p_T (bottom left), muon pseudorapidity (bottom right).	126
5.49 Effective mass plots in the signal region after unblinding. Top Left: $m(\mu, \tau_h, \cancel{E}_T)$. Top Right: $m(\tau_h, \tau_h, \cancel{E}_T)$. Bottom Left: $m(e, \tau_h, \cancel{E}_T)$. Bottom Right: $m(e, \mu, \cancel{E}_T)$. All plots in log scale.	130
5.50 Expected and observed limits for the $\tau_\mu\tau_h$, $\tau_h\tau_h$, $\tau_e\tau_h$, and $\tau_e\tau_\mu$ channels for the 13 TeV search. A k factor of 1.3 has been used to scale the leading order (LO) signal cross-section.	133
5.51 Combined expected limit for the $\tau_\mu\tau_h$, $\tau_h\tau_h$, $\tau_e\tau_h$, $\tau_e\tau_\mu$ channels for the 13 TeV search. A k factor of 1.3 has been used to scale the leading order (LO) signal cross-section.	134
5.52 $m(e, \tau_h, \cancel{E}_T)$ distribution for the SR in the 8 TeV search. A 1.5 TeV Z'_{SSM} sample is also included.	135
5.53 8 TeV 95% CL upper limit on $\sigma(pp \rightarrow Z') \times BR(Z' \rightarrow \tau\tau)$ as a function of Z' mass. The color bands on the expected limits represent one standard deviation (green) and two standard deviations (yellow).	135
6.1 1.5 TeV Z' generated energy (blue) compared to reconstructed visible mass, $m(\tau_1, \tau_2, \cancel{E}_T)$ (red). The $m(\tau_1, \tau_2, \cancel{E}_T)$ distribution is broader and peaked at a much lower energy than the MC truth.	138
6.2 A ditau event (top) will produce decay products whose tracks miss the PV, whereas leptonic decays from a shorter-lived parent produced in the pp collision, such as a W^+W^- event (bottom), will produce “prompt” leptons which can be traced back to the PV.	138

6.3	Plot showing the performance of the IP definition of the lifetime cut. The colored lines indicate MC efficiencies and correspond to the left axis. The black triangles indicate $\frac{s}{\sqrt{s+b}}$ and correspond to the right axis. A threshold value of 2 was chosen for the cut since it is the first value on the plateau of $\frac{s}{\sqrt{s+b}}$ maxima, thereby keeping signal MC acceptance as high as possible.	140
6.4	Plot showing the performance of the DCA definition of the lifetime cut. The colored lines indicate MC efficiencies and correspond to the left axis. The black triangles indicate $\frac{s}{\sqrt{s+b}}$ and correspond to the right axis. A threshold value of 2.5 was chosen for the cut since it is the first value on the plateau of $\frac{s}{\sqrt{s+b}}$ maxima, thereby keeping signal MC acceptance as high as possible.	141
6.5	Example of where the DCA cut could reject a valid signal event - the two lepton tracks have a low DCA despite a high IP	142
6.6	Plot showing the performance of the DCA OR IP-based definition of the lifetime cut. The colored lines indicate MC efficiencies and correspond to the left axis. The black triangles indicate $\frac{s}{\sqrt{s+b}}$ and correspond to the right axis. A threshold value of 2 was chosen for the cut since it is the first value on the plateau of $\frac{s}{\sqrt{s+b}}$ maxima, thereby keeping signal MC acceptance as high as possible. Note that the performance in this cut, in terms of signal acceptance and background rejection, is higher than either the IP-based or DCA cuts individually.	142
6.7	Limit calculated with standard 8 TeV selection sequence (top) compared with limit calculated with the standard selection sequence and a lifetime threshold of 2 applied (bottom). The addition of the lifetime cut improves the exclusion limit by about 70 GeV.	143
6.8	Distribution of combined IP/ σ for a 2.5 TeV $Z' \rightarrow \tau_h \tau_h$ sample in the SR (top) and data (used as a proxy for QCD) in the SS QCD CR (bottom). A threshold value of 3 is selected for the limit comparison.	144

6.9 Limit calculated with standard 13 TeV selection sequence (top) compared with limit calculated with the standard selection sequence and a lifetime threshold of 3 applied (bottom).	146
6.10 Limit comparison with the p_ζ requirement removed and the \cancel{E}_T minimum threshold increased to 50. The top plot has no lifetime threshold applied, while the bottom plot has a lifetime threshold of 2 applied. While the removal of p_ζ and the tightening of the \cancel{E}_T threshold increases the baseline limit, the limit with the lifetime cut added is still lower (worse) than that without it applied.	147

Chapter 1

Introduction

1.1 Introduction

The mission of particle physics is to increase our understanding of the most fundamental constituents of our universe and their interactions. These efforts have arguably been going on since the era of the Greek philosophers, who classified all matter into the elemental categories of earth, air, water, and fire. As the centuries progressed, our knowledge of the nature of the universe has been refined as new theories were proposed to explain the phenomena we observe in nature, and, at the same time, more and more sophisticated experiments were designed to test them. Particle physicists are broadly divided into two sub-fields: theorists and experimentalists. Theorists seek to develop models that offer a more complete explanation of particle interactions, and experimentalists are tasked with validating these models. Efforts in particle theory in the late-20th century culminated in the Standard Model (SM) of particle physics, which has proven to be tremendously successful at describing a large number of observed particle phenomena.

The SM is not without its deficiencies, however. Today, there remain many open questions in particle physics that are not sufficiently (or at all) addressed by the SM. To name just a few: Why is there more matter than antimatter in the universe? What is dark matter? What is dark energy? Can the strong interaction be unified with the electroweak interaction? Why is the Higgs mass 125 GeV and not at the Planck scale?

Attempts to answer these questions require new theories modifying and building on the SM. These theories must then be tested, which is where the experimentalists come in. While the SM

was perhaps the crowning achievement of high energy theory in the twentieth century, the Large Hadron Collider (LHC) is arguably the most noteworthy undertaking in high energy experiment so far in the twenty-first. Located in Geneva, Switzerland at the European Organization for Nuclear Research (CERN, from the French “Conseil Européen pour la Recherche Nucléaire”), the LHC is the highest-energy particle collider ever built and has been an invaluable tool in the quest to validate theories of new physics.

This thesis represents an effort using one of the primary experiments on the LHC, the Compact Muon Solenoid (CMS), to search for a new particle, called the Z' , predicted by many such “Beyond Standard Model” (BSM) theories. Chapter 2 lays the groundwork, offering an overview of the SM and then a discussion of the modifying theories which predict the Z' ’s existence. Chapters 3 and 4 discuss the tools used in the search: the LHC, the CMS experiment, and the substantial computing resources needed to conduct a full search for new physics. Chapter 5 discusses the searches for Z ’s decaying to tau pairs at both $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV. Chapter 6 features proposed improvements to the analysis exploiting the lifetime of the tau, and Chapter 7 offers an overview of future possibilities in this search effort as well as concluding remarks.

Chapter 2

The Standard Model and Beyond

The Standard Model (SM) of particle physics is the set of theories describing subatomic particles and their interactions according to the electromagnetic, weak, and strong forces. Describing these interactions individually are the theory of quantum electrodynamics (QED), which describes the electromagnetic interaction, electroweak theory (EW), which encompasses QED and adds the weak interaction to form a broader theory of both interactions, and finally the theory of quantum chromodynamics (QCD), which describes the strong interaction. These individual theories combine to form the symmetry group describing the SM:

$$SU(3)_C \times SU(2)_L \times U(1)_Y \quad (2.1)$$

where $SU(3)_C$ is the symmetry group describing QCD, while $SU(2)_L \times U(1)_Y$ together are the symmetry groups describing electroweak theory, the unification of the electromagnetic and weak interactions.

This chapter will begin with an overview of the fundamental particles in the SM, and then will provide a brief discussion of each of the constituent theories listed above. It will conclude with additional sections on specific elements of the SM which are especially relevant to searches for new physics at hadron colliders.

2.1 Particles of the Standard Model

Shown in Figure 2.1, the particles included in the SM are believed to be fundamental (that is, they cannot be broken down into smaller constituent particles). Broadly, these particles can be

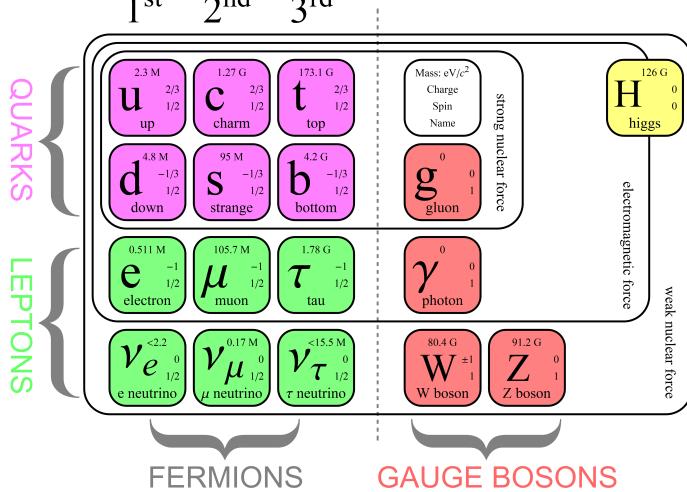


Figure 2.1: Table showing the quarks, leptons, gauge bosons (force carriers), and Higgs boson in the SM[82].

divided into spin-1/2 particles, known as fermions, and integer-spin particles, known as bosons.

Fermions make up the matter of our universe. They are further divided into quarks and leptons. Each of these particles also has an antiparticle partner with identical mass and spin but opposite electric charge (except for antineutrinos, which are electrically neutral). Quarks combine with other quarks to form *hadrons*. Hadrons are divided into *baryons* and *mesons*. Baryons are three-quark states such as the protons and neutrons which form atomic nuclei. Mesons are quark-antiquark pairs such as pions and kaons. Quarks have electric charge and interact via the electromagnetic and weak forces, and they also have color charge, meaning they interact via the strong force as well. This strong force is what keeps baryons and mesons held together and will be discussed further in the section on QCD. Quarks are divided into pairs of positive and negative charge across three “generations” of increasing mass. In each pair, the positively-charged quark has electric charge $+2/3$ and the negatively-charged quark has electric charge $-1/3$. The first generation consists of the positively-charged *up* (*u*) quark and the negatively-charged *down* (*d*) quark. These are the quarks most commonly found in matter as they make up protons and neutrons as well as the pions. The second generation consists of the positively-charged *charm* (*c*) quark and the negatively-charged *strange* (*s*) quark. These quarks form more exotic hadrons such as kaons and

lambdas. The third generation consists of the positively-charged *top* (*t*) quark and the negatively-charged *bottom* (*b*) quark. These are quite heavy, quite short-lived quarks. The *b* quark is found in some exotic bound states, but the *t* quark is so short-lived that it is never found in bound states as it decays too quickly.

While the quarks interact via both the strong and electroweak forces, the leptons only experience the electroweak interaction. The nature of this interaction will be discussed further in the sections on QED and EW theory. Like the quarks, the leptons are divided into three generations. Each generation consists of a negatively-charged particle with charge -1 , and a companion neutrino which is electrically-neutral. The first generation consists of the electron (*e*), the most commonly-found lepton forming shells around nuclei to create atoms and being responsible for electric current, and its corresponding neutrino, the electron neutrino (ν_e). The second generation consists of the muon (μ) and its corresponding neutrino, the muon neutrino (ν_μ). Muons are heavier than electrons and decay primarily into an electron, a muon neutrino, and an electron antineutrino. The mean lifetime of a muon is about $2 \mu\text{s}$ (in the reference frame of the muon). The third generation consists of the tau (τ) and its corresponding neutrino, the tau neutrino (ν_τ). Taus are heavier still and can decay either leptonically or hadronically. In both cases, the tau decays directly into a tau neutrino and a virtual *W* boson matching the sign of the tau. In the hadronic case, the *W* decays into a quark-antiquark pair which quickly hadronizes into jets, and in the leptonic case, the *W* decays into either an electron and electron antineutrino or a muon and a muon antineutrino. The mean lifetime of a tau is about 0.3 ps (in the reference frame of the tau)[72].

The bosons act as the force carriers of the SM. The photon (γ) mediates the electromagnetic interaction, the gluon (g) mediates the strong interaction, and the *W* and *Z* bosons mediate the weak interaction. The *W* can have electric charge $+1$ (W^+) or -1 (W^-), while the *Z*, photon, and gluon are all electrically neutral. The *W* and *Z* are massive, while the photon and gluon are massless. These bosons, known as gauge bosons or vector bosons, each have spin 1. The final boson, the scalar (spin 0) Higgs boson, is the final boson in the Standard Model and is a consequence of the Higgs field (discussed further in the section on the Higgs mechanism).

2.2 Quantum Electrodynamics

Quantum electrodynamics (QED) describes the interactions between particles having electric charge and the photons which mediate the electromagnetic force. A massive, unconstrained, fermionic field $\psi(x)$ with mass m (such as an electron) may be described according to the Dirac Lagrangian[78]:

$$\mathcal{L}_{\text{Dirac}} = i\bar{\psi}\gamma^\mu (\partial_\mu + ieA_\mu)\psi - m\bar{\psi}\psi \quad (2.2)$$

where the γ^μ are Dirac matrices representing Lorentz transformations on the Dirac spinors ψ , the A_μ correspond to the photon field, and e is the electron charge ($\approx 1.6 \times 10^{-19}$ Coulombs.) This Lagrangian corresponds to observable fields, so it must be invariant under gauge transformations. This Lagrangian is invariant under Lorentz transformations, and we define it to be invariant under the transformation

$$\psi(x) \rightarrow e^{ie\alpha(x)}\psi(x) \quad (2.3)$$

where $\alpha(x)$ is some position-dependent angle. This $U(1)_Q$ gauge invariance corresponds to conservation of electric charge.

To transform $\psi(x)$ in position to $\psi(y)$, we define the scalar operator $U(x, y)$ such that

$$\psi(y) = U(x, y)\psi(x) \quad (2.4)$$

In order for the Lagrangian to remain invariant, we must further have

$$A_\mu(x) \rightarrow A_\mu(x) - \frac{1}{e}\partial_\mu\alpha(x). \quad (2.5)$$

Now, if we define

$$D_\mu \equiv \partial_\mu + ieA_\mu \quad (2.6)$$

and the electromagnetic field strength tensor as

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu \quad (2.7)$$

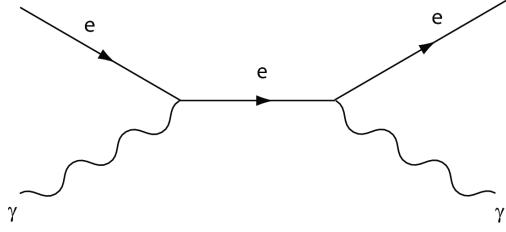


Figure 2.2: Typical QED interaction[51].

we can construct a Lagrangian that is fully gauge invariant. Note that this gauge invariance requires that there be no mass term $A_\mu A_\mu$, ensuring that the photon is massless. Putting everything together, we arrive at the final Lagrangian for QED [57]:

$$\mathcal{L}_{\text{QED}} = \bar{\psi} (i\gamma^\mu D_\mu - m) \psi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}. \quad (2.8)$$

Where $F_{\mu\nu} F^{\mu\nu}$ is the kinetic term for the photon. A typical QED interaction, a photon scattering off of an electron, is shown in Figure 2.2.

2.3 Electroweak Theory

The weak force is responsible for radioactive decay and is mediated by the W^\pm and Z bosons. It is understood in terms of its unification with electromagnetism known as Electroweak theory. Both quarks and leptons can interact via the weak force, with W^\pm and Z bosons decaying to both.

The symmetry group of EW theory is $SU(2)_L \times U(1)_Y$, where $SU(2)_L$ is non-Abelian with generators $I_i = \sigma_i$ where σ_i are the standard Pauli matrices, and where $U(1)_Y$ is Abelian with generator $Y/2 = Q - I_3$ where Q is the electric charge. The corresponding gauge transformation for $SU(2)_L$ is

$$U_L(\alpha) = e^{\frac{i\sigma_j \alpha_j}{2}} \quad (2.9)$$

The corresponding gauge bosons are the W_1 , W_2 , and W_3 from $SU(2)_L$, and the B from $U(1)_Y$ and are all massless. Through spontaneous symmetry breaking via the Higgs mechanism, these four massless bosons become the massive W^\pm and Z and the massless photon of the SM.

Before electroweak symmetry breaking, the EW Lagrangian is defined to be[78]

$$\mathcal{L}_{\text{EW}} = -\frac{1}{4}W_{\mu\nu}^i W_i^{\mu\nu} - \frac{1}{4}B_{\mu\nu} B^{\mu\nu} + \bar{f}_L \gamma^\mu \left(i\partial_\mu - \frac{g}{2}\sigma_i W^i - g' \frac{Y}{2} B_\mu \right) f_L + \bar{f}_R \gamma^\mu \left(i\partial_\mu - g' \frac{Y}{2} B_\mu \right) f_R \quad (2.10)$$

where we define:

$$W_{\mu\nu}^i \equiv \partial_\mu W_\mu^i - \partial_\nu W_\mu^i - g\epsilon_{ijk} W_\mu^j W_\nu^k \quad (2.11)$$

and

$$B_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu. \quad (2.12)$$

Here, g and g' are the respective coupling constants for $SU(2)_L$ and $U(1)_Y$, and f_R and f_L represent right- and left-handed fermionic fields, respectively. “Handedness,” or helicity, is simply the dot product of the momentum and spin unit vectors of the particle. Right-handed particles have helicity $+1$, and left-handed particles have helicity -1 .

After spontaneous symmetry breaking, the massless W_1 and W_2 combine to form the massive W^\pm via the relation

$$W^\pm = \frac{1}{\sqrt{2}} (W_1 \mp iW_2) \quad (2.13)$$

and the massless B and W_3 combine to form the massive Z and massless photon via the relation

$$\begin{pmatrix} A \\ Z \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix} \quad (2.14)$$

where θ_W is the weak mixing angle, or Weinberg angle, which defines the rotation by which spontaneous symmetry breaking transforms the $W_{1,2,3}$ and B into the W^\pm , Z , and photon of the SM. In terms of the coupling constants of $SU(2)_L$ (g) and $U(1)_Y$ (g'), the weak mixing angle may be defined as[78]:

$$\cos \theta_W = \frac{g}{\sqrt{g^2 + g'^2}} \quad (2.15)$$

and empirically as [57]

$$\cos \theta_W = \frac{m_W}{m_Z}. \quad (2.16)$$

2.4 Quantum Chromodynamics

Quantum chromodynamics (QCD) is the theory that describes the strong force. QCD introduces a new property, called “color,” that is only felt by particles which strongly interact. These particles are quarks (q , the constituents of baryons and mesons) and gluons (g , the mediator of the strong force). Particles can have color charge “red”, “green”, or “blue” according to rotations in the group $SU(3)_C$, the portion of the SM relevant to QCD.

The generators of $SU(3)$ are eightfold, and can be used to represent the gauge transformation $U(\alpha)$:

$$U(\alpha) = e^{i\alpha_j T^j} \quad (2.17)$$

They satisfy the commutation relation:

$$[T^a, T^b] = i f^{abc} T^c \quad (2.18)$$

where

$$f^{abc} \equiv -2i\text{Tr} \left(T^a T^b T^c - T^b T^a T^c \right). \quad (2.19)$$

We can write the Lagrangian of QCD as[57]:

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4} F_{\mu\nu}^\alpha F_{\mu\nu}^\alpha + \bar{q} (i\gamma^\mu D_\mu - m) q. \quad (2.20)$$

Here, we define

$$D_\mu \equiv \partial_\mu + ig_s A_a^\mu T^a \quad (2.21)$$

and

$$F_{\mu\nu}^\alpha \equiv \partial_\mu A_\nu^\alpha - \partial_\nu A_\mu^\alpha - g_s f^{abc} A_\mu^b A_\nu^c \quad (2.22)$$

where g_s is the coupling constant of the strong force and A^μ represents the gluon field. The first term in the Lagrangian represents the gluon propagator, but note that it also includes a gluon self-interaction. This is in contrast to the QED Lagrangian, which allows no photon self-interaction. Due to this, QCD in fact allows for a bound state made entirely of gluons known as a “glueball.” The second term represents the quark-gluon interactions as well as the free quark propagator.

No stable particle may have non-neutral overall color charge. Since quarks and gluons on their own have net color charge, they immediately pull quarks and gluons out of the vacuum in order to form color-neutral bound states. These bound states may either achieve color neutrality in the form of a quark-antiquark pair (e.g. red + antired = color-neutral), or as a three-quark baryon (red + green + blue = color-neutral). This process of pulling strongly-interacting particles out of the vacuum is called *hadronization*. Hadronization occurs because the strong force is extremely short-ranged (on the order of a proton radius). As two color-charged particles are pulled apart, the energy required to keep them in a bound state satisfying color-neutrality increases rapidly. Very quickly, it becomes energetically favorable to pull a quark-antiquark pair out of the vacuum to form new color-neutral states. This requirement of color-neutrality is called *color confinement*, and the property of the strong coupling strength decreasing rapidly at short distance/high energy is known as *asymptotic freedom*[57].

2.5 Spontaneous Symmetry Breaking and the Higgs Mechanism

Recall from 2.3 that, although the gauge bosons of the electroweak interaction, the $W_{1,2,3}$ and the B , are required by gauge invariance to be massless, some of their SM counterparts, the W^\pm and Z , are observed to be massive (80.4 GeV and 91.2 GeV, respectively). In order for these bosons to retain their mass without violating the gauge symmetry, the Higgs mechanism was introduced to explain this spontaneous symmetry breaking.

Developed in 1964 by Peter Higgs, Robert Brout, and François Englert, the Higgs mechanism introduces two new complex scalar fields ϕ^+ and ϕ^0 into the EW Lagrangian. These fields are invariant under $SU(3)$ and can be written in terms of an $SU(2)_L$ doublet

$$\Phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}_L \quad (2.23)$$

The Higgs potential $V(\Phi)$ is defined as

$$V(\Phi) \equiv \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad (2.24)$$

where $\lambda > 0$. The total addition to the EW Lagrangian is of the form

$$\mathcal{L}_{Add} = D_\mu \Phi D^\mu \Phi - V(\Phi) \quad (2.25)$$

Minimizing the potential $V(\Phi)$ in Φ gives the vacuum expectation value (VEV) of Φ . If we choose $\mu^2 < 0$, we arrive at multiple critical points in our minimization and a degenerate state is possible. The VEV can be chosen to take the form[78]

$$\Phi^\dagger \Phi = \frac{v^2}{2} = -\frac{\mu^2}{2\lambda} \quad (2.26)$$

This VEV has infinite degenerate minima all with the same magnitude but different phase. Choosing one phase leads to spontaneous symmetry breaking, and the W and Z bosons acquire mass via[57]

$$m_W = \frac{g}{2} \sqrt{\frac{-\mu^2}{\lambda}} \quad (2.27)$$

$$m_Z = \frac{m_W}{\cos \theta_W}. \quad (2.28)$$

The final piece of the SM, the Higgs boson, has a mass which is dependent entirely on the VEV:

$$m_H = \sqrt{-2\mu^2}. \quad (2.29)$$

On July 4, 2012, the discovery of the Higgs boson was announced concurrently by the ATLAS and CMS collaborations, and was measured to have a mass of 125.4 GeV [59] [54].

2.6 Beyond the Standard Model

As mentioned in the introduction, some of the most pressing questions in particle physics are the following: (1) What is the origin of the matter-antimatter asymmetry?; (2) What is the

origin of neutrino mass?; (3) Are there new fundamental forces in nature?; (4) What is the origin of dark energy; and (5) Is the Higgs boson solely responsible for electroweak symmetry breaking and the origin of mass? Much like the Higgs mechanism is introduced to account for the $SU(2)\times U(1)$ symmetry breaking, there are a plethora of theoretical models which incorporate additional gauge fields and interactions to address these questions.

For example, string theory is considered a promising candidate for describing gravitational systems at strong coupling and thus plays a prominent role in the description of black holes and evolution of the universe through the understanding of the origin of dark energy. Similarly, models with additional neutrino fields at the TeV scale provide a possible explanation for the mass of light neutrinos. Such models often manifest themselves as new heavy particles that could be observed at the LHC. Surprisingly, some of these new particles predicted on the basis of pure particle physics arguments can provide the correct dark matter relic density.

There are several ways new heavy gauge bosons can appear. The most natural possibility is one in which these heavy gauge bosons are the gauge field of a new local broken symmetry. Examples include models with a new $U(1)$ gauge symmetry[14], little Higgs models[64], and E_6 Grand Unified Theories (GUT)[67]. In models with a new $U(1)$ gauge symmetry, the Z' is the gauge boson of the broken symmetry. In little Higgs models, breaking of the global symmetry by gauge and Yukawa interactions generates Higgs mass and couplings at the TeV scale that cancel off the SM quadratic divergence of the Higgs mass from top, gauge, and Higgs loops. This results in one or more Z' bosons. In Kaluza-Klein models, the Z' bosons arise as a result of hidden extra dimensions. It should be noted that little Higgs and Kaluza-Klein models were not considered in the searches presented in the upcoming chapters, and therefore will not be discussed further in this thesis. Even so, from the breadth, scope, and implications of these models, it is apparent that probing these questions and puzzles gives ample hope of uncovering new physics at the TeV scale. As such, it is highly worthwhile to engage in searches for Z' candidates.

Of particular interest in such searches are models that include an extra Z' -like neutral gauge boson that decays to pairs of high- $p_T \tau$ leptons. Although many models with extra gauge bosons

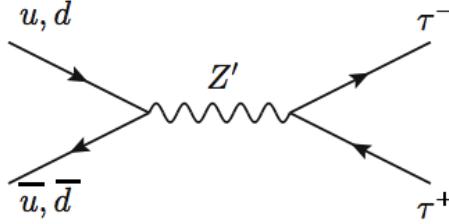


Figure 2.3: $Z' \rightarrow \tau\tau$ event at the LHC, wherein two quarks from colliding protons form a Z' which then decays to an oppositely-charged, back-to-back pair of τ leptons.

obey the universality of the couplings (meaning the bosons decay to each generation of fermions with equal frequency), some models include generational-dependent couplings resulting in extra neutral gauge bosons that preferentially decay to τ leptons, making this analysis an important mode for discovery. However, even if a new gauge boson decaying to $\mu\mu$ is discovered first, it will be critical to establish the $\tau\tau$ decay channel to establish the coupling relative to $\mu\mu$ channel. The primary model studied in this thesis is the Sequential Standard Model Z' , denoted Z'_{SSM} . This model assumes universality of the couplings (just like the SM Z) which makes it a useful benchmark, both for testing other models and for tuning the search methodology. The current CMS published limit on a Z'_{SSM} decaying to pairs of electrons or muons, established with the data taken during the 2012 run at $\sqrt{s} = 8$ TeV, excludes a Z'_{SSM} below a mass of 2.90 TeV[61].

In pp collisions at the LHC, the Z'_{SSM} is expected to be generated in much the same way as the SM Z , through Drell-Yan production via quark-antiquark interactions from the colliding protons. Even though protons (not antiprotons) are the particles being collided, there are significant quantities of antiquarks present in the collisions. This is because, at high energies, the protons radiate large numbers of gluons which promptly decay to quark-antiquark pairs. A diagram of the Z'_{SSM} creation process is shown in Figure 2.3.

2.6.1 Models Predicting New Neutral Bosons

2.6.1.1 Sequential Standard Model (SSM)

The SSM does not predict the existence of a Z' due to a larger symmetry group, but rather manually adds an extra neutral gauge boson, Z'_{SSM} , which is identical to the SM Z , with the same couplings to quarks and leptons. This model is not gauge invariant unless the Z'_{SSM} couples to additional, exotic fermions or unless it exists as an excitation of the SM Z in the case of models involving extra dimensions at the weak scale. Its decay width is given by

$$\Gamma_{Z'} = \Gamma_Z \times M_{Z'}/M_Z. \quad (2.30)$$

While the SSM is typically not gauge-invariant, it still serves as an exceptionally useful benchmark in the search for new physics, as it can be used as a baseline for comparison with other models should an excess be observed[14]. Given this, the Z'_{SSM} is the chief signal used in the physics searches described in this thesis.

2.6.1.2 Grand Unified Theory (GUT)-inspired Models

Models that try to merge the strong and electroweak interactions into one larger, all-encompassing symmetry group are called Grand Unified Theories (GUTs). One popular GUT uses the E_6 symmetry group. In the E_6 model, the E_6 group breaks in the following fashion:

$$E_6 \rightarrow SO(10) \times U(1)_\psi \rightarrow SU(5) \times U(1)_\chi \times U(1)_\psi \rightarrow SM \times U(1)_{\theta_{E_6}}. \quad (2.31)$$

The new gauge boson (Z' candidate) is predicted to arise as a mixing of the $U(1)_\psi$ and $U(1)_\chi$ groups with a mixing angle θ_{E_6} :

$$Z' = Z'_\chi \cos \theta_{E_6} + Z'_\psi \sin \theta_{E_6}, \quad (2.32)$$

where θ_{E_6} is a free parameter ranging from -90° to 90° . The choice of θ_{E_6} is dictated by the model, and the four most popular values are $\theta_{E_6} = 0^\circ$ (Z'_χ), $\theta_{E_6} = 90^\circ$ (Z'_ψ), $\theta_{E_6} = \sin^{-1} \sqrt{3/8}$ (Z'_η), and $\theta_{E_6} = \sin^{-1} \sqrt{5/8}$ (Z'_I)[67]. While these E_6 -inspired models were not included in the 13

TeV search, the Z'_ψ signal was one of the principal signal Monte Carlo samples used in the 8 TeV analysis.

Chapter 3

The Large Hadron Collider (LHC)

The Large Hadron Collider (LHC) is the world’s largest and highest-energy particle accelerator. Straddling the France-Switzerland border at CERN, near Geneva, Switzerland, the LHC is installed in a 26.7 km-long circular tunnel between 45 and 170 m underground. The LHC is a proton-proton (pp) collider, designed to accelerate protons to near-light speed (about 3 m/s slower than c). The protons are accelerated in two counter-rotating beams, each housed inside vacuum tubes and bent into a ring by 8.3 T superconducting dipole magnets. At four detector locations, the beams are crossed so that collisions may occur. An overview of the main experiments on the LHC ring can be seen in Figure 3.1.

The protons are divided into “bunches” by radio-frequency (RF) cavities, with about 10^{11} protons per bunch. Each beam contains up to 2808 bunches. The design center of mass collision energy is $\sqrt{s} = 14$ TeV, which corresponds to 7 TeV per colliding proton beam. The LHC is designed for bunch spacings of 25 ns, corresponding to a design collision rate of 40 MHz. The design instantaneous luminosity (number of protons in a collision area per unit time) of the LHC is $\mathcal{L} = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. Integrated over time, the instantaneous luminosity gives the *integrated luminosity*:

$$L = \int \mathcal{L}(t) dt. \quad (3.1)$$

Over the 2012 run, during which the LHC was running at a pp collision energy of $\sqrt{s} = 8$ TeV and a bunch spacing of 50 ns, the LHC delivered 23.3 fb^{-1} of integrated luminosity, of which

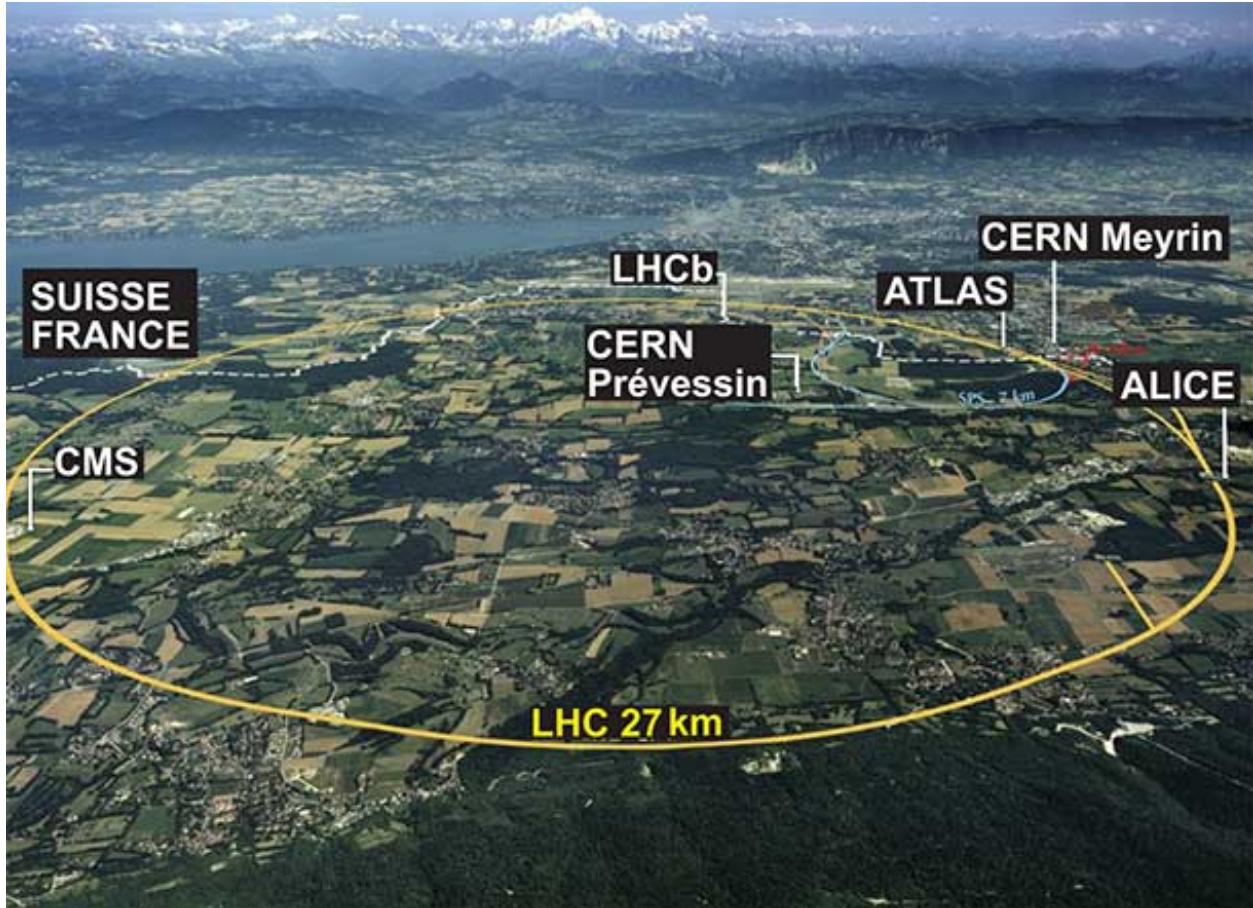


Figure 3.1: The Large Hadron Collider located underneath the Franco-Swiss border at CERN near Geneva, Switzerland. The CMS Experiment is located on the French side in Cessy, France

CMS recorded 21.8 fb^{-1} . During the 2015 run, during which the LHC was running at a pp collision energy of $\sqrt{s} = 13 \text{ TeV}$ and a bunch spacing of 25 ns, the LHC delivered 4.22 fb^{-1} of integrated luminosity, of which CMS recorded 3.81 fb^{-1} . As of this writing, the LHC is wrapping up its 2016 run, also with a pp collision energy of $\sqrt{s} = 13 \text{ TeV}$ and a bunch spacing of 25 ns, and so far 34.1 fb^{-1} of instantaneous luminosity have been delivered and 31.4 fb^{-1} have been recorded by CMS. A plot of the luminosity delivered so far during 2016 is shown in Figure 3.2.

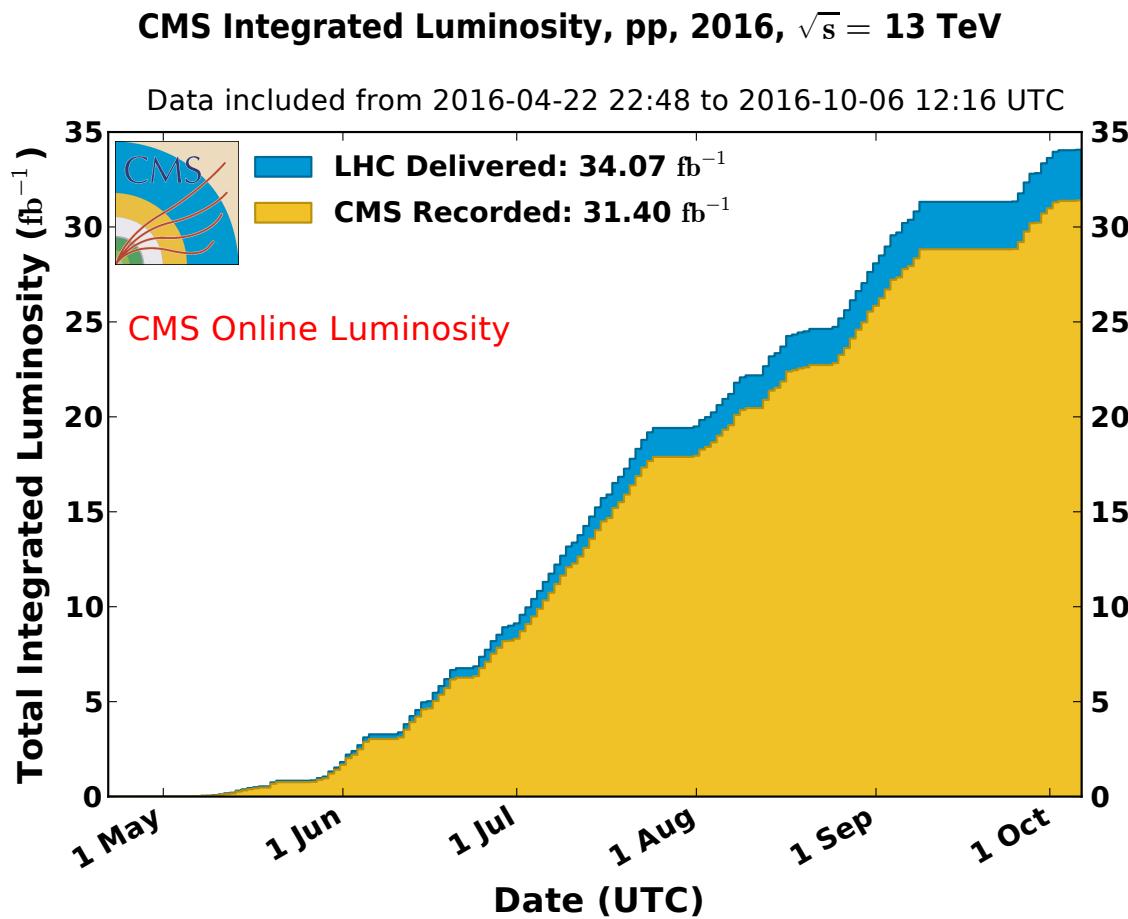


Figure 3.2: Total integrated luminosity delivered by the LHC (blue), and total integrated luminosity recorded by CMS (yellow) during the 2016 run. The difference is a measure of the experiment's efficiency.

Chapter 4

The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) Experiment is a multipurpose particle detector located in the LHC ring. The CMS experiment is located 100 meters underground in Cessy, France and is 28.7 meters long, 15.0 meters in diameter, and weighs approximately 14,000 tonnes. It is arranged in a cylindrical, multi-layered structure consisting of a “barrel” and two endcaps, with the LHC beam passing through the central axis of the cylinder. The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. A photograph of CMS can be seen in Figure 4.1, and a diagram highlighting the layout of CMS can be seen in Figure 4.2. The experiment uses a right-handed coordinate system: the origin is set at the nominal pp collision point, with the x -axis pointing towards the center of the LHC ring, the y -axis pointing up, and the z -axis pointing along the beam line in the counter-clockwise direction. CMS also uses a pseudo-polar coordinate system, with ϕ defined as the azimuthal angle measured from the x -axis in the xy plane, θ defined as the polar angle from the $+z$ axis, and η as the “pseudorapidity”, defined as

$$\eta = -\ln [\tan (\theta/2)] \quad (4.1)$$

Another important piece of terminology is the *transverse plane*, which is simply the xy plane

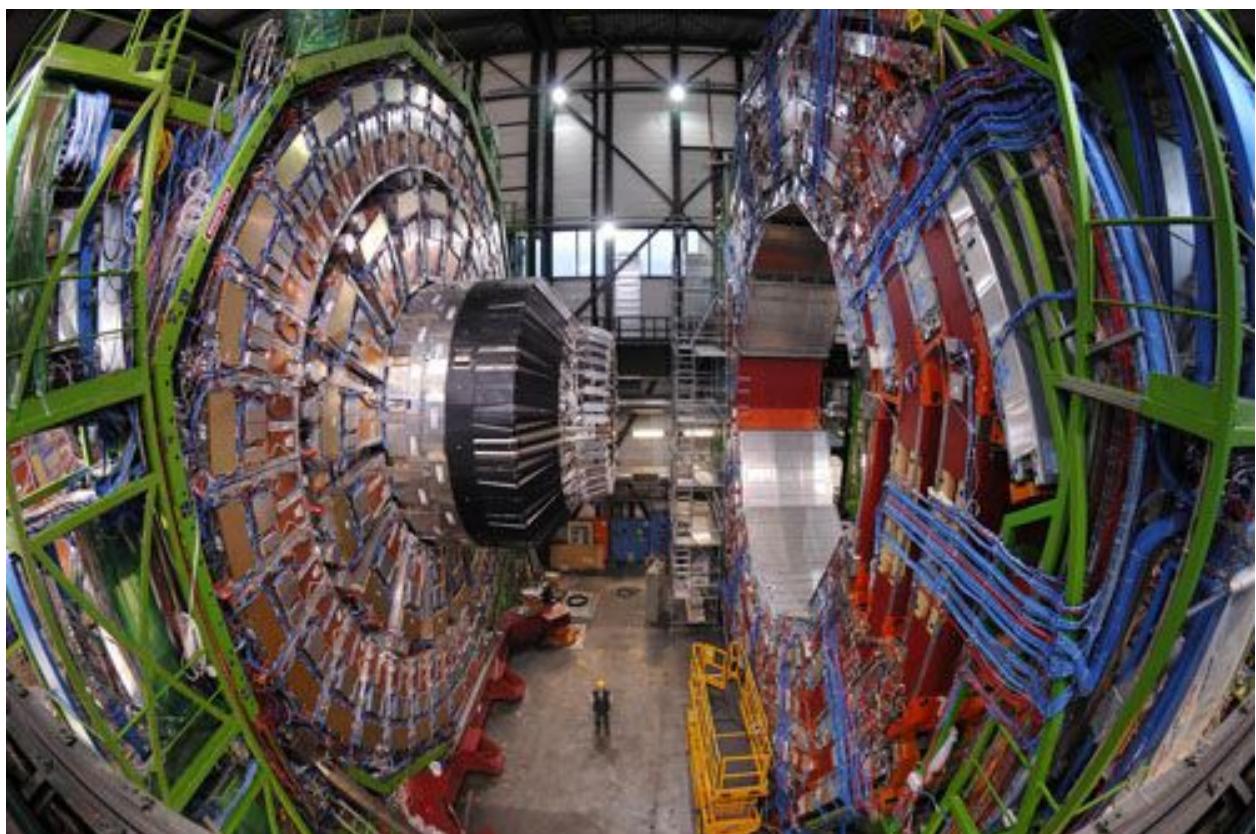


Figure 4.1: The CMS Experiment (open for maintenance)

perpendicular to the beam axis. Observable quantities are often given in terms of their transverse component, denoted by a “ T ” subscript. For example, particle momentum is often referred to in terms of its transverse component, p_T . The quantity $R = \sqrt{\eta^2 + \phi^2}$ gives longitudinal information.

4.1 The Silicon Tracker

The silicon tracker is the innermost detector element in CMS, and is designed to provide the highest resolution measurement of charged particle trajectories (such trajectories are referred to as “tracks”). The tracker is composed of approximately 200 m^2 of silicon, and includes arrays of silicon pixels in the inner layers and arrays of silicon strips in the outer layers. The total number of channels is 66 million pixels and 9.6 million strips[33]. The tracker is divided into a barrel segment and two forward endcaps. The pixel barrel consists of three concentric cylindrical layers at radii between 4.4 cm and 10.2 cm, and the strip barrel consists of a further ten cylindrical layers extending out to a radius of 110 cm. The pixel endcaps contain two small disks each, while the strip endcaps each consist of three small disks and nine large disks. A schematic of the tracker layout can be seen in Figure 4.3. Particle trajectories are reconstructed by fitting hits in the individual silicon pixels/strips to an interpolated track.

4.1.1 The Pixel Detector

The pixel detector, shown in Figure 4.4, consists of three barrel layers and two endcap disks on each side. The barrel has a length of 53 cm, and the endcap disks range from 6 cm to 15 cm in radius. The pixel modules are arranged in a ladder-like configuration in the barrel, with 768 total modules comprising the barrel. The endcap disks are arranged in a turbine-like fashion, with 24 “blades” per disk, and 7 pixel modules per blade for a total of 672 pixel modules in the endcaps. The size of an individual pixel is $100 \times 150\text{ }\mu\text{m}^2$. Each pixel module consists of several readout chips (ROCs), which are bump bonded to the module. In total, the pixel detector includes about 16,000 ROCs[47].

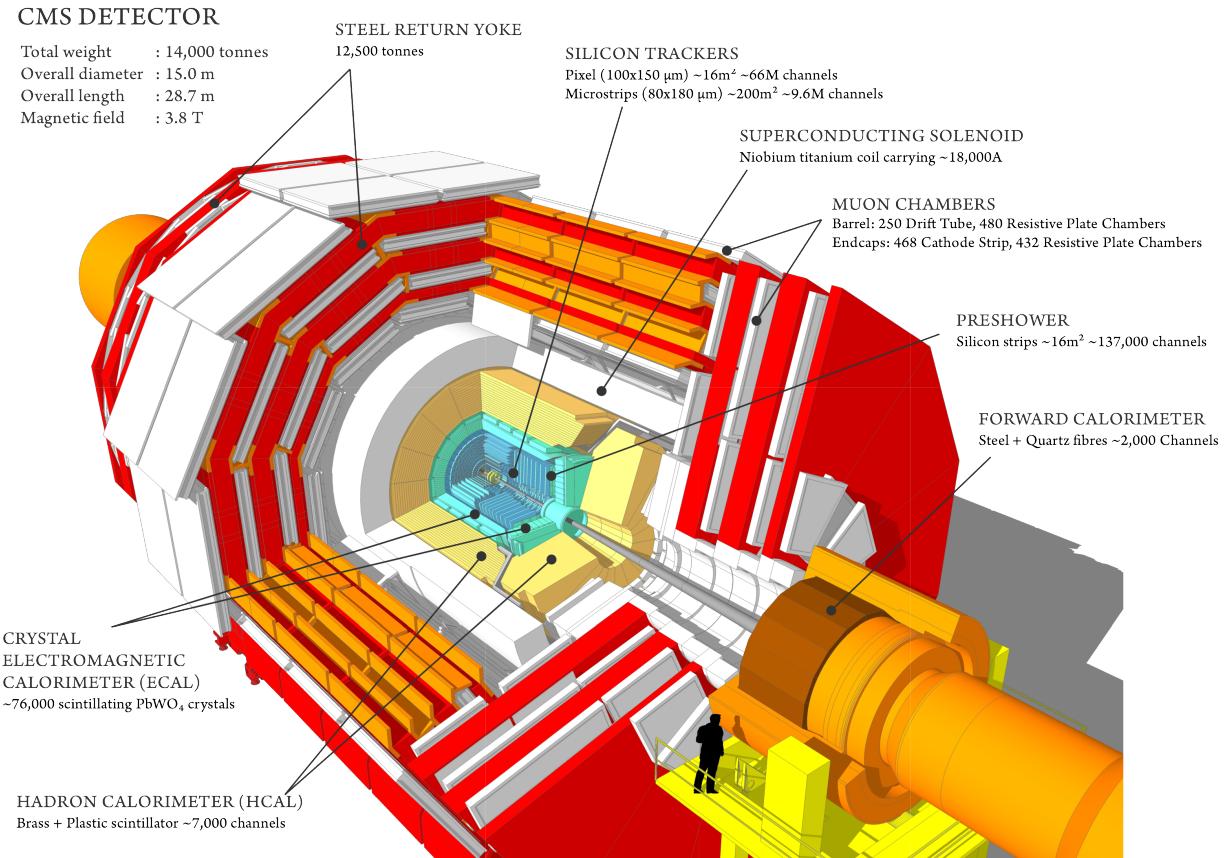


Figure 4.2: Schematic of the CMS Experiment showing the silicon trackers, electromagnetic calorimeter (ECAL), hadronic calorimeter (HCAL), forward calorimeters, superconducting solenoid, steel magnetic flux return yoke, and muon system.

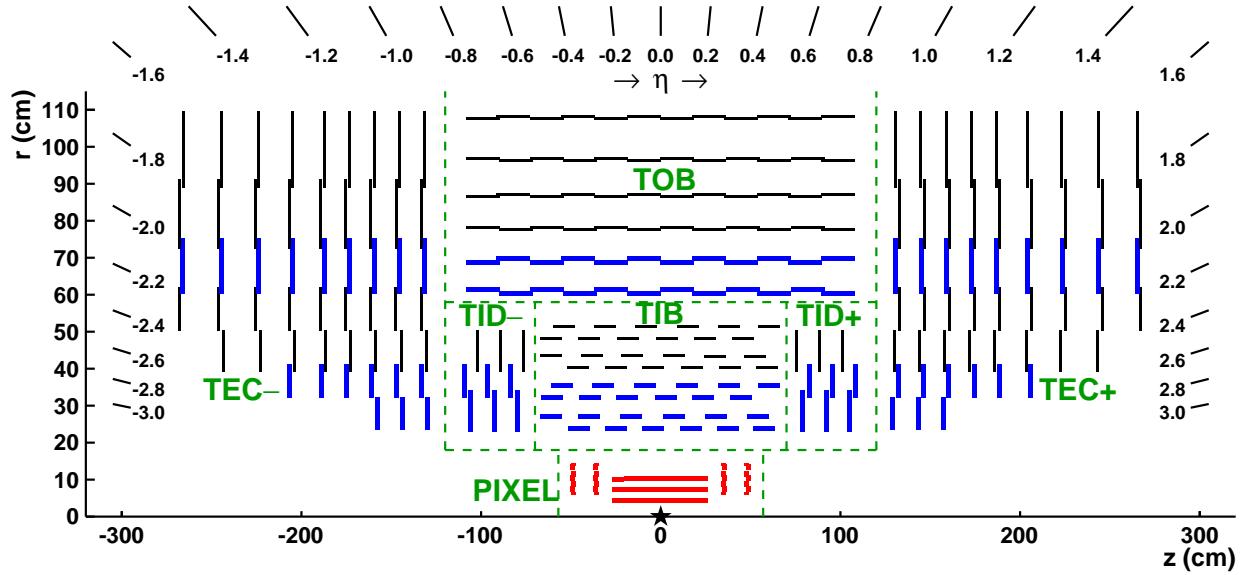


Figure 4.3: Overview of the silicon tracker layout (projected onto the Rz plane, including the pixel layers and disks (red), the TIB and TOB layers, and the TID and TEC disks.

