

Development of Large-Area GEM Detectors for the Forward Muon
Endcap Upgrade of the CMS Experiment
and
Search for SM Higgs Boson Decay in the
 $H \rightarrow \tau^+ \tau^- \rightarrow \mu^+ \mu^- \bar{\nu}_\mu \nu_\mu \bar{\nu}_\tau \nu_\tau$ Channel at $\sqrt{s} = 13$ TeV

by

Vallary Shashikant Bhopatkar

Bachelor of Science
Physics
University of Mumbai
2006

Master of Arts
in Physics
University of Mumbai
2008

A dissertation
submitted to the College of Science at
Florida Institute of Technology
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in
Physics

Melbourne, Florida
December 2017

© Copyright 2017 Vallary Shashikant Bhopatkar
All Rights Reserved

The author grants permission to make single copies _____

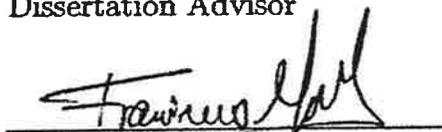
We the undersigned committee hereby recommend
that the attached document be accepted, as fulfilling in
part the requirements for the degree of
Doctor of Philosophy in Physics.

“Development of Large-Area GEM Detectors for the Forward Muon Endcap
Upgrade of the CMS Experiment and Search for SM Higgs Boson Decay in the
 $H \rightarrow \tau^+ \tau^- \rightarrow \mu^+ \mu^- \bar{\nu}_\mu \nu_\mu \bar{\nu}_\tau \nu_\tau$ Channel at $\sqrt{s} = 13$ TeV”,

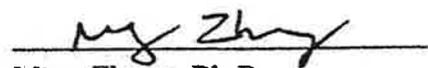
a dissertation by Vallary Shashikant Bhopatkar



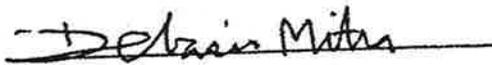
Marcus Hohlmann, Ph.D.
Professor, Physics and Space Sciences
Dissertation Advisor



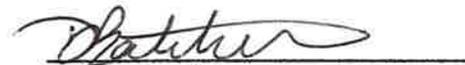
Francisco Yumiceva, Ph.D.
Associate Professor
Physics and Space Sciences



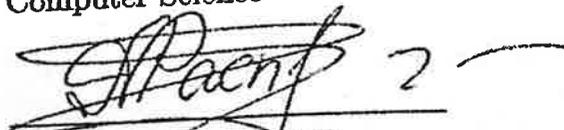
Ming Zhang, Ph.D.
Professor
Physics and Space Sciences



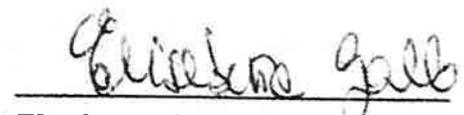
Debasis Mitra, Ph.D.
Professor
Computer Science



Daniel Batcheldor, Ph.D.
Department Head and Professor
Physics and Space Sciences



Alexei Raspereza, Ph.D.
Research Scientist at DESY
Dissertation Co-Advisor



Elisabetta Gallo, Ph.D.
Professor at UH and DESY
Dissertation Co-Advisor

Abstract

TITLE: Development of Large-Area GEM Detectors for the Forward Muon Endcap Upgrade of the CMS Experiment and Search for SM Higgs Boson Decay in the $H \rightarrow \tau^+ \tau^- \rightarrow \mu^+ \mu^- \bar{\nu}_\mu \nu_\mu \bar{\nu}_\tau \nu_\tau$ Channel at $\sqrt{s} = 13$ TeV

AUTHOR: Vallary Shashikant Bhopatkar

MAJOR ADVISOR: Marcus Hohlmann, Ph.D.

Gas Electron Multiplier (GEM) technology is being considered for the forward muon upgrade of the Compact Muon Solenoid (CMS) experiment in Phase II of the CERN LHC. The first GEM Endcap (GE1/1) is going to be installed in the $1.5 < |\eta| < 2.2$ region of the muon endcap mainly to control muon level-1 trigger rates after the second long LHC shutdown. A GE1/1 triple-GEM detector is read out by 3,072 radial strips with $453 \mu\text{rad}$ pitch arranged in eight η -sectors. A meter-long GE1/1 prototype-III was assembled at Florida Tech and tested in 20-120 GeV hadron beams at Fermilab using Ar/CO₂ 70:30 and the RD51 Scalable Readout System (SRS). Four GEM detectors with 2-D readout and an average measured azimuthal resolution of $36 \mu\text{rad}$ provided precise reference tracks. Construction of this GE1/1 prototype-III detector and its performance in the test beam are described. Strip cluster parameters, detection efficiency, and spatial resolution are studied with position and high voltage scans. The plateau detection efficiency is $[97.80 \pm 0.2 \text{ (stat)}]\%$. The azimuthal resolution is found to be $[123.5 \pm 1.6 \text{ (stat)}] \mu\text{rad}$ when operating in the center

of the efficiency plateau and using full pulse height information. The CMS upgrade design calls for readout electronics with binary hit output. When strip clusters are formed correspondingly without charge-weighting and with fixed hit thresholds, a position resolution of $[136.8 \pm 2.5 \text{ stat}] \mu\text{rad}$ is measured, consistent with the expected resolution from strip-pitch/ $\sqrt{12} = 131.3 \mu\text{rad}$. The eight η -sectors of the detector show a similar response and performance. VFAT3 electronics are being considered for the readout system of GE1/1 detectors. The charge that is induced on the GE1/1 readout strips by minimum-ionizing particles is an important parameter that informs the design of the amplifier-shaper input stage of the VFAT3 chip. To estimate the input charge range for these electronics, the most probable value, mean value, and 99th percentile value of the Landau distribution of the charge induced on a single strip are measured and found to be 4 fC, 11 fC, and 115 fC, respectively.

The $Z/\gamma^* \rightarrow \tau\tau$ cross section in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ is measured, using data recorded by the CMS experiment at the LHC during 2015 and corresponding to an integrated luminosity of 2.3 fb^{-1} . The product of the cross section and branching fraction is measured in the dimuon final state to be $1967 \pm 121 \text{ (stat.)} \pm 92 \text{ (syst.)} \pm 37 \text{ (lumi.) pb}$, in agreement with the standard model expectation, computed at next-to-next-to-leading order accuracy in perturbative quantum chromodynamics.

A search for Standard Model (SM) Higgs bosons decaying into pairs of tau leptons and then to two muons plus (anti)-neutrinos is performed using the data collected by the CMS detector in 2016 with 35.9 fb^{-1} of integrated luminosity. The upper limits on the cross section relative to SM prediction are calculated in three event categories with different jet multiplicities focusing on Higgs boson signal events produced via gluon-gluon fusion and vector boson fusion. A multivariate analysis with Boosted Decision Trees (BDT) is used to suppress the large Drell-Yan background. The di-tau mass is reconstructed using a Secondary-Vertex fit (SVFit) algorithm using a maximum likelihood approach. Experimental limits are presented in all three categories extracted from the maximum likelihood fit of reconstructed di-tau mass and the visible mass of the dimuon system. The signal strength for the combination of all three categories is estimated as -1.0 ± 1.7 . The expected and observed upper limits

with 95% CL is at 3.2 and 2.7, respectively, with respect to the SM cross section times branching fraction.

Contents

Abstract	iii
List of Figures	xvi
List of Tables	xix
Acknowledgments	xx
1 Introduction	1
2 The CMS Experiment at the LHC	4
2.1 The Large Hadron Collider	4
2.2 The CMS Experiment	8
3 Development of GEMs for Phase II Upgrade of Muon Endcap	15
3.1 Gaseous Electron Multiplier Detectors	15
3.1.1 Gain Measurements With Different GEM Readouts . . .	18
3.1.2 Experimental Setup	20
3.2 Overview of the GEM Endcap	29
3.3 Fermilab Test Beam Analysis	33
3.3.1 Construction of a GE1/1 GEM Detector Prototype . . .	33
3.3.2 Test Beam Setup	36
3.3.3 Beam Test Results	36
3.3.4 Correction of Non-linear Strip Response	73
3.3.5 Summary and Conclusion	78
4 The Standard Model and Higgs Boson	80
4.1 The Standard Model of Particles	80

4.2	The SM Scalar Sector	82
4.2.1	Higgs Boson Production Mode at the LHC	83
4.3	Drell-Yan Process	89
5	Physics Objects and Analysis Tools	91
5.1	Particle Objects and Reconstruction	91
5.1.1	Muons	92
5.1.2	Taus	94
5.1.3	Jets	95
5.1.4	Missing Transverse Energy	97
5.1.5	Di- τ Mass Reconstructions	98
5.2	Boosted Decision Trees	100
5.2.1	Gradient Boost	100
5.2.2	TMVA and Boosted Decision Trees	102
5.3	Statistical Analysis Using Combine Tools	103
5.3.1	Likelihood	104
5.3.2	Limits	107
5.3.3	Goodness of Fit	109
6	Cross Section Measurements of the Z Boson in the	
	$Z/\gamma^* \rightarrow \tau\tau$ Decay Channel	110
6.1	Introduction	110
6.2	CMS Data Samples and Monte Carlo Simulations	111
6.3	BDT Analysis	113
6.4	Event Selection	119
6.5	Corrections applied to MC	121
6.5.1	Muon identification and isolation efficiency	121
6.5.2	Muon Trigger Efficiency	123
6.5.3	E_T^{miss} resolution and response	123
6.5.4	b-tag efficiency and mistag rate	123
6.5.5	Rochester correction	125
6.6	Background Estimation	125
6.6.1	DY Background	126
6.6.2	Multijet Background	126

6.6.3	$t\bar{t}$ Background	128
6.7	Systematic Uncertainties	128
6.7.1	Event categorization	132
6.8	Cross Section Measurements	133
6.9	Results and Summary	138
7	Search for Neutral Higgs Boson Decays in the	
	$H \rightarrow \tau\tau \rightarrow \mu\mu$ Channel	142
7.1	Introduction	142
7.2	Simulated Samples and Run II Datasets	144
7.3	Event Weights and Data/MC Corrections	146
7.4	Event Categorization	153
7.5	Event Selections	154
7.6	Multivariate Techniques to Reduce the Drell-Yan Background	162
7.7	Signal Extraction	175
7.8	Systematic Uncertainties	183
7.9	Final ML Fit and Results	184
7.10	Future Work	193
8	Summary	197

List of Figures

2.1	The LHC accelerator complex.	7
2.2	The 3-D cross-sectional view of the CMS detector indicating the different sub-detector systems.	9
2.3	Sectional view of the CMS tracker detector.	12
2.4	A quadrant of the R - z plane of the CMS sub-detector systems, as it will appear after the second long LHC shutdown in 2019.	13
3.1	Electron microscope view of a GEM foil. The hole diameter and pitch are $70\ \mu\text{m}$ and $140\ \mu\text{m}$, respectively.	17
3.2	Electric field lines in a GEM foil. The dotted ring indicates the region of highest charge density.	18
3.3	Optical microscope view of the different GEM readout planes: (a) 1-D straight strips, (b) 1-D zigzag strips, (c) 2-D straight strip.	19
3.4	Schematics of a Triple-GEM detector.	19
3.5	Zigzag readout board with 84 strips divided into two sections, left with fine strips and right with coarser strips.	21
3.6	Experimental setup in the lead box for gain measurements.	21
3.7	(a) ^{55}Fe energy spectrum with photo peak ($5.9\ \text{keV}$) and escape peak ($3.1\ \text{keV}$), (b) x-ray energy spectrum with energy peak at $6.155\ \text{keV}$	22
3.8	The rate plateau for the four reference points starts at $4100\ \text{V}$	24
3.9	Gain of the $10\ \text{cm} \times 10\ \text{cm}$ GEM detector ion the order of $10^3 - 10^4$	24
3.10	Experiment setup for gain measurement of a $30\ \text{cm} \times 30\ \text{cm}$ triple-GEM detector with a 2D straight strip readout board.	25
3.11	Beam profile of the x-ray source at position (X2,Y2).	26
3.12	Rate plateau and gain of the $30\ \text{cm} \times 30\ \text{cm}$ GEM detector on the order of $10^3 - 10^4$	26

3.13	Rate and gain vs. V_{drift} measured at the center of the zigzag GEM detector with Ar/CO ₂ 70:30.	28
3.14	A quadrant of the CMS muon system with proposed upgrade (dashed box), showing different subsystems. GE1/1 is highlighted in red dashed box.	29
3.15	GEM and CSC systems enlarge the lever arm for the bending angle measurements.	30
3.16	(a) Simulation of the bending angle measurement in the first endcap station (GE1/1-ME1/1) for soft (~ 5 GeV) and hard (~ 20 GeV) muons. (b) Simulation of the inclusive muon trigger rate expected for the LHC Phase II as a function of the Level-1 p_T trigger threshold for $1.6 < \eta < 2.1$	31
3.17	<i>Left</i> : Super-chamber. <i>Right</i> : Long and short GEM chambers in one endcap.	32
3.18	Cross section through inner and outer chamber frames and GEM foils that shows how the GEM foils are mounted within the GE1/1 chamber so that they can be mechanically tensioned against the brass pull-out posts without deforming the drift or readout boards. The materials of all chamber components are specified.	34
3.19	Construction of the 1m-long large-area GE1/1-III prototype in three steps: (a) Step I: GEM foil assembly with inner frames. (b) Step II: GEM foils stretching. (c) Step III: Closed GEM detector with readout.	35
3.20	Experimental setup in the beam line at Fermilab. The zoomed picture show the GE1/1 detector placed on the moveable table between the four tracker detectors.	37
3.21	A GE1/1-III triple-GEM prototype detector with 24 APV25 hybrids connected to strips via 24 Panasonic connectors.	38
3.22	Cluster charge distribution at 3250 V fitted with Landau function.	38
3.23	Strip multiplicity in strip clusters size at 3250 V in the η -sector 5.	39
3.24	Strip cluster size increases with high voltage.	40
3.25	Charge-to-ADC count calibration for APV25 chip.	41

3.26 (a) ADC-to-fC conversion for lower ADC counts with linear fit. (b) ADC to fC conversion for higher ADC counts with exponential fit.	42
3.27 A large-area GEM detector with radial zigzag strips read out with eight APV hybrids connected to eight η -sectors through Panasonic connectors.	44
3.28 Charge distribution measured with GE1/1-III detector and fitted with a Landau function: (a) Total cluster charge distribution. (b) Individual strip charge distribution.	45
3.29 Charge distribution of the GEM detector with radial zigzag strip readout fitted with the Landau function: (a) Total cluster charge distribution. (b) Individual strip charge distribution.	47
3.30 GE1/1 Detector: Means of the total cluster charge distribution counts vs. drift voltage (a) in ADC. (b) in the fC unit.	48
3.31 GE1/1 Detector: Means of the individual strip charge distribution counts vs. drift voltage (a) in ADC. (b) in the fC unit.	49
3.32 GEM detector with zigzag readout: Means of the total cluster charge distribution counts vs. drift voltage (a) in ADC. (b) in the fC unit.	50
3.33 GEM detector with zigzag readout: Mean of the individual strip charge distribution counts vs. drift voltage (a) in ADC. (b) in the fC unit.	51
3.34 GE1/1 Detector: MPV of the total cluster charge distribution counts vs. drift voltage (a) in ADC. (b) in the fC unit. MPV of the individual strip charge distribution counts (c) in ADC. (d) in the fC unit from Landau fit.	53
3.35 GE1/1 Detector: Sigma of the total cluster charge distribution counts vs. drift voltage (a) in ADC. (b) in the fC unit. Sigma of the individual strip distribution vs. drift voltage (c) in ADC. (d) in the fC unit from Landau fit.	55

3.36	GEM Detector with the radial zigzag strip readout: MPV of the total cluster charge distribution counts vs. drift voltage (a) in ADC. (b) in the fC unit. MPV of the individual strip charge distribution (c) in ADC (d) in the fC unit from Landau fit. . . .	57
3.37	GEM detector with radial zigzag strip readout: Sigma of the total cluster charge distribution (a) in ADC (b) in the fC unit. Sigma of the individual strip distribution (c) in ADC (d) in the fC unit from Landau fit.	59
3.38	The GE1/1 Detector: 99 th percentile of strip charge distribution (a) in ADC (b) in fC unit with linear extrapolation of data that suffers little APV saturation.	61
3.39	The GEM detector with radial zigzag strip readout: 99 th percentile of strip charge distribution (a) in ADC (b) in the fC unit.	62
3.40	Detection efficiency with different cuts on pedestal width.	64
3.41	Response uniformity in seven η -sectors for three different APV positions, i.e., Upper, Middle, and Lower at 3250 V.	64
3.42	Charge uniformity for different η -sectors at 3250 V using single-strip, single-cluster events.	65
3.43	Beam profiles: (a) 32 GeV hadron beam, (b) 120 GeV proton beam.	66
3.44	Correlation of the GE1/1 detector hits with hits in the first tracker detector.	67
3.45	Residuals for tracker 1 in φ : (a) Inclusive residual with width $\sigma = 21 \mu\text{rad}$ and (b) Exclusive residual with width $\sigma = 75 \mu\text{rad}$.	67
3.46	Angular resolution of the tracking detectors.	68
3.47	Inclusive GE1/1 residual with $\sigma = 110.7 \mu\text{rad}$ using full pulse height information.	69
3.48	Exclusive GE1/1 residual with $\sigma = 137.9 \mu\text{rad}$ using full pulse height information.	70
3.49	Resolutions of the GE1/1 detector at the center of sector 5 for different voltages using full pulse height information.	70
3.50	Detection efficiency using VFAT-like binary hit data with three different thresholds.	71

3.51	(a) Inclusive GE1/1 residual $\sigma = 122.8 \mu\text{rad}$ and (b) Exclusive GE1/1 residual $\sigma = 152.5 \mu\text{rad}$ using binary hit reconstruction.	72
3.52	(a) and (b) Scatter plots of exclusive residual vs. η for 2-strip and 3-strip clusters respectively for combined HV scan data. (c) and (d) Scatter plots of exclusive residual vs. η for all strip cluster multiplicities before and after correction, respectively.	75
3.53	(a) Resolution after correction for 2-strip clusters, (b) Resolution after correction for 3-strip clusters, (c) Resolution after correction for 2-strip, 3-strip, and 4-strip clusters for the high voltage scan.	77
4.1	Feynman diagram for the SM Higgs production mode: (a) Gluon-Gluon Fusion (ggH), (b) Vector Boson Fusion (VBF), (c) Vector Boson Association Production (VH), and (d) in Association with a pair of Top Quarks.	84
4.2	The Higgs production cross section at the LHC.	86
4.3	The branching ratios for the decay of the SM Higgs boson with respect to different masses.	87
4.4	Feynman diagram for Drell-Yan process where quark and anti-quark annihilate by creating either virtual photon (electromagnetic process) or Z (electroweak process) which further decays into a muon and anti-muon pair.	90
5.1	Reconstructed di- τ mass with SVFit algorithm and visible dimuon mass in the $H \rightarrow \tau\tau$ decay using simulated 125 GeV Higgs Decays.	99
5.2	Schematic view of a decision tree. Starting from the root node, a sequence of binary splits using the discriminating variables x_i is applied to the data. Each split uses the variable that at this node gives the best separation between signal and background when being cut on. The same variable may thus be used at several nodes, while others might not be used at all. The leaf nodes of the tree are labeled “Sig” for signal and “Bkg” for background depending on the majority of events that end up in the respective nodes in training sample with known background and signal events.	101

6.1	The P_ζ variable.	115
6.2	Variables used as inputs to the BDT that separate the $Z/\gamma^* \rightarrow \tau\tau$ signal from the $Z/\gamma^* \rightarrow \mu\mu$ background. Simulated $Z/\gamma^* \rightarrow \tau\tau$ signal events are shown in red and simulated $Z/\gamma^* \rightarrow \mu\mu$ background events in blue. All distributions are normalized to unity.	118
6.3	The BDT response plotted using TMVA for the $Z/\gamma^* \rightarrow \tau\tau$ signal and the $Z/\gamma^* \rightarrow \mu\mu$ background.	119
6.4	Distributions of $m_{\tau\tau}$ in different jet categories: (a) 0-jet, (b) 1-jet low, (c) medium, and (d) high Z boson p_T	134
6.5	Distributions of $m_{\tau\tau}$ in different jet categories: (a) 2-jet VBF, (b) 1 b jet, and (c) 2 b jet.	135
6.6	Dependence of the profile likelihood ratio $-2 \ln \lambda(\xi)$ on the cross section ξ for DY production of τ pairs. The PLR is computed for the simultaneous ML fit to the observed $m_{\tau\tau}$ distributions in the $\tau_e\tau_h$, $\tau_\mu\tau_h$, $\tau_h\tau_h$, $\tau_e\tau_\mu$, and $\tau_\mu\tau_\mu$ channels. The three curves correspond to the case that statistical uncertainties, the uncertainty in the integrated luminosity, and other systematic uncertainties are included in the fit and when the nuisance parameters corresponding to the integrated luminosity and to other systematic uncertainties are successively fixed in the ML fit. The horizontal line represents the value of $-2 \ln \lambda(\xi)$ that is used to determine the 68% CI on ξ	138
6.7	Distributions of $m_{\tau\tau}$ in inclusive events selected in the $\tau_\mu\tau_\mu$ channel. Signal and background contributions are shown for the values of nuisance parameters obtained from the ML fit to the data.	139
6.8	Cross section $\sigma(\text{pp} \rightarrow Z/\gamma^* + X) \times \mathcal{B}(Z/\gamma^* \rightarrow \tau\tau)$ measured in individual channels and in the combination of all final states, compared to the theoretical prediction and to the cross section measured by CMS for DY production of electron and muon pairs.	141
7.1	Illustration of hadronic recoil.	149
7.2	Mean value of U_1 in the data and simulation as a function of Z p_T and jet multiplicity.	150

7.3	Resolution of the U_1 projection in the data and simulation as a function of $Z p_T$ and jet multiplicity.	151
7.4	Resolution of U_2 in the data and simulation as a function of $Z p_T$ and jet multiplicity.	151
7.5	Effect of applying recoil corrections to the E_T^{miss} distribution in the $Z \rightarrow \mu\mu$ selection. The recoil corrections improve the agreement between the data and MC samples.	152
7.6	Kinematic plots for 0-jet category	156
7.7	Kinematic plots for boosted category	157
7.8	Kinematic plots for VBF category	158
7.9	impact parameter in 0-jet category	159
7.10	impact parameter in boosted category	160
7.11	impact parameter in VBF category	161
7.12	BDT discriminant inputs for 0-jet category	164
7.13	BDT discriminant inputs for boosted category	166
7.14	BDT discriminant inputs for VBF category	167
7.15	BDT response by TMVA package for 0-jet category. Blue represents $H \rightarrow \tau\tau$ and red represents the DY background.	168
7.16	BDT response by TMVA package for boosted category. Blue represents $H \rightarrow \tau\tau$ and red represents the DY background.	168
7.17	BDT response by TMVA package for VBF category. Blue represents $H \rightarrow \tau\tau$ and red represents the DY background.	169
7.18	BDT response in data and simulated events for 0-jet category.	170
7.19	BDT response in data and simulated events for boosted category.	171
7.20	BDT response in data and simulated events for VBF category.	172
7.21	ROC curve for 0-jet category.	173
7.22	ROC curve for boosted category.	174
7.23	ROC curve for VBF category.	174
7.24	Invariant dimuon mass ($m_{\mu\mu}$) for three event jet categories.	176
7.25	Invariant di-tau mass ($m_{\tau\tau}$) for three event jet categories.	177
7.26	2D mass distribution for simulated signal and background events in the 0-jet category.	178

7.27	2D mass distribution for simulated signal and background events in the boosted category.	179
7.28	2D mass distribution for simulated signal and background events in the VBF category.	180
7.29	The prefit 2D mass distribution for the 0-jet category.	181
7.30	The prefit 2D mass distribution for the boosted category.	181
7.31	The prefit 2D mass distribution for the VBF category.	182
7.32	The post-fit 2D mass distribution for the 0-jet category.	185
7.33	The post-fit 2D mass distribution for the boosted category.	186
7.34	The post-fit 2D mass distribution for the VBF category.	187
7.35	The expected and the observed limits on the signal strength for different Higgs masses	188
7.36	GOF using the saturated model for the VBF category.	189
7.37	Pulls and impacts of first 30 nuisance parameters for combined categories.	190
7.38	Pulls and impacts of second 30 nuisance parameters for combined categories.	191
7.39	Pulls and impacts of nuisance parameters 61-89 for combined categories.	192
7.40	DCA significance of the two muon tracks in the transverse plane in the data and simulated MC events for the 0-jet category.	194
7.41	DCA significance of the two muon tracks in the longitudinal plane in the data and simulated MC events for the 0-jet category.	195

List of Tables

4.1	Leptons classified in three generations.	81
4.2	Three generations of quarks with their corresponding charge and mass.	81
4.3	Fundamental forces with interacting mediators.	81
4.4	The branching fraction of the $\tau\tau$ decay.	89
5.1	Muon ID and Isolation Cut used in the $Z \rightarrow \tau\tau$ analysis.	93
5.2	Muon ID and Isolation Cut used in the $H \rightarrow \tau\tau$ analysis.	94
5.3	Tau decay modes with respective branching fractions \mathcal{B} and the intermediate resonances. Here h stand for kaons or pions.	95
5.4	Global tags used to apply jet energy corrections to 2015 and 2016 data and corresponding simulated events.	97
6.1	Datasets used in the $\tau_\mu\tau_\mu$ channel with JSON file used to apply a good-run selection.	112
6.2	List of MC samples used to model the $Z/\gamma^* \rightarrow \tau\tau$ signal and $Z/\gamma^* \rightarrow ee$, $Z/\gamma^* \rightarrow \mu\mu$, W +jets, $t\bar{t}$, single top quark, and diboson background processes. Mass ranges for DY events are given as (a) $10 < m_{\mu\mu} < 50$ GeV and (b) $m_{\mu\mu} > 50$ GeV.	112
6.3	The BDT method-specific ranking of the input variables. The top variable is the best ranked variable	116
6.4	Trigger paths used by the $\tau_\mu\tau_\mu$ channel in data. The path given in the table remained un-prescaled during the entire data taking period.	120
6.5	Trigger paths in simulated events, used by the $\tau_\mu\tau_\mu$ channel.	120

6.6	Measured efficiencies for muons to pass the identification and isolation criteria applied in the $\tau_\mu\tau_\mu$ channel compared to MC predictions.	122
6.7	The efficiencies of the single-muon trigger used in the $\tau_\mu\tau_\mu$ channel and related correction factors applied to simulated events.	124
6.8	Expected number of background events in the $\tau_\mu\tau_\mu$ channels in the data corresponding to an integrated luminosity of 2.3 fb^{-1} . The numbers are rounded to a precision of two significant digits on the uncertainty.	126
6.9	Effect of experimental and theoretical uncertainties in the measurement of the $Z/\gamma^* \rightarrow \tau\tau$ cross section.	131
6.10	Event categories used to study the modeling of backgrounds. Similar categories have been used in previous $H \rightarrow \tau\tau$ analyses at the LHC.	132
6.11	Branching fraction \mathcal{B}_τ , signal acceptance \mathcal{A} , selection efficiency ε , and mass window correction factor f_{out} for $\tau_\mu\tau_\mu$ final state.	136
6.12	Yields in $Z/\gamma^* \rightarrow \tau\tau$ signal events and backgrounds in the $\tau_\mu\tau_\mu$ channel, obtained from the ML fit. The analyzed data corresponds to an integrated luminosity of 2.3 fb^{-1}	140
7.1	List of datasets included in the analysis.	144
7.2	List of Monte Carlo signal samples included in the analysis.	144
7.3	MC background samples generated for pp collisions at a center-of-mass energy of 13 TeV. Samples used in this analysis are reconstructed and stored in miniAOD format. A k-factor of 1.16 is considered for the Z+jets samples and 1.21 for the W+jets samples. All of these MC samples belong to the central CMS Summer16 production, with Moriond Premix conditions. When available, all sample extensions are used.	145
7.4	Medium BDT cuts used in each event category for SVFit di- τ mass calculation.	163
7.5	Final optimized BDT cuts used in each event category for signal extraction.	173

7.6	Expected pre-fit numbers of background events in the $\tau_\mu\tau_\mu$ channel in the data corresponding to an integrated luminosity of 35.9 fb^{-1} for the three jet categories. The numbers are rounded to a precision of two significant digits on the uncertainty.	175
7.7	Post-fit yields in $H \rightarrow \tau\tau$ signal events and respective background events in the $\tau_\mu\tau_\mu$ channel with data corresponding to an integrated luminosity of 35.9 fb^{-1} . The numbers are rounded to a precision of two significant digits on the uncertainty.	185
7.8	Upper limits at 95% CL on the signal strength relative to standard model prediction in the $\tau_\mu\tau_\mu$ channel.	186
7.9	Measurement of $H \rightarrow \tau\tau \rightarrow \mu\mu$ signal strength relative to the expected SM cross section times branching fraction.	187

Acknowledgements

I would like to express my sincere gratitude to my research and dissertation advisor, Dr. Marcus Hohlmann, for his guidance, patience, support, and encouragement. Thank you again for getting the best out of me and making me realize my own capabilities. I would like to thank Dr. Francisco Yumiseva, Dr. Ming Zhang, and Dr. Debasis Mitra for agreeing to be on my dissertation committee.

During the first few years, I worked on GEM development and learned about the detector. I would like to thank Mike Staib, Mike Phipps, and Jessie Twigger, former students of Dr. Hohlmann who helped and guided me in my learning phase about GEMs. I would like to thank Dr. Aiwu Zhang, former post-doc at Florida Tech, for guiding and supervising me in GEM tracking analysis. We had a successful beam test experience at Fermilab. The beam test preparation and actual data taking wouldn't have been possible without his guidance. I am also very grateful to the current GEM team at the Florida Tech, especially our postdoc Dr. Stefano Colafrancesci and current PhD student Mehedi Ramani for valuable discussions about the GEM upgrade.

I would like to thank Ankit Mohapatra for his guidance during my transition between the hardware studies and the physics analysis and for all the useful discussions about the physics and multivariate analysis. I would like to express my deepest gratitude to Dr. Alexei Raspereza and Dr. Elisabetta Gallo, my co-research advisors from DESY, Hamburg. I successfully finished the SM Higgs analysis in the tau decay mode with two muons in final states for this dissertation using CMS Run II 2016 data. This wouldn't have been possible without their constant guidance and encouragement. I would like to thank again all my advisors for this DESY-Florida Tech collaboration. Due to this collaboration, I

had an opportunity to visit and work at DESY as well as to meet and learn from the experts in high energy physics. I would like to thank all the DESY Higgs group, especially, Valeria, Yiwen, Francesco for their support and guidance. I would also like to thank the CMS Higgs groups for their valuable guidance and support.

The journey of getting a doctorate degree was not very easy. I am grateful to have such a wonderful family and friends who supported me all the time. I am truly grateful to my parents Megha and Shashikant Bhopatkar, who have always believed in me and encouraged me to follow my dreams. They taught me, if you decide anything in your heart, then nothing is impossible. Following their teaching, I am trying to lead with an example. I am grateful to my sisters, Manashri and Vaidehi, for their love, support, and encouragement. This doctorate wouldn't have been possible without my husband, Milind. I can't thank him enough for his great support and the sacrifices he made in his life to fulfill my dream. Also, I am very thankful to my in-laws, Mamta and Krishna Sali for supporting me in every decision. I would like to thank all my extended family, mostly my aunts, Poulomi, Sucheta, Sunetra, and, my sister-in-law, Rashmi, for their love and encouragement.

I would like to thank my friends, Sayali, Trisha, Cori, Titas, Omkar, Kanchan, and Gayatri for always being there for me when I needed them the most. Finally, I would like to express my gratitude to each and every person who has helped me throughout this journey!!!!

I would like to dedicate this dissertation to my late grandparents, Smt. Vijaya Bhopatkar (Grandmother) and Shri. Raghunath Bhopatkar (Grandfather) for their immense love for education and knowledge.

Chapter 1

Introduction

Currently the Large Hadron Collider (LHC) at CERN, the European Organization for Nuclear Research, is running in the Run II era of its first phase. The Compact Muon Solenoid (CMS) experiment is taking proton-proton (pp) collision data at a center-of-mass energy $\sqrt{s} = 13$ TeV. For Phase II of the LHC, several upgrades are planned in different detector systems. In Phase II, LHC will have increase its instantaneous luminosity and will be referred to as the High Luminosity (HL)-LHC. Several upgrades will take place during the second long LHC shutdown (LS2) beginning in 2019. One of the major CMS upgrade projects is the upgrade of the muon endcap (ME) which introduces the new Gaseous Electron Multiplier (GEM) endcap stations. GEM detectors will be used for this upgrade achieve high gain on the order of 10^4 - 10^5 . Also, they exhibit good timing resolution. Due to their high spatial resolution, these detectors can provide very accurate tracking. This installation will overall improve the performance of muon triggerig and tracking. The GEM detectors will be read out using VFAT3 front-end chips, which provide binary hit information.

Strip charge measurements are carried out to determine the required electronic input charge range for this front-end chip.

Also, during HL-LHC in addition to the GEM station installation, several other muon detector upgrades are planned [1]. These upgrades are necessary to keep the muon trigger and data acquisition as efficient as in Run I and Run II. This GEM station installation will make a positive impact not only on the trigger information but ultimately also on several physics analyses, including Higgs physics. This motivates us to study Higgs decay. In the $H \rightarrow \tau\tau$ analysis, Drell-Yan (DY) events are a very important background and hence, it is important to study and model this background carefully. The background estimation methods are established by measuring the Z cross section using decays into the τ leptons. This dissertation focuses on the $\mu\mu$ final state of the τ decays. Other groups in the CMS collaboration have studied the other hadronic and leptonic tau decays. Due to the similar background model, the background estimation in this channel helps to estimate the background for the $H \rightarrow \tau\tau \rightarrow \mu\mu$ decay. In this physics analysis, a boosted decision tree (BDT) multivariate method is employed to separate the signal from the background. This BDT training focuses on the DY background, as other backgrounds have a very small effect on the signal extraction. A BDT has been trained separately on 2015 data for Z cross section measurements and on 2016 data for the Higgs measurement. In the case of the Higgs analysis, BDTs are trained separately for three event categories based on the jet multiplicity in the final state. Finally, the signal strength in the cross section relative SM to the prediction is estimated as well as the expected and observed upper limit on the cross section times branching ratio for this channel is determined.

This dissertation explains in detail the GEM upgrade work as well as the two physics analyses. The structure of the dissertation is as follows:

1. This first chapter gives the overall idea and structure of the dissertation.
2. The second chapter introduces the LHC and the CMS experiment. The details of the CMS sub-detectors are explained.
3. Chapter three summarizes the Phase II upgrade of the muon endcap and describes measurement of overall performance characteristics of GE1/1 prototype detectors, which include gain, strip charge, detection efficiency, and spatial resolution.
4. In chapter four, the dissertation makes a transition from the hardware project to physics analysis. The basic theory of electroweak interaction is explained along with the Higgs production processes and decay modes.
5. Chapter five describes the basic physics objects needed for the analyses and their reconstruction methods used in CMS.
6. In chapter six, the BDT analysis for the signal extraction and the cross section measurement for Z boson in the $Z \rightarrow \tau\tau \rightarrow \mu\mu$ channel is explained in detail.
7. Chapter seven describes the $H \rightarrow \tau\tau \rightarrow \mu\mu$ analysis including the BDT methods for signal extraction.
8. Lastly, chapter eight summarizes the hardware development and physics analysis results and give an outlook towards potential future work.

Chapter 2

The CMS Experiment at the LHC

2.1 The Large Hadron Collider

The LHC [2] is the world's largest and most powerful accelerator. It is a circular hadron-hadron collider situated about 100 m underground, with an accelerator ring that spreads over 27 km in the circumference. The LHC was built inside the tunnel formally used for the Large Electron-Positron (LEP) collider and its construction was finished in early 2008. One of the main purpose of designing this collider was to look for the scalar boson particle predicted by the standard model (SM) and also to reveal the physics behind it. This collider was initially designed to collide beams at a center-of-mass energy (\sqrt{s}) of 14 TeV. The LHC is currently taking data in the last stage of Phase I with $\sqrt{s} = 13$ TeV with an instantaneous luminosity of about $1.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. The LHC consists of superconducting magnets with many accelerating structures to boost the energy

of the colliding particles. The LHC supports proton-proton (pp) collision along with heavy ion collisions. Details on the LHC machine can be found in [3]. The number of events generated per second in LHC collisions are given as:

$$N_{event} = \mathcal{L} \sigma_{event}, \quad (2.1)$$

where σ_{event} is the cross section of the process and \mathcal{L} is the nominal LHC luminosity. The instantaneous luminosity of the LHC depends on several beam parameters, and their relation is given by [4].

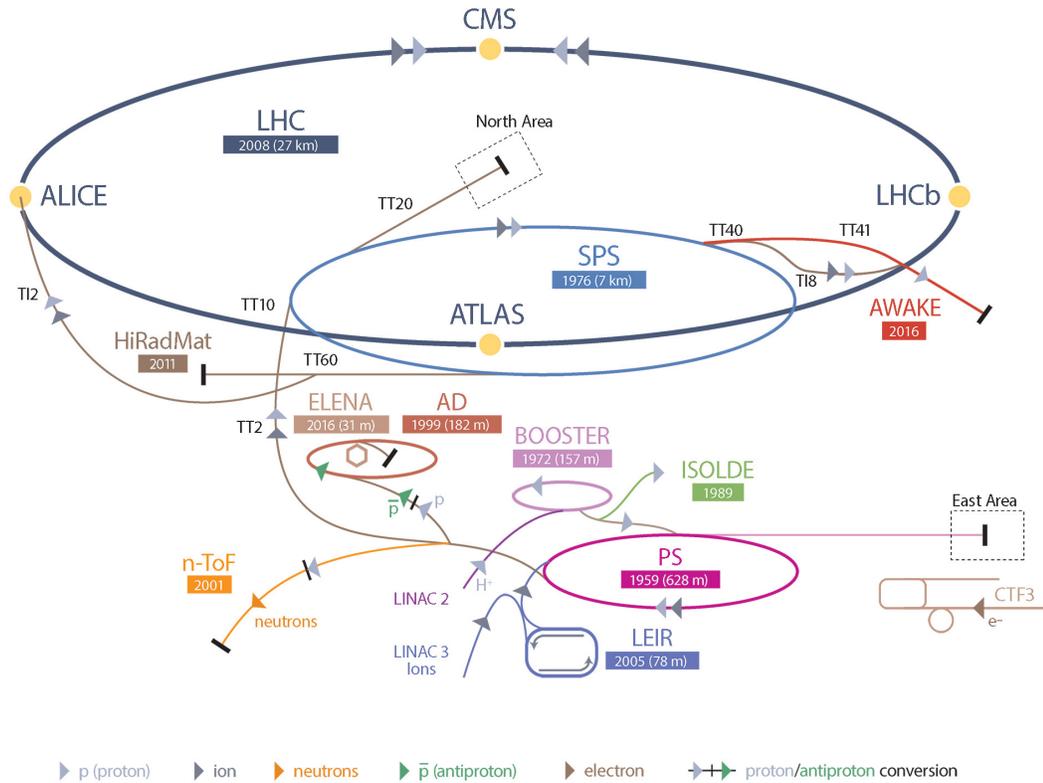
$$\mathcal{L} = \frac{N_1 N_2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} F \quad (2.2)$$

The beam parameters used in the above equation are as follows:

- $N_{1,2}$ is the number of particles per bunch in beams 1 and 2
- n_b is the number of bunches per beam
- f_{rev} is the revolution frequency
- γ_r is the Lorentz factor
- ϵ_n is the normalized transverse beam emittance
- β^* is the beta function at the collision point
- F is the reduction factor due to the crossing angle of the two beams.

A schematic diagram of the LHC accelerator complex and its underground layout is given in Figure 2.1. There are seven experiments at the LHC that use its collision data and they are all distinct and characterized by their detectors.

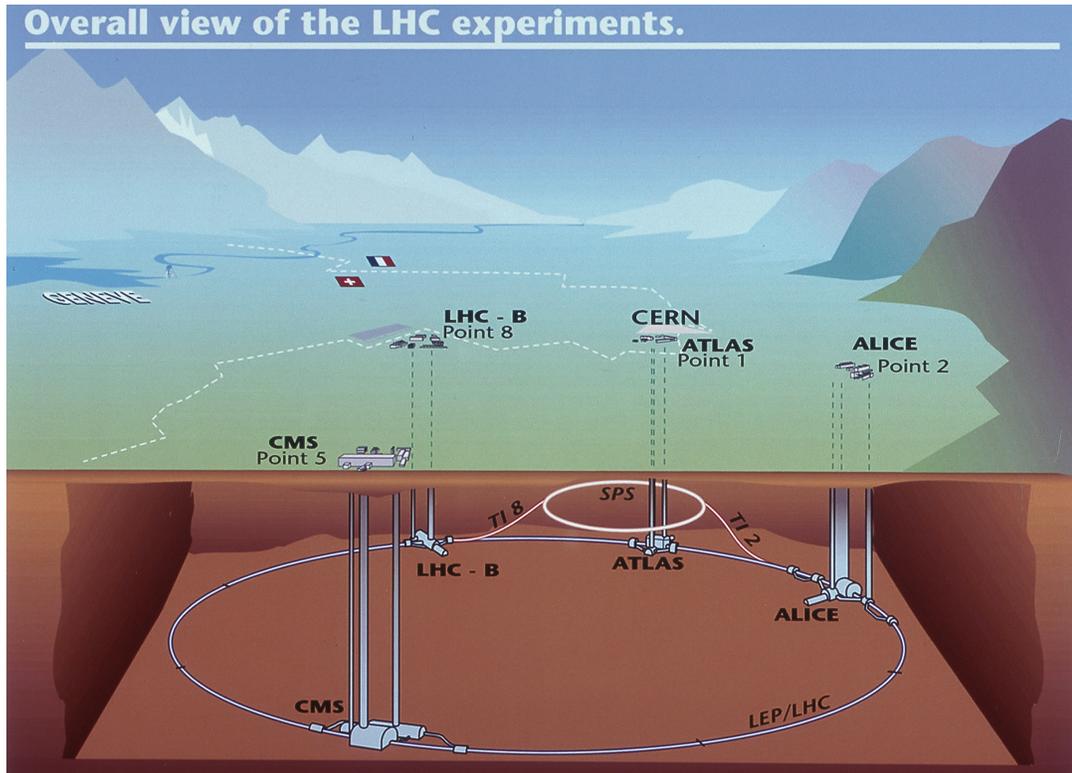
CERN's Accelerator Complex



(a) The schematic diagram of the position of the LHC.

Experiments at the LHC are run by the largest collaboration of scientists from all over the world. ATLAS and CMS are the general purpose experiments and are also the largest among them. Their data are used to study all the different possible physics concepts. The LHCb (Large Hadron Collider beauty) experiment uses the b quark properties to mainly study the matter-antimatter a symmetry, while ALICE focuses on the strong interaction physics in heavy-ion collision. These four detectors are situated at the four different colliding points at the LHC. The three smallest experiments are TOTEM, LHCf, and MoEDAL. Information about these experiments can be found in [5].

Phase I of the LHC will be ending in 2022 and Phase II, referred as the High



(b) Underground layout of the LHC.

Figure 2.1: The LHC accelerator complex.

Luminosity (HL) era, will start in 2023. For Phase II, there are several upgrades proposed in different experiments, which will be taking place during the Second and Third Long Shutdown (LS2, LS3) of the LHC. After the third shutdown, the instantaneous luminosity will exceed the current maximum by ultimately a factor of five. The work presented in this dissertation is conducted within the CMS collaboration. The details of the CMS experiment are described in detail in the following sections.

2.2 The CMS Experiment

The CMS experiment is one of the largest and most general purpose experiments at the LHC. The compact muon solenoid name comes from its relatively compact size, considering the complexity of the experiment as well as from its powerful solenoid coil and its sophisticated muon detection system. A 3-D cross-sectional view of the CMS detector is shown in Figure 2.2. The CMS detector is 28.7 m long and 15 m in height with a weight of more than 14,000 tons. The CMS experiment has one of the strongest large magnets in the world with a field strength of about 3.8 Tesla. The sub-detectors of the CMS are located in the inner cylindrical region, followed by the outer barrel region and sandwiched by the two endcaps. The CMS experiment place the origin of its right-handed Cartesian Coordinate system at the point of nominal collisions. The x-axis points towards the center of the LHC, the y-axis points up in the vertical direction, and finally the z-axis is along the beam direction.

The polar angle θ is measured from the z-axis, while the azimuthal angle ϕ is measured from the x-axis in the x-y plane. The geometry of the CMS experiment is often discussed in terms of the azimuthal angle and pseudo-rapidity η . Since a difference in rapidities of particles remains invariant under a Lorentz boost along the z-direction, using the pseudo-rapidity is preferred over the polar angle. The definition of pseudo-rapidity is given as:

$$\eta = -\ln \tan \left(\frac{\theta}{2} \right), \quad (2.3)$$

where as the rapidity of a particle of energy E and momentum \vec{p} is expressed

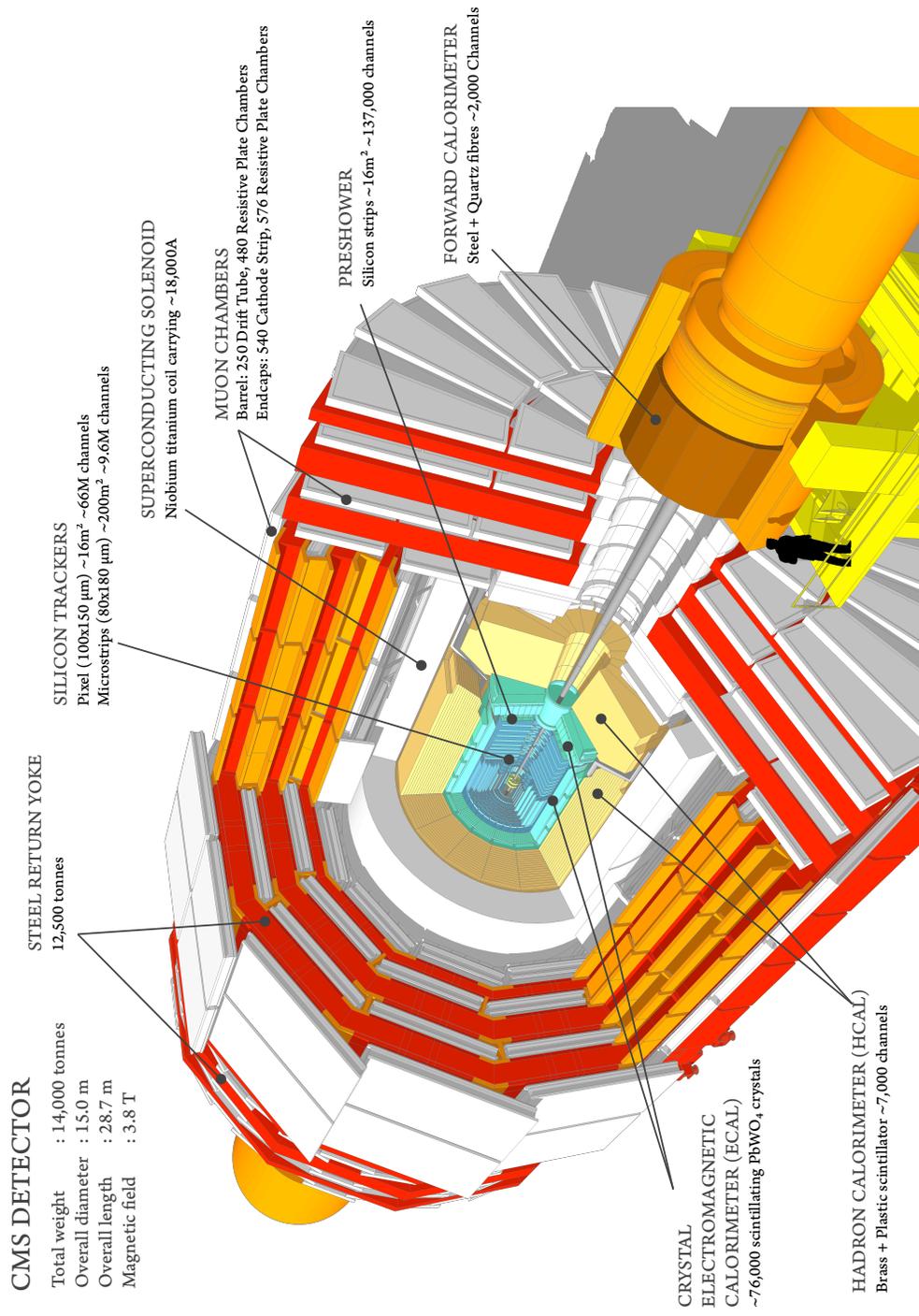


Figure 2.2: The 3-D cross-sectional view of the CMS detector indicating the different sub-detector systems.

as:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right). \quad (2.4)$$

In the η - ϕ plane, the angular separation between two particles is given by:

$$\Delta R = \sqrt{(\phi' - \phi)^2 + (\eta' - \eta)^2} \quad (2.5)$$

The sub-detectors of the CMS from the innermost to outermost parts are given as follows:

- The **Inner Tracker System** reconstructs the secondary vertices particle decays and estimates charged particle trajectories.
- The **Electromagnetic Calorimeter (ECAL)** estimates the energy of particles that interact electromagnetically, such as electrons and photons.
- The **Hadronic Calorimeter (HCAL)** absorbs hadrons and measures their energy.
- The **Solenoid** is a magnet with a field strength of about 3.8 T that is used to bend the tracks of charged particles for momentum measurements.
- The **Muon System** tracks and measures the momentum of muons.

In addition, the trigger system is developed to select the data relevant to interesting physics scenarios. Details on the CMS sub-detector systems can be found in [6]. This dissertation focuses on the upgrade of the muon endcap, and on H and Z decays in the τ channel, specifically those that further decay into muons. For these physics analyses, the important measurements come from the

tracker and muon system; hence these subsystems are detailed in the following subsections.

Inner Tracker Detectors

The innermost subsystem of the CMS detector is composed of the inner tracker detectors. These silicon trackers are cylindrical in shape with a 2.5 m diameter and a length of 5.8 m. Since these detectors are very close to the collision point, these systems are bombarded by a large number of particles. The primary purpose of these detectors is to reconstruct the charged particles from their track measurements and to estimate the momentum associated with them. Due to the very high particle density environment, it is essential to have high granularity. This will help to separate the hard collision events from the pileup events. For every bunch crossing there are more than one collision occur that know as the pileup. For better reconstruction performance, one should limit phenomena such as multiple scattering, bremsstrahlung, photon conversion, and nuclear interactions. The only way it is possible to limit these effects is by reducing the material interaction. Due to these requirements, the best suited technology to use for this system is silicon detector technology.

The sectional view of the tracker is shown in Figure 2.3 [7]. The core of the tracker detector consists of the silicon pixel detectors surrounded by silicon microstrip detectors. The size of the silicon pixels is $100 \times 150 \mu\text{m}^2$; they are arranged in three cylindrical barrels and two endcap disk layers. The silicon microstrip detectors are arranged in ten (4 + 6) barrel layers parallel to the beam axis and radially in 3 + 9 layers in the endcap. The total tracker systems can reconstruct the 3-dimensional vertex with a spatial resolution of about 15-

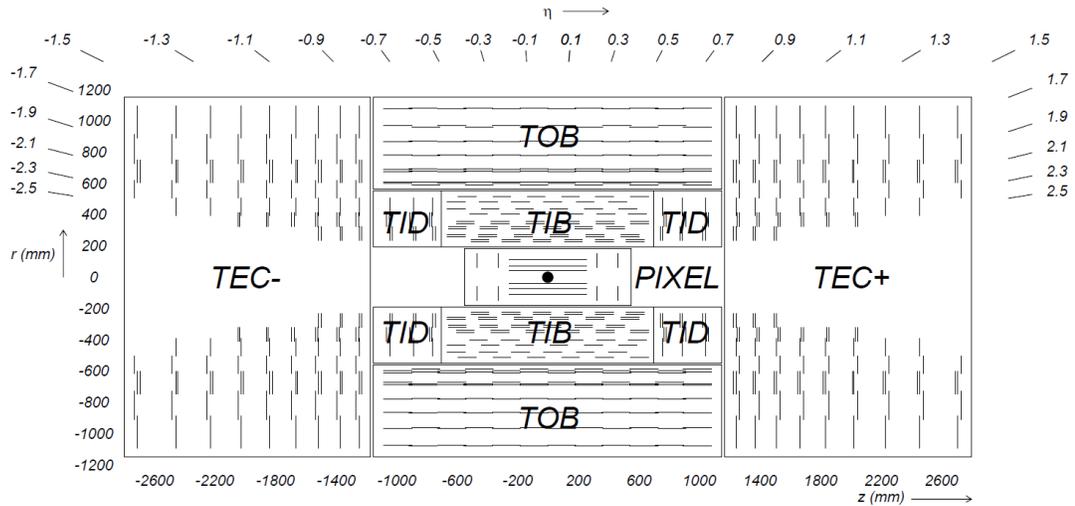


Figure 2.3: Sectional view of the CMS tracker detector.

20 μm and transverse momentum with an accuracy of about 2%. In total, the tracker detector represents a 200 m^2 active silicon area with 75 million readout channels.

Muon Detector

CMS has a sophisticated muon system, which provides measurement of the transverse momentum and the charge of muons with great precision as well as triggering. A quadrant of the R - z cross section view of the CMS detector is shown in Figure 2.4. The muon sub-detectors rely on three gaseous detector technologies just outside the magnet solenoid: (1) Drift Tubes (DT) in the barrel region cover the acceptance up to $|\eta| < 1.2$ and (2) Cathode Strip Chambers (CSC) are installed in the endcap covering the region of $1.0 < |\eta| < 2.4$. Additional Resistive Plate Chambers (RPC) are installed in both barrel and endcap region covering an acceptance up to $|\eta| < 1.6$. A detailed explanation of the muon system can be found in [8]. These detectors are operated with various gas

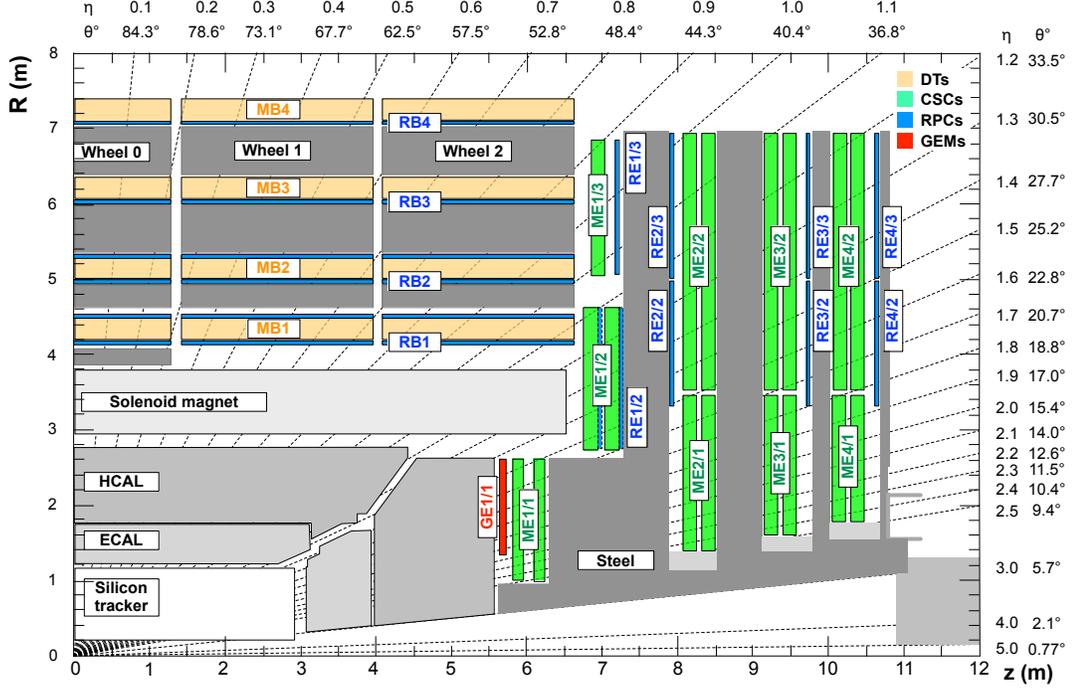


Figure 2.4: A quadrant of the R - z plane of the CMS sub-detector systems, as it will appear after the second long LHC shutdown in 2019.

mixture. Muons entering these detector ionize the gas mixtures. Ions produced in this process pass through the internal electric fields and produce an avalanche of secondary electron-ion pairs.

The DT provides a very high reconstruction efficiency for muon hits. Each chamber has a resolution of about $100 \mu\text{m}$ in the r - ϕ plane. The CSCs are trapezoidal multiwire proportional chambers with six anode wire planes, which provide a fine spatial resolution of $50 \mu\text{m}$ for muon tracks. Unlike the DT, the CSC can operate in a high-rate neutron induced background with a very high and non-uniform magnetic field. The RPCs provide excellent time resolution and are only used at the trigger level. They are double-gap chambers operated in avalanche mode for high-rate operations.

Trigger

At the LHC, bunches cross with 40 Hz frequency. In each bunch crossing ~ 20 pp collisions occur. It is an impossible task to store all the data coming from these events; therefore CMS uses a two-stage trigger process to store data only from events whose kinematics are relevant to specific physics analyses. The first stage trigger is the Level-1 (L1) trigger [9]. This trigger uses fast information from calorimeters and muon systems about local objects such as photons, electrons, jets, and muons. It also uses the global sum of E_T and E_T^{miss} . The maximum design trigger output rate for L1 is about 100 kHz. At this rate, the rejection of unwanted events is on the order of 10^4 . In the second step, this large trigger rate is further reduced to a few hundred Hz by using the High Level Trigger (HLT). The complete information about the HL trigger can be found in [10]. The HLT is based on software techniques that store only the relevant event information; full readout information can also be accessed during this stage. The events that pass the HLT requirements are kept in storage, which bears the information about the reconstruction of objects and events passing the criteria.

Chapter 3

Development of GEMs for Phase II Upgrade of Muon Endcap

3.1 Gaseous Electron Multiplier Detectors

A Gas Electron Multiplier (GEM) [11], a type of Micro-Pattern Gas Detector (MPGD), was introduced by Fabio Sauli in 1997 to find an alternative option for micro-strip gas chambers (MSGC) [12]. The MSGC had limitations on its gas gain due to the restriction on the maximum voltage that could be applied. As the gain of the detector increases with applied voltage, so does the risk of sparking and discharge. GEM technology helps to overcome this issue by providing one or more stages of pre-amplification in the gas. This helps to achieve a high gain by allowing the operation of the micro-strip readout at lower applied voltage. Consequently, it reduces the risk of sparking caused by

the ionizing particles.

A typical GEM detector consists of a GEM foil placed between the drift and the readout structure. The GEM foil is a 50 μm polymer foil such as kapton that is clad on both sides by a layer of copper and chemically perforated by high density holes. Typically, these holes are 70 μm in diameter and 140 μm apart from each other as shown in Figure 3.1 [11]. Double-sided etching gives double-conical shape to the holes. For small GEM foils, the double-mask manufacturing method is used while for large area foils the single-mask technique is used. The complete manufacturing process can be found in [13]. The voltage difference between the two metallic surfaces of the GEM foil produces a high electric field in the holes. As a result, electron multiplication occurs when the electron generated in the primary ionization enters the hole. Figure 3.2 [14] shows electric field lines in the amplification region in a GEM foil.

Amplifications process in the GEM foils along with small strip pitch preserve good spatial resolution. High density holes in the foils help to keep the gas gain unaffected in high radiation flux by maintaining the stable high voltage across the hole. These advantages are very important for tracking particles in high energy physics experiments. Another advantage of this technology is the freedom to choose the pattern of the charge collection and readout plane as required by the experiment. The pattern of the readout plane can be strips, pads, or a combination of the two. Figure 3.3 shows different readout strip patterns. Figure 3.3 (a) shows the 1-dimension (1-D) straight strip readout, (b) shows the 1-D zigzag strips while (c) show the 2-D straight strip readout.

For GEM detectors, larger gain can be achieved by cascading multiple GEM foils. The most popular and reliable configuration is a triple-GEM detector in

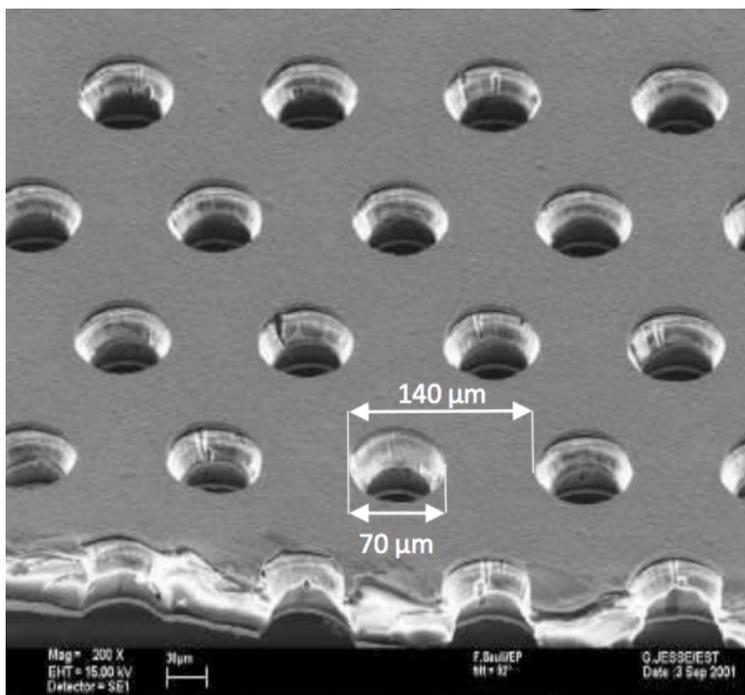


Figure 3.1: Electron microscope view of a GEM foil. The hole diameter and pitch are $70 \mu\text{m}$ and $140 \mu\text{m}$, respectively.

which three GEM foils are placed between the drift and readout plane. The schematic diagram of the triple-GEM is shown in Figure 3.4 [15]. GEM detectors can use various gas mixtures based on the performance requirement. For example, for a Cherenkov ring imaging application, pure carbon tetrafluoride gas is one of the factor to achieve a very high gain. A gas mixture with CF_4 also can reach a time resolution better than 5 ns , which is essential for a high rate collider such as the Large Hadron Collider (LHC) [16]. At Florida Institute of Technology (Florida Tech), a non-flammable Ar/CO_2 70:30 mixture is used typically. With this mixture, GEM detectors can reach high particle detection efficiency close to 100% [17] as well as a gain in the range of 10^4 . In the following sections, various characteristics measurements of detector, such as detector gain, efficiency, cluster size and spatial resolutions are discussed in detail. The

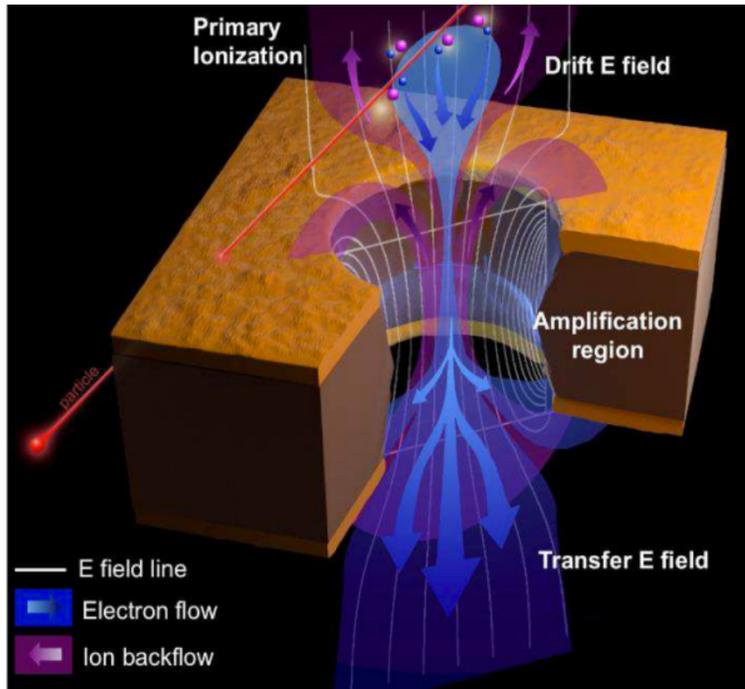


Figure 3.2: Electric field lines in a GEM foil. The dotted ring indicates the region of highest charge density.

muon group in the Compact Muon Solenoid (CMS) experiment is implementing one of the upgrades of the muon endcap in early phase II by introducing the a GEM endcap station officially know as GE1/1 station. All details of the full muon endcap phase II upgrade can be found in [18] and for the GE1/1 upgrade, details can be found in the technical design report [14].

