

ABSTRACT

**1 STUDIES OF VECTOR BOSON SCATTERING IN THE SEMILEPTONIC
2 ZV CHANNEL WITH THE CMS DETECTOR AT $\sqrt{s} = 13$ TEV**

3 Ramanpreet Singh, Ph.D.

4 Department of Physics

5 Northern Illinois University, 2022

6 Vishnu Zutshi, Director

7 About vector boson scattering (VBS)

NORTHERN ILLINOIS UNIVERSITY
DE KALB, ILLINOIS

SUMMER 2022

**STUDIES OF VECTOR BOSON SCATTERING IN THE SEMILEPTONIC
ZV CHANNEL WITH THE CMS DETECTOR AT $\sqrt{s} = 13$ TEV**

BY

RAMANPREET SINGH
© 2022 Ramanpreet Singh

A DISSERTATION SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS

Dissertation Director:
Vishnu Zutshi

ACKNOWLEDGEMENTS

10 Acknowledge people here.

DEDICATION

To my family.

TABLE OF CONTENTS

	Page
¹³ LIST OF TABLES	vii
¹⁴ LIST OF FIGURES	viii
¹⁵ LIST OF APPENDICES	xi
 Chapter	
¹⁶ 1 INTRODUCTION	1
¹⁷ 1.1 Standard model	2
¹⁸ 1.1.1 Quantum Electrodynamics (QED)	4
¹⁹ 1.1.2 Quantum Chromodynamics (QCD)	5
²⁰ 1.1.3 Electroweak Theory	6
²¹ 1.1.4 Electroweak Symmetry Breaking and Higgs Mechanism	8
²² 1.2 Vector Boson Scattering	11
²³ 1.2.1 Motivation	11
²⁴ 1.2.2 Topology of VBS in Semileptonic ZV Final State	13
²⁵ 2 THE LHC AND CMS EXPERIMENT	15
²⁶ 2.1 The Large Hadron Collider	15
²⁷ 2.1.1 Integrated Luminosity	17
²⁸ 2.2 The CMS Detector	18
²⁹ 2.2.1 The CMS Coordinate System	20
³⁰ 2.2.2 The Superconducting Magnet	20
³¹ 2.2.3 The Tracking System	22

	Page
32 Chapter	Page
33 2.2.4 The Electromagnetic Calorimeter	23
34 2.2.5 The Hadronic Calorimeter	25
35 2.2.6 Muon Detector	26
36 2.2.7 Level 1 Trigger	28
37 3 EVENT SIMULATION AND RECONSTRUCTION	30
38 3.1 Track Reconstruction and Calorimeter Clustering	30
39 3.2 Reconstructed Particles	31
40 3.2.1 Muons	32
41 3.2.2 Electrons and Photons	33
42 3.2.3 Hadrons and Jets	34
43 3.2.3.1 N-Subjetiness and Deep Taggers	36
44 3.2.3.2 Softdrop Mass	38
45 3.2.4 Missing transverse momentum	38
46 3.3 TBD Monte Carlo Simulation	39
47 3.3.1 TBD Generators	39
48 3.3.2 TBD Hadronization	39
49 4 VBS MEASUREMENT IN ZVJJ FINAL STATE	40
50 4.1 Dataset and Simulation	40
51 4.1.1 Data	41
52 4.1.2 MC Simulations	41
53 4.2 Event Selection	47
54 4.2.1 Boosted ZV DY+Jets Control Region	48
55 4.2.2 Resolved ZV DY+Jets Control Region	55

	Chapter		Page
57	4.3	Machine Learning Modeling	63
58	4.3.1	Algorithm: Gradient Boosted Decision Tree	63
59	4.3.2	Training and Results	63
60	4.3.3	MVA Score inference in Signal Region	70
61	4.4	Measurement	71
62	4.4.1	Statistical and Systematic Uncertainties	71
63	4.4.2	Significance	71
64	4.4.3	Impact Plots	72
65	4.4.4	Postfit Plots	75
66	5	HIGH GRANULARITY CALORIMETER UPGRADE	76
67	5.1	Technical Design	76
68	5.2	Scintillator Tiles	77
69	5.2.1	End of Life Scenario	77
70	6	CONCLUSIONS	79
71	REFERENCES		80
72	APPENDICES		85

LIST OF TABLES

	Table	Page
74	2.1 Standard physics luminosity for run-2	17
75	4.1 Trigger paths used to select events in CMS collision data	42
76	4.2 List of MC samples for Signal and Background modeling	44
77	4.3 List of MC samples for Signal and Background modeling	45
78	4.4 List of MC samples for Signal and Background modeling	46
79	4.5 Training and Testing Statistics	64
80	4.6 Training Input Variable Ranking	64
81	4.7 Models Metrics	65
82	4.8 Significance.	71
83	5.1 HGCAL scenarios comparison	78