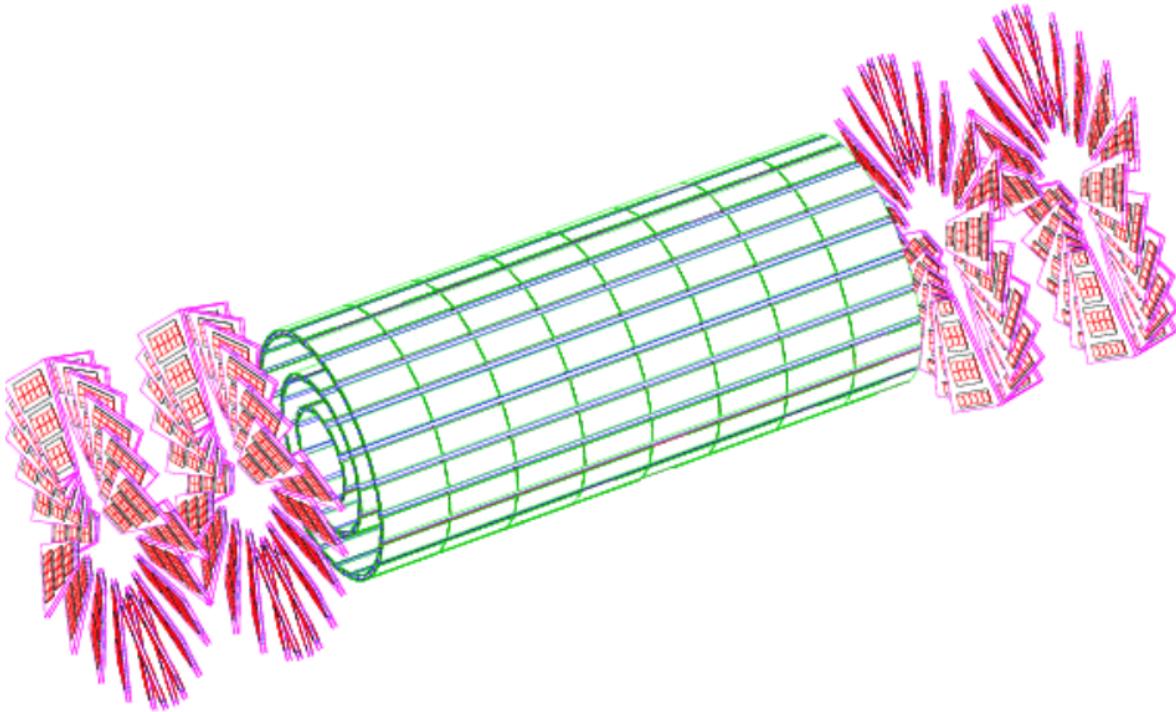


Figure 4.4: Diagram indicating the layout of the barrel pixel (BPIX) sensors (in green) and endcap pixel (FPIX) sensors (in pink)

4.1.1.1 Pixel Detector Performance

Common measurements of the detector performance are hit efficiency and hit resolution. The hit efficiency is defined as the probability of finding any hit clusters within a $500 \mu\text{m}^2$ area around an expected hit, where an expected hit is provided by a “good quality track” with an associated primary vertex (PV), small impact parameter with respect to that vertex, and $p_T > 1.0 \text{ GeV}$. “Primary vertex” refers to the reconstructed origin of the tracks corresponding to the pp collision, and will be discussed in further detail in Section 4.6.1.2. “Impact parameter” refers to the distance between the PV and the point of closest approach between the PV and the track. The hit resolution is measured by comparing the hit position on a given layer with an interpolated track and taking the residual difference. Typical results for the hit efficiency and hit resolution are shown in Figure 4.5 and Figure 4.6, respectively[79].

4.1.2 The Strip Tracker

The strip tracker is divided into barrel and endcap detectors. The barrel segment is composed of a Tracker Inner Barrel (TIB) section and a Tracker Outer Barrel (TOB) section. The TIB includes four layers of silicon strips each $320 \mu\text{m}$ thick and ranging in pitch from $80 \mu\text{m}$ and $120 \mu\text{m}$. In the TOB, the lower rate of particle flux allows for larger strips (each $500 \mu\text{m}$ thick and ranging in pitch from $120 \mu\text{m}$ to $180 \mu\text{m}$).

The endcaps are each comprised of a Tracker Endcap (TEC) and Tracker Inner Disks (TIDs). The TIDs are designed to fill the region between the TEC and the TIB. The TECs each contain nine disks, and each TID contains three disks. On each disk, the modules (for both TID and TEC) are arranged in rings centered on the beam line. Each TID disk contains three rings of modules, while each TEC disk has up to seven rings. The TID strips (and three innermost ring strips of the TEC) have thickness $320 \mu\text{m}$, while the rest of the TEC strips have thickness $500 \mu\text{m}$. In total the strip tracker contains about 15,400 strip modules. The modules in the innermost two layers of both the TIB and the TOB, as well as the modules in rings 1 and 2 of the TID, and 1, 2 and 5 of the

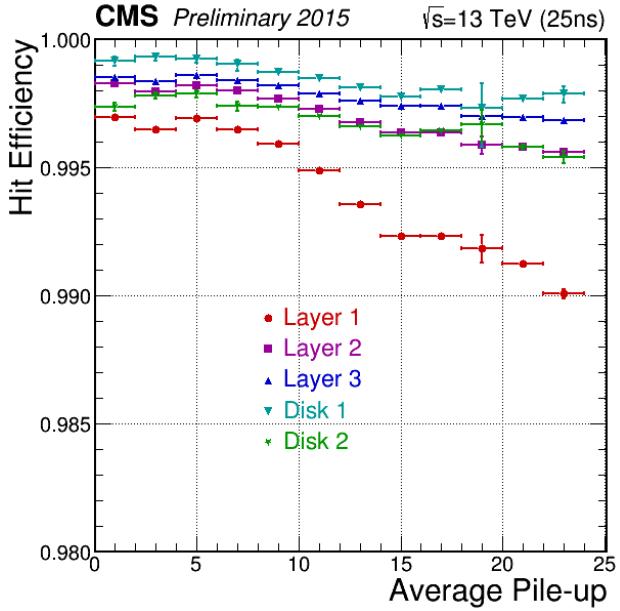


Figure 4.5: Hit efficiency vs. the average number of pp collisions per bunch crossing for pixel barrel layers and forward disks.

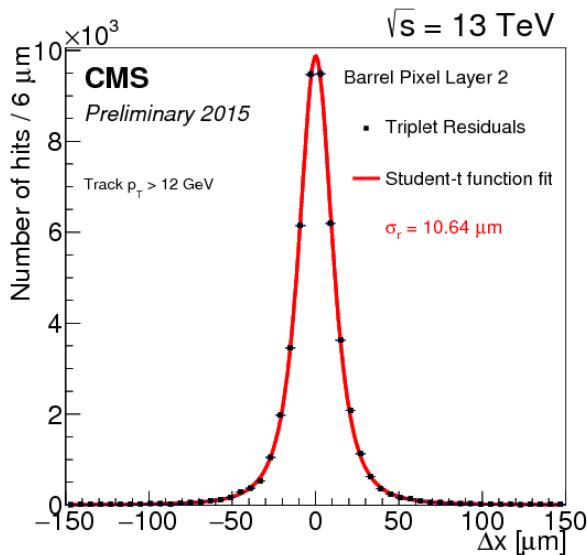


Figure 4.6: Distribution of hit residuals on the second layer of the pixel barrel in the transverse direction to the beam. The distribution is fitted with a student's t-function for which sigma is shown on the plot.

TEC, carry a second strip detector module, which is mounted back-to-back to the first and rotated in the plane of the module by a “stereo” angle of 100 mrad. The hits from these two modules, known as “ $r\phi$ ” and “stereo hits”, can be combined into matched hits that provide a measurement of the second coordinate (z in the barrel and r on the disks)[49][33].

4.2 The Electromagnetic Calorimeter

The next layer outward from the silicon tracker is the electromagnetic calorimeter (ECAL), which is designed primarily to measure energies of photons and electrons. The ECAL consists of an array of 75,848 lead tungstate (PbWO_4) crystals, and has a barrel (EB) segment as well as two endcap (EE) segments. The EB segment has 61,200 crystals, while each EE segment contains 7,324 crystals. Lead tungstate is highly transparent and is an effective scintillator for photons and electrons, producing light in short, fast, well-defined showers (cascades of photons and electrons arising from high energy electrons and photons interacting with the medium) that are picked up by photodetectors glued to each crystal. These photodetectors then convert the scintillation light into an electrical signal that can be read out for analysis.

Each barrel crystal presents an apparent cross-section (when viewed from the interaction vertex) of $22 \times 22 \text{ mm}^2$, and are 230 mm (25.8 radiation lengths) long. The barrel has an inner radius of 129 cm, and covers the pseudorapidity range $|\eta| < 1.479$ [33].

The endcap crystals are each arranged in two D-shaped semicircular aluminum plates. From these plates are cantilevered “supercrystal” structures consisting of 5x5 crystal blocks. Each crystal presents an apparent cross-section of $28.6 \times 28.6 \text{ mm}^2$, and are 220 mm (24.7 radiation lengths) long. The endcap crystals cover the pseudorapidity range $1.479 < |\eta| < 3.0$ [55].

Sitting in front of both EE segments are the *preshower detectors*, which offer a much finer granularity than ECAL with detector strips that are just 2 mm wide. Rather than lead tungstate crystals, the preshower uses a combination of lead planes and silicon strip sensors (similar to those in the tracker). Electromagnetically-interacting particles shower in the lead layers, and the showers are subsequently read out by the silicon sensors. Each preshower uses two layers of lead and silicon,

giving further resolution. The complete preshower each form a disk about 2.5 m in diameter, with a hole 50 cm in diameter for the beam pipe.

A diagram of the ECAL layout can be seen in Figure 4.7.

4.2.0.1 ECAL Performance

In the barrel section of the ECAL, an energy resolution of about 1% is achieved for unconverted or late-converting photons in the tens of GeV energy range. All other barrel photons have a resolution of about 1.3% up to a pseudorapidity of $|\eta| = 1$, rising to about 2.5% at $|\eta| = 1.4$. In the endcaps, the resolution of unconverted or late-converting photons is about 2.5%, while the remaining endcap photons have a resolution between 3 and 4% [63]. The energy resolution as a function of pseudorapidity is shown in Figure 4.8[55].

4.3 The Hadronic Calorimeter

Surrounding the ECAL is the hadronic calorimeter (HCAL). The primary function of the HCAL is to measure the energies of hadrons (particles composed of quarks). In most cases, electromagnetically-interacting particles are absorbed by ECAL and don't make it to the HCAL. Located just inside the solenoid magnet, the HCAL is primarily composed of brass panels made from melted-down artillery shells. Interspersed with the brass panels are plastic scintillation panels, in which are embedded wavelength-shifting (WLS) fibers, which carry the signal to clear fibers outside the scintillators for readout. As hadrons enter the HCAL, they produce secondary particles in the brass which in turn create further particles. These hadron "showers" then interact with the plastic scintillators, where the fibers carry the signal to hybrid photodiodes so the signal can be measured.

The HCAL is divided into barrel (HB) and endcap (HE) portions. Due to limited space between the ECAL and solenoid, the HCAL also includes material outside the solenoid: the outer HCAL (HO) lining the solenoid, and the forward calorimeter (HF) inside the muon endcap system. The forward hadron (HF) calorimeter uses steel as an absorber and quartz fibers as the sensitive

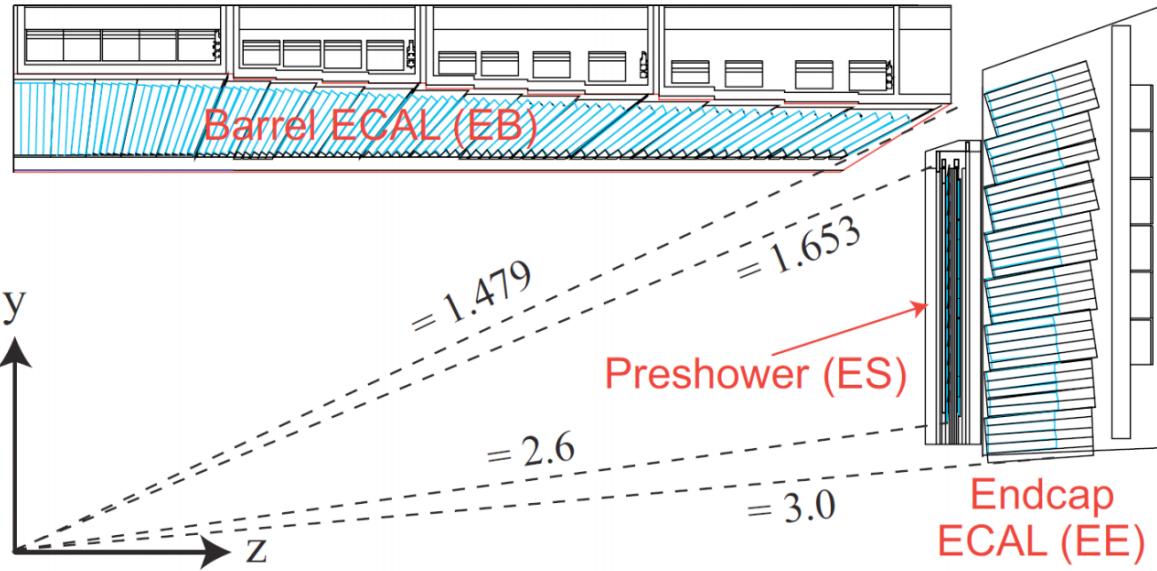


Figure 4.7: Layout of the electromagnetic calorimeter (ECAL) indicating the configuration of the barrel (EB) and endcap (EE) crystals. Numbers on the layout are in terms of η .

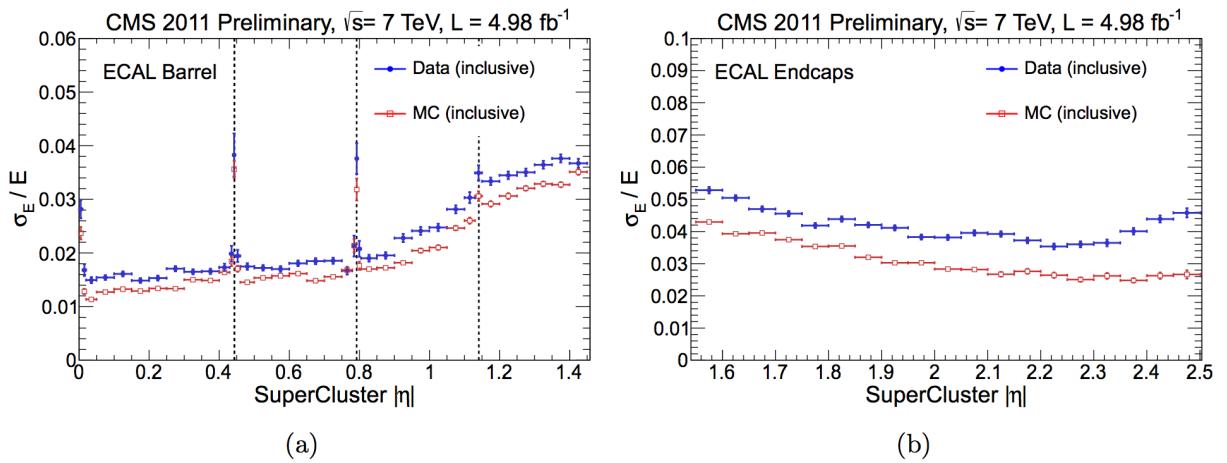


Figure 4.8: Relative energy resolution vs $|\eta|$ in EB (a) and EE (b) for both Monte Carlo (MC) and data. The vertical dashed lines indicate module boundaries in EB. The mismatch between data and MC is corrected by applying η -dependent correcting factors to subsequent models.

material. The two halves of the HF are located 11.2 m from the interaction region, one on each end, and together they provide coverage in the range $3.0 < |\eta| < 5.2$. They also serve as luminosity monitors[33].

In the region $|\eta| < 1.74$, the HCAL cells have widths of 0.087 in pseudorapidity and 0.087 in azimuth (ϕ). In the $\eta\text{-}\phi$ plane, and for $|\eta| < 1.48$, the HCAL cells map on to 5×5 arrays of ECAL crystals to form calorimeter towers projecting radially outwards from close to the nominal interaction point. For $|\eta| > 1.74$, the coverage of the towers increases progressively to a maximum of 0.174 in $\Delta\eta$ and $\Delta\phi$. Within each tower, the energy deposits in ECAL and HCAL cells are summed to define the calorimeter tower energies, subsequently used to provide the energies and directions of hadronic jets.

A diagram of the HCAL layout is shown in Figure 4.9.

4.4 The CMS Magnet

The need to measure high p_T muons has driven the requirements on the CMS magnetic field strength. With the goal to determine the momentum of $\approx 1\text{TeV}$ muons with a relative uncertainty of $\approx 10\%$, a superconducting solenoid with an interior magnetic field strength of 3.8T was chosen as the central design feature of CMS.

The solenoid is 13 m in length and 6 m in diameter, and is located outside the silicon tracker, ECAL, and HCAL systems, but inside the muon system. To generate the magnetic field, 18,160 amperes are passed through four layers of tightly-wound superconducting Nb-Ti wire, resulting in a stored energy of 2.3 gigajoules. The magnetic field generated by the solenoid bends charged particles according to the equation

$$R = \frac{p_T}{0.3eB} \quad (4.2)$$

where R is the radius of curvature (in meters), p_T is the transverse momentum (in GeV), e is the electron charge (in Coulombs), and B is the field strength (in Tesla). Thus, the transverse momentum (p_T) of the charged particle can be measured from the known field strength and the

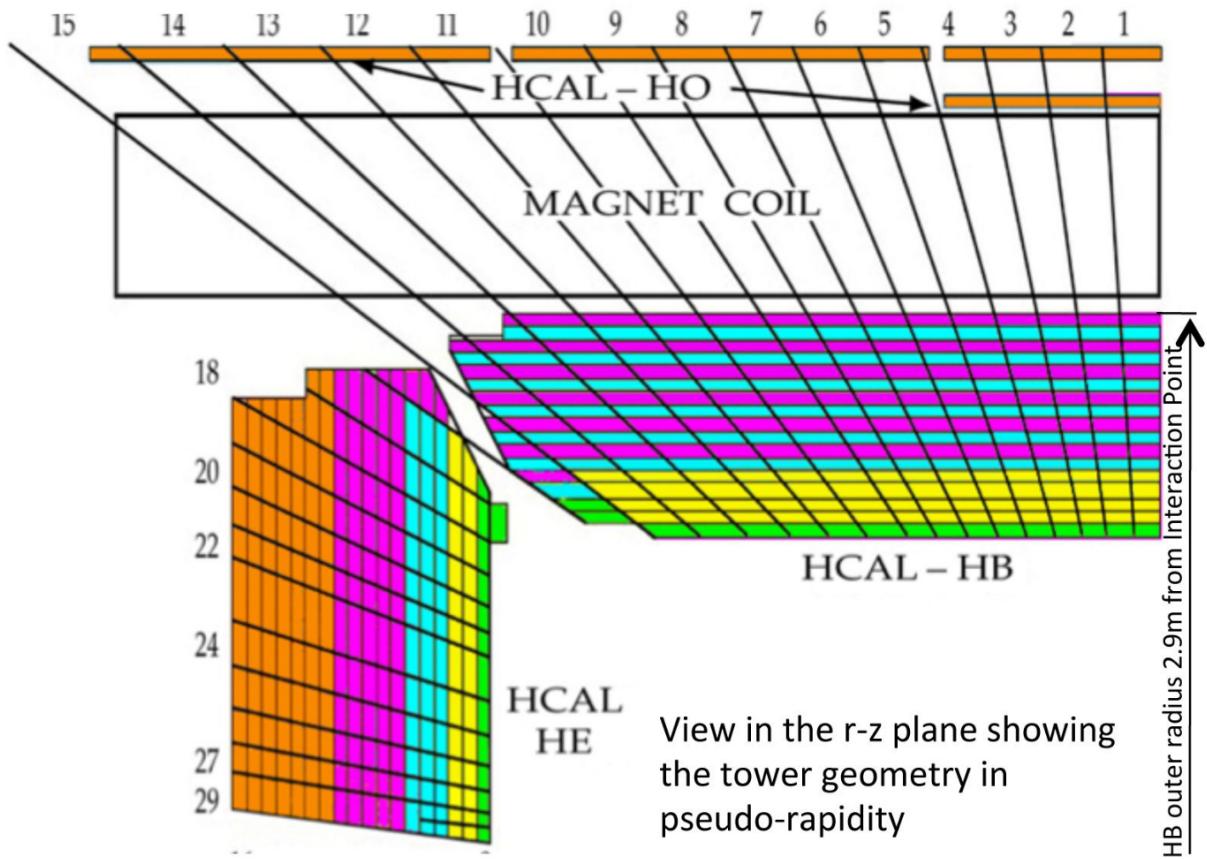


Figure 4.9: Layout of the hadronic calorimeter (HCAL) indicating the configuration of the barrel (HB) and endcap (HE) components as well as the outer HCAL (HO) lining the solenoid.

observed radius of curvature of the tracks measured in the tracker and muon systems[33].

4.5 The Muon System

The muon system is the outermost subdetector. Muons are measured in the pseudorapidity range $|\eta| < 2.4$, with detection planes made using three technologies: drift tubes, cathode strip chambers, and resistive plate chambers. Interspersed with the muon system is the steel flux return yoke, which contains the field outside the solenoid and provides a 2 T field to allow the muon system to measure charged particle momentum. Matching muons to tracks measured in the silicon tracker results in a relative transverse momentum resolution for muons with $20 < p_T < 100$ GeV of 1.3–2.0% in the barrel and better than 6% in the endcaps. The p_T resolution in the barrel is better than 10% for muons with p_T up to 1 TeV [32]. A diagram of one quadrant of the muon system can be seen in Figure 4.10. “MB” indicates the muon barrel subdetector, which is composed of DTs, “RB” indicates the RPC barrel subdetector, “ME” indicates the muon endcap subdetector (CSCs), and “RE” indicates the RPC endcap subdetector.

An overview of the layout of the muon system can be seen in Figure 4.10.

4.5.1 Drift Tubes (DTs)

The barrel of the muon system contains 250 layers of DTs arranged in four sections. In the inner three sections, DTs are arranged in 8 $R\phi$ -measuring planes, and 4 z -measuring planes. The other section only contains $R\phi$ -measuring planes.

Each DT is filled with a gaseous mixture of argon (85%) and carbon dioxide (15%). Charged particles passing through the DT will ionize the gaseous mixture and produce a current along a central filament (upon which a positive voltage is applied) to be read out. The maximum drift length per DT is 2.0 cm and the single point resolution is ≈ 200 μm . Since the drift velocity of electrons in the gaseous mixture is known, a second position coordinate can be measured for the muon by comparing the readout time with the time of the pp collision. A plot of the DT reconstruction efficiency is given in Figure 4.11[33]. The DT reconstruction efficiency is measured

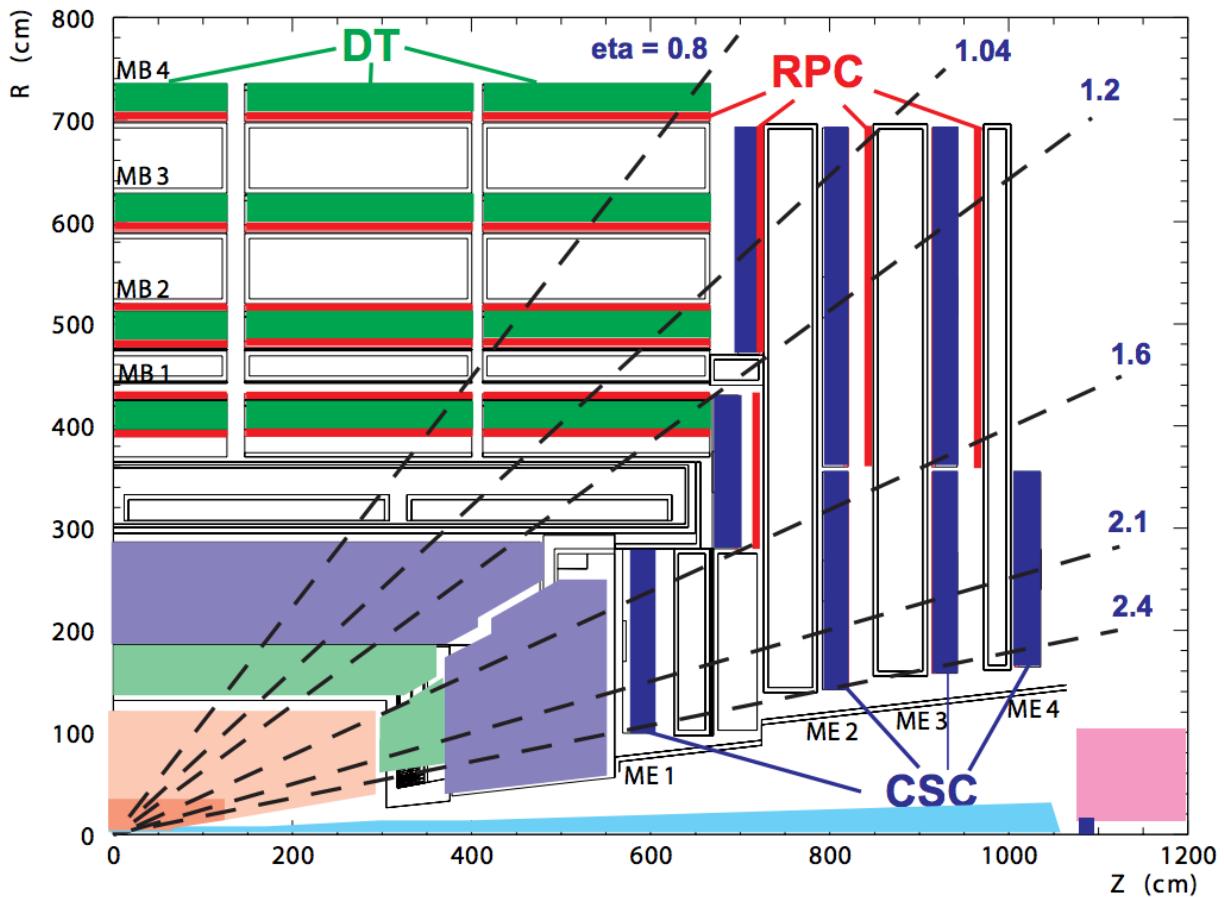


Figure 4.10: Layout of the muon system highlighting the layered structure of the drift tubes (DTs, green), cathode strip chambers (CSCs, blue), and resistive plate chambers (RPCs, red).

using the standard CMS “tag-and-probe” method, wherein dimuon pairs from Z and J/ψ decays are collected with a single-muon trigger. One muon (the tag) is required to pass standard CMS global muon identification requirements across both the tracker and muon system, while the other muon (the probe) is only required to pass the tracker requirements. The efficiency is measured as the number of “probe” muons passing the muon system identification requirements divided by the number of “tag” muons[48].

4.5.2 Cathode Strip Chambers (CSCs)

The muon endcap (ME) system is composed of 234 CSCs in each endcap. The CSCs are trapezoidal in shape and are arranged in four layers. The innermost layer consists of 108 CSCs arranged in three rings of 36 CSCs each. The other layers each have two rings (18 CSCs on the inner ring, 36 on the outer ring, for a total of 54 CSCs per layer)[48].

Each CSC is composed of six gas gaps, with each gap containing a radial array of negatively-charged cathode strips and a plane of positively-charged anode wires running perpendicular to the cathodes. The CSCs are overlapped in ϕ to avoid gaps in the particle acceptance. When a charged particle ionizes the gas, it leaves a charge on the anode wire and an image charge on a group of cathode strips. The spatial resolution provided from each CSC is $\approx 200\ \mu\text{m}$. Because the strips and wires are perpendicular, two position coordinates in (R, ϕ) are provided for each passing particle (as well as a z coordinate simply from the position of the CSC.)

The forward (endcap) region of the muon system can expect a higher muon flux, so CSCs were chosen due to their fast response time and fine resolution. Figure 4.12 shows the CSC efficiency, measured using the same “tag-and-probe” method described in the above section on DTs.

4.5.3 Resistive Plate Chambers (RPCs)

RPCs are present in both the barrel and endcaps of the muon system, and are used in parallel with the other two subsystems. The RPCs have timing resolution on the order of nanoseconds and are used primarily to resolve ambiguities in cases of multiple hits in the DTs or CSCs, thereby

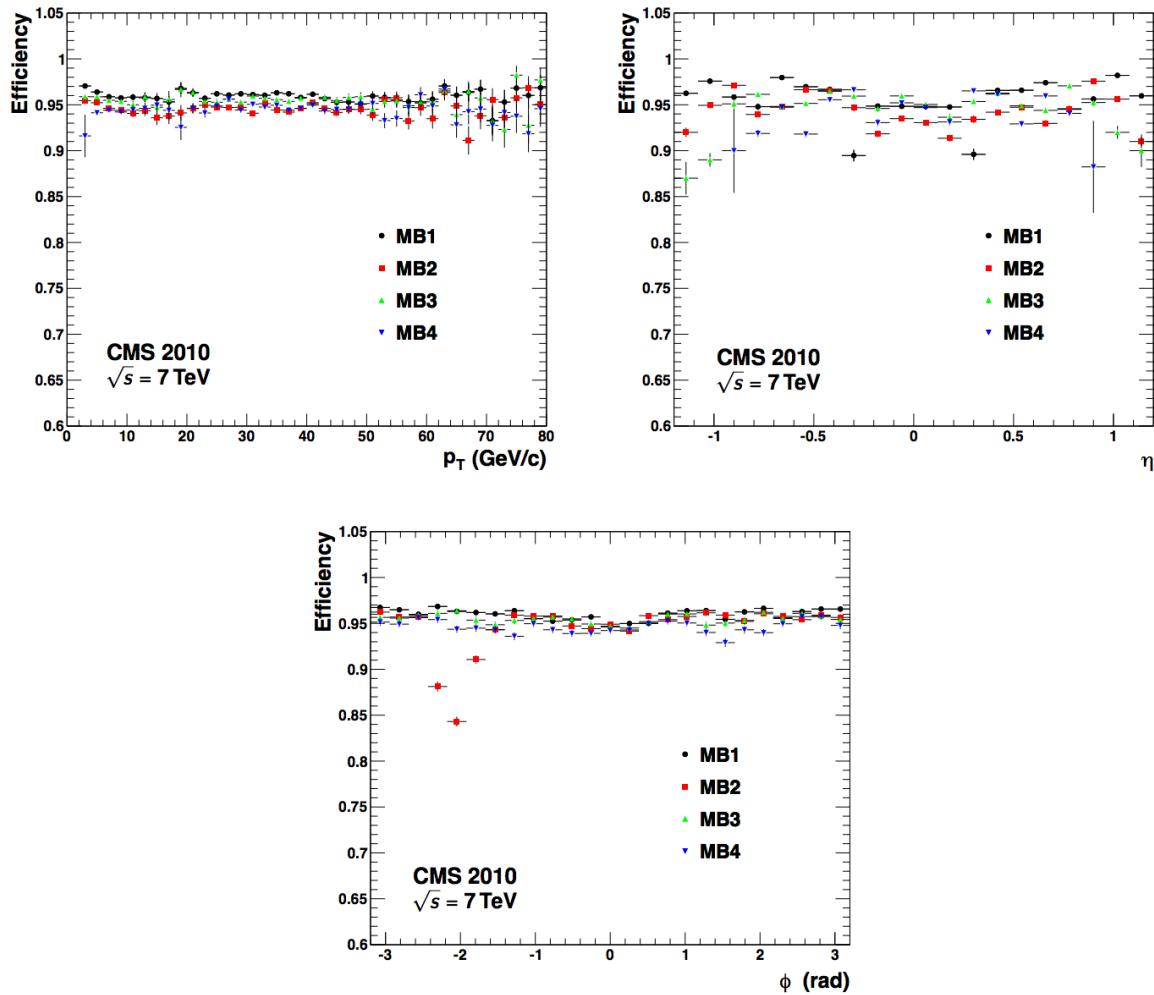


Figure 4.11: The measured DT efficiency as a function of the muon transverse momentum p_T , η , and ϕ . Results for the 4 stations are superimposed.

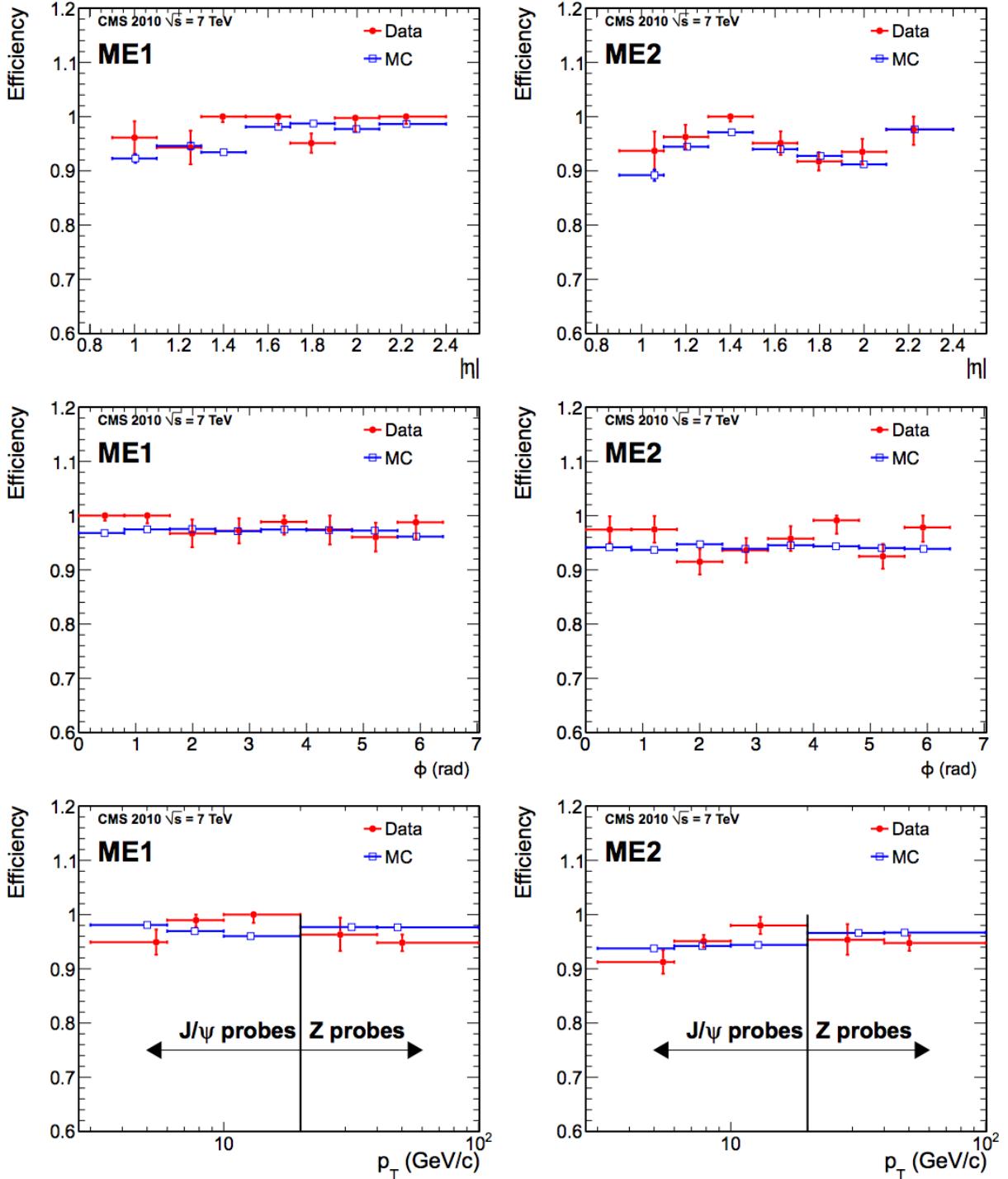


Figure 4.12: The measured CSC efficiency as a function of η , ϕ , and the muon transverse momentum p_T for ME1 (left) and ME2 (right) stations. The vertical lines on the p_T distributions divide the ranges between values covered by the J/ψ and Z probes.

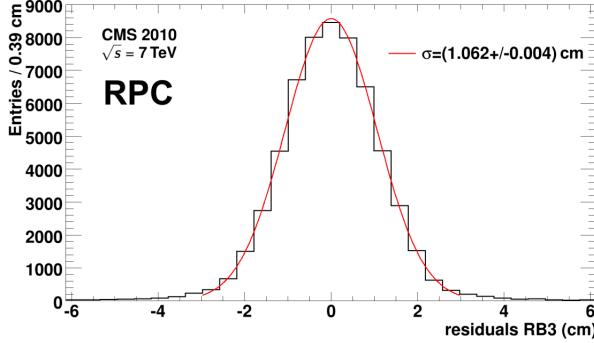


Figure 4.13: Residuals (distance between “good” track and RPC hit) distribution in RB3 (the RPC associated with MB3 in the barrel of the muon system), fitted to a Gaussian.

improving the overall accuracy of particle reconstruction.

There are two layers of RPCs “sandwiching” each of the first two DT layers (MB1 and MB2), while the outer DT layers (MB3 and MB4) are paired with one RPC each. The endcap CSCs are also paired with one RPC each.

Each RPC consists of parallel plates of bakelite enclosing two millimeter-thick gas gaps. Each gas gap contains a mixture of 96.2% R134a ($C_2H_2F_4$), 3.5% isobutane (C_4H_10), and 0.3% sulfur hexafluoride (SF_6). When a muon passes through the chamber, it ionizes the gas and produces a shower of electrons. The bakelite plates are highly transparent to the electrons, which are read out by external metallic strips.

Figure 4.13 shows an example plot indicating RPC resolution. Similar to pixel resolution, RPC resolution is measured as the residual difference between a “good” muon track in the muon system and the RPC hits[48].

4.6 Particle Reconstruction

Even after a given event is collected, its constituent particles only exist as hit patterns in the various subdetectors of CMS. In order for any kind of meaningful physics analysis to be performed, these hit patterns must be turned into physics objects which are more easily-read by the CMS software framework (CMSSW). Broadly, the six principal objects that are “reconstructed” from

raw detector data in this manner are: electrons, muons, taus, photons, jets, and missing transverse energy (\cancel{E}_T). The chief set of software tools used to reconstruct physics objects is a group of algorithms collectively referred to as Particle Flow (PF).

4.6.1 Particle Flow

The global event reconstruction (also called particle-flow event reconstruction [2, 35]) consists of reconstructing and identifying each single particle with an optimized combination of all subdetector information. In this process, the identification of the particle type (photon, electron, muon, charged hadron, neutral hadron) plays an important role in the determination of the particle direction and energy. Photons are identified as ECAL energy clusters not linked to the extrapolation of any charged particle trajectory to the ECAL. Electrons are identified as a primary charged particle track and ECAL energy clusters corresponding to this track extrapolation to the ECAL and to possible bremsstrahlung photons emitted along the way through the tracker material. Muons are identified as a track in the central tracker consistent with either a track or several hits in the muon system, associated with an energy deficit in the calorimeters. Charged hadrons are identified as charged particle tracks neither identified as electrons, nor as muons. Finally, neutral hadrons are identified as HCAL energy clusters not linked to any charged hadron trajectory, or as ECAL and HCAL energy excesses with respect to the expected charged hadron energy deposit.

The energy of photons is directly obtained from the ECAL measurements. The energy of electrons is determined from a combination of the track momentum at the main interaction vertex, the corresponding ECAL cluster energy, and the energy sum of all bremsstrahlung photons attached to the track. The energy of muons is obtained from the corresponding track momentum. The energy of charged hadrons is determined from a combination of the track momentum and the corresponding ECAL and HCAL energy, corrected for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding ECAL and HCAL energy measurements. For each event, hadronic jets are clustered from these reconstructed particles with the infrared and collinear safe anti- k_T algorithm, operated with a size

parameter of 0.5[31]. The jet momentum is determined as the vectorial sum of all particle momenta in this jet, and is found in the simulation to be within 5 to 10% of the true momentum over the whole p_T spectrum and detector acceptance. Jet energy corrections are derived from the simulation, and are confirmed with in situ measurements with the energy balance of dijet and photon + jet events [31]. The jet energy resolution amounts typically to 15% at 10 GeV, 8% at 100 GeV, and 4% at 1 TeV, to be compared to about 40%, 12%, and 5% obtained when the calorimeters alone are used for jet clustering.

An important quantity related to particle reconstruction is *isolation*. The isolation of a particle is determined by drawing a fixed cone of size parameter $R = \sqrt{\eta^2 + \phi^2}$ around a particle candidate and summing the energies of all of the other particles (originating from the collision point) as measured in the tracker, ECAL, and HCAL. The most-commonly used definition of isolation is *relative isolation*, I_{rel} , which divides the sum of the energy deposits by the p_T of the candidate:

$$I_{rel} = \frac{\sum E_{\text{Tracker}} + \sum E_{\text{ECAL}} + \sum E_{\text{HCAL}}}{p_T(\text{candidate})} \quad (4.3)$$

Isolation requirements are often standardized by physics object groups (POGs) at CMS. These groups are tasked with establishing the criteria used for identifying a given particle, of which isolation is often a significant factor. Typically the POG for a given class of particle (e.g. the muon POG), will define a group of isolation “working points” ranging from loose or very loose to tight. The difference between the working points is typically the maximum value of I_{rel} a candidate must have in order to qualify as a physics object, and the cone size parameter R in which other particles are considered in calculating I_{rel} .

The following subsections will detail the methods used to reconstruct each class of particle (except tau leptons, which will be discussed in greater detail in Section 5.3.4).

4.6.1.1 Track Reconstruction

The goal of the silicon tracker is to reconstruct tracks showing the trajectories of charged particles through the magnetic field, and furthermore to use these trajectories to “trace back” the

origin of the particle, known as the vertex. The vertices that correspond to pp collisions are known as primary vertices (PV). The presence of multiple proton-proton collisions in each bunch crossing, a phenomenon known as pileup, presents the possibility of misidentifying the correct PV for a given track or set of tracks. Therefore, there is a need for a vertexing and tracking algorithms exhibiting a high degree of granularity.

Tracks with the highest number of hits in the pixel and strip subsystems within the silicon tracker are recorded (and reconstructed) first. A “hit” in this case refers to a collection of neighboring strips or pixels registering signal which are grouped into clusters. A particle trajectory can be reconstructed from tracker hits according to the following procedure[49]:

1. Seed generation: First, all possible combinations of either three hits in the pixel detector or two hits and the beam spot are used to generate a trajectory helix assuming a uniform magnetic field. The *beam spot* is a fixed coordinate representing the average position in x, y, and z of the pp collision over many events. These track seeds are required to pass some minimum p_T threshold and maximum impact parameter threshold, where the *impact parameter* or d_{xy} is the minimum distance between the track and either the beam spot or PV.

2. Track finding: Next, the trajectories are extrapolated further into the tracker. At each subsequent layer, each hit within a 3σ region around the extrapolated trajectory are fitted using a Kalman filter[52]. The hit with the smallest χ^2 is accepted and the process is repeated in each subsequent layer until the end of the tracker is reached. In cases where a hit is not found in a layer, a “missing hit” is recorded. Track candidates with more than one missing hit are generally not accepted.

3. Track fitting: In this stage, the requirements on p_T and impact parameter from the seed generation step are removed, and each track candidate is again run through a Kalman filter (refitted). This process starts with the innermost hits and then moves outward through the list of hits, updating the track parameter estimates after each hit.

4. Track selection: This step is designed to reduce the rate of *fake tracks* or reconstructed tracks that are not associated with an actual charged particle. In order to do this, several “quality cuts”

are imposed on the track candidate, such as small impact parameter with respect to a PV (an event may have several PVs due to pileup — in these cases all PVs are considered), low χ^2/dof , and high number of tracker layers with hits.

This tracking procedure is repeated in six total iterations. The first has the most stringent requirements on track p_T and d_{xy} , and each following iteration loosens these requirements. After each iteration, the hits corresponding to the tracks reconstructed in the previous iteration are removed for the next iteration.

4.6.1.2 Primary Vertex Reconstruction

Once the tracks in an event are reconstructed, they are then used to estimate each PV in the event. This PV reconstruction is done according to the following procedure[49]:

- 1. Track selection:** The tracks to be used in the PV reconstruction are selected based on a series of requirements related to track quality as well as likelihood of having been produced in the primary interaction region. These include impact parameter with respect to the beam spot, high number of layers with hits, low χ^2/dof , and high p_T .
- 2. Track clustering:** A clustering algorithm known as *deterministic annealing* is used to assign the selected tracks to estimated PV candidates. These assignments are made on the basis of both the z coordinate of each track’s point of closest approach to the beam spot as well as each track’s uncertainty.
- 3. Vertex fitting:** Each track cluster forms a vertex candidate. Those vertex candidates which have at least two associated tracks are used as input to an *adaptive vertex fitter*[53], which outputs the vertex x , y , and z coordinates, position uncertainty, and χ^2 .

4.6.1.3 Electron Reconstruction

Electrons are reconstructed via the combination of a track in the silicon tracker and hits in the ECAL towers. These tower hits are called “superclusters” and their size in η and ϕ is dictated by various clustering algorithms selected according to the relevant HLT path. As the electron

interacts with the silicon in the tracker, it radiates photons via bremsstrahlung. This emission causes the electron to lose momentum and bend further in the magnetic field. To account for this, the supercluster size is tuned to catch these bremsstrahlung photons as well, thereby capturing the total energy of the electron before bending and leading to a more accurate calculation of the electron's energy and momentum.

To fully reconstruct an electron track, one starts with a track **seed**, which is a possible electron trajectory found by interpolating between layer hits in the pixel and strip detectors. These track seeds must be matched with a supercluster in ECAL. There are two methods to achieve this. The first, ECAL driven seeding, begins by reconstructing superclusters in ECAL and attempts to match them with track seeds in the innermost layers of the pixel detector. The track is then reconstructed from these track seeds. The second, tracker driven seeding, begins with a high-quality track seed and attempts to match it to an appropriate supercluster in ECAL, while also taking into account other clusters arising from bremsstrahlung photons. This is achieved by drawing straight lines tangent to the track towards the ECAL. If a corresponding cluster is found, it is added to the total energy of the electron.

4.6.1.4 Muon Reconstruction

Muons can be measured both in the silicon tracker and the muon system. In order to reconstruct a muon, hits in the muon system (**stand-alone muons**) are matched to hits in the tracker in order to reconstruct a complete muon object, known as a **global muon**. These muons are reconstructed using two different methods, **global muon reconstruction (outside-in)** and **tracker muon reconstruction (inside-out)**.

Global muon reconstruction begins in the muon system. Hits in the DTs and CSCs are allocated into segments, which are short stubs containing just enough hits to assign each a momentum and direction vector. These segments are grouped together to form tracks in the muon system. These tracks, known as stand-alone muon tracks, are matched, one by one, with tracks from the tracker by comparing parameters of the two tracks propagated onto a common surface.

Tracker muon reconstruction begins in the silicon tracker. All tracks with $p_T > 0.5$ GeV and $p > 2.5$ GeV are considered possible muon candidates and are propagated outward toward the muon system, taking into account the magnetic field, expected energy loss, and the possibility for multiple Coulomb scattering in the detector material. If a segment in the muon system is found within 3 cm in local x, y coordinates, the candidate qualifies as a tracker muon.

Tracker muon reconstruction is more efficient than global muon reconstruction at low energies ($p < 5$ GeV), since it only requires a single segment in the muon system. Global muon reconstruction is more appropriate for muons with higher energies penetrating through more than one station in the muon system.

4.6.1.5 Photon and Neutral Hadron Reconstruction

Since photons and neutral hadrons are chargeless, and therefore leave no hits in the silicon tracker, their reconstruction takes place in ECAL (photons) and HCAL (hadrons). The identification of photons and neutral hadrons takes place via *clustering algorithms* which are designed to separate photons from neutral as well as charged hadrons, and to differentiate photons originating from the primary or secondary vertices from those generated by electrons undergoing bremsstrahlung. Photons may also convert into e^+e^- pairs in the tracker, and further clustering algorithms have been developed in order to capture these conversions and include them in the photon reconstruction.

Clustering algorithms in ECAL involve three steps. First, a seed cluster is identified in one of the ECAL towers as a tower with a local energy maximum passing a threshold specified by the algorithm. Next, topological clusters are defined as clusters adjacent to (sharing at least one side in common with) the seed cluster and with cell energies passing another threshold. These thresholds are generally set to be two standard deviations above the electronics noise level in each of the separate ECAL segments (80 MeV in the barrel and 300 MeV in the endcaps). Finally, topological clusters then generate as many “particle flow” clusters as there are seed clusters, and the energy in each particle flow cluster is allocated according to its distance from the seed. Furthermore, in particle flow, ECAL clusters are required to **not** match any tracks in the tracker in order to qualify

as photon candidates. Similarly, neutral hadrons are identified by the presence of clusters in HCAL with no corresponding track.