3.1.1 Gain Measurements With Different GEM Read-outs

GEM detectors require higher gain to produce a large enough signal to allow detection of the single electron entering the GEM cascade [19]. In gaseous detectors, GEM detectors can provide higher gain with lower drift voltage. In this

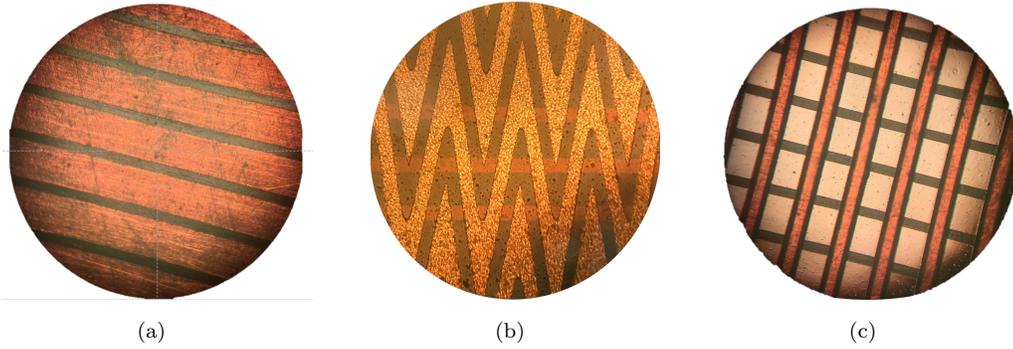


Figure 3.3: Optical microscope view of the different GEM readout planes: (a) 1-D straight strips, (b) 1-D zigzag strips, (c) 2-D straight strip.

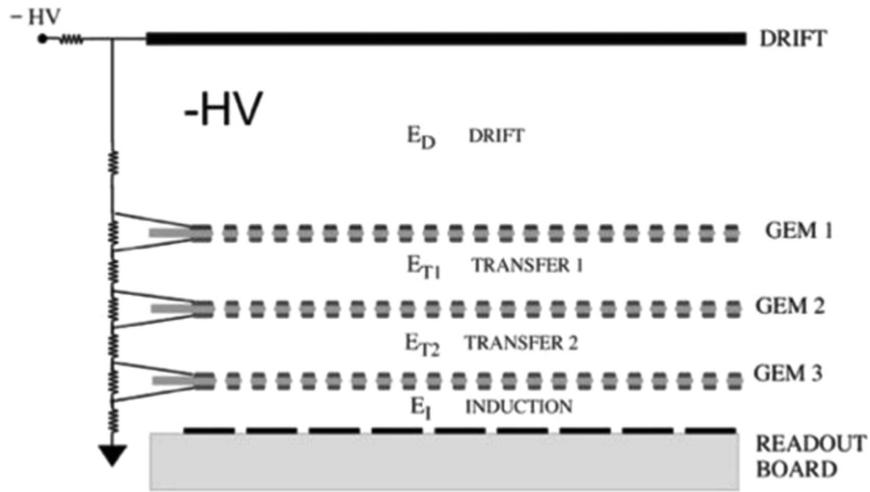


Figure 3.4: Schematics of a Triple-GEM detector.

section, the gain of the triple-GEM detector is measured with different readout sizes, patterns, as well as with two different internal gap structures. The $10\text{ cm} \times 10\text{ cm}$ and the $30\text{ cm} \times 30\text{ cm}$ size triple-GEM detectors are constructed using $3/2/2/2\text{ mm}$ internal gap configuration. The numbers represent the spacing between the drift, three GEM foils, and the readout: E_D , E_{T1} , E_{T2} , E_I as shown in Figure 3.4, respectively. In the 1-m long GE1/1 prototype GEM detector, the internal gap configuration is $3/1/2/1\text{ mm}$. Mini-X x-ray source with gold (Au) anode is used for the measurements. The gain measurement is performed

in three steps. The first step is an energy calibration to obtain the operating x-ray energy. In the second step, the event rate is measured for different voltages and a rate plateau is obtained. In the final step, the gain of the detector is calculated by measuring the anode current directly from the readout.

sub

3.1.2 Experimental Setup

The 10 cm \times 10 cm triple-GEM detector used in this measurement has a readout board with 84 zigzag strips with a 2 mm pitch. The 84 zigzag strips are divided into two sections as shown in Figure 3.5, one with 48 fine strips and the other 48 with coarser strips. An Ar/CO₂ gas mixture is used in 70:30 proportion. The x-ray source is operated at 10 kV voltage with 5 μ A current without any filters. The entire experimental setup, as shown in Figure 3.6, is shielded in a large lead box for radiation safety. A Panasonic connector with 128 channels is used to read the signal from the readout board. As shown in Figure 3.6, four different positions on the detector are considered, namely fine near end, fine far end, coarse near end, and coarse far end. The readout signal is amplified using a charge sensitive pre-amplifier and a linear amplifier. The amplified signal is fed to the a multi-channel analyzer (MCA) and scalar for energy calibration and count rate measurement, respectively.

Energy Calibration

An ⁵⁵Fe source is used for energy calibration. The exact energy value of the x-ray is required for accurate gain measurements. The photo peak energy (5.9 keV) and argon escape peak energy (3.1 keV) of ⁵⁵Fe are used to calibrate the MCA

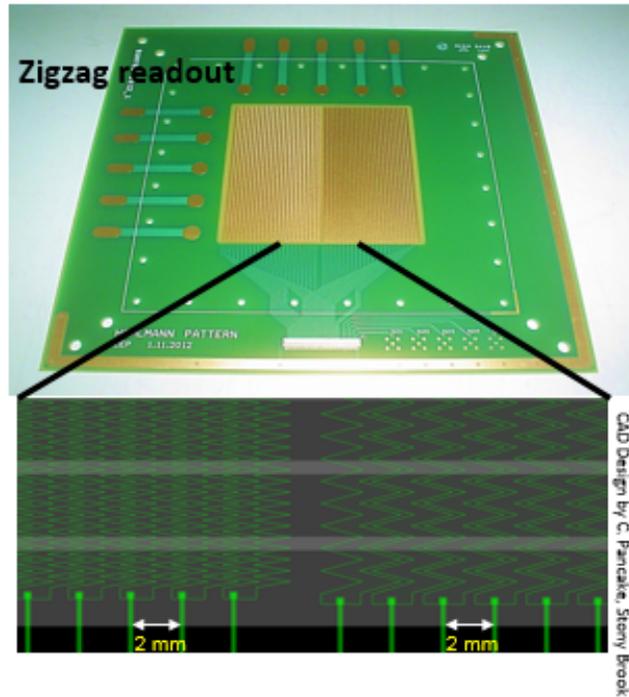


Figure 3.5: Zigzag readout board with 84 strips divided into two sections, left with fine strips and right with coarser strips.

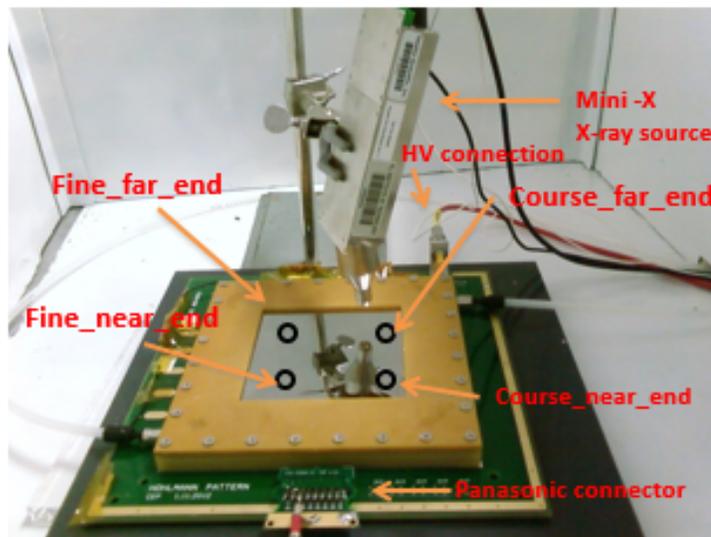
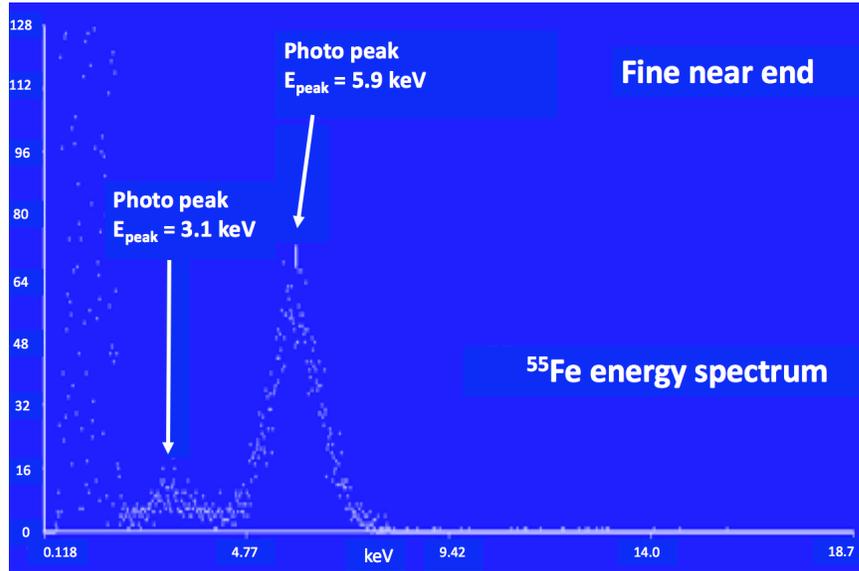
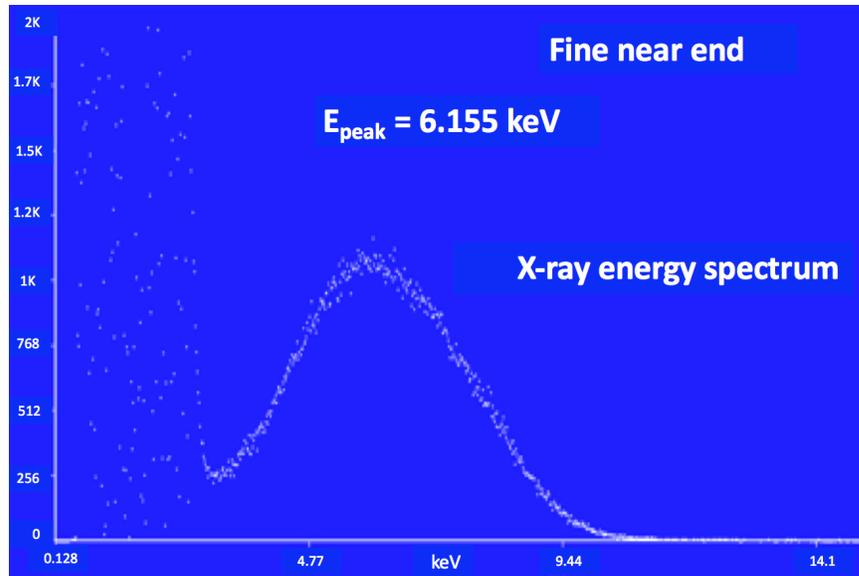


Figure 3.6: Experimental setup in the lead box for gain measurements.

scale. From the calibrated scale, the exact x-ray energy at all four reference points is calculated. Figure 3.7 shows the energy spectrum of the ^{55}Fe source and x-ray source at the fine near end position.



(a)



(b)

Figure 3.7: (a) ^{55}Fe energy spectrum with photo peak (5.9 keV) and escape peak (3.1 keV), (b) x-ray energy spectrum with energy peak at 6.155 keV.

Rate Plateau and Gain Measurements

The gain measurement requires the event rate at the voltage on the rate plateau. The event rate measured for different voltages varies from 3700 V to 4200 V in the interval of 50 V. The rate plateau is achieved above 4100 V for all four reference points as shown in Figure 3.8. Hence, the rate at 4150 V is used for the gain calculation. The gain of the detector is defined as the ratio of charges detected by the readout board to primary charges and mathematically expressed as follows:

$$Gain = \frac{I}{R \times n_{primary} \times e}, \quad (3.1)$$

where R is the particle rate, e is the electron charge and $n_{primary} = \frac{E_{x-ray}}{W_i}$; here, E_{x-ray} is the energy of the x-rays and W_i is the effective average energy required to produce one ion-electron pair in the gas. In a 10 cm \times 10 cm triple-GEM detector, an Ar/CO₂ 70:30 gas mixture is used. Therefore, in this particular case, the number of the primary ion-electron pairs for ionization are calculated by

$$n_{primary} = \frac{E_{x-ray}}{W_i} = E_{x-ray} \left(\frac{0.7}{W_i^{Ar}} + \frac{0.3}{W_i^{CO_2}} \right) \quad (3.2)$$

For the Ar/CO₂ gas mixture $W_i^{Ar} = 25$ eV and $W_i^{CO_2} = 34$ eV.

Using Equation 3.2, the $n_{primary}$ is calculated for all four reference points, and using these values, the gain of the 10 cm \times 10 cm GEM detector is calculated. The calculated gain is on the order of $10^3 - 10^4$, as shown in Figure 3.9.

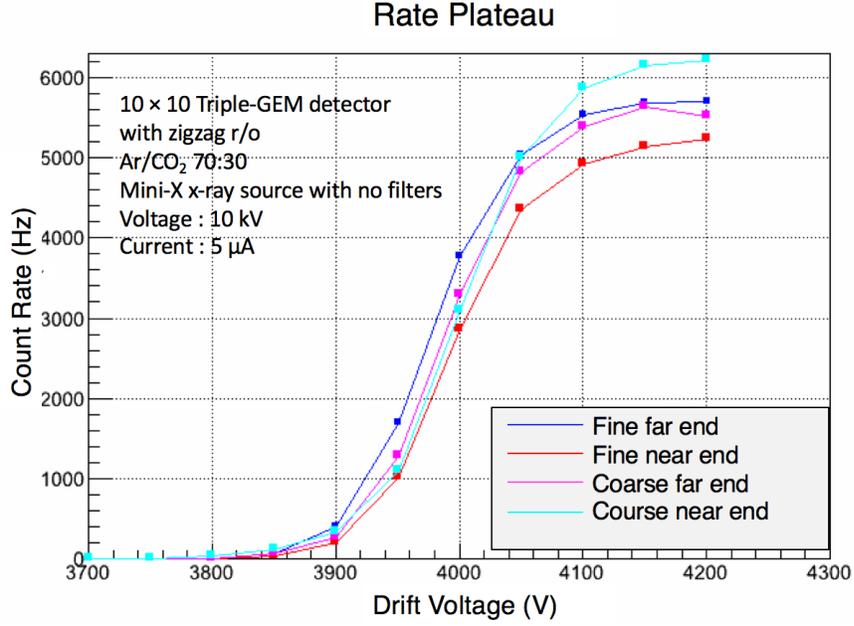


Figure 3.8: The rate plateau for the four reference points starts at 4100 V.

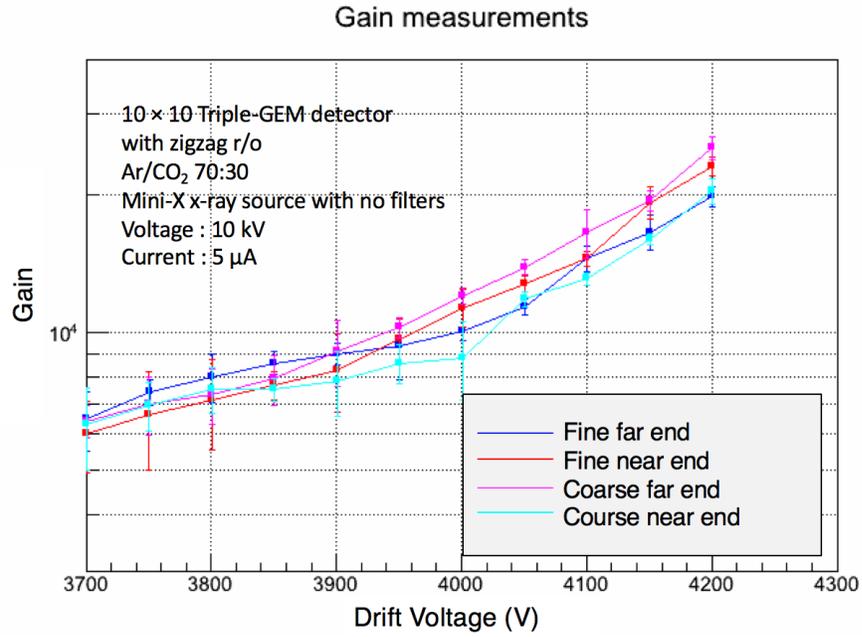


Figure 3.9: Gain of the 10 cm × 10 cm GEM detector ion the order of 10³ - 10⁴.

30 cm × 30 cm Triple-GEM Detector with 2D Straight Strip Readout

A 30 cm × 30 cm triple-GEM detector is assembled using internal mechanical foil stretching. The same technique is used for the CMS GE1/1-III prototype detector. The details are discussed later in Section 3.3.1. Each GEM foil has 10 high voltage sectors and the readout has a 2D structure. The signal is read out from a total of 1536 readout strips, 768 each in X- and Y-direction. For this measurement, the x-ray source is used with copper (Cu) filters and placed at a distance of 6.9 cm from the drift, as shown in Figure 3.10. The mini-X

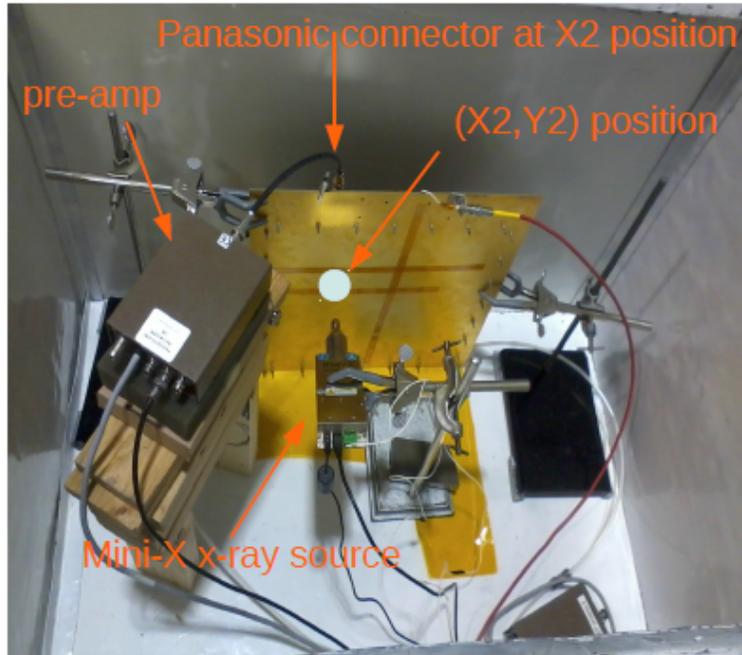


Figure 3.10: Experiment setup for gain measurement of a 30 cm × 30 cm triple-GEM detector with a 2D straight strip readout board.

x-ray source is operated at 49.7 kV voltage with 20.1 μA current. Due to the larger thickness of the drift of this detector, a higher setting on the mini-X x-ray is used for the signal detection. Since this detector has a 2D readout plane, the X-direction is considered in the vertical plane and Y-direction in the

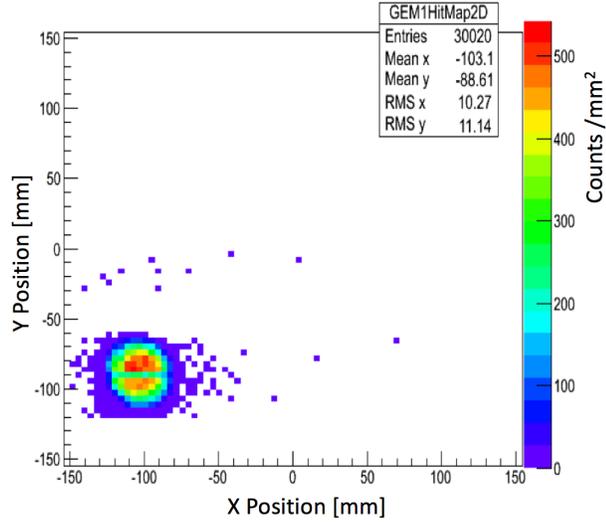


Figure 3.11: Beam profile of the x-ray source at position (X2,Y2).

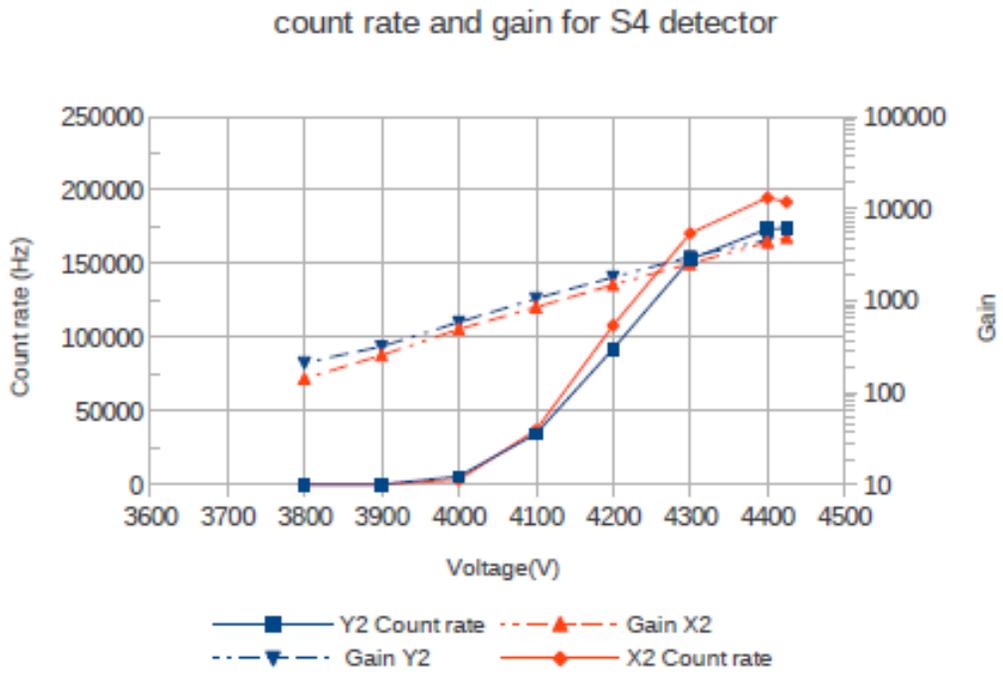


Figure 3.12: Rate plateau and gain of the 30 cm × 30 cm GEM detector on the order of $10^3 - 10^4$.

horizontal plane. The beam profile of the x-ray source at position (X2,Y2) is shown in Figure 3.11. The gain measurement is done using the same procedure as followed for the 10 cm \times 10 cm detector. Each plane of the detector has 6 Panasonic connectors, connected to APV25 hybrid chip. The scalable readout system (SRS) developed by RD51 collaboration [20] is used for data taking. SRS used with GEM detector typically consists of the Front-End Concentrator (FEC) which is built around a configurable FPGA with event buffer, Gigabit ethernet, I/O for trigger and clocks, and I/O for adapter cards such as ADCs. The rate plateau and gain measurement of the detector at (X2,Y2) position are shown in Figure 3.12. The gain of the detector is on the order of $10^3 - 10^4$.

1-m Long Trapezoidal Triple-GEM Detector with Zigzag Strips Readout

Finally, the gain of the 1-m long trapezoidal triple-GEM detector is measured using a zigzag radial strip readout as well as with a radial straight strip readout. The x-ray source with no filters is used with 10 kV accelerating voltage for the rate measurement, the signal is read out from the bottom of the third GEM foil. The gain is calculated again using Equation 3.1 and it is on the order of 10^4 . Figure 3.13 shows the gain of the 1 m long triple-GEM detector with the radial zigzag readout board [21].

Due to the high gain, good spatial resolution, and fast timing resolution, this detector is getting more popular in high energy physics experiments. For phase II upgrade of the muon endcap of the CMS experiments, the muon group is planning to install a GEM station (GE1/1) in the forward region. For this project, several prototypes of 1 m long triple-GEM detectors have been designed

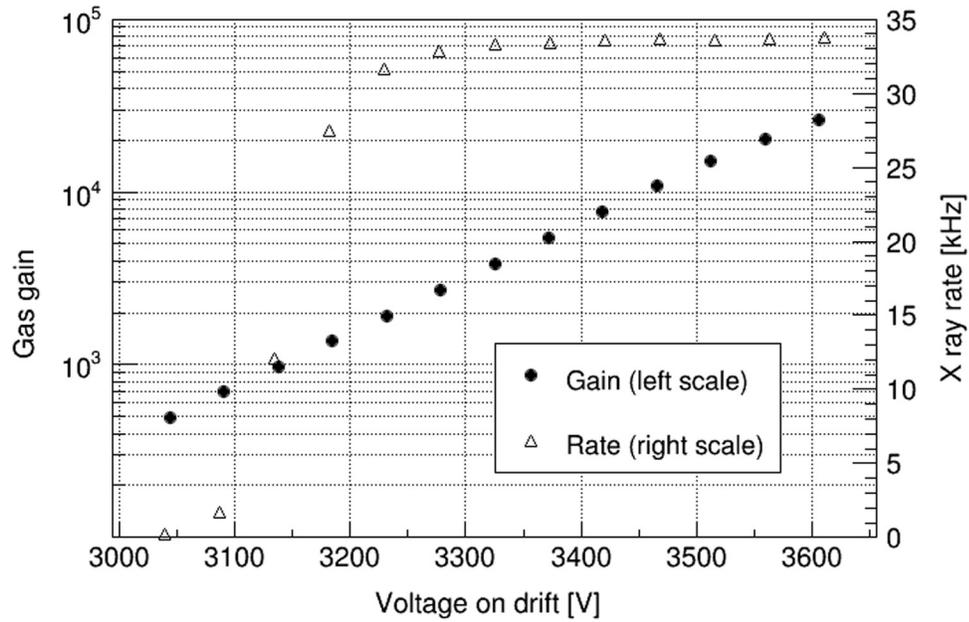


Figure 3.13: Rate and gain vs. V_{drift} measured at the center of the zigzag GEM detector with Ar/CO₂ 70:30.

and tested in several beams. At Florida Tech, a GE1/1 prototype-III was built and tested in a test beam at Fermilab. The characteristics were studied using the test beam data. These studies are detailed in the following sections.

3.2 Overview of the GEM Endcap

The CMS collaboration is going to install large-area GEM detectors in the forward muon endcap in the high- η region $1.5 < |\eta| < 2.2$. Figure 3.14 [22] shows the quadrant of the muon system, where installation of GE1/1, GE2/1, and ME0 detectors is proposed. After the LS2, the LHC will approach instantaneous luminosity of $2 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. This upgrade is important to maintain the high level performance and acceptable L1 trigger rate for muons without addition efficiency loss in endcaps as compared to the Run 1 and 2. This installation will help to restore redundancy for tracking and triggering in the muon system, as the GEM detectors provide very precise tracking information due to the high spatial resolution. They can also sustain high particle rates up to MHz/cm^2 . CSC alone misidentify lower p_T muons as high p_T muons, because they undergo multiple scattering in the steel absorbers.

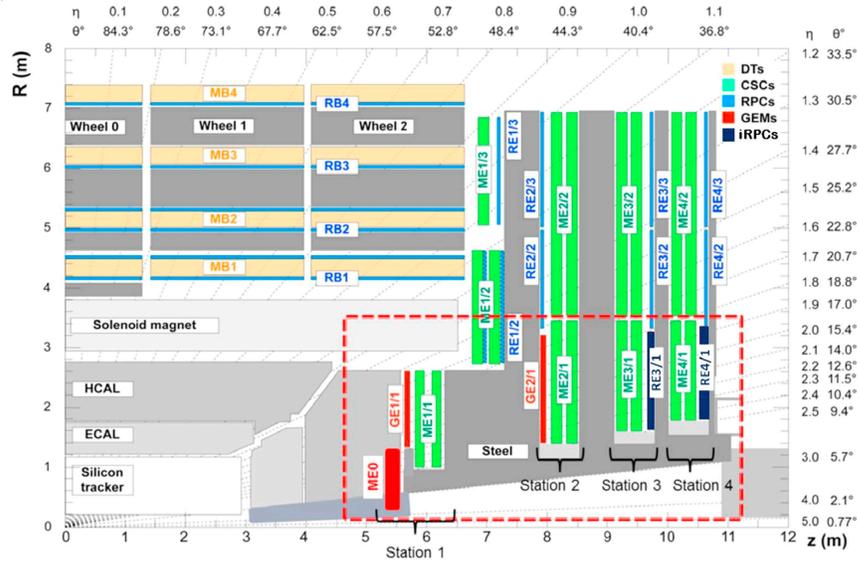


Figure 3.14: A quadrant of the CMS muon system with proposed upgrade (dashed box), showing different subsystems. GE1/1 is highlighted in red dashed box.

This problem can be overcome by using the GEM detectors with the CSC system as shown in Figure 3.15 [23], since there is only air between GE1/1 and ME1/1. Together they provide an accurate measurement of the muon bending angle unaffected by multiple scattering. This discriminates lower p_T muons from higher p_T muons and reduces the soft muon rate at the level-1 trigger, as shown in Figure 3.16 [24]. This will help to control the muon trigger rate at HL-LHC.

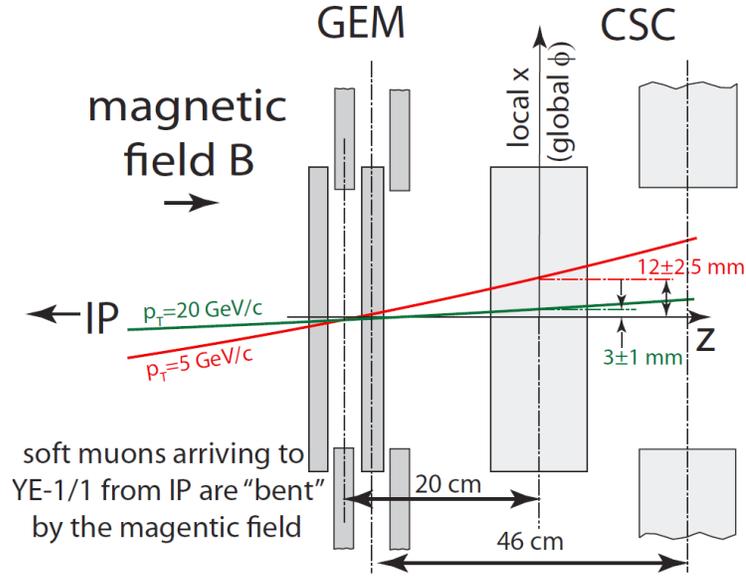
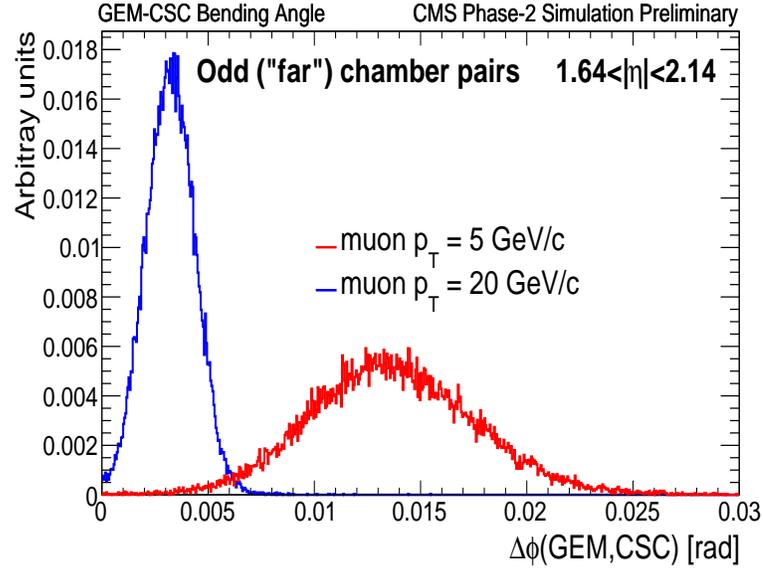
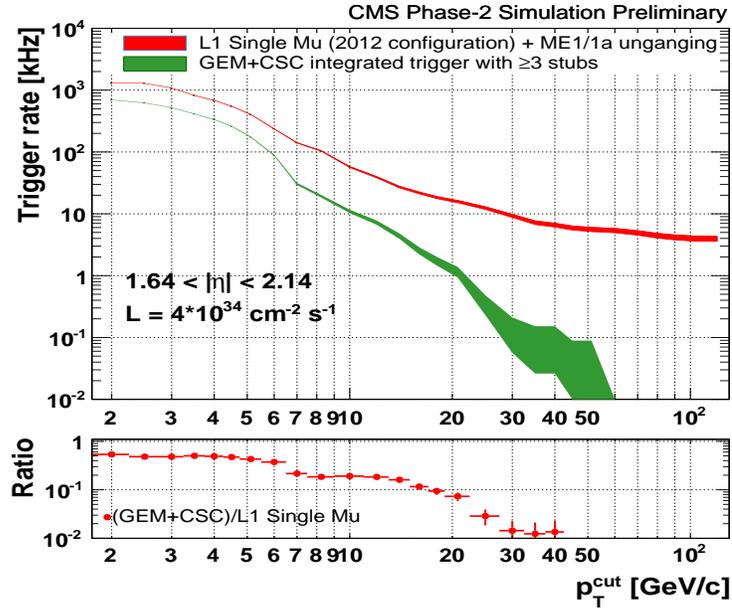


Figure 3.15: GEM and CSC systems enlarge the lever arm for the bending angle measurements.

In the GEM endcap, there will be 72 "super-chambers", each super-chamber is composed of two individual GE1/1 chambers, installed to complement the existing Muon Endcap (ME1/1) detectors to maximize the detection efficiency. Since each super-chamber covers a $\approx 10^\circ$ sector, there will be 36 super-chambers installed in each endcap. In each endcap long ($1.55 < |\eta| < 2.18$) and short ($1.61 < |\eta| < 2.18$) versions of these super-chambers alternate in the azimuthal



(a)



(b)

Figure 3.16: (a) Simulation of the bending angle measurement in the first endcap station (GE1/1-ME1/1) for soft (~ 5 GeV) and hard (~ 20 GeV) muons. (b) Simulation of the inclusive muon trigger rate expected for the LHC Phase II as a function of the Level-1 p_T trigger threshold for $1.6 < |\eta| < 2.1$.

direction (ϕ). The design of the super-chambers and their placement in the endcap is shown in Figure 3.17. The details about GE1/1 design and electronics

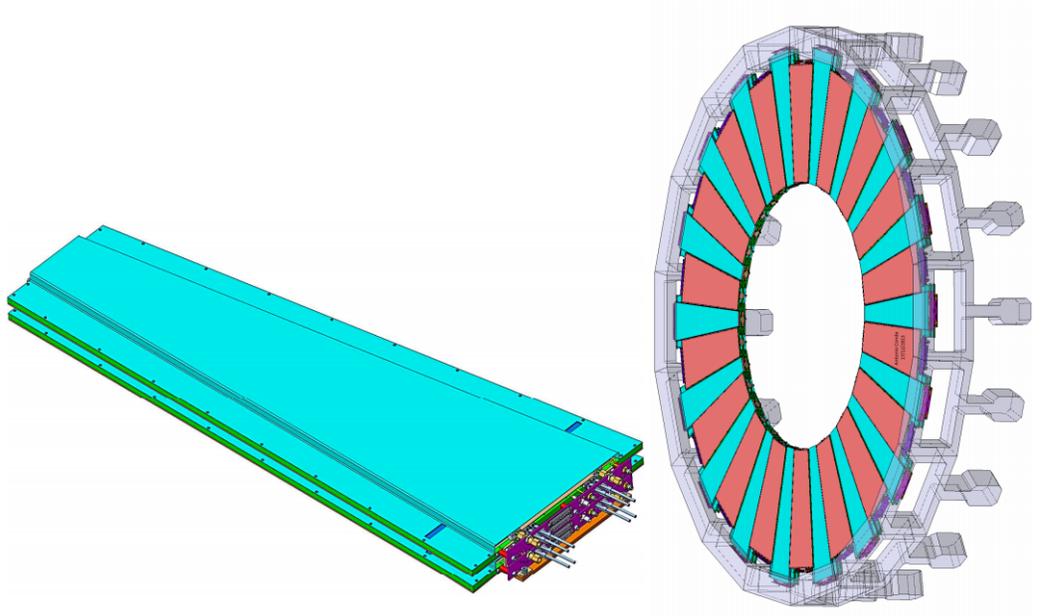


Figure 3.17: *Left*: Super-chamber. *Right*: Long and short GEM chambers in one endcap.

can be found in [14].

A third generation GE1/1 prototype GEM detector was constructed at Florida Tech and tested in a hadron beam at Fermilab. Its performance characteristics, such as strip cluster parameters, detection efficiency, and spatial resolution, have been studied using pulse height information and binary hit reconstruction. In the following section, the construction, experimental setup, and performance characteristics of the GE1/1 prototype-III detector with Fermilab test beam data are explained.

3.3 Fermilab Test Beam Analysis

3.3.1 Construction of a GE1/1 GEM Detector Prototype

The GE1/1-III prototype detector is a trapezoidal triple-GEM detector with an approximately $99 \times (28-45)$ cm² active surface area. This detector has a 3/1/2/1mm (drift, transfer 1, transfer 2, readout) internal electrode gap configuration. The GEM foils used in this detector are produced by a single-mask etching technique at CERN. These three GEM foils are mounted on the drift electrode and enclosed by the readout board. This detector is constructed using the internal mechanical stretching method [25] introduced in 2011. Each GEM foil is divided into 35 high voltage sectors that are transverse to the direction of the readout strips. There is a total of 3072 radial readout strips with a $455 \mu\text{rad}$ pitch along the length in the ϕ -direction and distributed over eight η -sectors. In each of the η -sectors, induced signals are read out via 384 radial strips through vias in the readout board to radial strips. Figure 3.18 [14] shows cross section through inner and outer chamber frames along with GEM foils and Figure 3.19 shows the main steps involved in the assembly of the GE1/1-III prototype detector by the author.

As shown in Figure 3.19 (a), the first step in the construction is to produce the stack of the three GEM foils. Using the inner frame pieces one foil is placed on the top of the other. The thickness of the inner frame pieces is chosen in such a manner that the internal gap configuration is maintained as 3/1/2/1 mm. Inner frame screws are used to hold the inner frame pieces and foils together. Once the stack is complete, excess kapton foil along the outer edge of the inner frames is removed and the foil stack is transferred to the drift board.

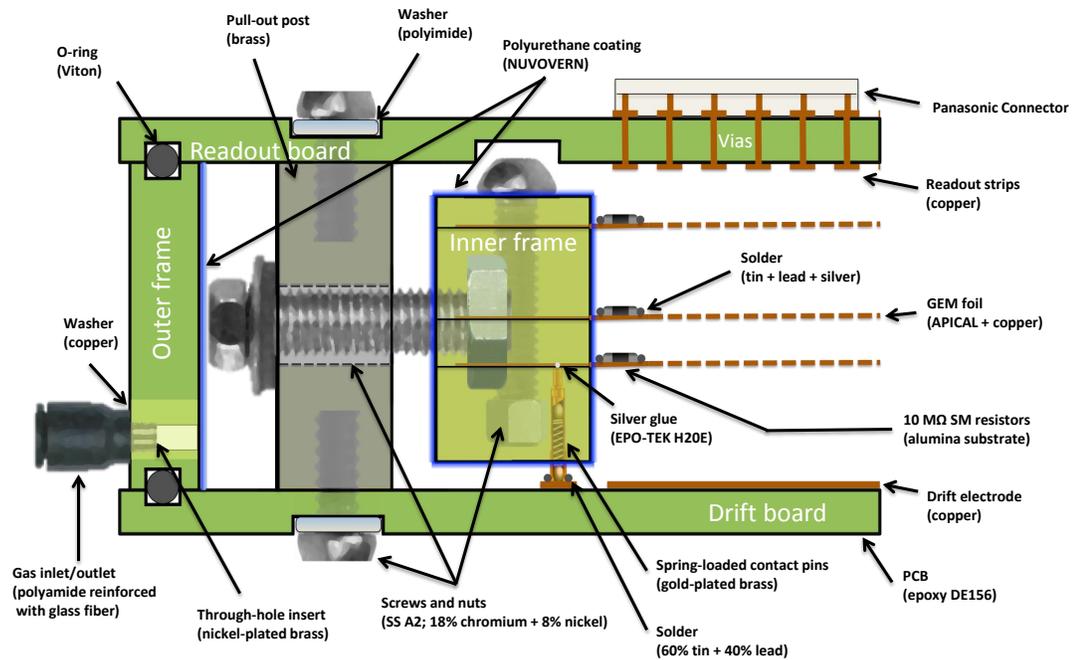
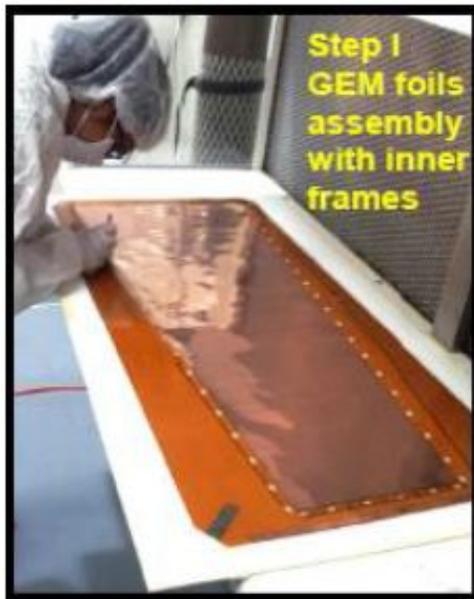


Figure 3.18: Cross section through inner and outer chamber frames and GEM foils that shows how the GEM foils are mounted within the GE1/1 chamber so that they can be mechanically tensioned against the brass pull-out posts without deforming the drift or readout boards. The materials of all chamber components are specified.

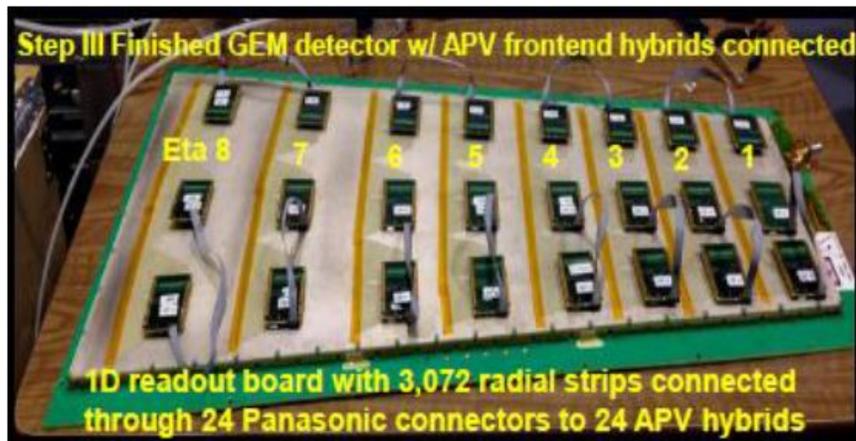
In second step, outer frame screws are inserted horizontally in an outer frame which is attached to the drift board, to connect to the inner frame pieces of the GEM stack. Foils are then stretched first along the long sides of the detector by tightening the outer frame screws evenly. Uneven stretching can affect the charge uniformity of the detector. Hence, it is very important to provide proper tension along all sides. Figure 3.19 (b) shows the stretching step. In the third step detector is closed by mounting the readout board on the drift. The outer frame on the drift has an o-ring on the surface to make the detector gas tight.



(a)



(b)



(c)

Figure 3.19: Construction of the 1m-long large-area GE1/1-III prototype in three steps: (a) Step I: GEM foil assembly with inner frames. (b) Step II: GEM foils stretching. (c) Step III: Closed GEM detector with readout.

Again, it is important to provide proper tension while closing the detector as uneven tension can change the gap between the drift, foils, and the readout and eventually affect electric fields and ultimately the gain of the detector. By measuring gas flow difference in the output and input gas line, gas tightness can

be checked. It is important to flush the detector with nitrogen gas at least for 4 volume exchanges to make sure the foils are moisture free. The gas is then swapped with the Ar/CO₂ gas mixture. It is important to wait again at least 4 times volume gas exchange before applying the high voltage.

3.3.2 Test Beam Setup

The GE1/1-III prototype detector was tested in a 32 GeV hadron beam at the Fermilab test beam facility in October 2013 as shown in Figure 3.20. For tracking studies, this detector was positioned on a movable table between four 2-D readout GEM detectors. Three of them were 10 cm × 10 cm GEM detectors and one was 50 cm × 50 cm detector with an active 10 cm × 10 cm area. These 2D GEM detectors contained 256 straight strips along each horizontal (y-coordinate) and vertical (x-coordinate) plane with a 0.4 mm pitch. The data were collected using RD51 scalable readout system (SRS) [26] with the external trigger provided by scintillators in coincidence. During this beam test all the detectors used an Ar/CO₂ gas mixture with a 70:30 ratio.

3.3.3 Beam Test Results

Performance Characteristics

The performance characteristics of the GE1/1 detector were studied using high voltage and position scan data. Figure 3.21 shows the detector with eight η -sectors, number from smaller side to longer side. The three row of APV25 are referred as Upper, Middle, and Lower APV rows for position scan. In a high voltage scan, the beam was focused on η -sector 5 and drift the voltage

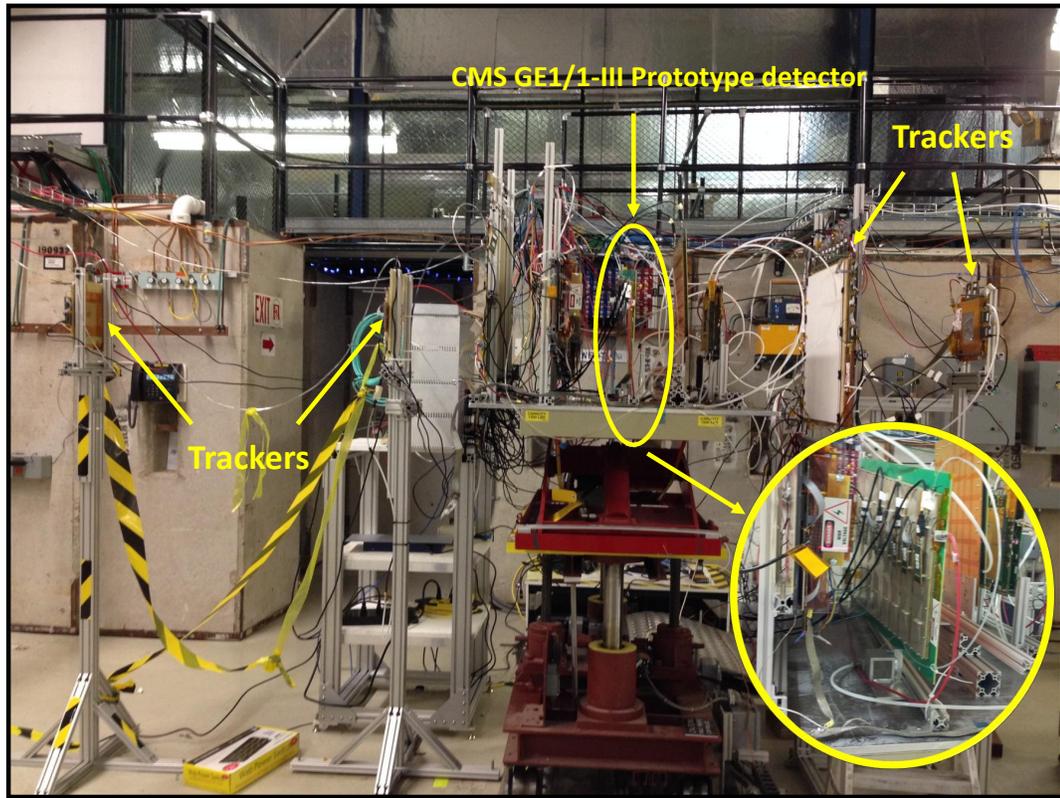


Figure 3.20: Experimental setup in the beam line at Fermilab. The zoomed picture show the GE1/1 detector placed on the moveable table between the four tracker detectors.

of the GE1/1 detector varied from 2900 V to 3350 V, whereas voltages of all tracker detectors were kept fixed in such a way that the efficiency was on plateau throughout all measurements. The charge collected from group of strips are referred as the cluster charge. Figure 3.22 shows the cluster charge distribution for η -sector 5 at 3250 V. The distribution is fitted with a Landau function and the Most Probable Values (MPV) obtained from this fitting are used for obtaining the uniformity results for the GE1/1 detector.

The number of strips fired in strip clusters define the strip multiplicity of strip clusters. For the GE1/1 detector, distribution of the strip multiplicity in

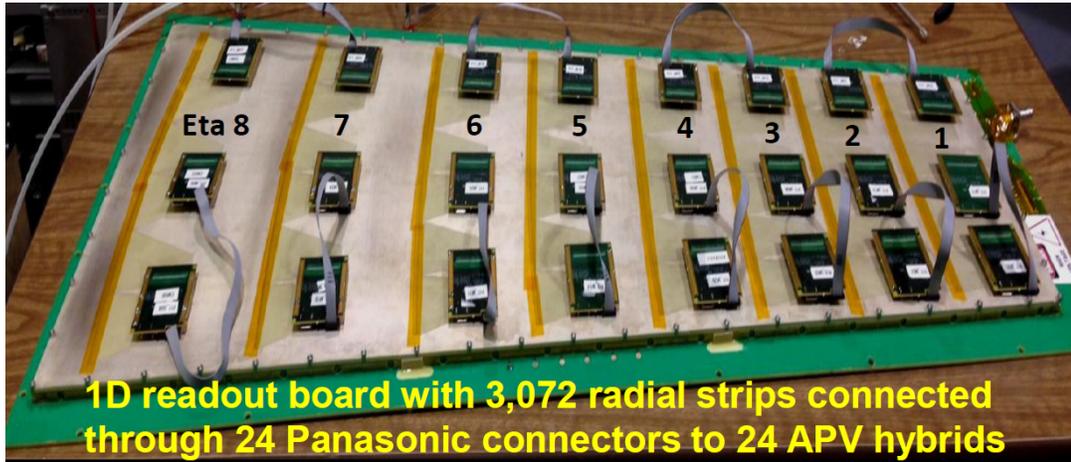


Figure 3.21: A GE1/1-III triple-GEM prototype detector with 24 APV25 hybrids connected to strips via 24 Panasonic connectors.

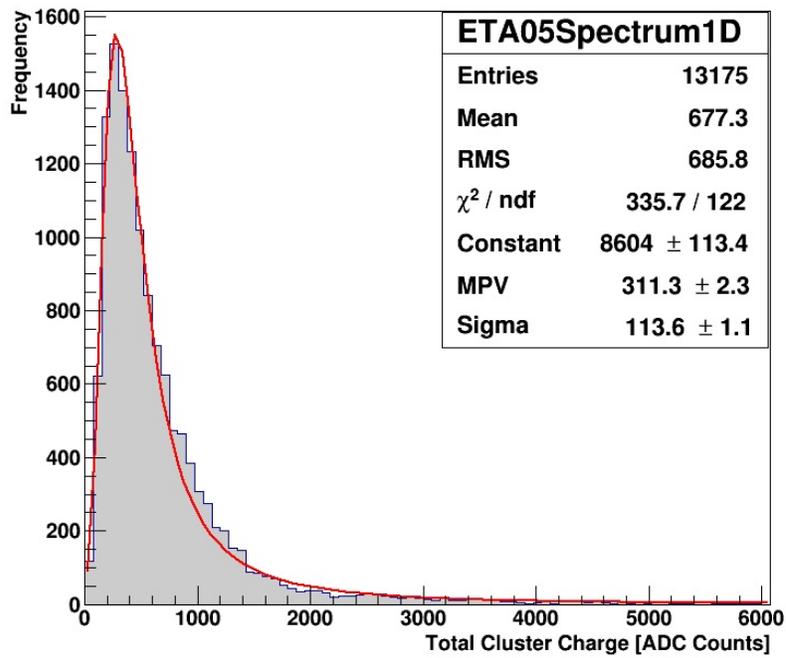


Figure 3.22: Cluster charge distribution at 3250 V fitted with Landau function.

strip clusters is shown in Figure 3.23 with the average strip multiplicity (cluster size) of 2.4 strips. Figure 3.24 shows that the strip multiplicity increases with

high voltage, i.e. with the gas gain.

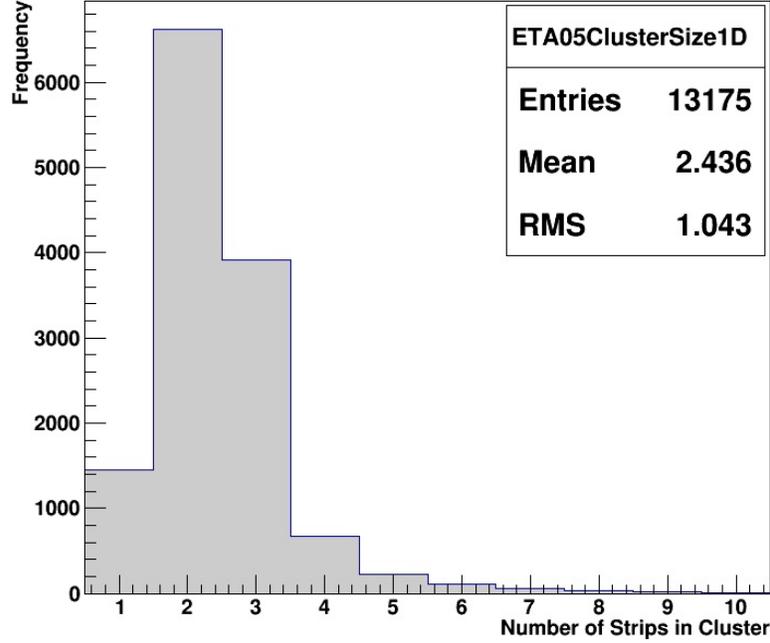


Figure 3.23: Strip multiplicity in strip clusters size at 3250 V in the η -sector 5.

Performance results of the GE1/1 detector are presented in three sections. Section one explains the charge measurement performed for the electronics upgrade. Efficiency results are described in the section two and finally, in section three tracking results are explained in details:

Measurement of the Charge Induced on the Readout Strips of a GE1/1 Detector

For the GE1/1 system, the CMS experiment proposes VFAT3 readout electronics, which produce binary hit output for each readout strip. Similar to APV25 hybrid, VFAT3 front-end chip also has 128 channels. The chip provides "fast

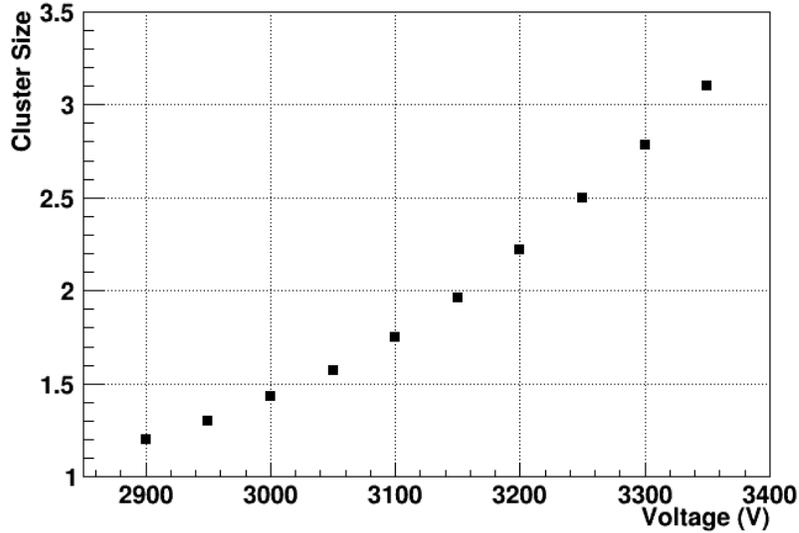


Figure 3.24: Strip cluster size increases with high voltage.

OR” fixed-latency trigger information and full granularity tracking information. The details about the VFAT3 and overall readout electronics for the GE1/1 system can be found in [14]. The matching of the dynamic range of the induced charge to the dynamic range of the chip input determines the quality with which the signal is read from each strip. Consequently, it is important to optimize the dynamic range of the chip with respect to the expected input charge. Following sections describe measurements of the most probable values (MPV), mean strip charges and cluster charges, as well as the 99th percentile of the charge distribution to determine the requirement on the dynamic range of the input charge for the VFAT electronics.

Charge Conversion from ADC to fC

To calculate the dynamic range for VFAT electronics, the charge must be converted from ADC counts to femto-Coulombs (fC). For the precise conversion from ADC to fC, the APV25 calibration data, shown in Figure 3.25 [27], is used. The APV25 shows a linear charge response up to around 800 ADC counts, and exponential response for higher ADC counts. For precise conversion, it is important to apply the non-linear correction for higher ADC counts (≥ 800).

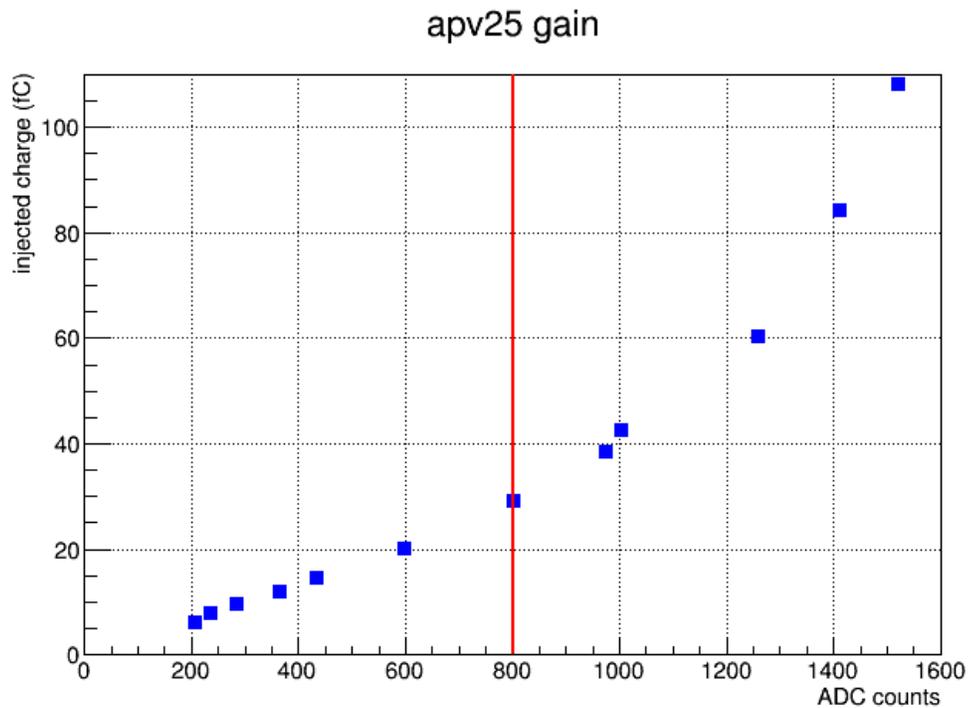
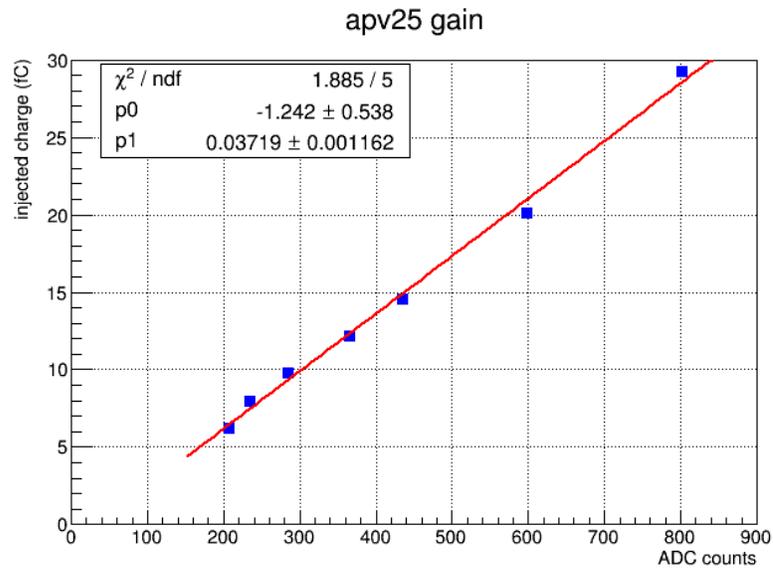
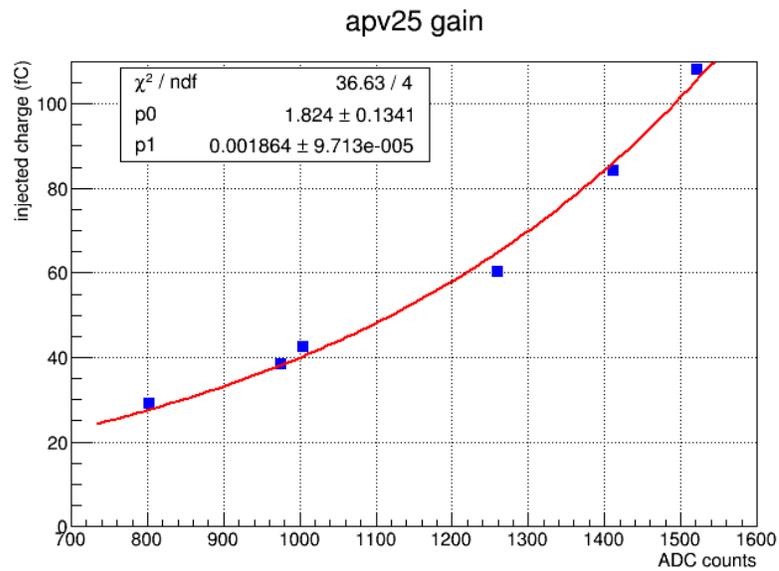


Figure 3.25: Charge-to-ADC count calibration for APV25 chip.

Figure 3.26 shows how the calibration data are split into two parts, one for the linear correction for lower ADC counts (< 800) and the other for the non-linear correction for higher ADC counts (≥ 800).



(a)



(b)

Figure 3.26: (a) ADC-to-fC conversion for lower ADC counts with linear fit. (b) ADC to fC conversion for higher ADC counts with exponential fit.

From the linear and exponential fit parameters, the ADC-to-fC conversions can be expressed as follows:

$$Q \text{ (in fC)} = 0.03719 \times \text{ADC} \quad \text{for ADC} < 800 \quad (3.3)$$

$$Q \text{ (in fC)} = \exp[1.8 + (0.0018 \times \text{ADC})] \quad \text{for ADC} \geq 800 \quad (3.4)$$

These factors give precise conversion values. Conversion factors are used in the following charge distribution studies to express the charge in both raw ADC counts and fC units.

Charge Distribution Measurements

The high voltage scan data is used for the charge distribution studies. For each voltage starting from 2900 V to 3350 V, the charge distribution is plotted and fitted with a Landau function. Fit parameters such as Most Probable Value (MPV), mean, and sigma are used for determining the range of the charges. The charge distribution is plotted for two cases, an individual strip charge and a total cluster charge. The strip multiplicity of the cluster is selected to compare the strip charge distribution as follows: all strip clusters, ≥ 2 -strips, 4-strip, 3-strip, 2-strip, and 1-strip clusters. Similarly, for the total cluster charge, all strips cluster, ≥ 2 -strips cluster, and 1-strip cluster were used for plotting the distribution. Finally, the 99th percentile of the maximum charge of the strip charge distribution is plotted to determine the required full dynamic input charge range for the VFAT electronics.

The cluster charge distribution is studied for this large-area triple GEM detector with two different readout designs, one with radial straight-strip readout

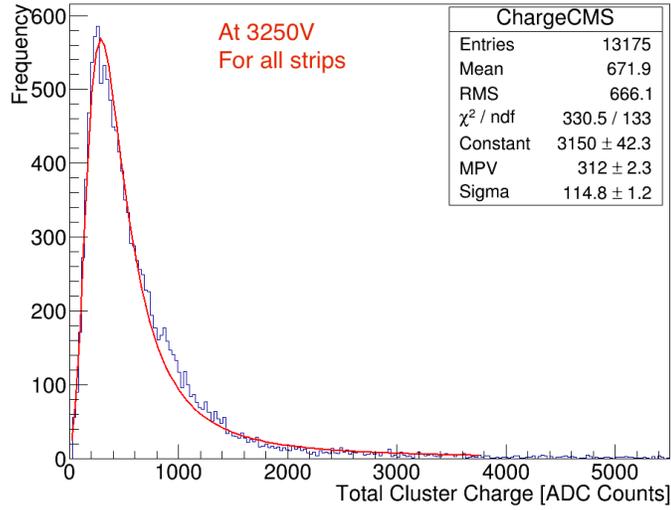
and the other with radial zigzag strips. A radial zigzag strip readout board [21] for a one-meter-long GEM detector was designed by Florida Tech for the Electron-Ion Collider (EIC) experiment. In the Fermilab test beam, the readout board of the GE1/1-III detector was replaced with this radial zigzag strip readout and its performance was tested again. As shown in Figure 3.27, this readout board has 1,072 radial zigzag strips distributed along the eight η -sectors with 128 strips per sector and the signal can be read out from the entire chamber using only eight APV25 hybrids. This readout board design is cost-effective because it reduces the number of channels for the readout electronics by a factor of three.