LIST OF FIGURES

	Figure	Page
85	1.1 Standard model list of matter and interaction particles	3
86	1.2 3D representation of Higgs potential	10
87	1.3 The cross-sections for longitudinal polarized VBS involving three and four- 88 boson coupling with and without Higgs coupling included.	12
89	1.4 Tree level Feynman diagram of ZV VBS process at LHC.	14
90	2.1 A schematic of the CERN accelerator complex.	16
91	2.2 Cumulative delivered and recorded luminosity versus time for 2015–2018 92 proton-proton collisions	17
93	2.3 The CMS detector cutaway view	18
94	2.4 The CMS detector slice view	19
95	2.5 The picture of the CMS detector central part when lowered in underground 96 cavern with superconducting magnet and iron yoke visible	21
97	2.6 The CMS pixel upgrade.	23
98	2.7 The CMS ECAL schematic layout	24
99	2.8 The ECAL endcap quadrant assembled view	24
100	2.9 The first half of the barrel HCAL inserted into the superconducting solenoid 101 (April 2006)	25
102	2.10 The HCAL depth segmentation after phase 1 upgrade	26
103	2.11 The quadrant view of CMS subdetectors layout, and the coverage of the 104 muon detector DTs, CSCs, and RPCs highlighted	27
105	3.1 Comparison of τ_{21} and τ_{32} shapes for signal and background in AK8 jets . . .	37

	Figure	Page
106		
107	3.2 The network architecture of DeepAK8	37
108		
109	4.1 DY+Jets Control Region: Leading electron kinematics in Boosted ZV Channel	48
110		
111	4.2 DY+Jets Control Region: Trailing electron kinematics in Boosted ZV Channel	49
112		
113	4.3 DY+Jets Control Region: Leading muon kinematics in Boosted ZV Channel	50
114		
115	4.4 DY+Jets Control Region: Trailing muon kinematics in Boosted ZV Channel	51
116		
117	4.5 DY+Jets Control Region: Hadronic boson kinematics in Boosted ZV Channel	52
118		
119	4.6 DY+Jets Control Region: Leading VBS tagged jet kinematics in Boosted ZV Channel	53
120		
121	4.7 DY+Jets Control Region: Trailing VBS tagged jet kinematics in Boosted ZV Channel	54
122		
123	4.8 DY+Jets Control Region: Leading electron kinematics in Resolved ZV Channel	55
124		
125	4.9 DY+Jets Control Region: Trailing electron kinematics in Resolved ZV Channel	56
126		
127	4.10 DY+Jets Control Region: Leading muon kinematics in Resolved ZV Channel	57
128		
129	4.11 DY+Jets Control Region: Trailing muon kinematics in Resolved ZV Channel	58
130		
131	4.12 DY+Jets Control Region: Hadronic boson leading jet kinematics in Resolved ZV Channel	59
132		
133	4.13 DY+Jets Control Region: Hadronic boson trailing jet kinematics in Resolved ZV Channel	60
134		
131	4.14 DY+Jets Control Region: Leading VBS tagged jet kinematics in Resolved ZV Channel	61
132		
133	4.15 DY+Jets Control Region: Trailing VBS tagged jet kinematics in Resolved ZV Channel	62
134		
134	4.16 Inputs Variables for training Boosted ZV BDT Classifier	66

	Figure	Page
135		
136	4.17 Inputs Variables for training Resolved ZV BDT Classifier	67
137	4.18 Correlation Matrix for Signal and Background	68
138	4.19 MVA Score ROC Curve	69
139	4.20 MVA Score in Signal Region for Boosted ZV Channel	70
140	4.21 MVA Score in Signal Region for Resolved ZV Channel	70
141	4.22 Impact Plots	72
142	4.23 Impact Plots	73
143	4.24 Impact Plots	74
144	4.25 MVA Score postfit in Signal Region for Boosted ZV Channel	75
145	4.26 MVA Score postfit in Signal Region for Resolved ZV Channel	75
146	5.1 Overview of CE	76
147	5.2 HGCAL scenarios	77

LIST OF APPENDICES

Appendix	Page
¹⁴⁹ A MACHINE LEARNING	85
¹⁵⁰ B DATA ANALYSIS	87

151

CHAPTER 1

152

INTRODUCTION

It doesn't matter how beautiful your theory is, it doesn't matter how smart you are.

If it doesn't agree with experiment, it's wrong. In that simple statement is the key to science.