4.6.1.6 Jet Reconstruction

Jets are primarily reconstructed in ECAL and HCAL, although their charged constituents usually leave tracks in the silicon tracker as well. The majority of reconstructed jet energy comes from charged particles such as pions and kaons, and a smaller but still sizable portion comes from photons from π^0 decays in ECAL. Jets are reconstructed offline from the energy deposits in the calorimeter towers, clustered by the anti- k_t algorithm [24, 26] with a size parameter of 0.5. In this process, the contribution from each calorimeter tower is assigned a momentum, the absolute value and the direction of which are given by the energy measured in the tower, and the coordinates of the tower. The raw jet energy is obtained from the sum of the tower energies, and the raw jet momentum by the vectorial sum of the tower momenta, which results in a nonzero jet mass. The raw jet energies are then corrected to establish a relative uniform response of the calorimeter in η and a calibrated absolute response in transverse momentum p_T .

4.6.1.7 Missing Transverse Energy

The colliding protons in the LHC have zero net transverse momentum. By conservation of momentum, it's expected that the particles generated from each pp collision will have net zero total transverse momentum. Some of these particles (particularly neutrinos) will evade CMS undetected, however. The sum of their momenta is known as missing transverse energy, denoted as \cancel{E}_T . Since we must have $\sum_{\text{detected particles}} \vec{p}_T + \sum_{\text{undetected particles}} \vec{p}_T = 0$, we can express the \cancel{E}_T as

$$\cancel{E}_T = - \sum_{\text{detected particles}} \vec{p}_T \quad (4.4)$$

This principle only works in the transverse plane, as overall momentum is not necessarily conserved in z due to the inelastic parton-parton scattering that takes place in pp collisions. CMS employs multiple definitions of \cancel{E}_T , but the one used in this thesis is *particle flow* \cancel{E}_T , or *PF-* \cancel{E}_T . PF- \cancel{E}_T

uses all particle-flow objects as well as all available detector information in calculating the \cancel{E}_T according to Equation 4.4.

4.7 The Trigger System

When running at full luminosity, the LHC collides proton bunches at a rate of 40 MHz. With each event requiring $\approx 1\text{-}2$ MB of digital storage space, this equates to a data-generation rate of ≈ 50 TB per second. Since this is obviously unsustainable with current computer storage technology, CMS employs a trigger system to reduce the rate of stored events. Two levels of triggering are used: the hardware-based Level-1 (L1) trigger, which reduces the rate from 40 MHz to approximately 100 kHz, and the hardware-and-software-based High Level Trigger (HLT), which reduces the rate from ≈ 100 kHz to less than 1 kHz[33].

4.7.1 The Level-1 (L1) Trigger)

The L1 trigger takes data from the muon system, HCAL, and ECAL in order to make a rapid decision about whether or not to keep the event and send it to the HLT for further review. The L1 trigger latency is $3.2 \mu\text{s}$, during which time the entirety of the data from the event is temporarily stored. Information from the silicon tracker is not used in this stage of triggering, as the time required by the tracker to reconstruct a track is outside this triggering time window.

The L1 trigger is divided into three major components: the calorimeter trigger, the muon trigger, and the global trigger. The calorimeter trigger uses the combined (summed) energy deposits in ECAL and HCAL. This information is then passed to the global trigger. The muon trigger uses the DTs, CSCs, and RPCs which in conjunction form the muon system. The muon trigger takes the “best” four muon candidates from the muon system, where a “good” muon candidate is defined as having high transverse momentum and a high-resolution track, and passes their information to the global trigger. Using a series of AND-OR algorithms, the global trigger makes the final decision to reject the event or to send it to the HLT.

4.7.2 The High level Trigger (HLT)

The HLT is software-based and uses software-defined “HLT paths” in order to determine whether or not a given event will pass a particular physics object selection determined by the path. The advantage of this setup is that different HLT paths can be used by different analysis groups studying different physics signatures.

The HLT reconstructs the event from the stored data passed to it by the L1 trigger, and determines whether the event passes a given trigger path. There are four broad categories of HLT path: electrons/photons and muons, jets, missing transverse energy (E_T), and taus. Hundreds of specific paths offering high levels of discrimination between required selection criteria are available and grouped into these four categories. As the paths are software-based and not hardware-based, new paths can be customized to the needs of any analysis effort.

Chapter 5

Search for $Z' \rightarrow \tau\tau$ events at $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV

5.1 Strategy

The τ lepton is the heaviest known lepton with a mass of 1.777 GeV and a lifetime of 2.9×10^{-13} seconds. Around one third of all τ leptons decay to e/μ , and the remainder decay into hadronic jets (τ_h). In the latter case, a τ_h corresponds to one, three, or (rarely) five charged mesons usually accompanied by one or more neutral pions. Each τ also decays into a τ neutrino and, in the case of the e and μ decays, an additional e or μ antineutrino.

We consider four distinct analyses for pairs of τ lepton decays, namely $e\mu$ (6.2%), $e\tau_h$ (23.1%), $\mu\tau_h$ (22.5%), and $\tau_h\tau_h$ (42%), where the percentages indicate the branching fraction (probability of decay) for each channel. We ignore the $Z' \rightarrow e^+e^-$ and $Z' \rightarrow \mu^+\mu^-$ channels due to the copious Drell-Yan $Z/\gamma^* \rightarrow e^+e^-,\mu^+\mu^-$ production, although there is ongoing development of algorithms to discriminate prompt e/μ from τ lepton decays to light leptons, which will be utilized in the 2016 data/analysis.

The overall strategy of the analysis is similar in all channels. In general, we identify events with two oppositely charged, nearly back-to-back objects. Because the $\tau\tau$ system decays with up to four neutrinos, we expect to have missing transverse energy (\cancel{E}_T) present in the event. In contrast to Z' searches in the e^+e^- and $\mu^+\mu^-$ channels, the visible $\tau^+\tau^-$ mass does not produce a narrow peak due to the missing neutrinos. Instead, we look for a broad enhancement in the $\tau^+\tau^- + \cancel{E}_T$ invariant mass distribution consistent with new physics. Our selections maintain high efficiency for signal events, provide strong background suppression, and reduce the influence of

systematic effects.

As $Z \rightarrow \tau^+ \tau^-$ is both background as well as an important validation signal, our final selection requirements are such that by removing or reversing just a few cuts we can obtain a clean sample of $Z \rightarrow \tau^+ \tau^-$ events. In order to ensure robustness of the analysis and our confidence in the results, whenever possible we rely on the data itself to understand and validate the efficiency of reconstruction methods as well as the estimation of the background contributions. For that purpose, we define control regions with most of the selection criteria similar to what we use in our main search, but enriched with events from background processes. Once a background enhanced region is created, we measure the selection efficiencies in those regions and extrapolate to the region where we expect to observe our signal. In cases where a complete data–driven method is not possible, we make use of scale factors, the ratio between observed data events and expected Monte Carlo (MC) events in the background enhanced region to estimate the background contribution in the signal region. Monte Carlo events are simulated physics events, wherein a specific physical process (e.g. $Z' \rightarrow \tau\tau$) is modeled via computer simulation and fed into a simulated CMS detector. MC simulation is extremely useful for particle physics analyses, as it lets analyzers predict the performance of the detector as well as tune their selection cuts. Although each individual channel could have its own set of requirements, we maintain, wherever possible, consistent definitions and selection criteria between channels.

To quantify the significance of any possible excess or set upper limits on the production rate, we perform a fit of the $m(\tau_1, \tau_2, \cancel{E}_T)$ mass distribution and employ the CL_s technique to interpret the results in terms of the upper 95% credibility level limits for each channel. The joint limit is obtained by combining the posterior probability density functions (likelihood) and taking into account correlation of systematic uncertainties within and across channels.

As there are substantial similarities between the analyses performed at 8 TeV and at 13 TeV, this chapter will present the more current 13 TeV analysis carried out in 2015. The 8 TeV results will be included in Section 5.10. The structure of this chapter is as follows: Sections 5.2 and 5.4 describe the data sets used in each analysis. Section 5.3 provides a brief discussion of the

reconstruction and identification of the objects used to reconstruct our $\tau\tau$ pairs. Sections 5.5–5.8 describe the specific selection criteria and background extraction methods applied to each individual channel. In Sections 5.9–5.10 we describe the statistical method used to extract the 95% C.L. upper limits and the results of the analysis.

5.2 Data and Monte Carlo Samples

The 13 TeV collision data collected by the CMS detector in year 2015 is used in this analysis. Table 5.1 shows the collision datasets used. The official CMS “good run list”, containing information on fill quality, is used to select “good” run ranges and lumi sections. The total integrated luminosity of the collision data samples is 2.11 fb^{-1} .

The official Spring 2015 MC samples are used for all Standard Model processes. The leading order generators, PYTHIA8 and MADGRAPH, were mainly used for signal and background MC production. The predicted background yields in simulation were determined using cross sections with the strong coupling constant expanded to next-leading-order (NLO) or next-next-leading-order (NNLO), while the signal yields and distributions in all plots shown in this analysis were normalized using the leading order cross-sections shown in Table 5.2. Table 5.2 shows the entire list of MC samples used for this analysis.

Table 5.1: Collision Data Samples

Physics Sample	Official CMS Datasets
Run 2015C SingleMu 05 Oct ReMiniAOD	<i>/SingleMuon/Run2015C_25ns-05Oct2015-v1/MINIAOD</i>
Run 2015D SingleMu 05 Oct ReMiniAOD	<i>/SingleMuon/Run2015D-05Oct2015-v1/MINIAOD</i>
Run 2012D SingleMu PromptReco v4	<i>/SingleMuon/Run2015D-PromptReco-v4/MINIAOD</i>
Run 2015C SingleElectron 05 Oct ReMiniAOD	<i>/SingleMuon/Run2015C_25ns-05Oct2015-v1/MINIAOD</i>
Run 2015D SingleElectron 05 Oct ReMiniAOD	<i>/SingleElectron/Run2015D-05Oct2015-v1/MINIAOD</i>
Run 2012D SingleElectron PromptReco v4	<i>/SingleElectron/Run2015D-PromptReco-v4/MINIAOD</i>
Run 2015C Tau 05 Oct ReMiniAOD	<i>/Tau/Run2015C_25ns-05Oct2015-v1/MINIAOD</i>
Run 2015D Tau 05 Oct ReMiniAOD	<i>/Tau/Run2015D-05Oct2015-v1/MINIAOD</i>
Run 2012D Tau PromptReco v4	<i>/Tau/Run2015D-PromptReco-v4/MINIAOD</i>

Because the MC simulated samples contain a pileup (PU) distribution that does not match

Table 5.2: MC Samples used in 13 TeV analysis. The multiplicative terms take the cross section from leading order to NLO.

Process	cross-section (pb)	Official CMS Datasets (MINIAODSIM)
$Z \rightarrow ll$	6025.2	$/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
HT binned	147.4*1.23	$/DYJetsToLL_M-50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1/$
LO samples	40.99*1.23	$/DYJetsToLL_M-50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1/$
	5.678*1.23	$/DYJetsToLL_M-50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v2$
	2.198*1.23	$/DYJetsToLL_M-50_HT-600toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
$Z \rightarrow ll$	6025.2	$/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
mass binned	7.67*0.987	$/DYJetsToLL_M-200to400_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
NLO samples	0.423*0.987	$/DYJetsToLL_M-400to500_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
	0.24*0.987	$/DYJetsToLL_M-500to700_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v3$
	0.035*0.987	$/DYJetsToLL_M-700to800_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
	0.03*0.987	$/DYJetsToLL_M-800to1000_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
	0.016*0.987	$/DYJetsToLL_M-1000to1500_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/RunIISpring15MiniAOv2-Asympt25ns_74X_mcRun2_asymptotic_v2-v1$
$W + Jets$	61526.7	$/WJetsToLNu_HT-100To200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
HT binned	1345*1.21	$/WJetsToLNu_HT-200To400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
LO samples	359.7*1.21	$/WJetsToLNu_HT-400To600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
	48.91*1.21	$/WJetsToLNu_HT-600ToInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
	18.77*1.21	$/WJetsToLNu_HT-700to800_TuneCUETP8M1_13TeV-madcatnloFXFX-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
$t\bar{t}$	831.76	$/TT_TuneCUETP8M1_13TeV-powheg-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-ext3-v1$
single Top samples	35.6	$/ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
	35.6	$/ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v2$
	136.02*0.108*3	$/ST_t-channel_top_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
	80.95*0.108*3	$/ST_t-channel_antitop_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v2$
VV	11.95	$/VVToLE2Nu_13TeV_amcatnloFXFX_madspin_pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
$ZZ \rightarrow 2l2q$	3.22	$/ZZTo2L2Q_13TeV_V_amcatnloFXFX_madspin_pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
$ZZ \rightarrow 4l$	1.212	$/ZZTo4L_13TeV_V_amcatnloFXFX-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
$WW \rightarrow ll2q$	49.997	$/WWTo1LNu2Q_13TeV_V_amcatnloFXFX_madspin_pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
$WZ \rightarrow 2l2q$	5.595	$/WZTo2L2Q_13TeV_V_amcatnloFXFX_madspin_pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
$WZ + Jets$	5.26	$/WZJets_TuneCUETP8M1_13TeV_amcatnloFXFX-pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
$WZ \rightarrow l3\nu$	3.05	$/WZTo1LNu3N_13TeV_V_amcatnloFXFX_madspin_pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
$WZ \rightarrow ll2q$	10.71	$/WZTo1LNu2Q_13TeV_V_amcatnloFXFX_madspin_pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
QCD samples	720648000	$/QCD_Pt200toInf_MuEnrichedPt15_TuneCUETP8M1_13TeV_pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
$Z'(500)$	9.33	$/ZprimeToTauTau_M_500_TuneCUETP8M1_tauola_13TeV_pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
$Z'(1000)$	0.468	$/ZprimeToTauTau_M_1000_TuneCUETP8M1_tauola_13TeV_pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
$Z'(1500)$	0.0723	$/ZprimeToTauTau_M_1500_TuneCUETP8M1_tauola_13TeV_pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
$Z'(2000)$	0.0173	$/ZprimeToTauTau_M_2000_TuneCUETP8M1_tauola_13TeV_pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
$Z'(2500)$	0.00554	$/ZprimeToTauTau_M_2500_TuneCUETP8M1_tauola_13TeV_pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$
$Z'(3000)$	0.00129	$/ZprimeToTauTau_M_3000_TuneCUETP8M1_tauola_13TeV_pythia8/RunIISpring15MiniAOv2-74X_mcRun2_asymptotic_v2-v1$

that of data, the MC needs to be properly weighted to fit the PU distribution observed in data. The reweighting of MC events is performed by determining the probabilities to obtain n interactions in data ($P_{data}(n)$) and MC ($P_{MC}(n)$) and using the event weights

$$w_{PU}(n) = \frac{P_{data}(n)}{P_{MC}(n)} \quad (5.1)$$

to reweight MC events based on the number of interactions. In Section 5.3.6, our understanding of PU and the performance of the PU reweighing method is validated.

5.3 Physics Object Reconstruction

5.3.1 Jet Reconstruction

The particle-flow (PF) technique [2, 3] is used to improve the jet p_T and angular resolution in this analysis. The PF technique combines information from different subdetectors to produce a mutually exclusive collection of particles (namely muons, electrons, photons, charged hadrons and neutral hadrons) that are used as input for the jet clustering algorithms. Jets are clustered using the anti- k_T algorithm [25], with a distance parameter of $\Delta R = 0.4$ in $\eta\text{-}\phi$ plane (defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$) by summing the four-momenta of individual PF particles.

The jets require energy corrections obtained using simulated events that are generated with PYTHIA, processed through a detector simulation based on GEANT4, and confirmed with in situ measurements of the p_T balance. The overall jet-energy corrections depend on the η and p_T values of jets. These jet-energy corrections are known as L1 FastJet, L2 Relative, and L3 Absolute corrections. In order to remove the extra energy in jets from underlying event (UE) and pileup (PU), the L1 FastJet corrections use the event-by-event UE/PU densities. The L2 and L3 corrections use jet balancing and photon+jet events to improve and provide a better energy response as a function of jet p_T and η . For data, additional residual corrections are applied.

In this analysis, jets are required to have $p_T > 30$ GeV and $|\eta| < 2.4$. For the identification of jets, the loose PF ID is used as recommended by the CMS Jet/ E_T physics object group (POG).

The “loose” working point requires a jet candidate to have $p_T > 10$ GeV, charged hadron fraction > 0.0 , neutral hadron fraction < 0.99 , charged EM fraction < 0.99 , and neutral EM fraction < 0.99 . “Hadron fraction” refers to the fraction of the jet’s energy taken from HCAL hits, and “EM fraction” refers to the fraction of jet energy taken from ECAL hits. Table 4 shows the selection criteria used for the recommended loose PF ID, which are validated in other studies [9]. The jet reconstruction and ID efficiency in simulation is $>98\%$.

Table 5.3: Loose Jet-ID Selections.

Selection	Cut
Neutral Hadron Fraction	< 0.99
Neutral EM Fraction	< 0.99
Number of Constituents	> 1
And for $ \eta < 2.4$ in addition apply	
Charged Hadron Fraction	> 0
Charged Multiplicity	> 0
Charged EM Fraction	< 0.99

5.3.1.1 b-Jet Tagging

In this analysis, b-tagged jets are used for two purposes: to reduce $t\bar{t}$ background in the signal region and to obtain $t\bar{t}$ enriched control samples used to estimate the signal rate.

The CSVv2 algorithm [29] is used to identify a jet as originating from hadronization of a b-quark. This algorithm combines reconstructed secondary vertex and track-based lifetime information to build a likelihood-based discriminator to distinguish between jets from b-quarks and those from charm or light quarks and gluons.

The minimum thresholds on these discriminators define loose, medium, and tight operating points with a b-jet misidentification probability of about 10%, 1%, and 0.1%, respectively, with an average jet p_T of about 80 GeV. The loose operating point with an efficiency of about 85% is used in this analysis. A sample of pair-produced top quark events is used to measure b-tagging efficiency using several methods [8]. A scale factor is applied to correct for differences in b-tagging efficiency

between data and simulation [30].

5.3.2 Electron Reconstruction and Identification

Electrons are reconstructed using information from the tracker and ECAL detectors. Electrons passing through the silicon tracker material lose energy due to bremsstrahlung radiation. The energy of the radiated photons is spread over several crystals of the ECAL detector along the electron trajectory, mostly in the ϕ direction (the magnetic field is in the z direction). Two algorithms based on energy clustering, “Hybrid” for the barrel and “Island” for the endcaps, are used to measure the energy of electrons and photons [17].

Electron tracks are reconstructed by matching trajectories in the silicon strip tracker to seed hits in the pixel detector. A pixel seed is composed of two pixel hits compatible with the beam spot. A Gaussian Sum Filter (GSF) is used for the reconstruction of trajectories in the silicon strips. In order to minimize the many possible trajectories due to different combinations of hits, the track that best matches an energy supercluster in the ECAL is chosen to be the reconstructed track.

The preselection of primary electron candidates requires good geometrical matching and good agreement between the momentum of the track and the energy of the ECAL supercluster. Two quantities used to estimate the geometrical matching are $\Delta\eta_{in} = \eta_{sc} - \eta_{vertex}^{Track}$ and $\Delta\phi_{in} = \phi_{sc} - \phi_{vertex}^{Track}$. The η_{sc} and ϕ_{sc} coordinates correspond to the supercluster position and are measured using an energy weighted algorithm. The η_{vertex}^{Track} and ϕ_{vertex}^{Track} coordinates are the position of the track at the interaction vertex extrapolated, as a perfect helix, to the ECAL detector. The good energy-momentum matching is measured by taking the ratio between the corrected energy E_{corr} in the ECAL supercluster and the momentum of the track P_{in} measured in the inner layers of the tracker.

The electron selection requirements have two main components: electron identification (eID) and electron isolation. In this analysis we use the non-triggering MVA (multivariate analysis) electron identification recommended by the e/ γ POG. The MVA cuts used to define the 80% and

90% signal efficiency working points are summarized in Table 5.4.

Table 5.4: Electron ID selection criteria

Category	MVA _{min cut} (80% signal eff)	MVA _{min cut} (90% signal eff)
Barrel ($\eta < 0.8$) $5 < p_T < 10$	0.287435	-0.083313
Barrel ($\eta > 0.8$) $5 < p_T < 10$	0.221846	-0.235222
Endcap $5 < p_T < 10$	-0.303263	-0.67099
Barrel ($\eta < 0.8$) $p_T > 10$	0.967083	0.913286
Barrel ($\eta > 0.8$) $p_T > 10$	0.929117	0.805013
Endcap $p_T > 10$	0.726311	0.358969

5.3.3 Muon Reconstruction and Identification

Muon reconstruction is a multistep process that begins with the information gathered from the muon subdetector. As a first step, standalone muons are reconstructed from hits in the individual drift tube (DT) and cathode strip (CSC) chambers. Hits from the innermost muon stations are combined with hits in the other muon segments using the Kalman fitting technique[52]. The standalone muon trajectory is extrapolated from the innermost muon station to the outer tracker surface. This standalone trajectory is then used to find a matching track reconstructed in the inner silicon tracker. Finally, standalone muons and matching silicon tracks are used to perform a global fit resulting in “global” muons.

Once a muon is required to have matching tracks in the inner and outer detectors, the main source of background consists of charged hadrons that leave a signature in the inner silicon tracker while also penetrating through the hadronic calorimeter and creating hits in the muon chambers. Charged hadrons that penetrate the hadronic calorimeter and leave hits in the muon system will deposit significant energy in the calorimeters. Therefore, a calorimeter compatibility algorithm can be used to significantly reduce the number of charged pion fakes. However, calorimeter compatibility is not used in this analysis due to our uncertainty of the performance of such algorithms in the presence of high PU. The presence of punch-throughs often occur due to pions from the hadronization of quarks and gluons. These punch-throughs can often be discriminated against by

requiring isolation. Similarly, non-prompt muons from heavy flavor decays and decays in flights (secondary decays) are expected to be within jets and can be discriminated against by imposing an isolation requirement. Muon identification is described in more detail in [39] and [43].

Isolated muons are required to have minimal energy from PF neutral and charged candidates in a cone of $\Delta R = 0.4$ around the lepton trajectory. PF charged candidates considered in the calculation of isolation are required to be near the primary vertex. Relative isolation is used for muons, and it is defined as:

$$I = \frac{\sum_i p_T^i}{p_T^\mu} \quad (5.2)$$

where the index i runs over PF neutral and charged candidates. Table 5.5 shows the complete list of for the “isMedium” μ identification criteria used in this analysis. In all channels, the identification and isolation used follows the POG recommended criteria.

Table 5.5: μ Identification requirements

Cut
<pre> muon::isLooseMuon(recoMu) recoMu.innerTrack()-> validFraction() > 0.8 AND recoMu.globalTrack()-> normalizedChi2() < 3 recoMu.combinedQuality().chi2LocalPosition < 12 recoMu.combinedQuality().trkKink < 20 muon::segmentCompatibility(recoMu) > 0.303 OR muon::segmentCompatibility(recoMu) > 0.451 RellIso < 0.15 </pre>

5.3.4 Tau Reconstruction and Identification

The challenge in identifying hadronically decaying taus is discriminating against generic quark and gluon QCD jets, which are produced with a cross section several orders of magnitude larger. CMS has developed several algorithms to reconstruct and identify hadronically decaying taus based on Particle Flow (PF) objects. For this analysis, the tau POG recommended Hadron Plus Strips algorithm (HPS) is used. HPS makes use of PF jets as inputs to an algorithm that uses strips

of clustered electromagnetic particles to reconstruct neutral pions. The electromagnetic strips (“neutral pions”) are combined with the charged hadrons within the PFJets to reconstruct the main tau decay modes outlined in Table 5.6.

Table 5.6: Reconstructed Tau Decay Modes

HPS Tau Decay Modes
Single Charged Hadron + Zero Strip
Single Charged Hadron + One Strip
Single Charged Hadron + Two Strips
Two Charged Hadrons
Three Charged Hadrons

The single hadron plus zero strips decay mode attempts to reconstruct $\tau \rightarrow \nu\pi^\pm$ decays or $\tau \rightarrow \nu\pi^\pm\pi^0$ decays where the neutral pion has very low energy. The single hadron plus one or two electromagnetic strips attempts to reconstruct tau decays that produce neutral pions where the resulting neutral pion decays produce collinear photons. Similarly, the single hadron plus two strips mode attempts to reconstruct taus that decay via e.g. $\tau \rightarrow \nu\pi^\pm\pi^0$ where the neutral pion decays to well-separated photons resulting in two electromagnetic strips. The three charged hadrons decay mode attempts to reconstruct tau decays that occur with three charged pions (“3-prongs”). Although it is possible to recover signal in the two charged hadron decay mode (in case one of the three prongs is not reconstructed), this mode is not considered as its inclusion reduces discrimination performance and hurts the limit. In all cases, electromagnetic strips are required to have $E_T > 1$ GeV. Additionally, the particle flow charged hadrons are required to be compatible with a common vertex and have a net charge of $|q| = 1$.

In order to enforce the isolation requirement on the reconstructed tau, a region of size $R = 0.5$ around the tau decay direction is defined. Any PF candidates not used for the reconstruction of electromagnetic strips and charged hadrons not involved in the reconstruction of the tau decay modes are used to calculate isolation. The “Tight”, “Medium”, and/or “Loose” 3-hits isolation working points are used as recommended by the tau POG.

Pileup can be a confounding problem in calculating isolation, so corrections for accidentally

including particles from the wrong PV in the isolation calculation are applied. It has been empirically determined that the total energy for neutral hadrons and photons coming from pileup is on average equal to about one half of the total energy of charged hadrons from the PV. The correction for this, known as the $\delta\beta$ correction, manifests itself as an extra term in the numerator of the isolation calculation, Equation 4.3. It is applied by subtracting off one half of the p_T sum of charged hadrons in the event *not* originating from the chosen PV. Unless otherwise stated, $\delta\beta$ corrections are applied to every isolation used in this analysis.

In order to discriminate against muons, HPS taus are required to pass the lepton rejection discriminator which requires the lead track of the tau not be associated with a global muon signature. In order to discriminate against electrons, HPS taus are required to pass a MVA discriminator which uses the amount of HCAL energy associated to the tau with respect to the measured momentum of the track (H/p). Additionally, the MVA discriminator considers the amount of electromagnetic energy in a narrow strip around the leading track with respect to the total electromagnetic energy of the tau. Finally, HPS taus must not reside in any gaps in the ECAL (“cracks”). In all channels, the identification and isolation used follows the Tau POG recommended criteria. The exact discriminator names and working points for each channel are listed and described in their respective sections.

5.3.4.1 Efficiency of Tau Identification discriminators

The efficiency of the τ_h ID discriminators used in the analysis are studied using Z'_{SSM} samples with $m(Z') = 2000$ GeV. For this purpose, we require the reconstructed τ_h to have $p_T > 45$ GeV and pseudorapidity $|\eta| < 2.1$. Further, a reconstructed τ_h is required to be matched to a generator-level (MC simulated) tau with $\Delta R(\tau_{reco}, \tau_{gen}) < 0.3$. The efficiency of the Decay Mode Finding (DMF) discriminator “DecayModeFinding” is found to be relatively flat at $\sim 80\%$ as shown in Figure 5.1. The individual efficiencies of anti-muon, anti-electron, and isolation discriminators, relative to the DMF criterion, are shown in Figures 5.2–5.4. The overall efficiency of the τ_h ID selection used in this analysis is $\sim 55\%$. We use the POG-recommended VLoose working point of the anti-electron

MVA5 discriminator as the tighter working points have poor performance (i.e. low efficiency that is also not constant vs. p_T). The relative flatness of the τ_h identification efficiency with p_T will facilitate the use of $Z \rightarrow \tau\tau$ events as a "standard candle" for comparison with signal.

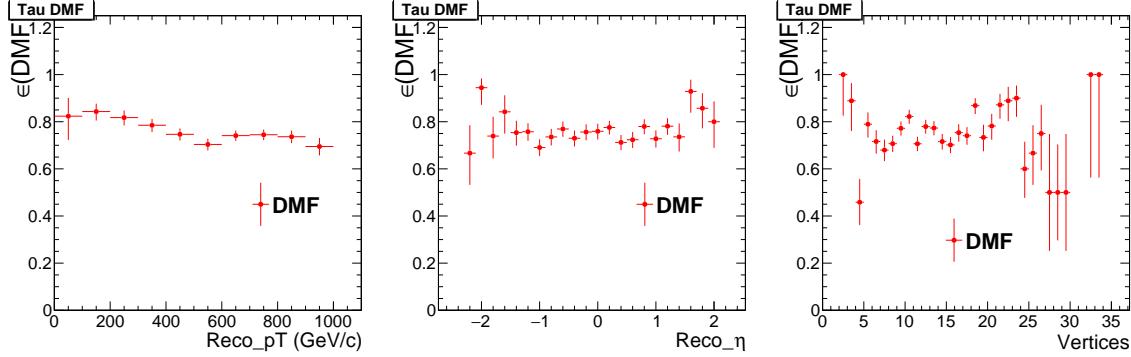


Figure 5.1: Tau Decay Mode Finding efficiency as a function of p_T , η , and number of reconstructed vertices

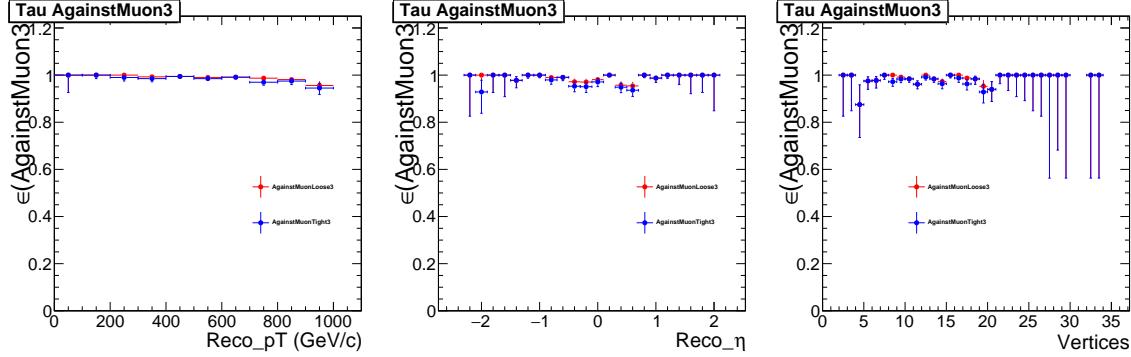


Figure 5.2: Efficiency of tau anti-muon discriminators ("Loose" working point in red and "Tight" in blue) as a function of p_T , η , and number of reconstructed vertices

5.3.4.2 Tau Energy Scale and Resolution

Since the resolution and scale of our mass reconstruction depends on the effectiveness of the τ_h reconstruction, in this section we summarize our studies on τ_h response and resolution. We define the response as the difference between the transverse momentum of a reconstructed tau (that has passed all of the tau ID discriminators) and the transverse momentum of a generated tau that has been matched ($\Delta R < 0.3$) to the reconstructed tau. We can see from Figure 5.5 (right) that the

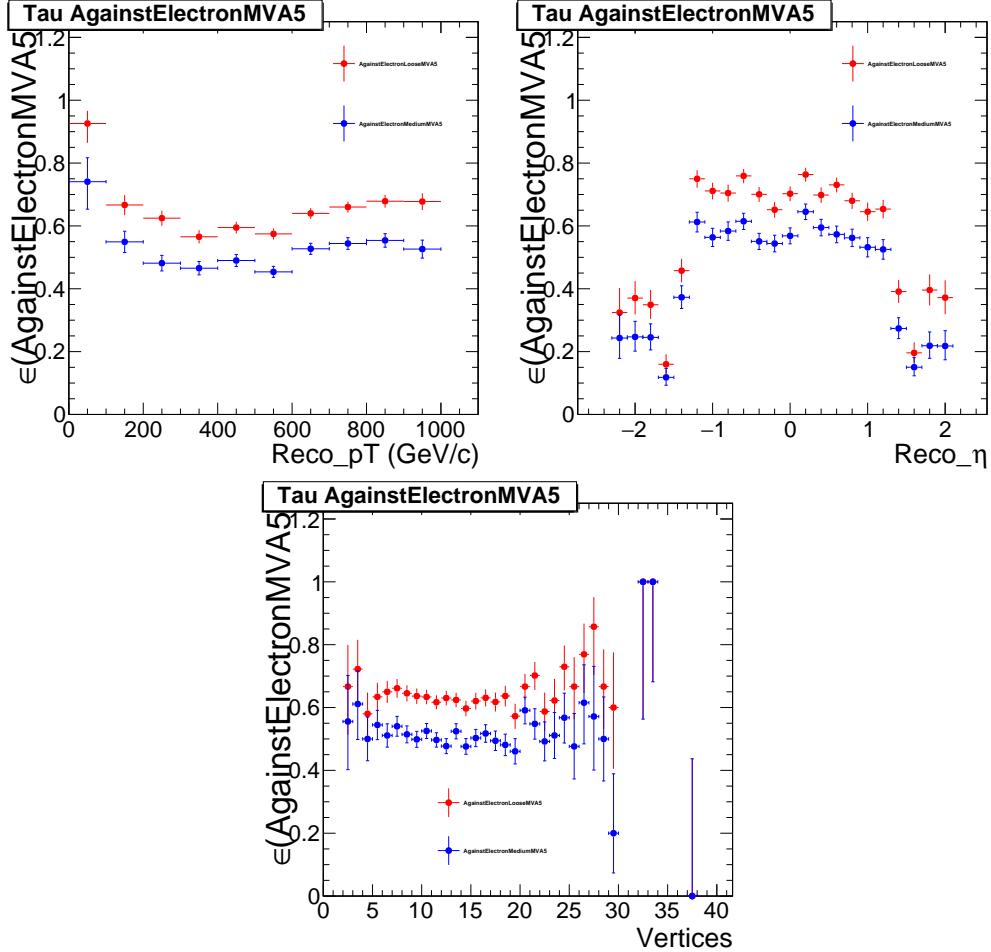


Figure 5.3: Efficiency of MVA-based anti-electron discriminators for taus (“Loose” working point in red and “Medium” in blue) as a function of p_T , η , and number of reconstructed vertices

response distribution contains a narrow Gaussian-like component in addition to a relatively long tail (compared to electrons and muons). While the tails become less substantial at high p_T , the Gaussian-like component of the response distributions broadens at high p_T .

5.3.5 Missing Transverse Energy (\cancel{E}_T)

The presence of neutrinos in the $\tau\tau$ decays must be inferred from the imbalance of total momentum because they escape from the detector without producing any direct response. The magnitude of the negative vector sum of the visible transverse momentum is known as missing transverse energy, denoted \cancel{E}_T .

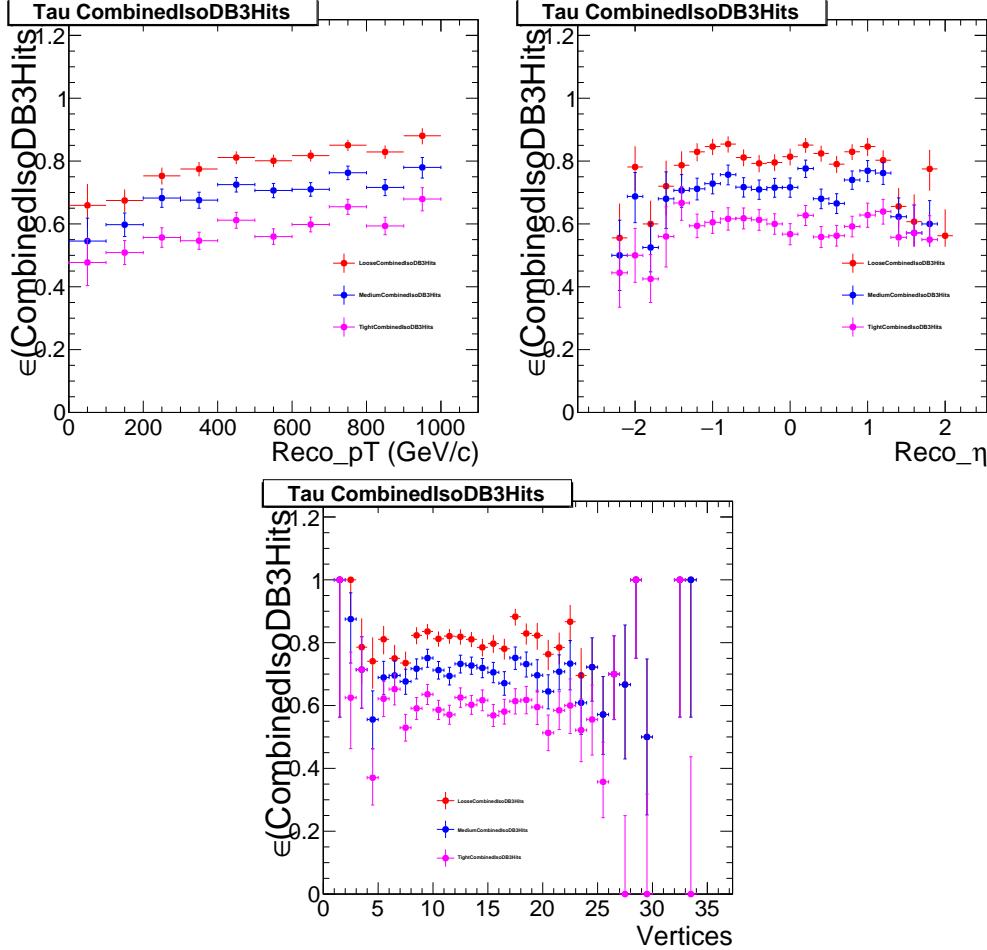


Figure 5.4: Efficiency of isolation discriminator “CombinedIsoDB3Hits” for taus (“Loose” working point in red, “Medium” in blue and “Tight” in purple) as a function of p_T , η , and number of reconstructed vertices

Missing transverse energy is one of the most important observables for discriminating the signal events from background events which do not contain neutrinos, such as QCD multijet and $\text{DY} \rightarrow \mu\mu/\text{ee}$ events. Because there are many factors which may result in artificial \cancel{E}_T , comprehensive studies of \cancel{E}_T as measured using the CMS detector have been performed [6]. We require events with $\cancel{E}_T > 30$ GeV in all channels.

Finally, the standard and recommended “MET Filters” are utilized in this search[2][6].

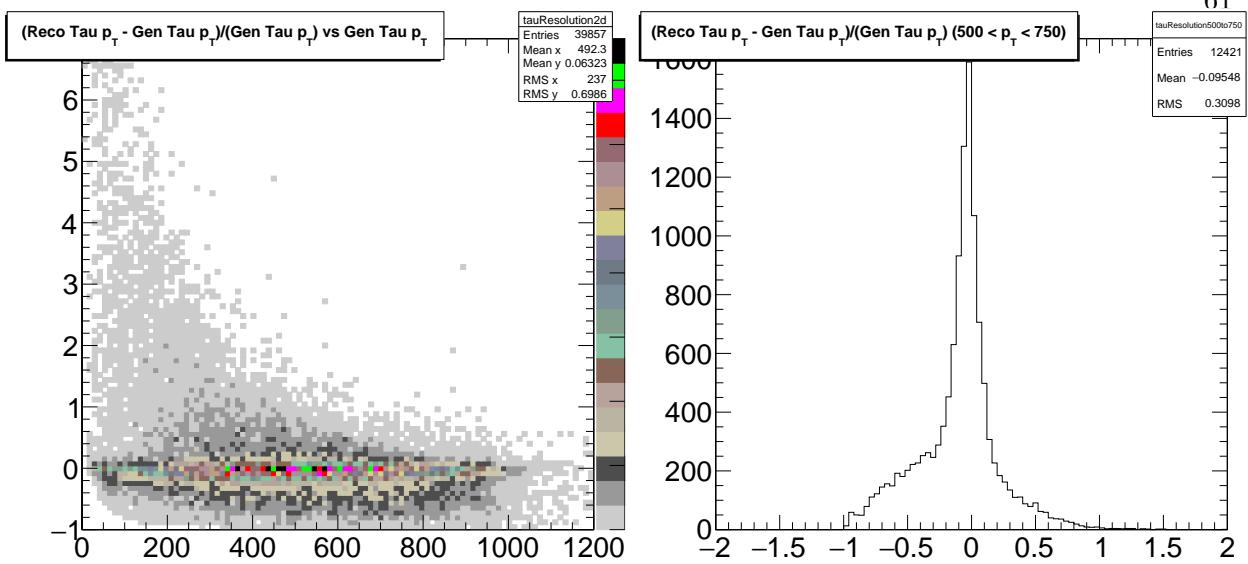


Figure 5.5: Energy response of reconstructed taus, raw and vs. generated tau p_T

5.3.6 Corrections for Pile-Up

Quantities such as \cancel{E}_T and isolation, where energy depositions are summed up over some range of the detector, can suffer large inefficiencies or systematic effects due to particles from pile-up interactions. Therefore, a simple and robust method must be employed to subtract off this contribution. In the case of \cancel{E}_T , this is done by using the FastJet corrections to determine the density of PU on an event by event basis. For isolation, the recommended $\delta\beta$ corrections are applied. The recommended PU corrections have been included for all objects in all the channels.

We apply event-by-event the “official” pile-up weights recommended by the Jet/MET POG. Figure 5.6 shows the distributions of the number of reconstructed vertices, before and after applying the pile-up weights.

5.4 Triggers

For the $e\tau_h$ and $\mu\tau_h$ final states, we use single electron or muon triggers instead of $l \times \tau_h$ cross-triggers (which require both an electron or muon and hadronic tau) to maintain a similar strategy across channels. This also allows us to use the τ_h isolation sidebands as control and validation

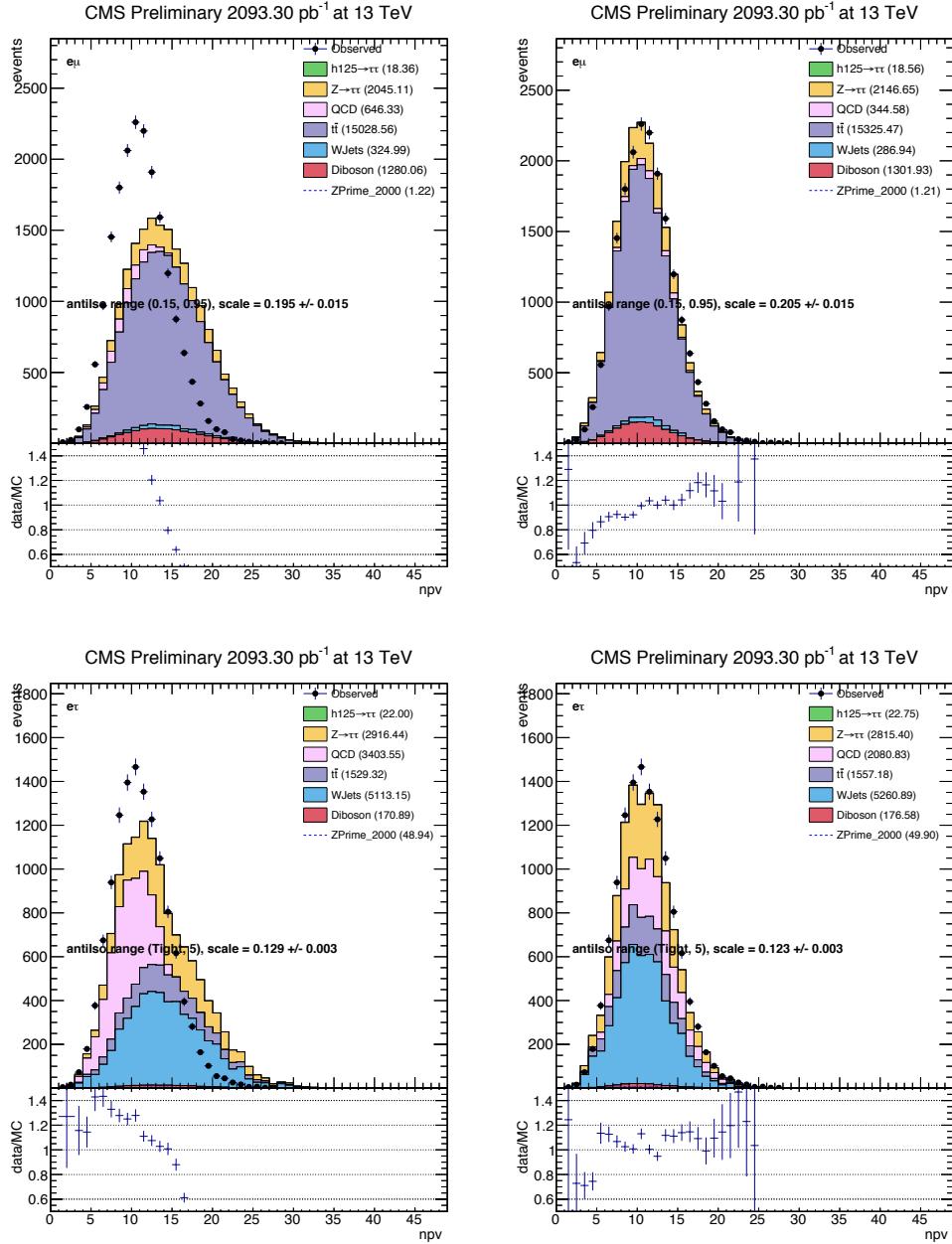


Figure 5.6: Distributions of the number of reconstructed primary vertices. Top: $e\mu$ channel. Bottom: $e\tau_h$ channel. Left: before reweighting. Right: after reweighting. The apparent substantial differences in the QCD “integrals” are artifacts of the plotting procedure, in which bins with negative contributions are zero-ed.

samples (see Section 5.6. For the $e\mu$ final state, any trigger with an isolation requirement on the muon, either an $e-\mu$ cross-trigger or single-muon trigger, would eliminate the isolation sideband for QCD estimation. Hence, we use the same single-electron trigger in $e\mu$ as for the $e\tau_h$ final state. Trigger paths are summarized in Table 5.7. Note these triggers constrain the object ID and phase space that we are studying. For example, the use of these triggers requires p_T cuts of 30 GeV ($\mu\tau_h$), 35 GeV ($e\tau_h$), and 60 GeV ($\tau_h\tau_h$) for our leading light leptons and τ_h 's across the different channels.

Table 5.7: The trigger paths used to collect the data. Emulated trigger paths, in particular those most similar to the paths used to collect the data, are applied to the simulated samples.

Channel	Trigger Path
$\tau_h\tau_h$ (data RunC, sim)	HLT_DoubleMediumIsoPFTau40_Trk1_eta2p1_Reg
$\tau_h\tau_h$ (data RunD)	HLT_DoubleMediumIsoPFTau35_Trk1_eta2p1_Reg
$\mu\tau_h$ (data RunC)	HLT_IsoMu24_eta2p1
$\mu\tau_h$ (data RunD)	HLT_IsoMu18
$\mu\tau_h$ (sim)	HLT_IsoMu17 (L1 $\mu p_T > 18$ GeV)
$e\tau_h$, $e\mu$ (data)	HLT_Ele27_eta2p1_WPLoose
$e\tau_h$, $e\mu$ (sim)	HLT_Ele27_eta2p1_WP75

5.4.1 Single Lepton Trigger Efficiency

The single electron and muon trigger efficiencies are measured using a tag and probe method, where we select events with at least one electron or muon pair. Each pair is required to satisfy a series of selection requirements:

For electrons:

- $p_T > 13$ GeV, $|\eta| < 2.1$, isolation < 0.15 , $d_{xy} < 0.045$ cm, $d_z < 0.2$ cm
- passing electron ID (we performed different measurements for “MVANonTrigWP80”, “HEEP”, “MVATrigWP80”)
- no matched conversions from photons
- number of missing hits in the silicon tracker = 0

For muons:

- $p_T > 20$ GeV, $|\eta| < 2.1$, isolation < 0.15 , $d_{xy} < 0.045$ cm, $d_z < 0.2$ cm
- passing “Medium” muon ID

For the pair:

- $\Delta R(e\mu) > 0.3$
- choose the most isolated pair in the event
- choose the opposite sign pair
- 3rd lepton veto (events must have fewer than three leptons)

5.4.2 Single Electron Trigger Efficiency

To measure the single-electron trigger efficiency, after the preselection mentioned above, we select (tag) events with a single-muon trigger (HLT_IsoMu18 for data and HLT_IsoMu17_eta2p1 for MC with L1 $\mu p_T > 18$ GeV) with the offline muon matching the HLT muon that fired the trigger. Then, the electron trigger efficiency is defined at the fraction of events which also pass (probe) the single-electron trigger (HLT_Ele27_eta2p1_WP Loose for data and HLT_Ele27_eta2p1_WP75 for MC).

The efficiency curves of the single-electron triggers, measured vs. electron p_T , are shown in Figure 5.7. A $p_T > 35$ GeV cut is motivated to avoid the turn-on region. As shown in Figure 5.8, the efficiency curves were further binned into two p_T bins containing similar numbers of events. A weight of 0.94 is measured, and used in the analysis, as the ratio between simulated and observed events in which the selected electron is in the endcap region.