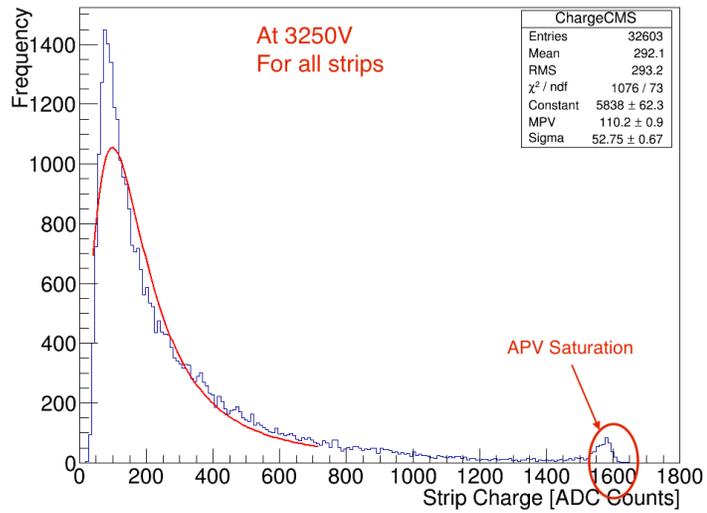
Figure 3.27: A large-area GEM detector with radial zigzag strips read out with eight APV hybrids connected to eight η -sectors through Panasonic connectors.

The total number of electrons produced in the ionization process follows the Landau distribution and the total charge produced in the detector can be characterized using its distribution parameters.

Figure 3.28 (a) shows the total strip cluster charge distribution for all strip clusters at the operating voltage of the detector, i.e. at 3250 V. The distribu-



(a)



(b)

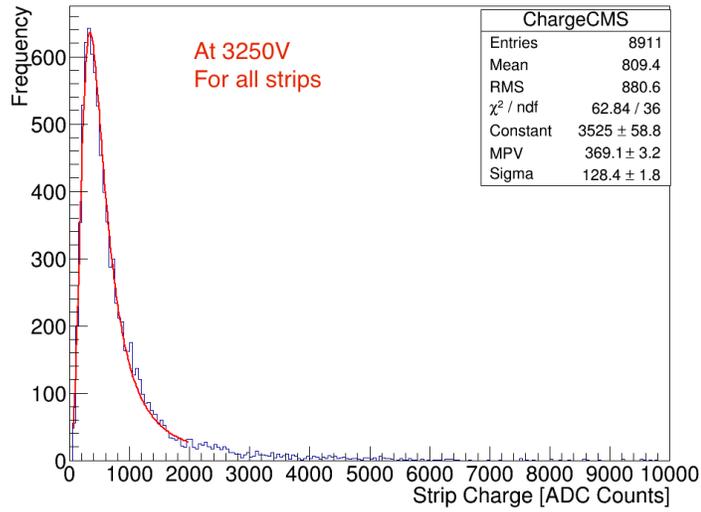
Figure 3.28: Charge distribution measured with GE1/1-III detector and fitted with a Landau function: (a) Total cluster charge distribution. (b) Individual strip charge distribution.

tion is fitted with a Landau function and the mean charge value is found to be 671.9 ADC counts, which is equivalent to 25 fC. For Ar/CO₂ 70:30 gas, the

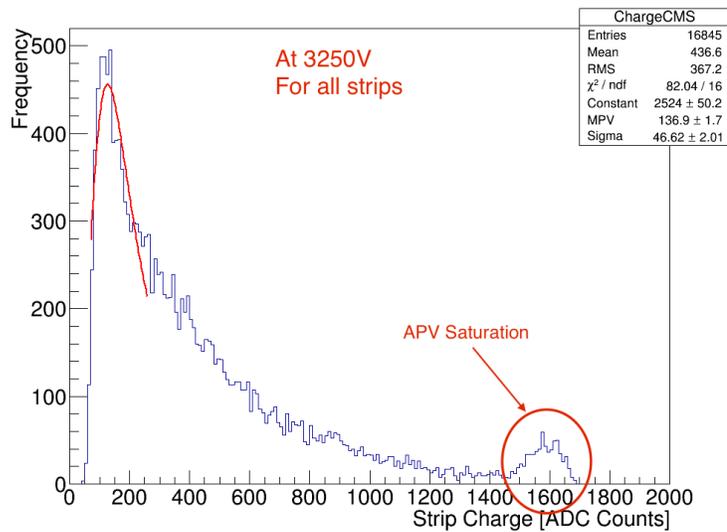
total mean number of primary electrons in the 3 mm drift gap is $\langle N_{tot} \rangle = 29$ electrons. This prototype detector was operated at a gain of around 8000 giving an expected mean charge ($\langle N_{tot} \rangle \times \text{Gain}$) of 232,000 e, equivalent to 37 fC or 964 ADC counts. The measured mean cluster charge is within 27% of the theoretical value, which is acceptable. The individual strip charge distribution is shown in Figure 3.28 (b). In this distribution, APV saturation is observed around 1600 ADC counts and these saturated events are excluded from the calculation of the 99th percentile of the maximum charge, which determines the dynamic range of the input charge for VFAT electronics. Similarly, the charge distribution is plotted for the GEM detector with the zigzag readout board. Figure 3.29 shows the total cluster charge and the individual strip charge distributions. These distributions are again fitted with the Landau function and the resulting mean, sigma, and MPV variable are plotted against the drift voltage. Since Ar/CO₂ gas is used in the same proportion and with the same drift and transfer gaps in the detector as above, the calculation of the expected mean charge for this configuration is the same i.e. 37 fC. From the Landau distribution of the total cluster charge, the measured mean charge for this detector at the operating voltage is about 26 fC, similar to the result for the straight strips. In Figure 3.29 (b) saturation of APVs is again seen to be around 1600 ADC counts and these events are again excluded from the analysis.

Mean Charge Distribution

The mean strip charge at operating voltage can be taken as the typical charge input for the electronics. A mean of the charge distribution for the above mentioned range of the drift voltages is plotted using different cuts on the



(a)

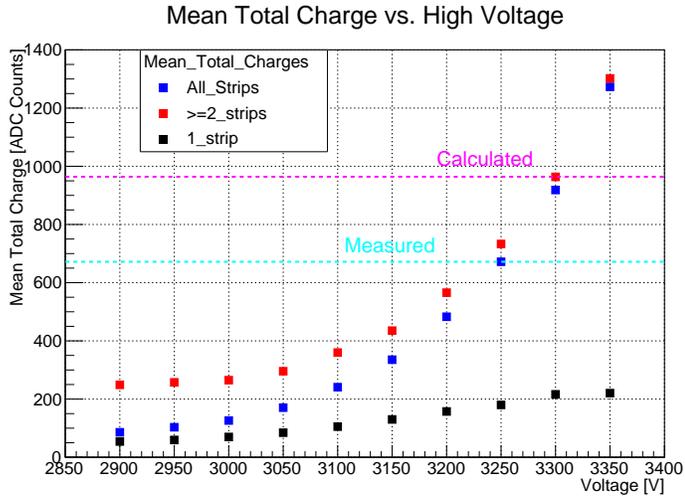


(b)

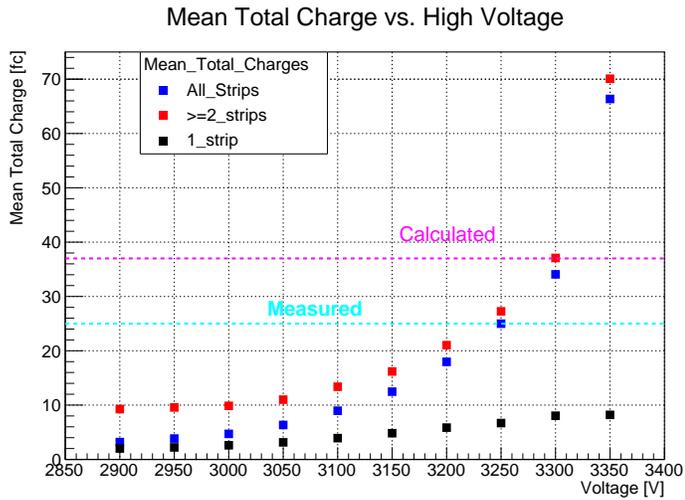
Figure 3.29: Charge distribution of the GEM detector with radial zigzag strip readout fitted with the Landau function: (a) Total cluster charge distribution. (b) Individual strip charge distribution.

cluster strip multiplicities. Figures 3.30 and 3.31 show mean charges for the GE1/1 detector in ADC counts and fC units for the total cluster charge and

individual strip charge. At 3250 V, the measured mean cluster charge is 25 fC, while the overall mean strip charge is 11 fC.



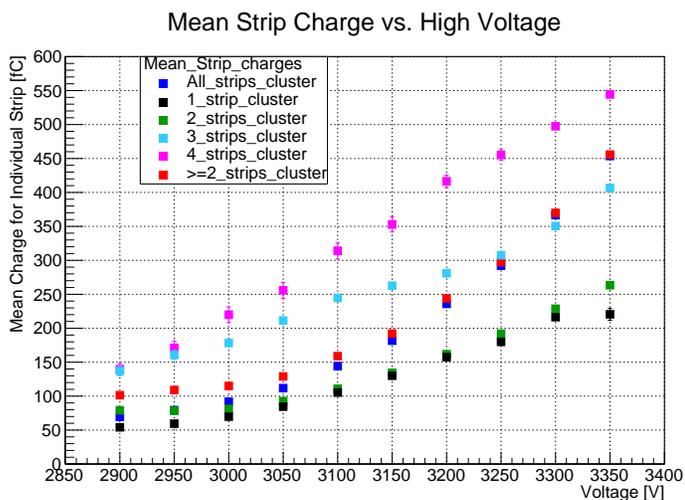
(a)



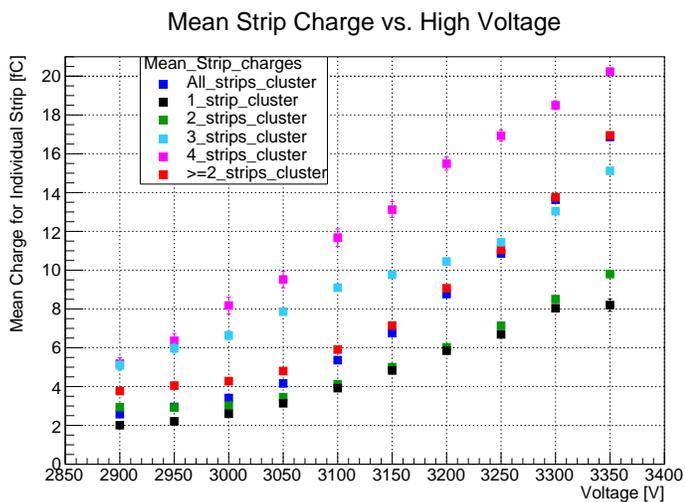
(b)

Figure 3.30: GE1/1 Detector: Means of the total cluster charge distribution counts vs. drift voltage (a) in ADC. (b) in the fC unit.

Similarly, mean charge values from Landau fits are plotted against the drift voltages of the GEM detector with radial zigzag strip readout. The mean of



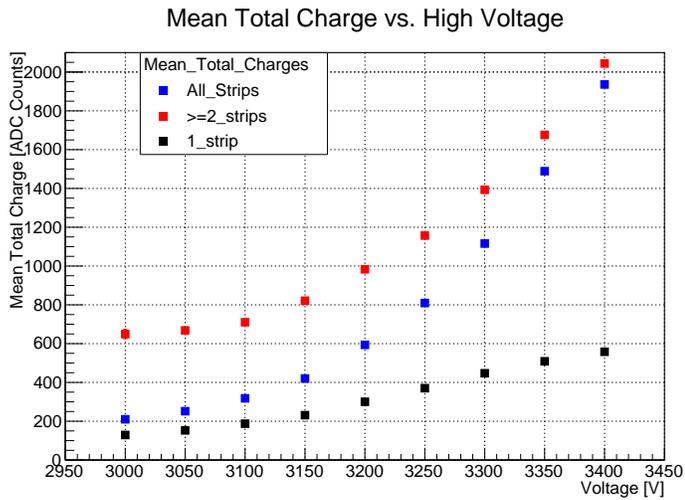
(a)



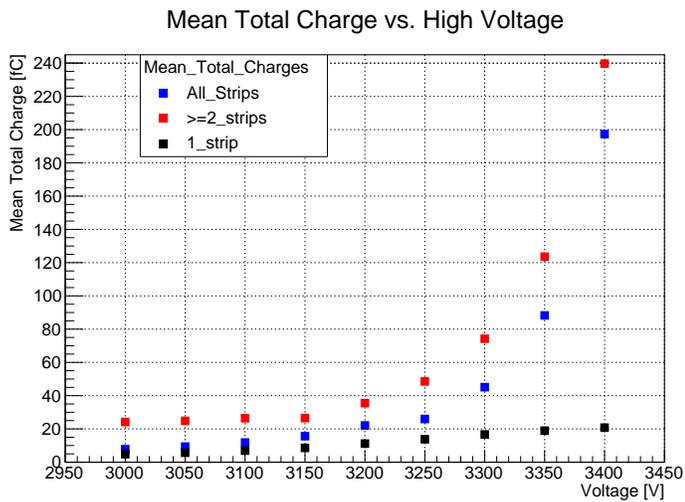
(b)

Figure 3.31: GE1/1 Detector: Means of the individual strip charge distribution counts vs. drift voltage (a) in ADC. (b) in the fC unit.

the total cluster charge and mean of the individual strip charge are shown in Figures 3.32 and 3.33, respectively. The mean charge value increases with higher drift voltage i.e higher gas gain. In Figure 3.33 (b), the mean strip charge is ~ 26 fC at the operating voltage. The mean cluster charge is similar to what



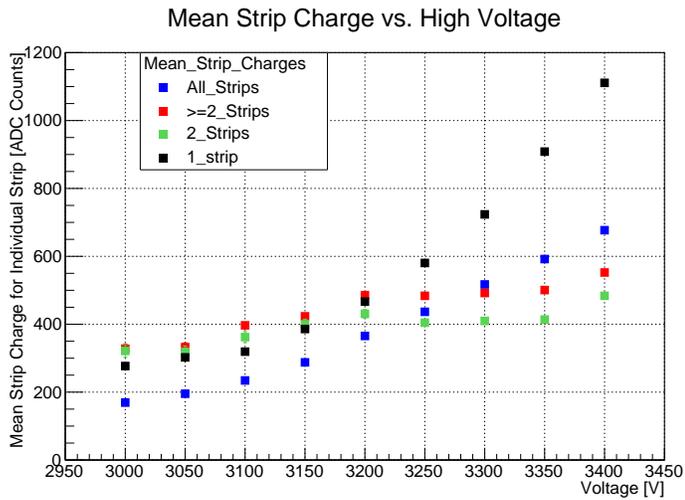
(a)



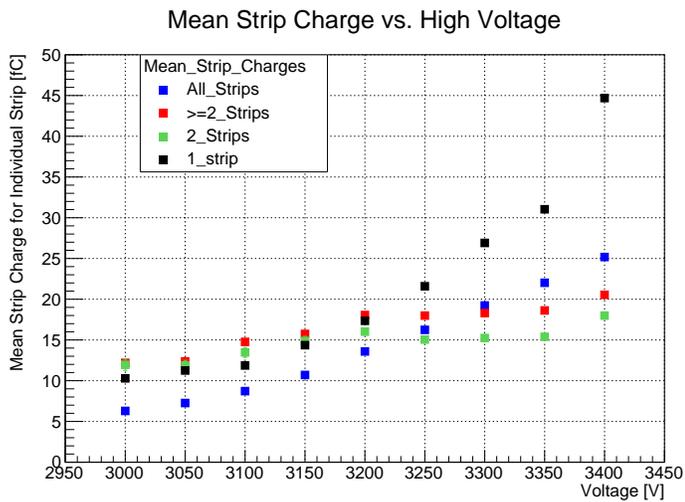
(b)

Figure 3.32: GEM detector with zigzag readout: Means of the total cluster charge distribution counts vs. drift voltage (a) in ADC. (b) in the fC unit.

is observed with the radial straight-strip readout, while the mean strip charge value for radial strips is higher than the straight-strip readout because more charge is induced on an individual zigzag strip due to its larger width.



(a)



(b)

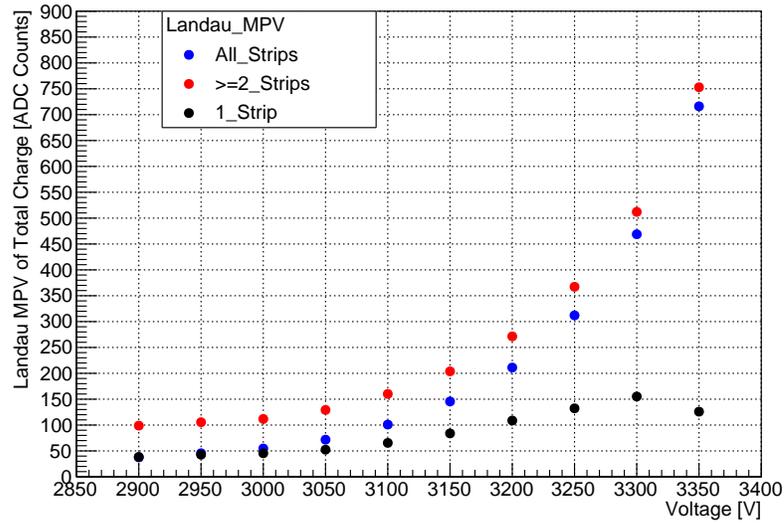
Figure 3.33: GEM detector with zigzag readout: Mean of the individual strip charge distribution counts vs. drift voltage (a) in ADC. (b) in the fC unit.

MPV and Sigma Distributions

In this section, all Landau MPV and Landau sigma value plots for the total cluster charge and the individual strip charge distribution are summarized for reference. Figures 3.34 and 3.35 show MPV and sigma vs. drift voltage for the

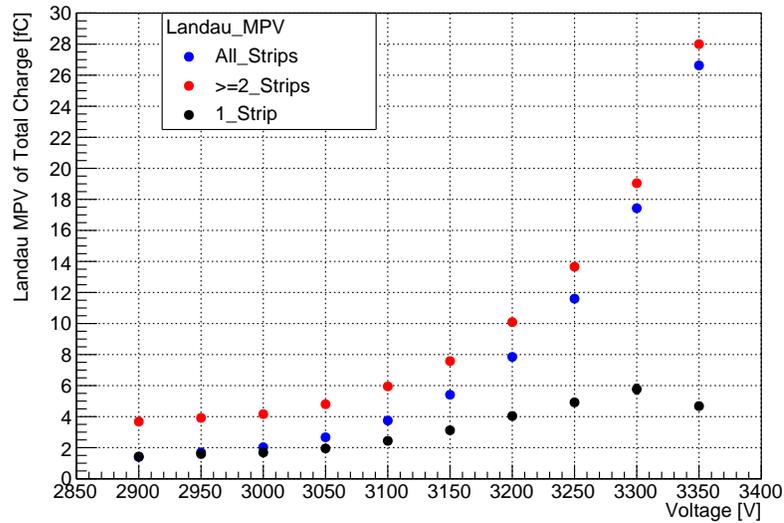
GE1/1 detector, while Figures 3.36 and 3.37 summarize the corresponding plots for the GEM detector with the radial zigzag strip readout.

Landau MPV vs. High Voltage

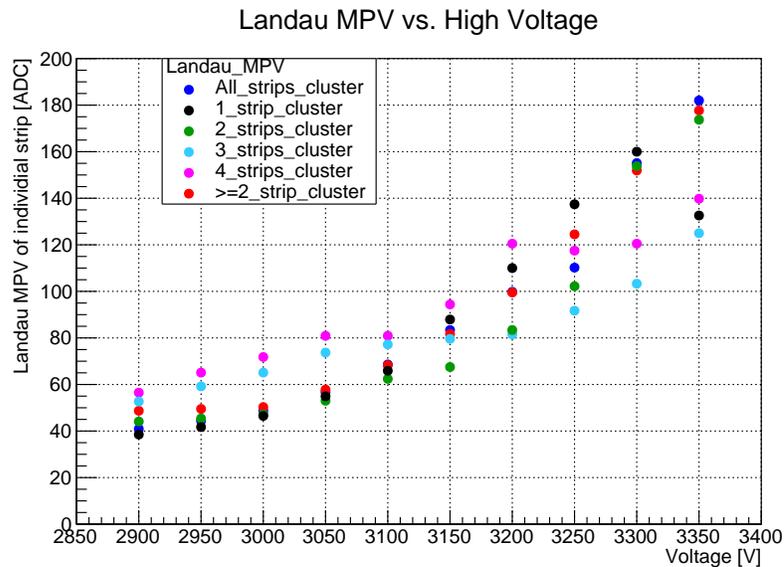


(a)

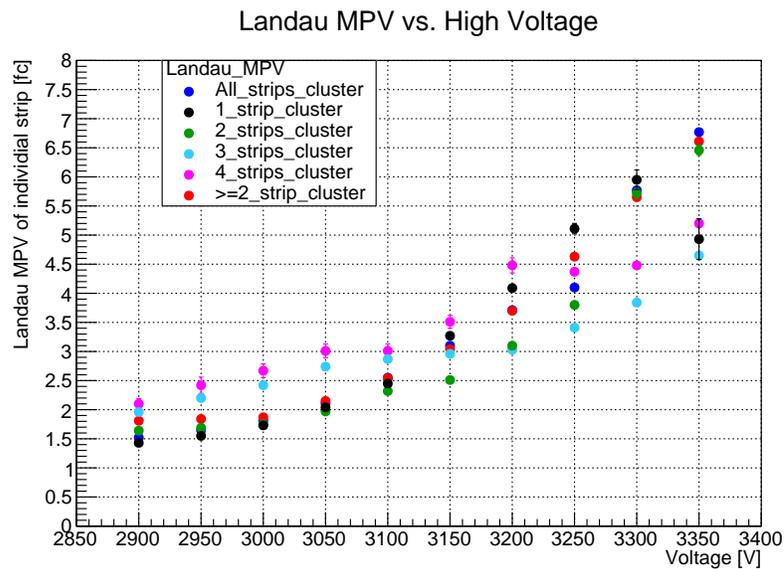
Landau MPV [fC] vs. High Voltage



(b)



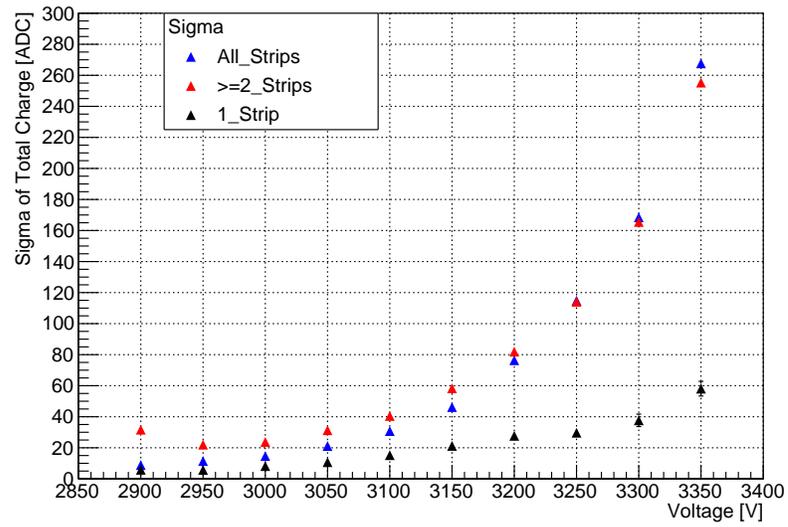
(c)



(d)

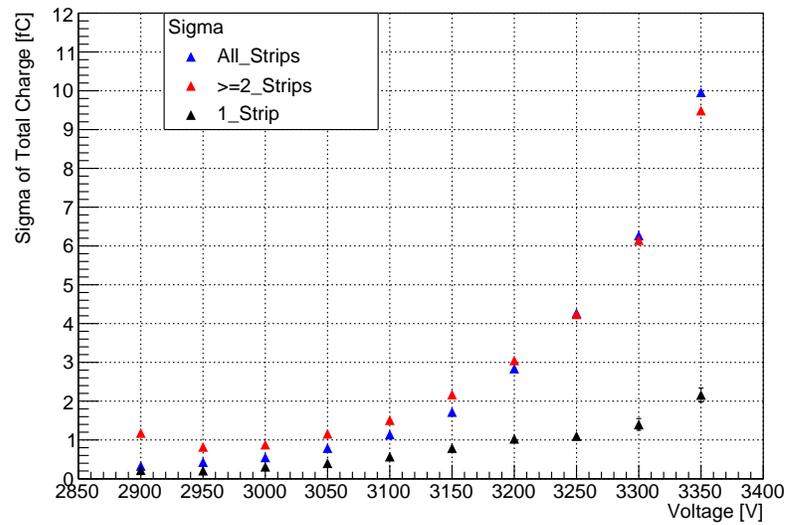
Figure 3.34: GE1/1 Detector: MPV of the total cluster charge distribution counts vs. drift voltage (a) in ADC. (b) in the fC unit. MPV of the individual strip charge distribution counts (c) in ADC. (d) in the fC unit from Landau fit.

Sigma vs. High Voltage

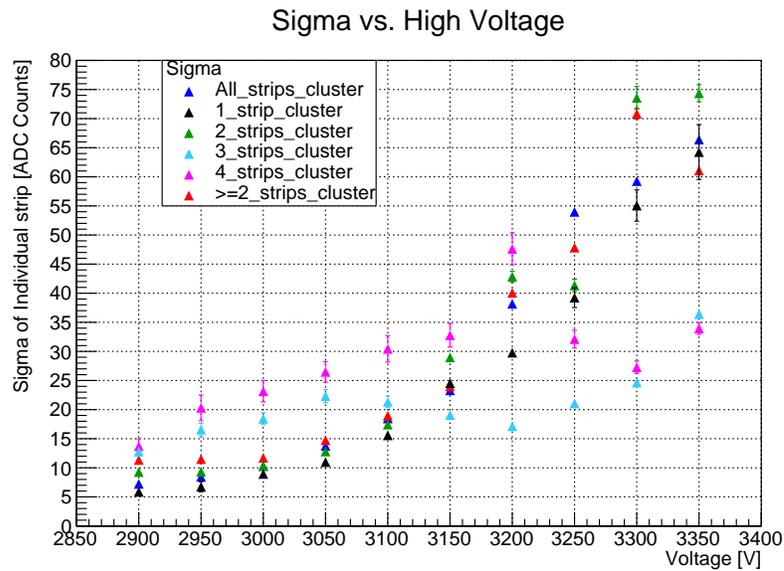


(a)

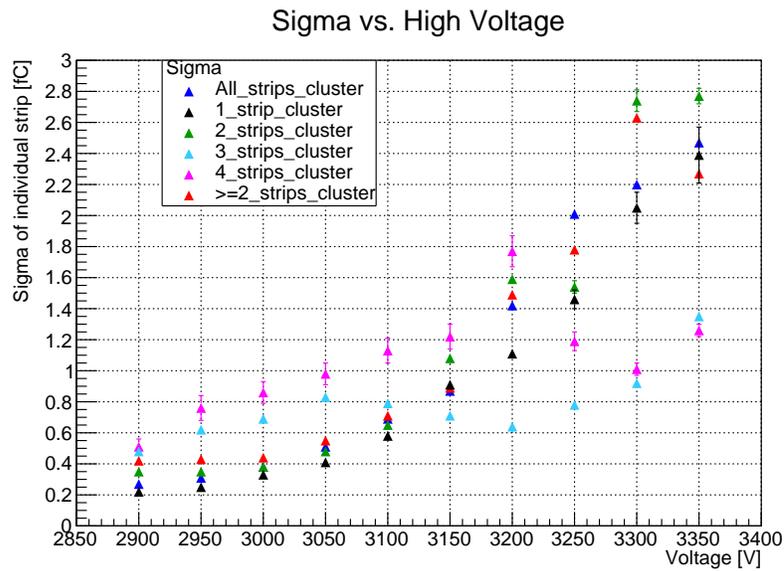
Sigma vs. High Voltage



(b)



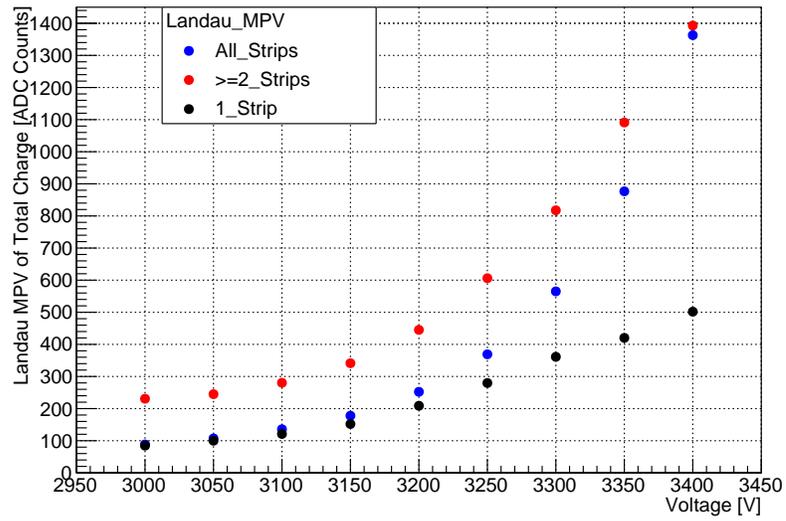
(c)



(d)

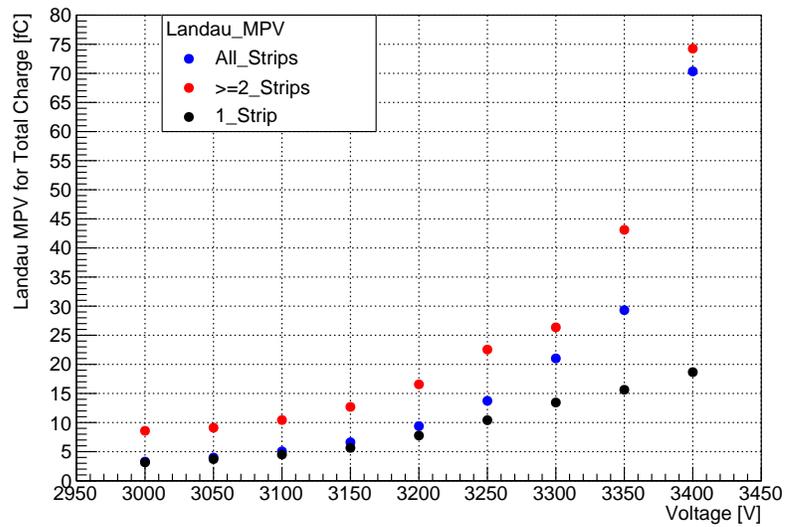
Figure 3.35: GE1/1 Detector: Sigma of the total cluster charge distribution counts vs. drift voltage (a) in ADC. (b) in the fC unit. Sigma of the individual strip distribution vs. drift voltage (c) in ADC. (d) in the fC unit from Landau fit.

Landau MPV vs. High Voltage

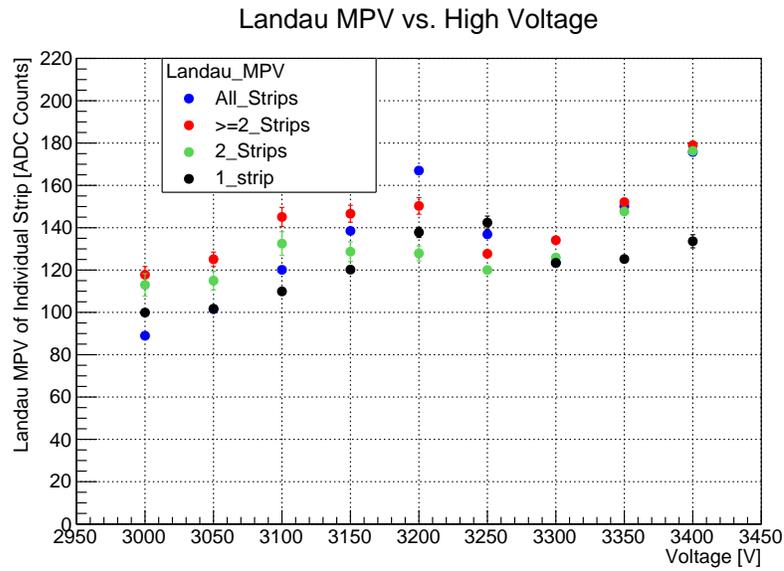


(a)

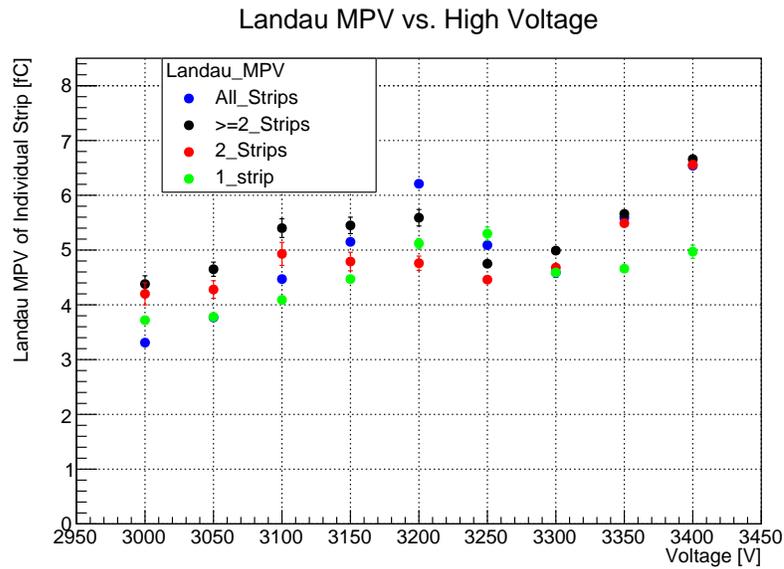
Landau MPV vs. High Voltage



(b)



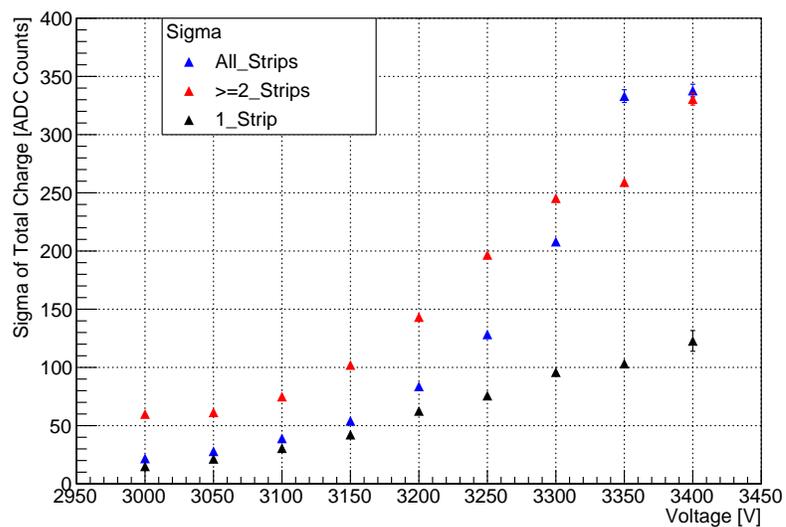
(c)



(d)

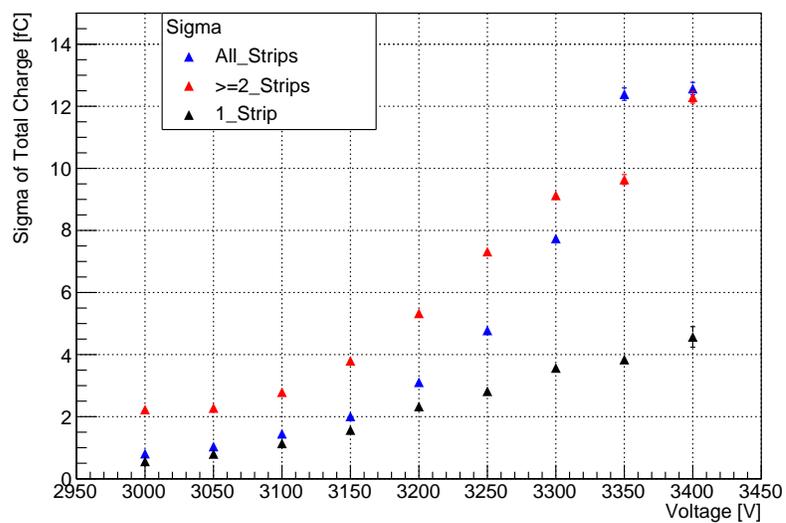
Figure 3.36: GEM Detector with the radial zigzag strip readout: MPV of the total cluster charge distribution counts vs. drift voltage (a) in ADC. (b) in the fC unit. MPV of the individual strip charge distribution (c) in ADC (d) in the fC unit from Landau fit.

Sigma vs. High Voltage

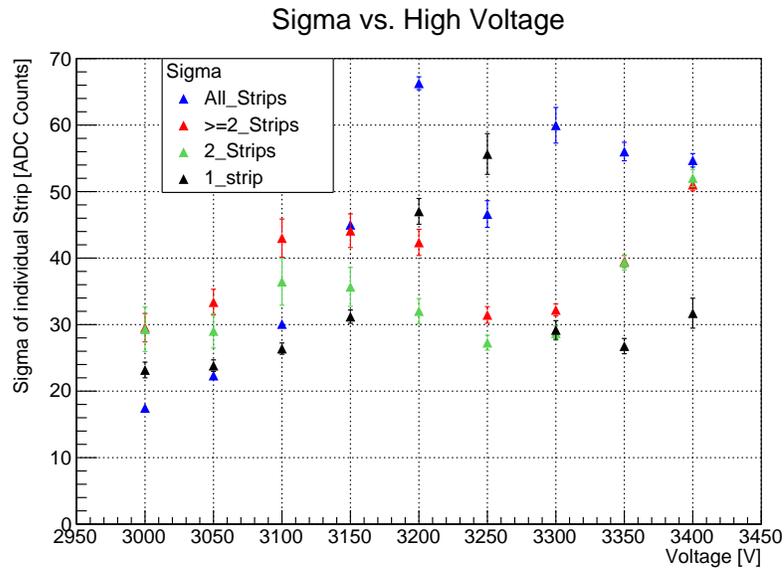


(a)

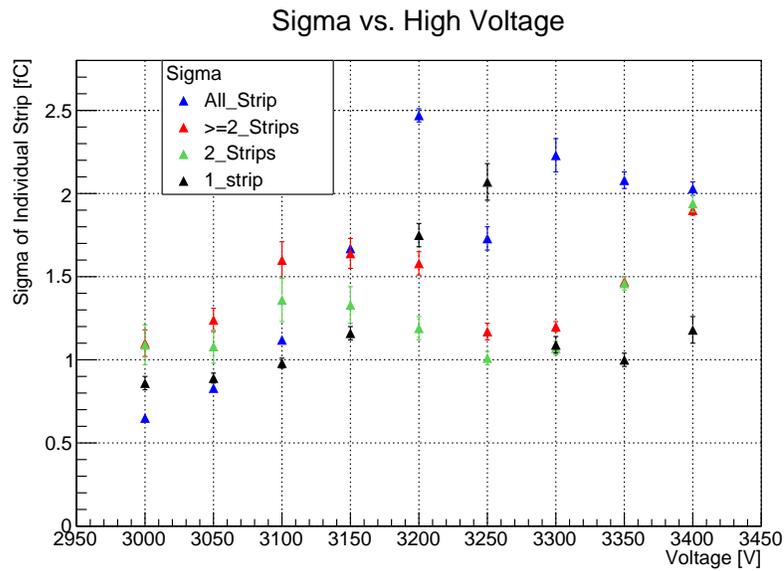
Sigma vs. High Voltage



(b)



(c)



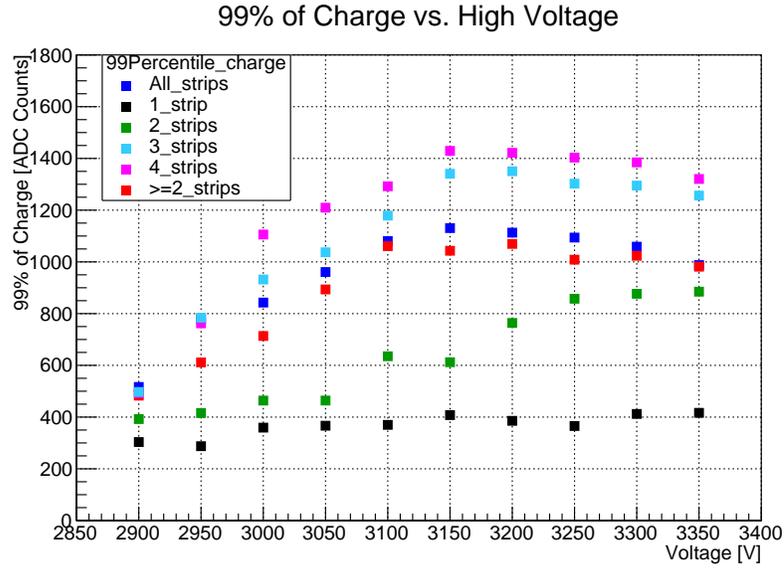
(d)

Figure 3.37: GEM detector with radial zigzag strip readout: Sigma of the total cluster charge distribution (a) in ADC (b) in the fC unit. Sigma of the individual strip distribution (c) in ADC (d) in the fC unit from Landau fit.

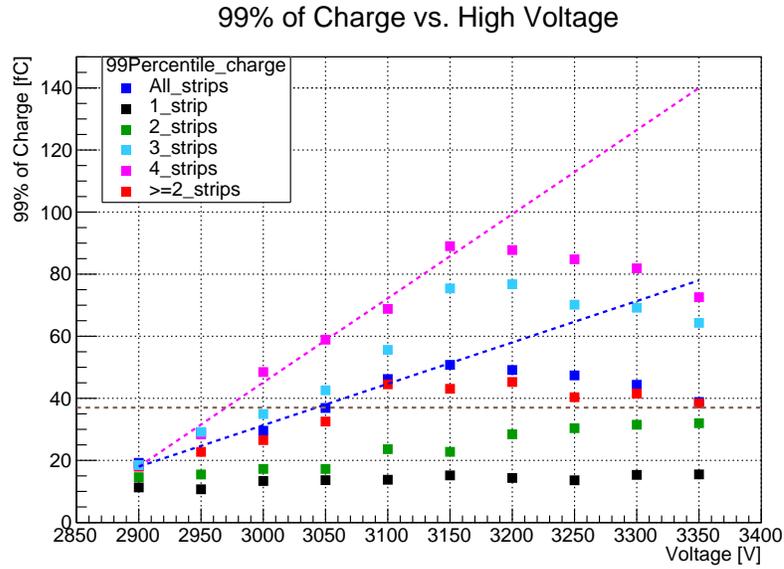
99th Percentile of the Maximum Charge of the Individual Strip Distribution

Finally, the 99th percentile of the strip charge distribution is calculated to find the maximum charge that can be expected at the input of a VFAT3 chip. For each voltage, the strip charge distribution is plotted with different cuts on the strip multiplicity of the clusters. Due to the limited dynamic range of the APV readout system, saturation of signal pulses is again observed causing the curves to flatten out at high drift voltages. Figure 3.38 (b), shows that the 99th percentile charge at the operating voltage is about 43 fC. The maximum range for strip charges is estimated by extrapolating a linear fit to compensate for the APV saturation at large operating voltages. With this approach, a maximum strip charge of 78 fC is estimated at 3350 V. To cover the extreme case, the fit to the 4-strips cluster data is extrapolated, which gives a maximum strip charge of 140 fC at 3350 V. Hence, the GE1/1 readout electronics should be prepared to have a dynamic range up to 140 fC for the input charge on one channel.

The 99th percentile of the maximum charge for the GEM detector with the zigzag strip readout is calculated in a similar way. The only difference in this case is that there are no 3-strip and 4-strip clusters due to the saturation effect. Figure 3.39 shows the 99th percentile of the maximum charge in ADC as well as in the fC units. The 58fC charge is estimated as 99% of the maximum charge for all strips at operating voltage and 75fC for the 2-strips cluster. These curves appear to be close to saturation for basically all drift voltages and are consequently only of limited value.



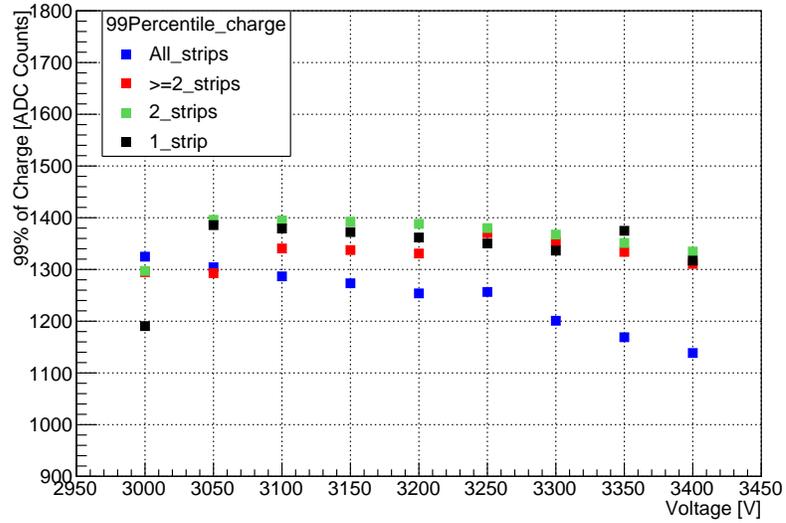
(a)



(b)

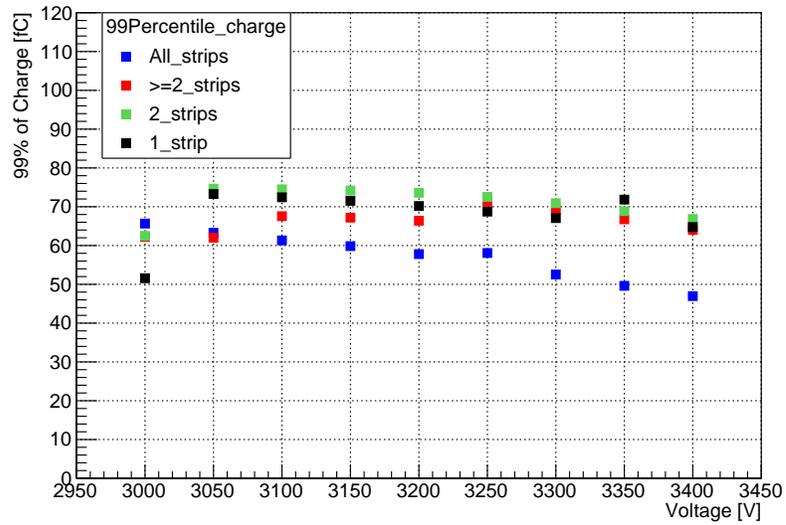
Figure 3.38: The GE1/1 Detector: 99th percentile of strip charge distribution (a) in ADC (b) in fC unit with linear extrapolation of data that suffers little APV saturation.

99% of Charge vs. High Voltage



(a)

99% of Charge vs. High Voltage



(b)

Figure 3.39: The GEM detector with radial zigzag strip readout: 99th percentile of strip charge distribution (a) in ADC (b) in the fC unit.

Efficiency Measurements

Detection efficiency is obtained as

$$\epsilon = \frac{N_1}{N - N_2} \quad (3.5)$$

where, N = Total number of triggered events; N_1 = Number of events with Cluster Multiplicity (CM) ≥ 1 for given sector; N_2 = Number of events with Cluster Multiplicity (CM) ≥ 1 for other sectors. A few strips from neighboring sectors are sometimes triggered due to scattering of particles or due to beam focus on the edge of the given sector. To obtain accurate efficiency of the given sector, it is required to eliminate such events from the total number. The overall detection efficiency of the detector is measured using four threshold cuts on a pedestal width, namely 3σ , 4σ , 5σ , and 6σ , where σ is the mean of RMS value of the noise distribution for all strips. Figure 3.40 shows efficiency curves fitted with a sigmoid function for these threshold cuts. The efficiency curves show a long plateau for higher voltages and are not affected by different threshold cuts. Detection efficiency of the GE1/1 detector with a 5σ cut on the pedestal width is $[97.80 \pm 0.2 \text{ (stat)}]\%$.

The position scan is used to measure the uniformity response of the detector for all η -sectors at an operating drift voltage 3250 V. The scan was performed for three APV positions, as mentioned earlier, namely Upper, Middle, and Lower APV position. For each sector, the charge uniformity was measured using the MPV of the Landau fit from the charge distribution. Figure 3.41 shows the uniformity response for the first seven sectors. The response varies by $\sim 25\%$ in different sectors. The variation in outer sectors 6 and 7 can be improved in

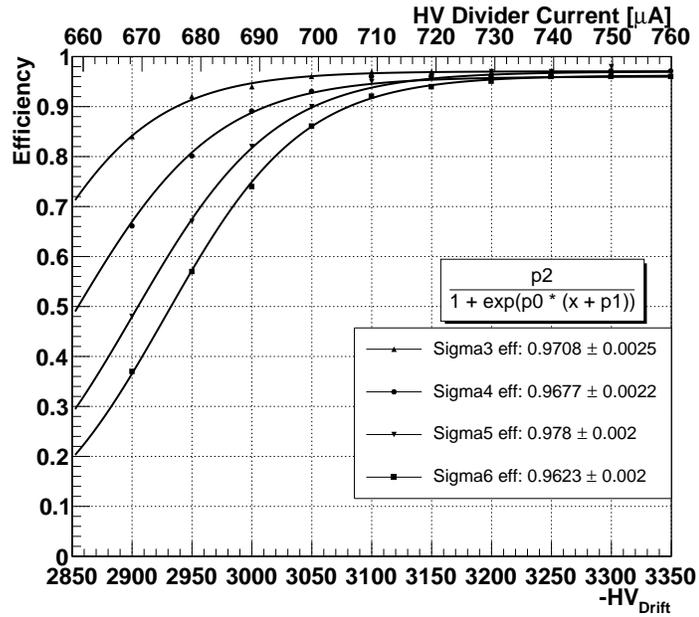


Figure 3.40: Detection efficiency with different cuts on pedestal width.

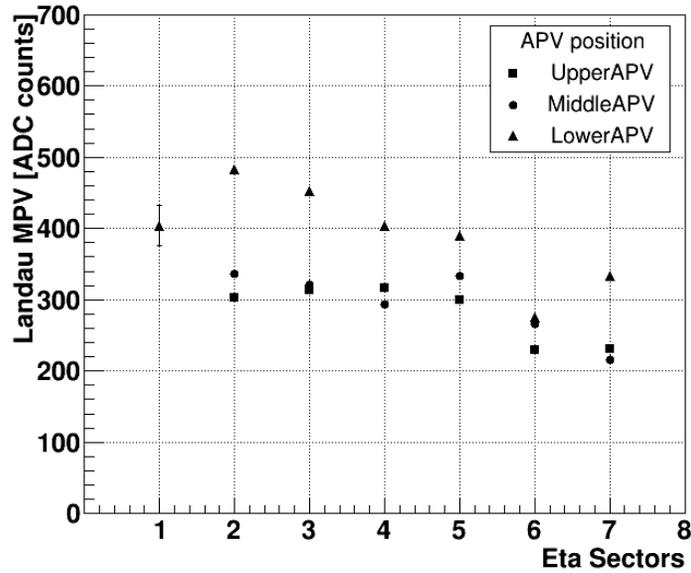


Figure 3.41: Response uniformity in seven η -sectors for three different APV positions, i.e., Upper, Middle, and Lower at 3250 V.

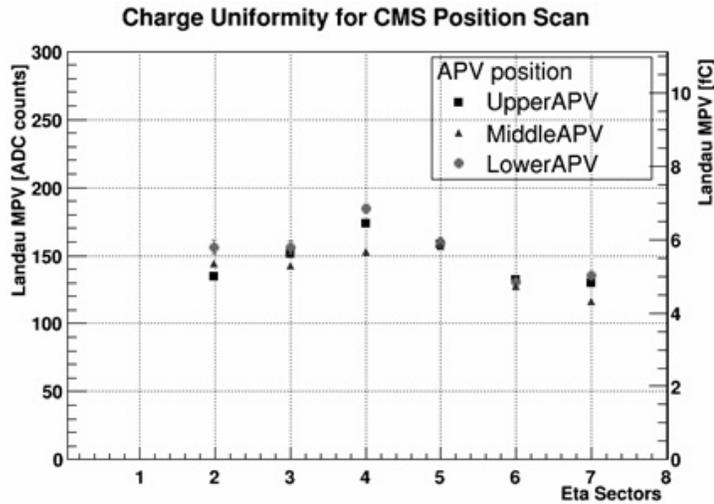


Figure 3.42: Charge uniformity for different η -sectors at 3250 V using single-strip, single-cluster events.

future assemblies by paying particular attention to stretching foils in that region. The response uniformity of the detector is again calculated by considering only single-strip single-cluster events. Figure 3.42 shows the uniformity response of the η -sectors of the detector.

Tracking Analysis

Overall performance characteristics of the GE1/1-III prototype detector were studied using a 32GeV mixed-hadron beam. Beam profiles can be obtained using 2D hit maps of the reference tracker detectors. The beam profile for the secondary 32 GeV mixed-hadron beam is oval, while the beam profile for the 120 GeV proton beam is circular and much tighter, as shown in Figure 3.43.

The tracking analysis is done in three steps to study the spatial resolution of the GE1/1 detector. The first step is alignment, which itself is executed in two steps. The first one is to shift all tracker detectors so that their origins in the x-y coordinate coincide and their residuals are centered at zero. In the second

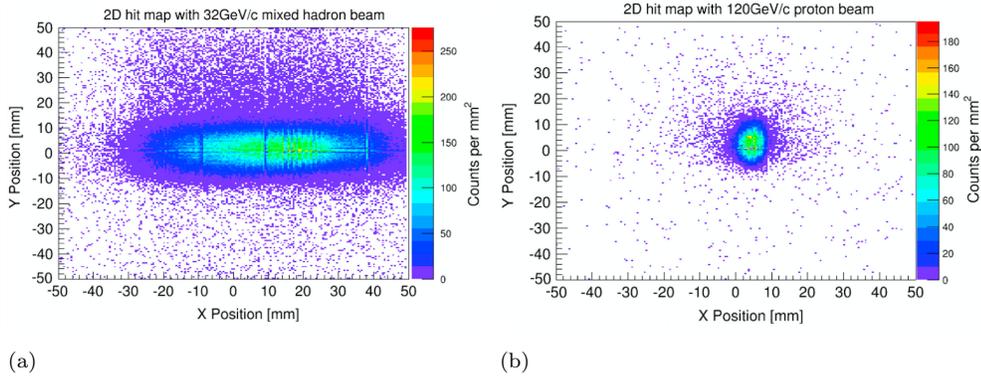


Figure 3.43: Beam profiles: (a) 32 GeV hadron beam, (b) 120 GeV proton beam.

step of alignment, the three downstream tracking detectors are rotated with respect to the first tracker detector by using initial shift parameters from step one. The rotating angle for each detector is optimized such that the residual width of each detector is minimized.

The next important step in the tracking analysis is a transformation from Cartesian (x,y) co-ordinates to polar (r, ϕ) co-ordinates, as the GE1/1 detector measures the azimuthal ϕ -coordinate with radial strips. Figure 3.44 shows correlated event hits in η -sector 5 of the GE1/1 detector and in the first tracker detector.

In the final step of the tracking analysis, both inclusive and exclusive track-hit residuals are calculated. The definition of an inclusive (exclusive) residual is the residual calculated by including (excluding) the probed detector in the track fitting. The spatial resolution of the detector is calculated from the geometric mean [28] of the widths of the inclusive and the exclusive residuals, i.e. $\sigma = \sqrt{\sigma_{inc} \times \sigma_{exc}}$ [29]. The geometric mean method has one drawback that it does not consider effect of Coulomb multiple scattering and results into esti-

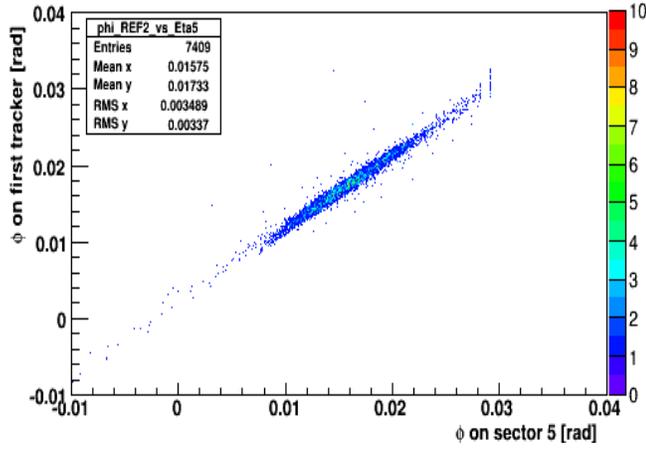


Figure 3.44: Correlation of the GE1/1 detector hits with hits in the first tracker detector.

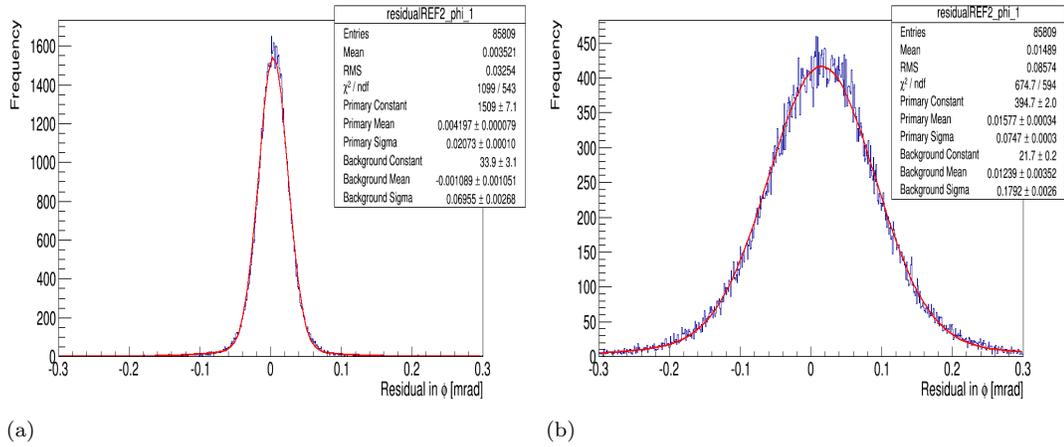


Figure 3.45: Residuals for tracker 1 in φ : (a) Inclusive residual with width $\sigma = 21 \mu\text{rad}$ and (b) Exclusive residual with width $\sigma = 75 \mu\text{rad}$.

estimating the resolution too optimistically by 10-50%. The accuracy in this measurement can be gained by considering the presence of the multiple Coulomb scattering [30]. Figure 3.45 shows both residual widths for the first tracker detector. The inclusive residual is $\sigma_{inc} = 21 \mu\text{rad}$ and the exclusive residual is $\sigma_{exc} = 75 \mu\text{rad}$, with a geometric mean of $\sim 40 \mu\text{rad}$. Similarly, the resolution

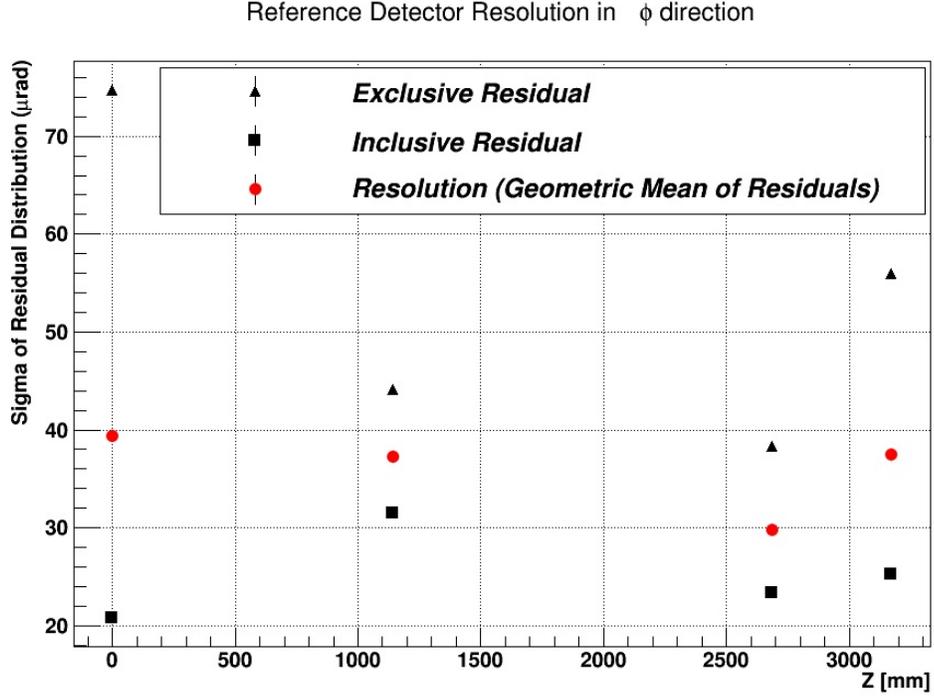


Figure 3.46: Angular resolution of the tracking detectors.

of the other two 10 cm×10 cm tracker detectors are calculated. Figure 3.46 summarizes the resolutions for all tracker detectors.

The resolution of the GE1/1-III prototype GEM detector is calculated using two different methods for obtaining the hit position. The barycentric method [31] uses the full pulse height information to find the strip cluster barycenter which is calculated as $s_b = \sum_{i=1}^n \frac{q_i \cdot s_i}{q_i}$, where s_i and q_i are the strip number and charge of the i^{th} strip, respectively. The binary method uses the reconstructed binary hits to emulate the VFAT [32] results. The results are presented in the following sections.

The Barycentric Method

Figures 3.47 and 3.48 show the inclusive and the exclusive residuals of the GE1/1 detector at the center of η -sector 5 at 3250 V using a 5σ cut on the pedestal width. The inclusive residual is $110.7 \mu\text{rad}$ and the exclusive residual is $137.9 \mu\text{rad}$ in the azimuthal (ϕ) direction; this corresponds to $208.1 \mu\text{m}$ and $259.3 \mu\text{m}$ respectively at $R = 1880.5 \text{ mm}$. The resolution of the GE1/1 detector using pulse height analog readout is $[123.3 \pm 1.6 \text{ (stat)}] \mu\text{rad}$, i.e. 27% of the strip pitch. Residuals for η -sector 5 plotted against the different voltages are shown in Figure 3.49. The best resolution is obtained on the efficiency plateau as expected.

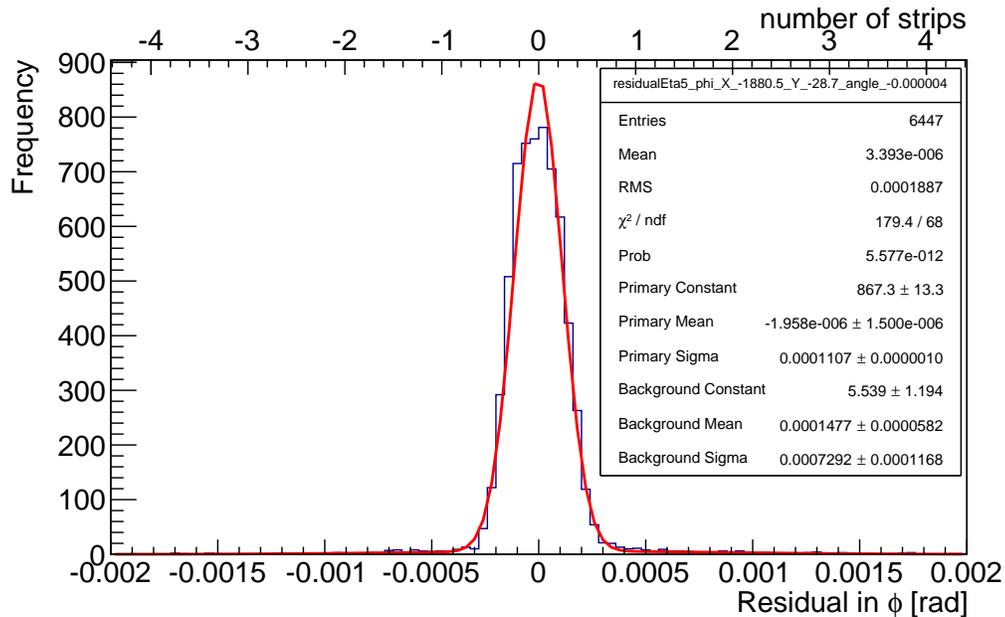


Figure 3.47: Inclusive GE1/1 residual with $\sigma = 110.7 \mu\text{rad}$ using full pulse height information.

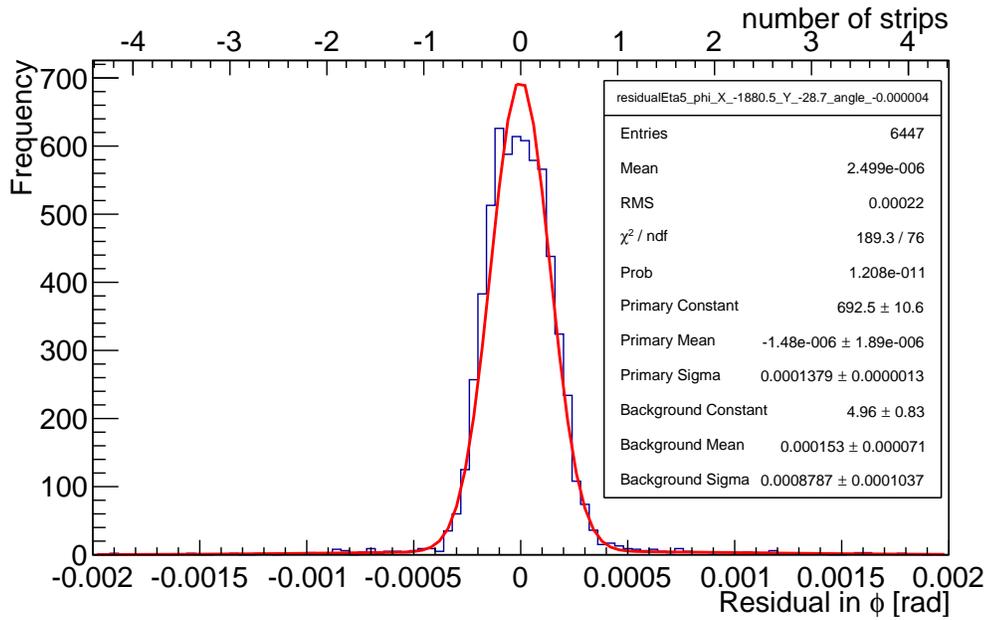


Figure 3.48: Exclusive GE1/1 residual with $\sigma = 137.9 \mu\text{rad}$ using full pulse height information.