— Richard P. Feynman

153 “Particle Physics” is the branch of physics which deals with the fundamental particles
154 and the interactions between them. Fundamental particles are the subatomic particles which
155 are not made of other particles. There are two types of fundamental particles “matter” and
156 “interaction” particles, as the name suggests matter particles are the fundamental constituent
157 of matter and the interactions between among them is governed by how they exchange
158 interaction particles. Standard Model (SM) of particle physics is the theory that classifies
159 these fundamental particles and describes three out of four fundamental interaction forces;
160 electromagnetic, weak, and strong.

161 This chapter introduces briefly to the theory of SM, Higgs mechanism, spontaneous
162 electroweak symmetry breaking (EWSB), vector boson scattering (VBS), and the motivation
163 for the search of VBS in semileptonic decay channel ZV with leptonic decay of Z, and hadronic
164 decay of V (W/Z) to pair of quarks.

165

1.1 Standard model

166 In SM, the matter particles are fermions, and the interaction particles are bosons. SM
 167 also includes anti-fermions, which are fermions with equal mass but opposite sign of charge.
 168 Figure 1.1 lists mass, electric charge, and spin of fermions and bosons in SM.

169 Fermions obey Fermi-Dirac statistics and have half integer spin. They can be further
 170 divided into leptons which have integral electric charge, and quarks which have fractional
 171 electric charge. There are three generations of quarks and leptons discovered to the date,
 172 each generation only differs by the mass. In addition to the electric charge, quarks also have
 173 three types of “color” charge (red, green and blue). Quarks cannot be isolated because of
 174 “color confinement”, which requires net color charge to be zero for an isolated particle, for
 175 this reason we can only have certain composition of quarks. Baryons (proton, neutrons, etc.)
 176 are made up of three quarks with each with different color charge, and mesons (pions, kaons,
 177 etc.) are made of two quarks with color and anti-color charge.

178 Bosons obey Bose-Einstein statistics and have integral spin. They are described by
 179 local gauge theory and are also called gauge boson. Photons are the interaction particle of
 180 electromagnetic force, they are massless and only interact with charged particles. Gluons
 181 are the mediator of strong force between quarks, they are massless and carries color charge.
 182 W^\pm and Z are the vector bosons and mediator of weak force, unlike photons and gluons they
 183 are massive. W^+ and W^- are antiparticles of each other, and Z is antiparticle of its own.
 184 The last gauge boson Higgs is a massive scalar boson with zero spin, zero electric and color
 185 charge. Higgs boson is not a force carrier, but rather explains why only some particles have
 186 mass.

187 The SM is built in the framework of quantum field theory (QFT), in which particles
 188 are excitation of the fields and interactions arise from local gauge invariance. The SM is a

¹⁸⁹ $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ gauge theory, $U(1)_Y$, $SU(2)_L$, $SU(3)_C$ are the gauge symmetries
¹⁹⁰ of quantum electrodynamics (QED), weak interaction and quantum chromodynamics (QCD)
¹⁹¹ respectively, where the indices stands for “hypercharge” (Y), “left-handed” (L) and “color”
¹⁹² (C).