5.4.3 Single Muon Trigger Efficiency

To measure the single-muon trigger efficiency, after the preselection mentioned above, we select (tag) events with a single-electron trigger (HLT_Ele27_eta2p1_WP Loose for data and HLT_Ele27

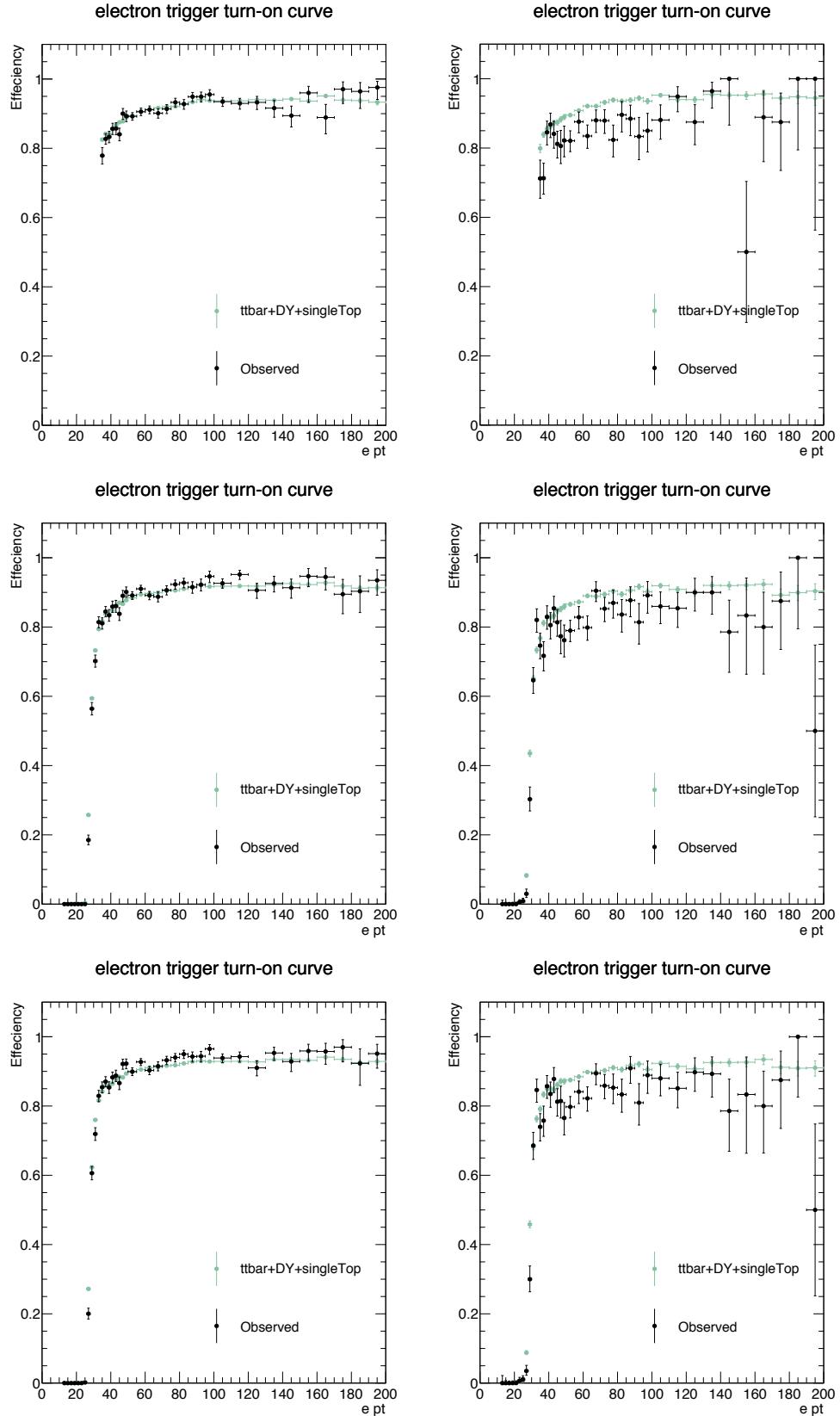


Figure 5.7: The efficiency vs. p_T curves of the single-electron triggers used. Left column: barrel. Right column: endcap. The offline electron ID requirements used are HEEP (top row), MVANonTrig80 (middle row), MVATrigWP80 (bottom row).

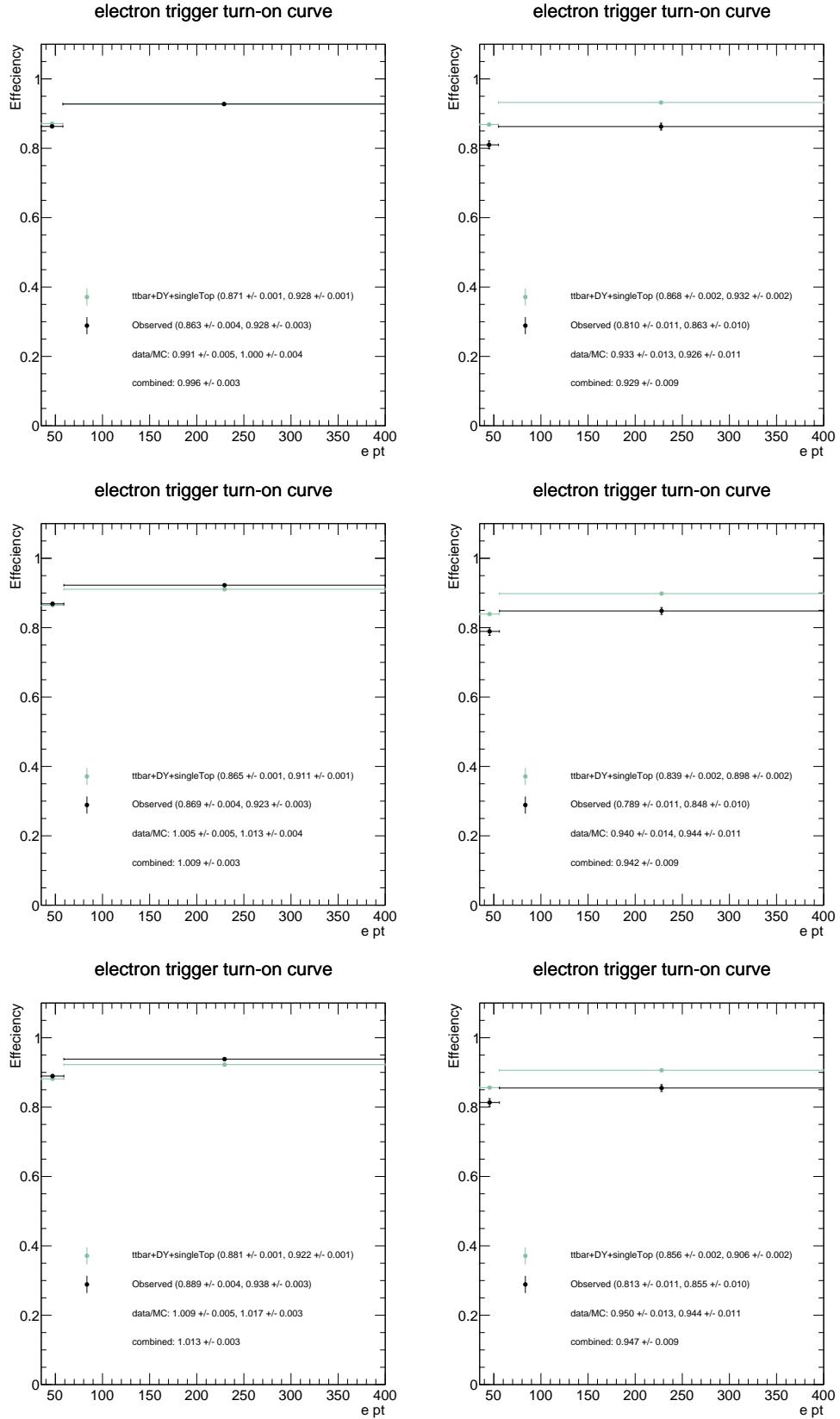


Figure 5.8: The efficiency vs. p_T curves of the single-electron triggers used. Left column: barrel. Right column: endcap. The offline electron ID requirements used are HEEP (top row), MVANonTrig80 (middle row), MVATrigWP80 (bottom row). The two bins each contain half of the events with $p_T > 35$ GeV.

`.eta2p1_WP75` for MC) with the offline electron match the HLT electron that fired the trigger. Then, the trigger efficiency is defined as the fraction of events which also pass (probe) the single-muon trigger (`HLT_IsoMu18` for data and `HLT_IsoMu17_eta2p1` for MC with $L1 \mu p_T > 18$ GeV).

The efficiency curves of the single-muon triggers, measured vs. muon p_T , are shown in Figure 5.9. A $p_T > 30$ GeV cut is motivated to avoid the turn-on region. As shown in the lower two plots, the efficiency curves were further binned into two bins containing similar numbers of events. Weights of 0.991 and 0.986 are measured, and used in the analysis, as the ratios between simulated and observed events in which the selected muon has $|\eta| \leq 1.2$ or $|\eta| > 1.2$.

5.4.4 Di-Tau Trigger Efficiency

The efficiency of the $\tau_h \tau_h$ trigger is measured using $Z \rightarrow \tau\tau \rightarrow \mu\tau_h$ events. The τ_h candidates reconstructed in the selected $Z \rightarrow \tau\tau \rightarrow \mu\tau_h$ events are required to pass the same τ_h identification used for the final analysis and which will be described in more detail in the sections to follow. The “denominator” selections used to define the $Z \rightarrow \tau\tau \rightarrow \mu\tau_h$ control sample are based on POG-recommended ID criteria and are summarized below:

- Events must fire the `HLT_IsoMu18` trigger (`HLT_IsoMu17_eta2p1` with HLT p_T cut of 18 GeV for MC)
- ≥ 1 global μ with $|\eta| < 2.1, p_T > 19$ GeV
- “`isMediumMuon`”
- muon best track $d_z < 0.2$ cm , $d_{xy} < 0.045$ cm w.r.t. PV
- Relative isolation (with $\delta\beta$ corrections) < 0.1
- ≥ 1 HPS τ_h with $|\eta| < 2.1, p_T > 20$ GeV
- Muon veto: “`againstMuonTight3`”
- Electron veto: “`againstElectronVLooseMVA5`”

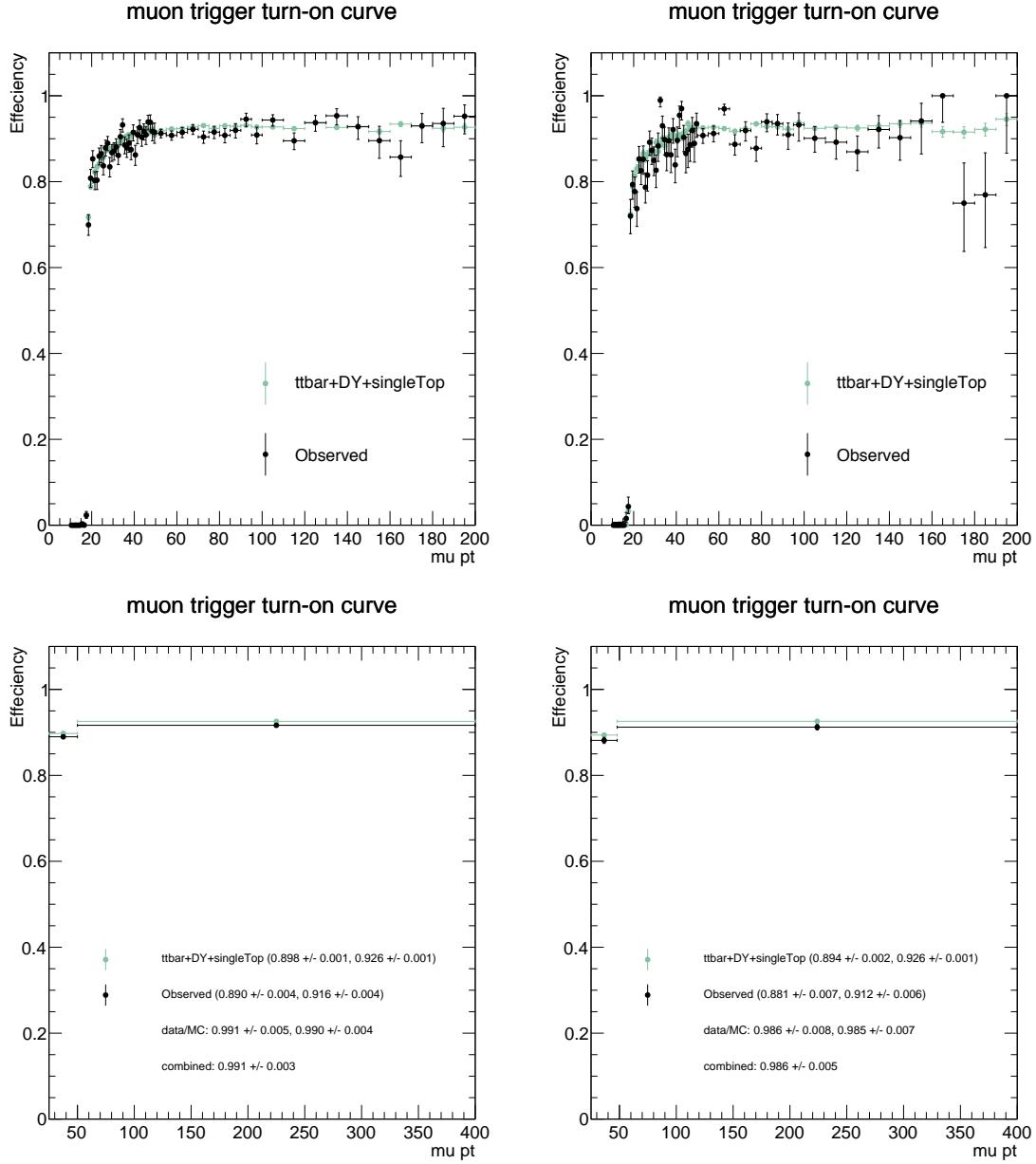


Figure 5.9: The efficiency vs. p_T curves of the single-muon triggers used. Left column: $|\eta| < 1.2$. Right column: $|\eta| > 1.2$. Bottom row: the two bins each contain half of the events with $p_T > 25$ GeV.

- “old” decay mode finding (1 or 3 charged hadrons)
- Isolation: “byTightCombinedIsolationDeltaBetaCorr3Hits”
- $p_T > 5$ GeV for leading Track of tau with $d_z < 0.2$ cm and $d_{xy} < 0.045$ cm w.r.t. PV
- $\Delta R(\mu, \tau_h) > 0.5$
- $Q(\mu) \times Q(\tau_h) < 0$
- $m_T(\mu, E_T) < 40$ GeV
- 0 jets tagged as b-jets
- 0 tagged electrons
- veto events with a second muon forming opposite-sign dimuon pair

The numerator is defined by additionally requiring those events to pass the HLT_IsoMu17_etaP2p1_MediumIsoPFTau40 trigger. The efficiency is measured for each leg individually and parameterized as function of p_T (Figure 5.10). From the plot it is clear that the emulation of the trigger in simulation does not provide the correct trigger efficiency observed in data. This is mainly due to a difference in the trigger definition in MC, specifically the L1 seed. Therefore, we do not apply the trigger in MC, but instead model the correct per leg trigger efficiency observed in data by weighing the predictions from simulation using the fit of the trigger efficiency curve in data (black solid curve in the plot). The trigger efficiency at the plateau is approximately 90%.

The trigger efficiency for data and all backgrounds is obtained by requiring $Z \rightarrow \tau_\mu \tau_h$ using the preselection cuts described above. Most of the backgrounds are taken from simulation except QCD whose shape and rate have been taken from same sign control region from data by subtracting the contribution of other backgrounds. The QCD normalization is then corrected by opposite-sign-to-like-sign (OS/LS) ratio which is taken as ~ 1.05 . Table 5.8 shows the event yield in this control region requiring the denominator selections. The “purity” is the fraction of DY+Jets events out of

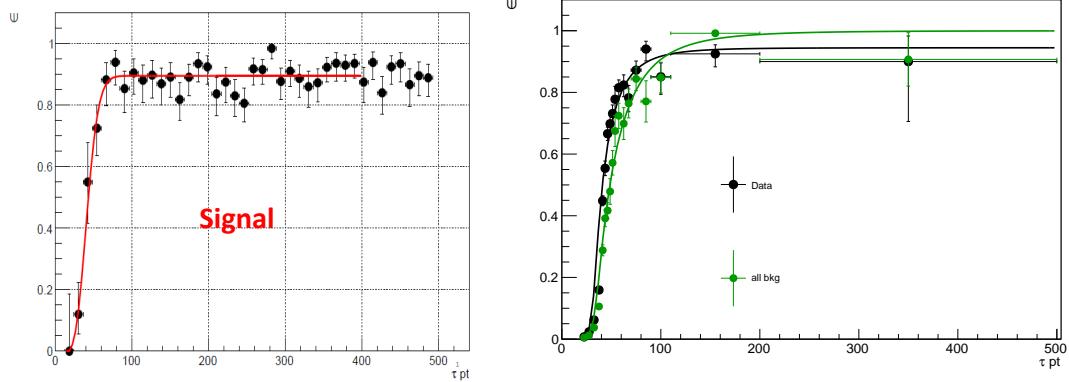


Figure 5.10: The per-leg τ_h trigger efficiency as a function of $p_T(\tau_h)$ for simulated Z' signal sample (left) and for data and all backgrounds obtained by requiring mu-tau region (right).

all backgrounds, and comes out to be $\sim 65\text{-}66\%$. The efficiency curves are fitted with crystal-ball (modified Gaussian) function and are shown in Figure 5.10.

Table 5.8: The number of events in the mu-tau control region after requiring the denominator level selection criteria

Sample	Events
Data	19578
$t\bar{t}$	171 ± 13
W+Jets	2128 ± 46
Z+Jets	12219 ± 111
QCD	4004 ± 63

5.5 Muon + Hadronic Tau Channel

The motivation for analyzing events where one τ lepton decays to a muon, while the other decays to hadrons is the same for all $\tau^+\tau^-$ -related analyses: because muons have the lowest jet misidentification among leptons, the mere requirement of a muon removes a substantial amount of background processes, especially the QCD multijet background. Once this requirement is made, a main source of background is due to the Drell-Yan (DY) process giving rise to τ leptons. Because we seek particles with masses much larger than that of the Z boson, this source of background can be easily discriminated against by looking at larger reconstructed $\tau^+\tau^-$ -mass regions (reconstructed

mass to be defined later). This process can also serve as a “standard candle” to validate the τ_h identification criteria and ensure the robustness of the analysis. Other main sources of background include (1) QCD multijet events where light-quark or gluon jets are misidentified as a τ_h , (2) W + Jets events where the W boson decays to a muon and a jet is misidentified as a τ_h , and (3) $t\bar{t}$ events where two leptons can come from the prompt decay of W bosons or one misidentified τ_h from a jet. The sum of the Drell-Yan and W + Jets backgrounds represent 89% of the total background (according to simulation) in this channel. The cuts used to select $\mu\tau_h$ pairs are factorized in to four categories: acceptance, μ identification, τ_h identification, and topological cuts. The acceptance criteria are completely driven by the limitations of the CMS detector and the need to maximize analysis sensitivity while also minimizing systematic effects. For example, to minimize systematic effects, the p_T threshold on the muon leg is chosen such that it falls on the plateau of the trigger turn-on curve ($p_T > 30$ GeV, see Section 5.4). Although it is possible to achieve slightly better sensitivity by increasing the thresholds, the selection criteria are also driven by the need to obtain a sample enriched with $Z \rightarrow \tau\tau$ events, with minimal modifications to the final selection criteria. Figure 5.11 shows the p_T distributions for signal and background processes relevant to this analysis. As discussed in Section 5.3, the muon identification criteria are designed mostly to discriminate against cosmic muons, punch-through pions, and muons associated to jets from decays in flight. The τ_h identification is described in Section 5.4 and is mostly designed to discriminate against hadronic jets produced from the fragmentation of quarks and/or gluons. Finally, topological cuts are principally used to minimize the remaining W + jet(s) and $t\bar{t}$ contributions after the muon and τ_h identification criteria have been imposed.

The production of W bosons in association with jets becomes a dominant background because a clean, well-reconstructed muon can be produced by the decay of the W boson. Therefore, the requirement of a clean muon signature does not provide additional discrimination. Additionally, the neutrino from the W boson decay acquires an average energy of approximately $m_W/2 \sim 40$ GeV. Because the neutrino will escape the CMS detector undetected, the average measurement of the momentum imbalance will be approximately 40 GeV. Therefore, our requirement on the momentum

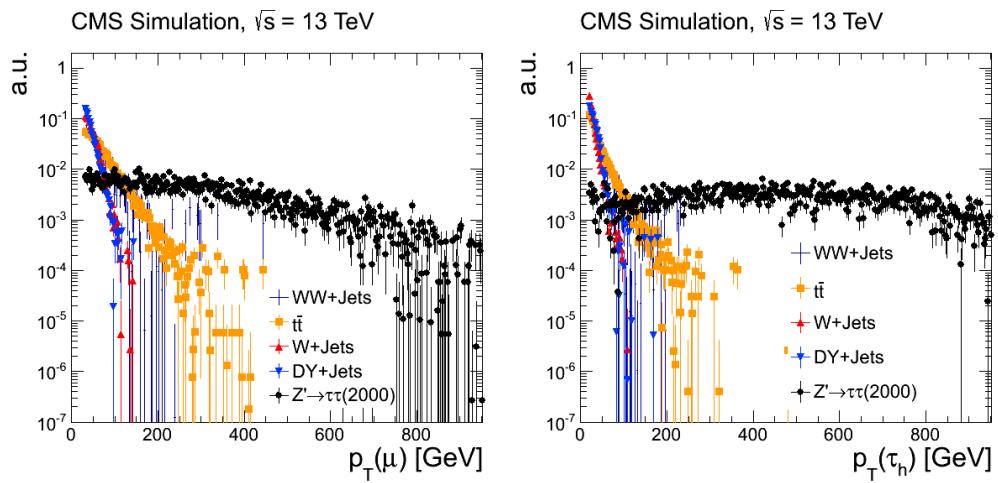


Figure 5.11: The distributions of $p_T(\mu)$ (left) and $p_T(\tau_h)$ (right) after acceptance, μ identification, and τ_h identification cuts.

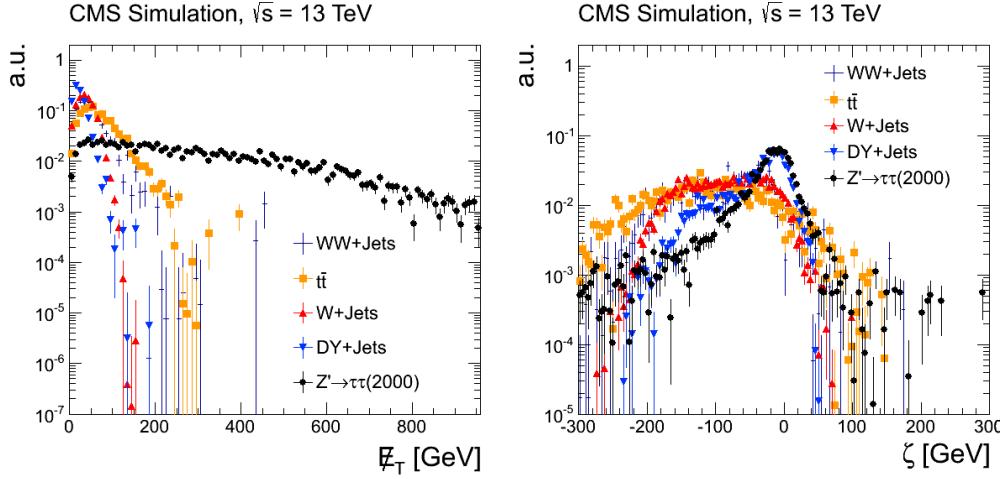


Figure 5.12: Left: \cancel{E}_T after acceptance, μ identification, and τ_h identification cuts. Right: $P_\zeta - 3.1 \times P_\zeta^{\text{vis}}$ after acceptance, μ identification, and τ_h identification cuts.

imbalance in the event ($\cancel{E}_T > 30$ GeV) does not provide significant discrimination against this background. Figure 5.12 (left) shows the \cancel{E}_T distribution for signal and background processes relevant to this analysis. Finally, the presence of “associated jets” means that the contamination from $W + \text{Jets}$ in the signal region is highly dependent on the jet $\rightarrow \tau_h$ fake rate, which is the largest among leptons. Therefore, reducing the $W + \text{Jets}$ background also requires additional topological requirements.

In $W + \text{jet(s)}$ events, unlike $X \rightarrow \tau\tau$ resonance production such as $Z' \rightarrow \tau\tau$, where the τ -lepton decay products are expected to be back-to-back in ϕ , the presence of the neutrino from the W decay and the uncorrelated jet gives rise to topologies where the jet and the lepton are not back-to-back (Figure 5.13). Therefore, one of the main discriminating variables against $W + \text{Jets}$ events is the difference in ϕ between the jet and muon directions. Figure 5.16 shows the $\cos\Delta\phi(\mu, \tau_h/\text{jet})$ distributions for $Z' \rightarrow \tau\tau$ and the relevant background samples. We require

$$\cos\Delta\phi(\mu, \tau_h/\text{jet}) < -0.95 \quad (5.3)$$

which is $> 90\%$ efficient for $Z' \rightarrow \tau\tau$ and $\approx 25\%$ efficient for $W+\text{Jets}$.

For $\tau^+\tau^-$ final states, the \cancel{E}_T in the event is due to the neutrinos from the τ -lepton decays and is expected to point in the direction collinear to the visible tau decay products. Furthermore,

Cross-sectional view of the detector (x-y plane)

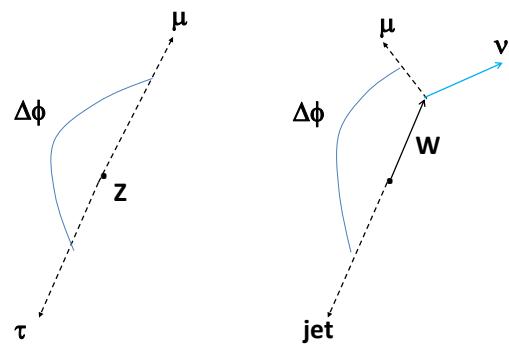


Figure 5.13: Sketch depicting the $W+J$ ets rejection power of a $\Delta\phi$ cut.

the measurement of \cancel{E}_T is completely correlated to the visible tau decay products. In $W + \text{jet(s)}$ events, the direction and magnitude of the momentum imbalance is completely correlated to the lepton from the W boson, but uncorrelated to the jet. We require events to be consistent with this signature of a particle decaying to two τ leptons by defining a unit vector along the bisector of visible tau decay products ($\hat{\zeta}$) and two projection variables, p_ζ and p_ζ^{vis} :

$$p_\zeta^{vis} = \vec{p}_{\tau_1}^{vis} \hat{\zeta} + \vec{p}_{\tau_2}^{vis} \hat{\zeta} \quad (5.4)$$

$$p_\zeta = p_\zeta^{vis} + \cancel{E}_T \hat{\zeta} \quad (5.5)$$

The sketch in Figure 5.14 displays the definitions for p_ζ^{vis} and p_ζ . Figure 5.15 shows the separation between $Z' \rightarrow \tau\tau$ and $W + \text{jet(s)}$ events in the p_ζ - p_ζ^{vis} plane. For the case of $W+\text{Jets}$, there is no strong correlation between p_ζ^{vis} and p_ζ due to the presence of a jet that is uncorrelated to the μ and ν_μ from the W boson. However, there is a strong correlation for the case of $Z' \rightarrow \tau\tau$. To discriminate against $W+\text{jet(s)}$ events, requirements on $\Delta\phi(\tau_1, \tau_2)$ and ζ are applied. ζ is defined as a linear combination of p_ζ and p_ζ^{vis} :

- $\cos \Delta\phi(\tau_1, \tau_2) < -0.95$
- $p_\zeta - 3.1p_\zeta^{vis} > -50$

[htbp]

Although events containing $t\bar{t}$ contribute to the expected background in all channels containing light leptons, its contribution to the $\mu\tau_h$ channel is only $\sim 1.4\%$ of the total background (according to simulation). For $\mu\tau_h$ final states the $t\bar{t}$ contribution comes in the form of a real light lepton from the semileptonic decay of the W^\pm and a fake τ_h from the hadronic decay of the second W^\pm . These events are characterized by an isolated light lepton, passing all lepton identification and isolation requirements, accompanied by a non-isolated “hadronic tau” due to the larger multiplicity of the hadronically decaying W boson. These events are suppressed with the use of topological cuts.

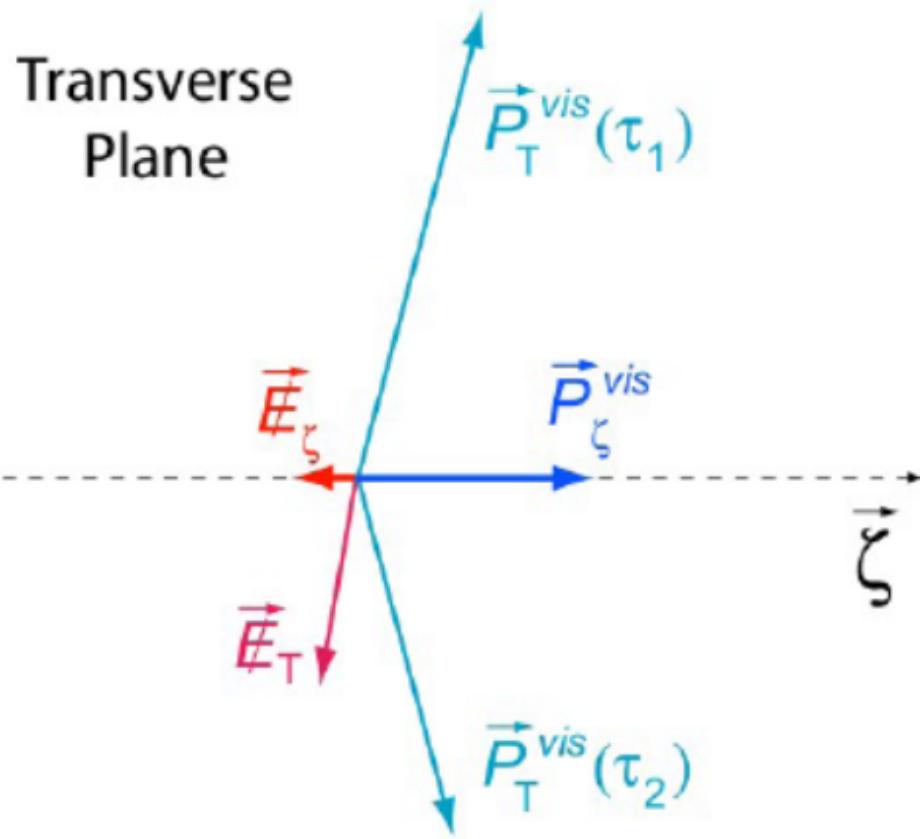


Figure 5.14: Definitions for p_ζ and p_ζ^{vis} .

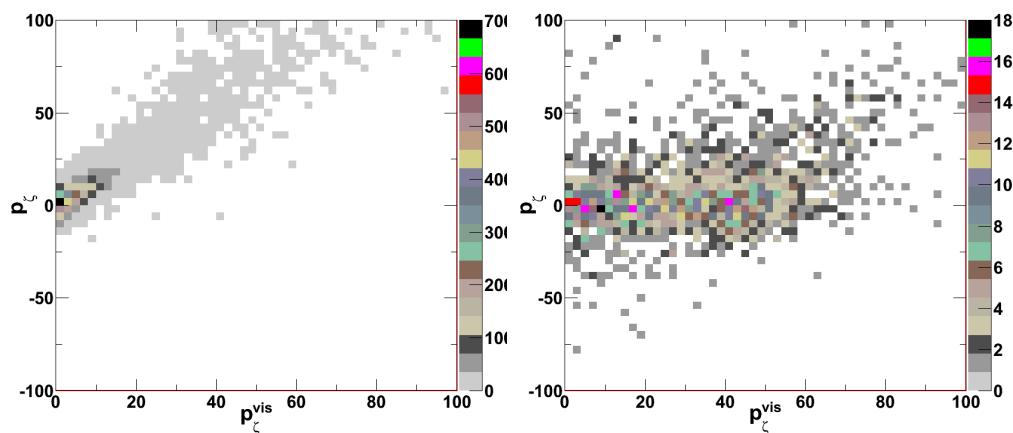


Figure 5.15: p_ζ vs. p_ζ^{vis} for $Z' \rightarrow \tau\tau$ (left) and for W+Jets (right).

After applying the lepton identification and isolation requirements, a non-negligible background contribution from $t\bar{t}$ events remain. These events can be further suppressed with cuts that take advantage of the very different topologies between $Z' \rightarrow \tau\tau$ and $t\bar{t}$ events. The first, and most important, of these differences is the presence of b-jets in the event. For our purposes, jets with $p_T > 30$ GeV and $|\eta| < 2.4$ are counted as b-tagged jets if the “combined secondary vertex” discriminator returns a value consistent with that of a b-jet.

The entire signal selection criteria are summarized below, while Table 5.10 is the cut flow efficiency table (yields normalized to lumi).

Acceptance Selection:

- Events must fire the HLT_IsoMu18 trigger
- ≥ 1 global μ with $|\eta| < 2.1, p_T > 30$ GeV
- ≥ 1 HPS τ_h with $|\eta| < 2.1, p_T > 20$ GeV
- $\Delta R(\mu, \tau_h) > 0.5$

μ Identification:

- “isMediumMuon”
- Relative isolation (with $\delta\beta$ corrections) < 0.15

τ_h Identification:

- Muon veto: “againstMuonTight3”
- Electron veto: “againstElectronMVA5LooseMVA5”
- “old” decay mode finding with 1 or 3 signal charged hadrons
- Isolation: “byTightCombinedIsolationDeltaBetaCorr3Hits”

Topological requirements:

- $\cos \Delta\phi(\mu, \tau_h) < -0.95$
- $Q(\mu) \times Q(\tau_h) < 0$
- $\cancel{E}_T > 30 \text{ GeV}$
- $P_\zeta - 3.1 \times P_\zeta^{vis} > -50$
- 0 jets tagged as b-jets

Because τ leptons decay to neutrinos, which leave the detector undetected, one cannot fully reconstruct the mass resonance with the visible τ decay products. Additionally, because the invariant mass for background processes such as QCD are typically steeply falling distributions in the tails (where new mass resonances are expected), it becomes important to make use of \cancel{E}_T to attempt to separate signal from background and reconstruct the true mass resonance. Historically, several methods such as the collinear approximation have been employed to reconstruct the true mass resonance. However, for the analysis presented, one of the main sources of backgrounds is $W + \text{jet(s)}$. In this case, the analysis achieves a high degree of sensitivity by requiring the μ and τ_h candidates to be back-to-back in ϕ . This is precisely the regime in which the collinear approximation fails. Therefore, the mass is reconstructed as follows:

$$m(\tau_1, \tau_2, \cancel{E}_T) = \sqrt{(E_{\tau_1} + E_{\tau_2} + \cancel{E}_T)^2 - (\vec{p}_{\tau_1} + \vec{p}_{\tau_2} + \vec{\cancel{E}}_T)^2} \quad (5.6)$$

where E_τ and \vec{p}_τ represent the energy and 4-momentum of the visible objects (in this case the μ and τ_h), respectively. As can be seen from Figure 5.16, this definition successfully distinguishes between lower mass production of τ -lepton pairs and high-mass τ -lepton pairs from new massive resonant particle production.

Figure 5.17 shows the $p_T(\tau_h)$, $\eta(\tau_h)$, \cancel{E}_T , and $m_T(\mu, \cancel{E}_T)$ (transverse mass) control plots using events satisfying the “preselection” requirements (i.e. acceptance + μ identification + τ_h

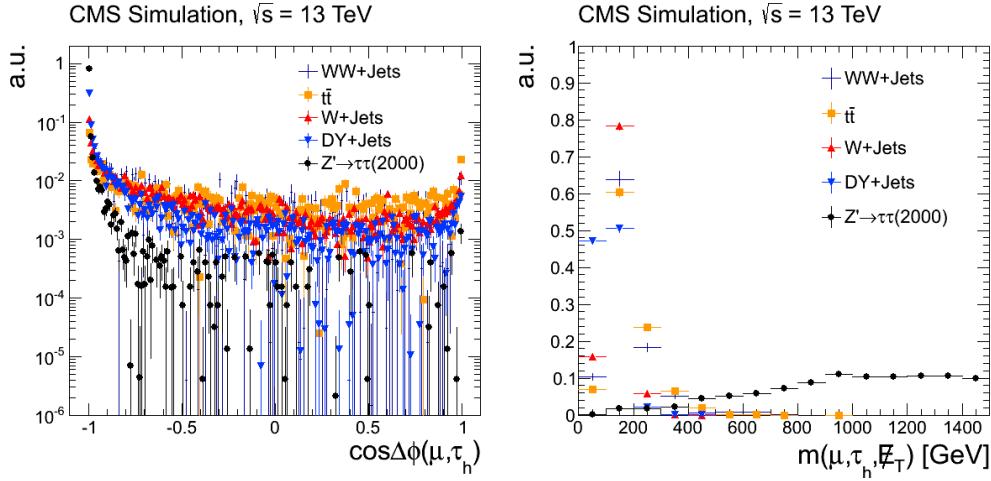


Figure 5.16: Left: $\cos\Delta\phi$ between the muon and tau/jet directions. Right: $m(\mu, \tau_h, \cancel{E}_T)$ for signal Z' and various backgrounds. Both plots normalized to unity.

Table 5.9: Event summary table after signal region selection

Process	$\tau_h \tau_h$	$\mu \tau_h$	$e \tau_h$	$e \mu$
Z' (500)	307.4 ± 35.3	502.3 ± 57.7	197.6 ± 22.7	218.6 ± 27.3
Z' (1000)	34.6 ± 2.6	40.8 ± 3.1	14.7 ± 1.1	19.0 ± 1.5
Z' (1500)	6.6 ± 0.3	7.2 ± 0.3	2.3 ± 0.1	3.6 ± 0.2
Z' (2000)	1.6 ± 0.07	1.8 ± 0.08	0.59 ± 0.03	0.91 ± 0.04
Z' (2500)	0.55 ± 0.02	0.60 ± 0.02	0.19 ± 0.01	0.30 ± 0.01
Z' (3000)	0.13 ± 0.01	0.14 ± 0.01	0.04 ± 0.00	0.07 ± 0.00
Drell-Yan	8.4 ± 3.1	882.4 ± 127.0	375.1 ± 117.6	321.2 ± 99.2
W+Jets	0.1 ± 0.1	916.2 ± 96.1	545.8 ± 85.6	18.9 ± 11.4
Diboson	0.5 ± 0.5	29.2 ± 7.4	18.0 ± 4.4	108.3 ± 17.4
$t\bar{t}$	—	26.1 ± 6.7	26.1 ± 7.5	222.8 ± 44.8
Multijet	48.7 ± 13.0	121.8 ± 83.5	116.7 ± 71.5	31.9 ± 24.3
Observation	55	1807	1113	728

Table 5.10: Signal and background yields after various stages of the $\mu\tau_h$ selection.

Process	Trigger: HLT_IsoMu18	1μ	$1\tau_h$	b-jet Veto	$E_T > 30 \text{ GeV}$	$\mu\tau_h$ topology cuts
Data	47386261	15478176	61718	53254	28131	1807
WW	24673 ± 77	20330 ± 70	432 ± 10	372 ± 10	257 ± 8	24 ± 2
WZ	7650 ± 28	6515 ± 26	124 ± 4	100 ± 3	64 ± 3	5 ± 1
ZZ	1826 ± 8	1643 ± 8	39 ± 1	32 ± 1	14 ± 1	1 ± 0.1
$t\bar{t}$	231246 ± 189	193283 ± 173	4369 ± 26	607 ± 10	511 ± 9	26 ± 2
W+Jets	16090363 ± 5271	11990572 ± 4538	30234 ± 215	28231 ± 211	18079 ± 168	758 ± 35
DY+Jets	2668493 ± 2584	2140959 ± 2086	12873 ± 181	11907 ± 175	3611 ± 99	883 ± 35
QCD	6296742 ± 17403	1058694 ± 7136	3992 ± 438	3030 ± 381	866 ± 204	143 ± 90
Total BG	25320993 ± 18368	15411993 ± 8712	52063 ± 522	44278 ± 470	23402 ± 282	1839 ± 103
$Z' \rightarrow \tau\tau(500)$	3521 ± 22	3220 ± 21	961 ± 11	869 ± 11	716 ± 10	502 ± 8
$Z' \rightarrow \tau\tau(1000)$	214 ± 1	202 ± 1	73 ± 1	65 ± 1	59 ± 1	41 ± 1
$Z' \rightarrow \tau\tau(1500)$	36 ± 0.1	34 ± 0.1	13 ± 0.1	12 ± 0.1	11 ± 0.1	7 ± 0.07
$Z' \rightarrow \tau\tau(2000)$	8.7 ± 0.05	8.2 ± 0.04	3.3 ± 0.03	2.9 ± 0.03	2.8 ± 0.03	1.8 ± 0.02
$Z' \rightarrow \tau\tau(2500)$	2.8 ± 0.01	2.6 ± 0.01	1.1 ± 0.01	0.9 ± 0.008	0.9 ± 0.008	0.6 ± 0.006
$Z' \rightarrow \tau\tau(3000)$	0.6 ± 0.003	0.6 ± 0.003	0.3 ± 0.002	0.2 ± 0.002	0.2 ± 0.002	0.1 ± 0.001

identification cuts described above). Figure 5.18 shows the $p_T(\tau_h)$, $\eta(\tau_h)$, E_T , and $m_T(\mu, E_T)$ control plots in the signal region. In these particular set of plots, the background predictions are entirely based on MC. In general, there is good agreement between the observed distributions and the predictions from simulation.

The Standard Model processes considered as backgrounds are Drell-Yan, diboson production, top quark single and pair production, W+Jets production, and QCD multijet production.

5.5.1 W + Jets Background Estimation

As discussed above, the main discriminators against W + Jets events are the topological variables ζ and $\cos\Delta\phi(\mu, \tau_h)$ in addition to τ_h isolation. Thus, it is natural to construct the W + Jets background estimation methodology using control samples obtained by inverting these requirements. We estimate this background using a data-driven approach which relies on the ABCD method. The regions ABCD are defined as follows:

- A: pass both the ζ and $\cos\Delta\phi$ cuts; pass “Tight” τ_h isolation (signal region)
- B: fail one or both of the ζ and $\cos\Delta\phi$ cuts; pass “Tight” τ_h isolation
- C: pass both the ζ and $\cos\Delta\phi$ cuts; fail “Tight” but satisfy a relaxed τ_h isolation of < 5 GeV (“loose”)
- D: fail one or both of the ζ and $\cos\Delta\phi$ cuts; fail “Tight” but satisfy a relaxed τ_h isolation of < 5 GeV (“loose”)

We estimate the W + Jets component N_W^i in regions $i = B, C, D$ by subtracting MC non-W backgrounds from data ($N_W^i = N_{\text{Data}}^i - N_{\neq W}^i$). We then estimate the W + Jets component in the signal region A, to be $N_W^A = N_W^C \cdot \frac{N_W^B}{N_W^D}$. Said differently, we take the yield of W + Jets events in data containing a non-isolated τ_h and then extrapolate to the signal region by correcting for the “Tight-to-Loose” ratio (also referred to as a “fake rate”), which is measured in a data sample enriched with W + Jets events and that is obtained by inverting the ζ and $\cos\Delta\phi$ cuts. The shape of the

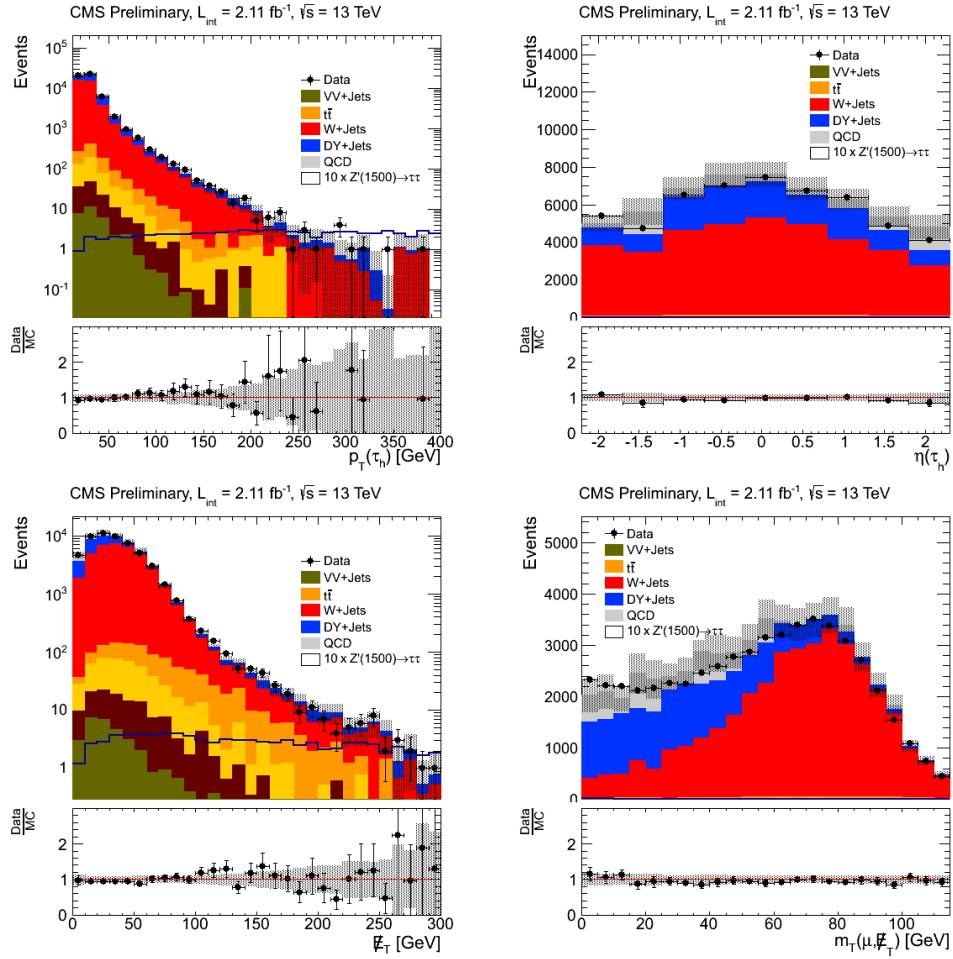


Figure 5.17: $p_T(\tau_h)$, $\eta(\tau_h)$, \cancel{E}_T , and $m_T(\mu, \cancel{E}_T)$ control plots using events satisfying the “preselection” requirements. The background predictions are entirely based on MC. Agreement between data and MC is good.

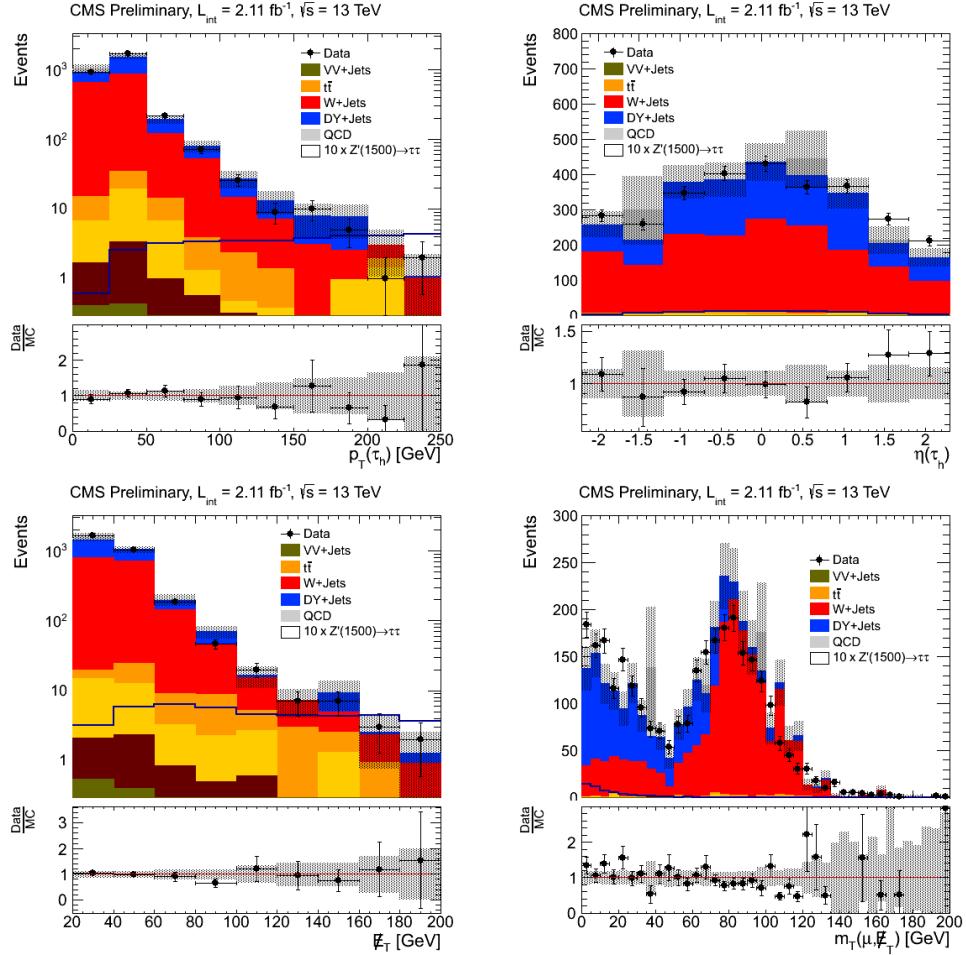


Figure 5.18: $p_T(\tau_h)$, $\eta(\tau_h)$, \cancel{E}_T , and $m_T(\mu, \cancel{E}_T)$ control plots using events satisfying the signal region requirements. The background predictions are entirely based on MC. Agreement between data and MC is good.