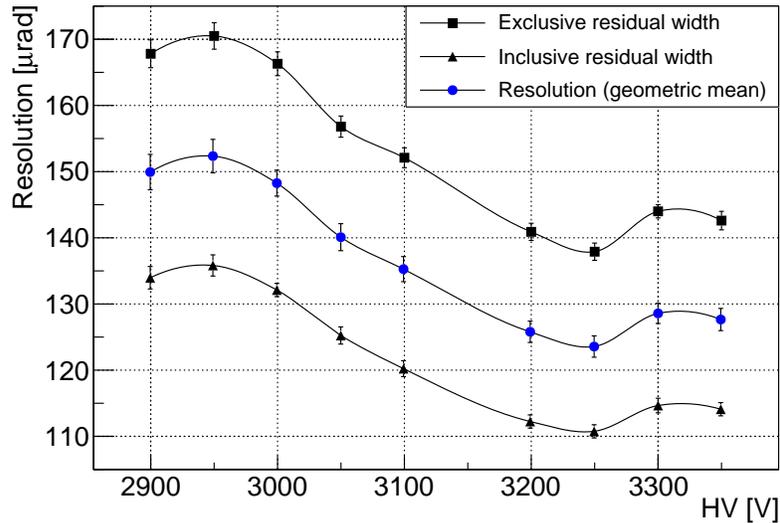


Figure 3.49: Resolutions of the GE1/1 detector at the center of sector 5 for different voltages using full pulse height information.

The Binary Method (Emulated VFAT threshold)

During this test beam, all data was collected using an APV25 [33] hybrid chip that provides the analog signal, whereas CMS upgrade electronics will use the VFAT3 chip [1] that produces a binary output for each readout strip (charge above or below threshold). Hence, it is important to study the characteristics of the GE1/1 detector using the binary hit method. By applying a fixed threshold cut on the pedestal width, binary hits are reconstructed offline from the pulse height data. The detector efficiency is calculated again by applying fixed cuts of 0.8fC, 0.98fC, and 1.2fC, which are equivalent to “10VFAT”, “12VFAT”, and “15VFAT” units, where “1VFAT” unit = 0.08fC. The efficiency curve is again fitted with the sigmoid function and from one of the fit parameters the efficiency of the detector is calculated which is $[96.9 \pm 0.2 \text{ (stat)}]\%$ on plateau.

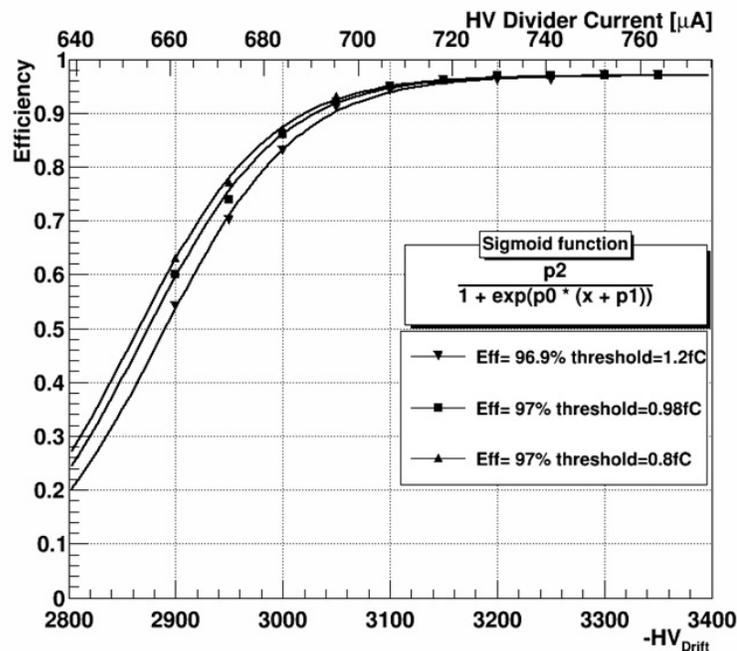
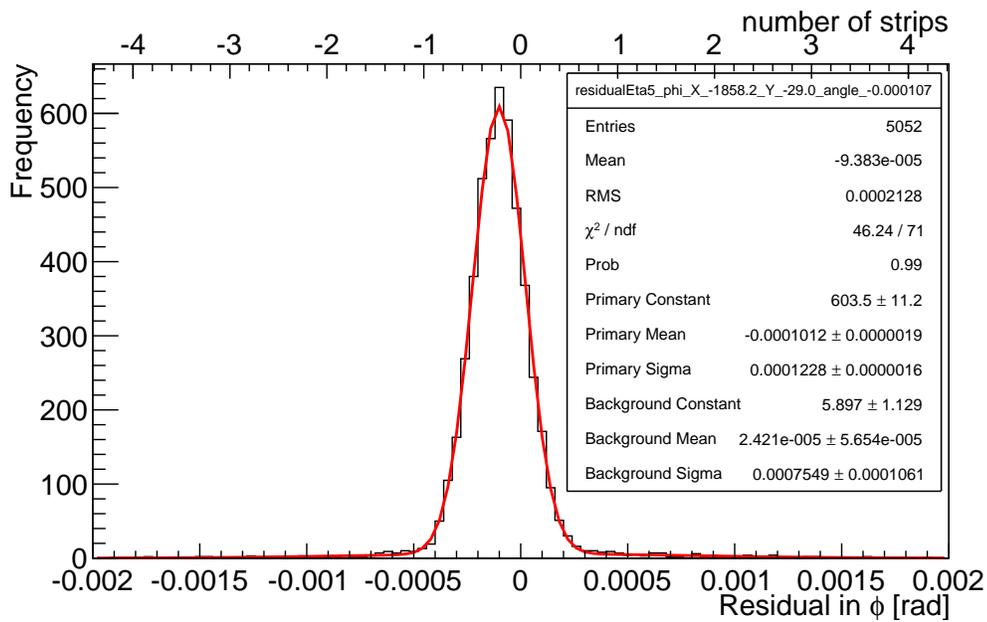
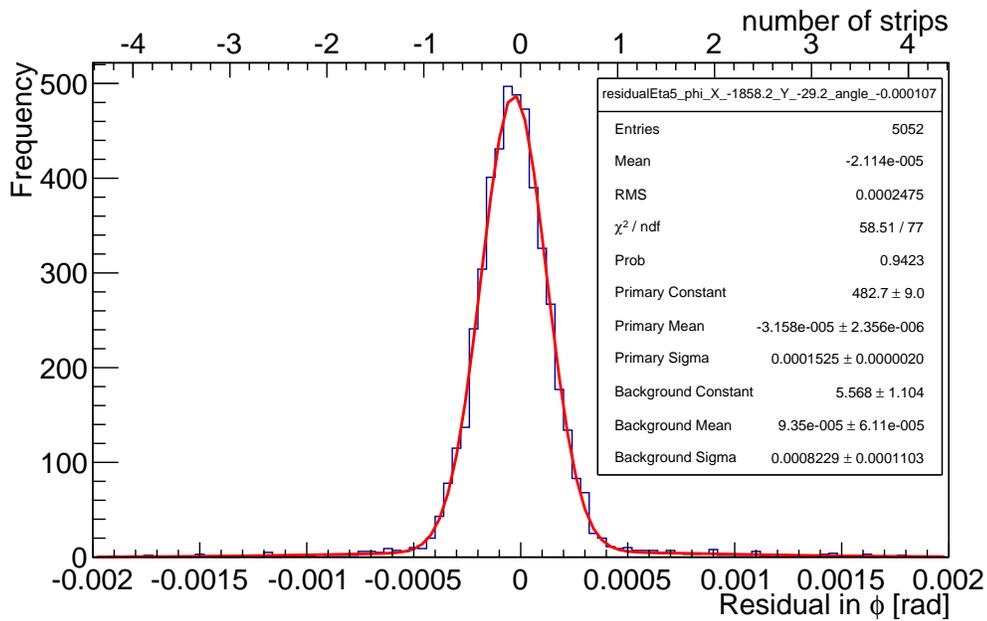


Figure 3.50: Detection efficiency using VFAT-like binary hit data with three different thresholds.



(a)



(b)

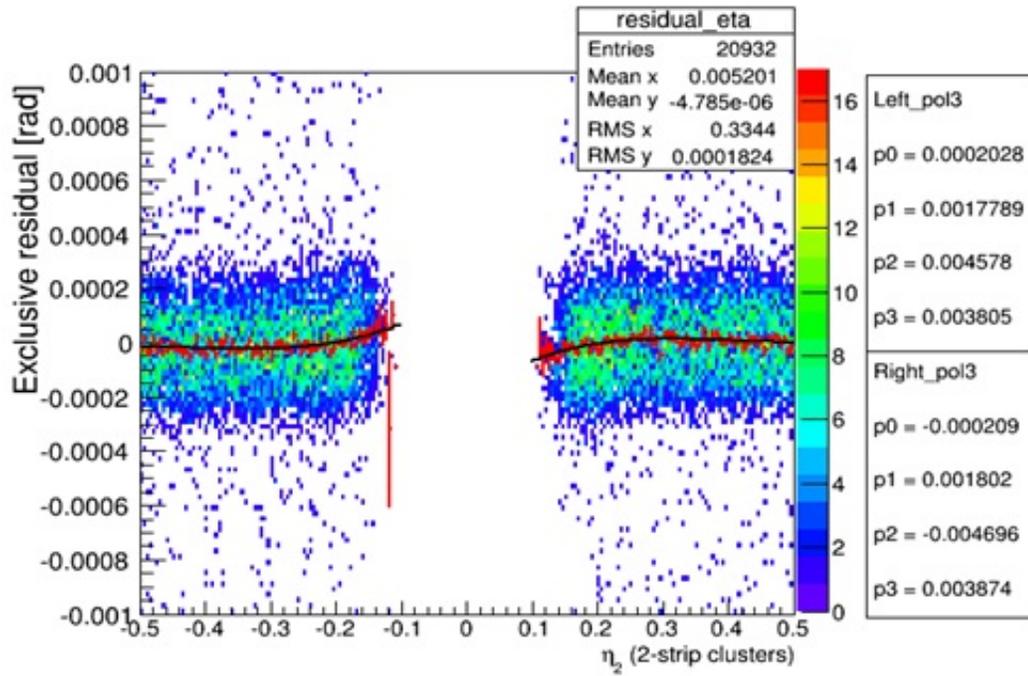
Figure 3.51: (a) Inclusive GE1/1 residual $\sigma = 122.8 \mu\text{rad}$ and (b) Exclusive GE1/1 residual $\sigma = 152.5 \mu\text{rad}$ using binary hit reconstruction.

Finally the inclusive and exclusive residuals are calculated. Figure 3.51 (a) shows that the inclusive residual width of the detector is $122.8 \mu\text{rad}$ ($228.1 \mu\text{m}$) and (b) shows that the exclusive residual width is $152.5 \mu\text{rad}$ ($283.3 \mu\text{m}$) with binary hits. The geometric mean of these widths is $136.8 \mu\text{rad}$ ($254.2 \mu\text{m}$), which is consistent with the expectation from $\text{pitch}/\sqrt{12} = 131.3 \mu\text{rad}$. Total radiation length in the test beam setup is about $\sim 14\%$ with $\sim 147 \mu\text{rad}$ RMS of the multiple scattering angle. This value is not negligible as it is within the range of the expected residual widths of the detector.

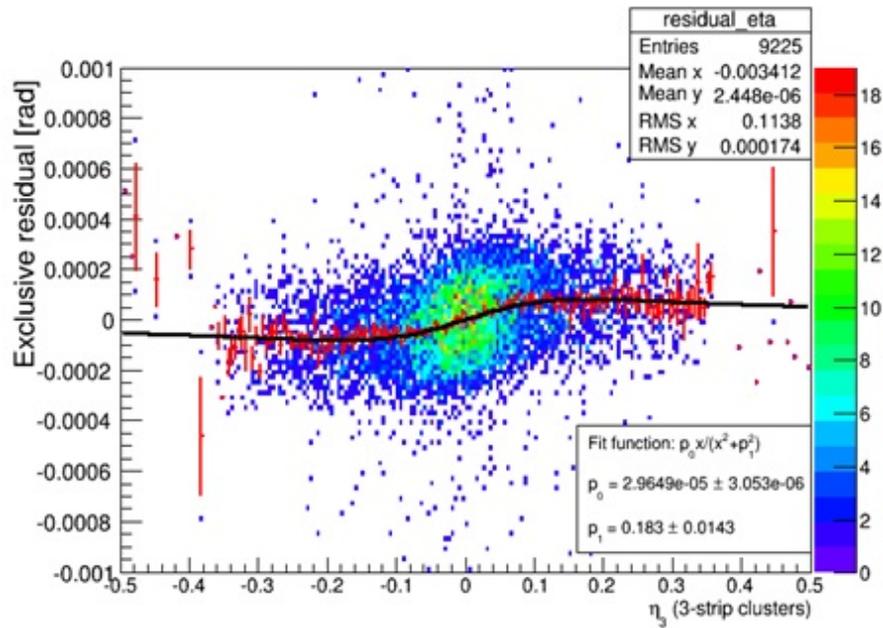
3.3.4 Correction of Non-linear Strip Response

The motivation for implementing this correction for the barycentric method is to further improve the spatial resolution of the GE1/1 detector. In this study, cluster positions are reconstructed in the detector based on the barycenter method. The η -correction factor was developed to correct the barycenter position in the GE1/1 detector. For a cluster with strip multiplicity greater than 1, η is defined as $\eta = s_b - s_{max}$ [34], where $s_b = \sum_{i=1}^n \frac{q_i \cdot s_i}{q_i} t_i$ gives the barycenter position in terms of strip numbers where s_i and q_i are the strip number and charge of the i^{th} strip respectively; s_{max} is the strip number with the maximum charge. This correction is done mostly for 2, 3, and 4-strip clusters. Figure 3.52 (a) and (b) show the scatter plots of the exclusive residual against η for 2-strips and 3-strips clusters, respectively. Both showed different behaviors and hence are fitted with different functions. Finally, the corrected resolution is obtained for the GE1/1 detector after subtracting the value of the fitted function at η from the original resolution. Figure 3.52 (c) shows the scatter plot for 2- 3-strip clusters before correcting the resolution and (d) shows the corrected resolution

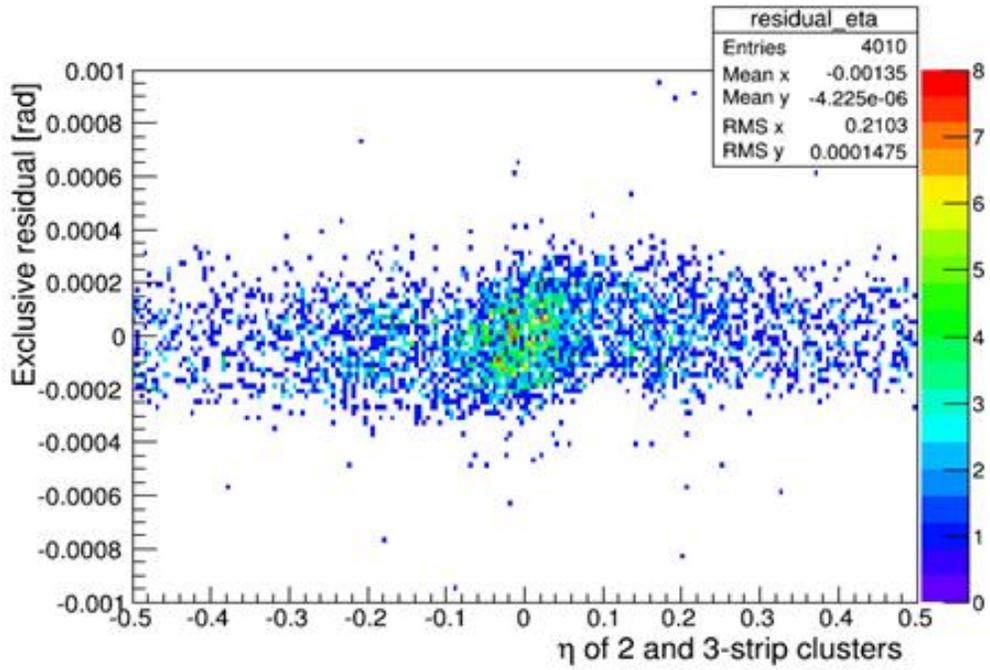
for these strip clusters.



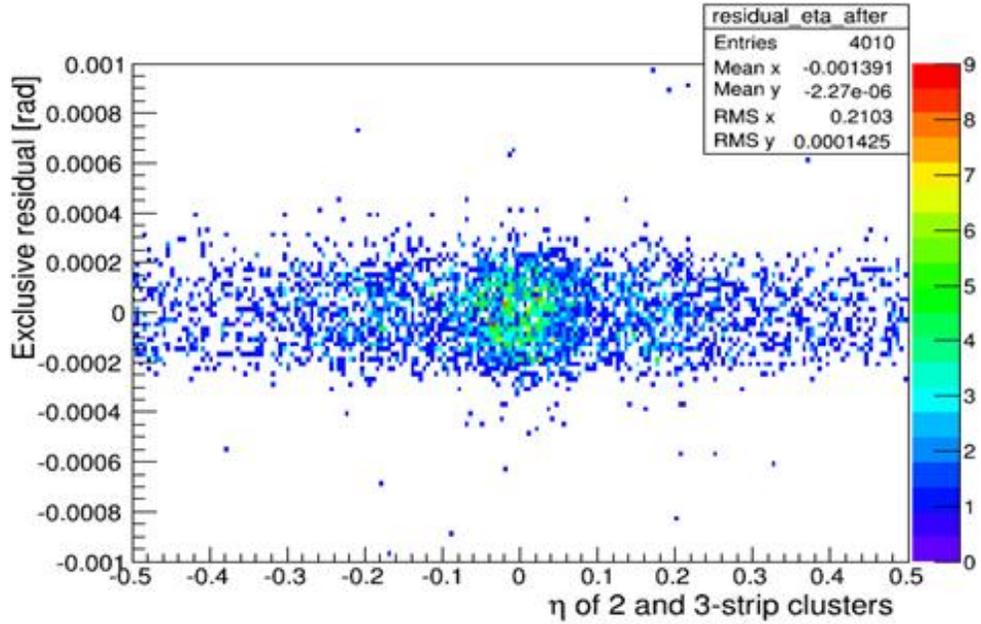
(a)



(b)



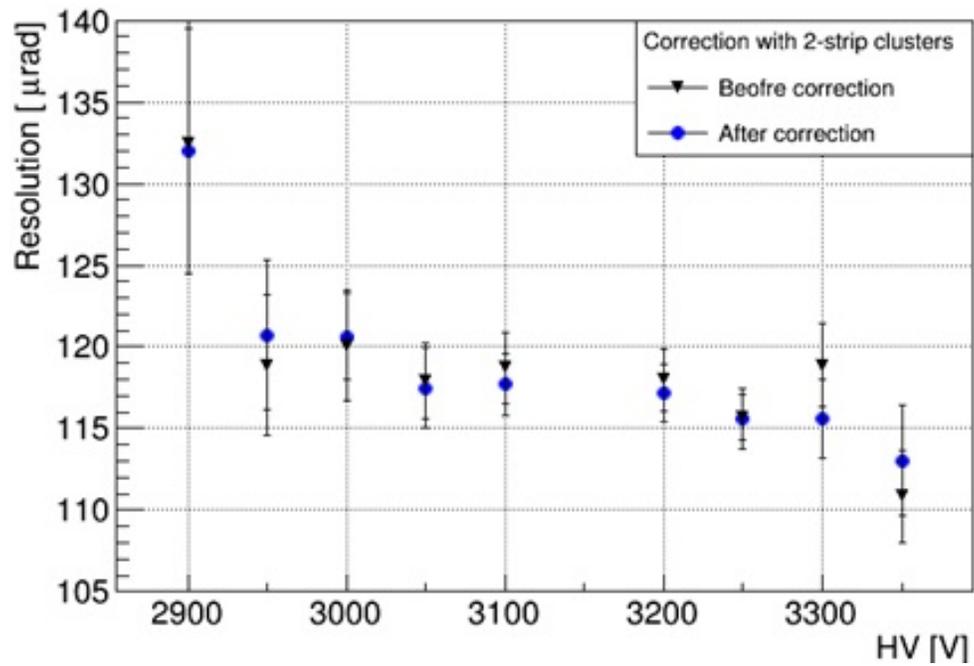
(c)



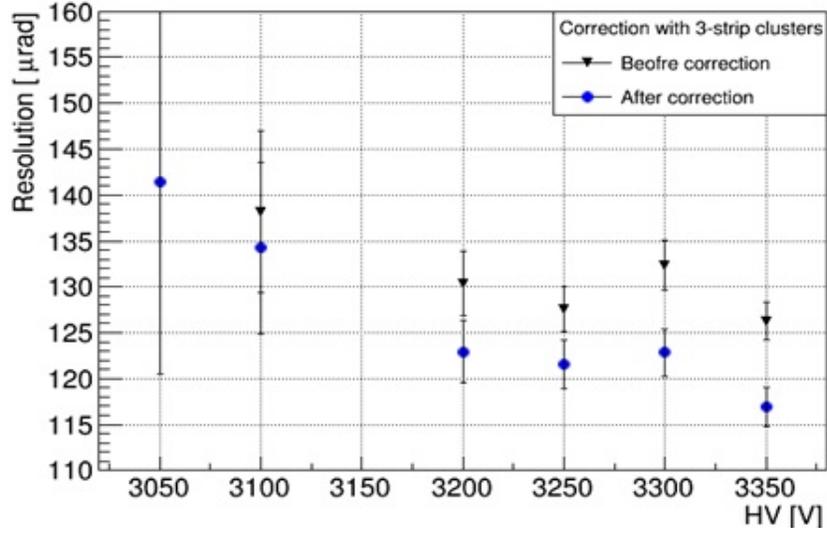
(d)

Figure 3.52: (a) and (b) Scatter plots of exclusive residual vs. η for 2-strip and 3-strip clusters respectively for combined HV scan data. (c) and (d) Scatter plots of exclusive residual vs. η for all strip cluster multiplicities before and after correction, respectively.

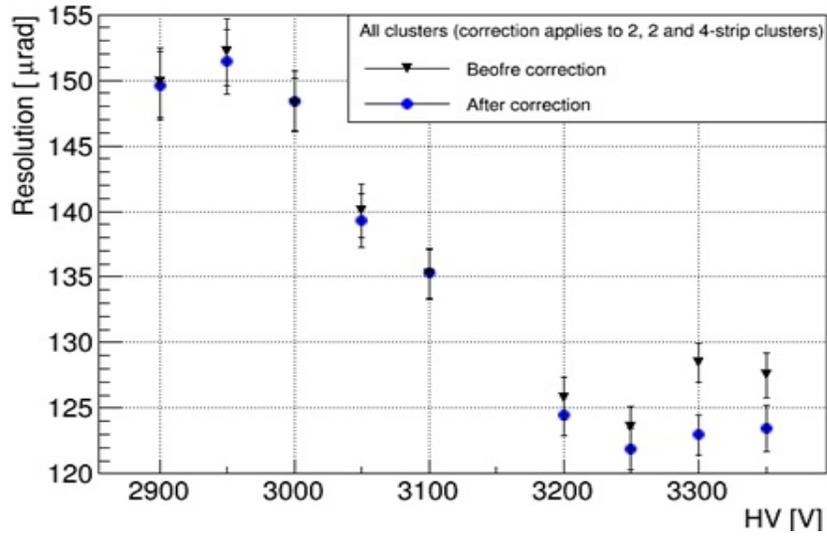
For a high voltage scan, the strip correction factor changes negligibly for the entire voltage range and hence, all HV scan data are combined before implementing the strip correction for more statistics. However, strip correction is performed individually on each sector for the position scan. Figure 3.53 shows the corrected resolutions for the high voltage scan. Since the readout of the GE1/1 detector has fine strip segmentation, the overall improvement factor is small (less than 8%, i.e., within $\sim 10 \mu\text{rad}$) compared to detectors having coarse readout strips.



(a)



(b)



(c)

Figure 3.53: (a) Resolution after correction for 2-strip clusters, (b) Resolution after correction for 3-strip clusters, (c) Resolution after correction for 2-strip, 3-strip, and 4-strip clusters for the high voltage scan.

3.3.5 Summary and Conclusion

A CMS GE1/1-III prototype GEM detector was successfully built and tested at Florida Tech and Fermilab, respectively, in 2013. The charge distribution is studied to estimate the expected input charge range for VFAT3 readout electronics for the GE1/1 detector of the muon endcap upgrade of the CMS experiment. The linear and non-linear conversion factors are used for converting charges from ADC counts to an fC unit for the analog readout system (APV25) used in the study. The expected mean charge in the GE1/1 detector is 37 fC at 3250 V drift voltage, which is calculated from the number of total electron-ion pairs in the drift region and from the gas gain. A direct measurement of this charge from the mean total cluster charge distribution using a Landau fit gives 25 fC, which is within 27% of the calculated charge value. When operating 50V above the start of the efficiency plateau in an Ar/CO₂ 70:30 gas mixture, i.e. with 3250V applied to the drift electrode, the most probable value, mean value, and 99th percentile value of the Landau distribution of the charge induced on a single strip are found to be 4 fC, 11 fC, and 115 fC, respectively. Measurements for the highest charges are somewhat hampered by the saturation of the APV chip. The largest input charge range for the VFAT readout electronics is estimated to be 140 fC based on an extrapolation of the 99th percentile of the individual measured strip charge distribution in 4-strip clusters. Therefore, the VFAT3 front-end chip should be designed in such a way that they can comfortably handle input charges over a range of 0-140 fC. Studies done for the same GEM detector equipped with a radial zigzag strip readout give similar results.

The detector performed well in terms of detector efficiency, which is greater than 97% with both methods for determining hit positions. It shows good

charge uniformity for all η -sectors except for sectors 6 and 7. The uniformity should be improved in future assemblies by making sure that all the foils have the same tension in all eight η -sectors. The spatial resolution of the GE1/1 detector is $\sim 123 \mu\text{rads}$ for the barycentric method and $\sim 136 \mu\text{rads}$ with the binary method, which meets the value of resolution expected from the pitch of the strip. The spatial resolution of the detector is improved by $\sim 10 \mu\text{rad}$ after correcting for a non-linear strip response. In conclusion, the GE1/1 detector meets performance expectations for use in a CMS muon endcap upgrade.

Chapter 4

The Standard Model and Higgs Boson

4.1 The Standard Model of Particles

In particle physics, the interaction between particles is described by the standard model (SM). According to the SM, all matter is composed of three kinds of elementary particles: leptons, quarks, and mediators [35]. Leptons and quarks are fermions with spin $1/2$, whereas the mediators are bosons with an integer spin. Overall, there are six leptons and six quarks that are distinguished by their charge, mass, and in the case of leptons by lepton number; they are then categorized into three generations. There are six leptons: electrons (e), muon (μ), tau (τ), and their corresponding neutrinos. Their classification is given in Table 4.1.

For each of these leptons, there exists an anti-lepton with an opposite charge. Similar to leptons there are six flavors of quarks, which are also classified into

Table 4.1: Leptons classified in three generations.

Generation	Leptons	Charge (Q)	Mass (MeV)	Lifetimes (s)
First	electron (e)	-1	0.51099	∞
	electron neutrino (ν_e)	0	$\simeq 0$	∞
Second	muon (μ)	-1	105.67	2.20×10^{-6}
	muon neutrino ν_μ	0	$\simeq 0$	∞
third	tau (τ)	-1	1776.99	2.91×10^{-13}
	tau neutrino ν_τ	0	$\simeq 0$	∞

three generations as shown in Table 4.2. Again, for each quark, there is an antiquark with an opposite charge. Each quark comes in three colors; therefore, in total there 36 quarks and anti-quarks.

Table 4.2: Three generations of quarks with their corresponding charge and mass.

Generations	Quarks	Charge (Q)	Mass
First	up (u)	2/3	$2.3^{+0.7}_{-0.5}$ MeV
	down (d)	-1/3	$4.8^{+0.5}_{-0.3}$ MeV
Second	charm (c)	2/3	1.275 ± 0.025 GeV
	strange (s)	-1/3	95 ± 5 MeV
Third	top (t)	2/3	$173.21 \pm 0.51 \pm 0.71$ GeV
	bottom (b)	-1/3	4.66 ± 0.03 GeV

Table 4.3: Fundamental forces with interacting mediators.

Interaction	Range	Relative Strength	Mediator
Strong	10^{-15} m	1	Gluons (g)
Electromagnetic	∞	10^{-3}	Photon (γ)
Weak	10^{-18} m	10^{-14}	W^\pm, Z
Gravitational	∞	10^{-43}	Graviton (?)

Finally, the mediator spin-1 particle explains the interaction between the elementary particles. These fundamental forces are given in Table 4.3. The electromagnetic interaction takes place via photons (γ), whereas the W^\pm and Z bosons are the mediators of the weak interaction. The strong interaction is mediated by eight colored gluons.

4.2 The SM Scalar Sector

The gauge symmetry of $SU(3)_C \times SU(2)_L \times U(1)_Y$ can perfectly describe the standard model, where C refers to color, L is the left chiral nature of $SU(2)$ couplings, and Y is the weak hypercharge. In this case, $SU(3)_C$ describes the strong interaction and $SU(2)_L \times U(1)_Y$ explains the electroweak interactions. The total Lagrangian for the SM can be expressed in terms of gauge, fermion, Yukawa, and scalar Lagrangians:

$$\mathcal{L}_{SM} = \mathcal{L}_{gauge} + \mathcal{L}_f + \mathcal{L}_{Yuk} + \mathcal{L}_\phi. \quad (4.1)$$

The gauge Lagrangian term regroups the gauge field for all three symmetries. The fermion Lagrangian describes the nature of the kinetic energy associated with leptons and quarks, whereas the Yukawa Lagrangian describes the interactions between the fermions and the scalar doublet $\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$, which are responsible for the fermion masses. The Yukawa coupling introduces the Cabbibo-Kobayashi-Masakawa (CKM) mixing matrix that leads to creating a large number of free parameters in the SM. The scalar Lagrangian focuses on the kinematic and potential component of the scalar field, as well as explain-

ing the spontaneous electroweak symmetry breaking, which gives masses to the W and Z bosons. The scalar Lagrangian is written as:

$$\mathcal{L}_\phi = (D^\mu \phi)^\dagger D_\mu \phi - V(\phi). \quad (4.2)$$

The potential term can be further expressed as:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2. \quad (4.3)$$

For electroweak symmetry breaking to occur, μ^2 has to be negative and real, and λ should have a positive value as well. The covariant derivative D_μ is defined as

$$D_\mu = \partial_\mu + ig\vec{T} \cdot \vec{W}_\mu + \frac{ig'}{2}B_\mu. \quad (4.4)$$

where g and g' are coupling strength of electromagnetic and weak interaction, respectively. This choice of parameters gives a potential with a shape known as “Mexican hat,” where the local maximum is at zero and the minimum has a non-zero value. Details on the spontaneous symmetry breaking via the Higgs mechanism can be found in [36]. The Brout-Englert-Higgs field couples universally to all quarks and leptons with a strength proportional to their masses and to gauge bosons with a strength proportional to the square of their masses [37].

4.2.1 Higgs Boson Production Mode at the LHC

At the LHC, Higgs boson production takes place via the following four processes and they are presented in the Feynman diagrams shown in Figure 4.1 [38].

- **Gluon-Gluon Fusion (ggH)** is the main Higgs production process at

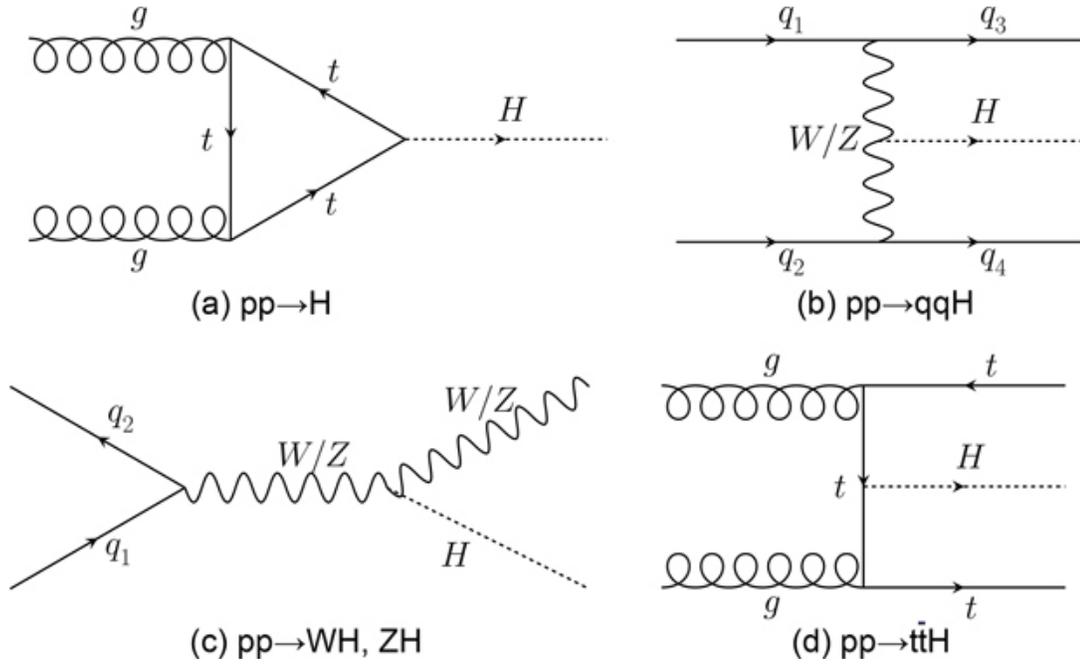


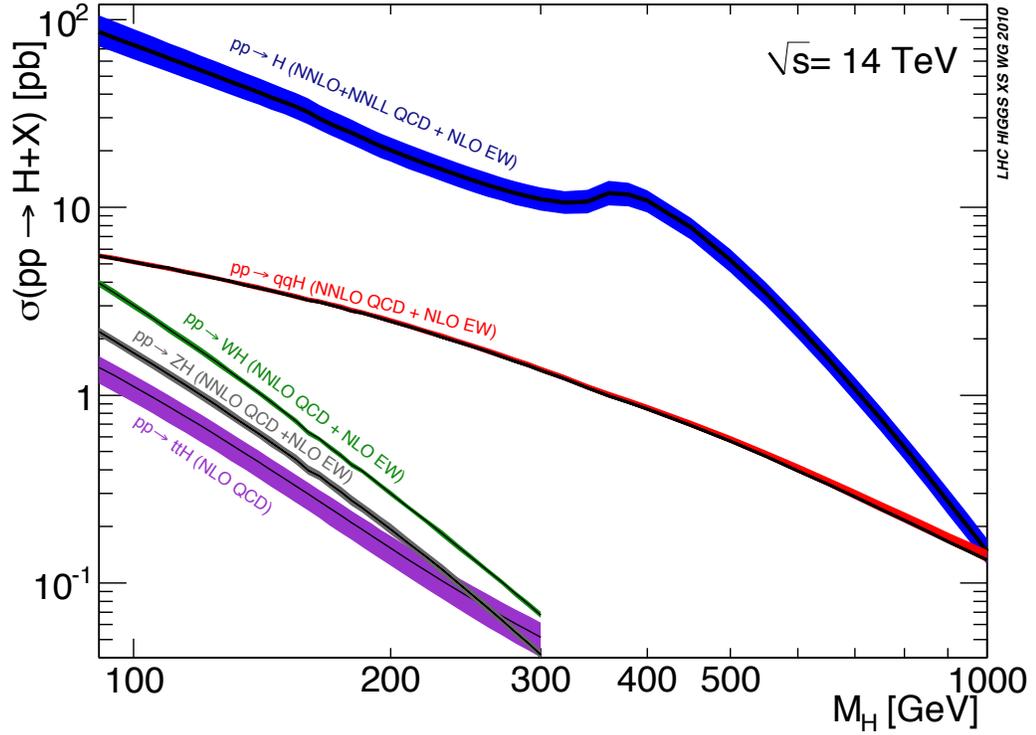
Figure 4.1: Feynman diagram for the SM Higgs production mode: (a) Gluon-Gluon Fusion (ggH), (b) Vector Boson Fusion (VBF), (c) Vector Boson Association Production (VH), and (d) in Association with a pair of Top Quarks.

the LHC because it has the largest cross section. In this process, two incoming gluons fuse to produce a Higgs boson through a triangular top quark loop.

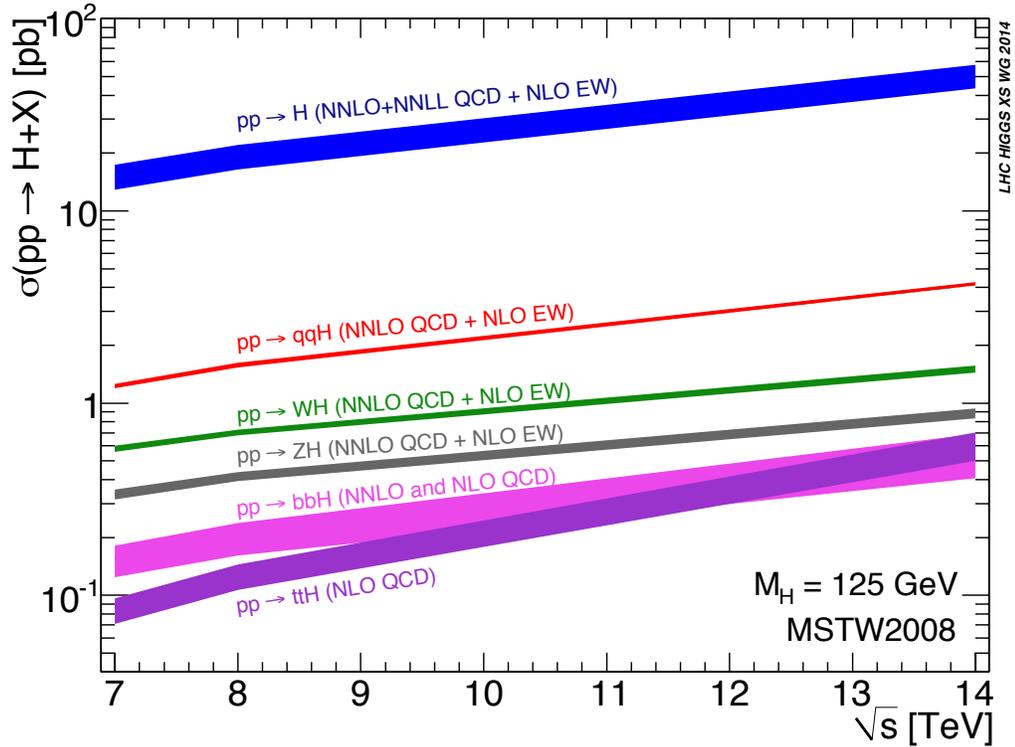
- **Vector Boson Fusion (VBF)** has the second largest production cross section at the LHC. In this process, the Higgs boson is produced via a vector boson (W or Z) interaction. These vector bosons are emitted by two interacting quarks, which further hadronize and produce high energy jets in the forward direction. The forward jets in the final state characterize this production process.
- **Vector Boson Associated Production (VH)** is also known as “Higgsstrahlung.” The Higgs boson is produced via the interaction of the vir-

tual boson V^* . The cross section of this production is very small compared to ggH and VBH production modes. The discrimination of the scalar boson H from the background is easily possible in the VH process due to the presence of additional leptons and quarks coming from the vector boson decay.

- **Production in Association with a Pair of Top Quarks ($t\bar{t}H$)** has the smallest production cross section. In Run I, there was no evidence of this production, but Run II has a higher luminosity and increased center-of-mass energy which is expected to benefit this mode. In October 2017, the ATLAS collaboration announced evidence for this production [39].



(a) Scalar boson production cross section as a function of m_H .



(b) Scalar boson production cross section as a function of center-of-mass energy.

Figure 4.2: The Higgs production cross section at the LHC.

The cross sections of the different Higgs boson production modes depend on the mass of the Higgs boson and the center-of-mass energy. Figure 4.2 [40] shows the expected production modes with respect to the different mass ranges and the center-of-mass energies.

The Higgs boson immediately decays into bosons or fermions. Figure 4.3 shows various Higgs decays with respective branching fractions. Bosonic Higgs decay channels are $\gamma\gamma$, ZZ , and W^+W^- , whereas fermionic decay channels are $b\bar{b}$, $\tau^+\tau^-$, $\mu^+\mu^-$, and so forth. The branching fractions can be computed using

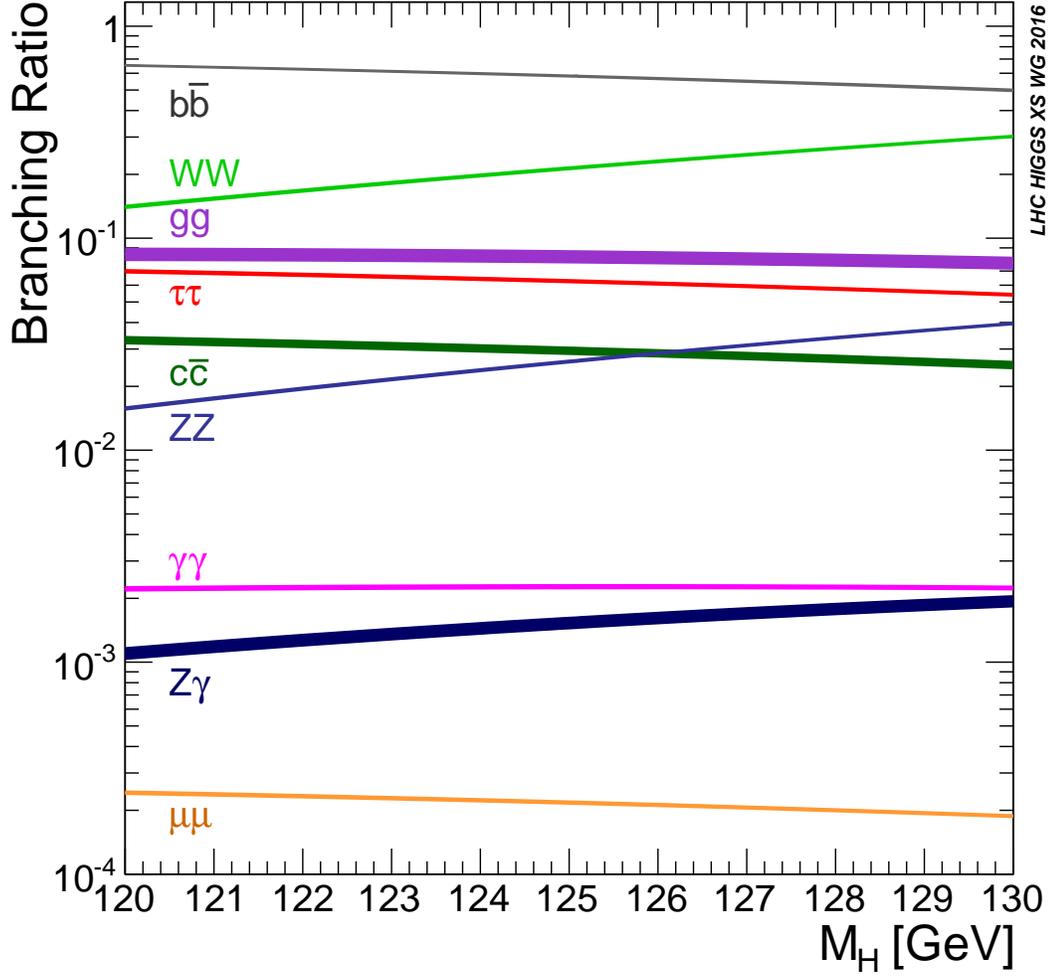


Figure 4.3: The branching ratios for the decay of the SM Higgs boson with respect to different masses.

the prescription explained in [41] and are given as:

$$\mathcal{B}(H \rightarrow XX) = \frac{\Gamma(H \rightarrow XX)}{\sum_{Y \in SM} \Gamma(H \rightarrow YY)}. \quad (4.5)$$

The $b\bar{b}$ and $\tau^+\tau^-$ channels are very important for understanding the Higgs interaction with the fermions. The $b\bar{b}$ channel has the largest branching ratio due to the high mass of this fermion, but the measurements of this channel

suffer a lot due to the large expected QCD background. The $\tau^+\tau^-$ channel has the second highest fermionic branching ratio, and due to effective background suppression, this channel is very promising to study. The fermionic SM Higgs boson coupling is directly proportional to the fermion masses, and its partial decay width can be expressed quantitatively using the Born approximation as follows:

$$\Gamma_{Born}(H \rightarrow f\bar{f}) = \frac{G_F N_C}{4\sqrt{2}\pi} m_H m_f^2 \beta^3, \quad (4.6)$$

where m_f is the fermion mass, m_H is the mass of the Higgs boson, and N_C is the QCD color factor; β is the velocity of the fermion in the final state which can be expressed as,

$$\beta = \sqrt{1 - \frac{4m_f^2}{m_H^2}}, \quad (4.7)$$

and finally, G_F is the Fermi coupling constant,

$$\frac{G_F}{\sqrt{2}} = \frac{\pi g'^2}{2m_W^2(1 - m_W^2/m_Z^2)}. \quad (4.8)$$

Here, m_W and m_Z are the masses of W and Z bosons, respectively. If the Higgs boson decays specifically into a pair of quarks, then the partial decay width is given as:

$$\Gamma_{NLO}(H \rightarrow q\bar{q}) \simeq \frac{3G_F}{4\sqrt{2}\pi} m_H m_q^2 \left[1 + \frac{4g_s}{3\pi} \left(\frac{9}{4} + \frac{3}{2} \ln \frac{m_q^2}{m_H^2} \right) \right], \quad (4.9)$$

where m_q is the mass of the quark.

This dissertation focuses on a fermionic Higgs decay, specifically into a pair of τ leptons, which further decay into a pair of muons. This Higgs decay, $H \rightarrow \tau^+\tau^-$, is henceforth referred to as $H \rightarrow \tau\tau$ and a pair of tau decay with

dimuon in the final state will be referred as $\tau_\mu\tau_\mu$. Together with the $\tau_e\tau_e$ channel, the $\tau_\mu\tau_\mu$ channel is the final state with the smallest branching factor. The branching fractions of all tau lepton pairs are given in Table 4.4.

Table 4.4: The branching fraction of the $\tau\tau$ decay.

Decay mode	BR in %
$\tau_h\tau_h$	42
$\tau_e\tau_h$	23
$\tau_\mu\tau_h$	23
$\tau_e\tau_\mu$	6
$\tau_e\tau_e$	3
$\tau_\mu\tau_\mu$	3

4.3 Drell-Yan Process

The Drell-Yan (DY) process is a process where a quark and anti-quark pair annihilates in a hadron-hadron collision annihilates by creating a virtual photon (γ^*) or Z boson which further decays into oppositely charged leptons:

$$q + \bar{q} \rightarrow Z/\gamma^* \rightarrow \ell + \bar{\ell} \quad (4.10)$$

The Feynman tree-level diagram for this process is shown in Figure 4.4 [42]. At the LHC, both the Drell-Yan channels with electrons or muons in the final state are being studied. Measurement of the Drell-Yan production cross section allows for example the extraction of the quark structure. At LHC, the $Z/\gamma^* \rightarrow \mu\mu$ is a very clear channel to study and often is referred to as the ‘‘Golden Channel’’ whereas the electron channel is very sensitive due to high background. The Drell-Yan process is explained in great detail in

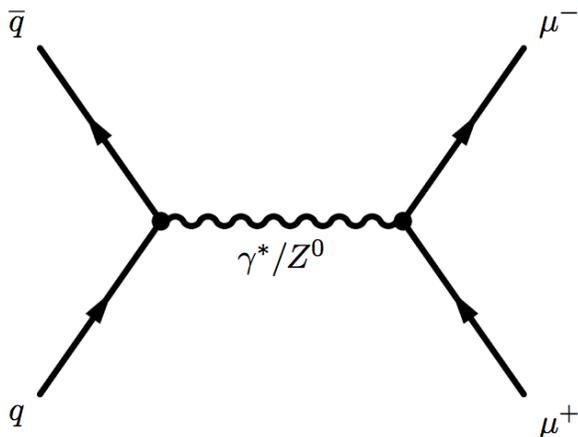


Figure 4.4: Feynman diagram for Drell-Yan process where quark and anti-quark annihilate by creating either virtual photon (electromagnetic process) or Z (electroweak process) which further decays into a muon and anti-muon pair.

Refs. [43, 44]. This Drell-Yan process is the dominant and irreducible background in the $Z \rightarrow \tau\tau$ and $H \rightarrow \tau\tau$ analyses and especially in the final state with a dimuon. Hence, in this dissertation this process has been intensively studied.

Chapter 5

Physics Objects and Analysis

Tools

5.1 Particle Objects and Reconstruction

In CMS, particle identification and reconstruction is performed using a particle-flow (PF) algorithm [45, 46, 47, 48] that combines the information from all of the CMS sub-detectors to identify and reconstruct each particle emerging from proton-proton collisions. These particles are classified as charged hadrons, neutral hadrons, photons, muons and electrons. By combining these PF objects, high-level objects such as jets, hadronic τ , or missing transverse momentum (E_T^{miss}) are reconstructed. The positions of reconstructed vertices of pp interactions fall within 24 cm of the nominal detector center and the radial coordinates within 2 cm from the beam spots, these vertices are referred to as primary vertices. The hard interaction vertex has the maximum $|\vec{p}_T|$ sum of all the tracks associated with it. All other vertices with lower p_T are coming from soft scat-

tering and they are referred as the pile-up (PU) events.

5.1.1 Muons

In the CMS detector, the muon reconstruction is first done independently in the muon chambers (CSC, DT, and RPC) using standalone tracks, and in the inner silicon trackers using the tracker tracks [49]. “Local reconstruction” is the first step of the reconstruction where the muons are reconstructed within one chamber. A combination of the reconstructed tracks within the muon chambers with inner tracker tracks is referred to as the “global muon reconstruction”. A detailed description of the local and global muon reconstruction can be found in [49] and [50], respectively. Standalone muons are fitted with Kalman-filter techniques [51]. So-called tracker muons are built “inside-out” by propagating the tracks from the inner tracker to the muon system by matching tracks to DT and CSC segments. The global muons are built “outside-in”, starting from standalone muons and extrapolating them to the inner tracks that match geometrically. During Run II, two specific reconstruction procedures are added to keep the muon reconstruction and isolation efficiency as high as possible in a high pile-up environment. In the first procedure, to increase the track hit efficiency, inner tracks identified as tracker muons are refitted by relaxing some quality constraints. In the second procedure, standalone muons with $p_T > 10$ GeV that fulfill a minimal set of quality requirements are used to seed an outside-in inner tracking reconstruction step. To reject muons from light hadron decays, one of the best variables is muon isolation which is the sum of the energy in trackers and calorimeters in a geometrical cone ($\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$) surrounding the muons. The other variable based on the muon reconstruction is χ^2 of

the track fit, number of hits on the track or the degree of matching between the inner and the standalone tracks. In different analyses, different identification working points are used such as “Loose”, “Medium”, “Tight”, “Soft”, or “High momentum” muon IDs. In this Run II analysis, the Medium muon ID [52] has been used. Medium muons are loose muons which are either global or tracker muons, reconstructed using the PF algorithm and have a segment compatibility probability between the tracker and muon tracks larger than 0.451. They also have well reconstructed tracker tracks with a strict requirement on the χ^2 matching as well as segment compatibility.

The isolation of the muons can be measured relative to their transverse momentum, p_T^μ , by summing over the transverse momenta of all PF particles within the ΔR cone around the muons. Equation 5.1 shows the mathematical expression for muon isolation.

$$I^\mu \equiv \frac{\sum_{\text{charged}} p_T + \max\left(0, \sum_{\text{neutral}} p_T - \frac{1}{2} \sum_{\text{charged, PU}} p_T\right)}{p_T^\mu}. \quad (5.1)$$

In the $Z \rightarrow \tau\tau$ analysis, the following muon ID and Isolation cuts were used:

Table 5.1: Muon ID and Isolation Cut used in the $Z \rightarrow \tau\tau$ analysis.

Identification Criteria	Isolation Criteria
Medium PF muon	$I_\mu < 0.15 \cdot p_T^\mu$

Table 5.2 referred to identification and isolation criteria used in the $H \rightarrow \tau\tau$ analysis.

Table 5.2: Muon ID and Isolation Cut used in the $H \rightarrow \tau\tau$ analysis.

Identification Criteria	Isolation Criteria
2016 (ICHEP) Medium muon (For Single Muon dataset Run B-F)	$I_\mu < 0.15 \cdot p_T^\mu$
Medium PF muon (For MC and Single Muon dataset Run G & H)	

5.1.2 Taus

The τ leptons are third generation leptons with a mass of 1776.86 ± 0.12 MeV, heavier than e and μ leptons. The lifetime of the τ leptons is very short (290.3×10^{-15} s) and hence, it decays immediately into lighter particles. In two-thirds of cases, τ leptons decay hadronically, mostly into one or three mesons along with neutral pions and ν_τ . In the other third, they decay either into muon or electron along with respective neutrinos. Table 5.3 summarizes the tau decays with respective branching ratios. From now onwards, leptonic τ decay, *i.e.* $\tau \rightarrow e\bar{\nu}_e\nu_\tau$ and $\tau \rightarrow \mu\bar{\nu}_\mu\nu_\tau$ are represented by symbols τ_e and τ_μ respectively, and hadronic τ decays are represented as τ_h . Muons originating from τ decays are reconstructed with methods described in Section 5.1.1. The hadronically decaying tau is reconstructed with the hadron plus strip (HPS) algorithm [53, 54], which is seeded with the anti- k_T [55] jets.

The HPS algorithm is carried out in two steps: reconstruction and identification. In the first step, the tau leptons, which decay hadronically are reconstructed from the tracks and ECAL strips in the $\eta - \phi$ plane with energy deposits in various hadronic decay modes. In Run I, the size of the strips was fixed, while in Run II, it is now allowed to vary dynamically to account for interaction between the tracker material and the low- p_T secondary particle. In

a second identification step, particles that can be misidentified as a hadronic tau are rejected. In the latter case, these particles are usually quark or gluon jets, electrons or muons. Multivariate (MVA) base discriminators are used to identify the quark or gluon jets from the identified hadronic tau candidates. Misidentified electrons and muons are rejected using MVA-based and cut-based discriminators, respectively.

Table 5.3: Tau decay modes with respective branching fractions \mathcal{B} and the intermediate resonances. Here h stand for kaons or pions.

Decay mode		Resonance	$\mathcal{B}[\%]$
leptonic	$\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau$		17.8
	$\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$		17.4
hadronic	$\tau^- \rightarrow h^- \nu_\tau$		11.5
	$\tau^- \rightarrow h^- \pi^0 \nu_\tau$	$\rho(770 \text{ MeV})$	26.0
	$\tau^- \rightarrow h^- \pi^0 \pi^0 \nu_\tau$	$a_1(260 \text{ MeV})$	10.9
	$\tau^- \rightarrow h^- h^+ h^- \nu_\tau$	$a_1(260 \text{ MeV})$	9.8
	$\tau^- \rightarrow h^- h^+ h^- \pi^0 \nu_\tau$		4.8
	Other hadronic modes		1.8

5.1.3 Jets

The quark and gluon jets produced during collisions are reconstructed using the anti- k_T clustering algorithm [55] implemented in the FASTJET library [56, 57]. The jets are built using the PF objects. One can summarize the definition of the algorithm by redefining the distance measured between two entities, such as particles and pseudojets. Lets d_{ij} be the distance between these two entities i^{th} (particle) and j^{th} (pseudojet). d_{iB} be the distance between the particle i^{th}

and the beam (B), then

$$d_{ij} = \min(k_{ti}^{-2}, k_{tj}^{-2}) \frac{\Delta_{ij}^2}{R^2}, \quad d_{iB} = k_{ti}^{-2}, \quad (5.2)$$

where k_{ti} and k_{tj} are the transverse momenta of i^{th} and j^{th} entities, respectively. $\Delta_{ij}^2 = (\phi_i - \phi_j)^2 + (\eta_i - \eta_j)^2$ and R is the chosen cone parameter. In the Run I, this parameter cone was 0.5, but in the Run II it is set to be 0.4. If the smaller distance between the entities is of d_{ij} -type, then the two entities are combined to form a new entity, but if the smaller distance is of type d_{iB} , then the i^{th} entity is considered a jet and is removed from the list of entities. This procedure continues until the entity list becomes empty. The anti- k_T jet algorithm produces jets with a conical shape with radius R . The jet boundary in the algorithm is flexible with respect to hard radiation compared to soft radiation. The jet energies are corrected to have uniform response in η and absolute calibration in p_T . An offset correction is applied to jet energies to take into account the contribution from additional pp interactions within the same or nearby bunch crossings. The energy of a jet is calibrated based on simulation and data through correction factors, formally known as Jet-Energy Corrections (JEC) [58]. Analyses in this dissertation use jets that are required to have $p_T > 30$ GeV and $|\eta| < 4.7$. These jets should be separated from the selected leptons at least by $\Delta R > 0.5$. The following table displays the global tags used to apply the JEC to 2015 and 2016 data, and corresponding simulated events used in the analyses.

The combined secondary vertex (CSV) algorithm is used to identify jets that are likely to originate from a b quark (“b-jets”). The algorithm uses the

Table 5.4: Global tags used to apply jet energy corrections to 2015 and 2016 data and corresponding simulated events.

Event type	Global tag
2015 Data	76X_dataRun2_16Dec2015_v0
Simulation	76X_mcRun2_asymptotic_RunIIFall15DR76_v1
2016 data	80X_dataRun2_2016SepRepro_v7

track-based lifetime information along with secondary vertices associated with the jet to provide a likelihood ratio discriminator for the b-jet identification. A set of p_T -dependent correction factors are applied to MC events to account for differences in the b-tagging efficiency between data and simulation. This algorithm selects real b-jets with an efficiency of 70% using the working points given by the JETMET POG [59] and only about 1% of light-flavor quark jets are being misidentified as b-jets.

5.1.4 Missing Transverse Energy

Neutrinos are not detected by the CMS detector, but information about their presence can be collected by studying the momentum imbalance in the transverse plane to the beam axis. The details of the measurement of the missing transverse energy (E_T^{miss}) can be found in [60]. The measurement strongly depends on the reconstruction of all physics objects. Also pile-up interaction, misidentification of physics objects, and detector malfunctioning can affect these measurements. There are two types of reconstruction methods of E_T^{miss} that are used in the Run II analyses. The first method is PF E_T^{miss} in which transverse momenta of all visible PF particles are used in reconstruction. The second method is referred to as MVA E_T^{miss} , which is designed to reduce the influence of

the pile-up interaction. In this dissertation, the MVA E_T^{miss} was used in Z cross section measurements, whereas the PF E_T^{miss} was used in the $H \rightarrow \tau\tau \rightarrow \mu\mu$ analysis. The bias in the measurement of E_T^{miss} is reduced by correcting the p_T of the jets with jet energy correction as mentioned in Section 5.1.3. This correction is formally known in the CMS as a “Type-I correction”.

5.1.5 Di- τ Mass Reconstructions

The invariant mass of a pair of τ leptons ($m_{\tau\tau}$) is reconstructed using the Secondary Vertex Fit (SVFit) algorithm. The details on the di- τ mass calculation can be found in [61]. The $m_{\tau\tau}$ is computed by reconstructing kinematic quantities, momenta of the visible decay products of the τ leptons, and reconstructed E_T^{miss} and its resolution, event-by-event using the Dynamic Likelihood Method (DLM) [62, 63]. The kinematics of the leptonic tau decays is defined by these three variables:

- the fraction X , of the tau lepton energy in the lab frame carried by the visible decay products,
- the angle ϕ , between the tau lepton vector and the momentum vector of the visible tau decay product,
- and the invariant mass $m_{\nu\nu}$, of the neutrino system produced in the decay.

In case of the $\tau_\mu\tau_\mu$ channel, the likelihood formalism depends on six kinematic parameters of two muons and transverse components of the missing transverse momentum (E_x^{miss} , E_y^{miss}) which constrain the momenta of the 4 neutrinos. Figure 5.1 shows the reconstructed di- τ mass distribution and the visible dimuon

mass distribution in the $H \rightarrow \tau\tau$ decay. Due to the missing transverse energy, the visible mass distribution peaks at lower mass range compared to the di- τ mass distribution.

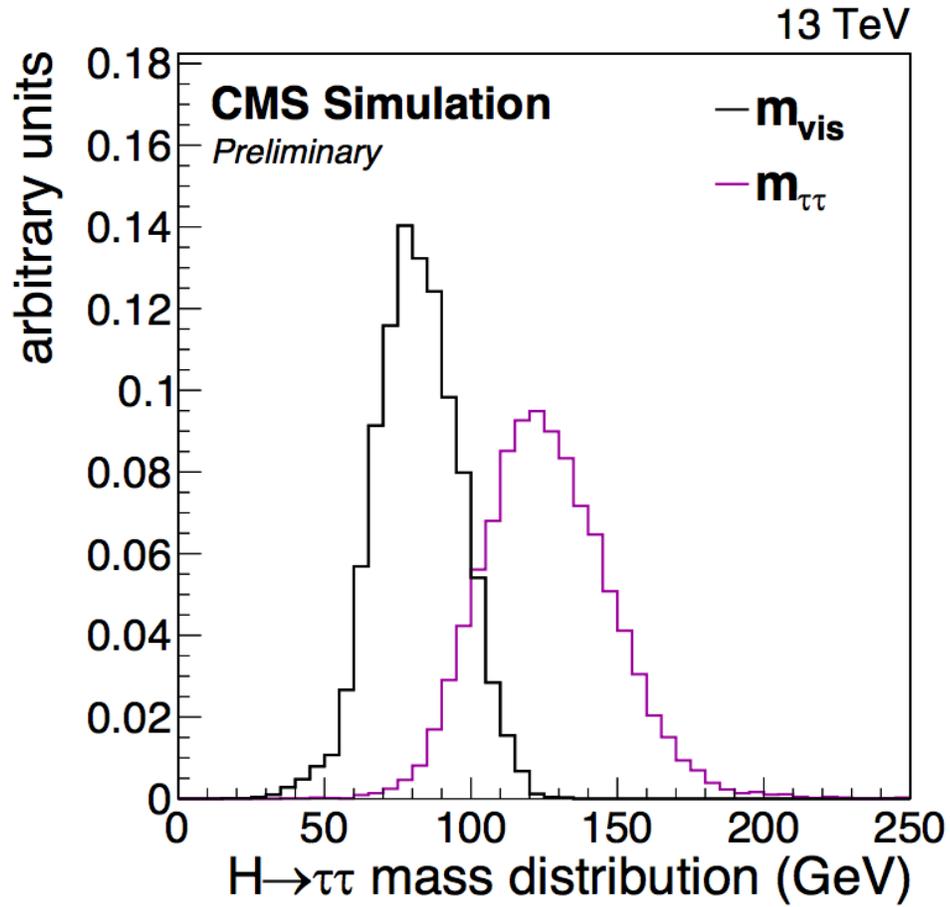


Figure 5.1: Reconstructed di- τ mass with SVFit algorithm and visible dimuon mass in the $H \rightarrow \tau\tau$ decay using simulated 125 GeV Higgs Decays.

5.2 Boosted Decision Trees

A decision tree is an event classifier with binary tree structure where events are classified as signal-like or background-like based on a single discriminating variable at each node until a stop criterion is achieved. A schematic diagram of the decision tree is shown in Figure 7.18 [64]. All events split into several regions until they reach the final leaf node, classified as a signal or background. With boosting, this concept is extended from a single tree to several trees, eventually creating a forest. At the end, all trees are combined into single classifier, which is given by an average of the individual decision trees. The boosting can enhance the performance of each single tree by stabilizing the response of the decision tree with respect to fluctuations in the training samples as well as improves the separation performance compared to the single decision tree. During training a decision tree, higher events weights are assigned to the misclassified events. Therefore subsequent tree is trained using a modified event sample where weights of previously misclassified events are multiplied by a common boost weight. This is the idea behind the boosting. Based on this concept, the boosted decision tree (BDT) classifier is defined. In the following chapters, the gradient boost is used for training the decision trees; a detailed explanation of the gradient boost is described in the following section.