Standard Model of Elementary Particles

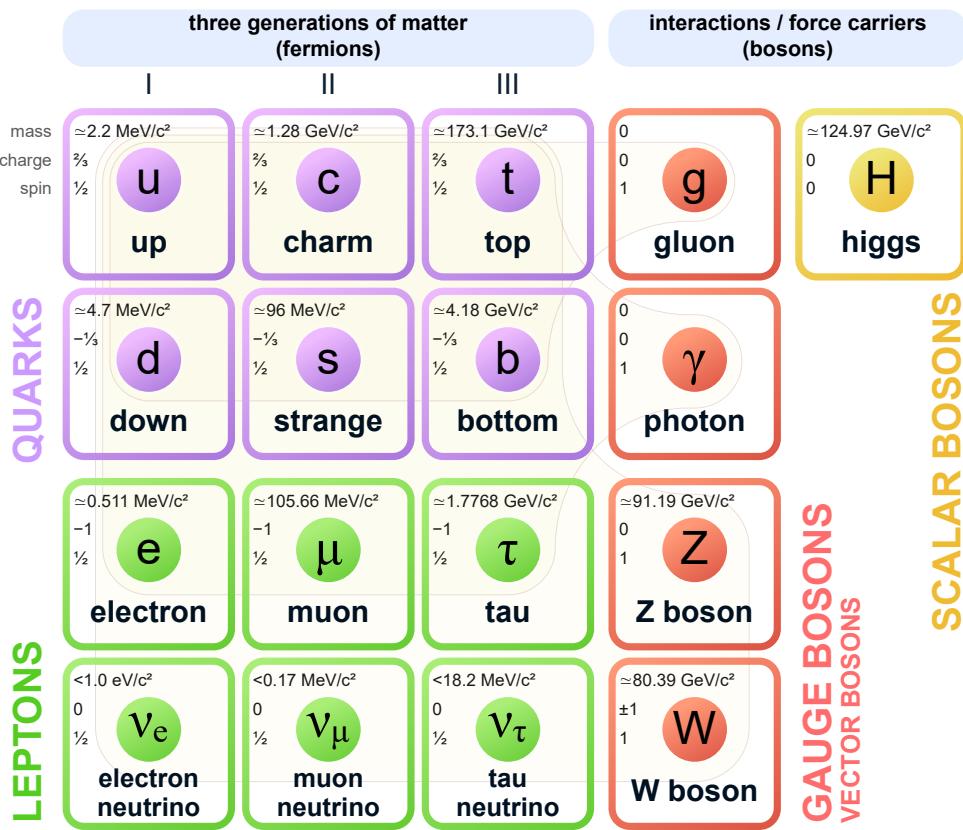


Figure 1.1: Standard model list of matter and interaction particles [1]

193

1.1.1 Quantum Electrodynamics (QED)

QED is a quantum field theory of electrodynamics, it describes the interaction of photons to the charged fermions. The QED is local gauge invariant and symmetric with $U(1)_Q$ group, defined as,

$$U(1)_Q = \exp(iQ\theta(x)) \quad (1.1)$$

194 where $\theta(x)$ is any spacetime function also called gauge parameter, and Q is coupling
195 constant of photon field to the fermions which is equivalent to the charge of fermion.

Under this transformation, fermion spinor $\psi(x)$ and four-potential A_μ electromagnetic tensor will transform as,

$$\psi(x) \rightarrow U(1)_Q \psi(x) \quad (1.2)$$

$$A_\mu \rightarrow A_\mu - \frac{1}{e} \partial_\mu \theta \quad (1.3)$$

The general Lagrangian of QED for fermions and their interaction with photon field is given by,

$$\mathcal{L}_{QED} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (1.4)$$

where m is the mass of fermion, D_μ is the covariant derivative, and $F_{\mu\nu}$ is the electromagnetic field tensor defined as,

$$D_\mu = \partial_\mu + iQA_\mu \quad (1.5)$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (1.6)$$

196

1.1.2 Quantum Chromodynamics (QCD)

197 The strong interactions are represented by $SU(3)_C$ gauge group, invariant under trans-
 198 formations of color charge degree of freedom, and it is based on Yang-Mills theory [2]. Since
 199 “electrodynamics” is the theory of electric charge, this theory of color (*chromo* in Greek)
 200 charge is called “chromodynamics”, hence the name quantum chromodynamics (QCD).

A quark spinor in initial state can be represented as,

$$\psi = \begin{pmatrix} \psi_{red} \\ \psi_{blue} \\ \psi_{green} \end{pmatrix} = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix} \quad (1.7)$$

$SU(3)_C$ is an exact symmetry, it means the difference between colors cannot be measured experimentally, thus the color labels in quark spinor are arbitrary. $SU(3)_C$ transformation is defined as,

$$SU(3)_C = \exp \left(i\theta^a(x) \frac{\lambda^a}{2} \right) \quad (1.8)$$

201 where λ^a for $a = 1, \dots, 8$, are the Gell-Mann matrices, and $\theta^a(x)$ are any gauge parameters. These eight generators of symmetry corresponds to eight gauge vector boson gluons.

Similar to QED, the covariant derivative for QCD can be formed as,

$$D_\mu = \partial_\mu + ig_s \frac{\lambda^a}{2} G_\mu^a \quad (1.9)$$

203 where g_s is the coupling constant of gluon to the quarks, and G_μ^a are the eight gauge
 204 fields corresponding to gluons.

Now the corresponding field strength tensor in QCD can be formed as,

$$F_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f^{abc} G_\mu^b G_\nu^c \quad (1.10)$$

where f^{abc} are the structure constants of $SU(3)_C$ which satisfy $[\lambda^a, \lambda^b] = i f^{abc} \lambda^c$ relation.

The full Lagrangian for QCD can now be constructed as,

$$\mathcal{L}_{QCD} = \bar{\psi}^i (i\gamma^\mu D_\mu{}^{ij} - m\delta^{ij}) \psi^j - \frac{1}{4} F_{\mu\nu}^a F^{a\mu\nu} \quad (1.11)$$

for mass m , indices i and j runs from 1 to 3.

The main difference of gluon field with respect to photon field is the presence of third term in field strength tensor which allows triplet and quartic self coupling of gluons.