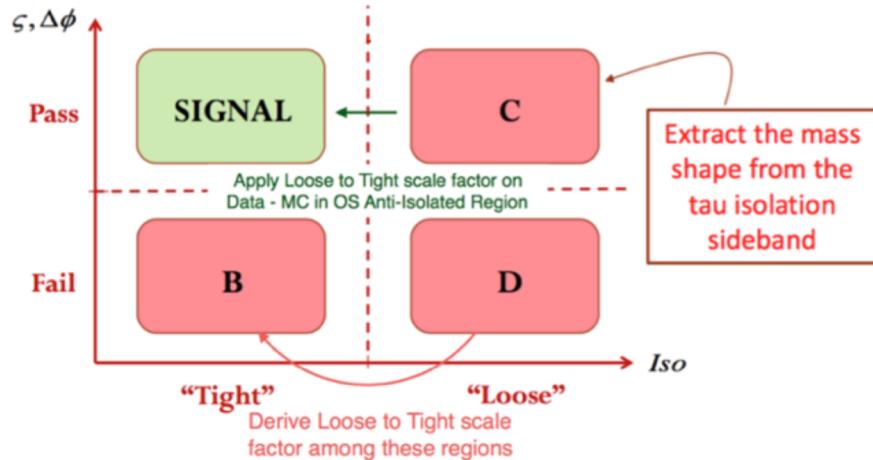


Figure 5.19: The four regions used in the “ABCD” method of estimating W+Jets background

$m(\mu, \tau_h, \cancel{E}_T)$ distribution is obtained from control region C (nominal selections with non-isolated τ_h).

Figure 5.20 shows a comparison of the ζ and $\cos\Delta\phi$ distributions in W + Jets MC, normalized to unity, under two different τ_h isolation conditions: (i) isolated (passing “Tight”), and (ii) non-isolated (failing “Tight” but passing loosened isolation). The τ_h isolation does not bias the ζ and $\cos\Delta\phi$ distributions, and thus we expect our ABCD background estimation method to accurately model the “true” W + jets yield in the signal region. A closure test (procedure for validating the estimation method) of the background estimation method outlined above is performed with MC. Two aspects are simultaneously tested: (1) closure on the normalization (i.e. $N_W^A = N_W^C \cdot \frac{N_W^B}{N_W^D}$); (2) correct determination of the $m(\mu, \tau_h, \cancel{E}_T)$ shape. Figure 5.21 shows the closure test in MC. The y-axis of the top left plot is in normal scale so the reader can focus on the level of agreement in the low mass region. The top right plot is similar, except the y-axis is in log scale in order to emphasize the high-mass tails. We observe very good agreement between the nominal yield/shape and the predicted yield/shape. The reader can perform his or her own cross-check of this conclusion by using the MC-based W + Jets yields in Table 5.11, which shows the data and MC background yields in the control samples, and plugging them into the equation $N_W^{\text{Prediction},MC} = N_W^{C,MC} \cdot \frac{N_W^{B,MC}}{N_W^{D,MC}}$. The MC-based prediction using the ABCD method is $N_W^{\text{Prediction},MC} = (1112 \pm 54) \cdot \frac{(5236 \pm 116)}{(7592 \pm 139)} = 767 \pm 43$, while the MC-based nominal yield is 758 ± 35 (see Table 5.10). Therefore, the MC-based nominal yield to predicted yield ratio is $\frac{N_W^{\text{Nominal},MC}}{N_W^{\text{Prediction},MC}} = \frac{758 \pm 35}{767 \pm 43} = 0.99 \pm 0.05$, which is statistically consistent with unity. Thus, no additional systematic uncertainties are applied due to closure.

As mentioned above, Table 5.11 shows the data and MC background yields in control regions B , C , and D . The purity of W + Jets, based on simulation, ranges from $\sim 65 - 81\%$, depending on the sample. The W + Jets scale factors, defined as $\text{SF} = \frac{N_W^{\text{Data}} - N_W^{MC}}{N_W^{MC}}$, show about a $\sim 20\%$ mismodeling in simulation. The top row of plots in Figure 5.22 are the $m_T(\mu, \cancel{E}_T)$ distributions in control regions C (top left) and A (top right). The previously mentioned scale factors have been applied to the plots.

A natural comment or question from the reader is that while the MC closure test looks good,

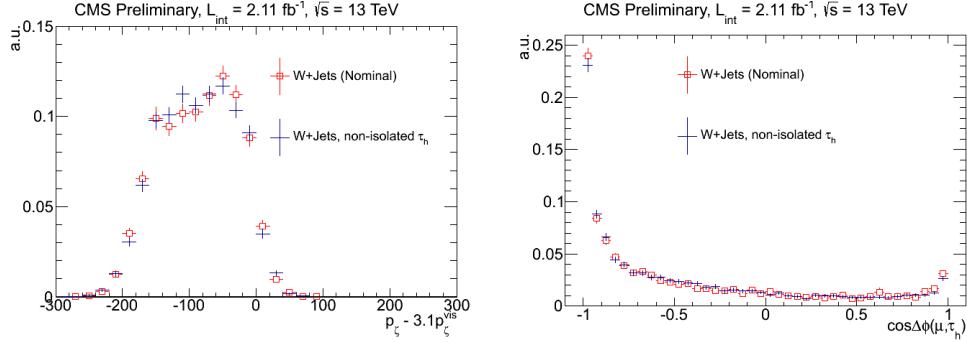


Figure 5.20: Left: Comparison of the ζ distribution in W + Jets MC, normalized to unity, for events with isolated and non-isolated τ_h . Right: Comparison of the $\cos\Delta\phi$ distribution in W + Jets MC, normalized to unity, for events with isolated and non-isolated τ_h .

Table 5.11: Background and data yields in W + Jets control regions B, C, D , under nominal \cancel{E}_T conditions ($\cancel{E}_T > 30 \text{ GeV}$).

Process	$\mu\tau_h$ W + Jets CR C	$\mu\tau_h$ W + Jets CR B	$\mu\tau_h$ W + Jets CR D
ZZ + Jets	9 ± 3	4 ± 2	0.7 ± 0.9
WZ + Jets	24 ± 5	18 ± 5	2 ± 1
WW + Jets	141 ± 13	86 ± 10	17 ± 5
QCD	189 ± 96	631 ± 177	141 ± 83
$t\bar{t}$	250 ± 17	190 ± 15	23 ± 5
Z + Jets	1278 ± 71	871 ± 47	426 ± 40
W + Jets	5236 ± 116	7592 ± 139	1112 ± 54
Data	8278	10434	1847
Purity	73.5%	80.8%	64.6%
Data - $\sum_{i \neq W} BG_i$	6388 ± 202	8633 ± 237	1238 ± 112
SF	1.22 ± 0.05	1.14 ± 0.04	1.11 ± 0.11

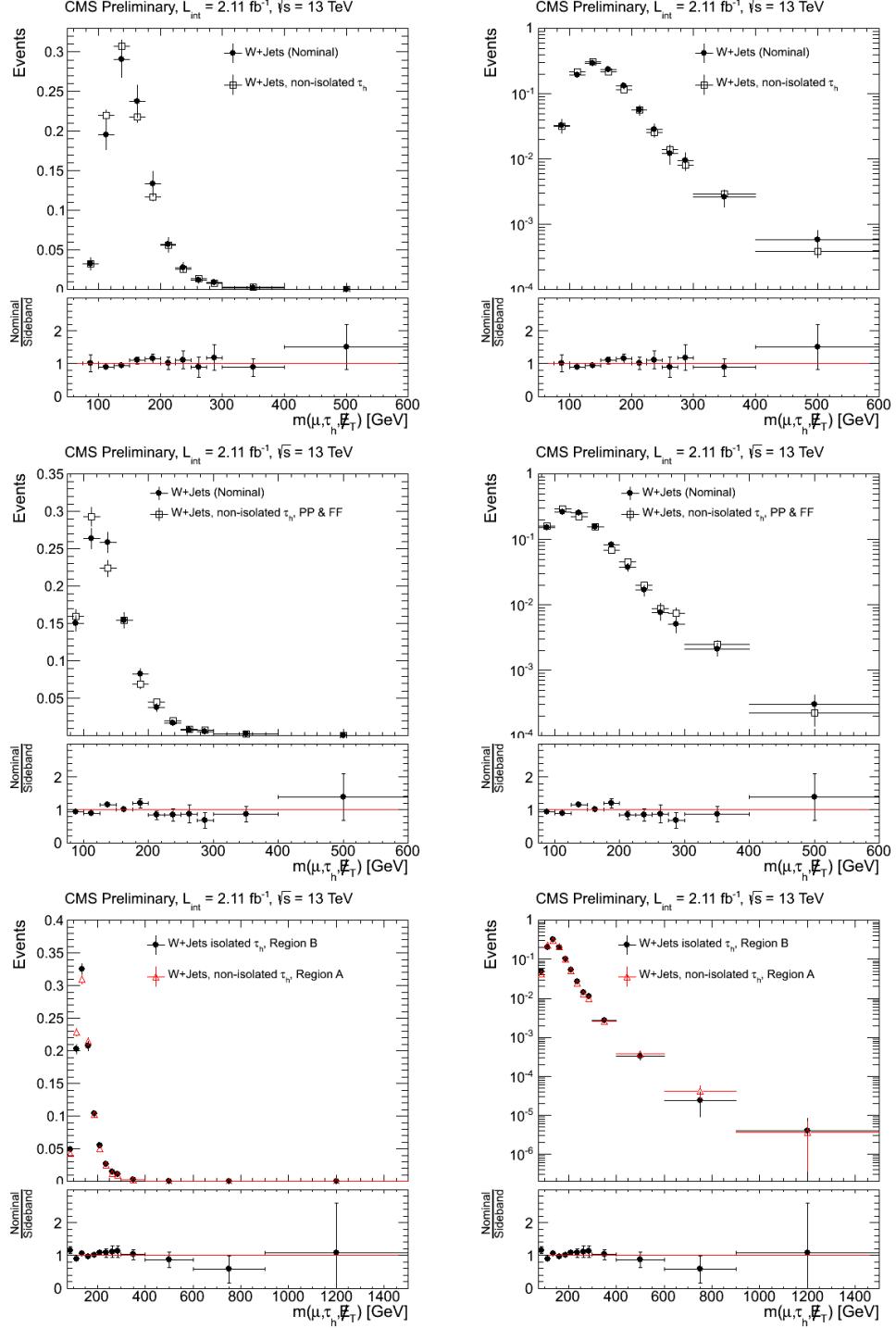


Figure 5.21: Top: W + Jets closure test performed with simulation (normal scale to focus on low mass) under nominal \cancel{E}_T conditions (i.e. $\cancel{E}_T > 30$ GeV). Middle: W + Jets closure test performed with simulation requiring $\cancel{E}_T > 0$ GeV. Bottom: Comparison of the reconstructed mass distributions in regions A and B. Left: normal scale to focus on low mass. Right: Log scale to focus on the high mass tails.

it is possible that mis-modeled \cancel{E}_T in simulation could perhaps pull the closure test in the wrong direction? In other words, how stable is the background estimation method with respect to \cancel{E}_T ? The middle row of plots in Figure 5.21 show similar closure tests, but without a \cancel{E}_T cut. We find that closure is observed even without the \cancel{E}_T cut, adding to the robustness of the methodology. Table 5.12 shows the data and MC background yields in control regions B , C , and D obtained without the \cancel{E}_T requirement. The purity of $W + \text{Jets}$, based on simulation, has decreased in comparison to the purity observed with $\cancel{E}_T > 30 GeV, due to the larger contribution from DY + Jets and QCD multijet.$

Table 5.12: Background and data yields in $W + \text{Jets}$ control regions B, C, D obtained without a \cancel{E}_T requirement.

Process	$\mu\tau_h$ W + Jets CR C	$\mu\tau_h$ W + Jets CR B	$\mu\tau_h$ W + Jets CR D
ZZ + Jets	16 ± 4	10 ± 3	2 ± 1
WZ + Jets	37 ± 6	26 ± 5	4 ± 2
WW + Jets	184 ± 15	114 ± 12	29 ± 6
$t\bar{t}$	284 ± 18	223 ± 16	30 ± 6
QCD	412 ± 142	1003 ± 222	442 ± 147
DY + Jets	3283 ± 101	2755 ± 85	1870 ± 84
W + Jets	7587 ± 139	10895 ± 167	2700 ± 85
Data	13896	17235	5984
Purity	64.3%	72.5%	53.2%
Data - $\sum_{i \neq W} BG_i$	9680 ± 394	13105 ± 401	3607 ± 268
SF	1.28 ± 0.06	1.20 ± 0.04	1.34 ± 0.11

The procedure outlined in this section yields a $W + \text{Jets}$ estimate of $N_W^{\text{Prediction,Data}} = (1238 \pm 112) \cdot \frac{(6388 \pm 202)}{(8633 \pm 237)} = 916 \pm 91$ (see Data - $\sum_{i \neq W} BG_i$ in Table 5.11). The uncertainty on the background prediction is $\sim 10\%$ and is mostly dominated by $\sqrt{N_{\text{Data}}}$ in control regions B , C , and D . We note that the main source of contamination in Table 5.11 is DY + Jets, which we've normalized using a data-to-MC correction factor described in the next section. However, even if DY + Jet(s) yields in the W control samples are taken entirely from simulation, it only provides a small effect on the $W + \text{Jets}$ prediction in the signal region. This is true because the systematic uncertainty for DY + Jets is typically $< 15\%$ (see Section 5.9), while the level of non- W background contamination

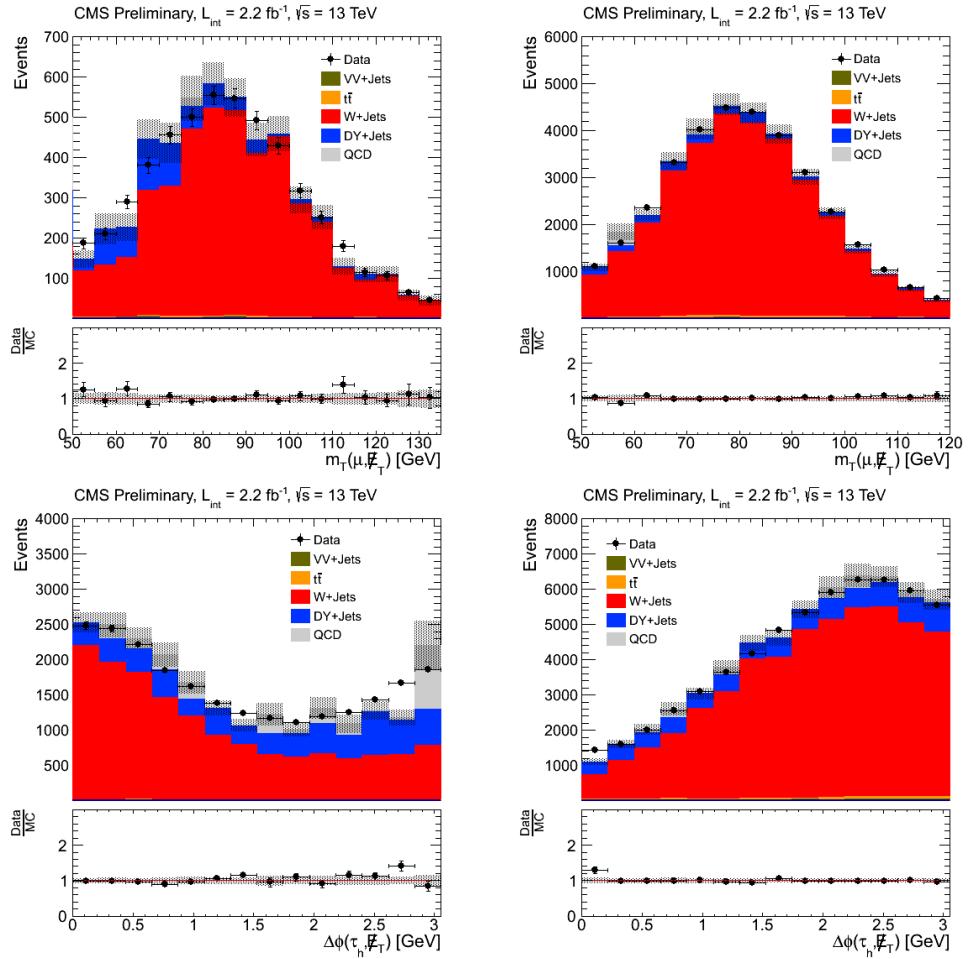


Figure 5.22: $m_T(\mu, \cancel{E}_T)$ distributions in W + Jets control region D (left) and B (right) with the nominal \cancel{E}_T cut of $\cancel{E}_T > 30$ GeV (top) and with no \cancel{E}_T cut (bottom).

is at most $\sim 35\%$ (i.e. upper bound on systematic effect is $0.35 \cdot 0.15 \sim 5\%$). The total predicted background $m(\mu, \tau_h, \cancel{E}_T)$ spectrum will be shown in Section 5.10.

5.5.2 DY + Jets Background Estimation

Obtaining a semi-clean sample of $Z \rightarrow \tau\tau$ events is one of the most important aspects of the analysis as it needs to be shown that the τ_h identification criteria employed is successful at identifying a τ_h . Unlike W + Jets, where a jet is misidentified as a τ_h and whose contribution to the signal region is expected to be mismodeled by simulation, there is no reason to expect significant disagreement between data and MC for $Z \rightarrow \tau\tau$. The efficiency for the requirement of a high quality muon and τ_h candidate is expected to be well modeled by simulation, and has indeed been validated by the POGs. Moreover, the \cancel{E}_T in $Z \rightarrow \tau\tau$ events is “real” missing momentum from the escaping neutrinos produced by the τ -lepton decay, and thus is also expected to be well modeled by simulation. Therefore, the estimate of the DY + Jets contribution ($\sim 98\% Z \rightarrow \tau\tau$) is based on both simulation and data. We utilize a scale factor, defined as the ratio of efficiencies measured in data and simulation, to normalize the predicted rate to the data by correcting the expected DY + Jets contribution obtained from MC. A semi-clean sample of $Z \rightarrow \tau\tau$ events is obtained with the following modifications to the final signal selections:

- $\cancel{E}_T < 30$ GeV
- $m(\mu, \tau_h, \cancel{E}_T) < 300$ GeV

It is important to note that the analysis has been designed to achieve sensitivity in the high mass region, while also enabling us to obtain a relatively clean $Z \rightarrow \tau\tau$ signature in the low mass region with only minor modifications to the final selection criteria. Figure 5.23 shows the $m(\mu, \tau_h, \cancel{E}_T)$ invariant mass in the $Z \rightarrow \tau\tau$ control region defined above, while the number of observed events in data and the predicted background yields are listed in Table 5.13 (The W + Jets and QCD multijet backgrounds are data-driven).

Table 5.13: Events in the $Z \rightarrow \tau\tau$ control region for data and MC.

Data	6658
W + Jets	2587 ± 168
Z + Jets	3806 ± 100
WZ + Jets	9 ± 3
WW + Jets	49 ± 8
ZZ + Jets	3 ± 2
QCD	320 ± 125
$t\bar{t}$	33 ± 6
Purity	55.9%
Data - $\sum_i BG_i$	3658 ± 225
SF	0.96 ± 0.06

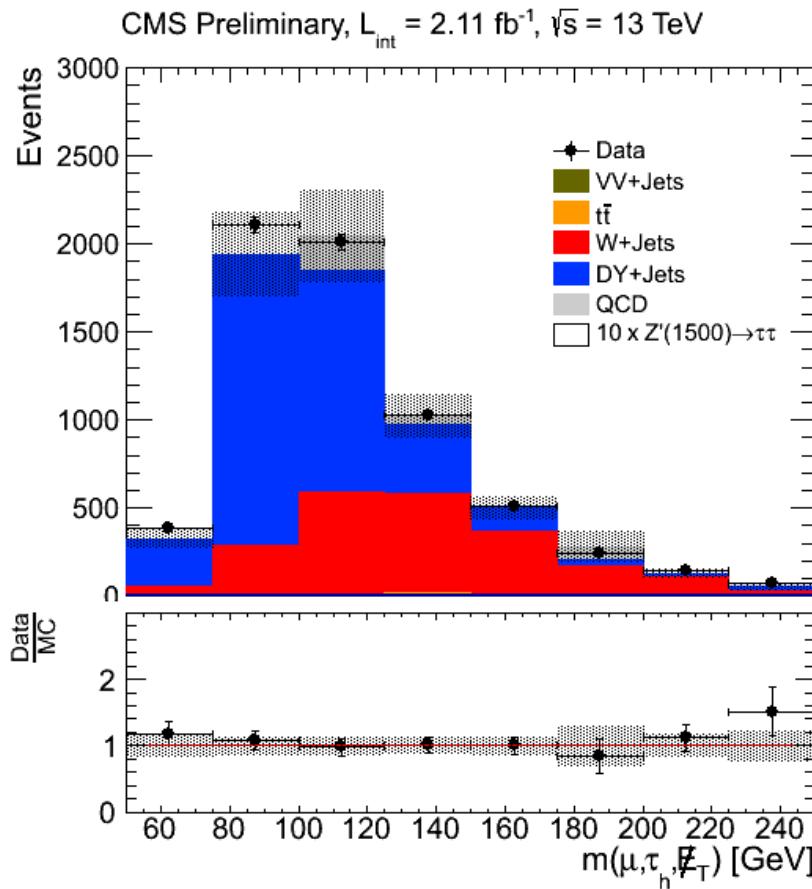
Figure 5.23: $m(\mu, \tau_h, \cancel{E}_T)$ distribution in the $Z \rightarrow \tau\tau$ control region defined by $\cancel{E}_T < 30 \text{ GeV}$ and $m(\mu, \tau_h, \cancel{E}_T) < 300 \text{ GeV}$, with otherwise similar selection criteria to the signal region. The $Z \rightarrow \tau\tau$ (DY+Jets) events are shown in blue.

Table 5.13 and Figure 5.23 clearly show that both the $Z \rightarrow \tau\tau$ rates and shapes are consistent between data and MC, with a SF of 0.96 ± 0.06 . Therefore, the expected $Z \rightarrow \tau\tau$ contribution in the signal region is taken from MC. The expected number of $Z \rightarrow \tau\tau$ events in the signal region is 883 ± 35 (see Table 5.10). Obtaining a semi-clean sample of $Z \rightarrow \tau\tau$ events where the number of observed events are consistent with the MC expectation ensures the robustness of the analysis and validates the ability to successfully identify a τ_h . Furthermore, since our τ_h identification efficiency is constant vs. $p_T(\tau_h)$, validating the correct modeling of $Z \rightarrow \tau\tau$ at low $p_T(\tau_h)$ indirectly validates that we can correctly identify a high- p_T τ_h . The total predicted background $m(\mu, \tau_h, \cancel{E}_T)$ spectrum will be shown in Section 5.10.

5.5.3 Other Small Backgrounds

As mentioned in the introduction to Section 5.5 and as can be calculated from Table 5.10, the sum of the Drell-Yan and W + jets backgrounds represent 89% of the total background in this channel. The remaining $t\bar{t}$, diboson, and QCD multijet processes compose $\sim 1.3\%$, 1.4% , and 4.4% of the total background, respectively. Therefore, these backgrounds are taken directly from simulation with appropriate systematic uncertainties, which are discussed and summarized in Section 5.9. Furthermore, the $t\bar{t}$ and diboson backgrounds are expected to be well modeled by simulation and thus, for added confidence, we only perform a validation of the “good enough” modeling of the QCD multijet background. For this purpose, we obtain a control sample enriched with $\sim 91\%$ purity of QCD multijet events by inverting the isolation on both the τ_h and μ legs as well as inverting the $\mu\tau_h$ charge requirement (select like-sign). Table 5.14 summarizes the observed data events and predicted background yields in this sample. The measured data-to-MC correction factor of $SF = 1.09 \pm 0.27$ is consistent with unity, and thus the QCD multijet yields (in the signal and control regions) are taken directly from simulation. The QCD multijet contribution to the signal region is conservatively assigned an additional 27% systematic uncertainty based on the uncertainty on the correction factor from this control sample (this is added in quadrature to the $\sim 63\%$ uncertainty from the limited number of MC events).

Table 5.14: Yields in the $\mu\tau_h$ QCD control region for data and MC.

Process	non-isolated τ_h , non-isolated μ , same-sign $\mu\tau_h$
Data	442
W + Jets	33 ± 9
Z + Jets	1 ± 1
WZ	0.3 ± 0.6
WW	0 ± 0
ZZ	0 ± 0
QCD	372 ± 91
$t\bar{t}$	2 ± 1
Purity	91.0%
Data - $\sum_i BG_i$	405 ± 23
SF	1.09 ± 0.27

5.6 The Fully-Hadronic Channel

The large branching ratio of $\tau\tau \rightarrow \tau_h\tau_h$ (about 42%) makes the contribution of the $\tau_h\tau_h$ channel to the sensitivity of the overall analysis highly important. Because of the resemblance of QCD jets with τ_h , the probability of misidentifying a QCD jet as a τ_h is at least an order of magnitude higher than that for a QCD jet to be misidentified as an electron or a muon. As a result, the final state is highly contaminated by QCD multijet background ($> 90\%$ of the background in the signal region). But, typical signal events are expected to appear at fairly high invariant mass values, where the QCD multijet contribution is strongly reduced in the high mass region. Thus QCD multijet background only moderately affects the sensitivity of the analysis. Apart from QCD multijets, the other prevailing background is the Drell-Yan process giving rise to τ leptons.

The selection criteria—designed to discriminate between the signal and background—are divided into: kinematic and geometric acceptance for selecting $\tau_h\tau_h$ pairs, τ_h identification, and topological requirements. The main difference with the analyses of the $e\mu$, $\mu\tau_h$ and $e\tau_h$ channels are the substantially tighter $p_T(\tau_h)$ requirements in order to stay efficient with respect to the trigger and suppressing the QCD multijet backgrounds. All surviving opposite-sign pairs of τ_h candidates are preserved at each intermediate stage in the selection. In events in which more than one pair of unique τ_h candidates pass all the requirements, only the pair with the highest $m(\tau_h, \tau_h, \cancel{E}_T)$ is selected. This requirement has a very high efficiency for both signal and backgrounds (the fraction of events with more than one pair is $\ll 1\%$).

Events fired by the HLT_DoubleMediumIsoPFTau35_Trk1 trigger are considered as the interesting events for offline analysis. After the trigger, events are required to have at least 2 HPS taus with p_T greater than 60 GeV. These taus are required to have pseudorapidity value of $|\eta| < 2.1$. A τ_h candidate is also required to satisfy the reconstruction and identification criteria described in Section 5.3. The τ_h candidates are required to pass the following discriminators : “DecayModeFindingNewDMs,” “TightCombinedIsoDB3Hits,” “againstMuonTight3,” and “againstElectronLooseMVA5.” The efficiency of these discriminators is shown in Section 5.3.4.1. These

discriminators ensure the proper identification of a τ_h against QCD jets, muons, and electrons. The candidate $\tau_h\tau_h$ pairs are required to be separated in $\eta - \phi$ space by $\Delta R(\tau_h, \tau_h) > 0.3$. Further, to reduce top pair contamination the event is required not to have any jet identified as a b-quark jet by the b-tagging algorithms using the “combined secondary vertex loose” (CSVL) working point. Further, each event must have at least 30 GeV of missing transverse energy to account for the neutrinos present in the signal and further discriminate against the QCD multijet background. For consistency with the other channels, only 1- and 3-prong taus are considered.

The $\tau_h\tau_h$ pair is expected to be back-to-back with $\cos(\delta\phi(\tau_h, \tau_h)) < -0.95$. In addition, a “ ζ ” cut is used ($p_\zeta - 3.1 \cdot p_\zeta^{vis} > -50$), which is explained in Section 5.5. These are topological cuts which reduce the contamination of backgrounds from $t\bar{t}$ and W+Jets processes to negligible levels.

The Standard Model processes considered as backgrounds are Drell-Yan, diboson production, top quark single and pair production, W+Jets production, and QCD multijet production.

5.6.1 QCD Background Estimation & Validation Strategy

The main background for the $\tau_h\tau_h$ final state is QCD multijet events ($> 90\%$) and it is evaluated by a data-driven approach as MC simulation is not expected to model the background properly, nor do the MC samples provide sufficient statistics. The number of events in the signal region is given by the following equation:

$$N_{\text{Signal}}^{\text{QCD}} = N_{\text{LS}}^{\text{QCD}} \cdot R_{OS/LS} \quad (5.7)$$

Where $N_{\text{LS}}^{\text{QCD}}$ and $R_{OS/LS}$ are evaluated using the following approach. For QCD estimation, we rely on counting events selected in a similar way to the signal events, but selecting $\tau_h\tau_h$ pairs with like-sign electrical charge. This criterion leads to events heavily dominated by the QCD multi-jet background. Assuming that QCD dijets are charge-blind, the number of like-sign events N_{LS} should be equal to the number of opposite-sign N_{OS} QCD multi-jet events after correcting N_{LS} measured in data for known contamination from electroweak backgrounds using simulation. However, the assumption of the charge symmetry in events with two jets is not always true. An asymmetry in

the charges of jets in multi-jet events can arise from the remaining correlation between the quark charge and the leading track charge of the jet in events where quark charges are correlated. Note that the correlation between the charge of the quark and the charge of the track becomes stronger in jets, in which the entire jet fluctuates into just a few high momentum tracks. Calculation of QCD N_{LS} and N_{OS} terms are made in control regions of the $\cancel{E}_T < 30$ GeV sidebands for the calculation of the asymmetry factor $R_{OS/LS}$ (shown in the next section). The contamination from signal in the LS control regions is small since the charge mis-measurement is small (1 – 5%) for $m(Z') < 2.5$ TeV.

The QCD estimation and validation strategy used in this analysis is shown in Figure 5.24. The shape is taken from region C and scaled to the signal estimation using a OS/LS scale factor derived in the $\cancel{E}_T < 30$ GeV sideband (regions B and D).

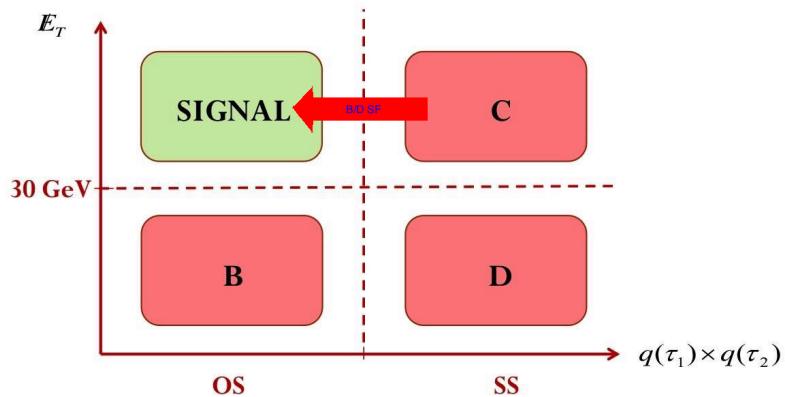


Figure 5.24: QCD estimation and validation strategy for the $\tau_h\tau_h$ channel.

The shape of the $m(\tau_h, \tau_h, \cancel{E}_T)$ distribution is obtained from control region C (same-sign $\tau_h\tau_h$ with nominal \cancel{E}_T). To extract the OS/LS ratio from data, two control regions, B and D, are obtained by keeping the same selections as signal selections, but inverting the \cancel{E}_T cut ($\cancel{E}_T < 30$ GeV) and requiring OS and LS $\tau_h\tau_h$ pairs respectively. The contribution of non-QCD MC backgrounds (Drell-Yan + Jets, $t\bar{t}$, W + Jets, and diboson) are subtracted from data in these control regions and then the ratio $R_{OS/LS}$ is measured:

$$\begin{aligned} N_{\text{OS}}^{\text{QCD}} &= N_{\text{OS}}^{\text{Data}} - N_{\text{OS}}^{\text{non-QCD MC}} \\ N_{\text{LS}}^{\text{QCD}} &= N_{\text{LS}}^{\text{Data}} - N_{\text{LS}}^{\text{non-QCD MC}} \end{aligned} \quad (5.8)$$

$$\begin{aligned} R_{\text{OS/LS}} &= N_{\text{OS}}^{\text{QCD}} / N_{\text{LS}}^{\text{QCD}} \\ & \end{aligned} \quad (5.9)$$

Table 5.15 shows the data and MC yields in controls regions B and D. The purity of QCD multijet, defined by Data – $\sum_i BG_i$, is approximately 96 – 99% depending on the sample. The measured OS/LS ratio is 1.64 ± 0.21 . The above equation shows the mathematical procedure used to obtain this ratio.

Closure and validation tests for the background estimation method outlined above are performed with real data, since there are insufficient statistics to perform such a test with simulation. Two aspects are simultaneously tested: (1) closure on the normalization (i.e. $N_{\text{QCD}}^A = N_{\text{QCD}}^C \cdot \frac{N_{\text{QCD}}^B}{N_{\text{QCD}}^D}$); (2) correct determination of the $m(\tau_h, \tau_h, \cancel{E}_T)$ shape. In order to check (with data) whether same-sign $\tau_h \tau_h$ events can correctly model the mass shapes in the opposite-sign regions, we perform a shape closure/validation test by taking the shape from region D (obtained as Data – $\sum_i BG_i$) and normalize it to the QCD yield in control region B. By comparing the shape for the QCD prediction in region B with the observed mass spectrum in the same region, we can determine whether same-sign $\tau_h \tau_h$ correctly models the mass shapes in the opposite-sign region. Furthermore, any disagreement in the shape between data and the QCD prediction can be used to assign a systematic uncertainty on the shape. Figure 5.25 shows the $m(\tau_h, \tau_h, \cancel{E}_T)$ mass distribution for this closure test in control region B. We observe very good agreement between the observed shape and the predicted shape and thus no additional systematic uncertainties are applied due to this particular closure test on the shape.

Two additional control samples are utilized in order to provide a further test of closure for this method. Control regions 2B and 2D are obtained by using the τ_h anti-isolation sidebands (passing the “loose” isolation working point and failing the “tight” isolation working point) with

Table 5.15: Yields in the control regions B and D used for calculation of OS/LS ratio.

Process	OS $\tau_h \tau_h$ isolated + $\cancel{E}_T < 30$ GeV	SS $\tau_h \tau_h$ isolated + $\cancel{E}_T < 30$ GeV
Data	207	113
W + Jets	0.9 ± 1.6	1.2 ± 1.3
Z + Jets	21.8 ± 4.7	0.1 ± 0.4
WZ + Jets	0 ± 0	0 ± 0
WW + Jets	0.5 ± 0.8	0 ± 0
ZZ + Jets	0.1 ± 0.2	0 ± 0
$t\bar{t}$	0.0 ± 0.1	0 ± 0
Data - $\sum_i BG_i$	184 ± 15	112 ± 11
OS/LS ratio	1.64 ± 0.21	

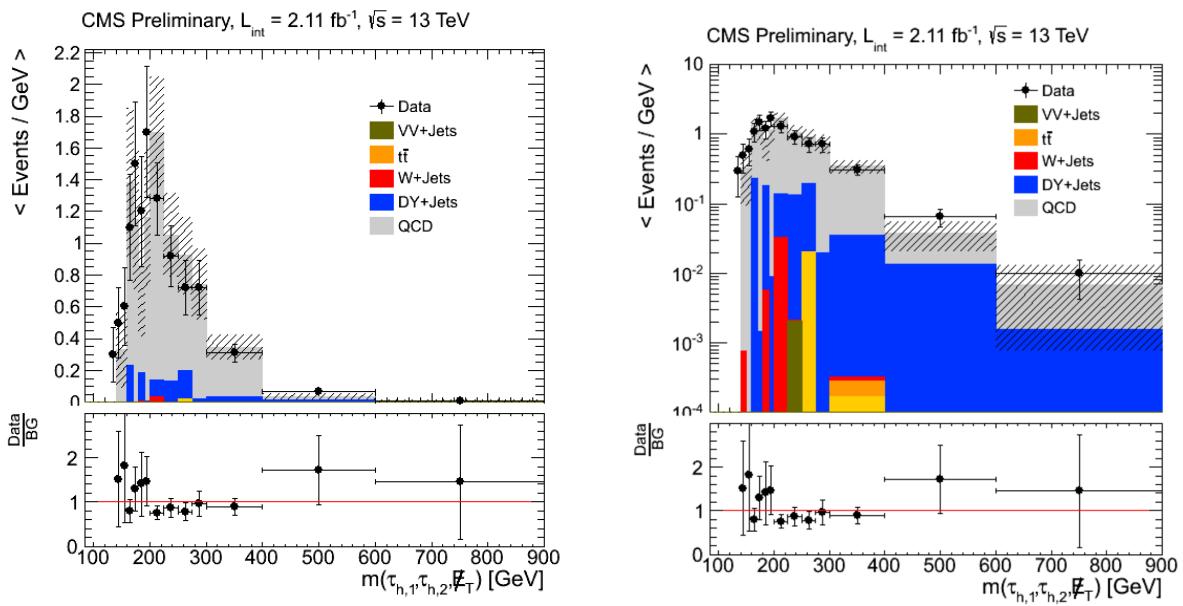


Figure 5.25: $\tau_h - \tau_h$ mass distribution in isolated OS, low- \cancel{E}_T sideband (region B). Left: linear scale. Right: log scale.

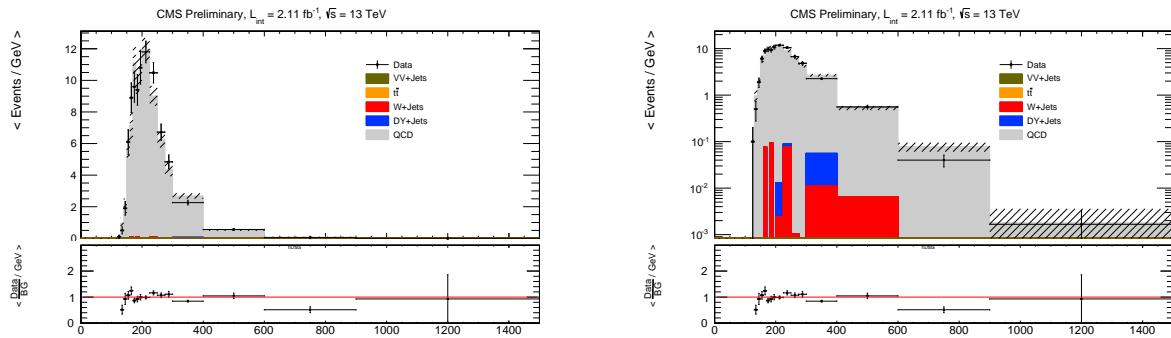


Figure 5.26: $m(\tau_h, \tau_h, \cancel{E}_T)$ mass distribution in anti-isolated OS, low- \cancel{E}_T sideband (region 2B) Left: linear scale. Right: log scale.

otherwise similar selections to control regions B and D (low \cancel{E}_T). The QCD multi-jet prediction in control region 2B is determined as data minus non-QCD backgrounds. The QCD shapes and rates extracted from control region 2D (after subtracting non-QCD backgrounds) are normalized to the QCD yield in control region 2B. Figure 5.26 summarizes this closure test. Good agreement between the data and predicted QCD rate in control region 2B (Figure 5.26) validates the ability to extract the shapes of distributions in OS from LS samples.

Next we perform the full closure test for the ABCD method by comparing the observed yield and mass spectrum in control region 2A (OS, non-isolated $\tau_h \tau_h$ with nominal \cancel{E}_T) with the QCD multijet prediction in that same region obtained by using the QCD shape from region 2C and normalizing it to $N_{\text{QCD}}^{2A} = N_{\text{QCD}}^{2C} \cdot \frac{N_{\text{QCD}}^{2B}}{N_{\text{QCD}}^{2D}}$. Figure 5.27 shows the closure test in region 2A, showing agreement between data and prediction. The QCD multijet prediction in control region 2A, using the ABCD method, is 427 ± 26 , while the observed yield is 429. Thus, no additional systematic uncertainties are applied due to closure.

Table 5.16 gives the yields of data and backgrounds in regions A (signal region) and C (same-sign, isolated $\tau_h \tau_h$ with nominal \cancel{E}_T). The QCD shapes/rates are extracted in data-driven way from control region C and scaled by the OS/LS ratio determined from control region B and D. The non-QCD backgrounds are taken directly from MC in Table 5.16 (see the next section for a validation of the DY + Jets background yield). Non-QCD processes such as W + Jets, $t\bar{t}$, and diboson represent

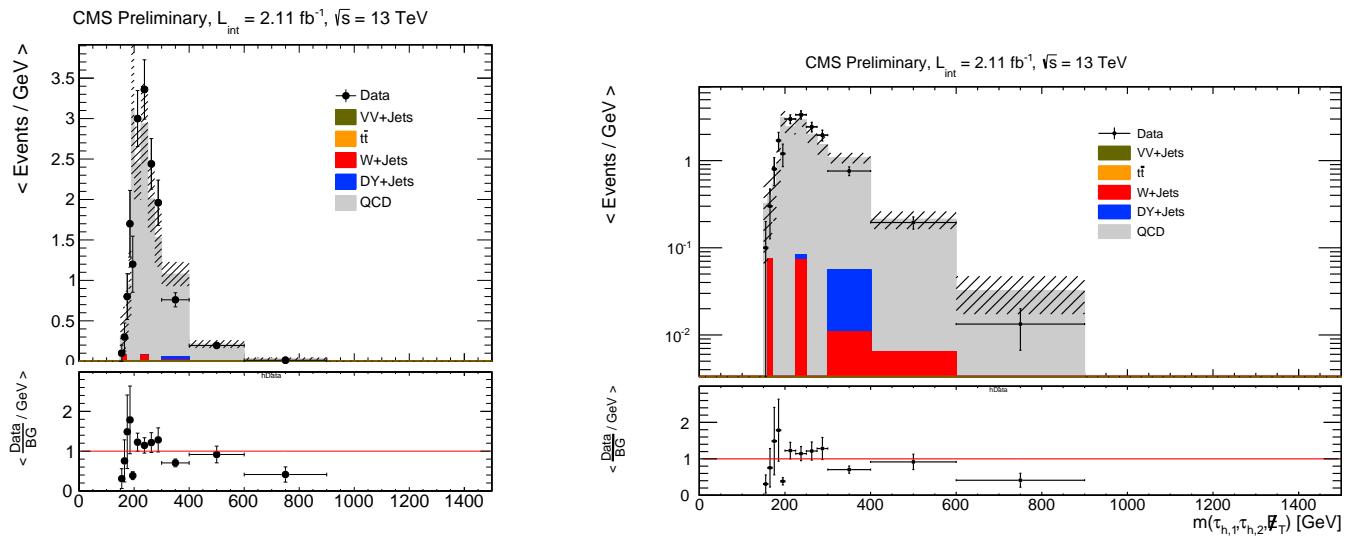


Figure 5.27: $m(\tau_h, \tau_h, \cancel{E}_T)$ mass distribution in anti-isolated OS region with nominal \cancel{E}_T (Left: linear scale. Right: log scale).

only $\sim 1\%$ of the total background rate in the signal region, and are thus taken directly from MC. The final prediction of QCD events in the signal region A is given in the right-most column of Table 5.17. The procedure outlined in this section yields a QCD estimate of $N_{\text{QCD}}^{\text{Signal}} = 48.7 \pm 11.0$. The uncertainty is based on the statistics of the data and MC samples. We stress that this is the QCD predicted rate over the entire $m(\tau_h, \tau_h, \cancel{E}_T)$ spectrum. As was mentioned in the strategy section of this note, we fit for a potential signal that would appear as an excess of events over the standard model expectation in the high $m(\tau_h, \tau_h, \cancel{E}_T)$ part of the distribution. The total predicted background $m(\tau_h, \tau_h, \cancel{E}_T)$ spectrum will be shown in the results section of this chapter.

Table 5.16: Background and data yields in QCD control regions A and C under nominal isolation and \cancel{E}_T conditions (i.e. isolated + $\cancel{E}_T > 30$ GeV).

Process	OS $\tau_h \tau_h$ isolated + $\cancel{E}_T > 30$ GeV (A)	SS $\tau_h \tau_h$ isolated + $\cancel{E}_T > 30$ GeV (C)
Data	55	30
W + Jets	0.09 ± 0.35	0.28 ± 0.61
Z + Jets	8.1 ± 2.8	0.10 ± 0.32
WZ + Jets	0 ± 0	0 ± 0
WW + Jets	0.49 ± 0.77	0 ± 0
ZZ + Jets	0 ± 0	0 ± 0
$t\bar{t}$	0 ± 0	0 ± 0
Data - $\sum_i BG_i$	46.3 ± 8.0	29.6 ± 5.5

Table 5.17: QCD yields in the isolated regions A (signal region), B , C , and D .

Region	OS $\tau_h \tau_h + \cancel{E}_T < 30$ GeV (B)	SS $\tau_h \tau_h + \cancel{E}_T < 30$ GeV (D)	SS $\tau_h \tau_h + \cancel{E}_T > 30$ GeV (C)	OS $\tau_h \tau_h + \cancel{E}_T > 30$ GeV (A)
isolated	184 ± 15	112 ± 11	30 ± 6	49 ± 11

5.6.2 Background Estimation for $Z(\rightarrow \tau\tau) + \text{Jets}$

We do not employ a complete data-driven measurement of the $Z \rightarrow \tau\tau + \text{Jets}$ contribution to the signal region. Instead, we take an approach based on both simulation and data. The efficiency for the requirement of at least two high quality τ_h s is expected to be well modeled by simulation. Therefore, the estimate of the $Z \rightarrow \tau\tau + \text{Jets}$ contribution is determined by obtaining a control sample used to validate the correct modeling of the requirement of at least two high quality τ_h s.

Since the DY + Jets background in this channel is < 10% of the total background in the signal region, this approach is sufficient.

The typical probability of misidentifying a QCD jet as a τ_h is at least an order of magnitude higher than that for a QCD jet to be misidentified as a light lepton. As a result the QCD multijet background in the $\tau_h\tau_h$ channel is substantially higher than in lepton plus tau or dilepton channels. One should note that the presence of large multijet background mainly complicates the definition of suitable control regions for validating the agreement between collision data and simulation for other backgrounds. For this purpose, events are selected using the “pre-selection” cuts, and additionally requiring $\tau_h\tau_h$ pairs with invariant mass less than 100 GeV in order to obtain a semi-clean sample of $Z \rightarrow \tau\tau$ events. Figure 5.28 shows the $m(\tau_h\tau_h)$ distribution for this validation sample where the QCD contribution has been determined using the method discussed in Section 5.6. One can see that the rate and shape between data and MC is consistent. The measured $Z \rightarrow \tau\tau$ data-to-MC scale factor is $SF_{\text{preselection}} = 0.97 \pm 0.19$. As this is consistent with unity, we take $Z \rightarrow \tau\tau$ directly from MC.

5.7 Electron + Hadronic Tau Channel

5.7.1 Event selection

Events must satisfy the single-electron trigger described in Section 5.4. We select reconstructed electrons satisfying:

- $p_T > 35$ GeV and $|\eta| < 2.1$
- distance of closest approach to the leading sum- p_T^2 primary vertex of less than 0.045 cm (transverse) and 0.2 cm (longitudinal)
- passing the tight working point of the e/ γ POG non-triggering MVA ID
- having no matched conversions to photons nor missing hits in the tracker
- within $\Delta R < 0.5$ of the HLT electron that fired the trigger

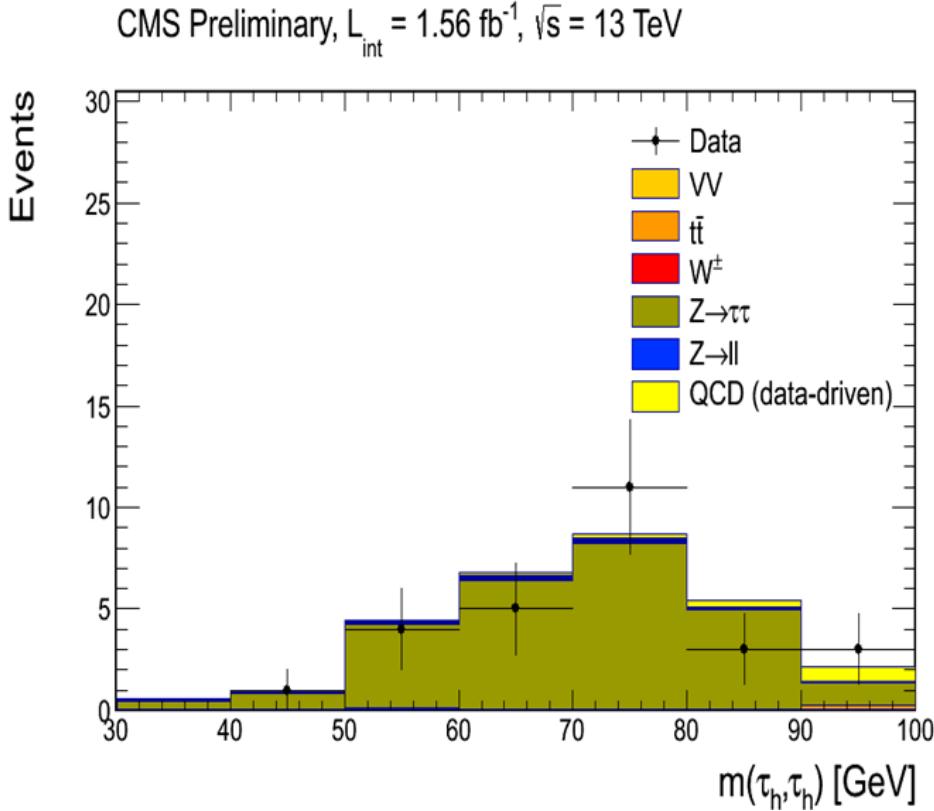


Figure 5.28: $m(\tau_h \tau_h)$ in the $Z \rightarrow \tau\tau$ control region obtained using the “pre-selection” cuts, and additionally requiring $\tau_h \tau_h$ pairs with invariant mass less than 100 GeV.

Offline τ_h 's are required to have:

- $p_T > 20 \text{ GeV}$ and $|\eta| < 2.1$
- distance of closest approach to the leading sum- p_T^2 primary vertex of less than 0.2 cm (longitudinal)
- pass the new Decay Mode Finding requirement as either a 1-prong or 3-prong τ_h
- pass the “againstElectronVLooseMVA5” and “againstMuonTight3” identification requirements

We build $e\tau_h$ pairs in which the electron and τ_h are separated by at least $\Delta R > 0.5$. In events with more than one such pair, we select the pair with the two most isolated leptons, considering first the electron, and then the τ_h . This criterion was seen to have good efficiency for signal samples.

In the rare case of multiple such pairs having identical isolation values, the reconstructed p_T 's are considered with preference given to higher values.