5.2.1 Gradient Boost

The simple additive expansion concept is used to estimate a function through the boosting method. Lets consider the function $F(x)$ which is a weighted sum of the parametrized functions $f(x; a_m)$ characterized by parameters $a = \{a_1, a_2, \dots, a_m\}$,

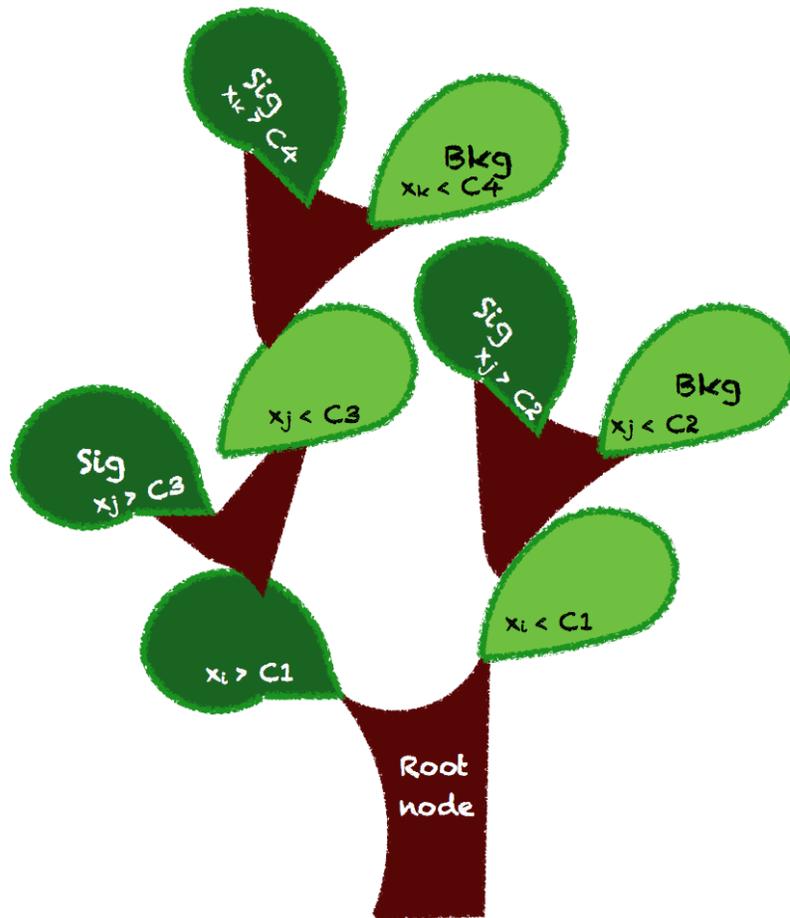


Figure 5.2: Schematic view of a decision tree. Starting from the root node, a sequence of binary splits using the discriminating variables x_i is applied to the data. Each split uses the variable that at this node gives the best separation between signal and background when being cut on. The same variable may thus be used at several nodes, while others might not be used at all. The leaf nodes of the tree are labeled “Sig” for signal and “Bkg” for background depending on the majority of events that end up in the respective nodes in training sample with known background and signal events.

also referred as “weak learners”, then the function can be expanded as follows:

$$F(\mathbf{x}; P) = \sum_{m=0}^M \beta_m f(x; a_m); P \in \{\beta_m; a_m\}_0^M. \quad (5.3)$$

where β_m is boost weight. While boosting, the parameters P are adjusted such that the loss function $L(F, y)$ obtained from the training sample is minimized. A simple loss function in boosting is given by $L(F, y) = (F(\mathbf{x}) - y)^2$, where y is the true value obtained from the training sample. In the case of the GradientBoost method, the algorithm uses a more robust binomial log-likelihood loss function for classification:

$$L(F, y) = \ln(1 + e^{-2F(\mathbf{x})y}) \quad (5.4)$$

Since, the minimization of the loss function is not a very straight forward method, one has to adopt the steepest-descent approach. This is done by calculating the current gradient of the loss function and then growing a regression tree whose leaf values are adjusted to match the mean value of the gradient in each region defined by the tree structure. The desired set of trees can be achieved by iterating this procedure to minimize the loss function. If the gradient calculation is feasible, then the GradientBoost can be adapted to any loss function.

5.2.2 TMVA and Boosted Decision Trees

The Toolkit for Multivariate Analysis (TMVA) [64] software package provides an integrated ROOT [65] environment for each type of the multivariate analyses. It provides training, testing, and a performance evaluation algorithm for each of the multivariate methods. This document focuses only on the Boosted Decision Tree (BDT) method. Various configurations and cuts can be applied to improve the performance of the training. These cuts can be found in the TMVA manual guide [64].

The ranking of each variable is based on how often that variable is used at each node for determining the maximum separation between the signal and the background. The variable that is used the maximum amount of times is referred as the best-ranked variable. The detailed performance of this multivariate method for the di-tau analyses is described in the Chapters 6 and 7.

5.3 Statistical Analysis Using Combine Tools

Statistical inferences are used in high energy physics to express the results of searches for a particle into a statement about the evidence. By defining how well the signal, backgrounds, and uncertainties are modeled in the given analysis, statistical significances can be estimated. At the LHC, the statistical methodology for Higgs Boson searches have been developed by the ATLAS and CMS collaborations, formally know as the LHC Higgs Combination Group [66]. In the following section the statistical methods used for the Z cross section measurement as well as for the searches of the SM Higgs Boson decay in the di-tau channel are explained. The likelihood is introduced to describe the systematic uncertainties. The maximum likelihood fit is applied to obtain the final value of the parameter of the interest (POI), such as the cross section or the signal strength. Also, the consistency of the background modeling with the data, the goodness-of-fit (GOF) is studied. It is basically a χ^2 test to measure the agreement between the observed and expected signal.

5.3.1 Likelihood

The likelihood is defined as the estimation of the combined probability density function (PDF) of all measurements in the data. The agreement between expectation and observation in a simple counting experiment can be quantified by the likelihood as:

$$\mathcal{L}(n|b) = \frac{e^{-b}b^n}{n!} \quad (5.5)$$

where n is the number of observed events and b is the number of expected events. If the data are binned, i.e if a histogram with N independent bins are considered, then the likelihood in this case can be defined as the product of the likelihood of the individual bins of the distribution:

$$\mathcal{L}(\vec{n}|\vec{b}) = \prod_{i=0}^N \frac{e^{-b_i}b_i^{n_i}}{n_i!} \quad (5.6)$$

here \vec{n} is the vector of the observed data in various individual bins and \vec{b} is the vector of corresponding expected events. If the data are not binned and instead described by some PDF $f_b(x)$ with some observable x , then the likelihood can express as follows [67]:

$$\mathcal{L}(\vec{x}|b, f_b(x)) = k^{-1} \prod_{i=0}^k b f_b(x_i) e^{-b} \quad (5.7)$$

where k is the number of events observed and b is the corresponding expected number of events in the full range of x .

In a physics analysis, systematic uncertainties are considered as nuisance parameters that impact the normalization and shape of the distribution. The following are examples of different sources of uncertainties considered in the

analyses discussed in the later chapters.

1. Theoretical uncertainties on cross section or parton distribution function uncertainties,
2. Statistical uncertainties in background estimate due to limited number of events in MC simulations or in observed events in a control region,
3. Experimental uncertainties on luminosity or trigger efficiency measurements.

If these uncertainties are included in the likelihood, then for L nuisance parameters it can be expressed as:

$$\mathcal{L}(\vec{n}|\vec{b}) = \prod_{i=0}^N \frac{e^{-b_i} b_i^{n_i}}{n_i!} \prod_{j=0}^L p(\bar{\theta}_j|\theta_j) \quad (5.8)$$

In Equation 5.8, $p(\bar{\theta}_j|\theta_j)$ represents the PDF for nuisance parameter θ , where $\bar{\theta}$ is considered the default value of the parameter and reflects the degree of belief on what the real value of parameter θ is. According to Bayes' theorem, the Bayesian probability $p(\bar{\theta}_j|\theta_j)$ can be expressed as a function of the frequentist probability $\rho(\theta|\bar{\theta})$, common choice [66] is usually a log-normal PDF with κ , the width of the log-normal distribution:

$$\rho(\theta|\bar{\theta}) = \frac{1}{\sqrt{2\pi \ln(\kappa)}} \exp\left(-\frac{(\ln(\theta|\bar{\theta}))^2}{2(\ln \kappa)^2}\right) \frac{1}{\theta} \quad (5.9)$$

Therefore, the likelihood function is the product of Poisson probabilities for all the bins in the range of x .

The uncertainties that affect the shape of the distribution of the POI [68]

can be model with a linear extrapolation method [69]. In CMS, these uncertainties are implemented by providing two alternating shapes and by varying the nuisance parameter by ± 1 standard deviation. If N shape nuisance parameters are modeled with the parameter $\vec{\theta} = (\theta_1, \theta_2, \dots, \theta_N)$, then the distribution as a function of the shape nuisance parameters $\vec{\theta}$ is given as [37]:

$$h(\vec{\theta}) = h_0 + \sum_{j=0}^n (a(\theta_j)h_j^+ + b(\theta_j)h_0 + c(\theta_j)h_j^-) \quad (5.10)$$

where h_0 is the nominal histogram distribution, h_j^+ is the histogram with a variation of +1 standard deviation of the j^{th} nuisance parameter, and h_j^- is the histogram with a variation of -1 standard deviation of the j^{th} nuisance parameter, and

$$a(\theta) = \begin{cases} \theta(\theta + 1)/2 & \text{if } |\theta| \leq 1 \\ 0 & \text{if } \theta < -1, \\ \theta & \text{if } \theta > +1 \end{cases} \quad (5.11)$$

$$b(\theta) = \begin{cases} -\theta^2 & \text{if } |\theta| \leq 1 \\ -(|\theta| - 1) & \text{if } |\theta| > 1, \end{cases} \quad (5.12)$$

$$c(\theta) = \begin{cases} \theta(\theta + 1)/2 & \text{if } |\theta| \leq 1 \\ 0 & \text{if } \theta > +1. \\ |\theta| & \text{if } \theta < -1 \end{cases} \quad (5.13)$$

Bin-by-bin (bbb) [70] uncertainties are used to determined the shape uncertainties on the number of MC events in every bin of the distribution.

Maximum Likelihood Fit

A maximum likelihood fit is performed to find the POI which gives the best agreement between the expected and observed values. Two common scenarios considered while fitting are the background-only-fit and the signal-plus-background-fit. In the first case, the nuisance parameters acting on the expected background distribution are varied to the value $\hat{\theta}$ that maximize the likelihood function $\mathcal{L}(\vec{n}|\vec{b}, \vec{\theta})$. In the latter case, the nuisance parameters and a freely floating signal strength μ of expected signal distribution are varied to their optimal values to maximize the likelihood $\mathcal{L}(\vec{n}|\mu\vec{s} + \vec{b}, \vec{\theta})$. The variations of the nuisance parameters after computing the maximum likelihood fit are called pulls, which describe the coherence between the signal and background modeling.

5.3.2 Limits

The confidence interval (CI) method [71] is used in CMS physics analyses to set an upper limit boundary on the production cross section of a signal when no excess in data over expected background is observed. The profile likelihood ratio (PLR) used to obtain the CI on the POI is given as:

$$q_\mu = -2 \ln \frac{\mathcal{L}(\vec{n}|\mu\vec{s} + \vec{b}, \hat{\theta}_\mu)}{\mathcal{L}(\vec{n}|\mu\vec{s} + \vec{b}, \vec{\theta})}, \text{ with a constraint } 0 \leq \hat{\mu} \leq \mu. \quad (5.14)$$

In this test statistic, signal strength is fixed in the numerator and the nuisance parameters are allowed to float to maximize the likelihood, whereas, in the denominator, both the signal strength and the nuisance parameters are allowed to freely float to maximize the likelihood. If $f(q_\mu|\vec{b}, \hat{\theta}_0^{obs})$ and $f(q_\mu|\mu\vec{s} + \vec{b}, \hat{\theta}_\mu^{obs})$ are the probability density functions that represent the test-statistic distributions

for the background-only and signal-plus-background hypotheses, respectively, with corresponding nuisance parameters θ_0^{obs} and θ_μ^{obs} , then the probabilities to obtain the observations with both hypotheses can be defined as:

$$p_{\mu s+b} = P(q_\mu \geq q_\mu^{obs} | \text{signal-plus-background}) = \int_{q_\mu^{obs}}^{\infty} f(q_\mu | \mu \vec{s} + \vec{b}, \hat{\theta}_\mu^{obs}) dq_\mu, \quad (5.15)$$

and

$$1 - p_b = P(q_\mu \geq q_\mu^{obs} | \text{background-only}) = \int_{q_\mu^{obs}}^{\infty} f(q_\mu | \vec{b}, \hat{\theta}_\mu^0) dq_\mu, \quad (5.16)$$

The ratio of these probabilities gives the CI value.

$$CI(\mu) = \frac{p_{\mu s+b}}{1 - p_b} \quad (5.17)$$

Equation 5.17 gives conservative limits, but this non-conventional definition allows to test cases, where the signal is very small and both hypothesis are compatible with the observation or where the signal strength is negative due to a deficit in data. To obtain the expected median limit, and the ± 1 , and ± 2 standard deviation bands, a large number of pseudo-datasets for background-only expectation (Asimov dataset) is generated and μ , the signal strength is computed until $CI = 0.05$. Then cumulative distribution functions can be built and the following limits can be extracted: median (50% quantile), $\pm 1\sigma$ (16% and 84% quantile) band, and $\pm 2\sigma$ (2.5% and 97.5% quantile) band. The Asimov dataset can be used to extract the observed rate corresponding to the expected in terms of χ^2 distribution using the asymptotic limits [72].

The local probability (p-value) for the background-only hypothesis is calcu-

lated if an excess is observed in data over the predicted background. The p-value corresponds to the probability that the fluctuating background creates excess events as large or larger than the observed events. This is usually converted into an equivalent significance.

5.3.3 Goodness of Fit

A GOF [73] test is a test of the null hypothesis when the alternative hypothesis is not specified. A likelihood ratio can be built independently without the parameters from the original model; such a model is referred as a “saturated model”. In this model, the alternate hypothesis has exactly the same observed data in each bin of the distribution and the likelihood corresponds to $\mathcal{L}_{sat}(\vec{n}|\vec{n})$. To test the compatibility of the expected and observed signal, likelihoods for the toy MC pseudo-datasets are generated for the given hypothesis. By minimizing the maximum likelihood, the minimal test statistic q_μ is obtained for each toy MC pseudo-dataset. If the observed values remain within the bulk of the distribution, then there is a good agreement between the data and expected events. However, if the observed values lie in the tail region, it indicates a mismodeling of the expected events.

Chapter 6

Cross Section Measurements of the Z Boson in the $Z/\gamma^* \rightarrow \tau\tau$ Decay Channel

6.1 Introduction

The DY production of Z boson and its decay into τ pairs ($q\bar{q} \rightarrow Z/\gamma^* \rightarrow \tau\tau$) plays an important role in the LHC physics studies. A cross section measurement of the Z boson is performed to validate the analysis techniques used later in the SM $H \rightarrow \tau\tau$ analysis. The process $Z/\gamma^* \rightarrow \tau\tau$ is used to study reconstruction and identification efficiencies of hadronic τ decays. In addition, the $Z/\gamma^* \rightarrow \tau\tau$ is a dominant background to SM $H \rightarrow \tau\tau$ analysis and the reducible backgrounds in the $Z/\gamma^* \rightarrow \tau\tau$ analysis are also relevant in the $H \rightarrow \tau\tau$ analysis.

In this chapter, the analysis focuses on the $\tau_\mu\tau_\mu$ decay channel and the

cross section is measured at $\sqrt{s} = 13$ TeV center-of-mass energy, using 2.3 fb^{-1} of LHC Run 2 data recorded by the CMS experiment in 2015. The CMS and ATLAS collaboration results on the $Z/\gamma^* \rightarrow \tau\tau$ cross section measurement using LHC Run 1 data at $\sqrt{s} = 7$ and 8 TeV can be found in Refs [74, 75]. Other decay channels i.e. $\tau_e\tau_h$, $\tau_\mu\tau_h$ and $\tau_e\tau_\mu$ have been studied by different groups in the CMS collaboration. The final cross section times branching fraction measurements for the $Z/\gamma^* \rightarrow \tau\tau$ process is obtained by combining individual cross sections measured in each decay channel.

6.2 CMS Data Samples and Monte Carlo Simulations

This analysis uses the data collected in the beginning of the Run II period in 2015 at a center of mass energy of $\sqrt{s} = 13$ TeV and with an integrated luminosity of 2.3 fb^{-1} . The bunch crossing period is 25 ns and on average there are approximately 20 inelastic pp interactions (pileup) taking place in each bunch crossing period. Events recorded with single muon triggers are used in this analysis. Only data taking periods with all detector systems are included are considered. The dataset corresponding JSON file:

`Cert_13TeV_16Dec2015ReReco_Collisions15_25ns_JSON.txt` [76] used in this analysis are shown in Table 6.1. The data have been reconstructed and analyzed using CMSSW version `CMSSW_7_6_x`.

The Monte Carlo (MC) simulated samples used in this analysis are summarized in Table 6.2.

Table 6.1: Datasets used in the $\tau_\mu\tau_\mu$ channel with JSON file used to apply a good-run selection.

$\tau_\mu\tau_\mu$ channel		
Dataset Name	Run-range	Luminosity [$\text{cm}^{-2}\text{s}^{-1}$]
SingleMuon (Run2015D)	256630–260627	2.26

Table 6.2: List of MC samples used to model the $Z/\gamma^* \rightarrow \tau\tau$ signal and $Z/\gamma^* \rightarrow ee$, $Z/\gamma^* \rightarrow \mu\mu$, W +jets, $t\bar{t}$, single top quark, and diboson background processes. Mass ranges for DY events are given as (a) $10 < m_{\mu\mu} < 50$ GeV and (b) $m_{\mu\mu} > 50$ GeV.

Process	Cross section [pb]
$Z/\gamma^* \rightarrow \mu\mu^a$	22634.0
$Z/\gamma^* \rightarrow \mu\mu^b$	5765.4
$Z/\gamma^* \rightarrow \tau\tau \rightarrow \mu\mu^b$	1967
W +jets	61526.7
$t\bar{t}$	831.8
Single top quark	
ST_tW_top	35.6
ST_tW_antitop	35.6
ST_t-channel_top	44.1
ST_t-channel_antitop	26.2
WW	50.0
WZ	5.6
WZ	5.3
WZ	3.1
WZ	10.7
ZZ	3.2
ZZ	1.2
$H \rightarrow \tau\tau$	
GluGluHToTauTau (M125)	2.8
VBFHToTauTau (M125)	0.2

6.3 BDT Analysis

A multivariate discriminant based on BDT is used to suppress the dominant $Z/\gamma^* \rightarrow \mu\mu$ background from the $Z/\gamma^* \rightarrow \tau\tau \rightarrow \mu\mu$ signal. The BDT cut is chosen by optimizing the signal-over-background ratio. The BDT analysis is performed using the TMVA package. The following variables are used as inputs to the BDT:

- **The ratio of the p_T of the dimuon system to the scalar sum of the p_T of the two muons, $p_T(2\mu)/\sum p_T(\mu)$**

The transverse momentum of the dimuon system is given as:

$$p_T(2\mu) = \sqrt{(p_x^+ + p_x^-)^2 + (p_y^+ + p_y^-)^2} \quad (6.1)$$

where $p_x^+(p_x^-)$ and $p_y^+(p_y^-)$ are the x and y components of the positive (negative) muon momentum, respectively. The scalar sum of the p_T of the two muons with opposite charge is given as $\sum p_T(\mu)$. This sum is larger for the $Z/\gamma^* \rightarrow \mu\mu$ events since the muons produced in this process have larger p_T than the muons produced in the $Z/\gamma^* \rightarrow \tau\tau \rightarrow \mu\mu$ signal events as shown Figure 6.2 (a). Hence, this variable can be useful to differentiate the background events from the signal.

- **The pseudo-rapidity of the dimuon system, $\eta_{2\mu}$**

The pseudo-rapidity of the dimuon system is defined as:

$$\eta_{2\mu} = \frac{1}{2} \times \ln \left(\frac{|p(2\mu)| + p_z(2\mu)}{|p(2\mu)| - p_z(2\mu)} \right) \quad (6.2)$$

where $|p(2\mu)|$ is the magnitude of the three-momentum of the dimuon

system and it is obtained as:

$$|p(2\mu)| = \sqrt{(p_x^+ + p_x^-)^2 + (p_y^+ + p_y^-)^2 + (p_z^+ + p_z^-)^2} \quad (6.3)$$

where p_x and p_y are transverse momenta of muons and p_z is the z component along the beam line. The z component of the momentum of the dimuon system is calculated as:

$$p_z(2\mu) = p_z^+ + p_z^-. \quad (6.4)$$

The eta of the dimuon system is inversely proportional to the missing transverse energy and the distribution is roughly symmetrical around the origin. In the signal events, the amount of momentum carried by the missing transverse energy is higher, hence, the distribution is more narrow compared to the background (Figure 6.2 (b)).

- **The E_T^{miss} reconstructed in the event**

The missing transverse energy is calculated as explained in Section 5.1.4. For the signal events, the τ decay products are mostly collinear, hence, they produce more observable E_T^{miss} compared to the background events. In background events most of the transverse moment is carried by the two muons. Therefore, background events have a slightly lower and more narrow distribution (Figure 6.2 (c)).

- **The P_ζ variable**

The P_ζ variable is defined as the difference in the projection of the visible transverse momentum of the τ decay products plus the missing transverse

momentum and the visible transverse momentum of the τ decay products on the $\vec{\zeta}$ axis, which is the angle bisector of the visible muon momenta as shown in the Figure 6.1 The linear combination of $P_\zeta - 1.85 \cdot P_\zeta^{\text{vis}}$ is

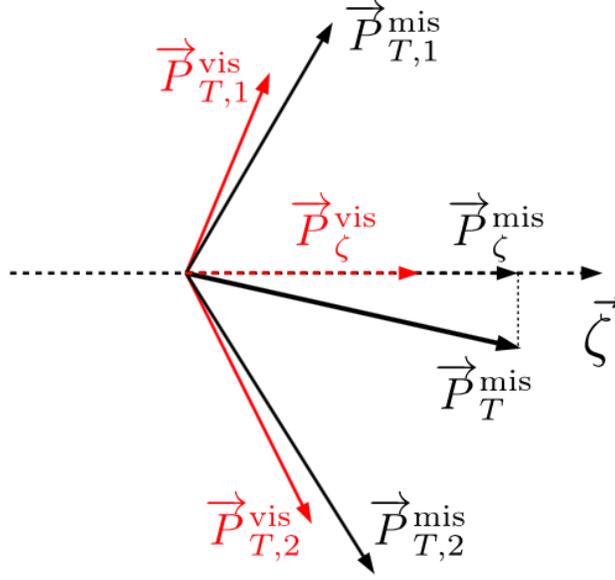


Figure 6.1: The P_ζ variable.

used as a discriminant to suppress the $t\bar{t}$ background from signal events.

Mathematically P_ζ and P_ζ^{vis} can be expressed as follows:

$$P_\zeta = \left(\vec{p}_T^e + \vec{p}_T^\mu + \vec{p}_T^{\text{miss}} \right) \cdot \frac{\vec{\zeta}}{|\vec{\zeta}|} \quad \text{and} \quad P_\zeta^{\text{vis}} = \left(\vec{p}_T^e + \vec{p}_T^\mu \right) \cdot \frac{\vec{\zeta}}{|\vec{\zeta}|} \quad (6.5)$$

The distribution for signal and background events are shown in Figure 6.2 (d).

- **The azimuthal angle between the direction of the three-momentum of the positively charged muon and the \vec{p}_T^{miss} vector, $\Delta\phi(\mu^+, \vec{p}_T^{\text{miss}})$**

The τ decay products are collinear as compared to the $Z/\gamma^* \rightarrow \mu\mu$ decay.

In the $Z/\gamma^* \rightarrow \mu\mu$ decay, two muons are produced back to back, therefore

produce a flat distribution as shown in Figure 6.2 (e). The $\Delta\phi(\mu^+, \vec{p}_T^{\text{miss}})$ and $\Delta\phi(\mu^-, \vec{p}_T^{\text{miss}})$ angles are strongly anti-correlated and show an identical discriminating power to separate signal and background events. Therefore, only one angle $\Delta\phi(\mu^+, \vec{p}_T^{\text{miss}})$ is used in the BDT analysis.

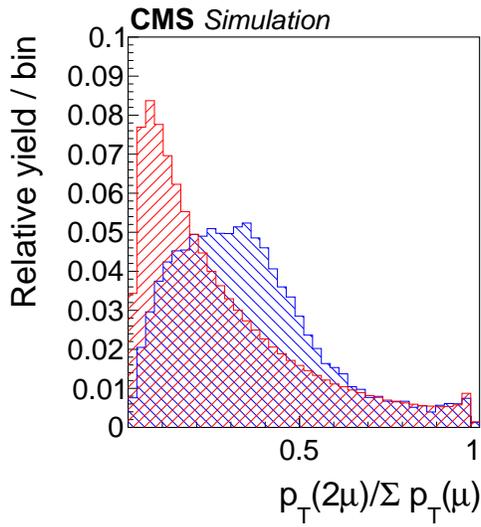
The above mentioned variables help to separate the $Z/\gamma^* \rightarrow \tau\tau \rightarrow \mu\mu$ signal events from the $Z/\gamma^* \rightarrow \mu\mu$ background events. Their level of separation in the signal and background events is shown in Figure 6.2 produced by the TMVA package.

BDT Response

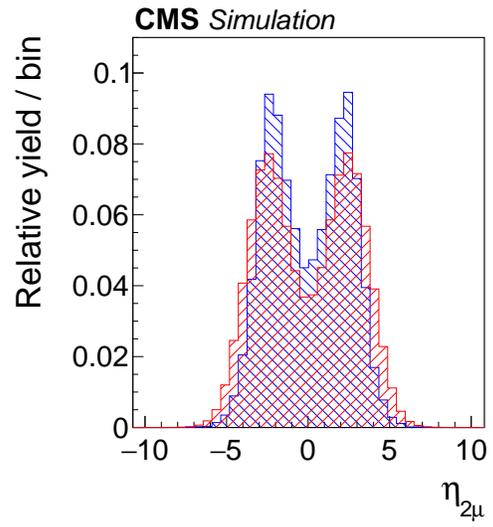
The BDT is constructed using the above input variables in terms of their separation power. The BDT method-specific ranking of the variables is displayed in Table 6.3. The variable with the higher separating power is ranked as the best variable. In this case, the P_ζ variable is ranked as the best variable followed by E_T^{miss} and $p_T(2\mu)/\sum p_T(\mu)$. The BDT response using for the $Z/\gamma^* \rightarrow \tau\tau$ signal

Table 6.3: The BDT method-specific ranking of the input variables. The top variable is the best ranked variable

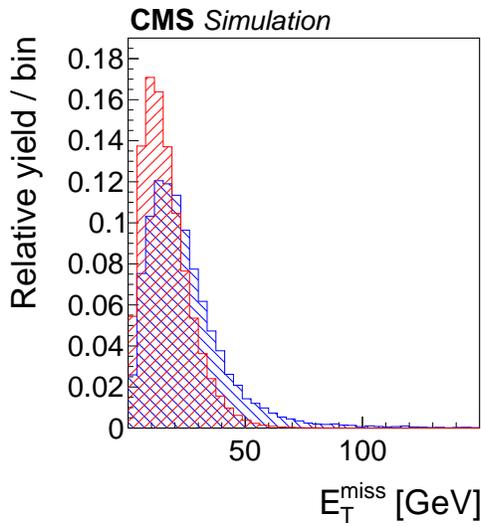
Rank	Variable	Variable Importance
1	P_ζ	2.980×10^{-1}
2	E_T^{miss}	2.878×10^{-1}
3	$p_T(2\mu)/\sum p_T(\mu)$	2.858×10^{-1}
4	$\Delta\phi(\mu^+, \vec{p}_T^{\text{miss}})$	1.238×10^{-1}
5	$\eta_{2\mu}$	4.631×10^{-3}



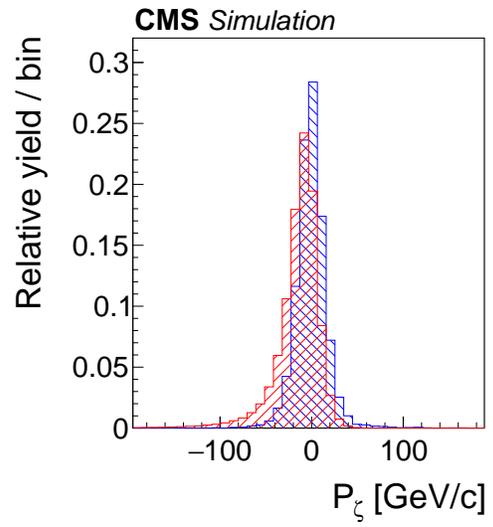
(a)



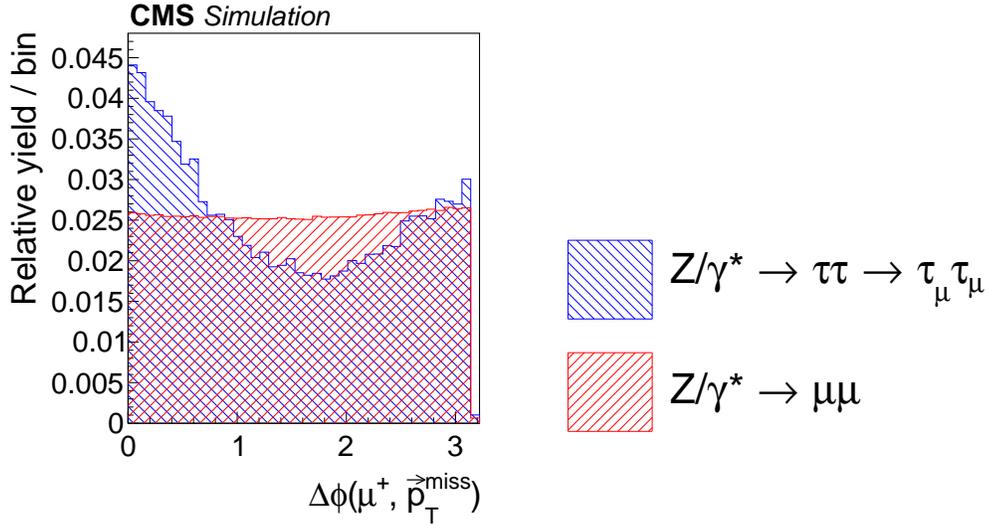
(b)



(c)



(d)



(e)

Figure 6.2: Variables used as inputs to the BDT that separate the $Z/\gamma^* \rightarrow \tau\tau$ signal from the $Z/\gamma^* \rightarrow \mu\mu$ background. Simulated $Z/\gamma^* \rightarrow \tau\tau$ signal events are shown in red and simulated $Z/\gamma^* \rightarrow \mu\mu$ background events in blue. All distributions are normalized to unity.

and the $Z/\gamma^* \rightarrow \mu\mu$ background is shown in Figure 6.3. The BDT response for signal and background is calculated using the GradientBoost algorithm which uses the binomial log-likelihood loss response as defined in equation 5.4. The MVA classifier response using GradientBoost has a range of $\{-1, +1\}$. A small value of the BDT response indicates background-like events whereas a large value indicates signal-like events. The signal-over-background ratio (S/\sqrt{B}) is maximized by optimizing the cut on the BDT response. The optimized BDT cut for this analysis is BDT response > 0.5 .

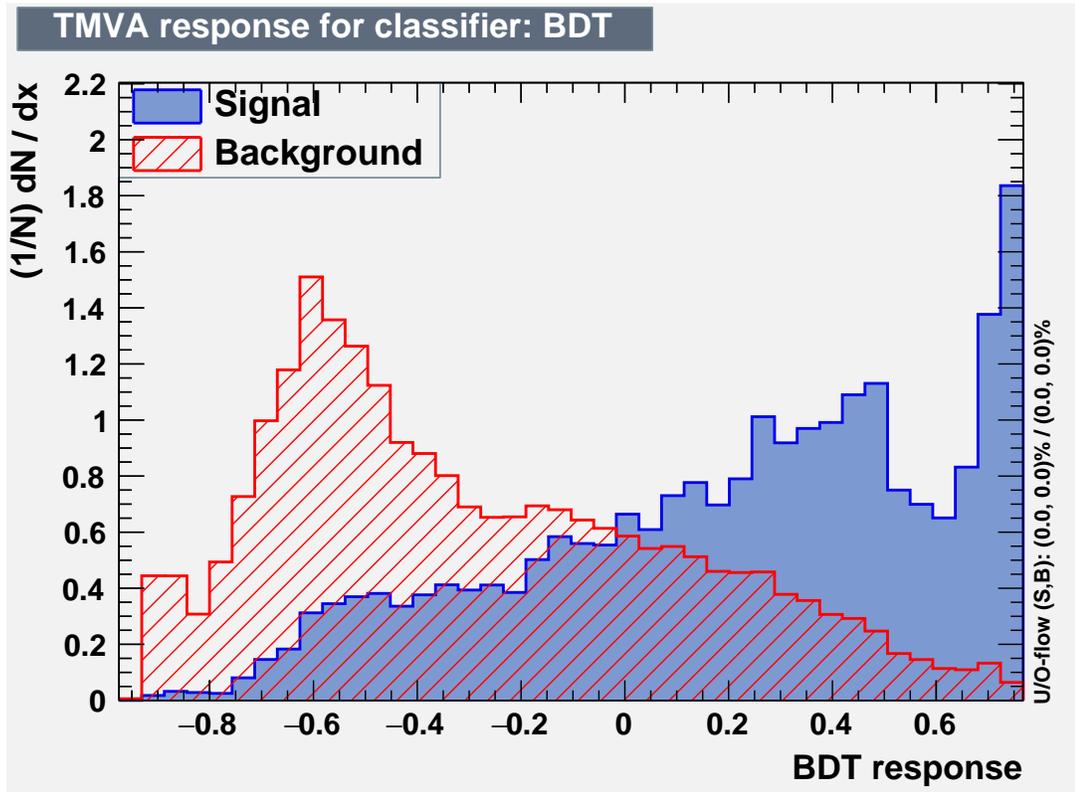


Figure 6.3: The BDT response plotted using TMVA for the $Z/\gamma^* \rightarrow \tau\tau$ signal and the $Z/\gamma^* \rightarrow \mu\mu$ background.

6.4 Event Selection

Events used in this analysis are recorded during the pp collision in stable beam condition, as well as when all the sub-detectors are fully operational. This provides the generic data-quality requirement. The event selection for this channel is based on two criteria, the generic data-quality requirement and the channel-specific selection criteria. The specific selection criteria based on the HLT trigger selection for the process and conditions defined to suppress the specific background.

The generic data-quality criteria applied to all channels are as follows:

- Events selected in the data are required to pass a good-run selection, depending on the JSON file given in Table 6.1.
- The JetMET group [77] has developed filters which reject significantly noisy events, beam background, and detector malfunctioning. Such “event cleaning” should be applied to data.
- In the 2015 data taking period, on average there was ≈ 15 vertices were reconstructed in the given pileup condition. The event vertex in hard-scattering with the highest $\sum p_T^2$ of tracks is considered as a primary vertex (PV) and this PV is reconstructed by the “deterministic annealing” algorithm [78] with $N_{\text{DoF}} \geq 4$ (degrees of freedom) and position $-24 < z_{\text{vtx}} < +24$ cm, $|r| < 2$ cm.

In the $\tau_\mu\tau_\mu$ channel, events are selected using a signal muon HLT trigger with the threshold of $p_T > 18$ GeV that is not pre-scaled. The HLT trigger and the L1 seed used in this channel are given in Table 6.4.

Table 6.4: Trigger paths used by the $\tau_\mu\tau_\mu$ channel in data. The path given in the table remained un-prescaled during the entire data taking period.

HLT Path	L1 Seed	Luminosity
HLT_IsoMu18_v*	L1_SingleMu16	2.26 fb ⁻¹

Table 6.5: Trigger paths in simulated events, used by the $\tau_\mu\tau_\mu$ channel.

HLT Path	L1 Seed
HLT_IsoMu18_v2	L1_SingleMu16

Channel-specific event selection is as follows:

- The event should pass the single-muon trigger as given in Tables 6.4 and 6.5 for the data and MC simulations, respectively.
- The two muons should have opposite charge and an invariant mass of $m_{\mu\mu} < 80$ GeV.
- Muons with $p_T > 10$ GeV and $|\eta| < 2.4$ must pass the medium PF muon identification criteria with a tight isolation of $I_\mu < 0.15 \cdot p_T^\mu$, computed by Equation 5.1.
- At least one muon is required to satisfy $p_T > 20$ GeV and to be matched to an HLT muon object within $\Delta R < 0.5$.
- The BDT multivariate method is used to discriminate the $Z/\gamma^* \rightarrow \mu\mu$ background events from the $Z/\gamma^* \rightarrow \tau\tau \rightarrow \mu\mu$ signal events as discussed in detail in Section 6.3. The output of the BDT discriminator is required to exceed 0.5.

6.5 Corrections applied to MC

6.5.1 Muon identification and isolation efficiency

The $Z/\gamma^* \rightarrow \mu\mu$ sample events in the data and simulations are used to measure the muon identification and isolation efficiencies using the “Tag-and-Probe” technique [79]. Events in the single-muon trigger dataset are selected that have two muons with opposite charges and that fall within the mass window $60 < m_{\mu\mu} < 120$ GeV. To select the tag muon, cuts are applied on p_T and $|\eta|$

variables, where $p_T > 20$ GeV and $|\eta| < 2.1$. In addition to these cuts, muons are required to pass the particle identification and isolation criteria described in Section 5.1.1. For probe muons, the required conditions for transverse momentum and pseudorapidity are $p_T > 10$ GeV and $|\eta| < 2.1$. Additional cuts are applied depending on whether identification or isolation efficiency is measured. For identification, muons are required to be reconstructed as RECO muons and for isolation, the criteria is described in Section 5.1.1. The ratio of particle identification and isolation efficiencies measured in the data and simulation as calculated in Table 6.6 [80] is applied as an event weight to the simulated signal and background events.

Table 6.6: Measured efficiencies for muons to pass the identification and isolation criteria applied in the $\tau_\mu \tau_\mu$ channel compared to MC predictions.

Muon identification and isolation efficiency corrections				
Kinematic range		Data	Simulation	Data/Simulation
$19 < p_T < 25$ GeV	$ \eta < 0.9$	0.813 ± 0.025	0.827 ± 0.017	0.983 ± 0.037
$25 < p_T < 30$ GeV	$ \eta < 0.9$	0.861 ± 0.021	0.877 ± 0.015	0.982 ± 0.028
$30 < p_T < 40$ GeV	$ \eta < 0.9$	0.915 ± 0.016	0.928 ± 0.011	0.986 ± 0.020
$40 < p_T < 60$ GeV	$ \eta < 0.9$	0.952 ± 0.014	0.965 ± 0.009	0.986 ± 0.018
$p_T > 60$ GeV	$ \eta < 0.9$	0.964 ± 0.012	0.976 ± 0.007	0.987 ± 0.015
$19 < p_T < 25$ GeV	$0.9 < \eta < 1.2$	0.836 ± 0.023	0.849 ± 0.019	0.986 ± 0.034
$25 < p_T < 30$ GeV	$0.9 < \eta < 1.2$	0.873 ± 0.021	0.884 ± 0.017	0.987 ± 0.031
$30 < p_T < 40$ GeV	$0.9 < \eta < 1.2$	0.926 ± 0.019	0.935 ± 0.012	0.990 ± 0.020
$40 < p_T < 60$ GeV	$0.9 < \eta < 1.2$	0.961 ± 0.016	0.971 ± 0.010	0.990 ± 0.017
$p_T > 60$ GeV	$0.9 < \eta < 1.2$	0.968 ± 0.013	0.978 ± 0.010	0.989 ± 0.015
$19 < p_T < 25$ GeV	$ \eta > 1.2$	0.871 ± 0.021	0.891 ± 0.016	0.978 ± 0.029
$25 < p_T < 30$ GeV	$ \eta > 1.2$	0.903 ± 0.020	0.918 ± 0.015	0.983 ± 0.026
$30 < p_T < 40$ GeV	$ \eta > 1.2$	0.938 ± 0.017	0.951 ± 0.011	0.987 ± 0.020
$40 < p_T < 60$ GeV	$ \eta > 1.2$	0.966 ± 0.015	0.979 ± 0.009	0.986 ± 0.017
$p_T > 60$ GeV	$ \eta > 1.2$	0.973 ± 0.013	0.984 ± 0.009	0.989 ± 0.015

6.5.2 Muon Trigger Efficiency

The efficiency of the single-muon trigger is measured using $Z/\gamma^* \rightarrow \mu\mu$ events. Selected events have two muons with opposite charges with a mass range $60 < m_{\mu\mu} < 120$ GeV. The tag muon is required to pass the single muon trigger with $p_T > 25$ GeV and $|\eta| < 2.1$. The probe muon is required to satisfy the condition where $p_T > 19$ GeV and $|\eta| < 2.1$. Both muons are required to pass the particle identification and isolation criteria as mentioned above. The trigger efficiency is measured separately for the data and simulated events as a function of the p_T in bins of η . The trigger efficiencies for $\tau_\mu\tau_\mu$ channel are summarized in Table 6.7 [80]. Signal and background events produced by the MC simulation are weighted by the ratio of the trigger efficiencies measured in the data and simulated events.

6.5.3 E_T^{miss} resolution and response

The discrepancy between the data and simulated MC events for the E_T^{miss} variable is corrected using the method described in [81]. It is a data-driven method. The hadronic recoil in $Z/\gamma^* \rightarrow \mu\mu$ events in the data is used to determine the corrections. These corrections are then applied to the simulated $Z/\gamma^* \rightarrow \tau\tau \rightarrow \mu\mu$ signal events as well as to the $Z/\gamma^* \rightarrow \mu\mu$ background events. These corrections are referred to as “Z-recoil corrections”.

6.5.4 b-tag efficiency and mistag rate

Correction factors and tools for the Btag efficiencies and the mistag rate are provided by the b-tag and Vertexing Physics Object Group (BTV POG) [82].

Table 6.7: The efficiencies of the single-muon trigger used in the $\tau_\mu\tau_\mu$ channel and related correction factors applied to simulated events.

Single-muon trigger efficiency corrections				
Kinematic range		Data	Simulation	Data/Simulation
$19 < p_T < 22$ GeV	$ \eta < 0.9$	0.842 ± 0.012	0.889 ± 0.012	0.947 ± 0.019
$22 < p_T < 25$ GeV	$ \eta < 0.9$	0.879 ± 0.012	0.921 ± 0.012	0.954 ± 0.018
$25 < p_T < 28$ GeV	$ \eta < 0.9$	0.898 ± 0.011	0.933 ± 0.011	0.963 ± 0.017
$28 < p_T < 31$ GeV	$ \eta < 0.9$	0.909 ± 0.011	0.942 ± 0.011	0.965 ± 0.016
$31 < p_T < 34$ GeV	$ \eta < 0.9$	0.919 ± 0.011	0.948 ± 0.010	0.970 ± 0.015
$34 < p_T < 37$ GeV	$ \eta < 0.9$	0.925 ± 0.010	0.950 ± 0.010	0.973 ± 0.015
$37 < p_T < 40$ GeV	$ \eta < 0.9$	0.928 ± 0.010	0.954 ± 0.010	0.973 ± 0.015
$p_T > 40$ GeV	$ \eta < 0.9$	0.931 ± 0.010	0.956 ± 0.010	0.974 ± 0.014
$19 < p_T < 22$ GeV	$0.9 < \eta < 1.2$	0.843 ± 0.012	0.896 ± 0.012	0.940 ± 0.018
$22 < p_T < 25$ GeV	$0.9 < \eta < 1.2$	0.867 ± 0.012	0.916 ± 0.011	0.946 ± 0.017
$25 < p_T < 28$ GeV	$0.9 < \eta < 1.2$	0.884 ± 0.012	0.923 ± 0.011	0.957 ± 0.017
$28 < p_T < 31$ GeV	$0.9 < \eta < 1.2$	0.882 ± 0.011	0.927 ± 0.011	0.951 ± 0.016
$31 < p_T < 34$ GeV	$0.9 < \eta < 1.2$	0.889 ± 0.011	0.931 ± 0.011	0.955 ± 0.016
$34 < p_T < 37$ GeV	$0.9 < \eta < 1.2$	0.889 ± 0.010	0.930 ± 0.010	0.956 ± 0.015
$37 < p_T < 40$ GeV	$0.9 < \eta < 1.2$	0.893 ± 0.010	0.932 ± 0.010	0.958 ± 0.015
$p_T > 40$ GeV	$0.9 < \eta < 1.2$	0.895 ± 0.010	0.933 ± 0.010	0.959 ± 0.014
$19 < p_T < 22$ GeV	$ \eta > 1.2$	0.809 ± 0.012	0.851 ± 0.012	0.951 ± 0.020
$22 < p_T < 25$ GeV	$ \eta > 1.2$	0.836 ± 0.012	0.878 ± 0.012	0.952 ± 0.019
$25 < p_T < 28$ GeV	$ \eta > 1.2$	0.858 ± 0.012	0.893 ± 0.012	0.961 ± 0.018
$28 < p_T < 31$ GeV	$ \eta > 1.2$	0.870 ± 0.011	0.900 ± 0.011	0.966 ± 0.017
$31 < p_T < 34$ GeV	$ \eta > 1.2$	0.883 ± 0.011	0.909 ± 0.011	0.972 ± 0.016
$34 < p_T < 37$ GeV	$ \eta > 1.2$	0.886 ± 0.010	0.917 ± 0.010	0.967 ± 0.016
$37 < p_T < 40$ GeV	$ \eta > 1.2$	0.894 ± 0.010	0.921 ± 0.010	0.971 ± 0.015
$p_T > 40$ GeV	$ \eta > 1.2$	0.903 ± 0.010	0.928 ± 0.010	0.973 ± 0.014

The discrepancy between the data and the MC simulations in the efficiency of the c-jets and b-jets to pass the Working Point (WP) of the CSV algorithm [83], as well as the difference in the mistag rate for lighter flavor (u,d,s) quarks and the gluon jets have been corrected by reclassifying jets in the MC events as a function of p_T and η .

6.5.5 Rochester correction

There is a discrepancy between the data and the MC simulations of the Z mass spectrums and p_T spectrum. Rochester corrections help to resolve this discrepancy. The details of these corrections can be found in [84].

6.6 Background Estimation

The dominant background for this analysis is DY production of the muon pair and it is taken from MC. The W+jets, single top quark, and diboson contributions together are considered as an “electroweak” background. The accuracy of the background estimate is improved by determining the contributions of the main backgrounds from the data as well as from backgrounds that are challenging to the model through the MC simulation. In particular, the background from multijet production is derived from the data and it is relevant in this channel. The normalization of the $t\bar{t}$ background is determined from the data, using a control region that contains events with a muon and one or more b-tagged jets. The $t\bar{t}$ normalization factor obtained from this control region is applied to this channel. Last, the negligible background contribution from the Higgs boson with a mass of $m_H = 125$ GeV, produced at the rate and with branching fractions predicted in the SM, is considered. The expected contributions from background processes are summarized in Table 6.8. The uncertainties noted in the table are derived as the quadratic sum of statistical and systematic sources. For signal extraction, these uncertainties are used as inputs to the maximum likelihood (ML) fit.

To prepare for the SM $H \rightarrow \tau\tau$ analysis, the validation of the background

Table 6.8: Expected number of background events in the $\tau_\mu\tau_\mu$ channels in the data corresponding to an integrated luminosity of 2.3 fb^{-1} . The numbers are rounded to a precision of two significant digits on the uncertainty.

Process	$\tau_\mu\tau_\mu$
$Z \rightarrow \mu\mu + \text{jets}$	$7\,650 \pm 300$
$t\bar{t}$	$1\,370 \pm 110$
Multijet	740 ± 140
Electroweak	312 ± 34
SM H	18 ± 4
Total expected background	$10\,100 \pm 390$

estimation is further tested using event categories based on the jet multiplicities, p_T of the τ lepton pairs, and on the multiplicity of b jets in the event. Details on the background estimation using event categories are summarized in Section 6.7.1.

6.6.1 DY Background

The event yield for DY background in this analysis is calculated from the $m_{\tau\tau}$ distribution after applying the BDT > 0.5 , mass cut $m_{\mu\mu} > 80 \text{ GeV}$, and all other required selection criteria as listed in Section 6.4.

6.6.2 Multijet Background

The contributions from the multijet background in the Signal Region (SR) are estimated using control regions containing events with two muons of the same charge. An estimate for the contribution from multijet events in the SR is obtained by scaling the yield of the multijet background in the same sign (SS) control region by a suitably chosen extrapolation factor, defined by the ratio

of $\mu\mu$ pairs with the opposite charge to those with the same charge. For this channel the extrapolation factor determined from the simulation is 2.06 ± 0.33 . The latter is measured in events in which at least one lepton passes an inverted isolation criterion of $I_\ell > 0.15 p_T^\ell$. This event sample is referred to as an isolation sideband region (SB). The requirement $I_\ell > 0.15 p_T^\ell$ ensures that the SB does not overlap with the SR. A complication arises from the fact that the ratio of opposite sign (OS) to SS pairs depends on the isolation criterion applied in the SB and on the lepton kinematics. The nominal OS/SS ratio is measured in an isolation sideband region (SB1) defined by requiring both leptons to satisfy a relaxed isolation criterion of $I_\ell < 0.60 p_T^\ell$, with at least one lepton passing the condition $I_\ell > 0.15 p_T^\ell$. The systematic uncertainty in the OS/SS ratio that arises from the choice of the upper limit on I_ℓ applied in SB1 is estimated by taking the difference between the OS/SS ratio computed in SB1 and the ratio computed in a different isolation sideband region (SB2). The latter is defined by requiring at least one lepton to pass the condition $I_\ell > 0.60 p_T^\ell$, without setting an upper limit on I_ℓ in the SB2 region. The criteria to select events in the isolation sidebands are optimized to ensure high statistics for the measurement of the OS/SS extrapolation factor independent of lepton kinematic variables and at the same to minimize differences in the lepton kinematic distributions between the SR and the SB. In both isolation sidebands, the OS/SS ratio is measured as function of p_T of the two leptons ℓ and ℓ' and of their separation $\Delta R(\ell, \ell') = \sqrt{(\eta_\ell - \eta_{\ell'})^2 + (\phi_\ell - \phi_{\ell'})^2}$ in the η - ϕ plane. The contributions to SB1 and SB2 from backgrounds other than multijet production are subtracted based on the MC simulation.

6.6.3 $t\bar{t}$ Background

The event yield of the $t\bar{t}$ background in the SR is determined from data using the CR dominated by the $t\bar{t}$ background, whereas the $m_{\tau\tau}$ distribution for the $t\bar{t}$ background is obtained from the MC simulation. The ratio of the $t\bar{t}$ event yield measured in data to the MC prediction is applied as a scale factor to simulated $t\bar{t}$ events to correct the $t\bar{t}$ background. The scale factor derived from the yeild observed in the control region is 1.01 ± 0.07 for this channel.

6.7 Systematic Uncertainties

Various imprecisely known or simulated effects can alter the normalization and the invariant mass distribution of the τ lepton pair of the $Z/\gamma^* \rightarrow \tau\tau$ signal and of background processes. The systematic uncertainties are categorized based on what factor affects the $m_{\tau\tau}$ distribution. They are as follows:

- **Normalization Uncertainties:** Affect the number of signal or background events.
- **Shape Uncertainties:** Affect the number of signal or background events in individual bins of the $m_{\tau\tau}$ distribution. Changes in normalization, given by the sum of signal or background events in all bins, are possible.

Most of the systematic uncertainties are represented by normalization uncertainties. An additional uncertainty arises from limited statistics available to model the shape of the $m_{\tau\tau}$ distribution for some of the backgrounds. These systematic and statistical uncertainties are explained as follows:

- **Trigger, identification, and isolation efficiency μ**

The uncertainties on muon trigger, identification, and isolation efficiencies are measured using $Z/\gamma^* \rightarrow \mu\mu$ events with the “tag-and-probe” method with an accuracy of 2% or muons of all p_T and η . The uncertainty in the efficiency of the single-muon trigger is about 2%.

- **Muon Energy Scale μ ES**

The energy scales for muons are calibrated using $J/\psi \rightarrow \mu\mu$, $\Upsilon \rightarrow \mu\mu$, and $Z/\gamma^* \rightarrow \mu$ events and have an uncertainty of 1%. The μ ES uncertainty affects the acceptance for the $Z/\gamma^* \rightarrow \tau\tau$ signal by less than 1%.

- **E_T^{miss} resolution and response**

The E_T^{miss} resolution and response are known within uncertainties of a few percent from studies performed using $Z/\gamma^* \rightarrow \mu\mu$, $Z/\gamma^* \rightarrow ee$, and γ +jets events [85]. The effect on the distribution in $m_{\tau\tau}$ is small. Uncertainties related to the modeling of the E_T^{miss} affect the acceptance for the $Z/\gamma^* \rightarrow \tau\tau$ signal because of the use of E_T^{miss} and P_ζ as input variables in the BDT that separate the $Z/\gamma^* \rightarrow \tau\tau$ signal from the $Z/\gamma^* \rightarrow \mu\mu$ background. The change in signal acceptance is $< 1\%$.

- **Background yields**

The uncertainty on the dominant background contributions is given by the statistical and systematic uncertainties of the data-driven background estimation methods. In the multijet background, the uncertainty is dominated by the OS/SS ratio and its contribution is about 20%. The CR in the $t\bar{t}$ background dominates the uncertainty defined in the normalization. For this channel, uncertainty in the $t\bar{t}$ normalization is 7%. In addition,

the shape uncertainties for this background are estimated by changing the weights applied to the $t\bar{t}$ MC samples to improve the modeling of the top quark p_T distribution, within no reweighting and the reweighting applied twice. The uncertainties in the single top quark and diboson background are estimated by uncertainties in the luminosity and on the cross section of the respective backgrounds. For both backgrounds, the uncertainty is 15%. Similarly, the 15% contribution comes from the yield of the W+jets background. Uncertainties on the background directly taken from [80]. The SM Higgs yield is assigned an uncertainty of 30%. This uncertainty is directly taken from the experimental uncertainty in the overall $H \rightarrow \tau\tau$ rate measured at $\sqrt{s} = 13$ TeV [86].

- **Signal acceptance**

The theoretical uncertainty in the product of signal acceptance and efficiency for the $Z/\gamma^* \rightarrow \tau\tau$ signal is $\approx 2\%$. This uncertainty includes the effect of missing higher order perturbative expansions, estimated through independent changes in the renormalization and factorization scales by factors of two and an half relative to their nominal equal values [87, 88], uncertainties in the NNPDF3.0 set of parton distribution functions (PDFs), estimated following the recommendations given in [89], and the uncertainties in the modeling of parton showers (PS) and the underlying event (UE).

- **Luminosity**

The uncertainty on the integrated luminosity amounts to 2.6% [90].

Table 6.9 [91] summarizes the systematic uncertainties. It also quantifies the

percent change in cross section when these uncertainties are fluctuated in the range of one standard deviation relative to their nominal values. The ML fit is used for the signal extraction. The impacts are calculated for the nuisance parameters obtained from this fit. The uncertainty in the integrated luminosity, in the cross section of the DY production of a muon pair has an impact of about 1.5%. The sizable impact of the systematic uncertainty in the production rate of $Z/\gamma^* \rightarrow ee$ and $Z/\gamma^* \rightarrow \mu\mu$ background processes is caused by the small statistical uncertainty in the $Z/\gamma^* \rightarrow \mu\mu$ background.

Table 6.9: Effect of experimental and theoretical uncertainties in the measurement of the $Z/\gamma^* \rightarrow \tau\tau$ cross section.

Source	Applies to	Change in \mathcal{A} or yield	Impact
Integrated luminosity	Simulation	2.3%	1.9%
Muon ID and trigger	Simulation	2%	1.6%
μ ES	Simulation	< 1%	< 0.1%
E_T^{miss} response and resolution	Simulation	1 – 10%	0.2%
Norm. $Z/\gamma^* \rightarrow \mu\mu$ ²	$Z/\gamma^* \rightarrow \mu\mu$ ²	unconstrained	1.8%
Norm. and distr. of multijet	Multijet	20%	0.2%
Norm. $t\bar{t}$	$t\bar{t}$	7%	1.0%
Distr. $t\bar{t}$	$t\bar{t}$	1 – 6%	< 0.1%
Norm. SM H	SM H	30%	< 0.1%
Norm. single top quark	Single top quark	15%	< 0.1%
Norm. diboson	Diboson	15%	0.2%
Norm. W+jets	W+jets	15%	< 0.1%
PDF	Signal	1%	1.0%
Scale dependence	Signal	< 6%	0.5%
UE and PS	Signal	1%	1.0%

6.7.1 Event categorization

The validity of the background estimation described in Section 6.6 is checked in event categories that are relevant for the SM $H \rightarrow \tau\tau$ analysis as well as in searches for new physics.

Event categories based on jet multiplicity, p_T of the τ lepton pair, and on the multiplicity of b jets are defined by the conditions given in Table 6.10.

Table 6.10: Event categories used to study the modeling of backgrounds. Similar categories have been used in previous $H \rightarrow \tau\tau$ analyses at the LHC.

Category	Selection
0-jet	No jets ¹ and no b jets ²
1-jet, low Z boson p_T	At least one jet ¹ , no b jets ² , $p_T^Z < 50$ GeV, and event not selected in 2-jet VBF category
1-jet, medium Z boson p_T	At least one jet ¹ , no b jets ² , $50 < p_T^Z < 100$ GeV, and not selected as 2-jet VBF
1-jet, high Z boson p_T	At least one jet ¹ , no b jets ² , $p_T^Z > 100$ GeV, and not selected as 2-jet VBF
2-jet VBF	At least one pair of jets ^{1,3} , no b jets ²
1 b jet	Exactly one b jet ²
2 b jet	Exactly two b jets ²

¹ With $p_T > 30$ GeV and $|\eta| < 4.7$

² With $p_T > 20$ GeV, $|\eta| < 2.4$, and identified by the CSV algorithm as originating from the hadronization of b quarks

³ Satisfying $m_{jj} > 500$ GeV and $\Delta\eta_{jj} > 3.5$

The transverse momentum of the Z boson (p_T^Z) is reconstructed by adding the momentum vectors from the visible τ decay products and the reconstructed \vec{p}_T^{miss} in the transverse plane. The observables m_{jj} and $\Delta\eta_{jj}$ are used to select signal events produced through the fusion of virtual vector bosons (VBF) in the SM $H \rightarrow \tau\tau$ analysis, and refer, respectively, to the mass and to the separation

in pseudorapidity of the two jets of highest p_T in events containing two or more jets.

The contributions of background processes that are modeled in the MC simulation to the different categories are affected by uncertainties in the jet energy scale and resolution. The energy scale of jets is measured using the p_T balance of jets with Z bosons and photons in $Z/\gamma^* \rightarrow ee$ and $Z/\gamma^* \rightarrow \mu\mu$ and γ +jets events and the p_T balance between jets in dijet events as described in [92]. The uncertainty in the jet energy scale is a few percent and depends on p_T and η . The impact of jet energy scale and resolution uncertainties on the yields of background processes is evaluated by varying the jet energy scale and resolution within their uncertainties, redetermining the multiplicity of jets and b jets, and reapplying the event categorization conditions given in Table 6.10. The distribution of the $m_{\tau\tau}$ mass in given event categories are shown in Figures 6.4 and 6.5.

6.8 Cross Section Measurements

The cross section times branching fraction of the $Z/\gamma^* \rightarrow \tau\tau$ process is obtained using a simultaneous ML fit to the $m_{\tau\tau}$ distribution in all five decay channels, namely, $\tau_e\tau_h$, $\tau_\mu\tau_h$, $\tau_h\tau_h$, $\tau_e\tau_\mu$, and $\tau_\mu\tau_\mu$ and can be expressed as follows:

$$\sigma(\text{pp} \rightarrow Z/\gamma^* + \text{jets}) \times \mathcal{B}(Z/\gamma^* \rightarrow \tau\tau) = \frac{1}{\mathcal{B}_\tau} \frac{N_{\text{sig}}^{\text{fit}} (1 - f_{\text{out}})}{\mathcal{A} \varepsilon \mathcal{L}}, \quad (6.6)$$

where the symbol \mathcal{B}_τ represents the branching fraction for the τ pair to decay into one of the final states mentioned above. The $N_{\text{sig}}^{\text{fit}}$ is the number of $Z/\gamma^* \rightarrow \tau\tau$ signal events passing the selection criteria as described in Section 6.4.

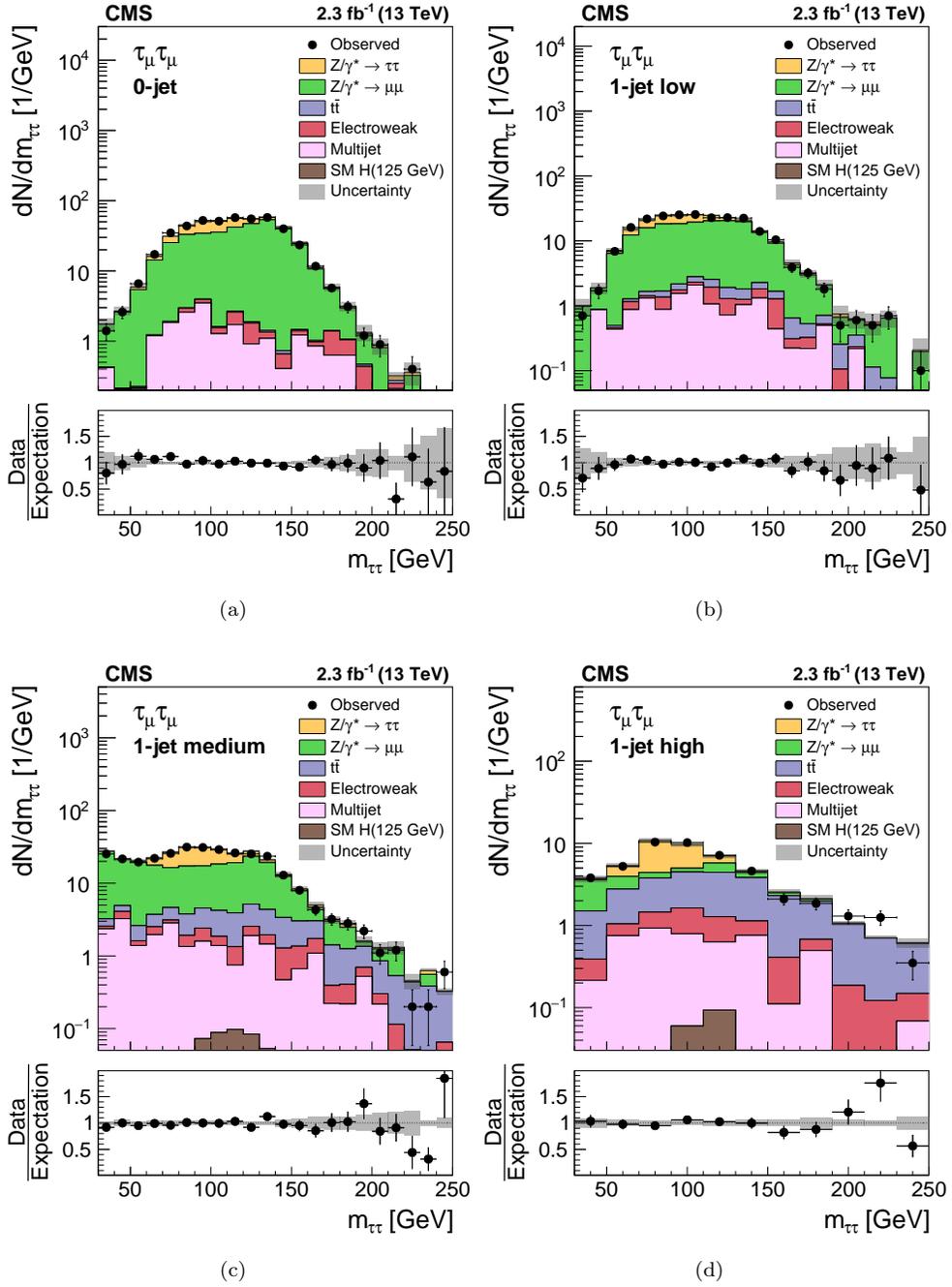
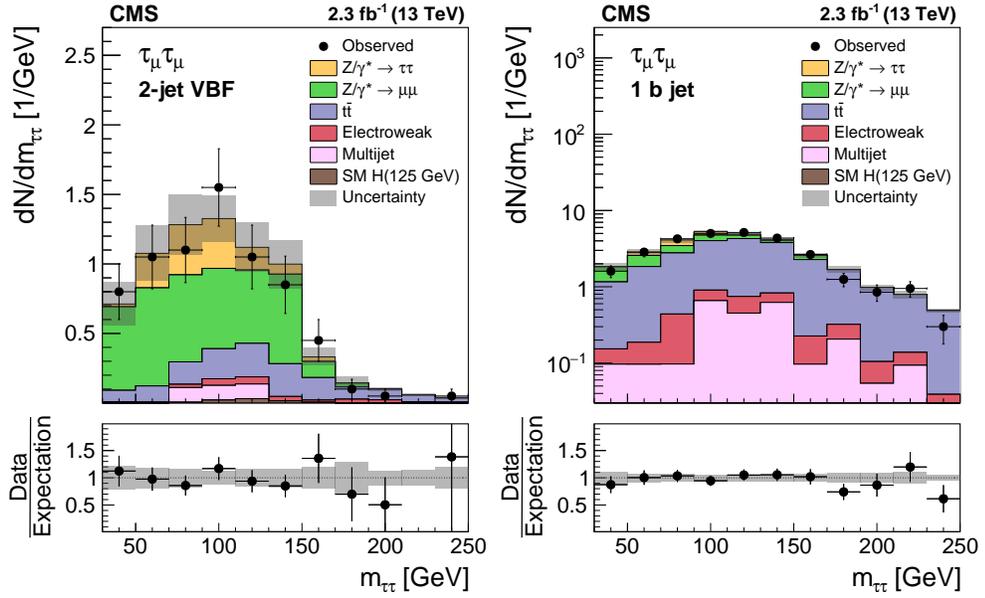
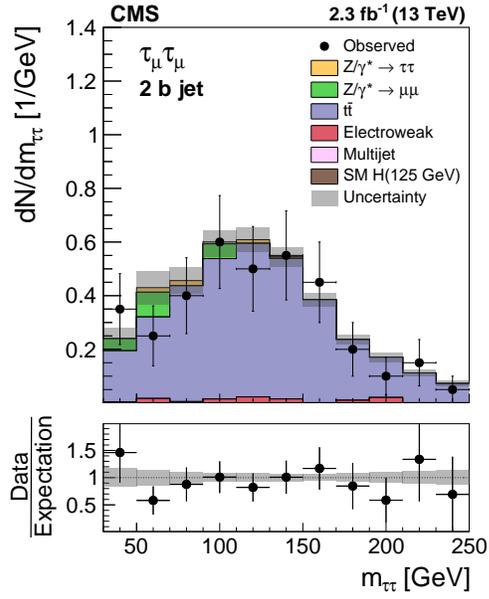


Figure 6.4: Distributions of $m_{\tau\tau}$ in different jet categories: (a) 0-jet, (b) 1-jet low, (c) medium, and (d) high Z boson p_T .