1.1.3 Electroweak Theory

The theory of weak interaction which changes the flavor of fermions is called quantum flavor dynamics (QFD). Since the unification of electromagnetic and weak interaction into electroweak (EW) interaction by Glashow, Weinberg, and Salam [3–5], the weak interaction is better understood in terms of EW theory.

Weak interaction only couples to left-handed fermions and it is same whether the fermion is charged or not. The underlying gauge group of EW interaction is $SU(2)_L \otimes U(1)_Y$ and has

two transformations one for the left-handed doublet L and the right handed singlet fermions ψ_R which are defined as,

$$SU(2)_L \otimes U(1)_Y = \exp \left(i\theta^a(x) \frac{\sigma^a}{2} + i\theta(x) \frac{Y}{2} \right), \quad (\text{doublet}) \quad (1.12)$$

$$= \exp \left(i\theta(x) \frac{Y}{2} \right), \quad (\text{singlet}) \quad (1.13)$$

where Y is the hypercharge (linear combination of electric charge and weak isospin component), and σ^a for $a = 1, 2, 3$ are the Pauli spin matrices generator of $SU(2)$ symmetry. Left-handed fermion L doublets are,

$$L = \begin{pmatrix} \nu_e \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau_L \end{pmatrix}, \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix} \quad (1.14)$$

and right-handed singlets are,

$$\psi_R = e_R, \mu_R, \tau_R, u_R, d_R, c_R, s_R, t_R, b_R \quad (1.15)$$

The covariant derivative of EW is then,

$$D_\mu L = \left(\partial_\mu + ig_w \frac{\sigma^a}{2} W_\mu^a + ig \frac{Y}{2} B_\mu \right) L \quad (1.16)$$

$$D_\mu \psi_R = \left(\partial_\mu + ig \frac{Y}{2} B_\mu \right) \psi_R \quad (1.17)$$

where W_μ^a and B_μ are the gauge fields. The EW Lagrangian can now written as,

$$\mathcal{L}_{EW} = i\bar{L}\gamma^\mu D_\mu L + i\bar{\psi}_R\gamma^\mu D_\mu \psi_R - \frac{1}{4}W_{\mu\nu}^a W^{a\mu\nu} - \frac{1}{4}B_{\mu\nu} B^{\mu\nu} \quad (1.18)$$

where $B_{\mu\nu}$ and $W_{\mu\nu}^a$ are fields strength, defined as,

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad (1.19)$$

$$W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a - g_w \epsilon^{abc} W_\mu^b W_\nu^c \quad (1.20)$$

the linear combination of B_μ and W_μ gauge field, with a weak mixing angle θ_w gives 4 vectors boson W^+ , W^- , Z , and γ of SM,

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \mp W_\mu^2) \quad (1.21)$$

$$Z_\mu = \cos \theta_w W_\mu^3 - \sin \theta_w B_\mu \quad (1.22)$$

$$A_\mu = \sin \theta_w W_\mu^3 + \cos \theta_w B_\mu \quad (1.23)$$

$$\tan \theta_w = g/g_w \quad (1.24)$$

214 Similar to QCD, the presence of third term in field strength tensor allows the self triple
215 (WWZ, WW γ) and quartic (WWWW, WWZZ, WWZ γ , WW $\gamma\gamma$) couplings.

216 **1.1.4 Electroweak Symmetry Breaking and Higgs Mechanism**

217 The *spontaneous symmetry breaking* is the phenomena which explains why the ground
218 state is not invariant under the symmetry of the Lagrangian. The “spontaneous” means the
219 symmetry breaking is not done by external agent but rather by Lagrangian itself in ground
220 state.

221 The EW theory unifies weak interaction and QED but the gauge boson in EW theory
222 are all massless, if we were to add mass terms like $-m^2 W_\mu W^\mu$ by hand, it will no longer
223 be gauge invariant. The solution to this without breaking gauge invariance is spontaneous

²²⁴ symmetry breaking, but this requires addition of new scalar field called Higgs field via
²²⁵ Brout–Englert–Higgs mechanism (BEH) [6, 7], and this symmetry breaking is known as
²²⁶ electroweak symmetry breaking (EWSB).