After a pair has been chosen for an event, we apply the following isolation requirements on the leptons, for an event to enter the signal region: electron relative isolation < 0.15 ; τ_h isolation “byTightCombinedIsolationDeltaBetaCorr3Hits.” In order to keep the different final states exclusive, an event is rejected if there is an additional electron satisfying the above identification requirements and with relative isolation < 0.3 , or a muon satisfying the identification requirements described in Section 5.8.1 with relative isolation < 0.3 . To reduce further possible di-electron events in the $e\tau_h$ channel, an event is rejected if there is an opposite-charge electron pair with $\Delta R > 0.15$ in which both of the electrons satisfy $p_T > 15$ GeV, $|\eta| < 2.5$, the e/ γ POG “veto” ID, and relative isolation < 0.3 . These requirements on electron and tau acceptance and ID are referred to henceforth as “preselection” cuts.

As for the other channels, the signal region is defined as having

- $\cos \Delta\phi(e, \tau_h) < -0.95$
- $Q(e) \times Q(\tau_h) < 0$
- $E_T > 30$ GeV
- $P_\zeta - 3.1 \times P_\zeta^{vis} > -50$ GeV
- no jet with $p_T > 30$ GeV tagged as a b-jet (CSV loose)

The distributions of a subset of these variables after preselection are shown in Figures 5.29, 5.30, 5.31.

The distributions of some variables after both the preselection and the selection requirements on the other variables, are shown in Figure 5.32.

The Standard Model processes considered as backgrounds are Drell-Yan, diboson production, top quark single and pair production, W+Jets production, and QCD multijet production.

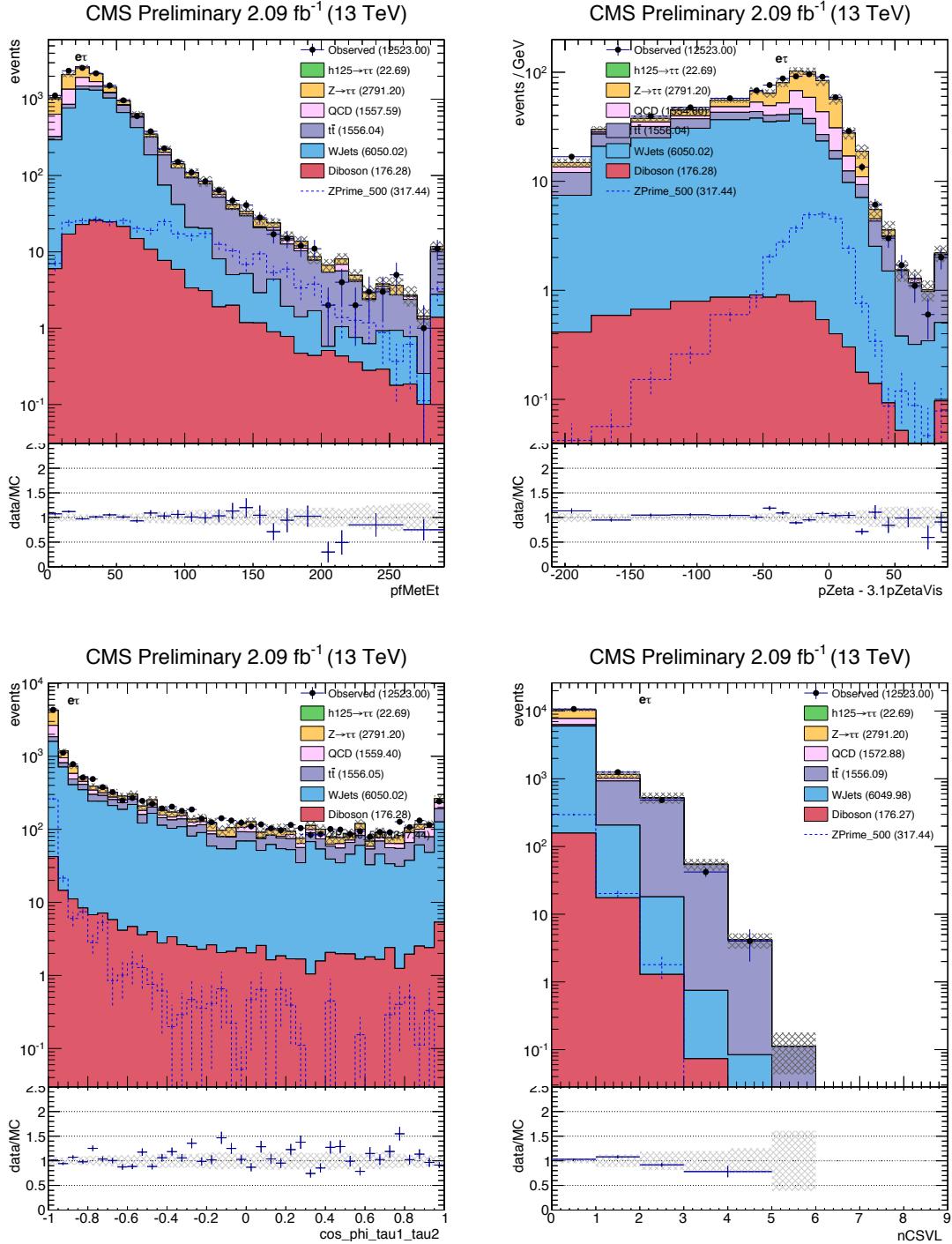


Figure 5.29: Distributions, after preselection, of the variables used for $e\tau_h$ signal selection: E_T (top left), “ p_ζ ” (top right), $\cos\Delta\phi(e, \tau_h)$ (bottom left), and n_b (bottom right).

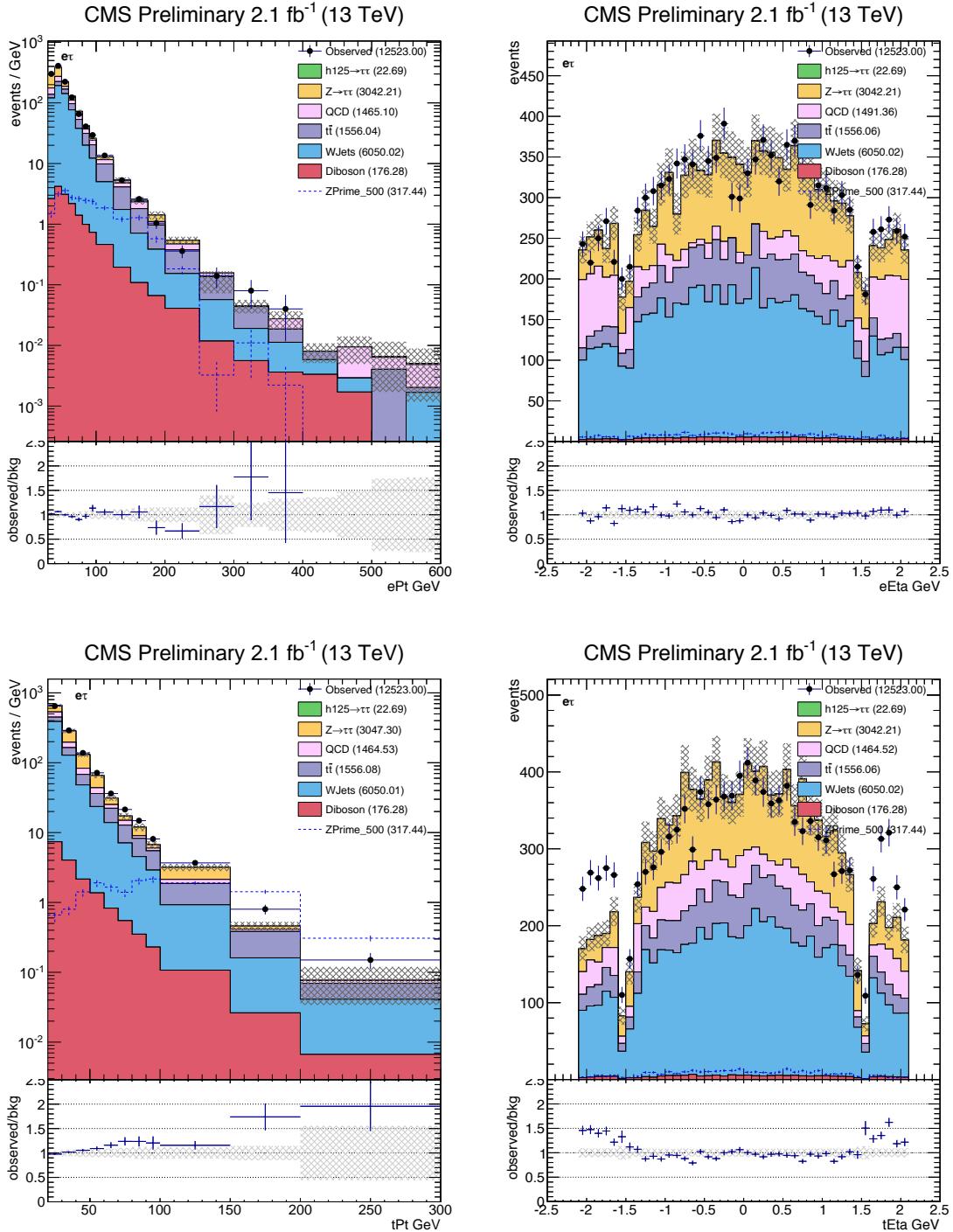


Figure 5.30: Distributions, after $e\tau_h$ preselection, of electron p_T (top left), electron pseudorapidity (top right), τ_h p_T (bottom left), τ_h pseudorapidity (bottom right).

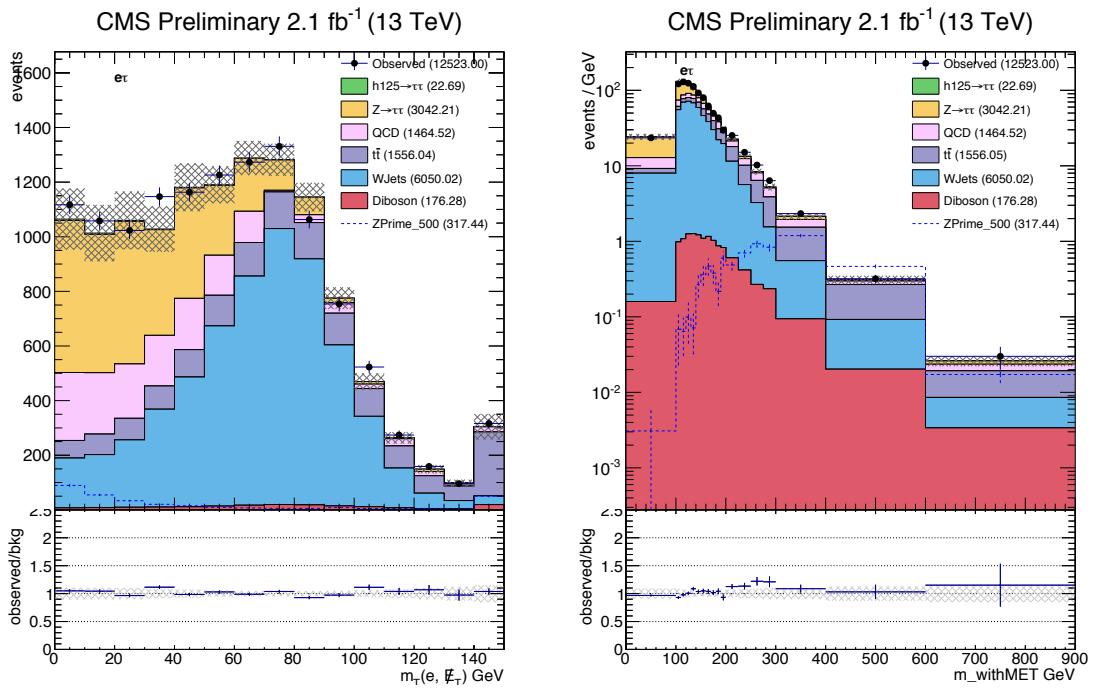


Figure 5.31: Distributions, after $e\tau_h$ preselection, of $m_T(e, \cancel{E}_T)$ (left), $m(e, \tau, \cancel{E}_T)$ (right).

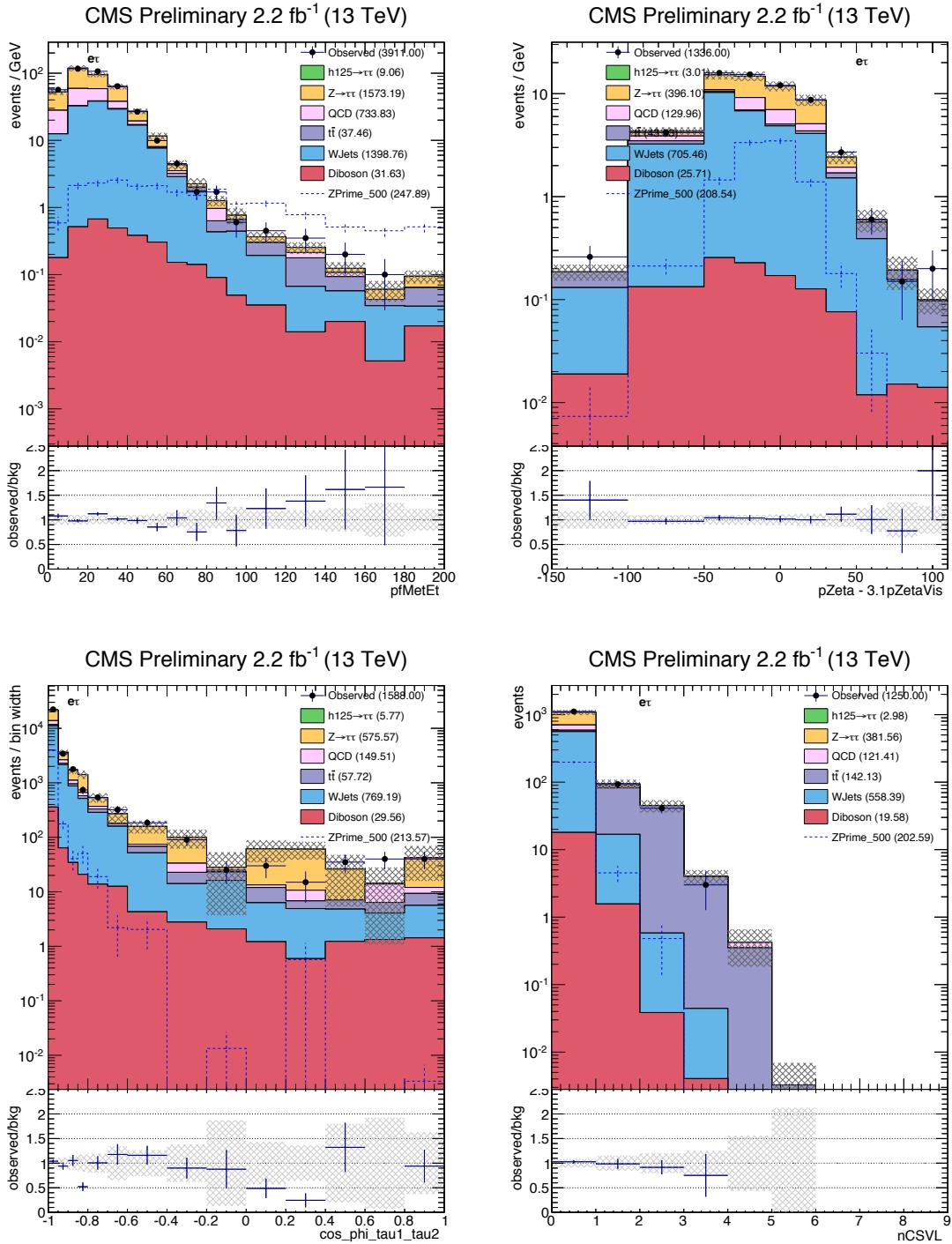


Figure 5.32: Distributions of the variables used for $e\tau_h$ signal selection, after all other signal selection requirements on variables other than the one plotted: \cancel{E}_T (top left), “ ζ ” (top right), $\cos\Delta\phi(e, \tau_h)$ (bottom left), and n_b (bottom right).

5.7.2 Genuine dilepton events

Studies of simulated events indicate that for the Drell-Yan process, top quark single and pair production, and diboson production, the reconstructed and selected electrons and hadronic taus are typically associated with genuine simulated leptons. The nominal expected event rates are estimated by scaling the simulated samples by the best available cross sections, listed in Table 5.2, and by the integrated luminosity of the data samples.

5.7.2.1 Drell-Yan processes

Due to large W+Jets and QCD contamination, as shown in the left panel of Figure 5.33 with the following selection criteria:

- $\cancel{E}_T < 30 \text{ GeV}$;
- no jet with $p_T > 30 \text{ GeV}$ tagged as a b-jet (CSV loose);
- $m(\tau_e, \tau_h, \cancel{E}_T) < 200 \text{ GeV}$

we use the Drell-Yan rate systematic uncertainty (12%) estimated from the $e\mu$ final state in Section 5.8.3.

5.7.2.2 $t\bar{t}$ and single top processes

For $e\tau_h$, we estimate the $t\bar{t} + \text{single top}$ production rate systematic uncertainty (8%) using the $e\mu$ final state, as described in Section 5.8.4, due to its higher top purity. However we cross check this in a $e\tau_h$ top-rich region, defined by the following selection criteria and shown in the right panel of Figure 5.33:

- $\cos \Delta\phi(e, \tau_h) < -0.95$
- $\cancel{E}_T > 30 \text{ GeV}$
- $P_\zeta - 3.1 \times P_\zeta^{vis} > -50 \text{ GeV}$

- at least one jet with $p_T > 30$ GeV tagged as a b-jet (CSV loose)

The $t\bar{t} + \text{single top}$ production rate systematic uncertainty is estimated to be:

$$t\bar{t} + \text{single top systematics} = \left| 1 - \frac{t\bar{t} + \text{single top}}{\text{Data} - \text{other backgrounds}} \right| = 5\% \quad , \quad (5.10)$$

in agreement with the estimate from $e\mu$.

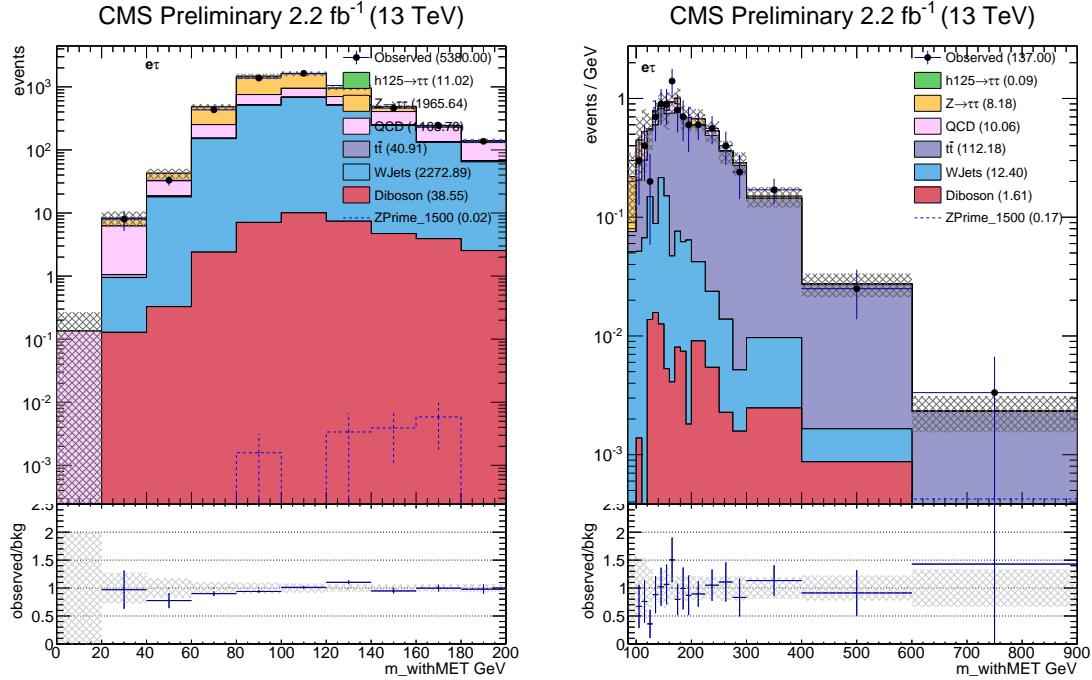


Figure 5.33: Distributions of $m(\tau_e, \tau_h, \cancel{E}_T)$. Left: validation region with $\cancel{E}_T < 30$ GeV, $n_b = 0$ $m(\tau_e, \tau_h, \cancel{E}_T) < 200$ GeV. Right: validation region with $n_b \geq 1$.

5.7.2.3 Diboson processes

We take diboson processes directly from simulation with a 15% production uncertainty. This number is a conservative estimate commonly used by many analyses in which the diboson contribution is a small background and is taken directly from MC.

5.7.3 QCD multijet background

For a given variable and binning, e.g. the effective mass variable used for signal extraction, we construct a data-driven template for the shape of the QCD multijet background, i.e. the processes lacking prompt leptons. Based on the charge of the final state and τ_h isolation, we split the events into four regions shown in Figure 5.34 and described below:

- A (Signal) Region: e and τ_h have opposite charge and τ_h pass “Tight” isolation requirement.
- B Region: e and τ_h have same charge and τ_h pass “Tight” isolation requirement.
- C Region: e and τ_h have opposite charge and τ_h pass anti-isolation requirement.
- D Region: e and τ_h have same charge and τ_h pass anti-isolation requirement.

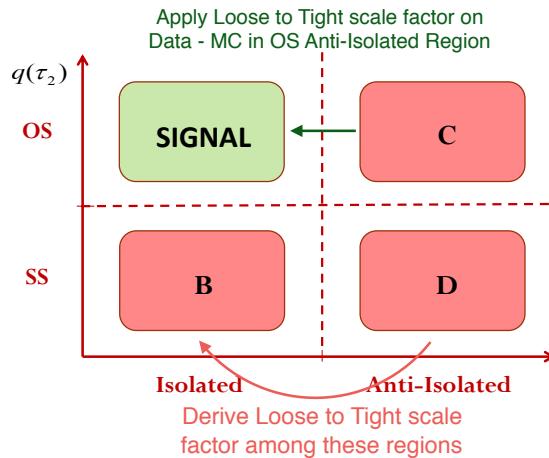


Figure 5.34: QCD estimation strategy for the $e\tau_h$ channel.

In each region (B, C, D), QCD events are estimated by subtracting events with genuine leptons (estimated by simulation) bin-by-bin from the data. QCD events are assumed to be charge blind, thus the number of QCD events in region B should be comparable to that in the signal region. However, with the freedom to define the anti-isolation region, we choose an anti-isolation definition such that there is much more QCD in region C compared to the signal region. Taking the QCD shape from region C will provide us a much smoother template for QCD estimation.

Hence, QCD events in the signal region are estimated with the shape from region C and multiplying a scale factor derived from regions B and D, defined as:

$$f_{\text{LT}}^{\text{QCD}} = (N_{\text{data}}^{\text{B}} - N_{\text{MC}}^{\text{B}}) / (N_{\text{data}}^{\text{D}} - N_{\text{MC}}^{\text{D}}) \quad (5.11)$$

where N_{MC}^{B} and N_{MC}^{D} are the number of non-QCD MC events.

This QCD estimation method is valid only if the QCD shape in the anti-isolated region correctly models the QCD shape in the isolated region. The check is performed by comparing the observation and background estimation in region B with the QCD shape taken from region D and normalized to the QCD in region B. An example of this test is shown in Figure 5.35 with the anti-isolation definition as: τ_h isolation failing the “Tight” working point but still < 5.0 GeV. After the signal region selection the “Loose-to-Tight” scale factor is estimated to be: 0.13 ± 0.02 where an additional 15% uncertainty is added to the QCD systematic uncertainty on top of the bin-by-bin systematics.

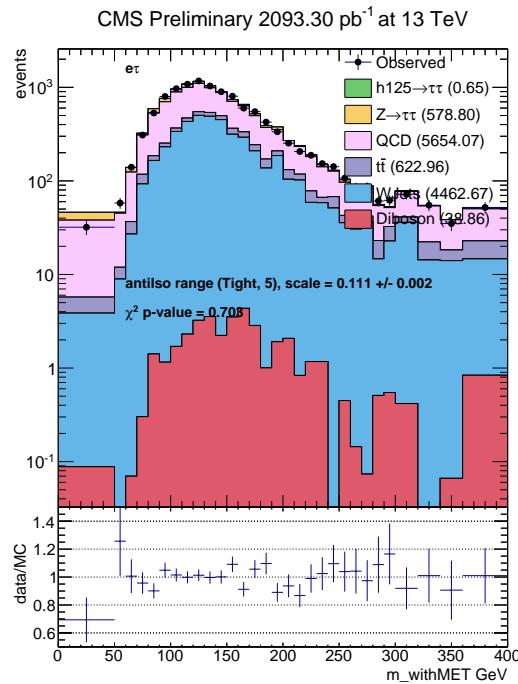


Figure 5.35: The distribution of reconstructed parent mass, $m(\tau_e, \tau_h, \cancel{E}_T)$, in the same-charge, tightly-isolated sample: $e\tau_h$ channel.

5.7.4 W+Jets background

The simulated W+Jets samples, especially at low HT, were not generated with large MC statistics. Using the samples directly would have two disadvantages: (a) non-smooth templates in the signal extraction restricts somewhat the choice of signal selection and binning; (b) poorly sampled shapes degrade the ability to validate the background estimates. As a workaround, we use a region of relaxed τ_h isolation to obtain the simulated shape: τ_h isolation is required to fail the “Tight” working point but be lower than 5 GeV within the isolation annulus. This shape template is then normalized to the integrated MC yield when requiring tight τ_h isolation.

Figure 5.36 compares the direct and relaxed predictions in the signal region and in the control region discussed below. It also shows the compatibility of the predicted shapes when the isolation sideband is split into two regions: τ_h isolation failing “Tight” working point but < 2.5 GeV; 2.5 GeV $< \tau_h$ isolation < 5 GeV.

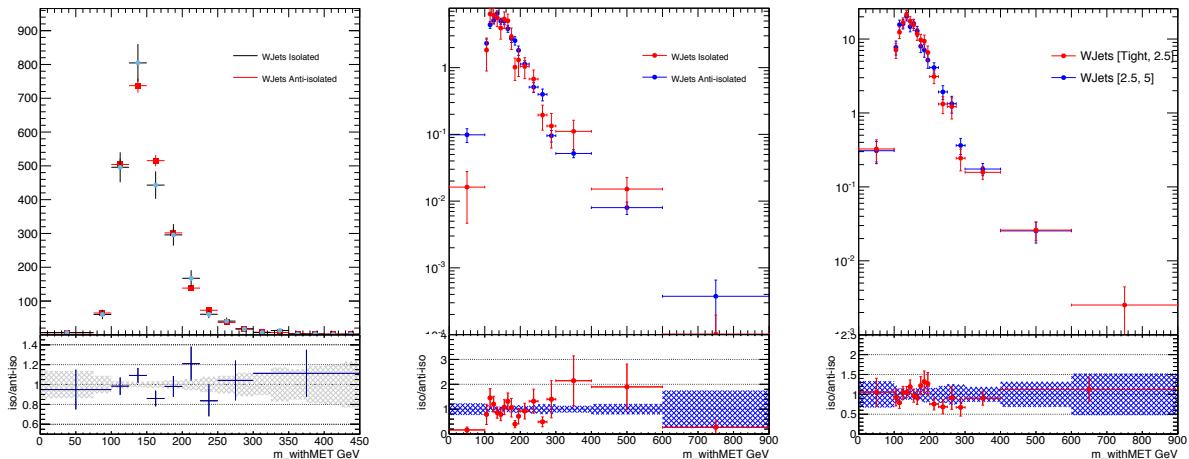


Figure 5.36: Left: comparison of the simulated distributions of $m(\tau_e, \tau_h, \cancel{E}_T)$, using tight τ_h isolation and with relaxed τ_h isolation, in the control region described in Section 5.7.4. Center: analogous comparison for the signal region. Right: comparison in the signal region of the “tighter” and “looser” halves of the sideband.

To evaluate the estimated background rate in a signal-depleted and W-enriched region, we use the same requirements as for the signal region, except requiring $p_\zeta - 3.1 \times p_{\zeta \text{vis.}} < -50$ GeV

(i.e., inverted w.r.t. the signal selection), and allow any value of $\cos \Delta\phi_{e\tau_h}$ (i.e., relaxed w.r.t. the signal selection). We then scan the W+Jets event rate from 1.0 to 1.25, in steps of 0.05. For each assumed rate, we redetermine the QCD prediction in this control region. Figure 5.37 shows the agreement between observations and background estimates at each scan point. A scale factor of 1.15 ± 0.10 gives the best agreement in this control region. The distribution of $m(\tau_e, \tau_h, \not{E}_T)$ in the W-enriched control region, requiring also $\cos \Delta\phi_{e\tau_h} < -0.8$ (to be somewhat closer to the signal region), is shown in Figure 5.38.

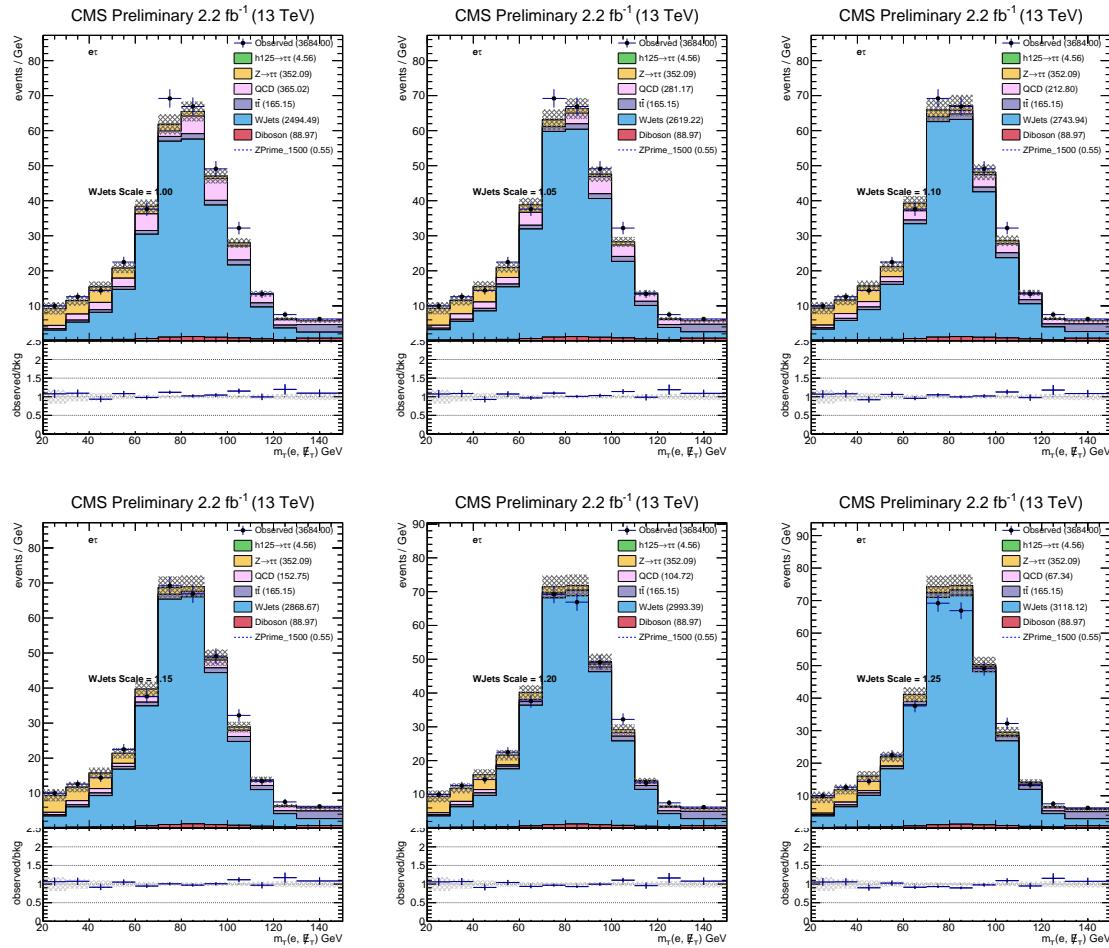


Figure 5.37: Distributions of $m_T(e, \not{E}_T)$ obtained in the W-enriched control region, when scanning the W+Jets event rate from 1.0 to 1.25 times its nominal value, in steps of 0.05. For each assumed W+Jets event rate, the QCD prediction is determined anew using the procedure described in Section 5.7.3.

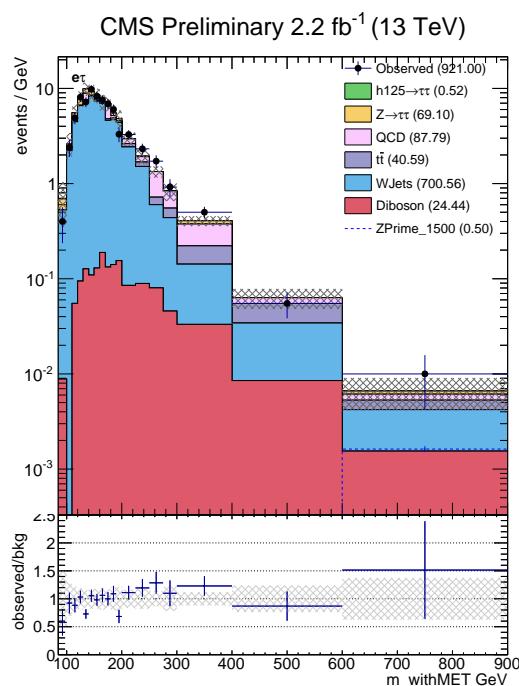


Figure 5.38: $e\tau_h$ channel: the distribution of reconstructed parent mass, $m(\tau_e, \tau_h, \cancel{E}_T)$, in the W-validation region described in Section 5.7.4.

5.7.5 Overlays of observations and SM predictions

The expected SM event yields in the signal region are shown in Figure 5.39.

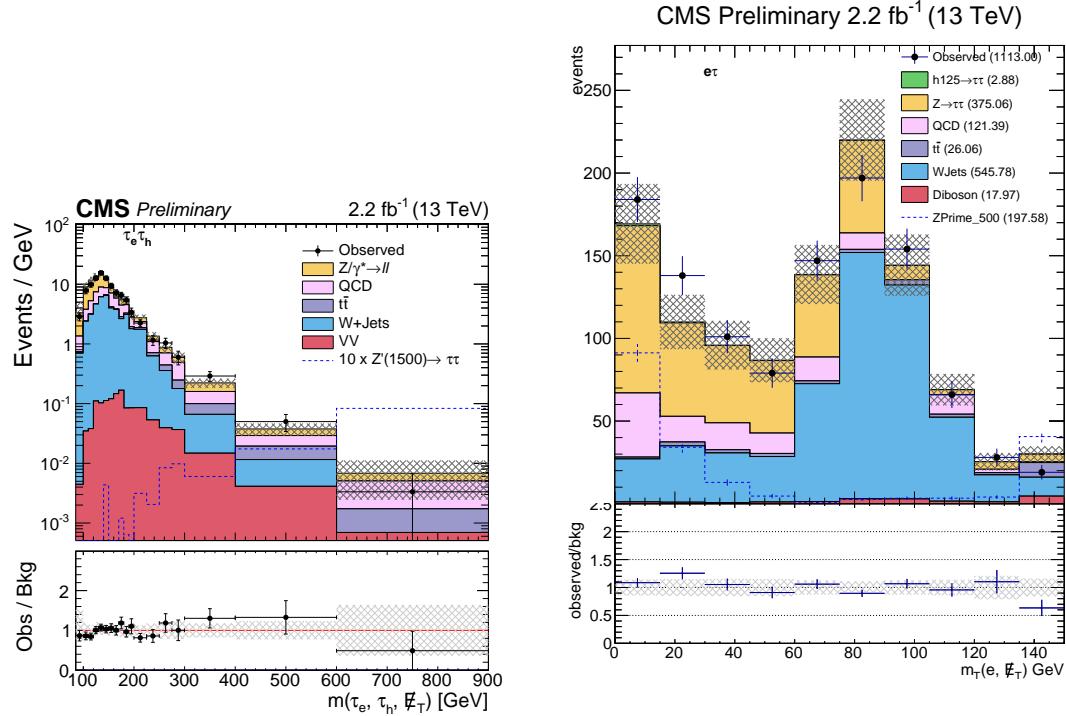


Figure 5.39: Left: predicted background yields and observed event yields in the $e\tau_h$ channel after signal selection. Right: the distribution of transverse mass.

Distributions of p_T and η after final selection are shown in Figure 5.40.

5.8 Electron + Muon Channel

5.8.1 Event selection

The electron selection is identical to that described in Section 5.7.1. Muons are required to

have:

- $p_T > 10$ GeV and $|\eta| < 2.1$
- distance of closest approach to the leading sum- p_T^2 primary vertex of less than 0.045 cm (transverse) and 0.2 cm (longitudinal)

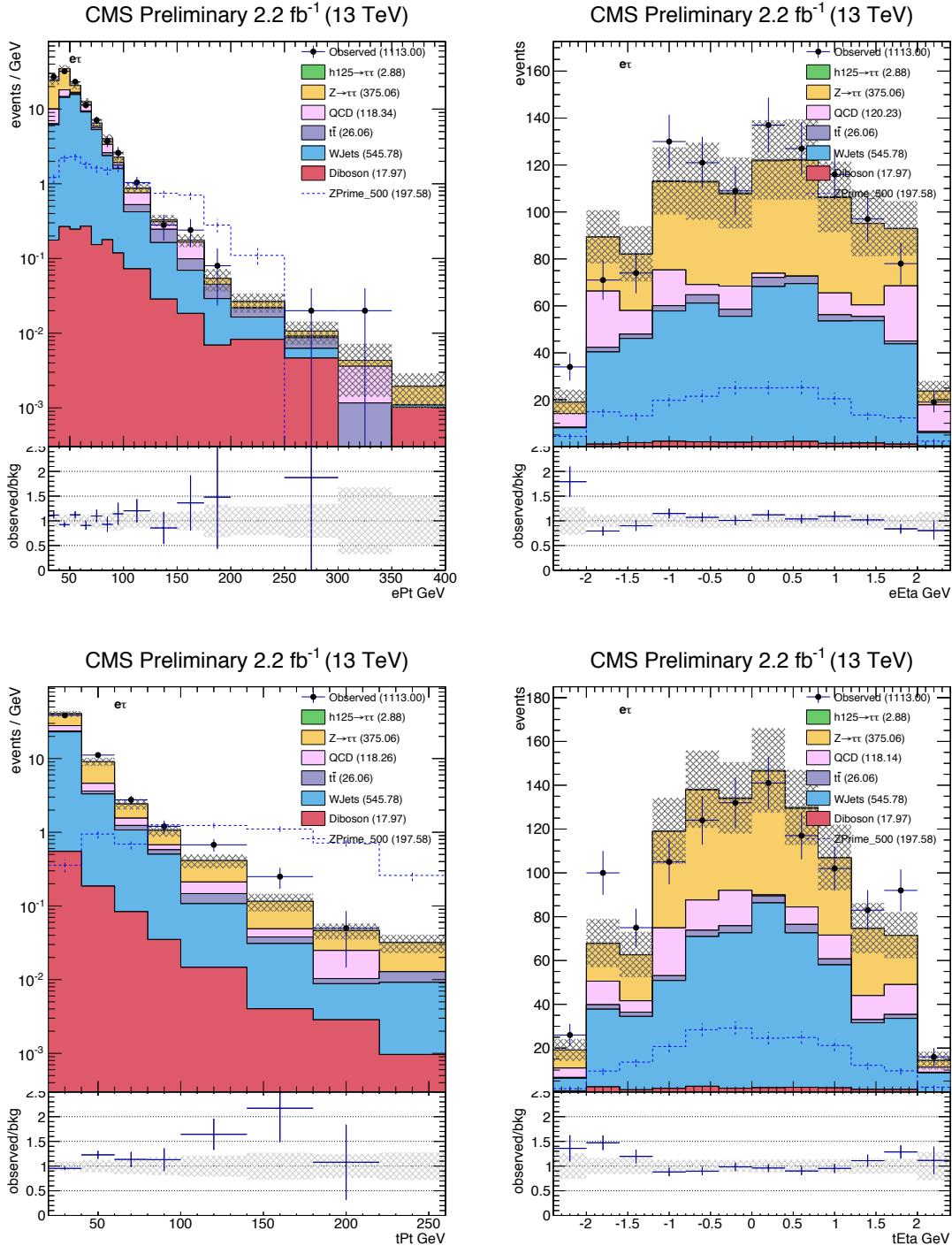


Figure 5.40: Distributions, after $e\tau_h$ final selection, of electron p_T (top left), electron pseudo-rapidity (top right), τ_h p_T (bottom left), τ_h pseudo-rapidity (bottom right).

- satisfy the muon POG medium muon requirement

We build pairs of electrons and muons in which the electron and muon are separated by at least $\Delta R > 0.3$. In events with more than one such pair, we select the pair with the two most isolated leptons, considering first the muon, and then the electron. This criterion was seen to have good efficiency for signal samples. In the rare case of multiple such pairs having identical isolation values, the reconstructed p_T 's are considered, with preference given to higher values.

After a pair has been chosen for an event, we require both the electron and muon relative isolations to be < 0.15 , for an event to enter the signal region. To reduce a possible Drell-Yan background, events are rejected if there is an additional electron satisfying the requirements described in Section 5.7.1 and with relative isolation < 0.3 , or an additional muon satisfying the above identification requirements with relative isolation < 0.3 .

As for the other channels, the signal region is defined as having (after preselection)

- $\cos \Delta\phi(e, \mu) < -0.95$
- $Q(e) \times Q(\mu) < 0$
- $E_T > 30 \text{ GeV}$
- $P_\zeta - 3.1 \times P_\zeta^{vis} > -50 \text{ GeV}$
- no jet with $p_T > 30 \text{ GeV}$ tagged as a b-jet (CSV loose)

The distributions of these variables after preselection are shown in Figures 5.41, 5.42, 5.43.

The distributions of these variables after preselection, and after all other selection requirements except on the variable in question (N-1 plots), are shown in Figure 5.44.

5.8.2 Genuine dilepton events

Studies of simulated events indicate that for Drell-Yan process, top quark single and pair production, and diboson production, the reconstructed and selected muons and electrons are typically associated with genuine simulated leptons. The nominal expected event rates are estimated

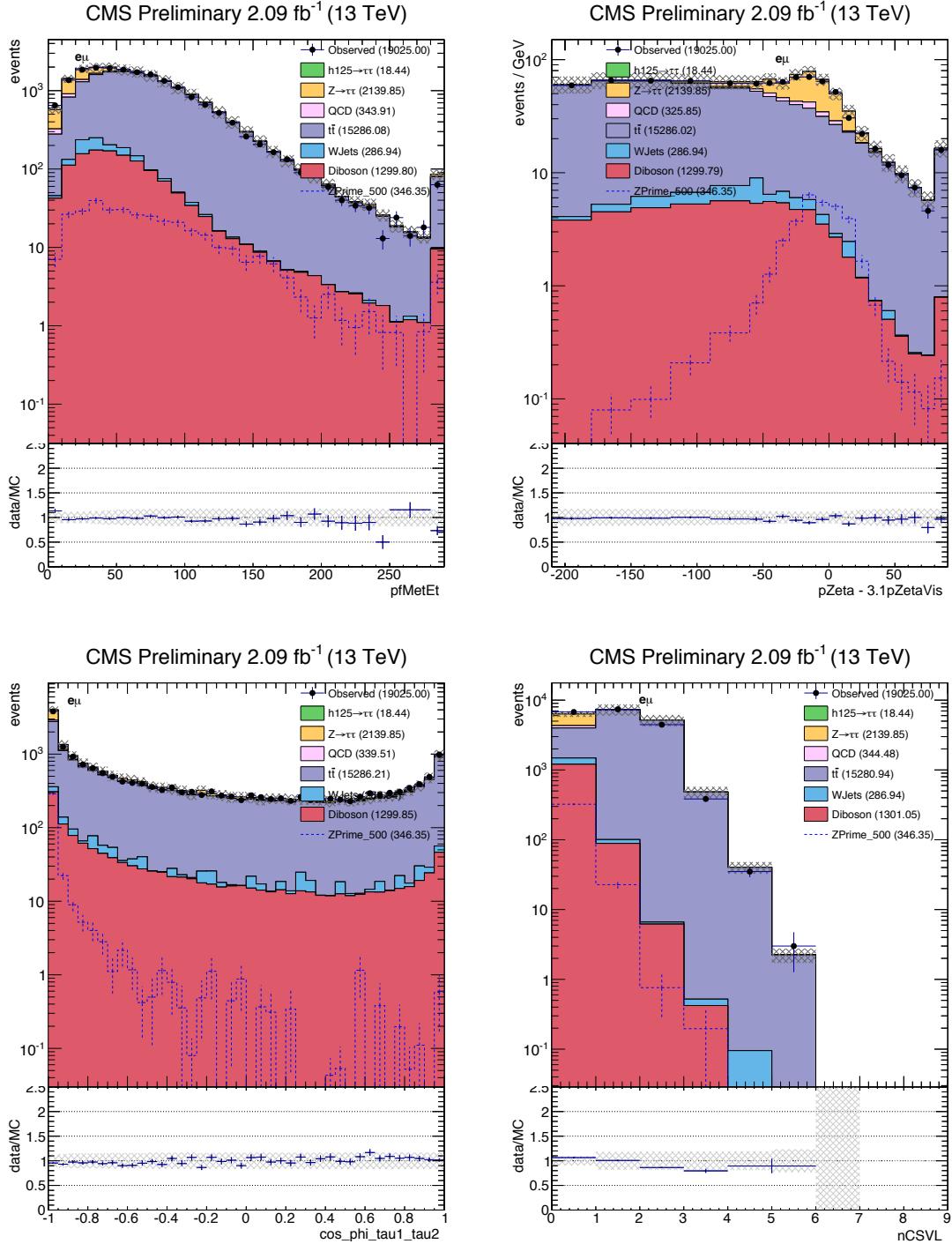


Figure 5.41: Distributions, after preselection, of the variables used for the $e\mu$ signal selection: E_T (top left), “ ζ ” (top right), $\cos \Delta\phi(e, \mu)$ (bottom left), and n_b (bottom right).

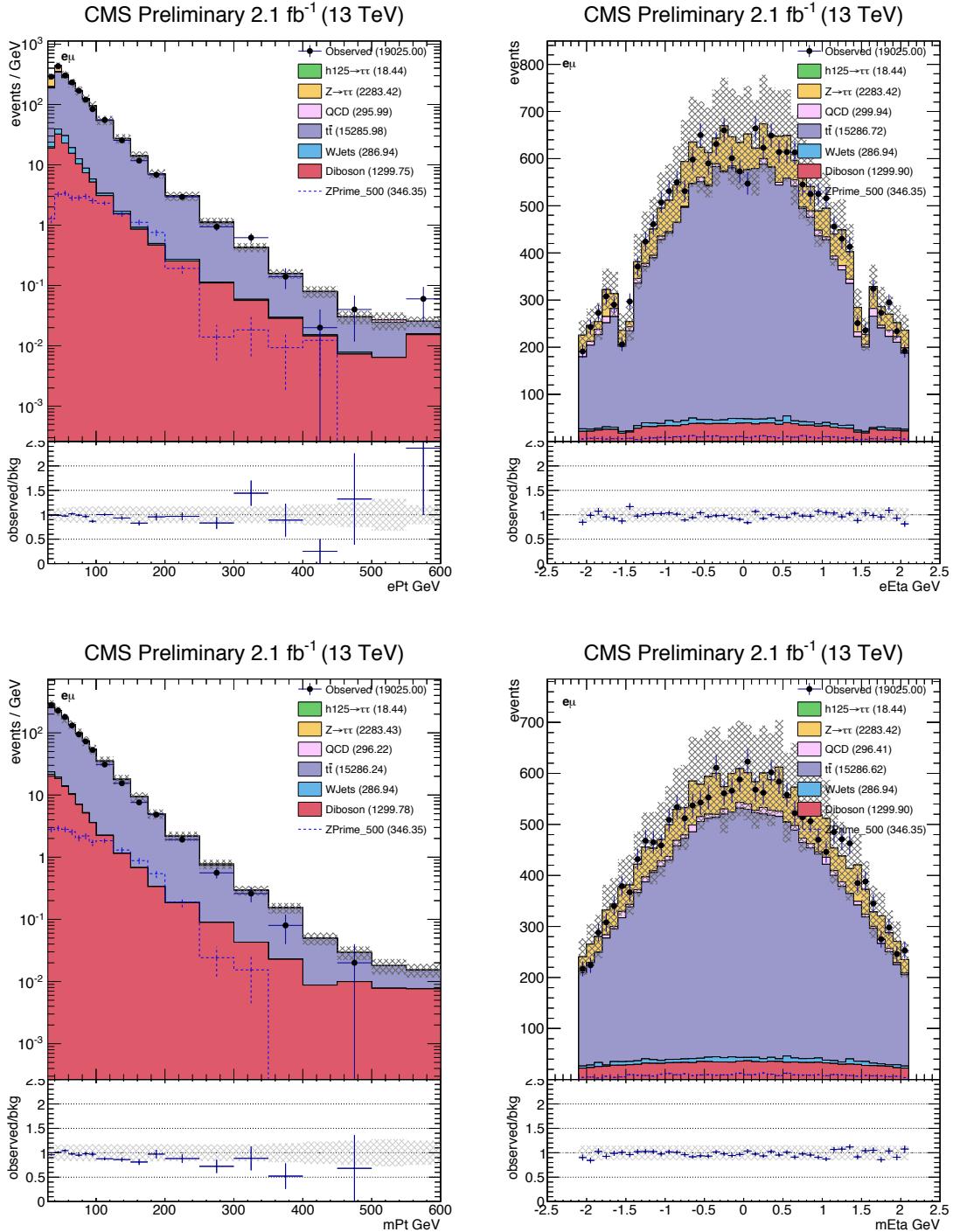


Figure 5.42: Distributions, after $e\mu$ preselection, of electron p_T (top left), electron pseudorapidity (top right), muon p_T (bottom left), muon pseudorapidity (bottom right).