(a)

(b)



(c)

Figure 6.5: Distributions of $m_{\tau\tau}$ in different jet categories: (a) 2-jet VBF, (b) 1 b jet, and (c) 2 b jet.

Table 6.11: Branching fraction \mathcal{B}_τ , signal acceptance \mathcal{A} , selection efficiency ε , and mass window correction factor f_{out} for $\tau_\mu\tau_\mu$ final state.

	$\tau_\mu\tau_\mu$
Branching fraction \mathcal{B}_τ	0.0318
Acceptance \mathcal{A}	0.1111
Selection efficiency ε	0.1434
Mass window correction f_{out}	0.0465

The fraction of $Z/\gamma^* \rightarrow \tau\tau$ signal events with $p_T > 24$ GeV and $|\eta| < 2.1$ cuts on the generator level quantities give the signal acceptance \mathcal{A} for the $\tau_\mu\tau_\mu$ channel. ε is the selection efficiency, which is defined as the number of events that simultaneously pass the acceptance cut as mentioned above as well as the selection criteria as mentioned in Section 6.4. Finally, \mathcal{L} is the integrated luminosity of the analyzed dataset. The signal acceptance \mathcal{A} and selected efficiency are obtained using the MC simulations and by considering only events that are generated within the mass range $60 < m_{\tau\tau}^{\text{gen}} < 120$ GeV. The factor f_{out} is also obtained from MC simulations and it is a correction factor that is applied to the events generated outside the mass range $60 < m_{\tau\tau}^{\text{gen}} < 120$ GeV and also passes the selection criteria as mentioned above. The branching fractions \mathcal{B}_τ for the different τ pair decay channels are directly taken from [93]. The values of \mathcal{B}_τ , \mathcal{A} , ε , and f_{out} for the $\tau_\mu\tau_\mu$ channel are given in Table 6.11.

The systematic uncertainties described in the earlier Section 6.7 can affect all the quantities mentioned in Equation 6.6. The likelihood function described in Section 6.3 depends on the cross section times the branching fraction and uses it as the parameter of interest (POI) ξ in the fit. In this fit, the systematic uncertainties derived in the earlier section are treated as the nuisance parameters

θ_k . Hence, in this case, the likelihood function can be written as

$$\mathcal{L}(\text{data} | \xi, \Theta) = \prod_i \mathcal{P}(n_i | \xi, \Theta) \prod_k \rho(\tilde{\theta}_k | \theta_k). \quad (6.7)$$

where index i denotes each separate bin of the $m_{\tau\tau}$ distribution in each of the five final states. The set of all nuisance parameters θ_k is denoted by the symbol Θ .

The best fit value $\hat{\xi}$ of the POI is the value that maximizes the likelihood $\mathcal{L}(\text{data} | \xi, \Theta)$ in Equation 6.7. A 68% confidence interval (CI) on the POI is obtained using the profile likelihood ratio (PLR) [67, 94, 95]:

$$\lambda(\xi) = \frac{\mathcal{L}(\text{data} | \xi, \hat{\Theta}_\xi)}{\mathcal{L}(\text{data} | \hat{\xi}, \hat{\Theta})}. \quad (6.8)$$

The symbol $\hat{\Theta}_\xi$ denotes the values of nuisance parameters that maximize the likelihood for a given value of ξ . The combination of $\hat{\xi}$ and $\hat{\Theta}$ correspond to the values of ξ and Θ for which the likelihood function reaches its maximum. The 68% CI is defined by the values of ξ for which $-2 \ln \lambda(\xi)$ increases by one unit relative to its minimum. To quantify the individual effects of statistical uncertainties, the uncertainty in the integrated luminosity of the analyzed data, and other systematic uncertainties, we ignore some single sources of uncertainties at a time and recompute the 68% CI. The nuisance parameters θ_k , corresponding to uncertainties that are ignored, are fixed at the values $\hat{\theta}_k$ that yield the best fit to the data. The square root of the quadratic difference between the CI, computed for all sources of uncertainties in the fit, and for the case that some given source is ignored, reflects the estimate of the uncertainty in the POI re-

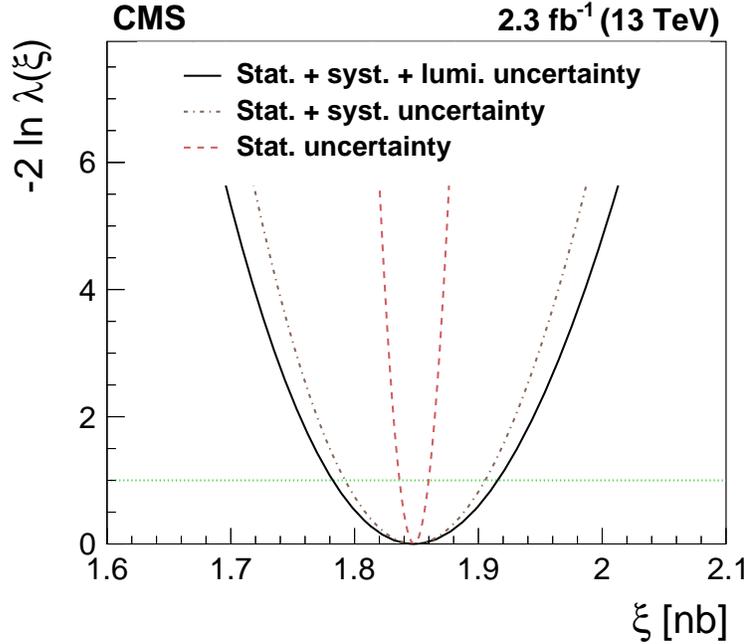


Figure 6.6: Dependence of the profile likelihood ratio $-2 \ln \lambda(\xi)$ on the cross section ξ for DY production of τ pairs. The PLR is computed for the simultaneous ML fit to the observed $m_{\tau\tau}$ distributions in the $\tau_e\tau_h$, $\tau_\mu\tau_h$, $\tau_h\tau_h$, $\tau_e\tau_\mu$, and $\tau_\mu\tau_\mu$ channels. The three curves correspond to the case that statistical uncertainties, the uncertainty in the integrated luminosity, and other systematic uncertainties are included in the fit and when the nuisance parameters corresponding to the integrated luminosity and to other systematic uncertainties are successively fixed in the ML fit. The horizontal line represents the value of $-2 \ln \lambda(\xi)$ that is used to determine the 68% CI on ξ .

sulting from a single source. The procedure is illustrated in Figure 6.6 [91] for the combined fit of all five final states. Correlations among different sources of uncertainty are estimated through this procedure.

6.9 Results and Summary

Figure 6.7 [91] shows the post ML fit $m_{\tau\tau}$ distribution for the $\tau_\mu\tau_\mu$ channel. The post ML fit yields in the $Z/\gamma^* \rightarrow \tau\tau$ signal and in background contributions

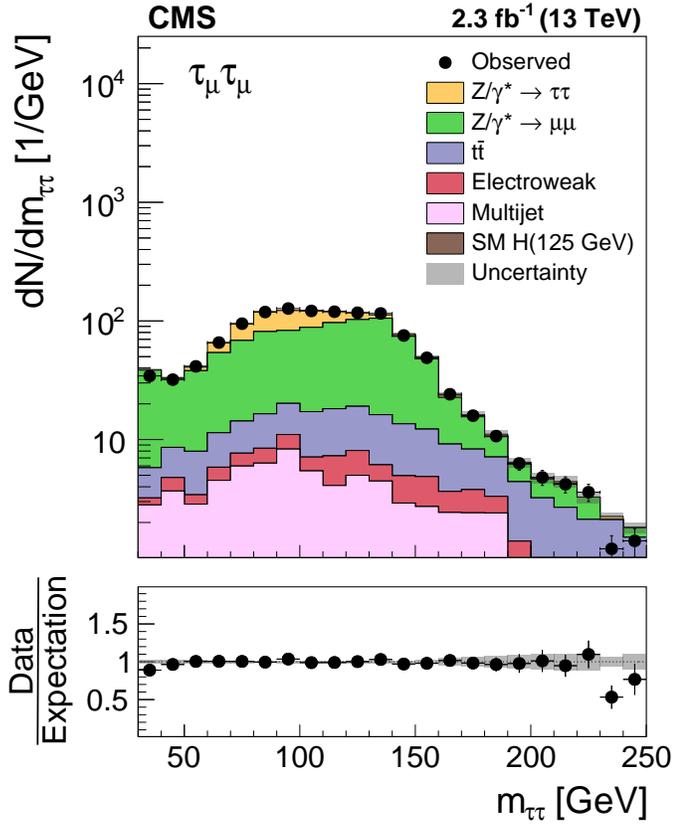


Figure 6.7: Distributions of $m_{\tau\tau}$ in inclusive events selected in the $\tau_\mu\tau_\mu$ channel. Signal and background contributions are shown for the values of nuisance parameters obtained from the ML fit to the data.

are summarized in Table 6.12.

The observed cross section for this channel is given in equation 6.9.

$$\sigma(\text{pp} \rightarrow Z/\gamma^* + X) \times \mathcal{B}(Z/\gamma^* \rightarrow \tau\tau) = 1967 \pm 121 \text{ (stat.)} \pm 92 \text{ (syst.)} \pm 37 \text{ (lumi.) pb.} \quad (6.9)$$

The total uncertainty in the cross section is derived from the statistical uncertainties, uncertainty in the integrated luminosity of the analyzed dataset, and the systematic uncertainties described in Section 6.7.

Table 6.12: Yields in $Z/\gamma^* \rightarrow \tau\tau$ signal events and backgrounds in the $\tau_\mu\tau_\mu$ channel, obtained from the ML fit. The analyzed data corresponds to an integrated luminosity of 2.3 fb^{-1} .

Process	$\tau_\mu\tau_\mu$
$Z/\gamma^* \rightarrow \tau\tau$	2067 ± 34
$Z/\gamma^* \rightarrow \mu\mu$	8010 ± 170
$t\bar{t}$	1239 ± 79
Multijet	710 ± 110
Electroweak	293 ± 30
SM H	18 ± 4
Total expected background	10270 ± 120
Total SM expectation	12340 ± 120
Observed data	12327

Finally, all five final states $\tau_e\tau_h$, $\tau_\mu\tau_h$, $\tau_h\tau_h$, $\tau_e\tau_\mu$, and $\tau_\mu\tau_\mu$ of the $Z/\gamma^* \rightarrow \tau\tau$ decay are combined and simultaneously fitted to obtain the cross section:

$$\sigma(\text{pp} \rightarrow Z/\gamma^* + X) \times \mathcal{B}(Z/\gamma^* \rightarrow \tau\tau) = 1848 \pm 12 \text{ (stat.)} \pm 57 \text{ (syst.)} \pm 35 \text{ (lumi.) pb.} \quad (6.10)$$

The results are compatible with the theory prediction as well as with the cross section measured by CMS for DY production and decay into electron and muon pairs. The theory prediction is computed at NNLO accuracy [96] by using, NNPDF3.0 PDF set and it is predicted as 1845_{-6}^{+12} (scale) ± 33 (PDF) pb. These results are summarized in Figure 6.8 [91]. The inner black error bars represent the statistical uncertainties and the outer blue error bars illustrate the quadratic sum of the statistical, systematic, and integrated luminosity uncertainties.

In conclusion, the cross section of the $Z/\gamma^* \rightarrow \tau\tau$ process is measured

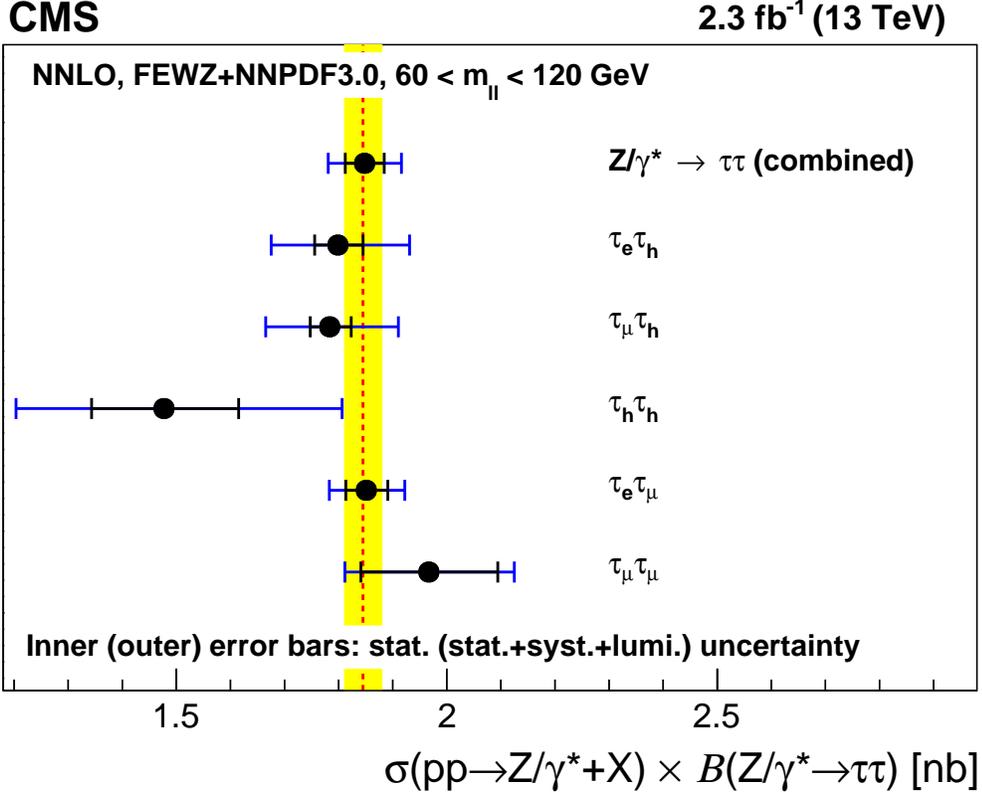


Figure 6.8: Cross section $\sigma(pp \rightarrow Z/\gamma^* + X) \times \mathcal{B}(Z/\gamma^* \rightarrow \tau\tau)$ measured in individual channels and in the combination of all final states, compared to the theoretical prediction and to the cross section measured by CMS for DY production of electron and muon pairs.

for the $\tau_\mu\tau_\mu$ channel using pp collisions recorded by the CMS experiment at $\sqrt{s} = 13$ TeV at the LHC with an integrated luminosity of 2.3 fb^{-1} . The $m_{\tau\tau}$ distribution is used to determine the signal yield. The measured cross section for the $\tau_\mu\tau_\mu$ channel is 1967 ± 121 (stat.) ± 92 (syst.) ± 37 (lumi.) pb, whereas the total cross section using all five decay channels: $\tau_e\tau_h$, $\tau_\mu\tau_h$, $\tau_h\tau_h$, $\tau_\mu\tau_\mu$, and $\tau_e\tau_\mu$ is measured as $\sigma(pp \rightarrow Z/\gamma^* + X) \times \mathcal{B}(Z/\gamma^* \rightarrow \tau\tau) = 1848 \pm 12$ (stat.) ± 57 (syst.) ± 35 (lumi.) pb. The measured cross section is in agreement with the standard model expectation computed at next-to-next-to-leading order (NNLO) accuracy.

Chapter 7

Search for Neutral Higgs Boson

Decays in the

$H \rightarrow \tau\tau \rightarrow \mu\mu$ Channel

7.1 Introduction

In the SM, the existence of a neutral scalar particle, namely the Higgs Boson (H), was predicted by the Brout-Englert-Higgs mechanism through electroweak symmetry breaking, as explained in detail in Chapter 4. During Run I of the LHC, both the ATLAS and CMS experiments observed this boson in ZZ , $\gamma\gamma$, and WW decay channels. The pp collision data was recorded in 2011 and 2012 at center-of-mass energies $\sqrt{s} = 7$ and 8 TeV, respectively. Combined ATLAS and CMS results can be found in [97, 98, 99]. The mass of the Higgs Boson is measured with precision to be $125.09 \pm 0.21(\text{stat.}) \pm 0.11(\text{syst.})$ GeV, from a combination of ATLAS and CMS measurements [100]. Fermions gain mass

due to the Yukawa coupling of the Higgs field to fermions. Therefore, fermionic decays of the Higgs Boson play an important role in understanding the properties of this newly discovered neutral Higgs Boson. The $H \rightarrow \tau^+\tau^-$ decay is the most accessible leptonic decay due to a higher event rate compared with other leptonic decays, e.g. $\mu^+\mu^-$, and a comparatively smaller background contribution than the $b\bar{b}$ decay channel. The SM branching ratio of $\mathcal{B}(H \rightarrow \tau^+\tau^-)$ is 6.3% for a mass of 125.09 GeV. The tau leptons decay further hadronically and leptonically. This chapter focuses on one of the leptonic decays in which $H \rightarrow \tau^+\tau^-$ further decays into a pair of muons. Henceforth, this particular decay channel will be referred as the $\tau_\mu\tau_\mu$ channel. The $\tau_\mu\tau_\mu$ channel has a very low branching fraction ($\sim 3\%$ of all $H \rightarrow \tau\tau$ decays) and a very high DY background. These challenges and low sensitivity of the channel make this analysis difficult.

The CMS experiments showed evidence for $H \rightarrow \tau\tau$ coupling at 3.2σ [101] in Run I of the LHC, while ATLAS and CMS observed the coupling at 5.5σ [102] and in Run II with an integrated luminosity of 35.9 fb^{-1} at center-of-mass energy $\sqrt{s} = 13 \text{ TeV}$ $H \rightarrow \tau\tau$ signal is established with a significance of 4.9 standard deviation [103]. In this chapter, the detailed measurement of the $H \rightarrow \tau\tau$ coupling in the $\tau_\mu\tau_\mu$ channel is discussed using Run II data taken in pp collisions in 2016 at the center-of-mass energy $\sqrt{s} = 13 \text{ TeV}$ with an integrated luminosity of 35.9 fb^{-1} . This analysis, considers Higgs production via the gluon-gluon fusion (ggH) and the vector boson fusion (VBF) production mechanisms.

7.2 Simulated Samples and Run II Datasets

In this analysis, the SingleMuon trigger dataset is used. Only collision runs for which entire detector system is working are considered. These datasets are certified by the CMS collaboration in the JSON selection file:

`Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16_JSON.json`.

The collision datasets with corresponding run ranges and integrated luminosities are summarized in Table 7.1.

Table 7.1: List of datasets included in the analysis.

Dataset	Run range	Integrated Luminosity
/SingleMuon/Run2016B-03Feb2017-ver2-v2/MINIAOD	272007–275376	5.788 /fb
/SingleMuon/Run2016C-03Feb2017-v1/MINIAOD	275657–276283	2.573 /fb
/SingleMuon/Run2016D-03Feb2017-v1/MINIAOD	276315–276811	4.248 /fb
/SingleMuon/Run2016E-03Feb2017-v1/MINIAOD	276831–277420	4.009 /fb
/SingleMuon/Run2016F-03Feb2017-v1/MINIAOD	277772–278808	3.102 /fb
/SingleMuon/Run2016G-03Feb2017-v1/MINIAOD	278820–280385	7.540 /fb
/SingleMuon/Run2016H-03Feb2017-ver2-v1/MINIAOD	280919–284044	8.606 /fb
/SingleMuon/Run2016H-03Feb2017-ver3-v1/MINIAOD	280919–284044	see above

Table 7.2: List of Monte Carlo signal samples included in the analysis.

Signal MC sample	Production cross section	$H \rightarrow \tau\tau$ branching fraction
/GluGluHToTauTau_M110_13TeV_powheg_pythia8	57.90 pb	0.0791
/GluGluHToTauTau_M120_13TeV_powheg_pythia8	52.22 pb	0.0698
/GluGluHToTauTau_M125_13TeV_powheg_pythia8	48.58 pb	0.0627
/GluGluHToTauTau_M130_13TeV_powheg_pythia8	45.31 pb	0.0541
/GluGluHToTauTau_M140_13TeV_powheg_pythia8	36.00 pb	0.0360
/VBFHToTauTau_M110_13TeV_powheg_pythia8	4.434 pb	0.0791
/VBFHToTauTau_M120_13TeV_powheg_pythia8	3.935 pb	0.0698
/VBFHToTauTau_M125_13TeV_powheg_pythia8	3.782 pb	0.0627
/VBFHToTauTau_M130_13TeV_powheg_pythia8	3.637 pb	0.0541
/VBFHToTauTau_M140_13TeV_powheg_pythia8	3.492 pb	0.0360

Signal and background processes are modeled using simulated MC events centrally by CMS Higgs group. The Higgs signal samples of the ggH and VBF production processes are generated at next-to-leading order (NLO) in perturbative quantum chromodynamics (pQCD) using the POWHEG 2.0 [104, 105, 106, 107, 108] generator. The production cross sections and branching ratios for the ggH and VBF signal processes, as well as the uncertainties related to them, are taken directly from [109, 110, 111]. The signal samples used in this analysis are given in Table 7.2.

Table 7.3: MC background samples generated for pp collisions at a center-of-mass energy of 13 TeV. Samples used in this analysis are reconstructed and stored in miniAOD format. A k-factor of 1.16 is considered for the Z+jets samples and 1.21 for the W+jets samples. All of these MC samples belong to the central CMS Summer16 production, with Moriond Premix conditions. When available, all sample extensions are used.

Background MC simulations	Cross section [pb]
/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	4954.0 (LO)
/DY1JetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1012.5 (LO)
/DY2JetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	332.8 (LO)
/DY3JetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	101.8 (LO)
/DY4JetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	54.8 (LO)
/DYJetsToLL_M-150_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6.657 (LO)
/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	18610 (LO)
/TT_TuneCUETP8M1_13TeV-powheg-pythia8	831.76 pb
/WJetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	50380 (LO)
/W1JetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	9644.5 (LO)
/W2JetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	3144.5 (LO)
/W3JetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	954.8 (LO)
/W4JetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/	485.6 (LO)
/ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.6
/ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.6
/ST_t-channel_antitop_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1	80.95
/ST_t-channel_top_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1	136.02
/WZTo1L3Nu_13TeV_amcatnloFXFX_madspin_pythia8	3.05
/WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	10.71
/WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.595
/WWTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	1.212
/VVTo2L2Nu_13TeV_amcatnloFXFX_madspin_pythia8	11.95
/ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.22
/ZZTo4L_13TeV-amcatnloFXFX-pythia8	1.212

The Z + jets and W + jets processes are simulated at leading order (LO) with MLM jet matching and merging [112] with the MG5_aMC@NLO generator. The diboson samples are produced using the same generator at NLO with FxFx jet matching and merging [113]. The $t\bar{t}$ and single top quark production events are simulated with the POWHEG 2.0 and 1.0 generators, respectively. The decay of the τ leptons and the parton showering and fragmentations are modeled by interfacing respective generators with PYTHIA 8.212 [114] with setting CUETP8M1 tune [115]. The MC background simulation processes and their respective cross sections used in this analysis are summarized in Table 7.3.

The simulated MC samples include the pile-up effect observed in the pp collisions; this effect is generated using the PYTHIA generator. The number of pileup interactions generated in the MC matches the number of interactions in data that are measured from the instantaneous luminosity for each bunch crossing. In Run II, approximately 27 interactions on average occurred per bunch crossing. All the generated events are processed through the CMS detector simulation based on GEANT 4 [116].

7.3 Event Weights and Data/MC Corrections

The PF algorithm is used for reconstruction of the observed and simulated events. Reconstruction of the PF objects, such as charged hadrons, neutral hadrons, photons, and muons, as well as the higher-level objects like E_T^{miss} and jets are explained in detail in Chapter 5. For better agreement between the data and MC events, various event corrections and weights are applied. These corrections are explained in detail in the following sections.

Generated Events

The generator event weights are applied to MC simulated samples on an event-by-event basis. These weights are scaled to the number of expected yield for each simulated signal and background sample, using its cross-section and the integrated luminosity of the corresponding data. The AMC@NLO generator produces both positive and negative weights. Negative weights can reduce the effective statistics for the respective sample; this case is observed for diboson samples.

Pile-up Reweighting

In Run II, CMS uses the luminosity-based pile-up (PU) estimate, where the PU interactions are measured from the luminosity for each bunch crossing. The PU reweighting is applied to the MC simulated events in order to match the number of PU interactions in the data. A minimum-bias cross section of 69.2 mb is used with the number of primary vertices ranging from 0 to 80 with 800 bins.

Lepton ID/Isolation Scale Factors

The lepton ID and isolation efficiencies in this channel are measured using the $Z \rightarrow \mu\mu$ events, in bins of muon p_T and $|\eta|$. The “tag and probe” technique is used to derive the efficiency ratio $\epsilon_{data}/\epsilon_{MC}$. This ratio is applied as a scale factor to correct the simulated events on an event-by-event basis. For probe muons, the p_T binning, $p_T = [10, 15, 20, 25, 30, 40, 60, >60]$ GeV is used, while the binning in η is taken as $|\eta| = [0.0, 0.9, 1.2, 2.1, 2.4]$. The ID/isolation criteria for muons are given in Section 5.1.1.

Trigger Efficiency

In this analysis, the SingleMuon trigger dataset with HLT_IsoMu22 trigger is used. The efficiency is again calculated using the “tag and probe method”. The procedure is similar to the one used for the ID/isolation scale factor, with an additional requirement that both the “tag and probe” muons are required to pass the ID/isolation criteria. The MC samples used in this analysis do not reflect the trigger requirements and hence, the trigger efficiencies in MC samples are considered equal to one. Therefore, the trigger efficiencies only in the data are used as scale factors.

Reweighting of LO Madgraph DY

The generator-level Z p_T , dimuon mass $m_{\mu\mu}$, and dimuon- η distributions show disagreement in the data/MC simulations due to mismodeling of these variables in the LO DY madgraph samples. Weights are extracted from the dimuon region using the three-dimensional distribution created by these variables. In this particular analysis, these weights are derived separately for each event category, i.e. 0-jet, boosted, and VBF, as defined in Section 7.4. The event selection criteria used here are described in Section 7.5. The weights are computed in such a manner that there is good agreement in the three-dimensional distributions of the Z p_T , dimuon mass $m_{\mu\mu}$, and dimuon- η in the data and MC samples. This reweighting affects only the DY shape distribution, but not the normalization and hence, only shape uncertainties are introduced for each event category for this reweighting.

Recoil Corrections

The mismodeling of $E_T^{\vec{miss}}$ in the simulated samples for single boson production such as Z+Jets, W+Jets, and Higgs production are rectified by applying recoil corrections. These corrections are applied on the vectorial difference between the measured $E_T^{\vec{miss}}$ and the total transverse momentum of the neutrino system originated from these boson decays. In the case of leptonic decays, the variable can be given as:

$$\vec{U} = -p_{T,B}^{\vec{}} - \vec{H}_T, \quad (7.1)$$

where $p_{T,B}^{\vec{}}$ is the leptonic recoil, i.e. the transverse momentum of the leptonically decaying Z, W, or Higgs Boson, and \vec{H}_T is transverse momentum of the hadronic recoil. The projection of the hadronic recoil onto the transverse momentum of

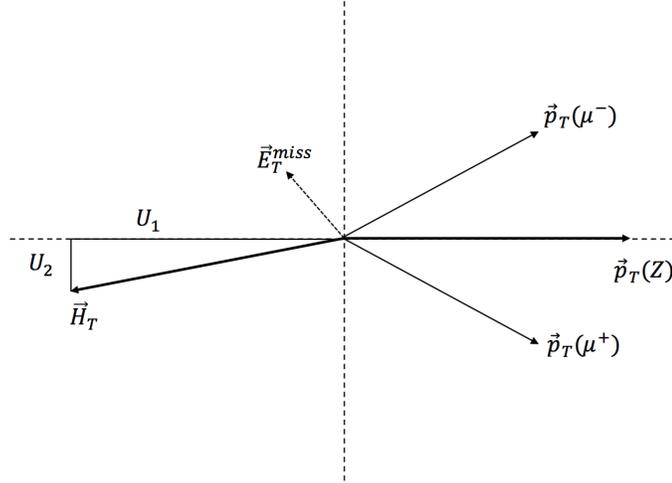


Figure 7.1: Illustration of hadronic recoil.

the boson is represented by \vec{U} as shown in Figure 7.1.

The recoil correction is measured using $Z \rightarrow \mu\mu$ events, as there are no neutrinos in leptonic recoil and the four-momentum of the Z boson can be calculated

precisely. Events are selected as described later in Section 7.5. In addition, all data/MC corrections discussed previously are applied to the selected events. The mean value of the U_1 projection and the resolution of the $U_{1,2}$ projections are calculated in both the $Z \rightarrow \mu\mu$ data and MC events. Corrections to $U_{1,2}$ variables are calculated as a function of $Z p_T$ and jet-multiplicity, given by:

$$U'_1 = \langle U_1 \rangle_{data} + (U_1 - \langle U_1 \rangle_{MC}) \frac{\sigma(U_1)_{data}}{\sigma(U_1)_{MC}} \quad (7.2)$$

and

$$U'_2 = U_2 \frac{\sigma(U_2)_{data}}{\sigma(U_2)_{MC}}. \quad (7.3)$$

The mean value of U_1 and the resolution of both the parallel and perpendicular components of the recoil projection are shown in Figures 7.2–7.4 [117]. In this analysis, these recoil corrections are applied to the DY, W+Jets, and Higgs production MC samples on the event by event basis. The effect of the recoil correction is shown in Figure 7.5.

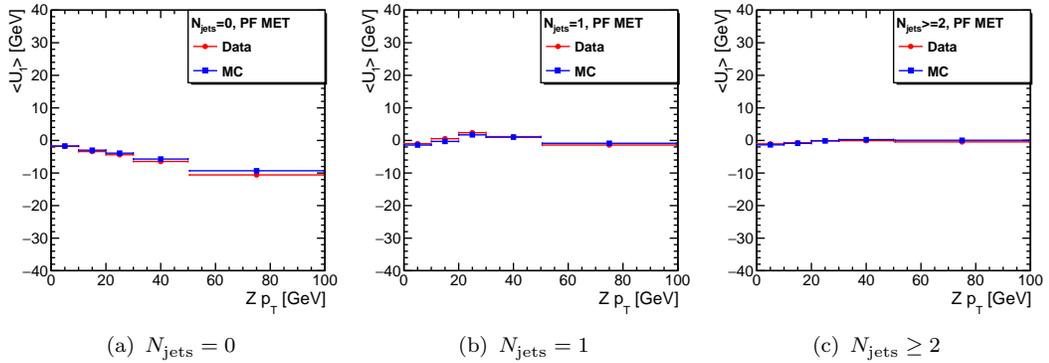


Figure 7.2: Mean value of U_1 in the data and simulation as a function of $Z p_T$ and jet multiplicity.

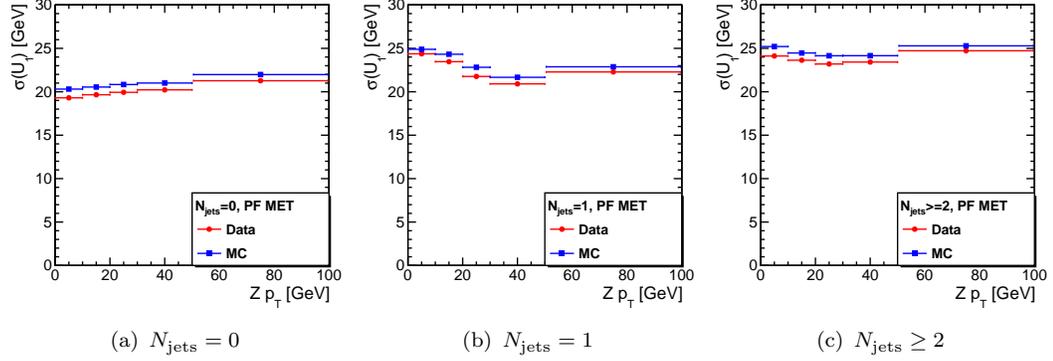


Figure 7.3: Resolution of the U_1 projection in the data and simulation as a function of $Z p_T$ and jet multiplicity.

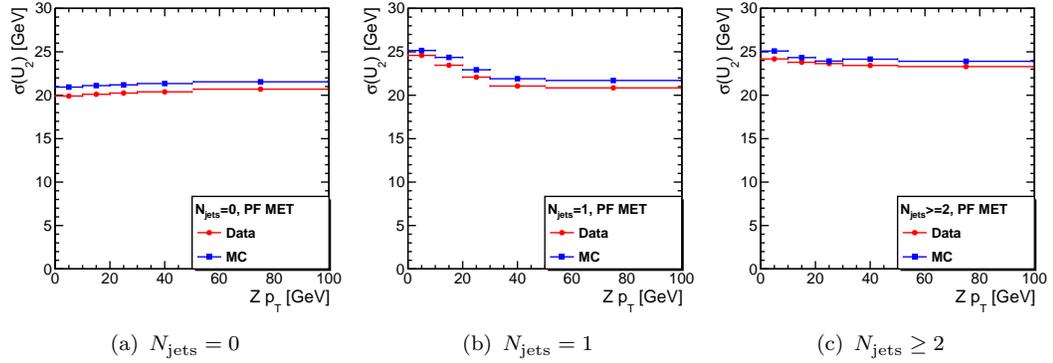


Figure 7.4: Resolution of U_2 in the data and simulation as a function of $Z p_T$ and jet multiplicity.

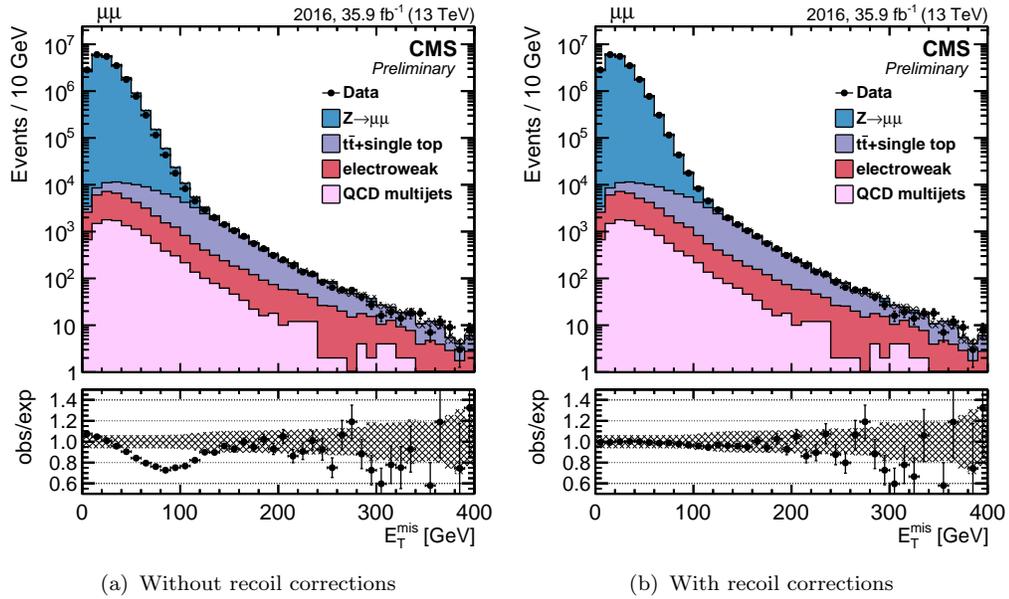


Figure 7.5: Effect of applying recoil corrections to the E_T^{miss} distribution in the $Z \rightarrow \mu\mu$ selection. The recoil corrections improve the agreement between the data and MC samples.

Top p_T Reweighting

The top p_T reweighting is applied only to the $t\bar{t}$ events, as the top p_T distribution in the data is softer than in MC events. This mismodeling is rectified by applying the reweighting. The uncertainties associated with the reweighting are contained within the parameter itself.

7.4 Event Categorization

Final limits on the signal cross section times branching ratio are derived in three exclusive event categories based on the jet multiplicity. The signal is extracted independently for each category based on the fitting of a two-dimensional mass distribution. These three jet categories are 0-jet, Boosted, and VBF.

- **0-jet:** This category classifies the events with zero jet multiplicity and focuses on the Higgs production via gluon fusion. Events with jets with $p_T > 30$ GeV and $|\eta| < 4.7$ and reconstructed with loose PF identification criteria are excluded. Though this Higgs production has the largest cross section rate at the LHC, the measurements suffer due to the very large background contribution, especially the irreducible DY background. Measurements in this category are useful in constraining the backgrounds with large statistics.
- **Boosted:** This category contains events where Higgs events are produced via boosted gluon fusion with recoiling one or more jets. VBF Higgs events with one jet or with low di-jet mass (< 300 GeV) are also included.
- **VBF:** Lastly, the VBF category is defined VBF Higgs production with two jets. Again, these two jets are required to pass the threshold of $p_T > 30$ GeV and $|\eta| < 4.7$ with loose PF identification. These events are also required to have di-jet mass greater than 300 GeV.

Henceforth, the analysis results are estimated independently for these three categories and finally, results for all three categories are combined statistically to extract the signal strength and limits on the SM Higgs cross section.

7.5 Event Selections

In the $\tau_\mu\tau_\mu$ channel, the events are selected based on the trigger requirement and the offline baseline selections. This channel uses the single-muon HLT triggers in the Run II analysis. The events are recorded using the HLT_IsoMu24_v* single muon trigger during the 2016 Run (B-H) era. The selected muon must match the HLT muon within $\Delta R < 0.5$.

The offline event selection requirements for the $\tau_\mu\tau_\mu$ channels are as follows:

- The event should pass the single-muon trigger for the data.
- The two muons should have opposite charge and an invariant mass of $m_{\mu\mu} > 20$ GeV.
- Muons with $p_T > 10$ GeV and $|\eta| < 2.4$ must pass the 2016 (ICHEP) medium PF muon identification criteria for the Run B-F single muon datasets, and the standard medium PF muon ID criteria for the Run G-H datasets and the simulated MC events.
- The $\Delta\beta$ -corrected relative isolation should be $I_\mu < 0.15$ for the isolated cone, $\Delta R < 0.4$. This is referred to as the tight isolation and it is computed by Equation 5.1.
- At least one muon is required to satisfy $p_T > 24$ GeV and to match an HLT muon object within $\Delta R < 0.5$.
- The distance of the closest approach of each muon to the primary vertex must be within $|d_z| < 0.2$ cm along the beam direction and $|d_{xy}| < 0.045$ cm in the transverse plane.

- b-jets are excluded from the analysis.

The kinematic plots for the leading and sub-leading muons are shown in Figures 7.6-7.8 and impact parameters are shown in Figures 7.9-7.11. Currently due to the mismodeling of the impact parameter in data, these variables are not included in the BDT analysis. In the future, after correcting the discrepancies, the impact parameters can be used in the BDT analysis to improve the sensitivity of the signals. In Section 7.10, the impact of these variables on the signal extraction is explained in detail.

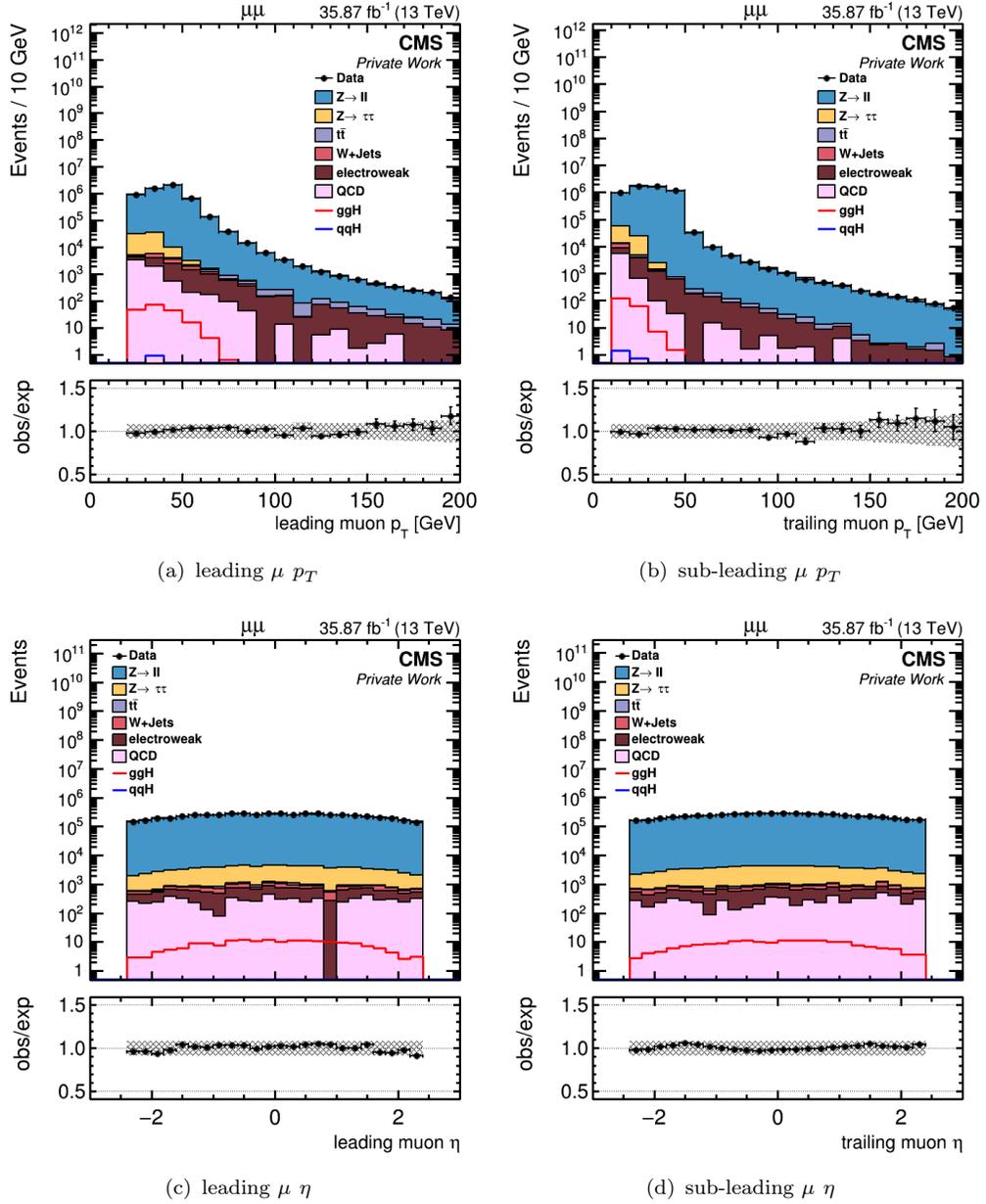


Figure 7.6: Kinematic plots for 0-jet category

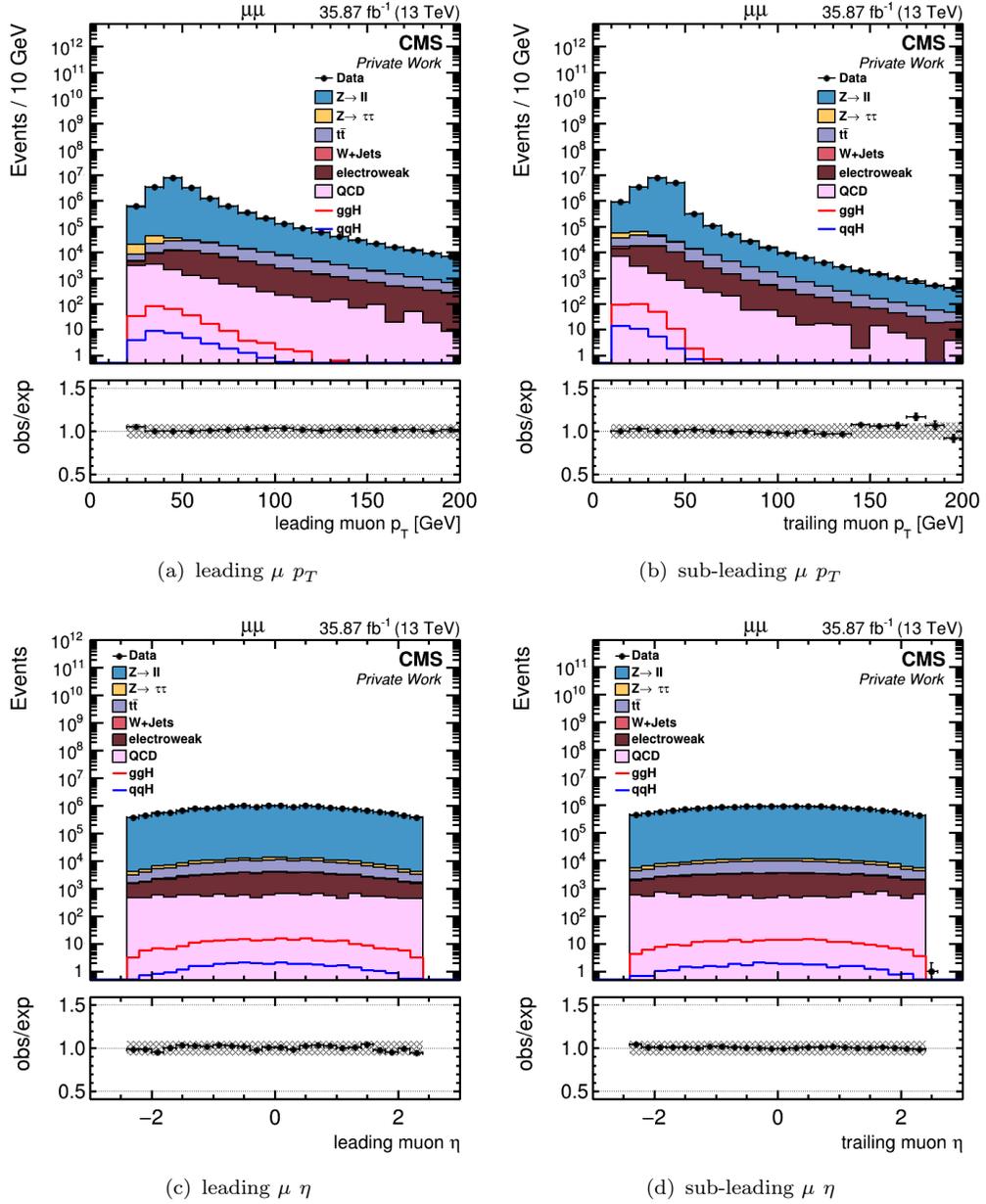


Figure 7.7: Kinematic plots for boosted category

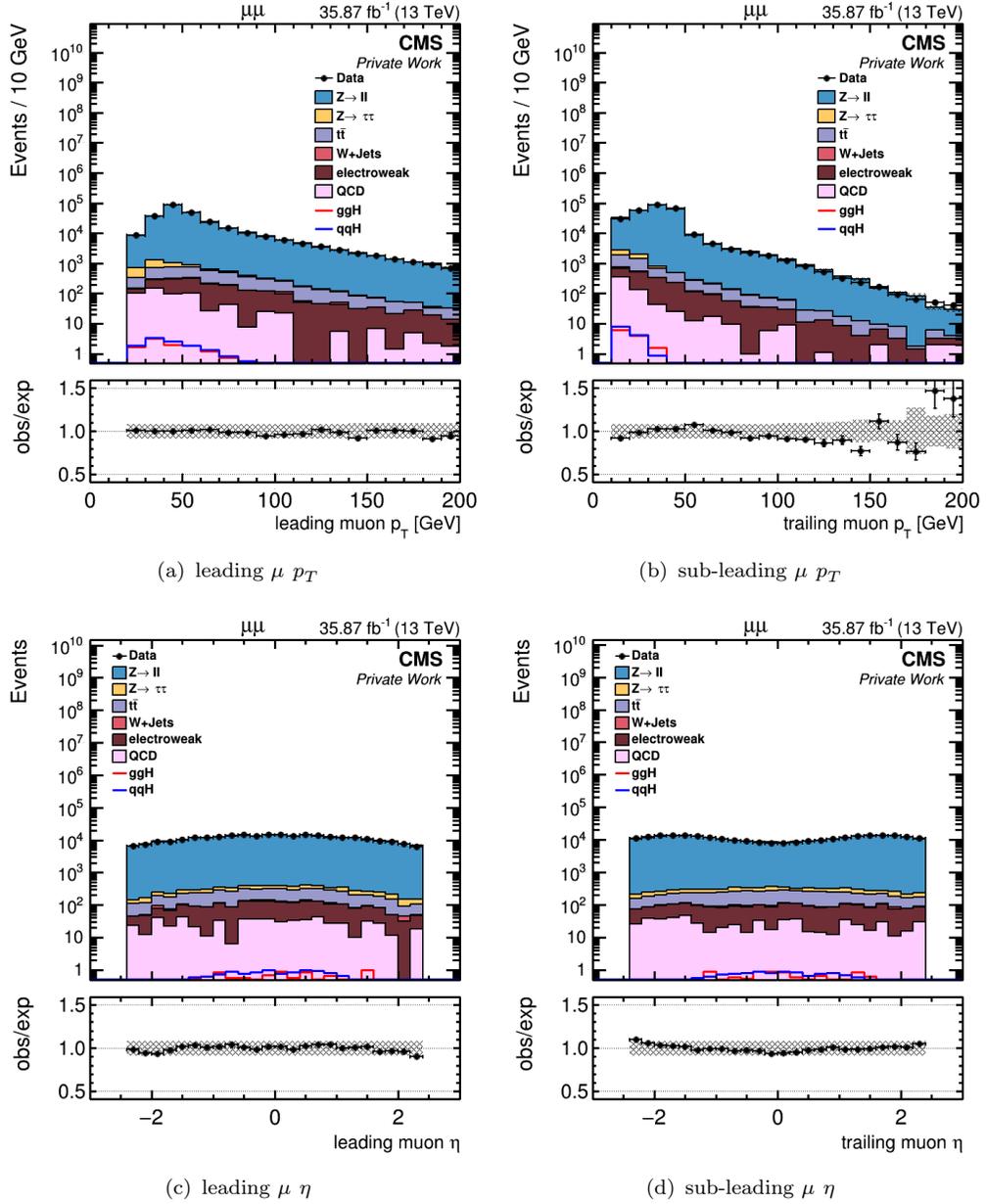
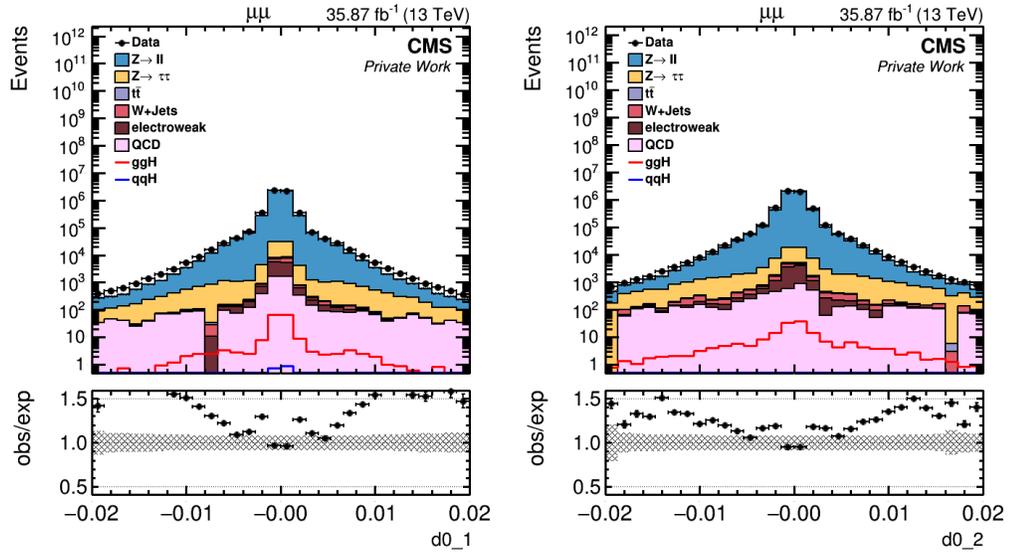
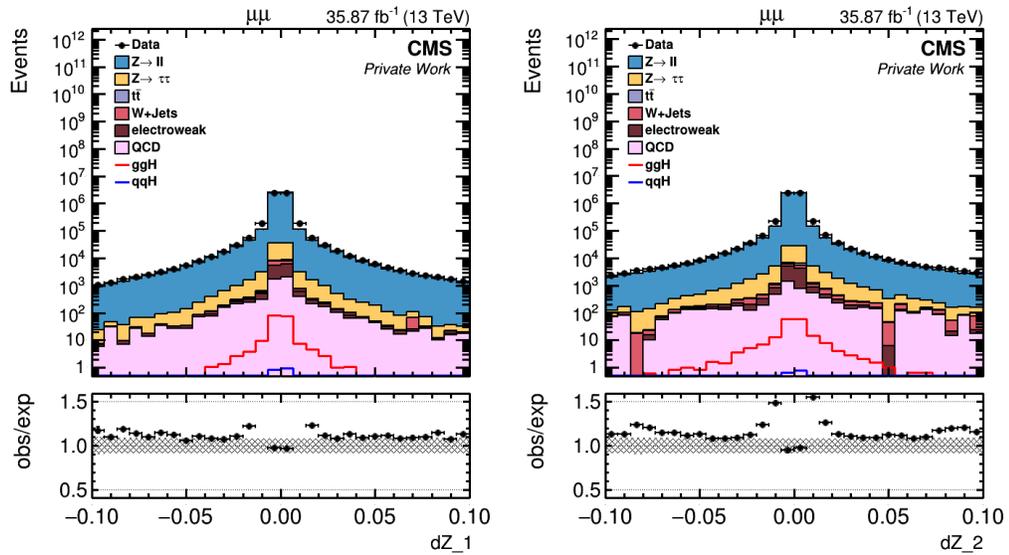


Figure 7.8: Kinematic plots for VBF category

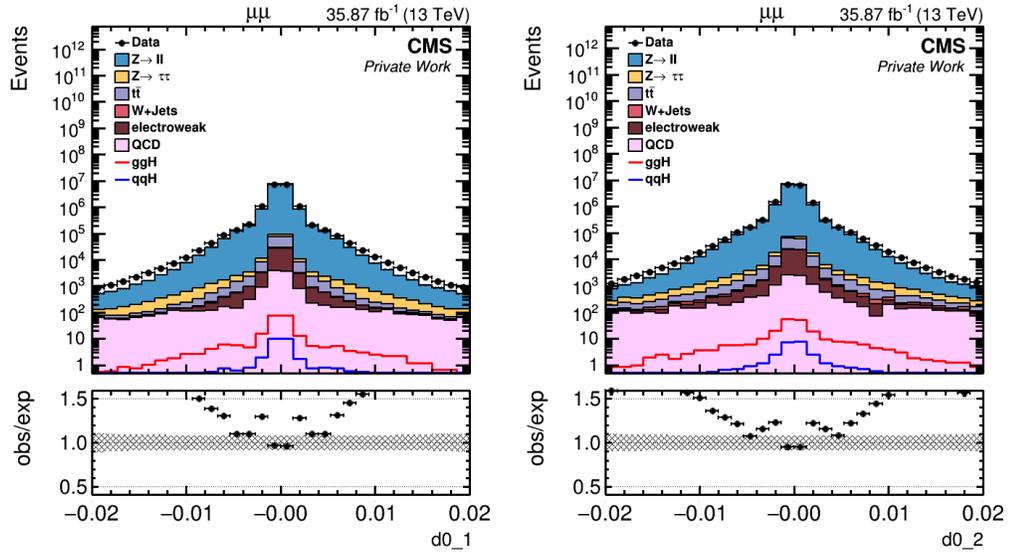


(a) leading μ impact parameter in transverse plane (b) sub-leading μ impact parameter in transverse plane

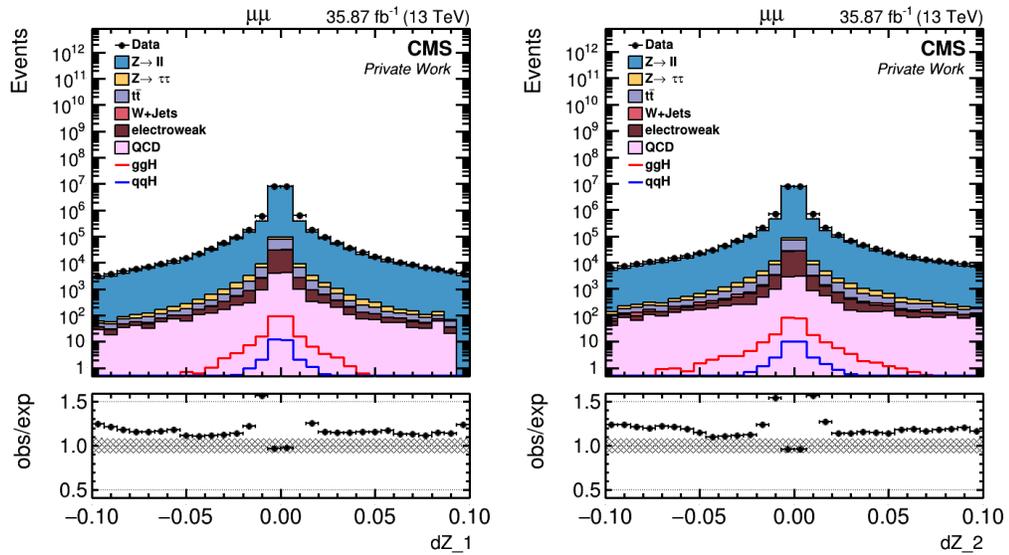


(c) leading μ impact parameter in longitudinal plane (d) sub-leading μ impact parameter in longitudinal plane

Figure 7.9: impact parameter in 0-jet category

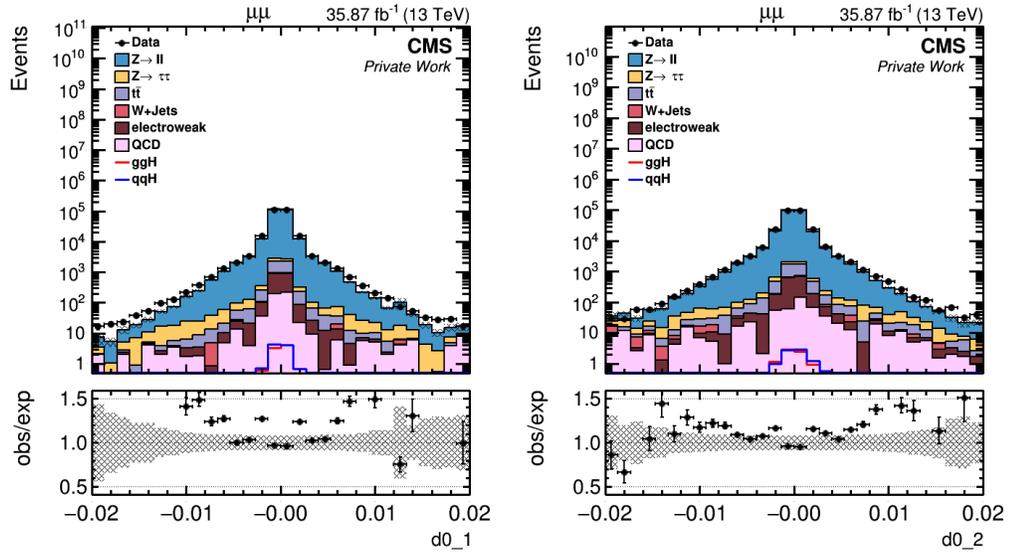


(a) leading μ impact parameter in transverse plane (b) sub-leading μ impact parameter in transverse plane

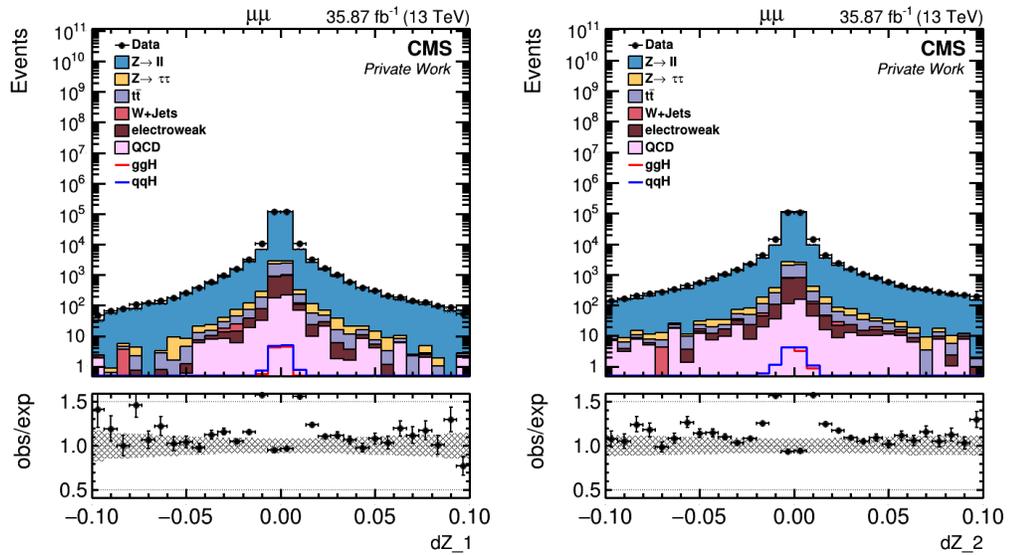


(c) leading μ impact parameter in longitudinal plane (d) sub-leading μ impact parameter in longitudinal plane

Figure 7.10: impact parameter in boosted category



(a) leading μ impact parameter in transverse plane (b) sub-leading μ impact parameter in transverse plane



(c) leading μ impact parameter in longitudinal plane (d) sub-leading μ impact parameter in longitudinal plane

Figure 7.11: impact parameter in VBF category

7.6 Multivariate Techniques to Reduce the Drell-Yan Background

This section describes the multivariate BDT method used to suppress the largest and irreducible DY background for this $\tau_\mu\tau_\mu$ channel. BDTs are trained separately for each category using simulated $H \rightarrow \tau\tau$ signal and DY (including both $Z \rightarrow \mu\mu$ and $Z \rightarrow \tau\tau$ events according to their relative branching ratios) background MC samples. There are six common input discriminants used in all three categories:

- the pseudo-rapidity of the dimuon system, $\eta_{2\mu}$,
- the ratio of the p_T of the dimuon system to the scalar sum of the p_T of the two muons, $p_T(2\mu)/\sum p_T(\mu)$,
- the E_T^{miss} ,
- the P_ζ variable,
- the azimuthal angle between the direction of the three-momentum of the positively charged muon and the \vec{p}_T^{miss} vector, $\Delta\phi(\mu^+, \vec{p}_T^{\text{miss}})$, and
- the decay angle θ^* of the positively charged muon in the rest frame of the dimuon system, $\cos\theta^*$.

The first five variables are explained in detail in Section 6.3. The last variable, $\cos\theta^*$, is added in this analysis, as it helps to discriminate the $Z \rightarrow \mu\mu$ from the $Z/H \rightarrow \tau\tau$ events. In the rest frame of the dimuon system, the two muons coming from the Z decay are back to back, creating an angle for the positive muon to be either 0 or π , in the rest frame of the dimuon system.

For training the BDTs for the boosted category, an additional discriminating variable, the total p_T of the dimuon system, is introduced whereas in the VBF category, an additional di-jet mass variable is introduced, as the category definition is based on this mass cut. This BDT method is formally referred to as the “Single-BDT method”, since BDTs are only trained once.

The variables used as a BDT input are shown in Figure 7.12-7.14. In BDT, the variables are ranked based on their performance to separate the signal events from the background. For the 0-jet category, $p_T(2\mu)/\sum p_T(\mu)$ variable is the best-ranked variable. The P_ζ variable is top ranked in the boosted category, while E_T^{miss} performs best ifor the VBF category. The medium BDT cuts are used in each category for reducing computational time when calculating the SVFit di- τ mass. The details on the reconstruction of this mass are explained in Section 5.1.5. The medium cuts used in the SVFit mass calculations are shown in Table 7.4.

Table 7.4: Medium BDT cuts used in each event category for SVFit di- τ mass calculation.