BEH introduces a complex scalar field as $SU(2)_L$ doublet with non-zero vacuum expectation value (VEV),

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^1 + i\phi^2 \\ \phi^3 + i\phi^3 \end{pmatrix} \quad (1.25)$$

and BEH Lagrangian is,

$$\mathcal{L}_{BEH} = |D_\mu \phi|^2 - V(\phi) \quad (1.26)$$

where D_μ is same as EW covariant derivate in Equation 1.16, and $V(\phi)$ is,

$$V(\phi) = \mu^2 |\phi|^2 + \lambda(|\phi|^2)^2 \quad (1.27)$$

the parameter λ is required to be positive, for $\mu^2 > 0$ the minima is at 0, which is not an interesting case, but for $\mu^2 < 0$ vacuum state energy is given by,

$$\phi^\dagger \phi = -\frac{\mu^2}{2\lambda} \quad (1.28)$$

by the choice of non-zero VEV v , scalar field can be parameterized as,

$$v = \sqrt{\frac{-\mu^2}{\lambda}} \quad (1.29)$$

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ h(x) + v \end{pmatrix} \quad (1.30)$$

where $h(x)$ is the Higgs field and BEH spontaneously breaks electroweak symmetry,

$$SU(2)_L \otimes U(1)_Y \rightarrow U(1)_{EM} \quad (1.31)$$

²²⁷ Visually the Higgs potential is shown in Figure 1.2. The ball position at the center
²²⁸ represents unbroken symmetry, and at the minima represents spontaneous broken symmetry
in the ground state of potential.

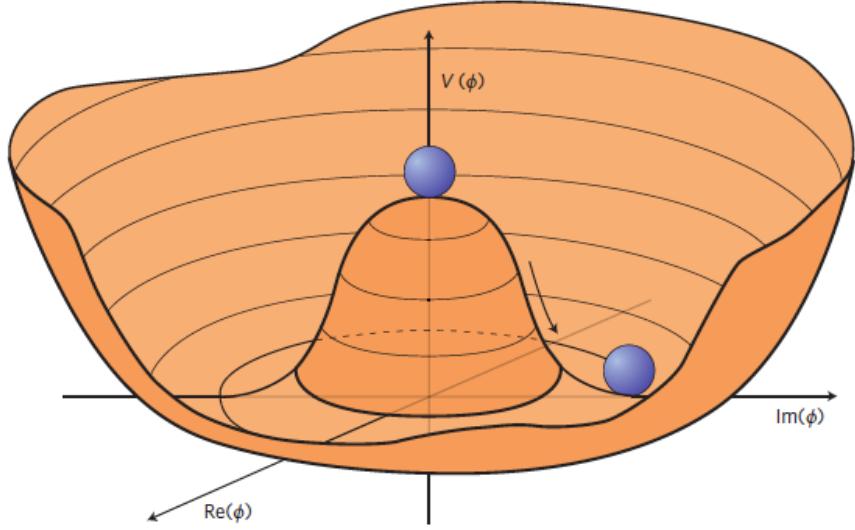


Figure 1.2: 3D representation Higgs potential [8].

²²⁹

After EWSB, the BEH Lagrangian contains the following mass terms,

$$m_W^2 W_\mu^+ W^{-\mu}, \quad m^2 Z Z_\mu Z^\mu, \quad m^2 h^2 \quad (1.32)$$

and gauge fields,

$$-\frac{1}{4} A_{\mu\nu} A^{\mu\nu}, \quad -\frac{1}{4} W_{\mu\nu}^+ W^{-\mu\nu}, \quad -\frac{1}{4} Z_{\mu\nu} Z^{\mu\nu}, \quad -\frac{1}{4} (\partial_\mu h)(\partial^\mu h) \quad (1.33)$$

²³⁰ thus explaining existence of three massive vector boson (W^\pm , Z), one massless vector
²³¹ boson (γ), and one massive scalar boson Higgs (H).

With experimentally measured value of VEV v approximately 246 GeV, The masses of bosons can be written in terms of v as,

$$m_A = 0, \quad m_W = \frac{g_W v}{2}, \quad m_Z = \frac{\sqrt{g_W^2 + g^2} v}{2}, \quad m_H = \sqrt{2\lambda} v \quad (1.34)$$

²³² **1.2 Vector Boson Scattering**

The vector boson scattering (VBS) is,

$$VV \rightarrow VV \quad (1.35)$$

²³³ that is, when you have two vector bosons in initial state and two vector bosons in final
²³⁴ state.

²³⁵ This section describes the motivation behind studying VBS, and the topology of scatter-
²³⁶ ing studied in this dissertation.