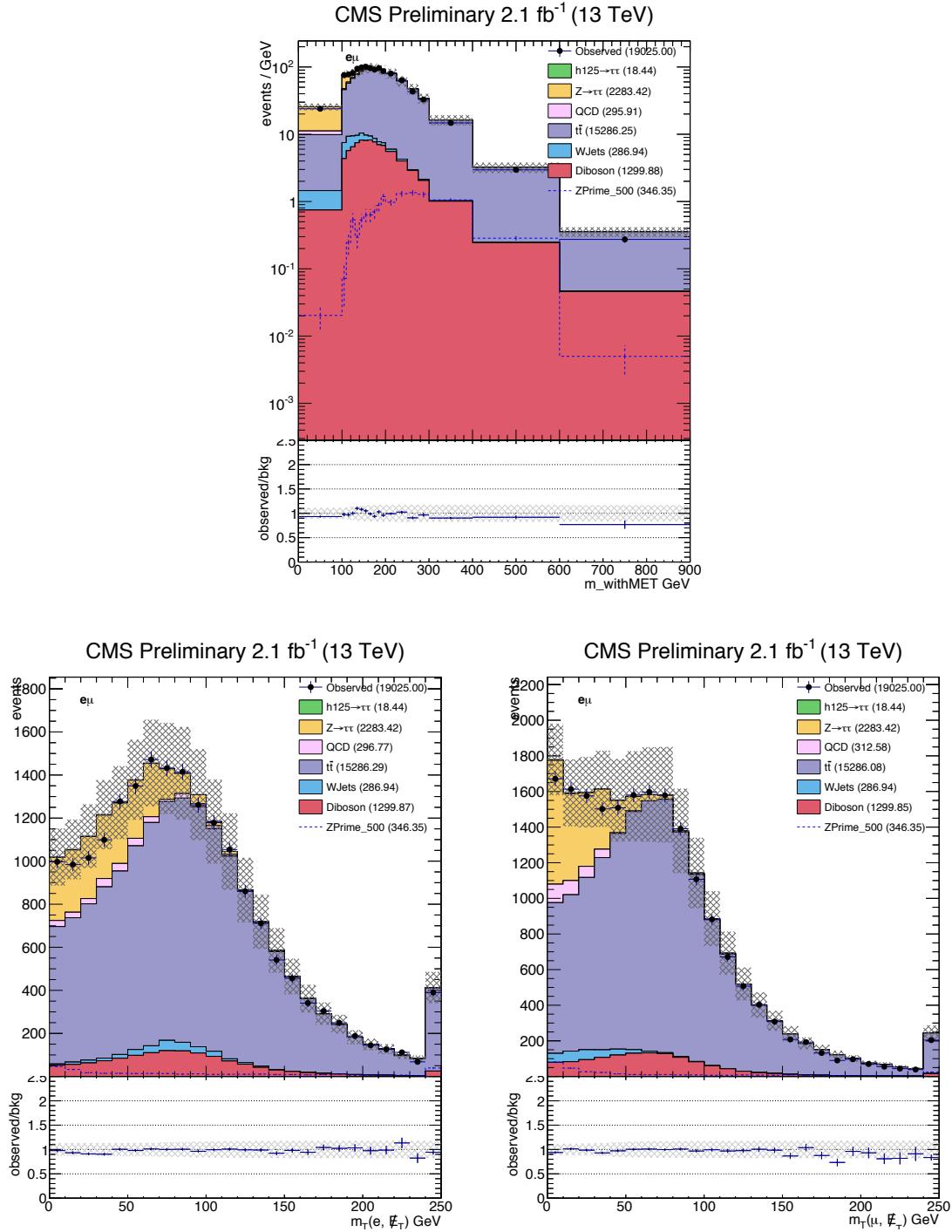


Figure 5.43: Distributions, after $e\mu$ preselection, of $m(e, \mu, \cancel{E}_T)$ (top), $m_T(e, \cancel{E}_T)$ (bottom left), and $m_T(\mu, \cancel{E}_T)$ (bottom right).

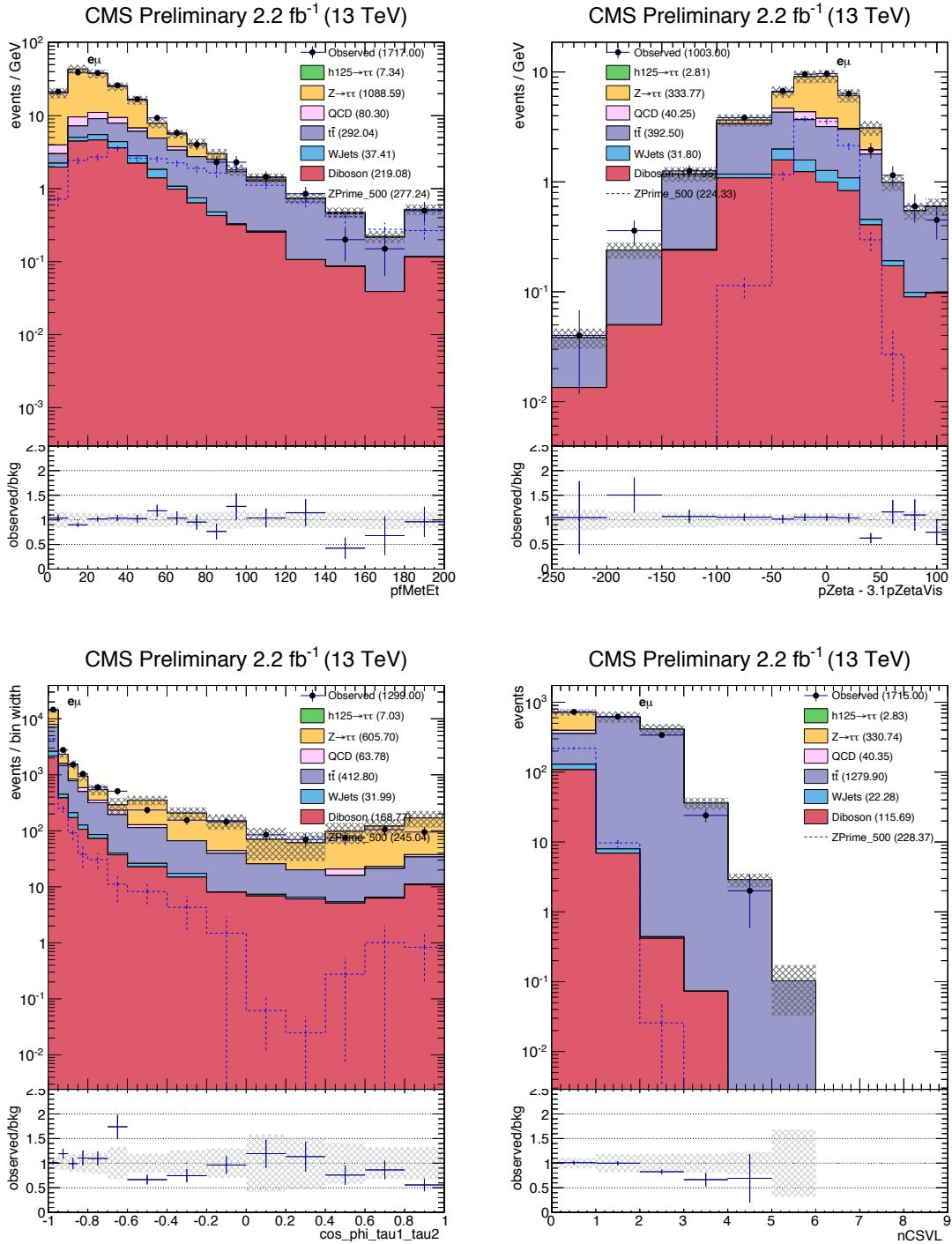


Figure 5.44: Distributions of the variables used for $e\mu$ signal selection, after all other signal selection requirements on variables other than the one plotted: \cancel{E}_T (top left), “ ζ ” (top right), $\cos \Delta\phi(e, \mu)$ (bottom left), and n_b (bottom right).

by scaling the simulated samples by the best available cross sections, listed in Table 5.2, and by the integrated luminosity of the data samples.

5.8.3 Drell-Yan process

Systematics for Drell-Yan process is estimated in a Drell-Yan rich region with the following selections and shown in the left panel of Figure 5.45:

- $Q(e) \times Q(\mu) < 0$
- $\cancel{E}_T < 30 \text{ GeV}$
- no jet with $p_T > 30 \text{ GeV}$ tagged as a b-jet (CSV loose)
- $m(\tau_e, \tau_\mu, \cancel{E}_T) < 125 \text{ GeV}$

The Drell-Yan production rate systematic uncertainty is estimated to be:

$$\text{Drell-Yan systematics} = \left| 1 - \frac{\text{Drell-Yan}}{\text{Data - other backgrounds}} \right| = 12\% \quad (5.12)$$

which we apply both to $e\mu$ and $e\tau_h$ final states.

5.8.4 $t\bar{t}$ and single top processes

Systematics for $t\bar{t}$ and single top processes are estimated in a top quark rich region with the following selections and are shown in the right panel of Figure 5.45:

- $Q(e) \times Q(\mu) < 0$
- $\cancel{E}_T > 30 \text{ GeV}$
- $P_\zeta - 3.1 \times P_\zeta^{\text{vis}} > -50 \text{ GeV}$
- at least one jet with $p_T > 30 \text{ GeV}$ tagged as a b-jet (CSV loose)

The $t\bar{t} +$ single top production rate systematic error is estimated to be:

$$t\bar{t} + \text{single top systematics} = \left| 1 - \frac{t\bar{t} + \text{single top}}{\text{Data - other backgrounds}} \right| = 8\% \quad (5.13)$$

which we apply both to $e\mu$ and $e\tau_h$ final states.

5.8.5 Diboson process

We take diboson processes directly from simulation with a 15% production uncertainty. As in $e\tau_h$, this 15% is taken as a commonly-used conservative estimate for small backgrounds taken directly from MC.

5.8.6 QCD multijet background

The estimation of the QCD background for the $e\mu$ channel is directly analogous to that in the $e\tau_h$ channel, except that the sideband is defined by the muon isolation. The range of relative isolation from 0.15 to 0.95 was chosen as the sideband. After the signal region selection the “Loose-to-Tight” scale factor is estimated to be: 0.20 ± 0.08 where this 40% rate uncertainty is applied to the QCD process (in addition to the bin-by-bin systematic uncertainties).

5.8.7 W+Jets background

The W background is very small. However, as in the $e\tau_h$ channel, the W+Jets simulated sample was not generated with high statistics. As a workaround, the W+Jets shape is taken from the simulated sample in the muon isolation sideband, and scaled to match the simulated yield in the tight muon isolation. The “loose-to-tight” factor is 0.07 ± 0.03 .

5.8.8 Overlays of observations and SM predictions

The expected SM event yields in the signal region are shown in Figure 5.47.

Distributions of p_T and η are shown in Figure 5.48.

5.9 Systematic Uncertainty

The following systematic effects have been considered (summarized in Table 5.18):

- **Luminosity:** We include a 5% uncertainty on the measured luminosity[5]. It is considered 100% correlated across MC based backgrounds within a channel. It is also considered 100%

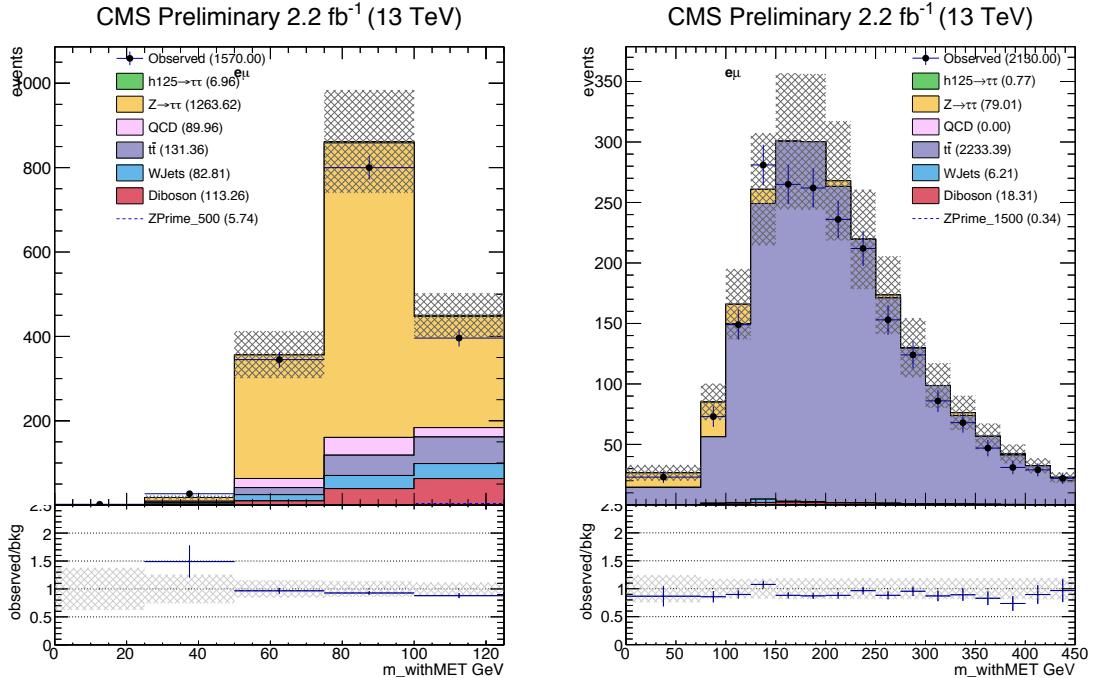


Figure 5.45: Distributions of $m(\tau_e, \tau_\mu, \cancel{E}_T)$. Left: validation region with $\cancel{E}_T < 30 \text{ GeV}$, $n_b = 0$ and $m(\tau_e, \tau_\mu, \cancel{E}_T) < 125 \text{ GeV}$. Right: validation region with $\cancel{E}_T > 30 \text{ GeV}$, $n_b \geq 1$ and $P_\zeta - 3.1 \times P_\zeta^{\text{vis}} > -50 \text{ GeV}$.

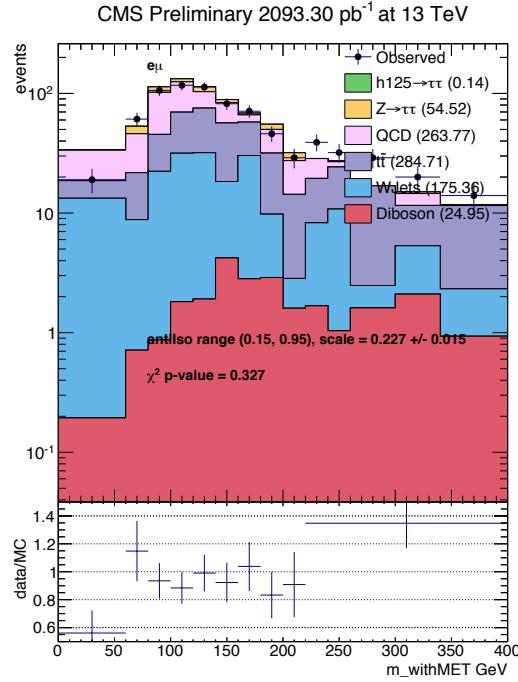


Figure 5.46: The distribution of reconstructed parent mass, $m(\tau_e, \tau_\mu, \cancel{E}_T)$, in the same-sign, tightly-isolated sample: $e\mu$ channel.

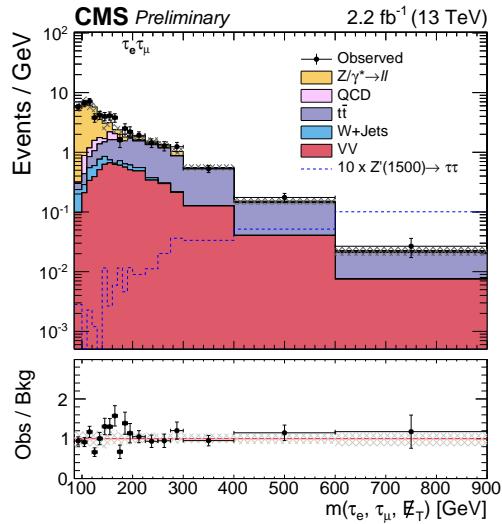


Figure 5.47: Predicted background yields and observed event yields in the $e\mu$ channel after signal selection.

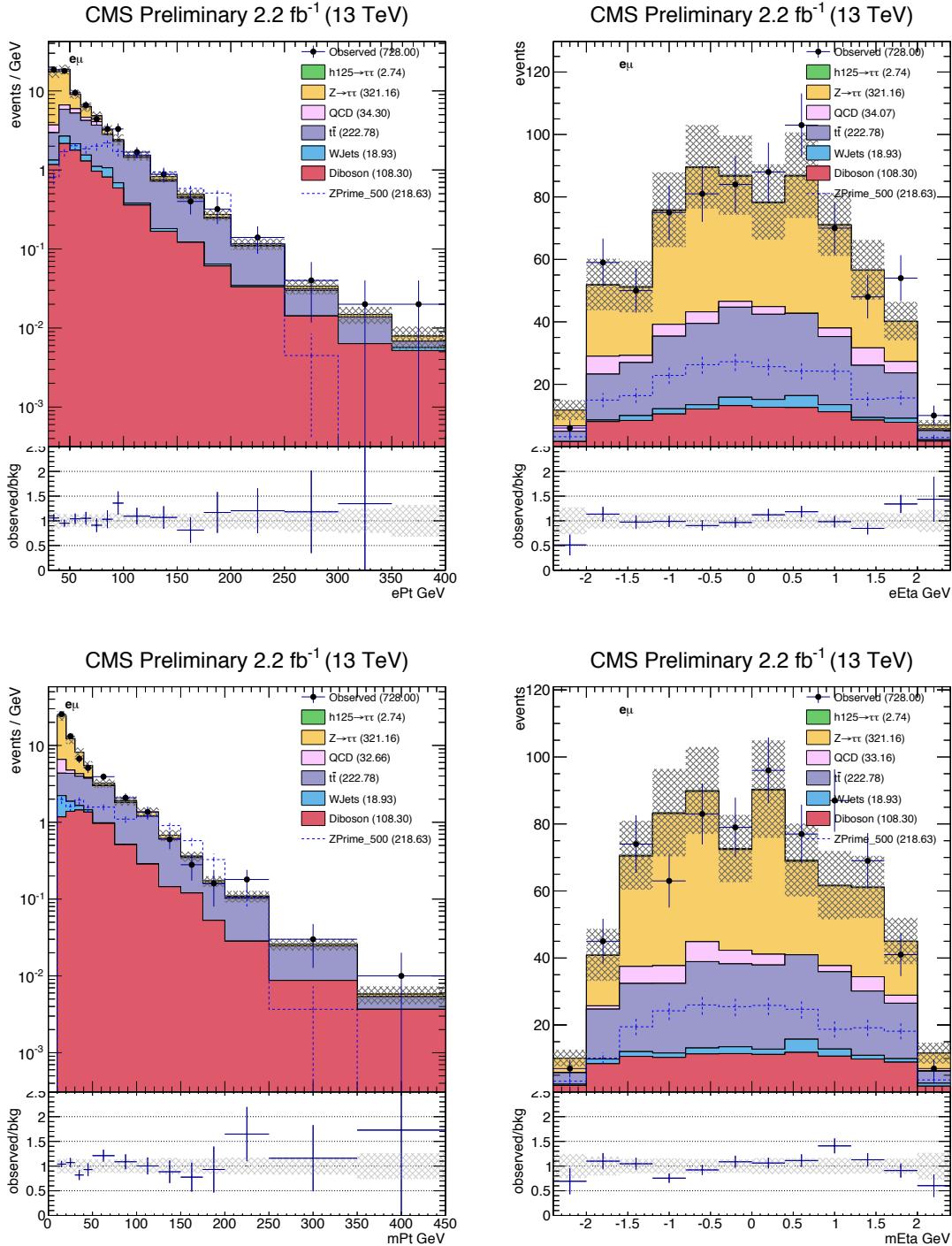


Figure 5.48: Distributions, after $e\mu$ final selection, of electron p_T (top left), electron pseudorapidity (top right), muon p_T (bottom left), muon pseudorapidity (bottom right).

correlated across channels (for MC-based backgrounds).

- **Trigger, Reconstruction, and Selection:** An overall uncertainty is applied for the trigger uncertainties determined on the correction factors described in Section 3 and which are measured using tag-and-probe methods. The uncertainty for muons and electrons is estimated to be 1% each independent of η or p_T , while we consider 5% uncertainty per hadronic tau leg in the $\tau_h\tau_h$ channel (assuming each leg is 100% correlated). The trigger uncertainty is considered 100% correlated across MC based backgrounds within a channel. It is also considered 100% correlated across channels using the same trigger. For the case of the $\tau_h\tau_h$ trigger, where the trigger efficiency uncertainty is measured per τ_h leg, the total trigger uncertainty is calculated by assuming both legs are 100% correlated. For example, if the per leg τ_h trigger uncertainty is 5%, the total trigger uncertainty for the $\tau_h\tau_h$ channel will be 10%.
- ***b*-Tagging Efficiency (**b ID**):** We consider a 30% uncertainty on the mis-tag rate as measured by the b-tagging POG[7]. For the case of our signal, the systematic uncertainty on the requirement of 0 jets mis-tagged as b-jets is determined by propagating the 30% uncertainty on the mis-tag rate through the following equation (which represents the signal efficiency for requiring 0 jets mis-tagged as b-jets):

$$\epsilon^{\text{NBtag}<1} = 1 - \sum_{n=1}^{\infty} P(n) \cdot \sum_{m=1}^n C(n, m) \cdot f^m \cdot (1-f)^{n-m} \quad (5.14)$$

where $P(n)$ is the probability to obtain n additional jets (non-tau and non-lepton) in the event, $C(n, m)$ the combinatorial of n choose m , and f the mis-tag rate. The probability to obtain at least one additional jet in the event is $\sim 10\%$. Therefore, based on the above equation, the mis-tag rate and uncertainty, and the probability to obtain at least one additional jet we calculate a systematic effect of $\sim 5\%$ on our signal due to the mis-tag rate. The b-tagging/mis-tagging systematics are considered 100% correlated across MC based backgrounds with similar composition (e.g. W + Jets and DY + Jets where there

are typically no real b-jets), but completely uncorrelated to backgrounds that have different composition (e.g. $t\bar{t}$ vs. DY + Jets).

- **Electron Energy Scale (EES):** We consider the effect on the signal acceptance efficiency of a 1% (2.5%) shift on the electron energy scale in the barrel (endcap) region. The resultant systematic uncertainty on signal and MC-based backgrounds is < 1%.
- **Muon Momentum Scale (MMS):** We consider the effect on the signal acceptance efficiency of a 1% momentum scale uncertainty on the muon momentum. The resultant systematic uncertainty on signal and MC-based backgrounds is < 1%.
- **Tau Energy Scale (TES):** We consider the effect of the 5% tau energy scale uncertainty measured by the tau POG on the signal acceptance. The energy component of the tau 4-momentum is scaled by a factor of $k = 1.05$, so that $p_{smeared} = k \cdot p_{default}$ and variables are recalculated using $p_{smeared}$. We find that by using $p_{smeared}$ calculated with a factor of $k = \pm 1.05$, the signal and MC-based backgrounds varies by up to 11%.
- **Jet Energy Scale (JES):** We consider the effect of a 3-5% jet energy scale uncertainty on the signal acceptance (depending on the η and p_T of the considered jet as prescribed by the *JetMET* POG). The jet 4-momentum is scaled by a factor of $k = 1.05$, so that $p_{smeared} = k \cdot p_{default}$ and variables are recalculated using $p_{smeared}$. We find that by using $p_{smeared}$ calculated with a factor of $k = \pm 1.05$, the signal and MC-based backgrounds varies by up to 12%.
- **Background Estimates:** The uncertainty on the data-driven background estimations are driven by the statistics in data in the various control samples. There is also a mostly negligible contribution from the level of contamination from other backgrounds in the control regions. In cases where MC based backgrounds must be subtracted off, the uncertainties in the MC backgrounds due to the above listed systematic uncertainties are propagated throughout the subtraction and used to assign a systematic uncertainty on the background

prediction.

Table 5.18: Summary of systematic uncertainties. Values are given in percent. “s” indicates template variations (“shape” uncertainties).

Source	QCD	W	DY	$t + \bar{t}$	VV	Signal
	$hh, \mu h, eh, e\mu$					
Luminosity	-,-,-,-	5,-,5,5	5,5,5,5	5,5,5,5	5,5,5,5	5,5,5,5
μ Trig	-,-,-,-	-,-,-,-	-,1,-,-	-,1,-,-	-,1,-,-	-,1,-,-
μ ID	-,-,-,-	-,-,-,1	-,1,-,1	-,1,-,1	-,1,-,1	-,1,-,1
e Trig	-,-,-,-	-,-,1,1	-,-,1,1	-,-,1,1	-,-,1,1	-,-,1,1
e ID	-,-,-,-	-,-,1,1	-,-,1,1	-,-,1,1	-,-,1,1	-,-,1,1
τ_h Trig	-,-,-,-	10,-,-,-	10,-,-,-	10,-,-,-	10,-,-,-	10,-,-,-
τ_h ID	-,-,-,-	30,-,6,-	12,6,6,-	12,6,6,-	12,6,6,-	12,6,6,-
b ID	-,-,s,s	10,-,s,s	3,3,s,s	10,12,s,s	3,3,s,s	3,3,s,s
JES	-,-,s,s	12,-,s,s	8,s,s,s	12,s,s,s	8,s,s,s	2,2,s,s
MMS	-,-,-,-	-,-,-,1	-,1,-,1	-,1,-,1	-,1,-,1	-,1,-,1
EES	-,-,-,-	-,-,1,1	-,-,1,1	-,-,1,1	-,-,1,1	-,-,1,1
TES	-,-,s,s	11,-,s,s	11,s,s,s	11,s,s,s	8,s,s,s	3,3,s,s
Closure+Norm.	8,68,16,37	5,8,10,41	19,7,12,12	8,8,8,8	15,15,15,15	

5.10 Results

5.10.1 13 TeV Results

Figure 5.49 shows the background predictions as well as the observed $m(\tau_1, \tau_2, \cancel{E}_T)$ spectrum, in log scale, for the four channels considered in this thesis: $\mu\tau_h$ (top left), $\tau_h\tau_h$ (top right), $e\tau_h$ (bottom left), $e\mu$ (bottom right). Table 5.19 lists the number of estimated background events compared with the total number of observed events in data for each final state considering the whole mass spectrum, while Table 5.20 lists those considering $m(\tau_1, \tau_2, \cancel{E}_T) > 300$ GeV. The observed $m(\tau_1, \tau_2, \cancel{E}_T)$ spectrum in the signal region does not reveal any evidence for $Z' \rightarrow \tau\tau$ production

An upper bound at 95% confidence level (CL) is set on $\sigma \cdot BR$, where σ is the cross-section for pair production of $pp \rightarrow Z'$ and BR the branching fraction for $Z' \rightarrow \tau\tau$.

The calculation of the exclusion limit is obtained by using each bin of the $m(\tau_1, \tau_2, \cancel{E}_T)$ distribution to construct one bin of the likelihood and computing the 95% confidence level (CL)

Table 5.19: Number of observed events in data and estimated background events for the entire mass range. The uncertainties quoted on the number of background events represent the combined statistical and systematic uncertainty.

Process	$\tau_h \tau_h$	$\mu \tau_h$	$e \tau_h$	$e \mu$
Drell-Yan	8 ± 3	882 ± 127	375 ± 118	321 ± 99
W+Jets	0.1 ± 0.1	916 ± 96	546 ± 86	19 ± 11
Diboson	0.5 ± 0.5	29 ± 7	18.0 ± 4	108 ± 17
$t\bar{t}$	—	26 ± 7	26 ± 8	223 ± 45
Multijet	49 ± 13	122 ± 84	117 ± 72	32 ± 24
Total BG	58 ± 13	1976 ± 180	1082 ± 162	703 ± 113
Observed	55	1807	1113	728

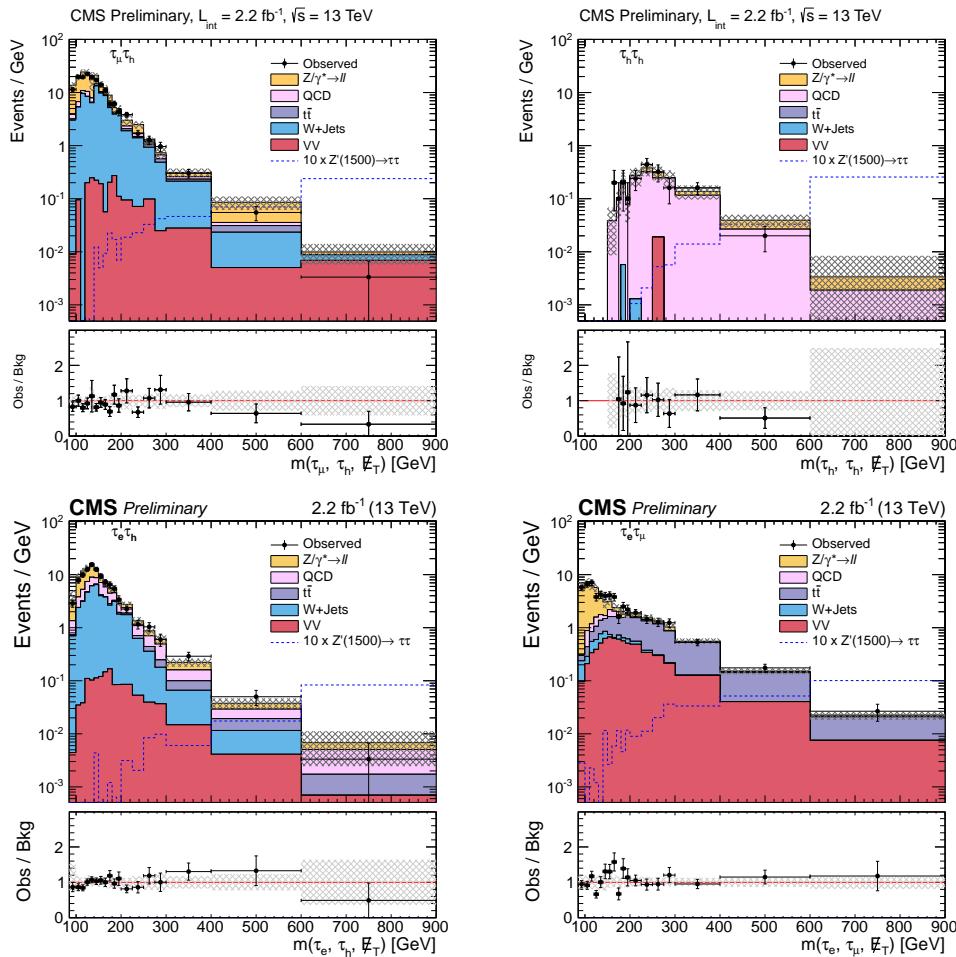


Figure 5.49: Effective mass plots in the signal region after unblinding. Top Left: $m(\mu, \tau_h, \cancel{E}_T)$. Top Right: $m(\tau_h, \tau_h, \cancel{E}_T)$. Bottom Left: $m(e, \tau_h, \cancel{E}_T)$. Bottom Right: $m(e, \mu, \cancel{E}_T)$. All plots in log scale.

Table 5.20: Number of observed events in data and estimated background events in the region $M(\tau_1, \tau_2, \cancel{E}_T) > 300$ GeV. The uncertainties quoted on the number of background events represent the combined statistical and systematic uncertainty.

Process	$\tau_h \tau_h$	$\mu \tau_h$	$e \tau_h$	$e \mu$
Drell-Yan	5 ± 2	16 ± 4	9 ± 3	4 ± 2
W+Jets	0.004 ± 0.004	23 ± 9	17 ± 7	1 ± 1
Diboson	0.02 ± 0.02	6 ± 3	2 ± 1	23 ± 4
$t\bar{t}$	—	4 ± 2	5 ± 1	65 ± 13
Multijet	18 ± 6	4 ± 3	6 ± 2	1 ± 1
Total BG	23 ± 6	51 ± 11	39 ± 8	94 ± 14
Observed	20	42	40	96

upper limit on the signal cross-section using the asymptotic CL_s method. Said differently, a shape based analysis is performed, using the $m(\tau_1, \tau_2, \cancel{E}_T)$ distribution as the fit discriminant to determine the likelihood of observing signal in the presence of the predicted background rate, given the observed yield in data. Systematic uncertainties are represented by nuisance parameters, which are profiled, assuming a log normal prior for normalization parameters, and Gaussian priors for mass-spectrum shape uncertainties.

The above procedure is performed using the Higgs limit calculation tool “combine”[77]. The tool takes as input data cards with the yields and nuisance parameters in each $m(\tau_1, \tau_2, \cancel{E}_T)$ bin of the search channels. These data cards are provided for the four channels being considered. The only further input that is required is the correlations within and across channels. As only an example, if all the channels had 10 bins from $0 < m(\tau_1, \tau_2, \cancel{E}_T) < 5000$ GeV (500 GeV/bin), this means there will be 10 cards per channel and therefore 40 cards in total for the four channels (if all the channels have the same bin size). The cards corresponding to a specific final state, e.g. $\mu \tau_h$, were then combined using the “CombineCards.py” tool provided by the Higgs limit tool, resulting in a single *combined* data card. This procedure was performed for all the final states considered in the analysis, resulting in X *combined* cards. The individual limits, per channel, were obtained by running the combine tool over each *combined* card separately. The final combined limit was obtained by combining the four resulting *combined* cards per channel described above, using the

“CombineCards.py” tool. In order to handle correlations within and across channels, the following approach was used. Each nuisance parameter was defined with a convention of two indices: the first index, i , denoted the channel ($i = \mu\tau_h = 0$, $i = \tau_h\tau_h = 1$, $i = e\tau_h = 2$, $i = e\mu = 3$) and the second one the type of process ($j = \text{Signal} = 0$, $j = \text{W+Jets} = 1$, etc.). Since the limit tool handles nuisance parameters with the same name as fully correlated, correlations across channels and processes were specified by utilizing the same i and process index j , respectively.

Figure 5.50 shows the expected limits as well as the theoretical cross-section as a function of $m(Z')$ mass for each channel. Figure 5.51 show the combined limit. A k factor of 1.3 has been used to scale the leading order (LO) signal cross-section. We exclude Z' (decaying through $Z' \rightarrow \tau\tau$) masses below approximately 2.1 TeV (where the theory line crosses the observed line). Due to a downward fluctuation in data, the observed limit is approximately 200 GeV higher than the expected limit.

Table 5.21: Event summary table after signal region selection

Process	$\tau_h\tau_h$	$\mu\tau_h$	$e\tau_h$	$e\mu$
Z' (500)	307.4 ± 35.3	502.3 ± 57.7	197.6 ± 22.7	218.6 ± 27.3
Z' (1000)	34.6 ± 2.6	40.8 ± 3.1	14.7 ± 1.1	19.0 ± 1.5
Z' (1500)	6.6 ± 0.3	7.2 ± 0.3	2.3 ± 0.1	3.6 ± 0.2
Z' (2000)	1.6 ± 0.07	1.8 ± 0.08	0.59 ± 0.03	0.91 ± 0.04
Z' (2500)	0.55 ± 0.02	0.60 ± 0.02	0.19 ± 0.01	0.30 ± 0.01
Z' (3000)	0.13 ± 0.01	0.14 ± 0.01	0.04 ± 0.00	0.07 ± 0.00
Drell-Yan	8.4 ± 3.1	882.4 ± 127.0	375.1 ± 117.6	321.2 ± 99.2
W+Jets	0.1 ± 0.1	916.2 ± 96.1	545.8 ± 85.6	18.9 ± 11.4
Diboson	0.5 ± 0.5	29.2 ± 7.4	18.0 ± 4.4	108.3 ± 17.4
$t\bar{t}$	–	26.1 ± 6.7	26.1 ± 7.5	222.8 ± 44.8
Multijet	48.7 ± 13.0	121.8 ± 83.5	116.7 ± 71.5	31.9 ± 24.3
Total BG	57.7 ± 13.4	1975.7 ± 240.6	1081.7 ± 162.3	703.1 ± 113.4
Observation	55	1807	1113	728

5.10.2 8 TeV Results

The search for $Z' \rightarrow \tau\tau$ events was also carried out during the 2012 run of the LHC, when the center-of-mass collision energy was lower at $\sqrt{s} = 8$ TeV. 19.7 fb^{-1} of collision data were

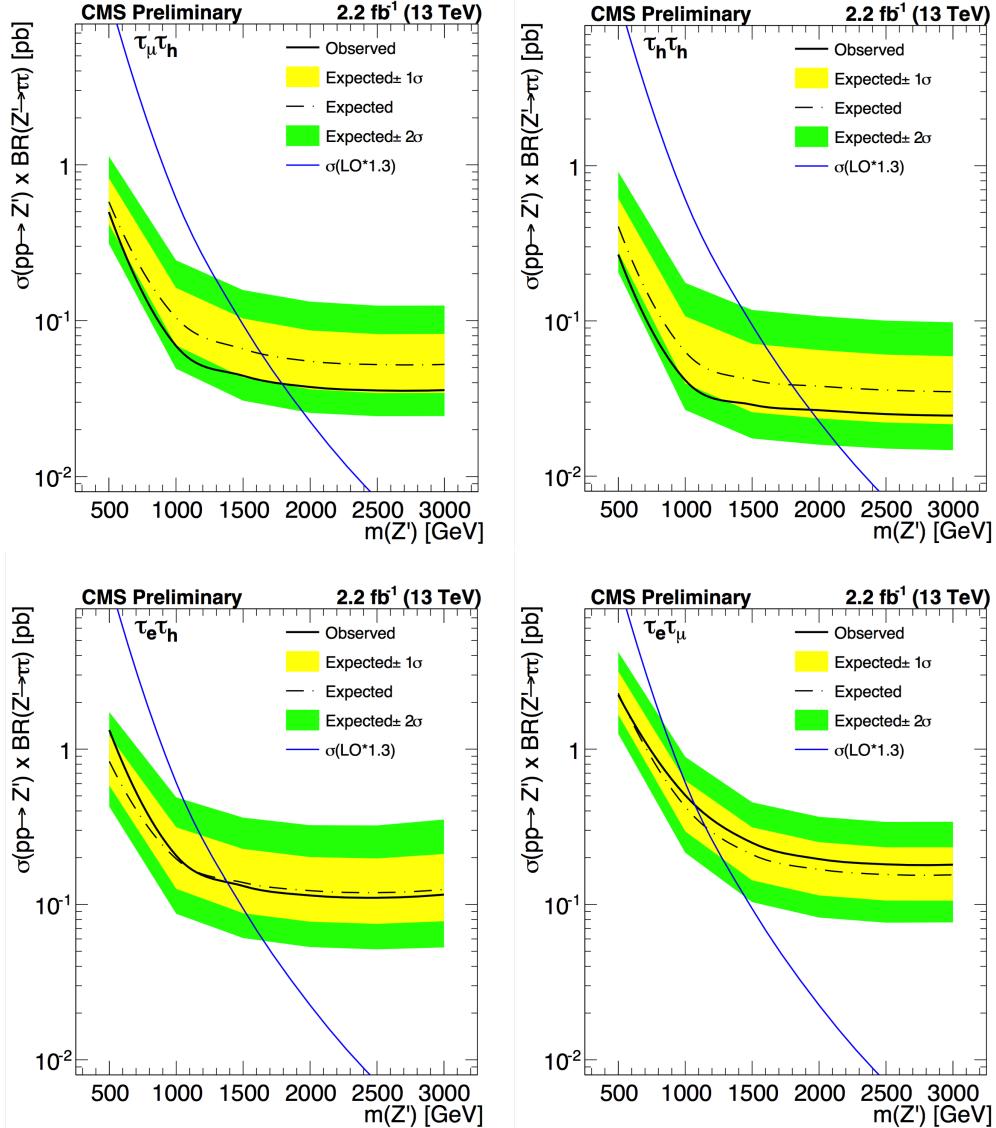


Figure 5.50: Expected and observed limits for the $\tau_\mu\tau_h$, $\tau_h\tau_h$, $\tau_e\tau_h$, and $\tau_e\tau_\mu$ channels for the 13 TeV search. A k factor of 1.3 has been used to scale the leading order (LO) signal cross-section.

collected during this run. Due to poor performance of the Tau POG recommended hadronic tau identification and reconstruction algorithms, the only channel which was studied to completion was the $e\mu$ channel. As the analysis strategy, trigger studies, and background estimation techniques were nearly identical to those presented for the $e\mu$ channel at $\sqrt{s} = 13$ TeV, only the results of the 8 TeV $e\mu$ analysis are shown here.

Figure 5.52 and Table 5.22 show the unblinded signal region and background rates after all

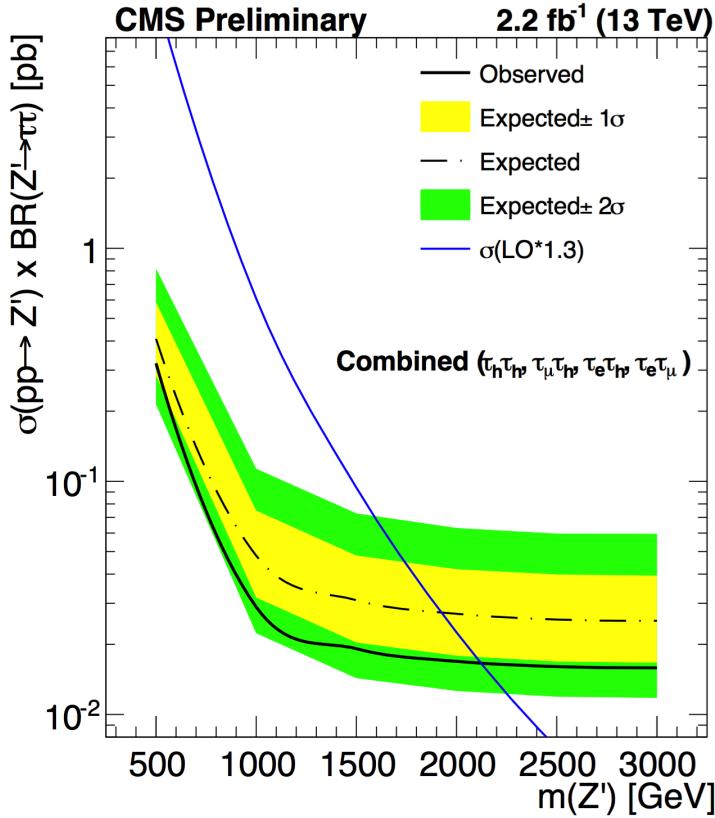


Figure 5.51: Combined expected limit for the $\tau_\mu\tau_h$, $\tau_h\tau_h$, $\tau_e\tau_h$, $\tau_e\tau_\mu$ channels for the 13 TeV search. A k factor of 1.3 has been used to scale the leading order (LO) signal cross-section.

backgrounds are estimated. Data agrees with SM expectation, so limits are set on the Z'_{SSM} visible mass and the E_6 -inspired Z'_ψ visible mass.

Table 5.23 shows the predicted rates in the SR of several Z'_{SSM} signal mass points considered in this analysis.

Finally, the limit on a Z'_{SSM} decaying to two taus are given in Figure 5.53. Upper limits are placed on $\sigma(pp \rightarrow Z') \times BR(Z' \rightarrow \tau\tau)$ as a function of mass. The limit is set as the point at which the experimental value of $\sigma(pp \rightarrow Z') \times BR(Z' \rightarrow \tau\tau)$ exceeds the theoretical value. Below this point, we exclude the existence of Z' -like particles decaying to tau pairs. Figure 5.53 shows that, at 95% confidence level, we exclude the Z'_{SSM} below a mass of 1300 GeV, and that we exclude the Z'_ψ below a mass of 810 GeV.

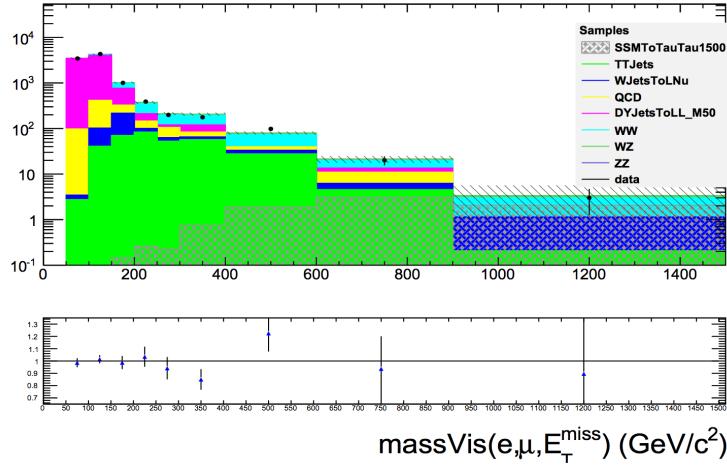


Figure 5.52: $m(e, \tau_h, E_T)$ distribution for the SR in the 8 TeV search. A 1.5 TeV Z'_{SSM} sample is also included.

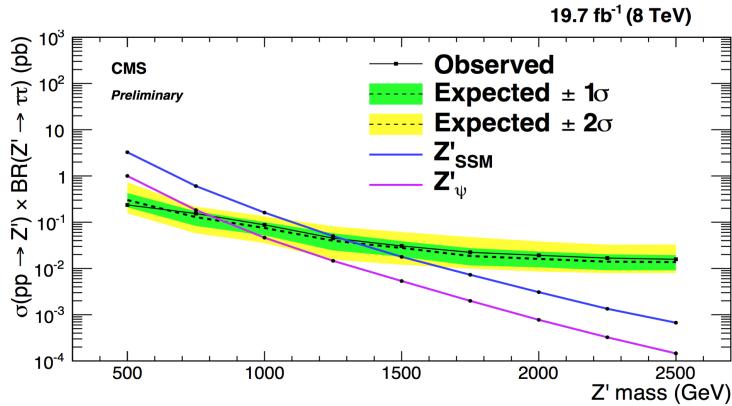


Figure 5.53: 8 TeV 95% CL upper limit on $\sigma(pp \rightarrow Z') \times BR(Z' \rightarrow \tau\tau)$ as a function of Z' mass. The color bands on the expected limits represent one standard deviation (green) and two standard deviations (yellow).

Table 5.22: Expected background rates compared to observed rates in bins of $m(e, \tau_h, E_T)$. Uncertainties are statistical in each bin.

$m(e, \tau_h, E_T)$ (GeV)	[50,100]	[100,150]	[150,200]	[200,250]	[250,300]
BG expected in SR	3480 ± 120	4230 ± 120	1010 ± 40	375 ± 23	210 ± 14
Data observed in SR	3430 ± 60	4300 ± 70	1000 ± 30	388 ± 20	198 ± 14
$m(e, \tau_h, E_T)$ (GeV)	[300,400]	[400,600]	[600,900]	[900,1500]	[0,1500]
BG expected in SR	207 ± 13	80 ± 5	21 ± 4	3.3 ± 2.3	9620 ± 180
Data observed in SR	176 ± 13	98 ± 10	20 ± 4	3.0 ± 1.7	9600 ± 100

Table 5.23: Expected signal rates and associated significances in bins of $m(e, \tau_h, \not{E}_T)$.

$m(e, \tau_h, \not{E}_T)$ (GeV)	[300,400]	[400,600]	[600,900]	[900,1500]	[0,1500]
Z'_{SSM} (750) Sig expected in SR $s/\sqrt{s+b}$	49 ± 3 3.19	76 ± 4 5.91	25.2 ± 2.2 3.83	0.6 ± 0.3 0.31	199 ± 6 2.01
Z'_{SSM} (1000) Sig expected in SR $s/\sqrt{s+b}$	10.3 ± 0.7 0.74	22.6 ± 1.1 2.13	19.2 ± 1.0 3.15	1.6 ± 0.3 0.71	62.3 ± 1.8 0.63
Z'_{SSM} (1250) Sig expected in SR $s/\sqrt{s+b}$	2.4 ± 0.2 0.17	6.2 ± 0.3 0.63	8.2 ± 0.4 1.59	3.2 ± 0.2 1.25	21.6 ± 0.6 0.22
Z'_{SSM} (1500) Sig expected in SR $s/\sqrt{s+b}$	0.8 ± 0.1 0.74	1.9 ± 0.1 2.13	3.1 ± 0.1 3.15	2.1 ± 0.1 0.71	8.4 ± 0.2 0.63
Z'_{SSM} (1750) Sig expected in SR $s/\sqrt{s+b}$	0.2 ± 0.0 0.02	0.7 ± 0.0 0.07	1.1 ± 0.1 0.26	1.3 ± 0.1 0.58	3.5 ± 0.1 0.04

Chapter 6

Tau Lifetime Studies

6.1 Motivation

One of the most significant challenges faced in searches for heavy resonances decaying to tau pairs is the presence of neutrinos in the tau decays. Hadronic tau decays produce one neutrino, and leptonic tau decays produce two neutrinos so that, depending on the channel studied, each ditau event may have two, three, or four neutrinos present in the decay products. These neutrinos do not interact with the detector, and carry energy away from the event. Information about the neutrinos can only be inferred from \cancel{E}_T , an event-level (as opposed to particle-level) quantity which estimates the net (event-wide) neutrino momentum in the transverse direction. If, for example, two neutrinos are produced back-to-back, the only recoverable information about them is the net difference in transverse momentum. Given that the taus, and by extension their decay products, are generated back-to-back (in the transverse plane) in Z' decays, \cancel{E}_T alone does not provide sufficient information to precisely model the di-tau mass. The current mass estimator used in the 8 TeV and 13 TeV $Z' \rightarrow \tau\tau$ searches, $m(\tau_1, \tau_2, \cancel{E}_T)$, depends on \cancel{E}_T and, as shown in Figure 6.1, suffers due to this loss of information.