Event Category	BDT ResponseCuts
0-jet	> 0.2
boosted	> 0.0
VBF	> 0.8

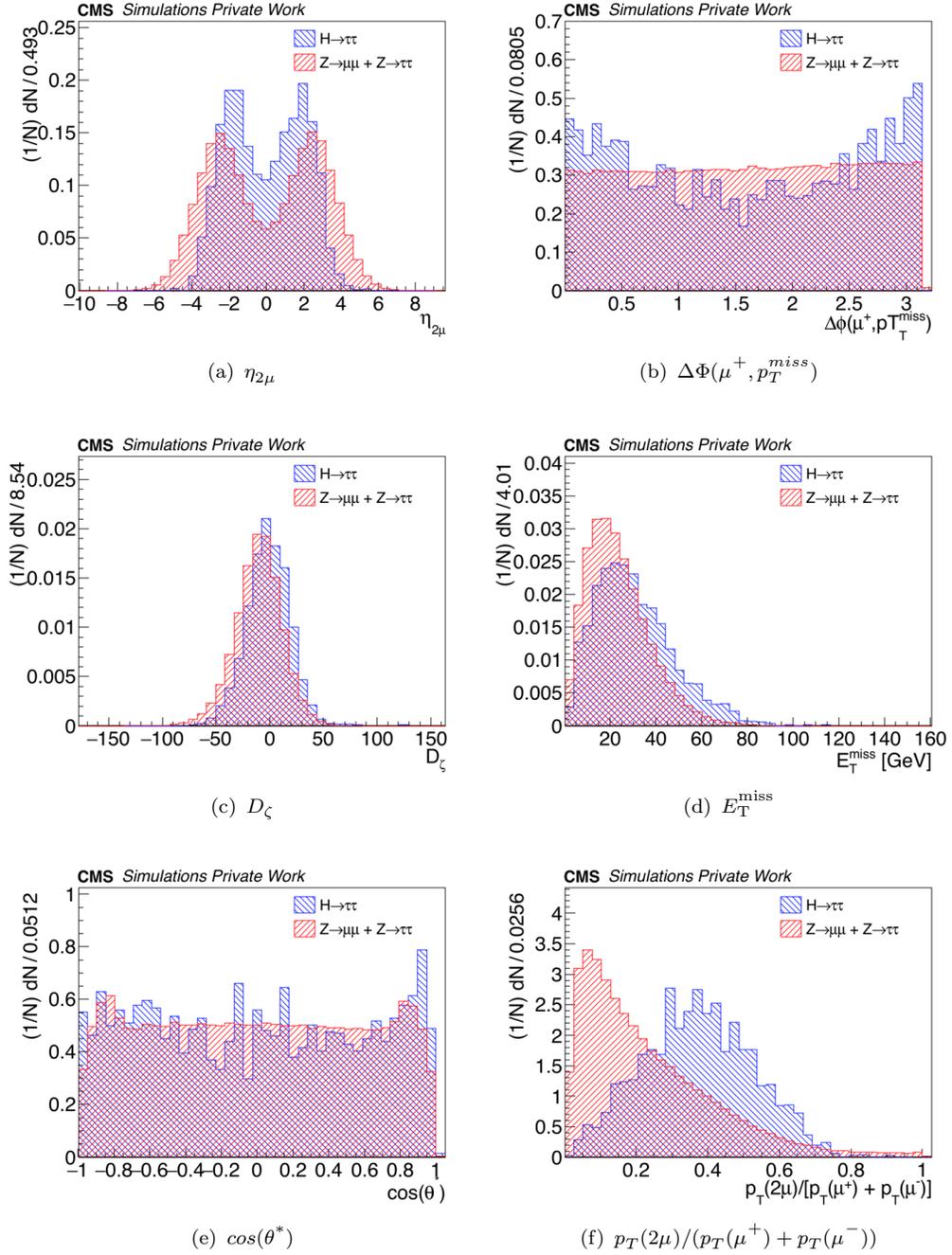
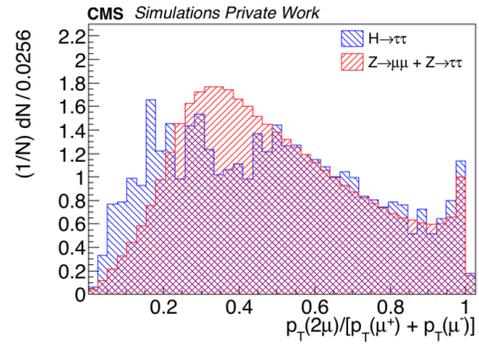
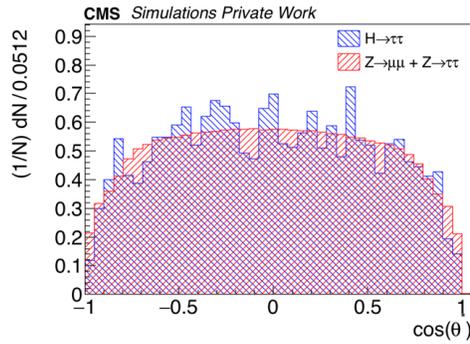
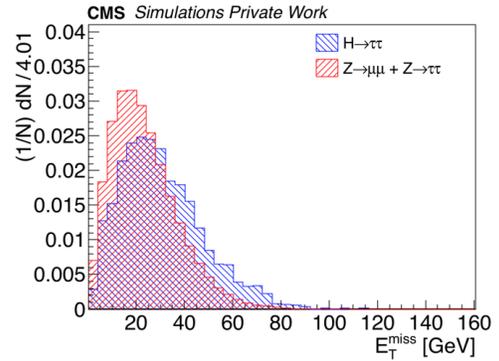
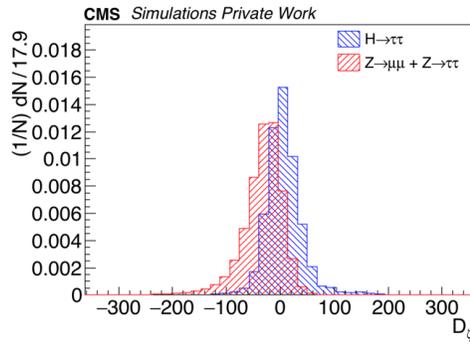
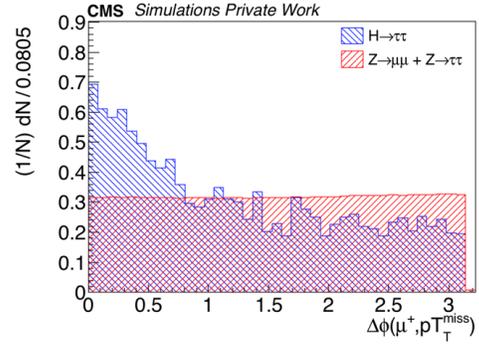
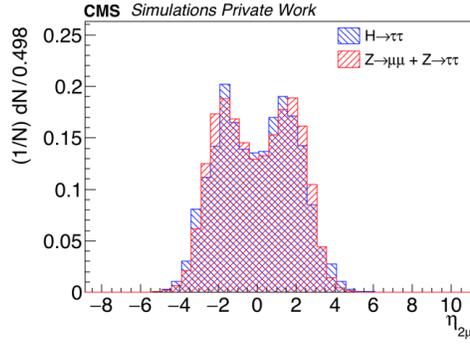
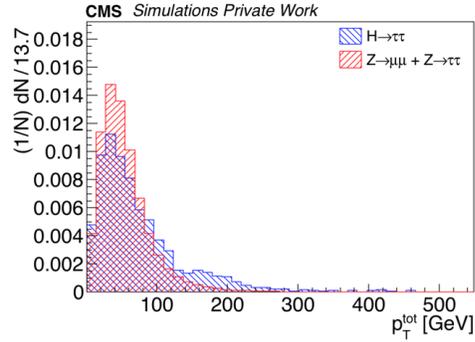


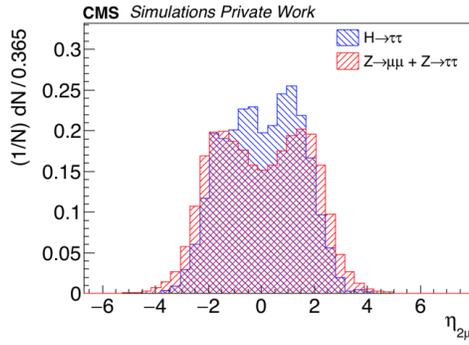
Figure 7.12: BDT discriminant inputs for 0-jet category



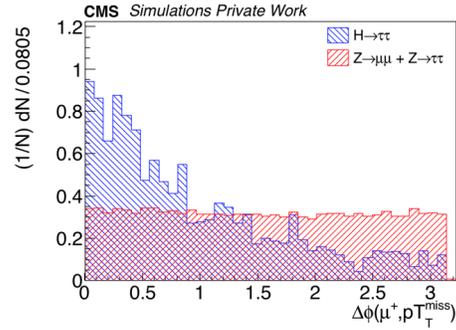


(g) p_T^{tot}

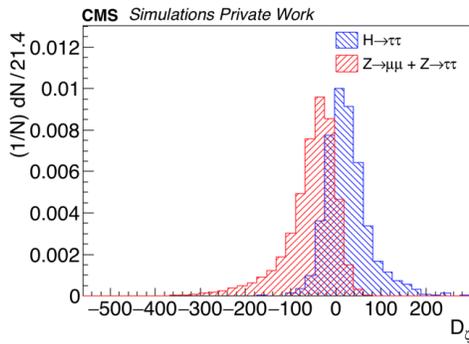
Figure 7.13: BDT discriminant inputs for boosted category



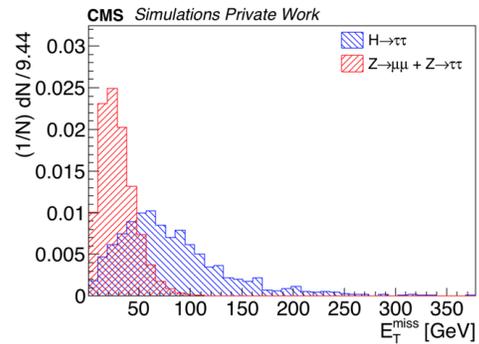
(a) $\eta_{2\mu}$



(b) $\Delta\Phi(\mu^+, p_T^{miss})$



(c) D_c



(d) E_T^{miss}

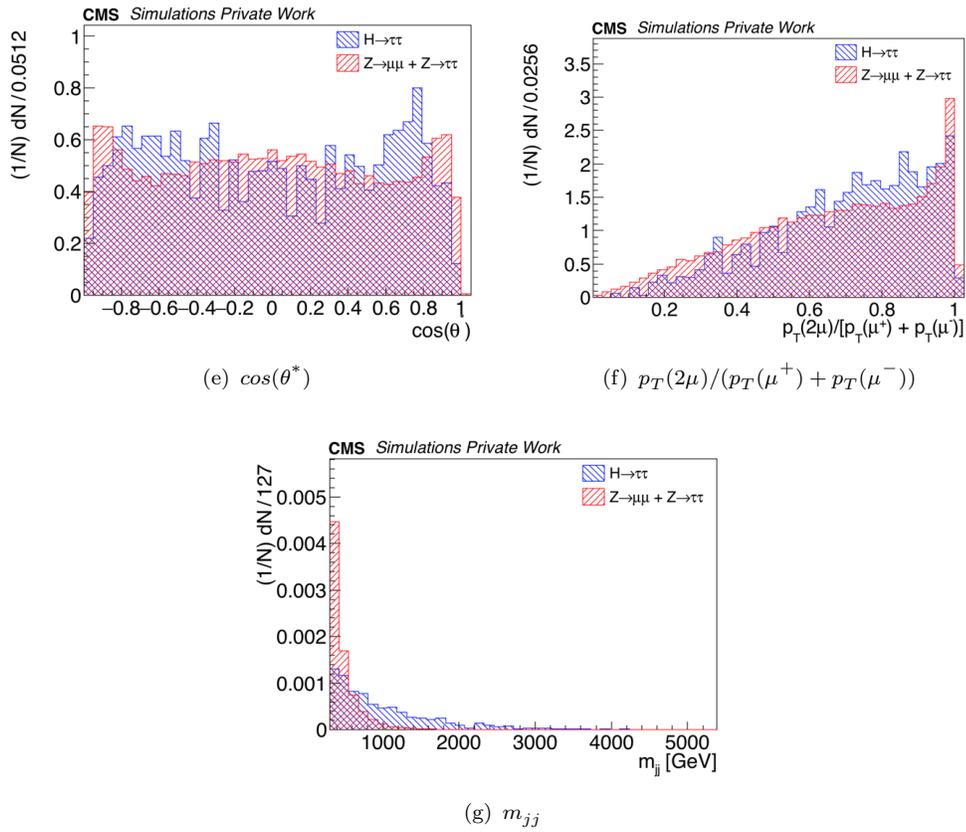


Figure 7.14: BDT discriminant inputs for VBF category

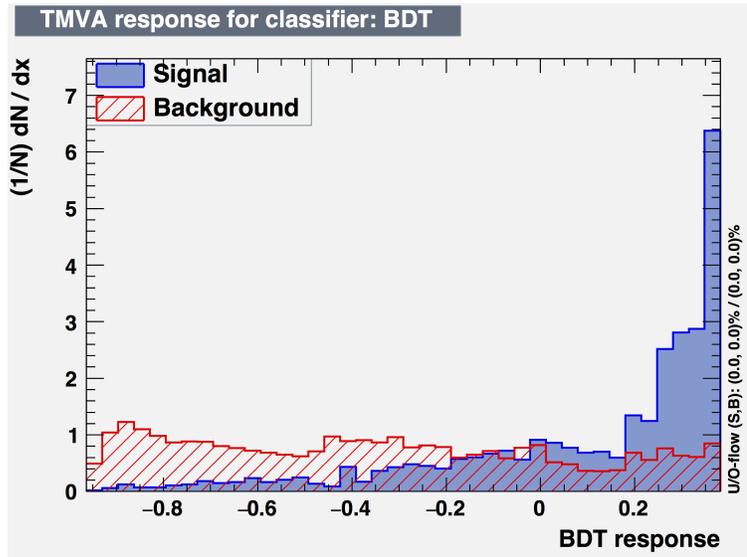


Figure 7.15: BDT response by TMVA package for 0-jet category. Blue represents $H \rightarrow \tau\tau$ and red represents the DY background.

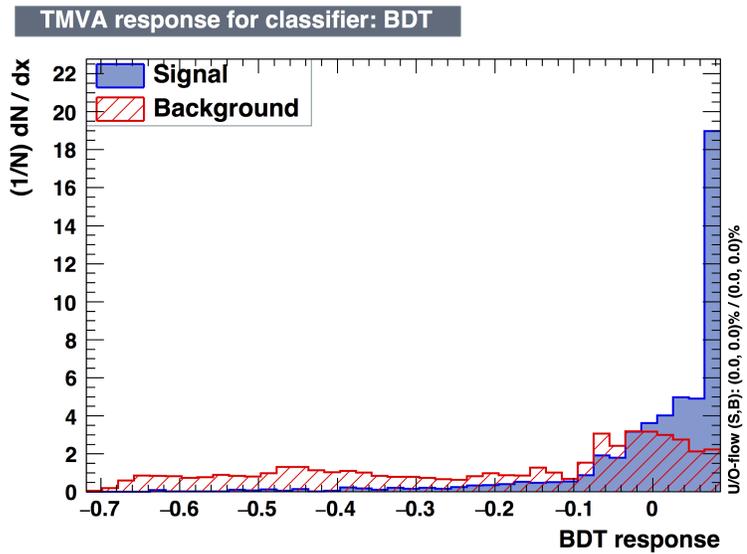


Figure 7.16: BDT response by TMVA package for boosted category. Blue represents $H \rightarrow \tau\tau$ and red represents the DY background.

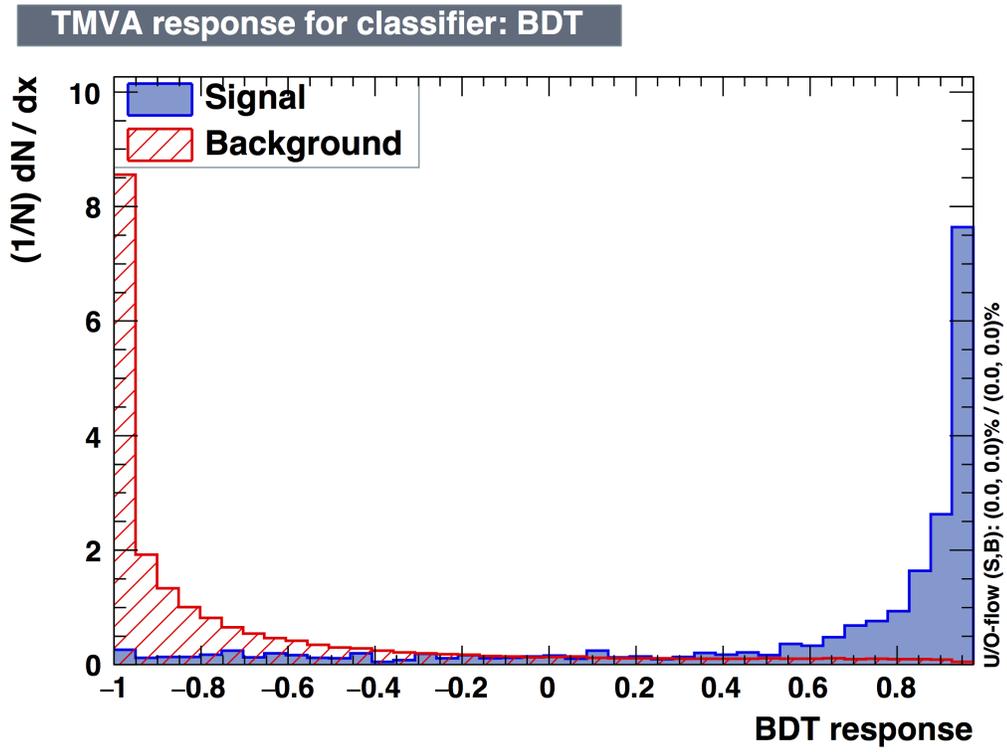


Figure 7.17: BDT response by TMVA package for VBF category. Blue represents $H \rightarrow \tau\tau$ and red represents the DY background.

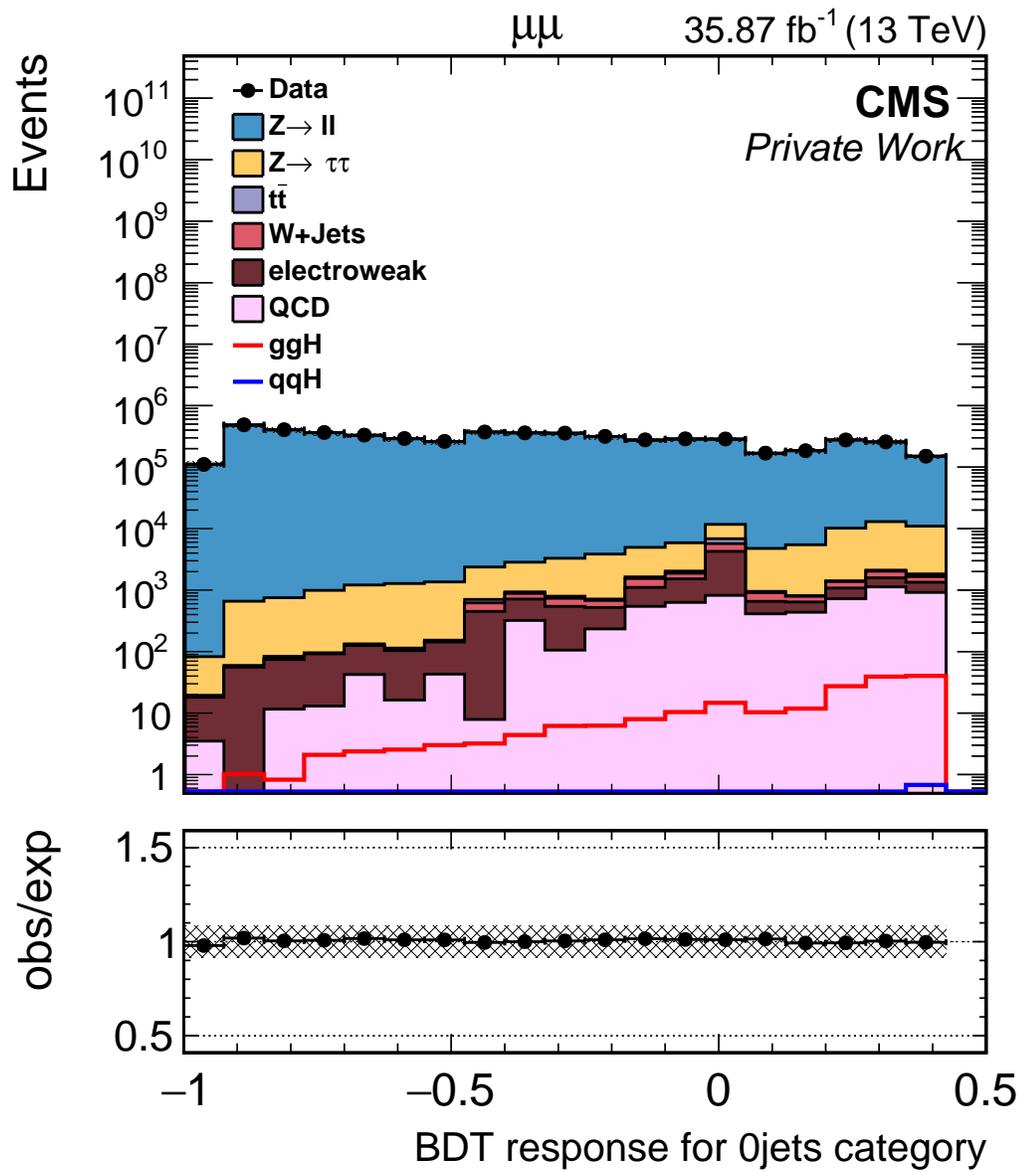


Figure 7.18: BDT response in data and simulated events for 0-jet category.

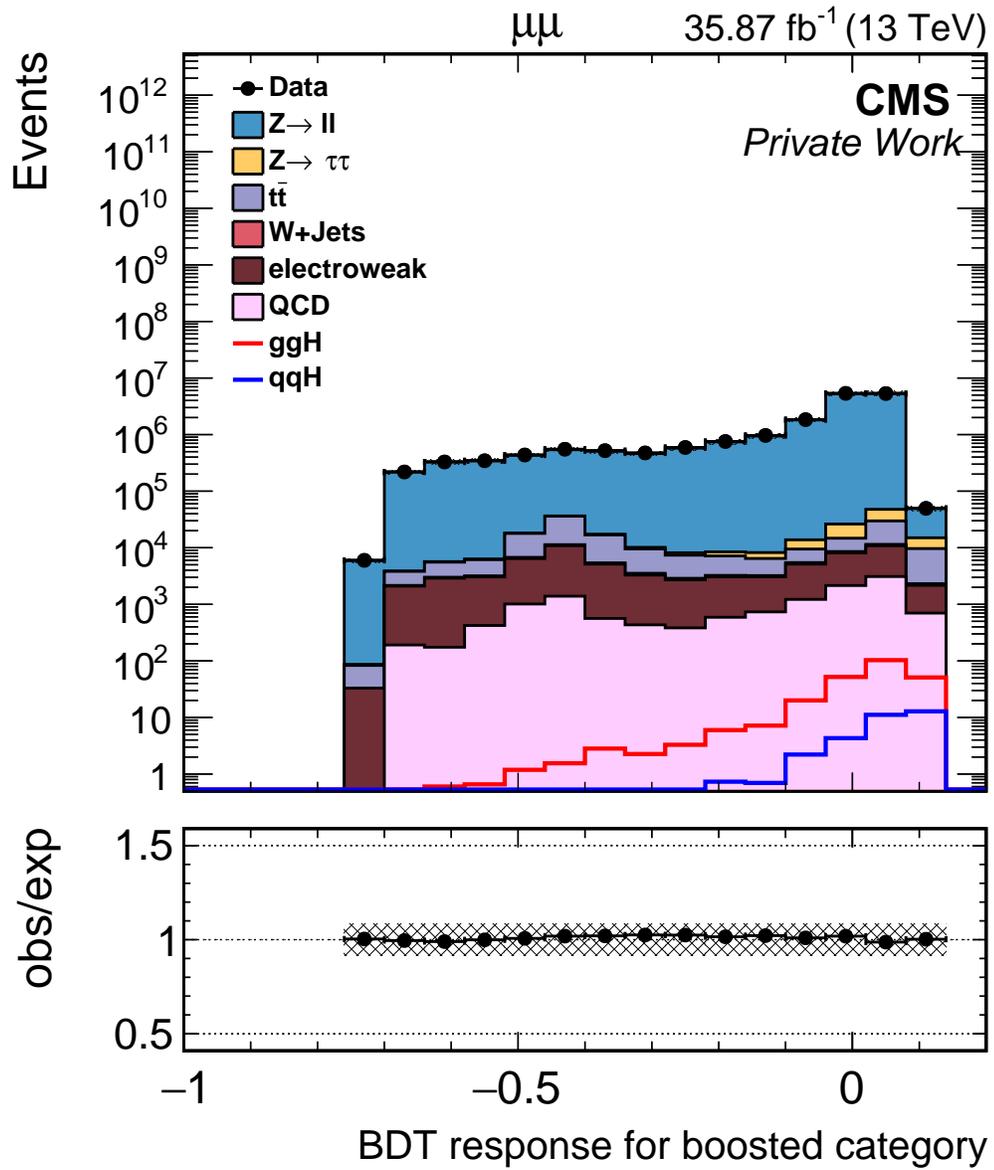


Figure 7.19: BDT response in data and simulated events for boosted category.

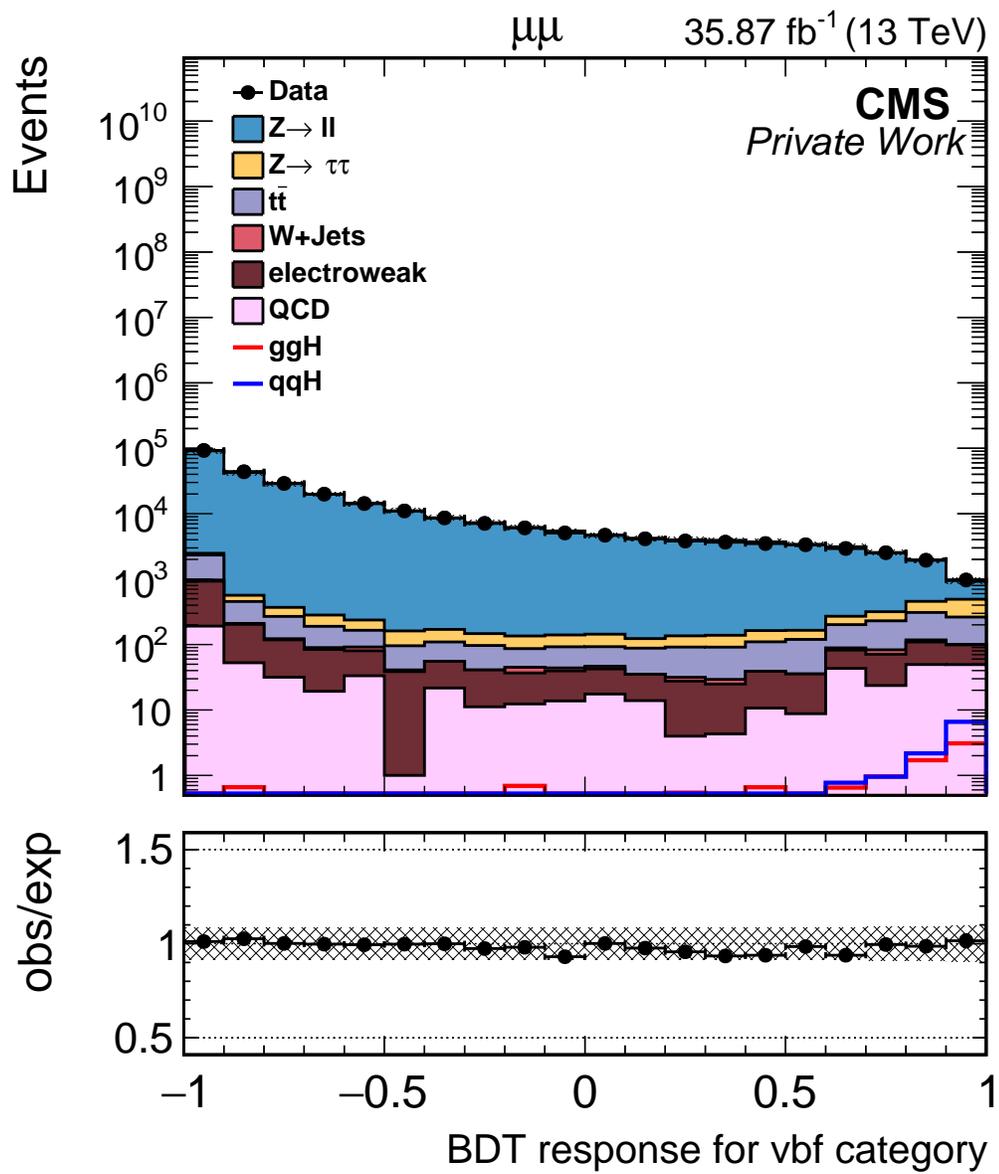


Figure 7.20: BDT response in data and simulated events for VBF category.

The BDT responses evaluated by TMVA are shown in Figure 7.15-7.17 and the responses in the data and the MC samples are shown in Figures 7.18-7.20. These BDT cuts are optimized by maximizing the signal to the square-root of the background ratio in the two-dimensional distribution of the dimuon mass and SVFit mass, and also by minimizing the expected limits on the expected SM cross section of the respective Higgs production. These optimized cuts for signal extraction are shown in Table 7.5. The ROC curves for each category are shown in Figure 7.21-7.23.

Table 7.5: Final optimized BDT cuts used in each event category for signal extraction.

Event Category	Optimized BDT Response Cuts
0-jet	+0.270
boosted	+0.085
VBF	+0.920

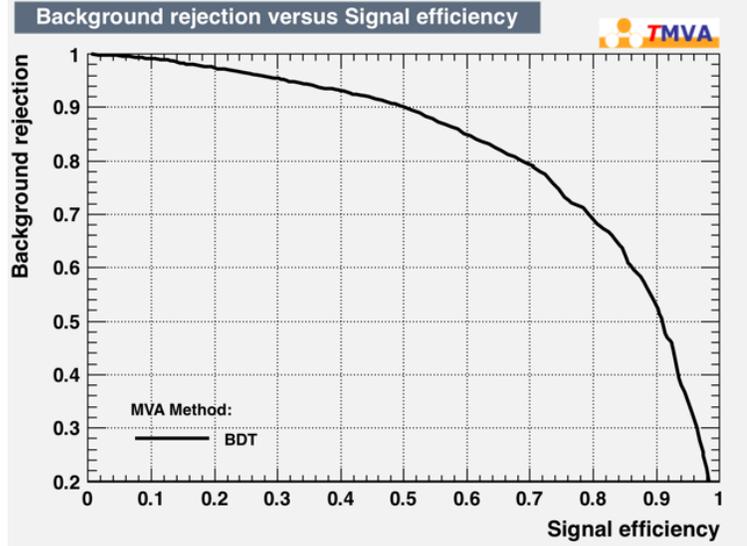


Figure 7.21: ROC curve for 0-jet category.

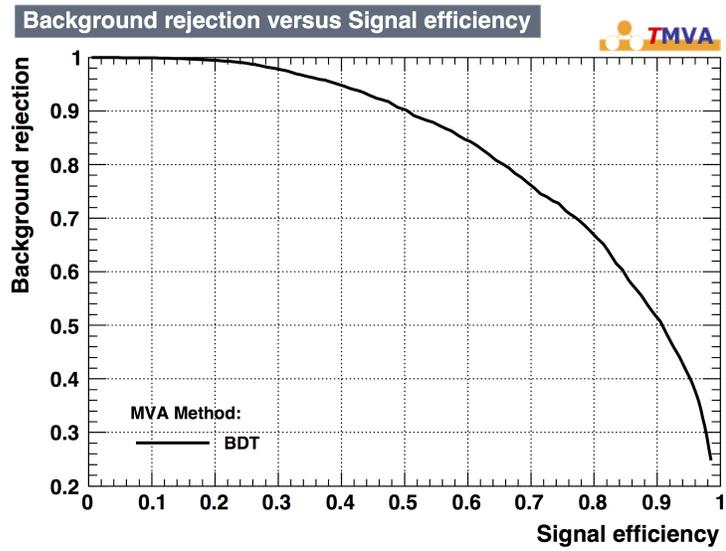


Figure 7.22: ROC curve for boosted category.

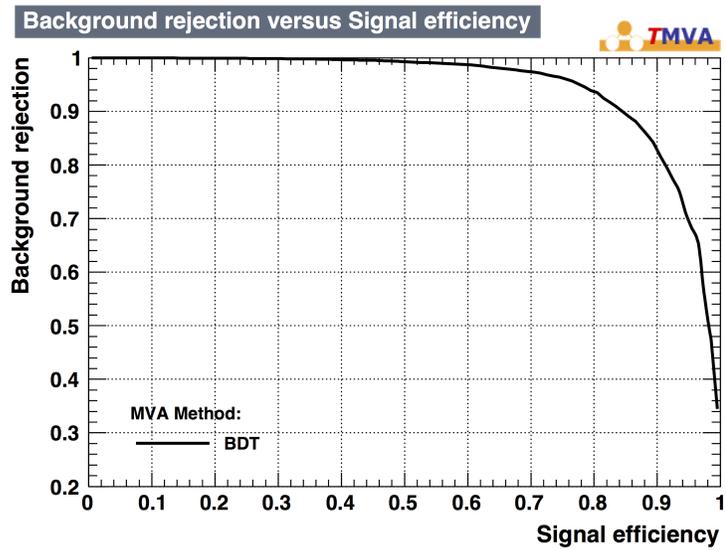


Figure 7.23: ROC curve for VBF category.

7.7 Signal Extraction

The 1D invariant visible mass of dimuons and reconstructed invariant SVFit di-tau mass distribution after applying all required cuts and corrections are shown in Figures 7.24 and 7.25, respectively. The final results are extracted with a maximum likelihood fit based on the 2D distributions of the dimuon mass $m_{\mu\mu}$ and SVFit mass $m_{\tau\tau}$ in each category for events that pass the final BDT cuts.

The 2D distribution can separate the signal from the background due to the separation in the resolution of the mass peaks. This can be confirmed by observing the mass distribution for the signal and background cases as shown in Figures 7.26-7.28. Signal and backgrounds peak in different bins in these distributions.

For bin-by-bin (bbb) fitting purpose, the 2D distributions are “unrolled” into 1D histograms. The prefit distributions for each category are shown in Figures 7.29-7.31 and the corresponding background event yields are summarized in Table 7.6.

Table 7.6: Expected pre-fit numbers of background events in the $\tau_\mu\tau_\mu$ channel in the data corresponding to an integrated luminosity of 35.9 fb^{-1} for the three jet categories. The numbers are rounded to a precision of two significant digits on the uncertainty.

Process	0-jet	boosted	VBF
$Z \rightarrow \mu\mu$	$28\,886 \pm 93$	$2\,027 \pm 37$	155 ± 11
$Z \rightarrow \tau\tau$	$20\,502 \pm 16$	$2\,585 \pm 37$	171 ± 12
Multijet	$1\,900 \pm 15$	182 ± 36	22 ± 10
Electroweak	560 ± 7	240 ± 7	16 ± 2
$t\bar{t}$	111 ± 7	$1\,266 \pm 23$	60 ± 5
Total Expected Background	$312\,495 \pm 561$	$6\,336 \pm 80$	425 ± 21
Signal $H \rightarrow \tau\tau$	86 ± 3	28 ± 2	6 ± 0
Observed in Data	308 013	6 342	405

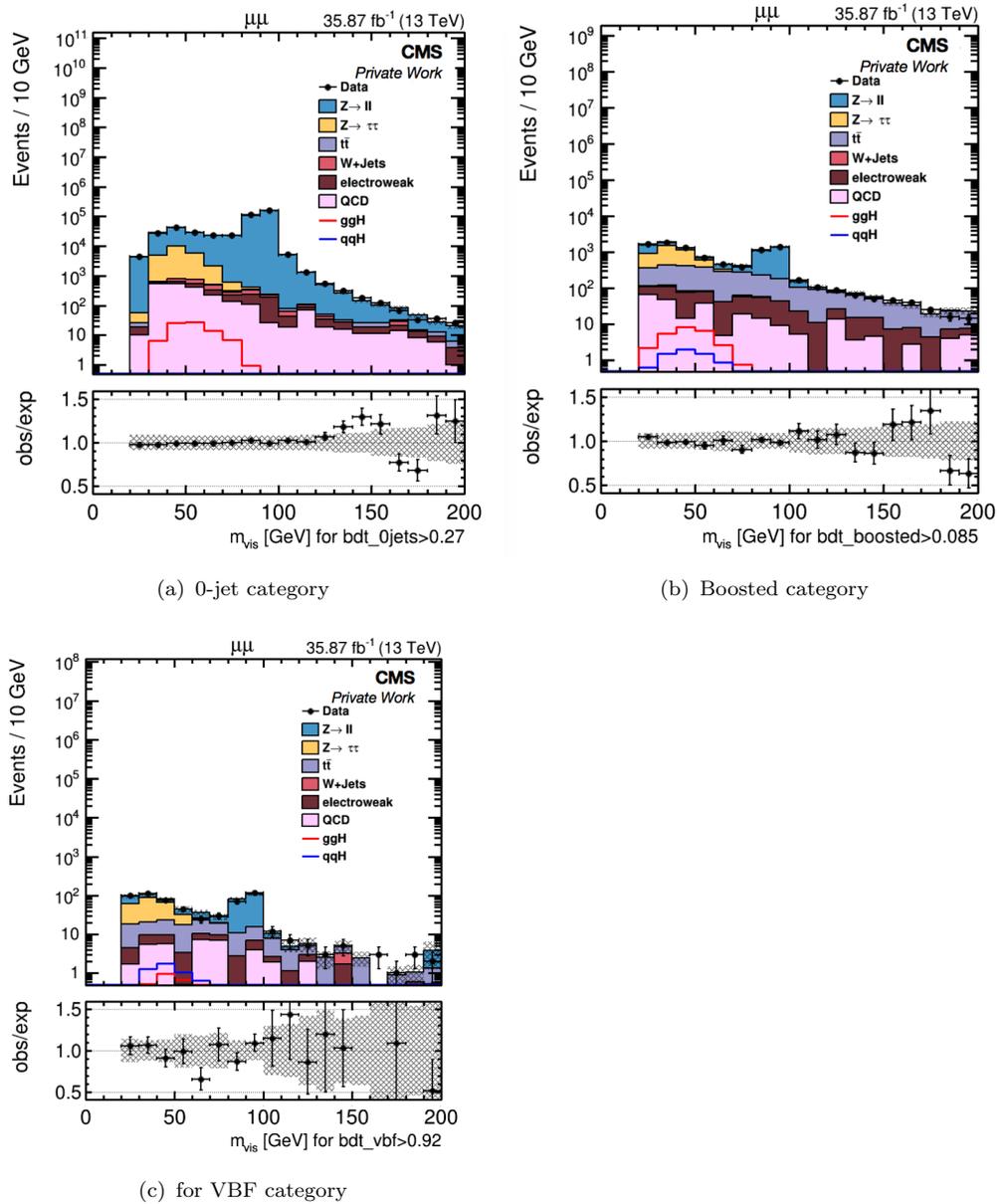


Figure 7.24: Invariant dimuon mass ($m_{\mu\mu}$) for three event jet categories.

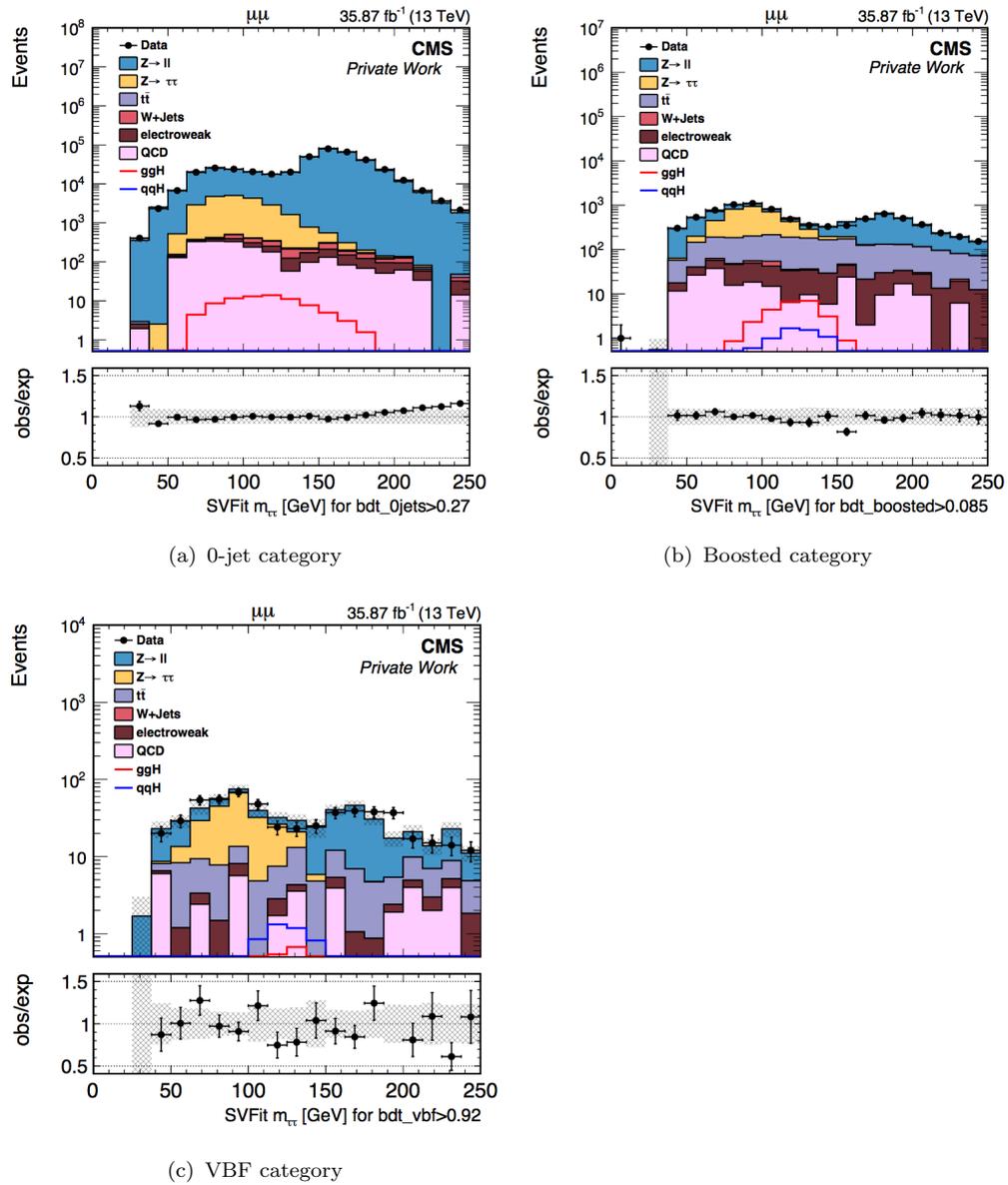


Figure 7.25: Invariant di-tau mass ($m_{\tau\tau}$) for three event jet categories.

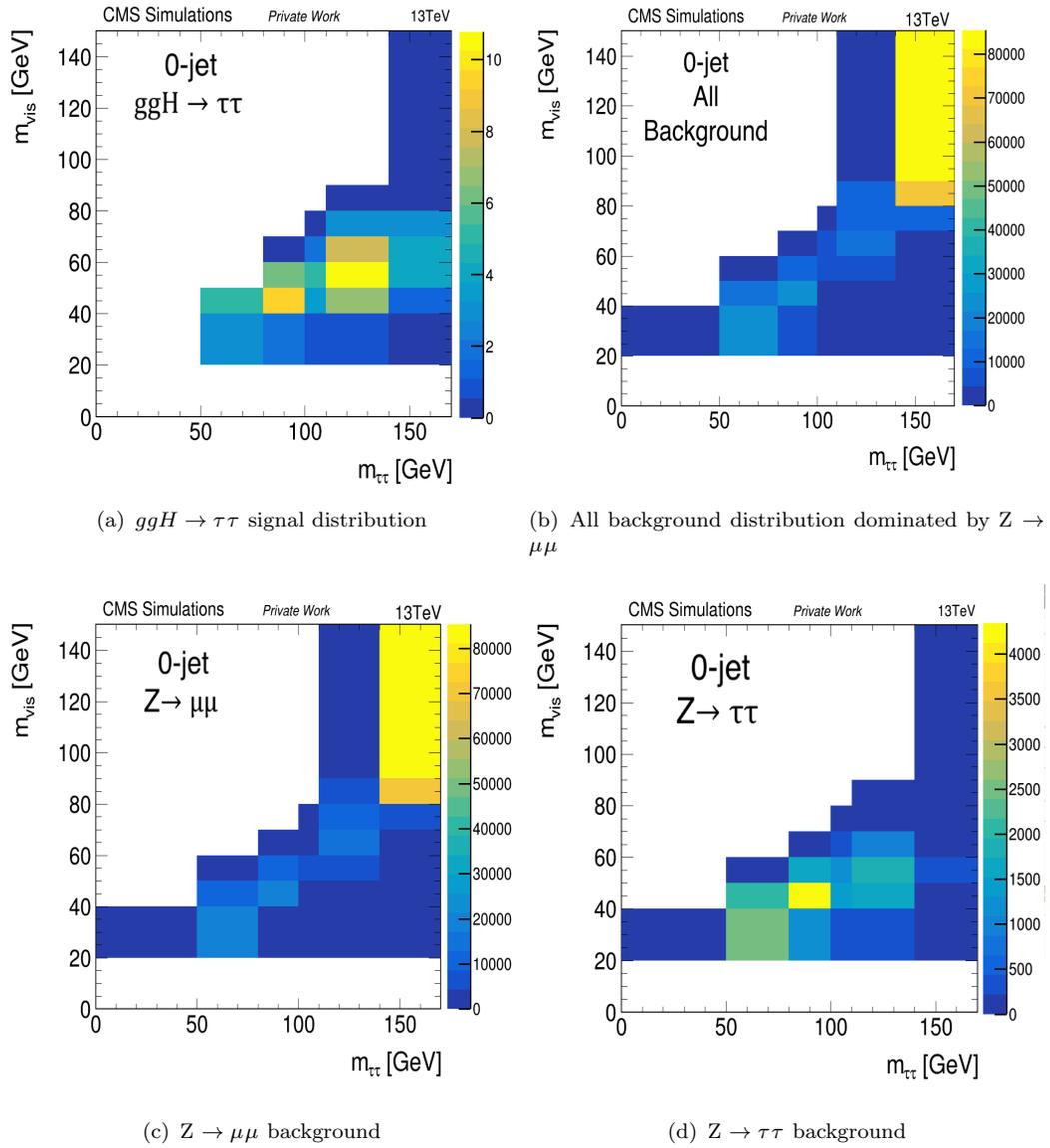


Figure 7.26: 2D mass distribution for simulated signal and background events in the 0-jet category.

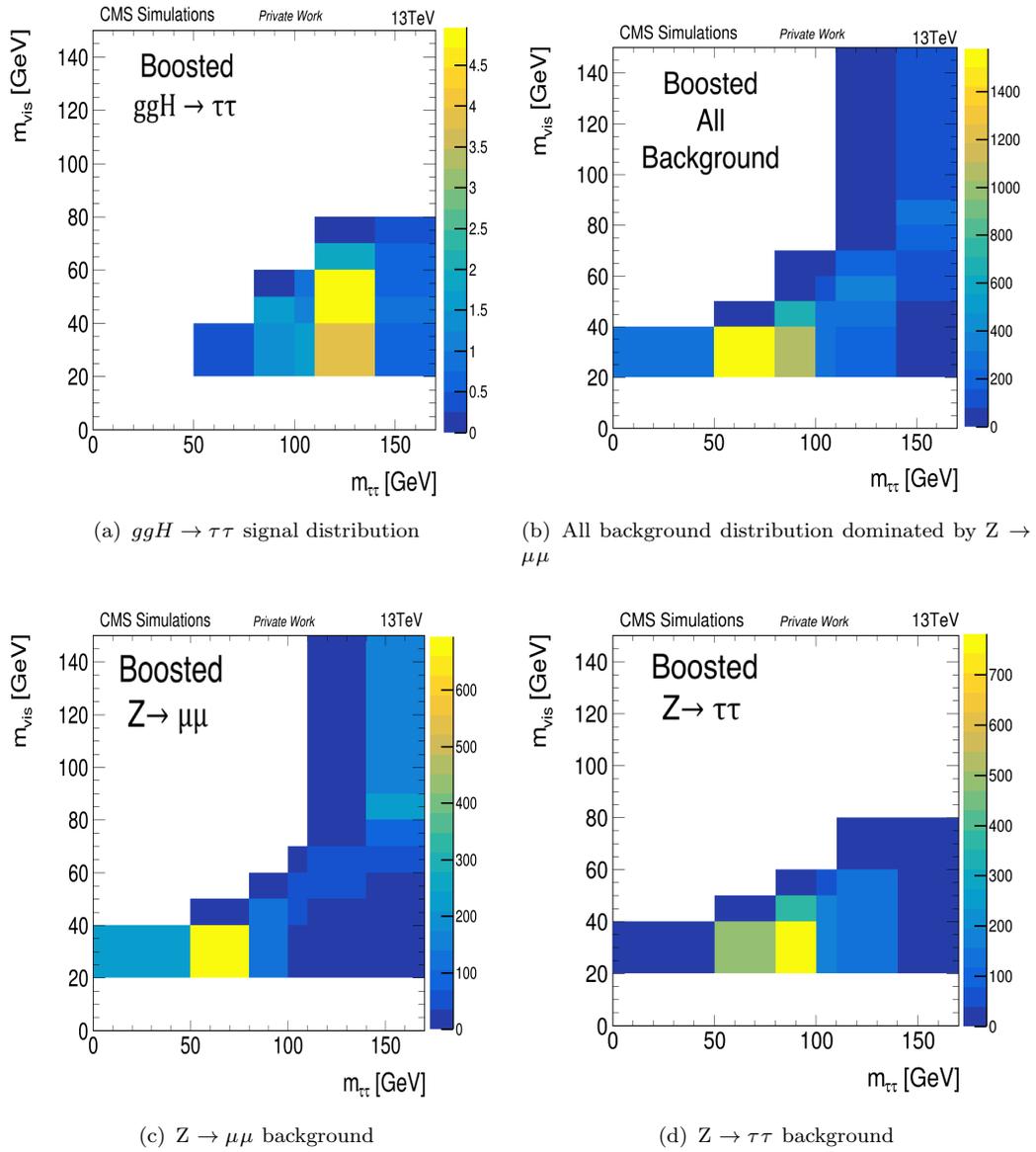


Figure 7.27: 2D mass distribution for simulated signal and background events in the boosted category.

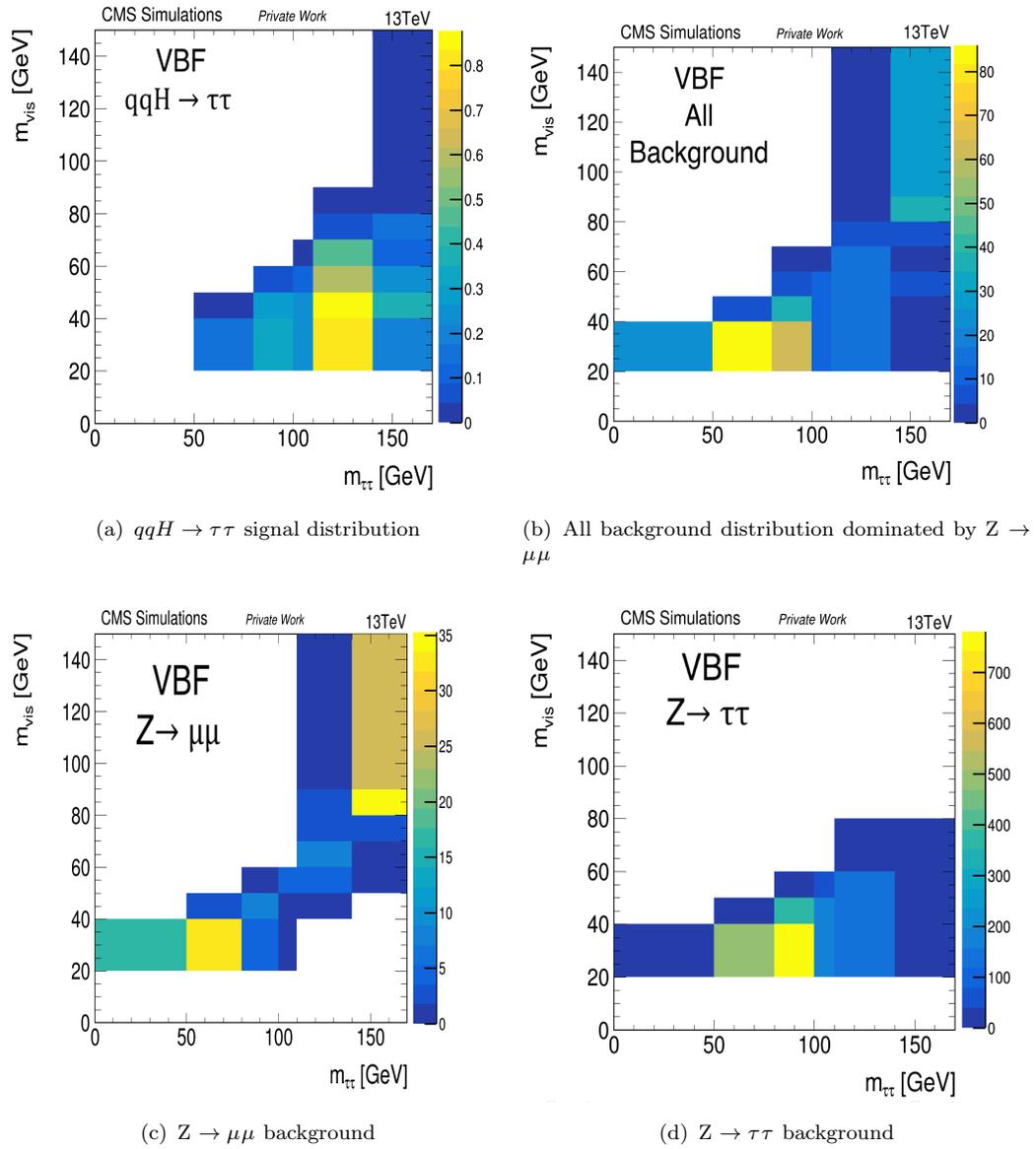


Figure 7.28: 2D mass distribution for simulated signal and background events in the VBF category.

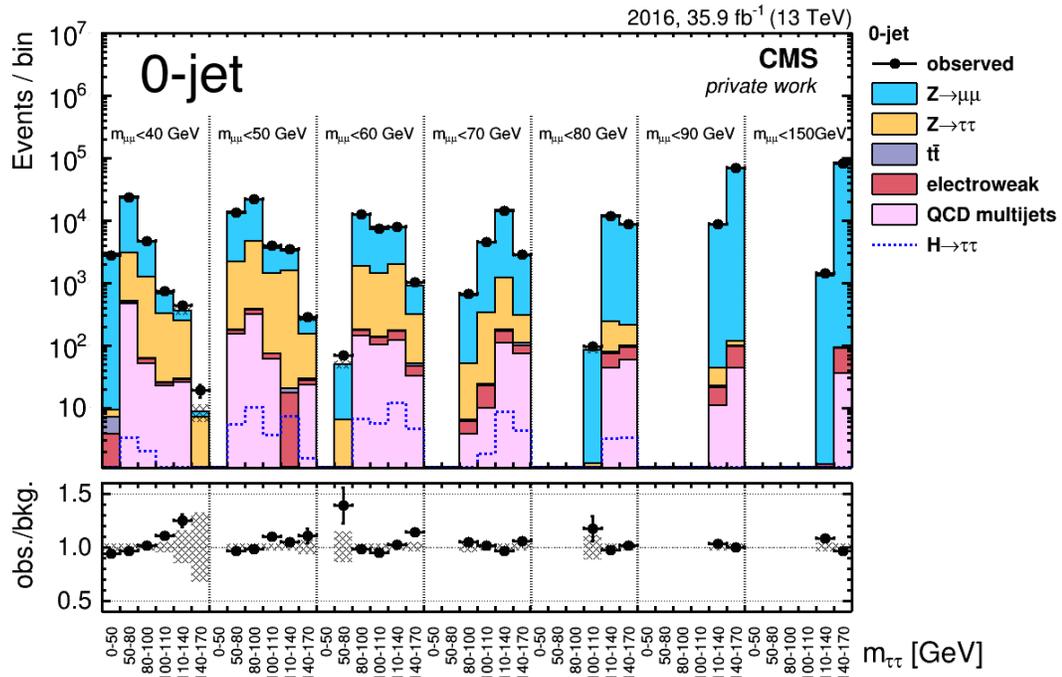


Figure 7.29: The prefit 2D mass distribution for the 0-jet category.

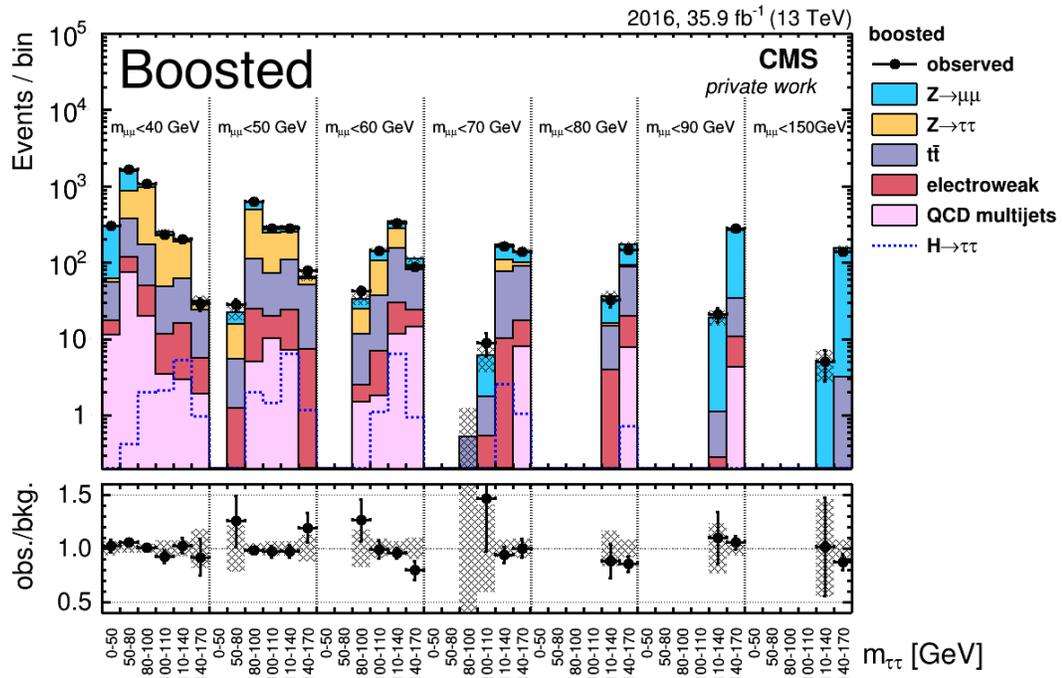


Figure 7.30: The prefit 2D mass distribution for the boosted category.

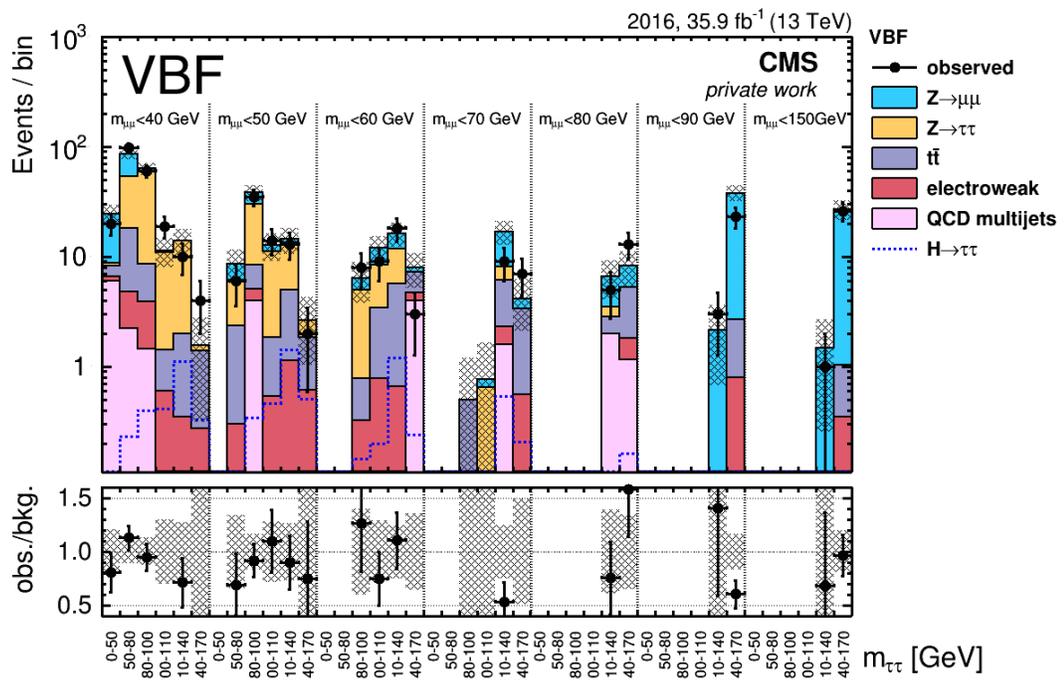


Figure 7.31: The prefit 2D mass distribution for the VBF category.

7.8 Systematic Uncertainties

Before the 2D mass distribution can be used to extract the signal strength uncertainties with a maximum likelihood fit, the uncertainties must be quantified as they enter into the fit as nuisance parameters. Details on the statistical strategies are discussed in Section 5.3. There are several factors which can affect the shape or normalization of the 2D distributions. The systematic uncertainties for the $\tau_\mu\tau_\mu$ channel are described by the following points.

- **Object reconstruction and identification**

The uncertainty in the trigger efficiency for this channel is about 2%. Similarly, the uncertainties on the identification and isolation of muons are measured with an accuracy of 2%. The uncertainties of the muon momentum scale reflects the 1% uncertainty in its measurement with $Z \rightarrow \mu\mu$ events. Overall, the signal acceptance for the $H \rightarrow \tau\tau$ events is not affected by this uncertainty. The change in the signal acceptance is even less than 1%. The uncertainties in the E_T^{miss} modeling directly affects the signal strength as it propagates through the $m_{\tau\tau}$ mass distribution and through the other three discriminant inputs used in the BDT that depend on the E_T^{miss} contribution. The effect of these uncertainties on the overall mass distribution is very small, but it introduces about 10% uncertainty in the signal acceptance. The uncertainty on the jet-energy scale depends on p_T and η of the jets and is reflected in the jet multiplicity in the events. Hence, these uncertainties affect the categorization of the events between the three categories. Also, it can affect the estimation of the di-jet mass, which is used as an additional BDT input for the VBF category. Uncer-

tainties in discarding b-jet events in the $t\bar{t}$ background are very small for this channel and uncertainties in the mistagging rate of gluon and light-flavor jets are also negligible.

- **Background yields**

The uncertainties in the QCD multijet backgrounds are dominated by the OS/SS ratio and their contribution is about 20%. The uncertainty in the $t\bar{t}$ normalization is about 7% and the shape uncertainties are introduced by varying the applied weight with either no reweighting or with the reweighting applied twice. The normalization uncertainties from the event yields for single top, $W+$ jets, and diboson backgrounds are about 15%. Shape uncertainties due to DY reweighting are introduced by applying 1.1 times the corrections and they are correlated within the categories. Uncertainties are directly taken from [118].

- **Other**

The uncertainty in the integrated luminosity is about 2.5%.

7.9 Final ML Fit and Results

The final result for this channel is extracted using the maximum likelihood fit based on the 2D distribution of masses by combining all three event categories. The uncertainties mentioned above are treated as nuisance parameters in the binned likelihood function given by Equation 5.8. The nuisance parameters affecting the normalization are considered as a log-normal PDF while the ones affecting the shape are represented with a Gaussian PDF. The 2D mass distri-

butions post ML fit are shown in Figures 7.32-7.34. The signal and background events obtained from the ML fit for each category are summarized in Table 7.7.

Table 7.7: Post-fit yields in $H \rightarrow \tau\tau$ signal events and respective background events in the $\tau_\mu\tau_\mu$ channel with data corresponding to an integrated luminosity of 35.9 fb^{-1} . The numbers are rounded to a precision of two significant digits on the uncertainty.