²³⁷ **1.2.1 Motivation**

A massless spin-1 boson can exists in two transverse polarization as,

$$\varepsilon_\pm^\mu = \mp \frac{1}{\sqrt{2}} (0, 1, \pm i, 0) \quad (1.36)$$

and massive vector bosons can also exists in one longitudinal polarization,

$$\varepsilon_L^\mu = \frac{1}{m}(p_z, 0, 0, E) \quad (1.37)$$

This means the longitudinal polarized VBS will scale as E/m , whereas the scattering of transverse polarized boson remains constant. The Figure 1.3 shows the cross-section of longitudinal polarized VBS $V_L V_L \rightarrow V_L V_L$ for low to high energies. Perturbatively the cross-section of longitudinal polarized VBS will scale with center of mass energy \sqrt{s} and eventually the unitarity is violated at ≈ 1.2 TeV scale [9, 10]. The Figure 1.3 also shows how the existence of light Higgs boson and inclusion of Higgs to vector boson coupling diagrams in longitudinal polarized VBS can restore unitarity violation, and since the discovery of Higgs boson $m_H = 125$ GeV in July 2012, the VBS studies became important and complementary to direct measurement of Higgs coupling in SM, and test for EWSB at TeV scale.

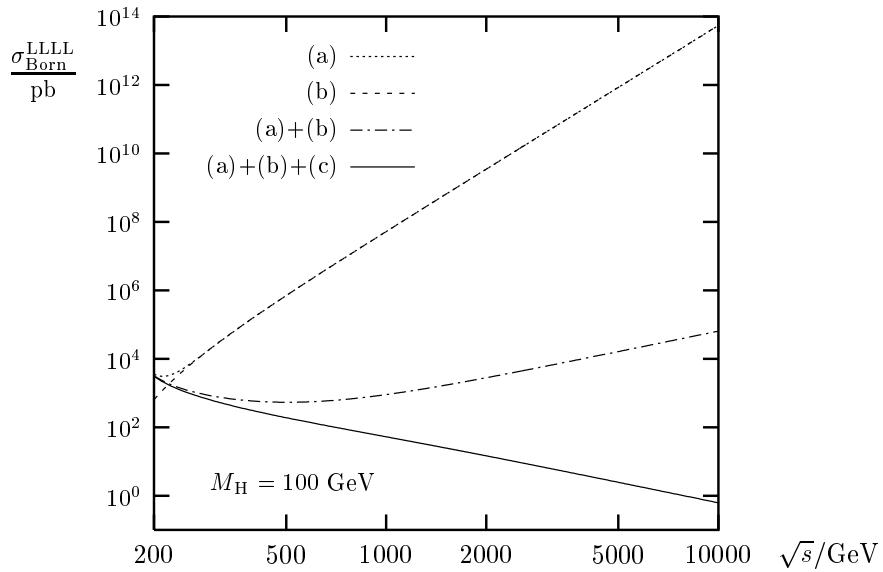


Figure 1.3: The cross-sections for longitudinal polarized VBS involving three and four-boson coupling with and without Higgs coupling diagram included. (a) is for the diagrams with three-boson coupling, (b) is for diagrams with four-boson coupling, and (c) is for diagrams with Higgs to vector boson coupling. [11]

1.2.2 Topology of VBS in Semileptonic ZV Final State

247 In proton-proton collisions, the actual interaction happens with the constituent quarks.
248 For VBS to happen, the incoming (colliding) quarks have to radiate vector boson, then the
249 scattering process between those vector boson can proceed via exchange of vector boson,
250 Higgs boson, or quartic coupling. The tree level Feynman diagram of a VBS process in
251 proton-proton collision is shown in Figure 1.4.
252

253 The outgoing quarks are the signature of VBS in hadron collider experiments because
254 they will have large pseudorapidity difference between them, and will also have large invariant
255 mass of outgoing quark pair. Generally the jets coming from these outgoing quarks are first
256 tagged as “VBS Jets” to filter out most of the QCD background.

257 The type of leptonically decaying vector boson can be determined, i.e. whether it was W
258 or Z, but for the hadronically decaying vector boson it is challenging and generally denoted
259 by V. This analysis looks for the VBS signature with ZV in final state with Z decaying to
260 two opposite sign same flavor (OSSF) leptons, and V decaying to pair of quarks.