One proposal to improve the discriminating power of these searches is to add additional selection criteria taking advantage of the lifetime of the tau. At 2.9×10^{-13} s, the tau lifetime is quite short, but it is still long enough to distinguish decay products originating from the primary vertex (PV) from those originating from the tau decay vertex. Essentially, “prompt” particles coming directly from the PV leave tracks that may be traced back to the PV, while tracks from

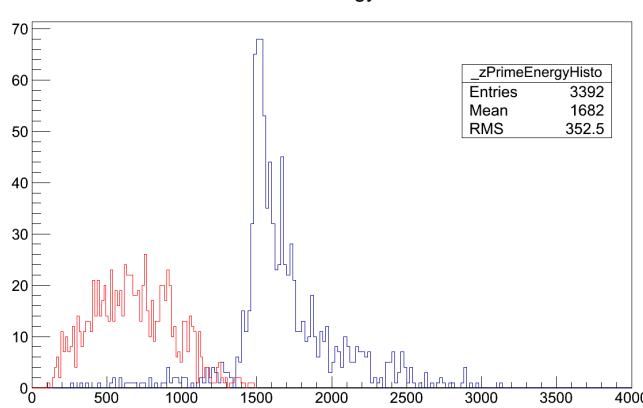


Figure 6.1: 1.5 TeV Z' generated energy (blue) compared to reconstructed visible mass, $m(\tau_1, \tau_2, \cancel{E}_T)$ (red). The $m(\tau_1, \tau_2, \cancel{E}_T)$ distribution is broader and peaked at a much lower energy than the MC truth.

particles coming from tau decays will “miss” the PV due to the distance the tau traveled from the PV before decaying. This concept is illustrated in Figure 6.2. These additional selection criteria, referred to as “lifetime cuts,” are based on tracking information collected from the tau decay products.

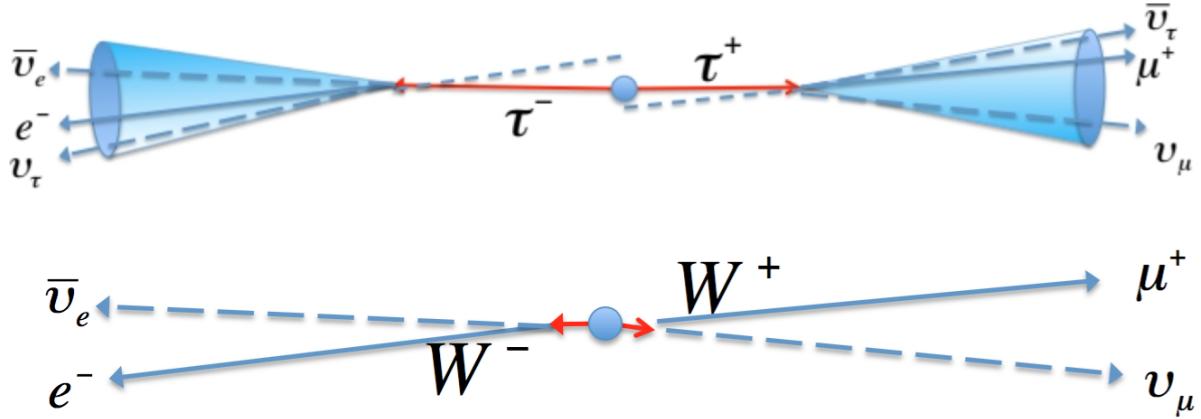


Figure 6.2: A ditau event (top) will produce decay products whose tracks miss the PV, whereas leptonic decays from a shorter-lived parent produced in the pp collision, such as a W^+W^- event (bottom), will produce “prompt” leptons which can be traced back to the PV.

6.2 Methods

Efforts to study the efficacy of the tau lifetime cuts began in the $e\mu$ channel with a simple cut on the individual impact parameters (IPs) of the electron and muon in each ditau event. In this case, IP is defined as the closest distance between the track and a point with the x- and y-coordinates of the beam spot (BS) and the z-coordinate of the PV. The beam spot is a fixed quantity indicating the coordinates of the beam, and the PV is a quantity that is reconstructed from tracks in each event. In this channel, the principal prompt backgrounds are $t\bar{t}$, where the b quarks from t decays decay leptonically into an electron and muon, and W^+W^- , where the W s decay leptonically into an electron and muon. Drell-Yan events where virtual Z bosons decay into tau pairs are in principle an irreducible background, but the rate falls off significantly in the high-mass (signal) region being considered.

The cut was first defined as the sum of the absolute values of the IPs of each lepton:

$$\text{Lifetime}_\tau = |\text{IP}_e| + |\text{IP}_\mu| \quad (6.1)$$

To account for the large variance in track resolution, and thus preferentially-select “clean” tau decays, the lifetime definition was modified to include the track measurement error:

$$\text{Lifetime}_\tau = \sqrt{\frac{(|\text{IP}_e| + |\text{IP}_\mu|)^2}{\sigma_{\text{IP}_e}^2 + \sigma_{\text{IP}_\mu}^2}} \quad (6.2)$$

where the error on the track IP is defined according to the recommendations laid down by the Tracker POG[58]:

$$\sigma_e = \text{abs}(\text{theGSFElec.gsfTrack}() \rightarrow \text{dxy}(\text{theBeamSpot})) \quad (6.3)$$

$$\sigma_\mu = \text{abs}(\text{thePatMuon.track}() \rightarrow \text{dxy}(\text{theBeamSpot})) \quad (6.4)$$

The cut is placed at the very end of the selection sequence, immediately following the b-jet veto. The quantity in Equation 6.2 is required to exceed a value chosen based on optimization

studies. These studies, based in signal and background MC, compare the acceptance of signal events with the rejection of background events across several values of this threshold. The threshold value is chosen based on its associated value of $\frac{s}{\sqrt{s+b}}$, where s is the signal rate and b is the aggregated background rate. Figure 6.3 shows the performance of the IP-based lifetime cut in MC. A cut threshold of 2 is chosen on the basis of maximizing both $\frac{s}{\sqrt{s+b}}$ and signal MC acceptance.

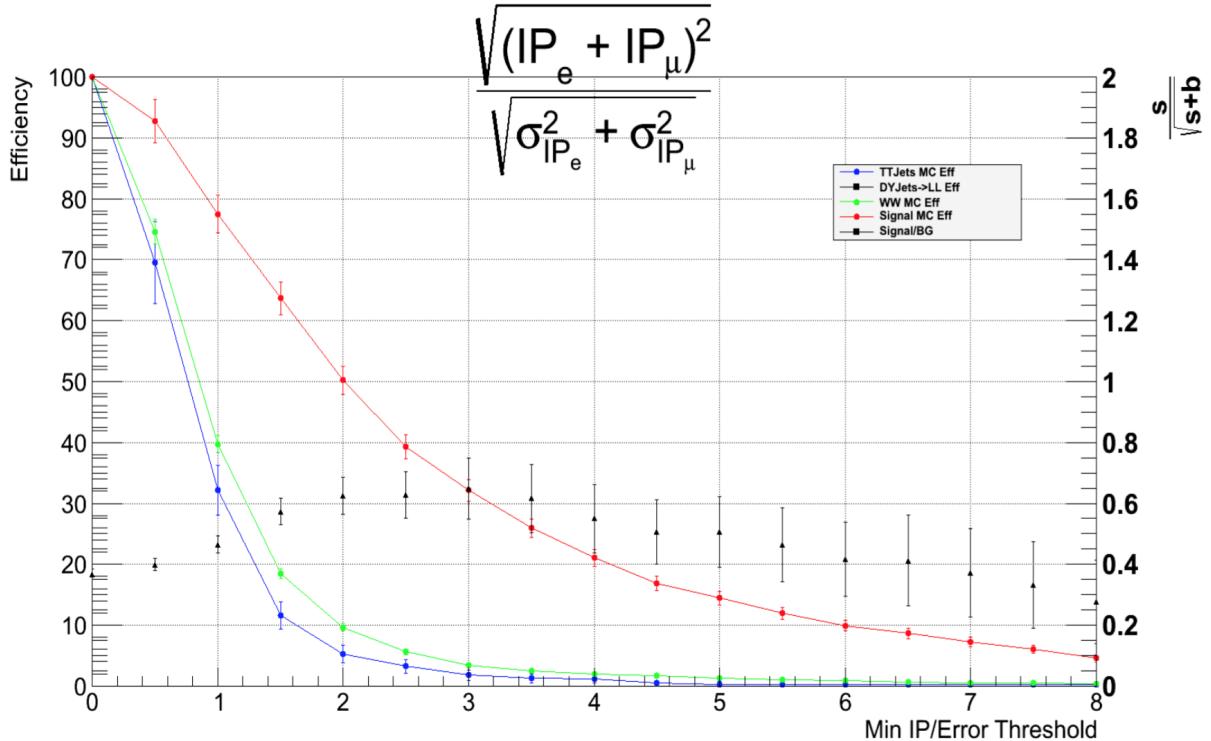


Figure 6.3: Plot showing the performance of the IP definition of the lifetime cut. The colored lines indicate MC efficiencies and correspond to the left axis. The black triangles indicate $\frac{s}{\sqrt{s+b}}$ and correspond to the right axis. A threshold value of 2 was chosen for the cut since it is the first value on the plateau of $\frac{s}{\sqrt{s+b}}$ maxima, thereby keeping signal MC acceptance as high as possible.

Another cut exploiting tau lifetime is one based on the distance-of-closest approach (DCA) between the two lepton tracks. DCA is defined to be the transverse distance between the two tracks at their point of closest approach. DCA is also divided by its error, where error on DCA is again defined according to the Tracker POG recommendation[58]:

$$\sigma_{DCA} = \sqrt{\frac{D\vec{C}A^T \times M_\sigma \times D\vec{C}A}{|D\vec{C}A|^2}} \quad (6.5)$$

where M_σ is a 3x3 covariance matrix containing the x, y, and z-errors summed over the electron and muon points of closest approach. Figure 6.4 shows the performance of the DCA/σ cut.

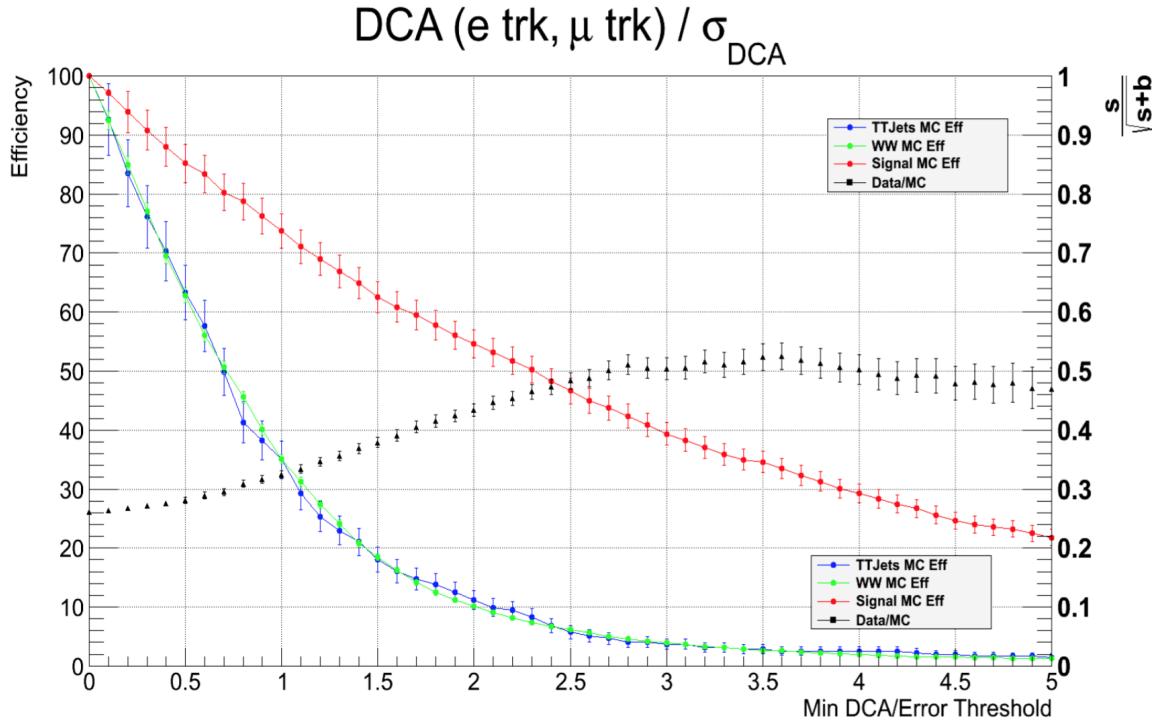


Figure 6.4: Plot showing the performance of the DCA definition of the lifetime cut. The colored lines indicate MC efficiencies and correspond to the left axis. The black triangles indicate $\frac{s}{\sqrt{s+b}}$ and correspond to the right axis. A threshold value of 2.5 was chosen for the cut since it is the first value on the plateau of $\frac{s}{\sqrt{s+b}}$ maxima, thereby keeping signal MC acceptance as high as possible.

The DCA cut has the potential to veto valid signal events in cases where the two lepton tracks are far from the BS/PV but close to one another (illustrated in Figure 6.5). In such cases, the IP-based cut would still retain the signal event. Therefore, the final definition of the lifetime cut is the requirement that events pass the IP-based cut **OR** the DCA cut. The performance of this “OR” cut is shown in Figure 6.6.

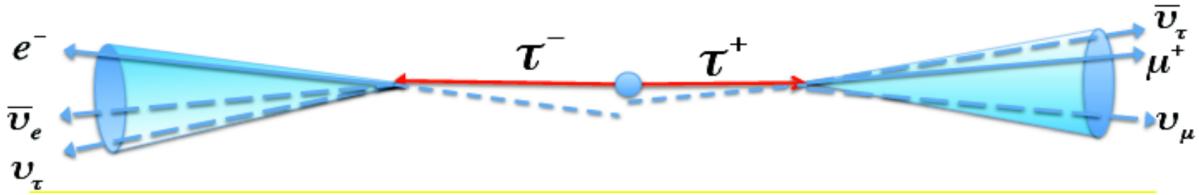


Figure 6.5: Example of where the DCA cut could reject a valid signal event - the two lepton tracks have a low DCA despite a high IP

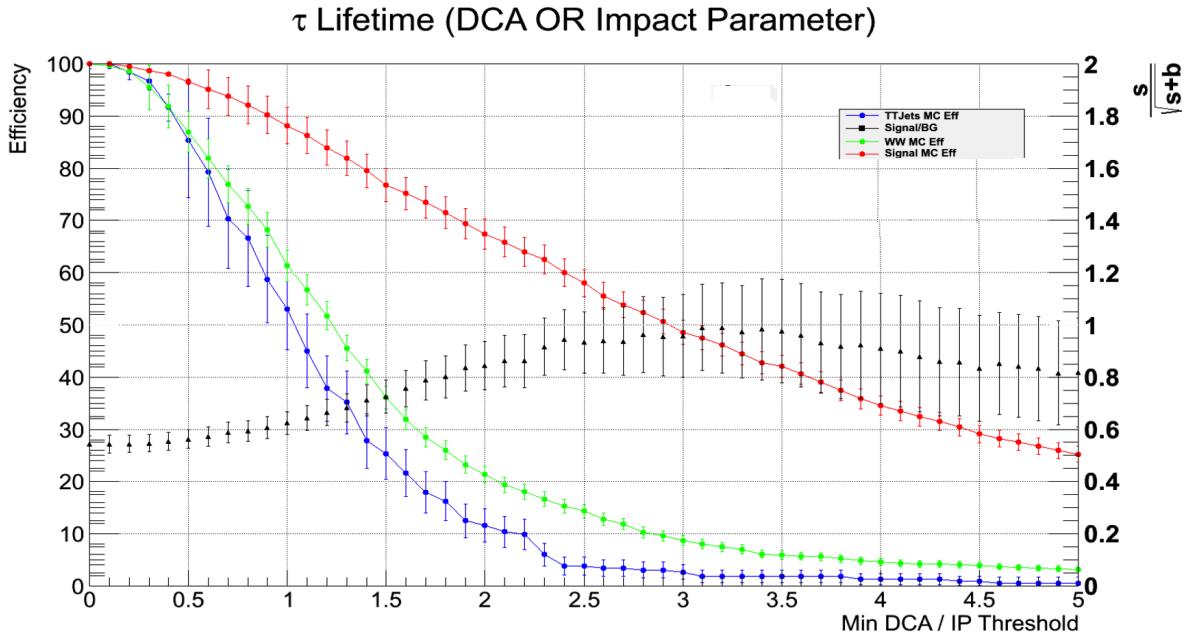


Figure 6.6: Plot showing the performance of the DCA **OR** IP-based definition of the lifetime cut. The colored lines indicate MC efficiencies and correspond to the left axis. The black triangles indicate $\frac{s}{\sqrt{s+b}}$ and correspond to the right axis. A threshold value of 2 was chosen for the cut since it is the first value on the plateau of $\frac{s}{\sqrt{s+b}}$ maxima, thereby keeping signal MC acceptance as high as possible. Note that the performance in this cut, in terms of signal acceptance and background rejection, is higher than either the IP-based or DCA cuts individually.

6.3 Impact on Limit

6.3.1 8 TeV Results in the $e\mu$ Channel

Figure 6.7 shows the impact of the addition of the lifetime cut (“OR” definition) on the overall limit. Note that the limit without the lifetime cut shown in Figure 6.7 is different from the final limit published in EXO-12-046 and shown in Section 5.10. This is due to the fact that the lifetime

study was performed early in the 8 TeV analysis when a different technique was being used for limit-setting. Namely, the limits shown here were produced with the Asymptotic CL_s method[77] using one datacard per background as input, while the final limits shown in EXO-12-046 used the full CL_s method[77] using one datacard per bin for each background. Despite the difference, it is still apparent that the addition of the lifetime cut improves the limit by ~ 70 GeV.

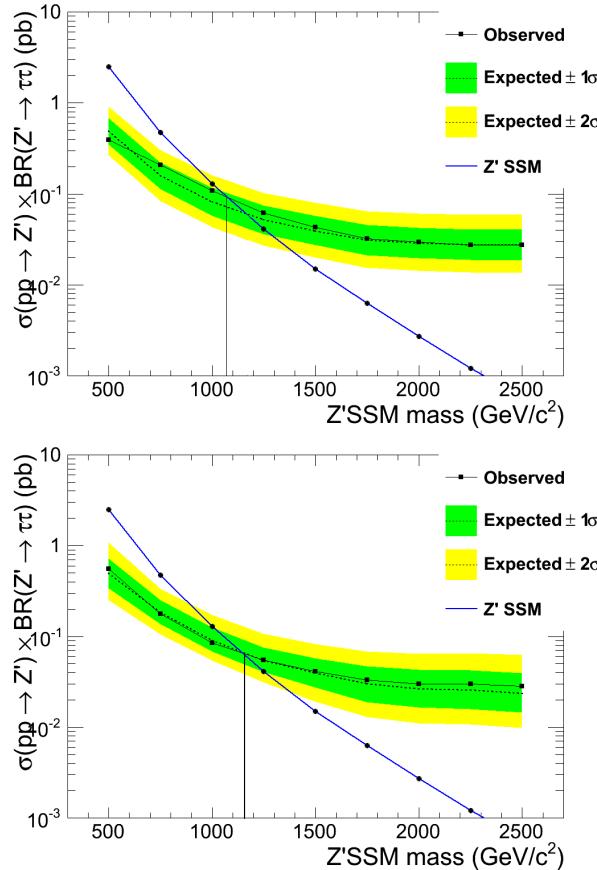


Figure 6.7: Limit calculated with standard 8 TeV selection sequence (top) compared with limit calculated with the standard selection sequence and a lifetime threshold of 2 applied (bottom). The addition of the lifetime cut improves the exclusion limit by about 70 GeV.

6.3.2 13 TeV Results in the $\tau_h\tau_h$ Channel

For the 13 TeV analysis, the focus of the lifetime study is shifted to the fully-hadronic $\tau_h\tau_h$ channel, where the leading charged pion tracks from the hadronic tau decays are used in place of the electron and muon tracks. QCD is the overwhelmingly dominant background in this channel

(see Section 5.6). Since QCD MC is not available, a full optimization study to determine the appropriate lifetime cut value is more difficult. Instead of a plot like Figure 6.6, the distributions of combined IP/ σ are compared for a 2.5 TeV $Z' \rightarrow \tau_h\tau_h$ sample in the signal region and for data in the same-sign QCD control region. In this case, data in the SS QCD CR is used as an approximation for QCD, given the relative dominance of QCD in this channel. Figure 6.8 shows this comparison, and a lifetime cut threshold of 3 is selected for this analysis.

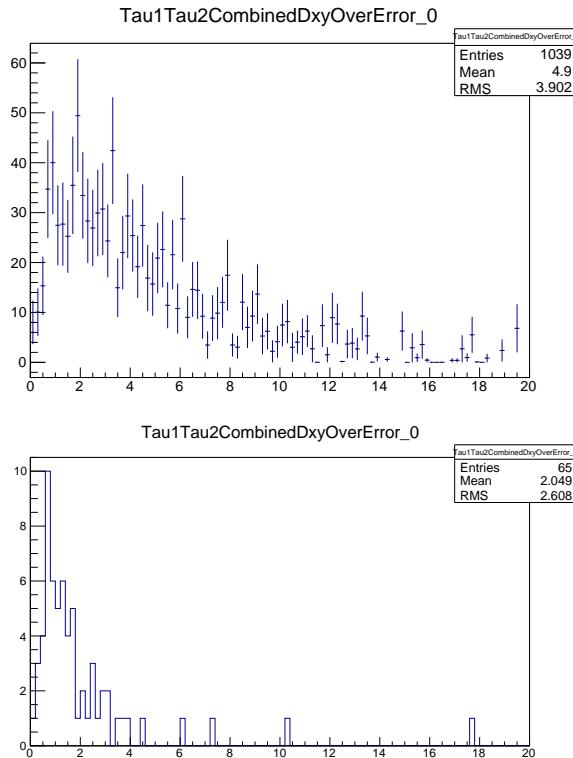


Figure 6.8: Distribution of combined IP/ σ for a 2.5 TeV $Z' \rightarrow \tau_h\tau_h$ sample in the SR (top) and data (used as a proxy for QCD) in the SS QCD CR (bottom). A threshold value of 3 is selected for the limit comparison.

Figure 6.9 shows the impact of the addition of the lifetime cut (“OR” definition) on the overall limit. In this channel and at this collision center-of-mass energy, the addition of the lifetime cut appears to decrease both the expected and observed limits by ~ 100 GeV. It’s likely that the extremely low event rates in the high mass region (i.e. the search region with the highest sensitivity) is what drives this behavior. Since the background rate in the high mass region is already quite

low, the fact that the lifetime cut rejects $\sim 50\%$ of the signal is likely to have a larger impact on the limit than boosting the signal-to-background ratio. In an effort to increase signal events in this region, the p_ζ cut was removed. However, this alone permitted a significant increase in QCD background, so the \cancel{E}_T cut was tightened to $\cancel{E}_T > 50$. The results are shown in Figure 6.10. These modifications appear to increase the baseline limit. However, even with increased statistics, the addition of the lifetime cut appears to decrease the limit relative to the limit without the lifetime cut. The conclusion is that, in the high-mass region, the loss of signal events due to the lifetime cut affects the limit sensitivity more than the concurrent suppression of background. The lifetime cuts will once again be tested as statistics accumulate during the 2016 run and beyond. Multivariate analysis (MVA) techniques will also be applied so that these new cuts can be fully integrated and properly “tuned” with the rest of the $Z' \rightarrow \tau\tau$ selection cuts.

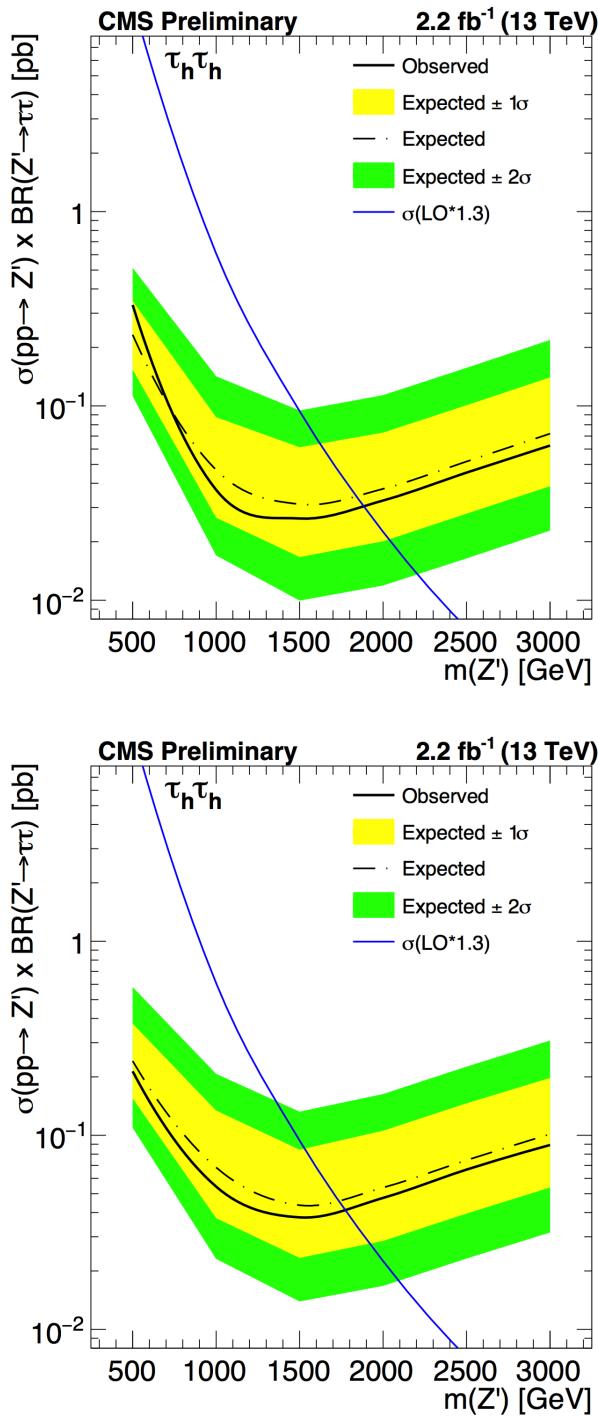


Figure 6.9: Limit calculated with standard 13 TeV selection sequence (top) compared with limit calculated with the standard selection sequence and a lifetime threshold of 3 applied (bottom).

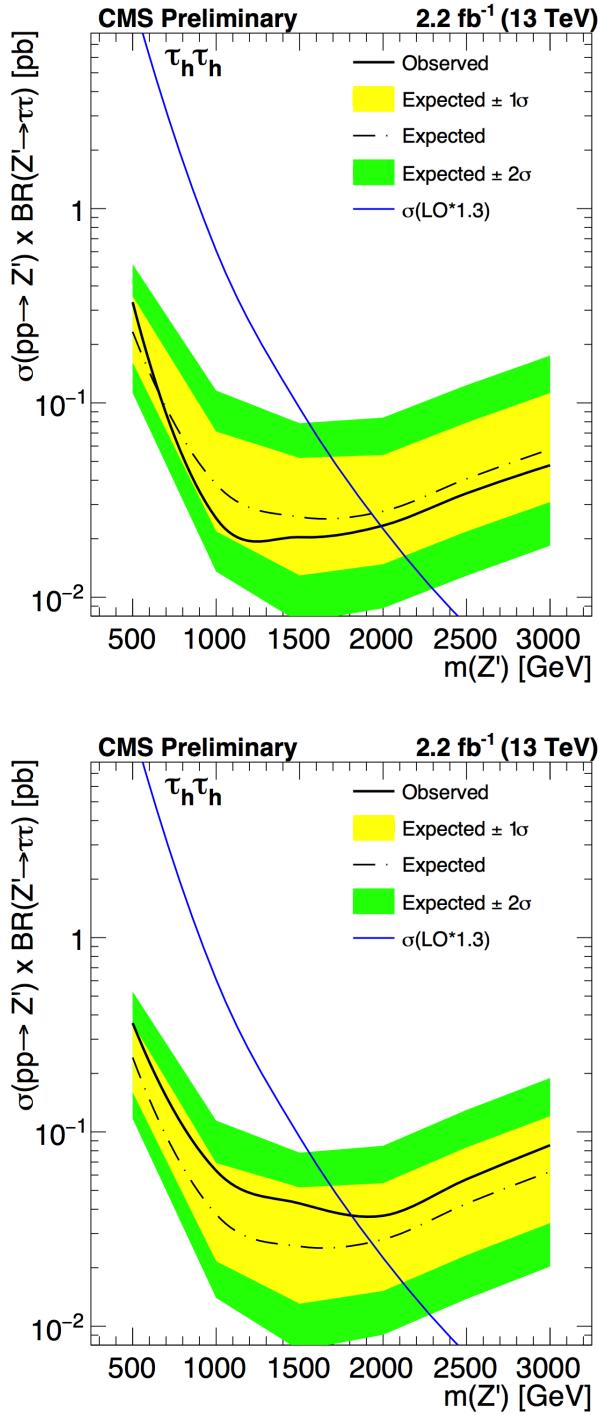


Figure 6.10: Limit comparison with the p_ζ requirement removed and the \cancel{E}_T minimum threshold increased to 50. The top plot has no lifetime threshold applied, while the bottom plot has a lifetime threshold of 2 applied. While the removal of p_ζ and the tightening of the \cancel{E}_T threshold increases the baseline limit, the limit with the lifetime cut added is still lower (worse) than that without it applied.

Chapter 7

Summary, Conclusion, and What's Next for $Z' \rightarrow \tau\tau$

This thesis has presented a brief overview of the Standard Model of particle physics, followed by a description of the CMS experiment at the Large Hadron Collider, followed by an in-depth description of the work done by the author and his collaborators over two major multi-year, multi-channel searches for new particles at CMS. Both analyses were framed with high mass neutral resonances, Z' 's, decaying to $\tau^+\tau^-$ pairs as the searched-for signal in question. Motivated by multiple beyond-Standard Model theories, namely those that add an additional U(1) symmetry to the SM, these searches focused on the $Z' \rightarrow \tau\tau$ decay channel.

The first search used 19.7 fb^{-1} of data collected during the 2012 CMS run at the LHC, during which time the center-of-mass energy was $\sqrt{s} = 8 \text{ TeV}$. With a combination of Monte Carlo-based and data driven background estimation techniques, the distribution of ditau visible mass + \cancel{E}_T was studied and no excess was observed. Due to issues related to reconstruction of hadronic taus, limits were only able to be set in the $e\mu$ channel. At 95% confidence level, no ditau excess decaying to an electron and muon is observed below 1.3 TeV for the Z'_{SSM} model and below 810 GeV for the E_6 -inspired model.

The second search used 2.2 fb^{-1} of data collected during the 2015 CMS run at the LHC, during which time the center-of-mass energy was $\sqrt{s} = 13 \text{ TeV}$. This analysis considered all four decay channels, with the bulk of the author's work focused on the $\tau_h\tau_h$ channel. Similar techniques for background estimation as used in the 8 TeV analysis were employed here, and once again the distribution of ditau visible mass + \cancel{E}_T was studied. No excess was observed, and limits were set on

each individual channel. The combined limit across all channels excludes a Z'_{SSM} boson decaying to pairs of taus below 2.1 TeV at 95% confidence level, the highest limit observed for a Z' decaying to taus. This analysis has been documented in a CMS analysis note (AN), a CMS physics analysis summary (PAS), and is currently undergoing collaboration-wide review (CWR) for publication as a CMS paper.

In addition to the core analyses, the author has devoted a large portion of his research career towards the development of additional selection criteria dedicated towards improved identification of tau leptons. This work was motivated by a central hypothesis that the tau, while having quite a short lifetime on the order of 10^{-13} s, is sufficiently long-lived to make its decay products, electrons, muons, and hadronic jets, distinguishable from the background electrons, muons, and jets originating from the primary vertex. This difference would manifest itself in those charged tracks originating from tau decays having a larger impact parameter with respect to the primary vertex (and later the beam spot) than those from “prompt” decays. This concept was first tested in the $e\mu$ channel with the 2012 8 TeV data, first with the simple sum of electron and muon impact parameters and then with more intricate combinations taking the track uncertainties into account. The effectiveness of cutting on the distance-of-closest-approach between the tracks rather than the impact parameter was studied, and the conclusion was reached that the two approaches yielded the highest overall signal-to-noise ratio when used in concert, requiring events to pass one or the other in the selection process. The application of these “lifetime cuts” led to a significant suppression of prompt backgrounds and subsequent improvement of the 8 TeV $e\mu$ limit by nearly 100 GeV. This study was again applied to the $\tau_h\tau_h$ channel in the 2015 13 TeV analysis, where the same cut definitions were modified to act on the leading charged tracks in hadronic tau candidates. Similar performance was found in terms of prompt background suppression, but the sacrifice of signal events in the high-mass region (the region most sensitive to the $Z' \rightarrow \tau\tau$ search) was severe enough that the limit suffered. Despite this, the lifetime cuts still have significant potential for future analyses involving taus.

The LHC and CMS have each performed beautifully this year. As of this writing, the scientists

and engineers running the accelerator and the detector have recorded more than 30 fb^{-1} of collision data for the 2016 run. By the time the LHC switches from pp collisions to heavy ion collisions in early November, the total integrated luminosity will be closer to 40 fb^{-1} [76]. To put this in perspective, the total data collected during the 2015 run was 3.81 fb^{-1} . With this much data, the hope is that sufficient statistics will be available in the space of high-mass taus that future searches will be able to afford the loss in signal events and reap the high rate of background suppression demonstrated by the lifetime cuts. To put it succinctly: given enough events, the lifetime cuts should be quite effective at pushing the limit higher.

In addition to the lifetime cuts, there are additional planned modifications to the $Z' \rightarrow \tau\tau$ search as the analysis is repeated with the 2016 data and beyond. A new series of topological cuts involving \cancel{E}_T have been proposed. Since the neutrinos in the signal come from the tau decays, it is expected that the \cancel{E}_T will be collinear with one of the tau legs (depending on the balance of p_T). One such topological cut requires the cosine of the angle between the \cancel{E}_T and the difference in p_T between the tau legs to be less than -0.9 (\cancel{E}_T and Δp_T pointing in opposite directions). Another requires the cosine of the angle between the \cancel{E}_T and one of the tau legs to be greater than 0.9 (\cancel{E}_T pointing along one of the tau decays). Finally, one recently-proposed cut would require the transverse mass of one of the tau legs and the \cancel{E}_T to be greater than 150 GeV. Each of these cuts is designed to suppress $t\bar{t}$, QCD, and W+Jets backgrounds, as these are all processes in which the \cancel{E}_T is not necessarily collinear with the leptons or jets faking tau decays.

The current LHC schedule cites a plan to increase the pp center-of-mass collision to $\sqrt{s} = 14 \text{ TeV}$, the full design energy of the collider by 2018. By the end of 2017, 100 fb^{-1} is expected to be delivered to both CMS and ATLAS by the LHC. The LHC will then undergo its second long shutdown from mid-2018 to 2019, during which time upgrades to the injector and cryogenics systems are planned. By the end of 2022, a full 300 fb^{-1} of collision data is expected to have been delivered. The third long shut down, from 2022 to 2025, will feature a host of upgrades to the LHC, CMS, and ATLAS in preparation for “Phase 2” of LHC operation, also known as “HL-LHC” (High-Luminosity LHC)[65]. During this phase, the LHC is expected to operate at 5- to 7-times

the nominal luminosity. To handle this significant increase in luminosity, the Phase 2 upgrade for CMS in particular will involve several upgrades. These include a new pixel and strip trackers with integrated triggering, new endcap calorimeters in both ECAL and HCAL, a new L1 trigger system with improved latency, and new high- η muon stations[81]. In particular, the CMS upgrade includes four additional pixel stations expected to deliver much higher impact parameter resolution. This has tremendous potential for the lifetime study as it could offer much better discrimination between tracks originating from tau decays and those originating from prompt vertices. By 2035, the HL-LHC is expected to deliver up to 3000 fb^{-1} of data[65].

Beyond the LHC, there are several additional colliders planned for construction around the world. The International Linear Collider (ILC) is a 500 GeV linear e^+e^- collider with a proposal to begin construction in 2019 in Japan. The ILC would serve as a “Higgs factory” with the chief goal of studying in greater detail the mass, spin, couplings, and other properties of the Higgs boson. In direct competition with the ILC is the Compact Linear Collider (CLIC), which is also a linear e^+e^- collider with a proposed collision energy of 3 TeV. Further in the future, proposed colliders include a circular muon collider, which would allow higher lepton collision energies since the heavier muons would lose less energy due to synchotron radiation than the lighter electrons, and Very Large Hadron Collider (VLHC), a 100 TeV pp collider which would open the door to the discovery of new particles with masses an order of magnitude heavier than currently possible with the LHC.

With higher collision energies, new accelerators on the horizon, and thousands of femtobarns of scheduled data-taking, there is a rich and diverse space of unexplored physics to investigate. All of this, together with an abundance of talented young researchers, points to a future in high energy physics that is indeed bright.

Bibliography

- [1] Search for neutral higgs bosons of the minimal supersymmetric standard model decaying to τ pairs in $p\bar{p}$ collisions at $\sqrt{s} = 1.96 TeV$. [Phys. Rev. Lett.](#), 96:011802, 2006.
- [2] Particle-flow event reconstruction in cms and performance for jets, taus, and met. [CMS Physics Analysis Summary](#), CMS-PAS-PFT-09-001, 2009.
- [3] Commissioning of the particle-flow event reconstruction with the first lhc collisions recorded in the cms detector. [CMS Physics Analysis Summary](#), CMS-PAS-PFT-10-002, 2010.
- [4] Performance of muon identification in pp collisions at $\sqrt{s} = 7$ tev. [CMS Physics Analysis Summary MUO-10-002](#), 2010.
- [5] Absolute luminosity normalization. [Detector Performance Summary CMS-DPS-2011-002](#), 2011.
- [6] Missing transverse energy performance of the cms detector. [JINST](#), 6:09001, 2011.
- [7] Performance of b-jet identification in cms. [JINST](#), 8:P04013, 2013.
- [8] Performance of b tagging at $\sqrt{s} = 8$ tev in multijet, $t\bar{t}$ and boosted topology events. [CMS Physics Analysis Summary](#), CMS-PAS-BTV-13-001, 2013.
- [9] Measurement of electroweak production of two jets in association with a z boson in proton-proton collisions at $\sqrt{s} = 8$ tev. [EPJC](#), 75:66, 2015.
- [10] G. Abbiendi et al. Muon reconstruction in the cms detector. [CMS Note](#), AN2008/097, 2008.
- [11] N. Akchurin et al. Cms missing transverse energy performance in photon+xevents for pp collisions at 7 tev. [CMS Note](#), AN2010/432, 2010.
- [12] M. Aldaya, P. Arce, J. Caballero, B. de la Cruz, P. Garcia-Abia, J. M. Hernandez, M. I. Josa, and E. Ruiz. Discovery potential and search strategy for the standard model Higgs boson in the $H \rightarrow ZZ^* \rightarrow 4\mu$ decay channel using a mass-independent analysis. [CMS Note](#) 2006/106, 2006.
- [13] ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, and the SLD Electroweak and Heavy Flavour Groups. Precision electroweak measurements and constraints on the Standard Model. 2010.

- [14] G. Altarelli, B. Mele, and M. Ruiz-Altaba. Searching for new heavy vector bosons in $p\bar{p}$ colliders. *Zeitschrift für Physik C Particles and Fields*, 45(1):109–121, 1989.
- [15] ATLAS Collaboration. Search for the Higgs boson in the $H \rightarrow WW(*) \rightarrow \ell^+\nu\ell^-\bar{\nu}$ decay channel in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. Submitted to *Phys. Rev. Lett.*, 2011.
- [16] ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Phys. Lett. B*, 716:1–29, 2012.
- [17] S. Baffioni et al. Electron reconstruction in cms. *Eur. Phys. J. C* 49, 1099 – 1116, 2007.
- [18] R. Barbier et al. R-parity-violating supersymmetry. *Physics Reports*, 420:1–195, 2005.
- [19] I. Belotelov, I. Golutvin, D. Bourilkov, A. Lanyov, E. Rogalev, M. Savina, and S. Shmatov. Search for ADD extra dimensional gravity in di-muon channel with the CMS detector. CMS Note 2006/076, 2006.
- [20] I. Bertram, G. Landsberg, J. Linnemann, R. Partridge, M. Paterno, and H. B. Prosper. A recipe for the construction of confidence limits. Technical Report TM-2104, Fermilab, 2000.
- [21] A. Brandt et al. Measurements of single diffraction at $\sqrt{s} = 630$ GeV: Evidence for a non-linear $\alpha(t)$ of the pomeron. *Nucl. Phys. B*, 514:3, 1998.
- [22] W. Buchmüller and D. Wyler. Constraints on SU(5)-type leptoquarks. *Phys. Lett. B*, 177:377, 1986.
- [23] Jon Butterworth, Stefano Carrazza, Amanda Cooper-Sarkar, Albert De Roeck, Joël Feltesse, Stefano Forte, Jun Gao, Sasha Glazov, Joey Huston, Zahari Kassabov, Ronan McNulty, Andreas Morsch, Pavel Nadolsky, Voica Radescu, Juan Rojo, and Robert Thorne. Pdf4lhc recommendations for lhc run ii. *Journal of Physics G: Nuclear and Particle Physics*, 43(2):023001, 2016.
- [24] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti- k_t jet clustering algorithm. *JHEP*, 04:063, 2008.
- [25] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti- k_t jet clustering algorithm. *JHEP*, 04:063, 2008.
- [26] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet user manual. *Eur. Phys. J. C*, 72:1896, 2012.
- [27] Marcela Carena, Stefania Gori, NausheenR. Shah, CarlosE.M. Wagner, and Lian-Tao Wang. Light stau phenomenology and the higgs $\gamma\gamma$ rate. *Journal of High Energy Physics*, 2012(7):1–32, 2012.
- [28] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.
- [29] S. Chatrchyan et al. Identification of b-quark jets with the cms experiment. *JINST*, 8:04013, 2013.
- [30] S. Chatrchyan et al. Recommendation for using b-tag objects in physics analyses. *CMS Physics Analysis Summary*, CMS-PAS-BTV-13-001:04013, 2013.

- [31] Serguei Chatrchyan et al. Determination of jet energy calibration and transverse momentum resolution in CMS. *JINST*, 6:P11002, 2011.
- [32] Serguei Chatrchyan et al. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7\text{TeV}$. *JINST*, 7:P10002, 2012.
- [33] CMS Collaboration. CMS Physics: Technical Design Report Volume 1: Detector Performance and Software. 2006.
- [34] CMS Collaboration. CMS technical design report, volume II: Physics performance. *J. Phys. G*, 34:995–1579, 2007.
- [35] CMS Collaboration. Commissioning of the particle-flow event reconstruction with the first LHC collisions recorded in the CMS detector. CMS Physics Analysis Summary CMS-PAS-PFT-10-001, 2010.
- [36] CMS Collaboration. Jet performance in pp collisions at $\sqrt{s}=7\text{ TeV}$. CMS Physics Analysis Summary CMS-PAS-JME-10-003, 2010.
- [37] CMS Collaboration. Search for pair production of first-generation scalar leptoquarks in pp collisions at $\sqrt{s} = 7\text{ TeV}$. Submitted to *Phys. Rev. Lett.*, 2010.
- [38] CMS Collaboration. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Phys. Lett. B*, 716:30–61, 2012.
- [39] CMS Collaboration. Performance of cms muon reconstruction in pp collision events at $\sqrt{(s)} = 7\text{ tev}$. *JINST*, 7:P10002, 2012.
- [40] CMS Collaboration. Performance of cms muon reconstruction in pp collision events at $\sqrt{s} = 7\text{ tev}$. Submitted to *J. Inst.*, 2012.
- [41] CMS Collaboration. Performance of τ -lepton reconstruction and identification in cms. *JINST*, 5:P01001, 2012.
- [42] CMS Collaboration. Measurement of the $b_s^0 \rightarrow \mu\mu$ branching fraction and search for $b^0 \rightarrow \mu\mu$ with the cms experiment. Submitted to *Phys. Rev. Lett.*, 2013.
- [43] CMS Collaboration. The performance of the cms muon detector in proton-proton collisions at $\sqrt{(s)} = 7\text{ tev}$ at the lhc. Submitted to *JINST*, 2013.
- [44] CMS Collaboration. Muon reconstruction and identification improvements for run-2 and first results with 2015 run data. CMS Detector Performance Summary CMS-DP-15-015, 2015.
- [45] CMS Collaboration. Performance of electron reconstruction and selection with the cms detector in proton-proton collisions at $\sqrt{s} = 8\text{ tev}$. *JINST*, 10:P06005, 2015.
- [46] CMS Collaboration. Search for a narrow resonance produced in 13 tev pp collisions decaying to electron pair or muon pair final states. CMS Physics Analysis Summary CMS-PAS-EXO-15-005, 2015.
- [47] CMS Collaboration. Commissioning and performance of the cms pixel tracker with cosmic ray muons. *Journal of Instrumentation*, 5(03):T03007, 2010.

- [48] The CMS collaboration. The performance of the cms muon detector in proton-proton collisions at $\sqrt{s} = 7$ tev at the lhc. *Journal of Instrumentation*, 8(11):P11002, 2013.
- [49] The CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the cms tracker. *Journal of Instrumentation*, 9(10):P10009, 2014.
- [50] B. Dutta, A. Gurrola, W. Johns, T. Kamon, P. Sheldon, and Sinha. Kuver. Vector boson fusion processes as a probe of supersymmetric electroweak sectors at the lhc. *Phys. Rev. D*, 87, 2013.
- [51] Physics Stack Exchange. Feynman diagram for annihilation, 2014.
- [52] R. Frühwirth. Application of kalman filtering to track and vertex fitting. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 262(2):444 – 450, 1987.
- [53] R Frühwirth, Wolfgang Waltenberger, and Pascal Vanlaer. Adaptive Vertex Fitting. Technical Report CMS-NOTE-2007-008, CERN, Geneva, Mar 2007.
- [54] Fabiola Gianotti. Status of standard model higgs searches in atlas. Latest update in the search for the Higgs boson, 2012.
- [55] Federico De Guio and the CMS collaboration. Performance of the cms electromagnetic calorimeter and its role in the hunt for the higgs boson in the two-photon channel. *Journal of Physics: Conference Series*, 455(1):012028, 2013.
- [56] A. Gurrola et al. Search for pair production of scalar top quarks decaying to a tau lepton and a b quark via r-parity violation. *CMS AN-2012/158*.
- [57] Francis Halzen and Alan D. Martin. *Quarks & Leptons*. John Wiley & Sons, Inc., 1984.
- [58] Andre Holzner. Reconstruction of the beam spot, 2011.
- [59] Joe Incandela. Status of the cms sm higgs search. Latest update in the search for the Higgs boson, 2012.
- [60] T. Junk. Confidence level computation for combining searches with small statistics. *Nucl. Instrum. Meth. A*, 434:435, 1999.
- [61] V. Khachatryan et al. Search for physics beyond the standard model in dilepton mass spectra in proton-proton collisions at $\sqrt{s} = 8$ tev. *Journal of High Energy Physics*, (4):1–49, 2015.
- [62] Vardan Khachatryan et al. Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8\text{TeV}$. *JINST*, 10:P06005, 2015.
- [63] Vardan Khachatryan et al. Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at $\sqrt{s} = 8\text{TeV}$. *JINST*, 10:P08010, 2015.
- [64] Andrea Lami and Pablo Roig. $H \rightarrow \ell\ell'$ in the simplest little Higgs model. *Phys. Rev.*, D94(5):056001, 2016.
- [65] Mike Lamont. Longer term lhc schedule, 2016.

- [66] LHCb Collaboration. Measurement of the $b_s^0 \rightarrow \mu^+ \mu^-$ branching fraction and search for $b^0 \rightarrow \mu^+ \mu^-$ decays at the lhcb experiment. Submitted to *Phys. Rev. Lett.*, 2013.
- [67] David London and Jonathan L. Rosner. Extra gauge bosons in e6. *Phys. Rev. D*, 34:1530–1546, Sep 1986.
- [68] G. Miu et al. W production in an improved parton-shower approach. [hep-ph/9812455](#), 1998.
- [69] L. Moneta, K. Belasco, K. S. Cranmer, A. Lazzaro, D. Piparo, G. Schott, W. Verkerke, and M. Wolf. The RooStats Project. In 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT2010). SISSA, 2010. PoS(ACAT2010)057.
- [70] P. Nadolsky et al. Implications of cteq global analysis for collider observables. *Phys. Rev. D* 78 (2008) 013004, arXiv:0802.0007.
- [71] G. Nanava et al. How to use sanc to improve the photos monte carlo simulation of bremsstrahlung in leptonic w-boson decays. [hep-ph/0303260](#), 2003.
- [72] K. A. Olive et al. Review of Particle Physics. *Chin. Phys.*, C38:090001, 2014.
- [73] M. Papucci et al. Natural susy endures. *J. High Energy Phys.*, 09:035, 2012.
- [74] M. Pioppi et al. Tau reconstruction and identification with particle-flow techniques using the cms detector at lhc. *Nucl. Phys. B, Proc. Suppl.* 189 (2009) 311 - 316.
- [75] A. L. Read. Presentation of search results: the cls technique. *J. Phys. G*, 28:2693, 2002.
- [76] Jakob Salfeld. Public cms luminosity information, 2016.
- [77] Rostyslav Shevchenko. Documentation of the roostats-based statistics tools for higgs pag, 2016.
- [78] Mark Srednicki. Quantum Field Theory. Cambridge University Press, 2007.
- [79] Viktor Veszpremi. Cms pixel detector 2015 offline performance results - for lhc committee meeting, 2015.
- [80] Z. Was et al. Tauola the library for tau lepton decay. [arXiv:hep-ph/0011305v1](#), 2000.
- [81] Wolfram Zeuner. Ls2 and ls3: The largest challenges as known today, 2014.
- [82] Universitat Zurich. Standard model, 2015.