Process	0-jet	boosted	VBF
$Z \rightarrow \mu\mu$	$28\,397 \pm 533$	$2\,044 \pm 45$	134 ± 12
$Z \rightarrow \tau\tau$	$20\,851 \pm 14$	$2\,585 \pm 51$	167 ± 13
Multijet	$1\,973 \pm 45$	194 ± 14	26 ± 5
Electroweak	550 ± 23	229 ± 15	16 ± 4
$t\bar{t}$	110 ± 11	$1\,248 \pm 35$	59 ± 8
Total Background	$308\,032 \pm 56$	$6\,332 \pm 80$	402 ± 20

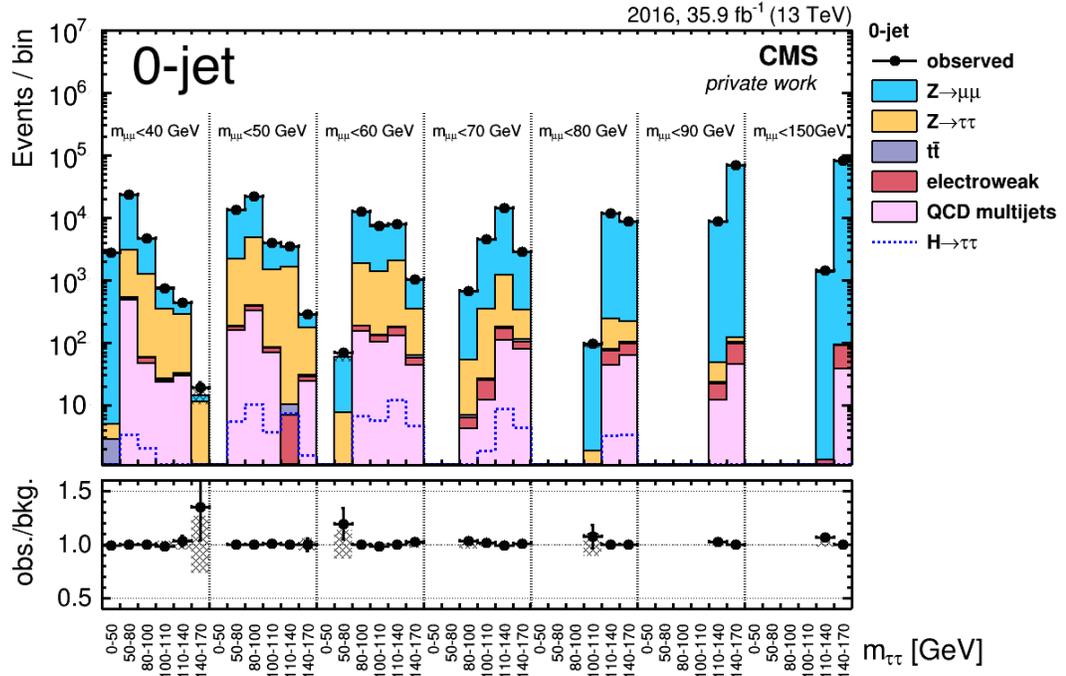


Figure 7.32: The post-fit 2D mass distribution for the 0-jet category.

Using these ML fit values, the expected and the observed limits are extracted. The detailed method is explained in Section 5.3.2. The upper limits

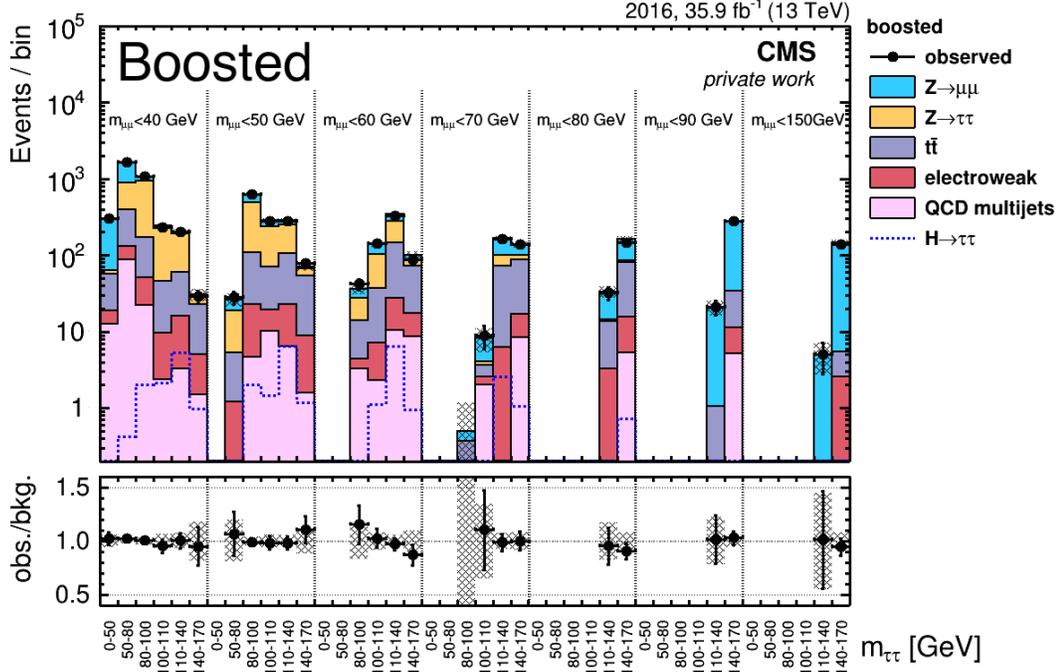


Figure 7.33: The post-fit 2D mass distribution for the boosted category.

at 95% CL on the signal strength for $m_H = 125 \text{ GeV}$ are displayed in Table 7.8 and Figure 7.35 show these limits for mass range 110-140 GeV.

Table 7.8: Upper limits at 95% CL on the signal strength relative to standard model prediction in the $\tau_\mu \tau_\mu$ channel.

Category	-2σ	-1σ	exp.	$+1\sigma$	$+2\sigma$	observed
0-jet	12.7	17	23.8	33.6	45.8	34.9
boosted	2.8	3.8	5.3	7.4	10.0	4.0
VBF	2.1	2.8	4.0	5.7	8.0	4.2
Combination	1.7	2.3	3.2	4.6	6.2	2.7

The Goodness-of-Fit (GOF) test is performed in each category using the saturated model, as explained in Section 5.3.3. Figures 7.36 show the GOF test responses in each category. The values from the ML fit are used to study the pulls and impacts of the various nuisance parameters on the signal strengths. These pulls and impacts are given in Figures 7.37-7.39.

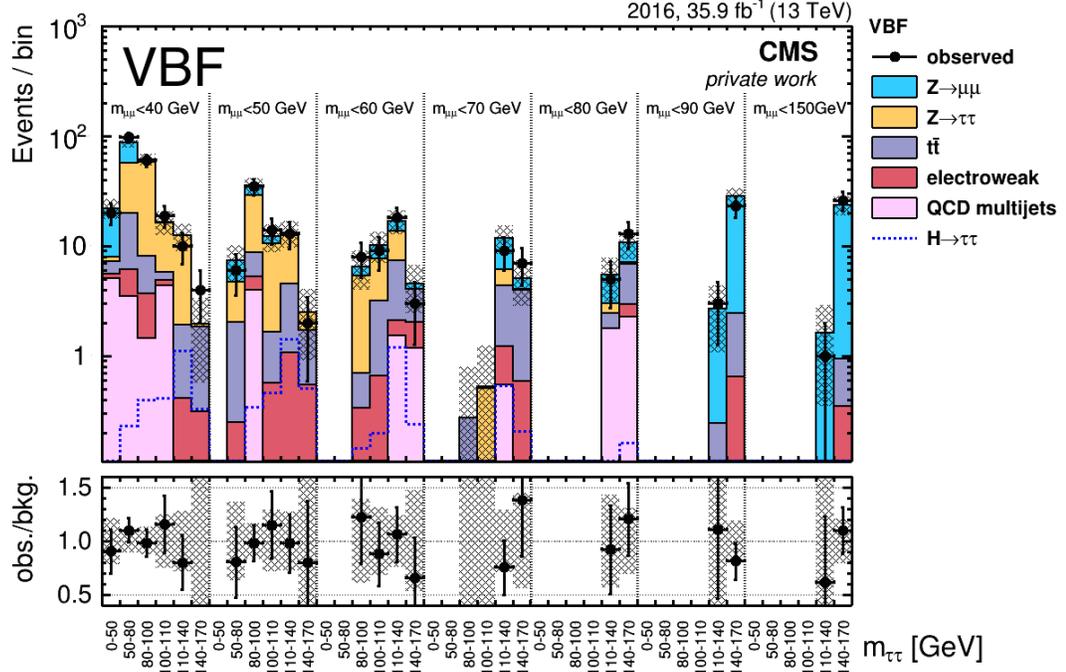


Figure 7.34: The post-fit 2D mass distribution for the VBF category.

Using the maximum likelihood fits of the signal and background to the data, the signal strength for Higgs production in the $\tau_\mu\tau_\mu$ channel can be estimated relative to the SM model cross section expectation times the branching ratio for the given decay channel. In this case the signal strength for the combination of all categories is estimated as -1.0 ± 1.7 . The signal strength for individual categories is tabulated as follows:

Table 7.9: Measurement of $H \rightarrow \tau\tau \rightarrow \mu\mu$ signal strength relative to the expected SM cross section times branching fraction.

Category	Best Fits
0-jet	12.1 ± 13.0
boosted	-2.28 ± 2.81
VBF	0.41 ± 1.92
Combination	-1.05 ± 1.67

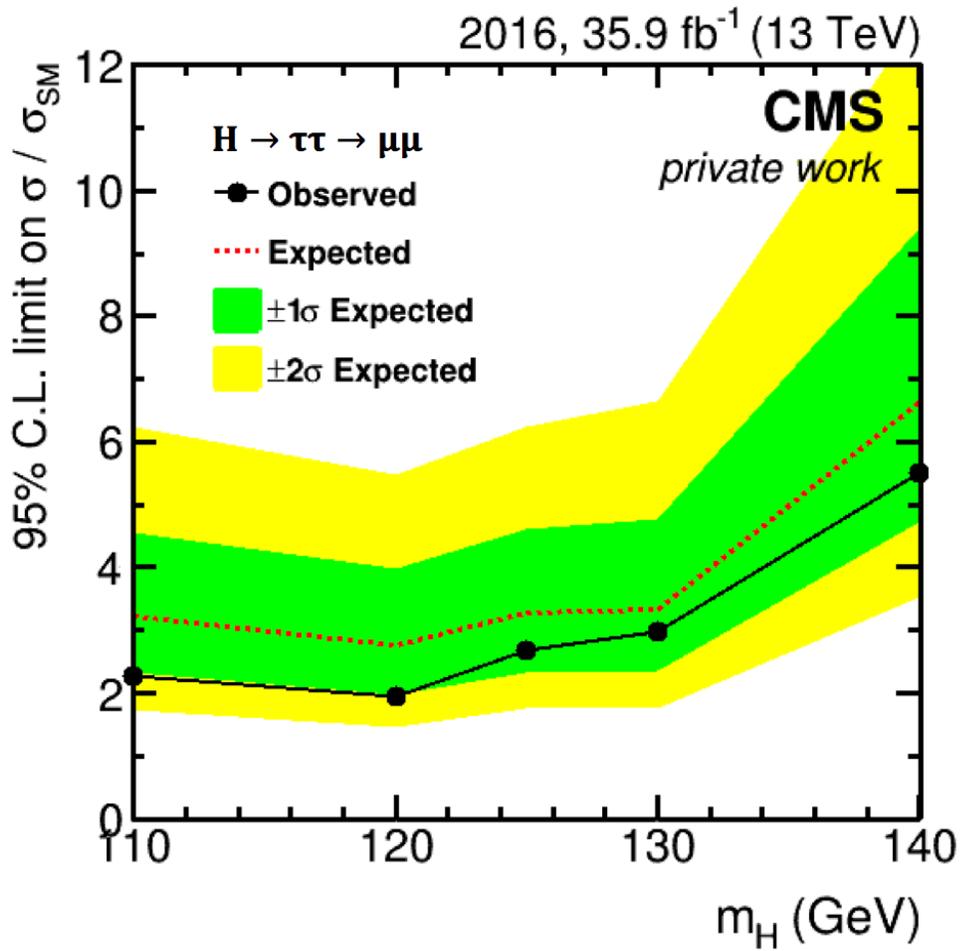


Figure 7.35: The expected and the observed limits on the signal strength for different Higgs masses

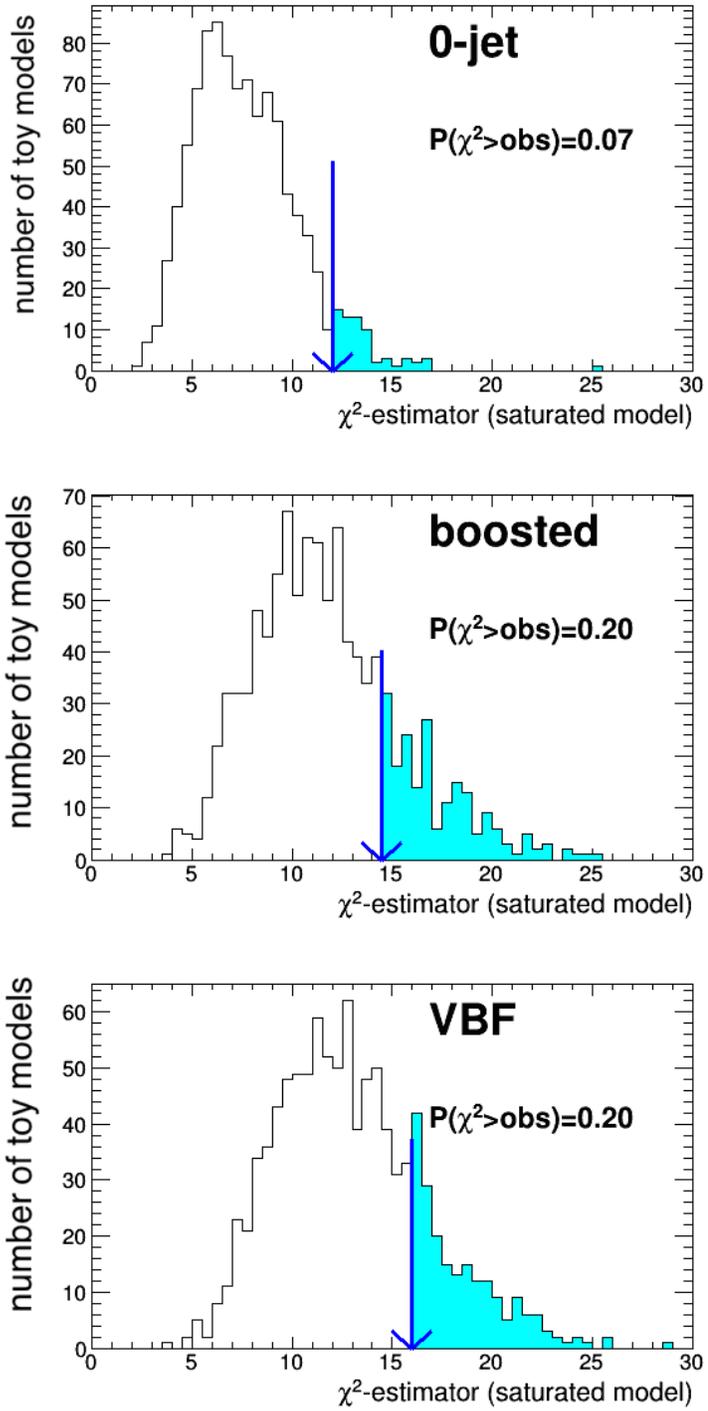


Figure 7.36: GOF using the saturated model for the VBF category.

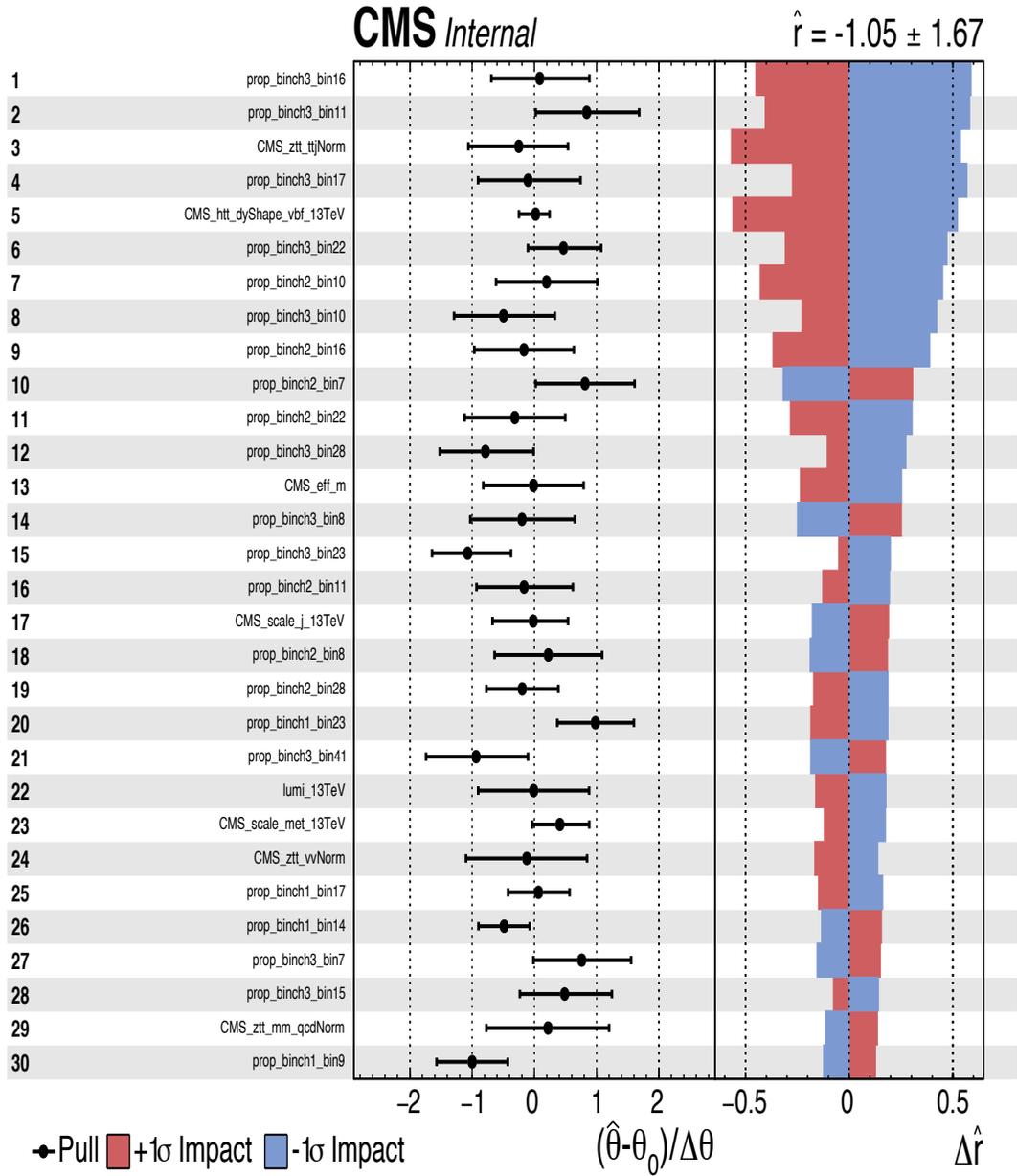


Figure 7.37: Pulls and impacts of first 30 nuisance parameters for combined categories.

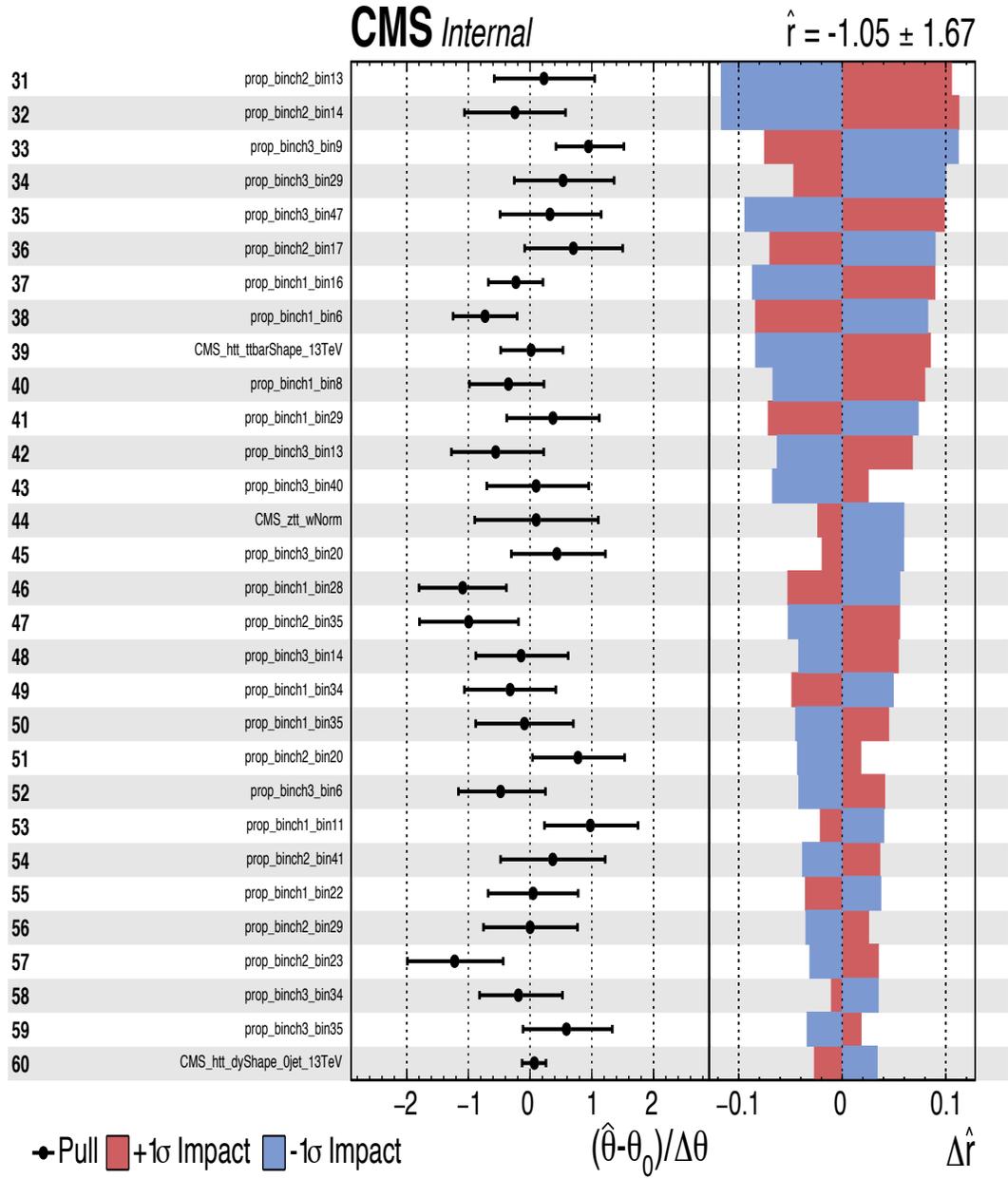


Figure 7.38: Pulls and impacts of second 30 nuisance parameters for combined categories.

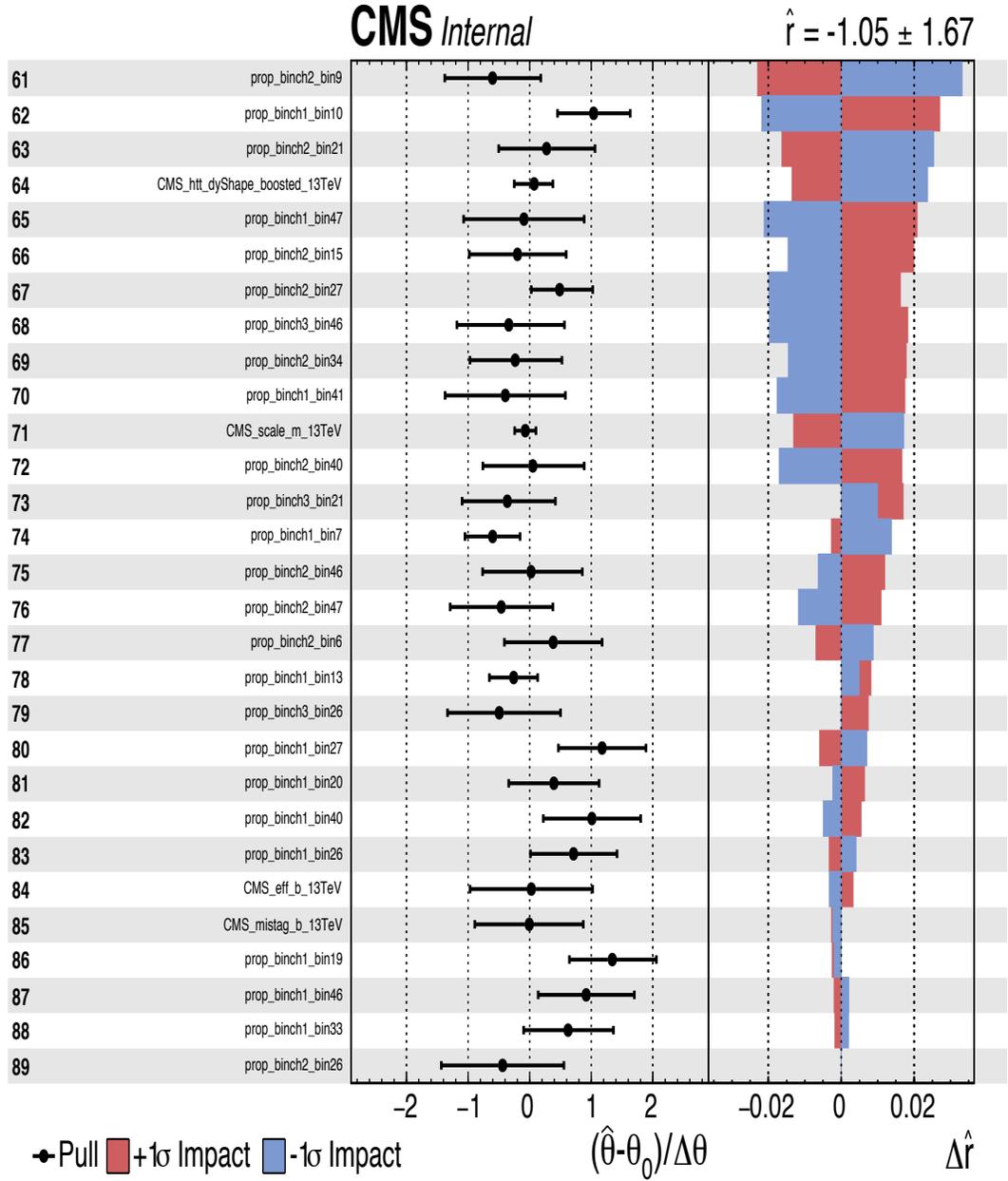
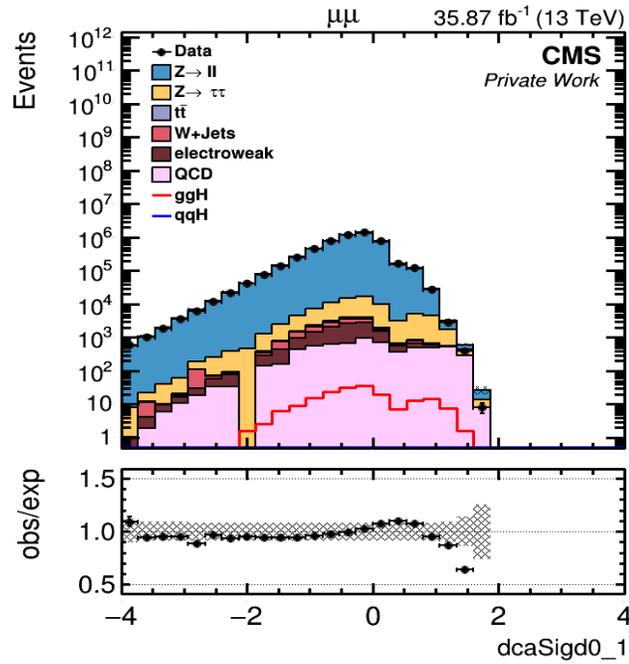


Figure 7.39: Pulls and impacts of nuisance parameters 61-89 for combined categories.

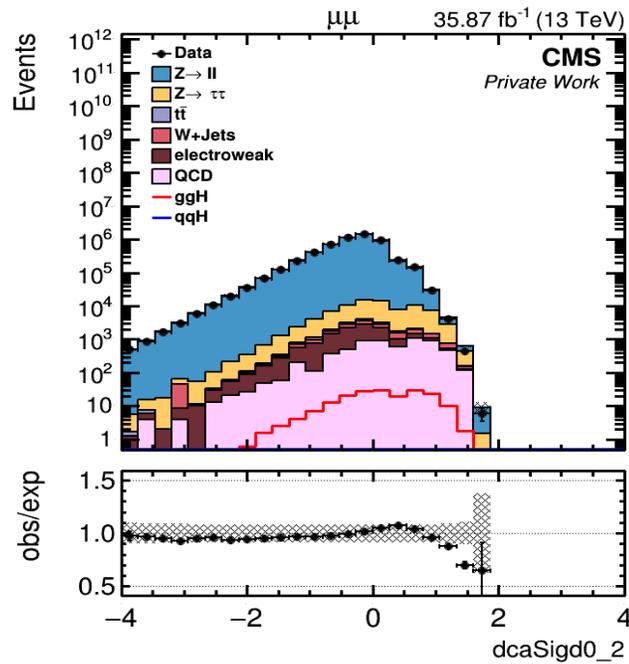
7.10 Future Work

The results of the $H \rightarrow \tau\tau \rightarrow \mu\mu$ analysis using Run II data are comparable with Run I results [101]. Compared to the Run I analysis procedure, in Run II, two major changes occurred. The first change is the exclusion of the distance of closet approach (DCA) parameters. In Run II, the misalignment in the tracker detectors affects the measurement of the impact parameters. The mismodeling of these variables in the data causes a high discrepancy in the data/MC agreement; this discrepancy is propagated in the measurement of the DCA parameter. Figure ?? shows this distribution for the 0-jet category events in the transverse plane and Figure 7.41 shows distribution in the longitudinal plane. In this analysis, the QCD multijet background is derived from the data events. DCA parameters are crucial parameters in this analysis, especially the d_z that is defined in the direction of the beam. They have the highest discriminating power to separate the Higgs decay from the Z decay due to their kinematics; this eventually improves the overall expected limit. Due to the mismodeling of the DCA variable in the MC, this variable is excluded from the analysis. This affects the overall expected and observed limits. To quantify this effect, the multivariate analysis is performed using the mismodeled impact variables and the limits are extracted using the same method to quantify the potential improvement. With the mismodeled parameters, the limits improve by about 10-15%. In conclusion, the result of this analysis can be improved in the future by using the corrected impact parameters in the multivariate analysis.

The second difference is the approach used for the multivariate analysis. In the Run II analysis, the single-BDT method, is used, whereas in Run I, two separate BDT's were used. In the double BDT approach, the first BDT is eval-

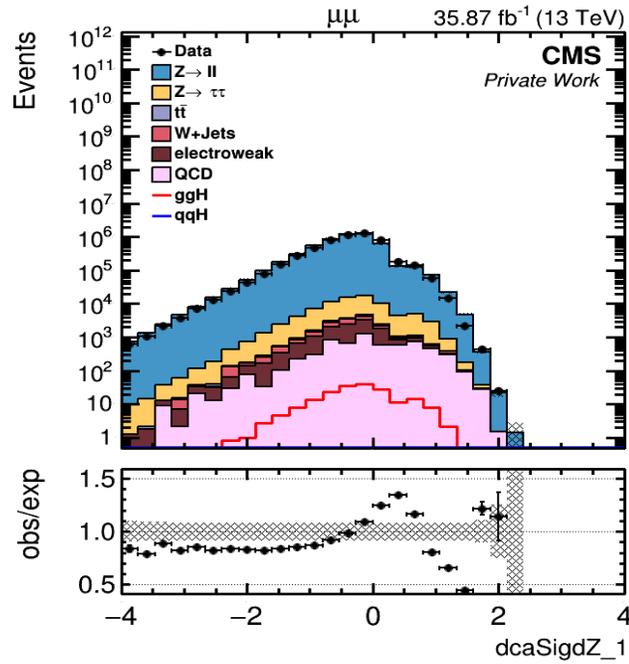


(a) Leading muon distribution

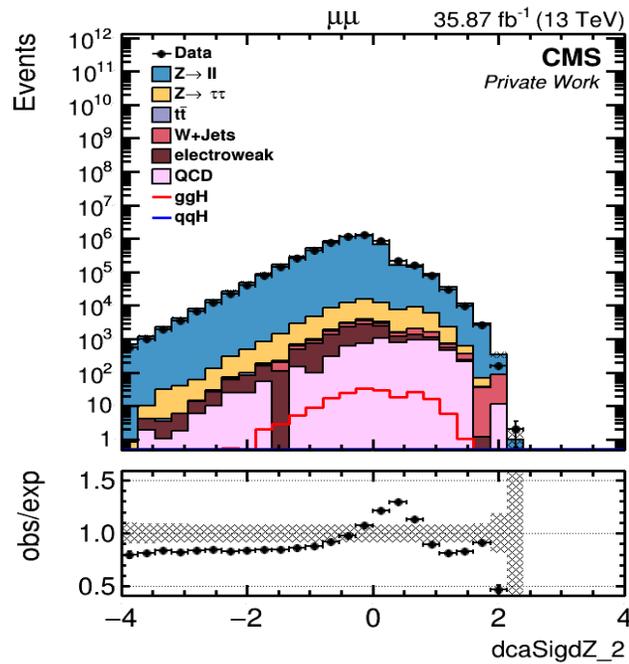


(b) Sub-leading muon distribution

Figure 7.40: DCA significance of the two muon tracks in the transverse plane in the data and simulated MC events for the 0-jet category.



(a) Leading muon distribution



(b) Sub-leading muon distribution

Figure 7.41: DCA significance of the two muon tracks in the longitudinal plane in the data and simulated MC events for the 0-jet category.

uated by training the Higgs samples against the $Z \rightarrow \mu\mu$ background. For the second BDT, the Higgs sample selected by the first BDT is trained against the $Z \rightarrow \tau\tau$ samples. A new discriminant is constructed by combining these two BDT variables. Double BDT approach improves the observed significance for Run II analysis but shows statistical constrained while obtaining the pulls and impact. To have a robust method, one needs to optimize these two BDT's simultaneously. This procedure is time consuming and due to the time constraint it was not possible to do this in this dissertation.

Chapter 8

Summary

This doctoral dissertation focuses on a CMS hardware upgrade for the HL-LHC as well as a physics analysis with Run II 2015-2016 CMS data. For the Phase II upgrade of the CMS muon endcap, GEM technology is selected and the installation of the GE1/1 station will start during LS2 in 2019. For the approval of this technology, several test beam efforts took place to thoroughly study the performances and characteristics of the triple-GEM detectors and over a period of time several updated versions of the detectors evolved. This dissertation focuses on the GE1/1-III prototype triple-GEM detector, built at Florida Tech and tested in the Fermilab test beam in 2013. The main studies involved the gain response, charge measurements, and the spatial resolution. The GE1/1-III prototype detector performed very well by providing a detection efficiency greater than 97% for the barycentric and binary hit-position determination methods. Typically, the gas gain of this detector is on the order of 10^4 , which provides a mean expected charge of 37 fC at operating drift voltage 3250 V. The measured charge using the Landau fit is within 25% of the calculated

charge value. These charge measurements are used to determine the dynamic charge range for the VFAT3 electronic readout system of this detector. The estimated charge input range is found to be upto 140 fC based on the extrapolation of the 99th percentile of the individual measured strip charge for the highest strip cluster size (i.e. 4-strip). This detector system will provide precise tracking of muons along with the CSC as it provides a good spatial resolution $\sim 23 \mu\text{rads}$ for the barycentric method and about $\sim 136 \mu\text{rads}$ with the binary method, which is comparable to the estimation of the resolution from the pitch of the strips. The overall spatial resolution is improved after applying a correction for the non-linear strip response. In conclusion, the GE1/1 detector meets the performance expectation for this upgrade.

My contribution towards the GEM upgrade project is as follows: I have assembled a 1-m long GE1/1 prototype III detector at Florida Tech. This is the first long detector built outside the CERN. After successfully assembling this detector, I made all the necessary preparation for the Fermilab beam test that included successfully assembling and testing $10 \text{ cm} \times 10 \text{ cm}$ GEM detectors for tracking purposes. I played an essential role in the Fermilab beam test. This project responsibility included the beam line assembly of the detectors, data taking and analyzing. I presented performance characteristics and tracking results in annual meetings, national, and international conferences such as American Physical Society (APS) and NSS IEEE. The conference proceedings were submitted to the IEEE conference record. The results of this beam test contributed towards the Technical Design Report (TDR) for the GE1/1 GEM upgrade. I have also studied the performance characteristics of this 1-m long detector with zigzag readout designed at Florida Tech by former post-doc Aiwu

Zhang. These results were published in the NIM journal.

The latter part of the dissertation focuses on the Higgs physics analysis, where the SM neutral Higgs boson decays into a pair of τ leptons and further into two muons. This decay channel is especially challenging due to its low branching fraction and very high irreducible DY background. To study this important DY background and to make sure that it has been modeled correctly in the analysis, the DY process and its cross section measurements were studied in detail. The BDT multi-variate method is implemented to suppress the $Z/\gamma^* \rightarrow \mu\mu$ background in the cross section measurements of the $Z/\gamma^* \rightarrow \tau\tau \rightarrow \mu\mu$ channel. The modeling of this background is good practice for the Higgs measurements, as both analyses have similar backgrounds. For this particular channel, the Z cross section times branching ratio is 1967 ± 121 (stat.) ± 92 (syst.) ± 37 (lumi.) pb measured with 2015 Run II CMS data at the center-of-mass energy $\sqrt{s} = 13$ TeV with an integrated luminosity of 2.3 fb^{-1} . The total cross section using all five decay channels: $\tau_e\tau_h$, $\tau_\mu\tau_h$, $\tau_h\tau_h$, $\tau_\mu\tau_\mu$, and $\tau_e\tau_\mu$ is measured as $\sigma(\text{pp} \rightarrow Z/\gamma^* + X) \times \mathcal{B}(Z/\gamma^* \rightarrow \tau\tau) = 1848 \pm 12$ (stat.) ± 57 (syst.) ± 35 (lumi.) pb, which is in good agreement with the SM prediction.

In the overall $Z \rightarrow \tau\tau$ cross section measurement, my contribution was towards the $\tau_\mu\tau_\mu$ channel. I successfully analyzed CMS 2015 Run II data and provided all the necessary measurements to the Higgs group for cross section calculation. A highlight of this analysis is the suppression and precision estimation of the irreducible $Z \rightarrow \mu\mu$ background using BDT multivariate techniques. I contributed in writing the $\tau_\mu\tau_\mu$ sections in the paper and analysis note for this cross section times branching fraction measurements. This paper will

be submitted to EPJC journal on 12 December 2017. For the physics analysis, I initiated the collaboration with the CMS group at DESY, Hamburg.

Finally, the methods developed using the Z cross section measurements were used to estimate the backgrounds in the $H \rightarrow \tau\tau \rightarrow \mu\mu$ decay channel. The measurements in this channel are carried out using the CMS 2016 Run II data at the center-of-mass energy $\sqrt{s} = 13$ TeV with an integrated luminosity of 35.9 fb^{-1} . Again, to suppress the DY background, the BDT multivariate method is used. In this particular case, the BDT was trained using MC Higgs signal samples and DY ($Z/\gamma^* \rightarrow \tau\tau$ and $Z/\gamma^* \rightarrow \mu\mu$) background samples in ratios according to their branching fractions. In this analysis, the signal is extracted independently in three event categories: 0-jets, boosted, and VBF based on the jet multiplicity in the final state. All corrections and BDT responses are evaluated separately for each category. At the end, all categories are combined to place an upper limit on the cross section as well as to determine the signal strength relative to SM cross section times branching fraction for this particular decay channel. For the $H \rightarrow \tau\tau \rightarrow \mu\mu$ channel, the signal strength for a combination of all categories is obtained as -1.0 ± 1.7 with the expected and observed upper limits with 95% CL at 3.2 and 2.7, respectively, with respect to the SM cross section times branching fraction.

I successfully analyzed the CMS 2016 data to obtain the results for the $H \rightarrow \tau\tau \rightarrow \mu\mu$ decay channel. I tested different BDT approaches for gaining sensitivity in this channel. The above results are derived using the “Single-BDT” approach. This complete analysis is performed by me with help of the CMS and DESY $H \rightarrow \tau\tau$ collaboration.

Bibliography

- [1] D. Contardo, M. Klute, J. Mans, L. Silvestris, and J. Butler, “Technical Proposal for the Phase-II Upgrade of the CMS Detector,” Tech. Rep. CERN-LHCC-2015-010, LHCC-P-008, CMS-TDR-15-02, Geneva, 2015.
- [2] L. Evans and P. Bryant, “LHC Machine,” *Journal of Instrumentation*, vol. 3, no. 08, p. S08001, 2008.
- [3] “The Large Hadron Collider,” URL <http://cds.cern.ch/record/1998498>, 2014.
- [4] B. Muratori, “Luminosity and luminous region calculations for the LHC,” Tech. Rep. LHC-PROJECT-NOTE-301, CERN, Geneva, 2002.
- [5] “Experiments,” URL <http://cds.cern.ch/record/1997374>, 2012.
- [6] S. Chatrchyan et al., “The CMS Experiment at the CERN LHC,” *JINST*, vol. 3, p. S08004, doi:10.1088/1748-0221/3/08/S08004, 2008.
- [7] S. Chatrchyan, et al., “Description and performance of track and primary-vertex reconstruction with the CMS tracker,” *JINST*, vol. 9, no. CMS-TRK-11-001, CERN-PH-EP-2014-070. CMS-TRK-11-001, p. P10009. 80 p, URL <https://cds.cern.ch/record/1704291>, 2014.

- [8] CMS Collaboration, The CMS muon project: Technical Design Report, Technical Design Report CMS, CERN, Geneva, URL <https://cds.cern.ch/record/343814>, 1997.
- [9] G. L. Bayatyan et al., CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems, Technical Design Report CMS, URL <https://cds.cern.ch/record/706847>.
- [10] S. Cittolin et al., CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project, Technical Design Report CMS, CERN, Geneva, URL <https://cds.cern.ch/record/578006>, 2002.
- [11] F. Sauli, “GEM: A new concept for electron amplification in gas detectors,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 386, no. 2, pp. 531 – 534, doi:[https://doi.org/10.1016/S0168-9002\(96\)01172-2](https://doi.org/10.1016/S0168-9002(96)01172-2), 1997.
- [12] B. Schmidt, “Microstrip gas chambers: Recent developments, radiation damage and long-term behavior,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 419, no. 2, pp. 230 – 238, doi:[https://doi.org/10.1016/S0168-9002\(98\)00796-7](https://doi.org/10.1016/S0168-9002(98)00796-7), 1998.
- [13] R. De Oliveira and S. Pinto, “Method of manufacturing a gas electron multiplier,” URL <https://www.google.com/patents/US8597490>, uS Patent 8,597,490, 2013.

- [14] A. Colaleo et al., “CMS Technical Design Report for the Muon End-cap GEM Upgrade,” Tech. Rep. CERN-LHCC-2015-012. CMS-TDR-013, 2015.
- [15] F. Sauli, “The gas electron multiplier (GEM): Operating principles and applications,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 805, no. Supplement C, pp. 2 – 24, doi:<https://doi.org/10.1016/j.nima.2015.07.060>, 2016.
- [16] M. Alfonsi, et al., “High-rate particle triggering with triple-GEM detector,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 518, no. 1, pp. 106 – 112, doi:<https://doi.org/10.1016/j.nima.2003.10.035>, 2004.
- [17] B. Ketzer, Q. Weitzel, S. Paul, F. Sauli, and L. Ropelewski, “Performance of triple GEM tracking detectors in the COMPASS experiment,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 535, no. 1, pp. 314 – 318, doi:<https://doi.org/10.1016/j.nima.2004.07.146>, 2004.
- [18] C. Lourenco, “The Phase-2 Upgrade of the CMS Muon Detectors,” Tech. Rep. CERN-LHCC-2017-012. CMS-TDR-016, CERN, Geneva, URL <http://cds.cern.ch/record/2283189>, 2017.

- [19] A. Sharma, “Detection of single electrons emitted by internal photocathodes with the Gas Electron Multiplier (GEM),” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 462, no. 3, pp. 603 – 606, doi: [https://doi.org/10.1016/S0168-9002\(00\)01310-3](https://doi.org/10.1016/S0168-9002(00)01310-3), 2001.
- [20] J. Toledo, et al., “The Front-End Concentrator card for the RD51 Scalable Readout System,” *Journal of Instrumentation*, vol. 6, no. 11, p. C11028, URL <http://stacks.iop.org/1748-0221/6/i=11/a=C11028>, 2011.
- [21] A. Zhang, et al., “Performance of a large-area GEM detector read out with wide radial zigzag strips,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 811, pp. 30 – 41, 2016.
- [22] S. Costantini et al., “Radiation background with the CMS RPCs at the LHC,” *JINST*, vol. 10, no. 05, p. C05031, doi:10.1088/1748-0221/10/05/C05031, 2015.
- [23] D. Abbaneo, et al., “Studies on the upgrade of the muon system in the forward region of the CMS experiment at LHC with GEMs,” *Journal of Instrumentation*, vol. 9, no. 01, p. C01053, URL <http://stacks.iop.org/1748-0221/9/i=01/a=C01053>, 2014.
- [24] D. Abbaneo, et al., “Status of the Triple-GEM project for the upgrade of the CMS Muon System,” *Journal of Instrumentation*, vol. 8, no. 12, p. C12031, 2013.

- [25] S. Colafranceschi and M. Hohlmann, “Beam test results for new full-scale GEM prototypes for a future upgrade of the CMS high- η Muon System,” in 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC), pp. 1172–1176, 2012.
- [26] K. Gnanvo et al., “Detection and imaging of high-Z materials with a muon tomography station using GEM detectors,” in IEEE Nuclear Science Symposium Medical Imaging Conference, pp. 552–559, 2010.
- [27] K. Gnanvo, “Calibration of the gain and measurement of the noise for the APV25 electronics,” URL <http://hallaweb.jlab.org/12GeV/SuperBigBite/SBS-minutes/2012/apvGain20121107.pdf>, 2012.
- [28] R. Carnegie et al., “Resolution studies of cosmic-ray tracks in a TPC with GEM readout,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 538, no. 1–3, pp. 372 – 383, 2005.
- [29] D. Arogancia et al., “Study in a beam test of the resolution of a Micromegas TPC with standard readout pads,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 602, no. 2, pp. 403 – 414, 2009.
- [30] A. Zhang and M. Hohlmann, “Accuracy of the geometric-mean method for determining spatial resolutions of tracking detectors in the presence of multiple Coulomb scattering,” *Journal of Instrumentation*, vol. 11, no. 06, p. P06012, URL <http://stacks.iop.org/1748-0221/11/i=06/a=P06012>, 2016.

- [31] G. Landi, “Properties of the center of gravity as an algorithm for position measurements,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 485, no. 3, pp. 698 – 719, 2002.
- [32] P. Aspell et al., “The VFAT Production Test Platform for the TOTEM Experiment,” URL <https://cds.cern.ch/record/1159891/files/p544.pdf>, 2008.
- [33] L. Jones, APV25-S1: User guide version 2.2, RAL Microelectronics Design Group, Chilton, URL <https://cds.cern.ch/record/1069892>, 2001.
- [34] M. Villa, “Developing and evaluating new micropattern gas detectors,” CERN-THESIS-2013-284, Bonn U., 2013.
- [35] D. J. Griffiths, Introduction to elementary particles; 2nd ed., Physics textbook, Wiley, New York, NY, 2008.
- [36] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons,” *Phys. Rev. Lett.*, vol. 13, pp. 508–509, doi:10.1103/PhysRevLett.13.508, URL <https://link.aps.org/doi/10.1103/PhysRevLett.13.508>, 1964.
- [37] C. Caillol, B. Clerbaux, and A. Mohammadi, “Scalar boson decays to tau leptons: in the standard model and beyond,” URL <https://cds.cern.ch/record/2222392>, 2016.

- [38] Castillo and L. R. Flores, “Preparing for the search at the LHC,” in *The Search and Discovery of the Higgs Boson*, 2053-2571, pp. 5–1 to 5–15, Morgan & Claypool Publishers, doi:10.1088/978-1-6817-4078-2ch5, URL <http://dx.doi.org/10.1088/978-1-6817-4078-2ch5>, 2015.
- [39] X. Poveda, “Evidence for $t\bar{t}H$ Production with ATLAS Detector,” URL https://indico.cern.ch/event/656771/attachments/1542665/2426952/ATLAS_ttH_seminar.pdf.
- [40] LHC Higgs Cross Section Working Group, URL <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG>.
- [41] A. Djouadi, “The anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model,” *Phys. Rept.*, vol. 457, pp. 1–216, doi:10.1016/j.physrep.2007.10.004, 2008.
- [42] P. Connor, L. Favart, and T. Ševa, “Study of the Drell-Yan process at high invariant mass at the LHC at $\sqrt{s} = 8$ TeV.” URL <https://cds.cern.ch/record/2291027>, 2014.
- [43] I. R. Kenyon, “The Drell-Yan process,” *Reports on Progress in Physics*, vol. 45, no. 11, p. 1261, URL <http://stacks.iop.org/0034-4885/45/i=11/a=002>, 1982.
- [44] S. D. Drell and T.-M. Yan, “Massive Lepton-Pair Production in Hadron-Hadron Collisions at High Energies,” *Phys. Rev. Lett.*, vol. 25, pp. 316–320, doi:10.1103/PhysRevLett.25.316, URL <https://link.aps.org/doi/10.1103/PhysRevLett.25.316>, 1970.

- [45] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and E_T^{miss} ,” CMS Physics Analysis Summary CMS-PAS-PFT-09-001, URL <http://cdsweb.cern.ch/record/1194487>, 2009.
- [46] 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, URL <http://cdsweb.cern.ch/record/1247373>, 2010.
- [47] CMS Collaboration, “Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV,” CMS Physics Analysis Summary CMS-PAS-PFT-10-002, URL <http://cds.cern.ch/record/1279341>, 2010.
- [48] CMS Collaboration, “Particle-flow commissioning with muons and electrons from J/Psi and W events at 7 TeV,” CMS Physics Analysis Summary CMS-PAS-PFT-10-003, URL <http://cds.cern.ch/record/1279347>, 2010.
- [49] CMS collaboration, “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV,” *Journal of Instrumentation*, vol. 7, no. 10, p. P10002, URL <http://stacks.iop.org/1748-0221/7/i=10/a=P10002>, 2012.
- [50] CMS Collaboration, “Performance of CMS muon reconstruction in cosmic-ray events,” *Journal of Instrumentation*, vol. 5, no. 03, p. T03022, URL <http://stacks.iop.org/1748-0221/5/i=03/a=T03022>, 2010.

- [51] 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*, vol. 262, no. 2, pp. 444 – 450, doi:[https://doi.org/10.1016/0168-9002\(87\)90887-4](https://doi.org/10.1016/0168-9002(87)90887-4), URL <http://www.sciencedirect.com/science/article/pii/0168900287908874>, 1987.
- [52] A. M. Sirunyan et al., “Particle-flow reconstruction and global event description with the CMS detector,” *JINST*, vol. 12, no. 10, p. P10003, doi:10.1088/1748-0221/12/10/P10003, 2017.
- [53] CMS Collaboration, “Performance of τ -lepton reconstruction and identification in CMS,” *Journal of Instrumentation*, vol. 7, no. 01, p. P01001, URL <http://stacks.iop.org/1748-0221/7/i=01/a=P01001>, 2012.
- [54] CMS collaboration, “Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS,” *Journal of Instrumentation*, vol. 11, no. 01, p. P01019, URL <http://stacks.iop.org/1748-0221/11/i=01/a=P01019>, 2016.
- [55] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm,” *JHEP*, vol. 04, p. 063, doi:10.1088/1126-6708/2008/04/063, 2008.
- [56] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet user manual,” *Eur. Phys. J. C*, vol. 72, p. 1896, doi:10.1140/epjc/s10052-012-1896-2, 2012.
- [57] M. Cacciari and G. P. Salam, “Dispelling the N^3 myth for the k_t jet-finder,” *Phys. Lett. B*, vol. 641, p. 57, doi:10.1016/j.physletb.2006.08.037, 2006.

- [58] S. Chatrchyan et al., “Determination of jet energy calibration and transverse momentum resolution in CMS,” *JINST*, vol. 6, p. 11002, doi:10.1088/1748-0221/6/11/P11002, 2011.
- [59] CMS Collaboration, “Jet Identification,” URL https://twiki.cern.ch/twiki/bin/view/CMS/JetID#Recommendations_for_13_TeV_data.
- [60] V. Khachatryan et al., “Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV,” *JINST*, vol. 10, no. 02, p. P02006, doi:10.1088/1748-0221/10/02/P02006, 2015.
- [61] L. Bianchini, J. Conway, E. K. Friis, and C. Veelken, “Reconstruction of the Higgs mass in $H \rightarrow \tau\tau$ Events by Dynamical Likelihood techniques,” *Journal of Physics: Conference Series*, vol. 513, no. 2, p. 022035, URL <http://stacks.iop.org/1742-6596/513/i=2/a=022035>, 2014.
- [62] K. Kondo, “Dynamical Likelihood Method for Reconstruction of Events with Missing Momentum. I. Method and Toy Models,” *Journal of the Physical Society of Japan*, vol. 57, no. 12, pp. 4126–4140, doi:10.1143/JPSJ.57.4126, URL <https://doi.org/10.1143/JPSJ.57.4126>, 1988.
- [63] K. Kondo, T. Chikamatsu, and S.-H. Kim, “Dynamical Likelihood Method for Reconstruction of Events with Missing Momentum. III. Analysis of a CDF High p_T $e\mu$ Event as $t\bar{t}$ Production,” *Journal of the Physical Society of Japan*, vol. 62, no. 4, pp. 1177–1182, doi:10.1143/JPSJ.62.1177, URL <https://doi.org/10.1143/JPSJ.62.1177>, 1993.
- [64] A. Hoecker, et al., “TMVA: Toolkit for Multivariate Data Analysis,” *PoS*, vol. ACAT, p. 040, 2007.

- [65] I. Antcheva, et al., “ROOT — A C++ framework for petabyte data storage, statistical analysis and visualization,” *Computer Physics Communications*, vol. 180, no. 12, pp. 2499 – 2512, doi:<https://doi.org/10.1016/j.cpc.2009.08.005>, URL <http://www.sciencedirect.com/science/article/pii/S0010465509002550>, 2009.
- [66] ATLAS and CMS Collaborations and LHC Higgs Combination Group, “Procedure for the LHC Higgs boson search combination in Summer 2011,” Tech. Rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, URL <https://cds.cern.ch/record/1379837>, 2011.
- [67] ATLAS and CMS Collaborations and LHC Higgs Combination Group, “Procedure for the LHC Higgs boson search combination in summer 2011,” Tech. Rep. ATL-PHYS-PUB-2011-011, CMS-NOTE-2011-005, URL <https://cds.cern.ch/record/1379837>, 2011.
- [68] J. Conway, “Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra,” , no. arXiv:1103.0354, pp. 115–120. 6 p, URL <https://cds.cern.ch/record/1333496>, 2011.
- [69] A. L. Read, “Linear interpolation of histograms,” *Nucl. Instrum. Meth.*, vol. A425, pp. 357–360, doi:10.1016/S0168-9002(98)01347-3, 1999.
- [70] R. Barlow and C. Beeston, “Fitting using finite Monte Carlo samples,” *Computer Physics Communications*, vol. 77, no. 2, pp. 219 – 228, doi:[https://doi.org/10.1016/0010-4655\(93\)90005-W](https://doi.org/10.1016/0010-4655(93)90005-W), URL <http://www.sciencedirect.com/science/article/pii/001046559390005W>, 1993.

- [71] A. L. Read, “Presentation of search results: The CL(s) technique,” *J. Phys.*, vol. G28, pp. 2693–2704, doi:10.1088/0954-3899/28/10/313, 2002.
- [72] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics,” *Eur. Phys. J.*, vol. C71, p. 1554, doi:10.1140/epjc/s10052-011-1554-0,10.1140/epjc/s10052-013-2501-z, 2011.
- [73] R. Cousins, “Generalization of chisquare goodness-of-fit test for binned data using saturated models, with application to histograms,” URL http://www.physics.ucla.edu/~cousins/stats/cousins_saturated.pdf.
- [74] S. Chatrchyan et al., “Measurement of the inclusive Z cross section via decays to τ pairs in pp collisions at $\sqrt{s} = 7$ TeV,” *JHEP*, vol. 08, p. 117, doi:10.1007/JHEP08(2011)117, 2011.
- [75] G. Aad et al., “Measurement of the Z $\rightarrow \tau\tau$ cross section with the ATLAS Detector,” *Phys. Rev.*, vol. D 84, p. 112006, doi:10.1103/PhysRevD.84.112006, 2011.
- [76] Software guide PhysicsTools, “How to work with files for Good Luminosity Sections in JSON format,” URL <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideGoodLumiSectionsJSONFile>.
- [77] CMS Collaboration, “JetMET POG,” URL <https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFilters>.
- [78] W. Erdmann and PSI, “Offline Primary Vertex Reconstruction with Deterministic Annealing Clustering,” URL <https://cds.cern.ch/record/1365063>, 2011.

- [79] CMS Collaboration, “Generic Tag and Probe tool for measuring efficiency at CMS with early data,” 2009.
- [80] V. Bhopatkar et al., “Validation of analysis techniques relevant for studies of τ lepton production and application to measurement of the $Z/\gamma^* \rightarrow \tau\tau$ cross section in pp collisions at $\sqrt{s} = 13$ TeV,” CMS Analysis Note CMS-AN-16-027, 2017.
- [81] G. Bauer et al., “Modeling of $W \rightarrow \ell\nu E_T^{\text{miss}}$ with boson recoil,” *CMS Analysis Note*, URL http://cms.cern.ch/iCMS/jsp/openfile.jsp?tp=draft&files=AN2010_332_v2.pdf, 2010.
- [82] BTV POG, “Usage of b Tag Objects for 13 TeV Data with 25ns bunch spacing and 76X ReReco,” URL <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation76X>.
- [83] C. Weiser, “A Combined Secondary Vertex Based B-Tagging Algorithm in CMS,” Tech. Rep. CMS-NOTE-2006-014, CERN, Geneva, URL <https://cds.cern.ch/record/927399>, 2006.
- [84] A. Bodek, A. Khukhunaishvili, and J. Han, “Rochester Coorection in 76X reprocessing,” URL https://indico.cern.ch/event/507695/contributions/1186900/attachments/1241153/1825193/rochcor_run2_MuonPOG_030916.pdf.
- [85] CMS Collaboration, “Performance of missing energy reconstruction in 13 TeVpp collision data using the CMS detector,” CMS Physics Analysis Summary CMS-PAS-JME-16-004, URL <http://cds.cern.ch/record/2205284>, 2016.

- [86] A. M. Sirunyan et al., “Observation of the Higgs boson decay to a pair of τ leptons,” URL <https://cds.cern.ch/record/2276465>, submitted to Phys. Lett. B. All figures and tables can be found at <http://cms-results.web.cern.ch/cms-results/public-results/publications/HIG-16-043>, 2017.
- [87] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi, “The $t\bar{t}$ cross-section at 1.8 TeV and 1.96 TeV: a study of the systematics due to parton densities and scale dependence,” *JHEP*, vol. 04, p. 068, doi:10.1088/1126-6708/2004/04/068, 2004.
- [88] S. Catani, D. de Florian, M. Grazzini, and P. Nason, “Soft gluon resummation for Higgs boson production at hadron colliders,” *JHEP*, vol. 07, p. 028, doi:10.1088/1126-6708/2003/07/028, 2003.
- [89] J. Butterworth et al., “PDF4LHC recommendations for LHC Run II,” *J. Phys. G*, vol. 43, p. 023001, doi:10.1088/0954-3899/43/2/023001, 2016.
- [90] CMS Collaboration, “CMS luminosity measurement for the 2015 data-taking period,” CMS Physics Analysis Summary CMS-PAS-LUM-15-001, URL <http://cds.cern.ch/record/2138682>, 2015.
- [91] CMS Collaboration, “Measurement of the $Z/\gamma^* \rightarrow \tau\tau$ cross section in proton-proton collisions at $\sqrt{s} = 13$ TeV and validation of analysis techniques relevant for studies of $\tau\tau$ lepton production,” Submitting to EPJC, 2017.

- [92] V. Khachatryan et al., “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV,” *JINST*, vol. 12, no. 02, p. P02014, doi:10.1088/1748-0221/12/02/P02014, 2017.
- [93] Particle Data Group, “Review of particle physics,” *Chin. Phys.*, vol. C 38, p. 090001, doi:10.1088/1674-1137/38/9/090001, 2014.
- [94] S. Chatrchyan et al., “Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV,” *Phys. Lett. B*, vol. 710, p. 26, doi:10.1016/j.physletb.2012.02.064, 2012.
- [95] V. Khachatryan et al., “Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV,” *Eur. Phys. J. C*, vol. 75, no. 5, p. 212, doi:10.1140/epjc/s10052-015-3351-7, 2015.
- [96] Y. Li and F. Petriello, “Combining QCD and electroweak corrections to dilepton production in FEWZ,” *Phys. Rev. D*, vol. 86, p. 094034, doi:10.1103/PhysRevD.86.094034, 2012.
- [97] S. Chatrchyan et al., “Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV,” *JHEP*, vol. 06, p. 081, doi:10.1007/JHEP06(2013)081, 2013.

- [98] G. Aad, T. Abajyan, B. Abbott, J. A. S. A. Khalek, and et. al., “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC,” *Physics Letters B*, vol. 716, no. 1, pp. 1 – 29, doi:<https://doi.org/10.1016/j.physletb.2012.08.020>, URL <http://www.sciencedirect.com/science/article/pii/S037026931200857X>, 2012.
- [99] S. Chatrchyan, et al., “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,” *Physics Letters B*, vol. 716, no. 1, pp. 30 – 61, doi:<https://doi.org/10.1016/j.physletb.2012.08.021>, URL <http://www.sciencedirect.com/science/article/pii/S0370269312008581>, 2012.
- [100] G. Aad et al., “Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments,” *Phys. Rev. Lett.*, vol. 114, p. 191803, doi:[10.1103/PhysRevLett.114.191803](https://doi.org/10.1103/PhysRevLett.114.191803), 2015.
- [101] S. Chatrchyan et al., “Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons,” *JHEP*, vol. 05, p. 104, doi:[10.1007/JHEP05\(2014\)104](https://doi.org/10.1007/JHEP05(2014)104), 2014.
- [102] G. Aad et al., “Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV,” *JHEP*, vol. 08, p. 045, doi:[10.1007/JHEP08\(2016\)045](https://doi.org/10.1007/JHEP08(2016)045), 2016.

- [103] A. M. Sirunyan et al., “Observation of the Higgs boson decay to a pair of τ leptons,” , no. CMS-HIG-16-043. CMS-HIG-16-043-003, URL <https://cds.cern.ch/record/2276465>, submitted to Phys. Lett. B. All figures and tables can be found at <http://cms-results.web.cern.ch/cms-results/public-results/publications/HIG-16-043>, 2017.
- [104] P. Nason, “A new method for combining NLO QCD with shower Monte Carlo algorithms,” *JHEP*, vol. 11, p. 040, doi:10.1088/1126-6708/2004/11/040, 2004.
- [105] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with parton shower simulations: the POWHEG method,” *JHEP*, vol. 11, p. 070, doi:10.1088/1126-6708/2007/11/070, 2007.
- [106] S. Alioli, P. Nason, C. Oleari, and E. Re, “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX,” *JHEP*, vol. 06, p. 043, doi:10.1007/JHEP06(2010)043, 2010.
- [107] S. Alioli, K. Hamilton, P. Nason, C. Oleari, and E. Re, “Jet pair production in POWHEG,” *JHEP*, vol. 04, p. 081, doi:10.1007/JHEP04(2011)081, 2011.
- [108] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO Higgs boson production via gluon fusion matched with shower in POWHEG,” *JHEP*, vol. 04, p. 002, doi:10.1088/1126-6708/2009/04/002, 2009.

- [109] D. de Florian et al., “Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector,” CERN Report CERN-2017-002-M, doi:10.23731/CYRM-2017-002, 2016.
- [110] A. Denner, S. Heinemeyer, I. Puljak, D. Rebuzzi, and M. Spira, “Standard model Higgs-boson branching ratios with uncertainties,” *Eur. Phys. J. C*, vol. 71, p. 1753, doi:10.1140/epjc/s10052-011-1753-8, 2011.
- [111] R. D. Ball et al., “Impact of heavy quark masses on parton distributions and LHC phenomenology,” *Nucl. Phys. B*, vol. 849, p. 296, doi:10.1016/j.nuclphysb.2011.03.021, 2011.
- [112] J. Alwall, et al., “Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions,” *Eur. Phys. J. C*, vol. 53, p. 473, doi:10.1140/epjc/s10052-007-0490-5, 2008.
- [113] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO,” *JHEP*, vol. 12, p. 061, doi:10.1007/JHEP12(2012)061, 2012.
- [114] T. Sjostrand, et al., “An introduction to PYTHIA 8.2,” *Comput. Phys. Commun.*, vol. 191, p. 159, doi:10.1016/j.cpc.2015.01.024, 2015.
- [115] V. Khachatryan et al., “Event generator tunes obtained from underlying event and multiparton scattering measurements,” *Eur. Phys. J. C*, vol. 76, p. 155, doi:10.1140/epjc/s10052-016-3988-x, 2016.
- [116] S. Agostinelli et al., “GEANT4 — a simulation toolkit,” *Nucl. Instrum. Meth. A*, vol. 506, p. 250, doi:10.1016/S0168-9002(03)01368-8, 2003.

- [117] I. Ojalvo et al., “Search for the standard model Higgs boson decaying to a pair of tau leptons in 2016 data,” CMS Analysis Note CMS-AN-16-355, 2017.
- [118] The H $\rightarrow \tau\tau$, “Theoretical Uncertainties on Higgs Production via VBF and Gluon Fusion Processes for the H $\rightarrow \tau\tau$ Analysis at $\sqrt{s} = 13$ TeV,” CMS Analysis Note CMS-AN-16-387, 2017.