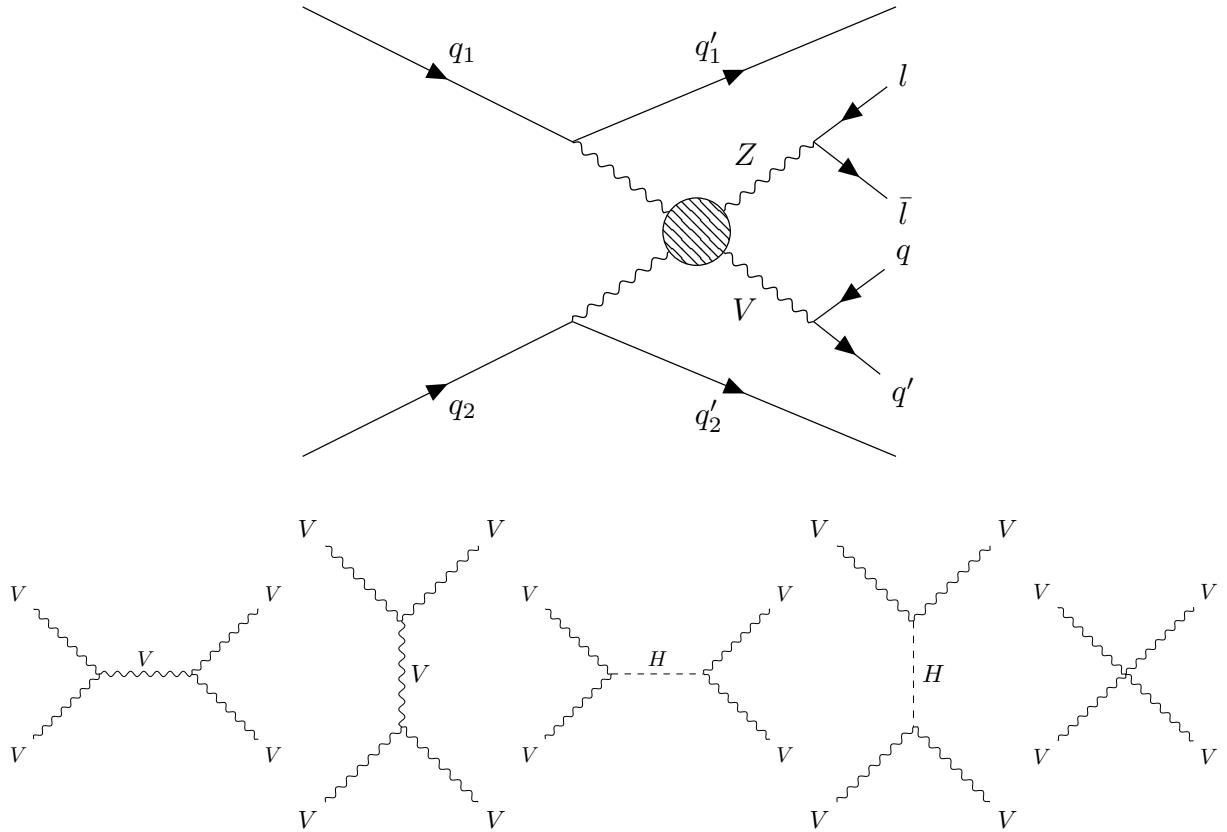


Figure 1.4: Tree level Feynman diagram of ZV VBS process at LHC. The top diagram shows the production of two vector boson being radiated from incoming quarks, in final state after scattering (blob), Z decays to pair of leptons, V (W/Z) decays to pair of quarks, and plus two outgoing quarks. The bottom row of diagram shows the tree level processes that can happen in scattering represented by blob in top diagram, starting from s and t-channel exchange of vector boson, Higgs boson, and the last one is quartic coupling of vector bosons.

261

CHAPTER 2

262

THE LHC AND CMS EXPERIMENT

263

The physics analysis is carried out using Compact Muon Solenoid (CMS) experiment at European Council for Nuclear Research (CERN) Large Hadron Collider (LHC) accelerator. This chapter provides overview of LHC and detail of CMS experiment and its sub-detectors for particle tracking and calorimetry.

267

2.1 The Large Hadron Collider

268

The LHC is the largest accelerator located at CERN in Geneva, Switzerland. The main LHC ring is 27 km in circumference and around 50 to 175 m underground. The LHC is built to collide protons at 14 TeV center-of-mass energy, LHC delivered proton-proton collisions at 7 and 8 TeV during run-1 (2010–2012), and at 13 TeV center-of-mass energy during run-2 (2015–2018) [12].

273

274

275

276

277

278

279

The Figure 2.1 describes CERN accelerator complex. The protons are sourced by ionizing hydrogen atoms and then fed into linear accelerator (LINAC). The LINAC accelerates the protons to 50 MeV and sent to the booster. Then the booster increases energy of protons to 1.4 GeV and feeds it to the proton syncrotron (PS) which further increases energy to 25 GeV and starts bunching them together with bunches 25 ns apart. Then the proton bunches are passed through super proton synchrotron (SPS) which increases energy to 450 GeV and finally sent to main LHC clockwise and counterclockwise rings where they are accelerated to

final energy required which is 6.5 TeV for both bunches going clockwise and counterclockwise to obtain collisions at 13 TeV center-of-mass energy.

The proton-proton collisions occurs at four different location where two general purpose detectors CMS and A Toroidal LHC Apparatus (ATLAS), and two specific purpose detector A Large Ion Collider Experiment (ALICE) and LHC-beauty (LHCb) are located.

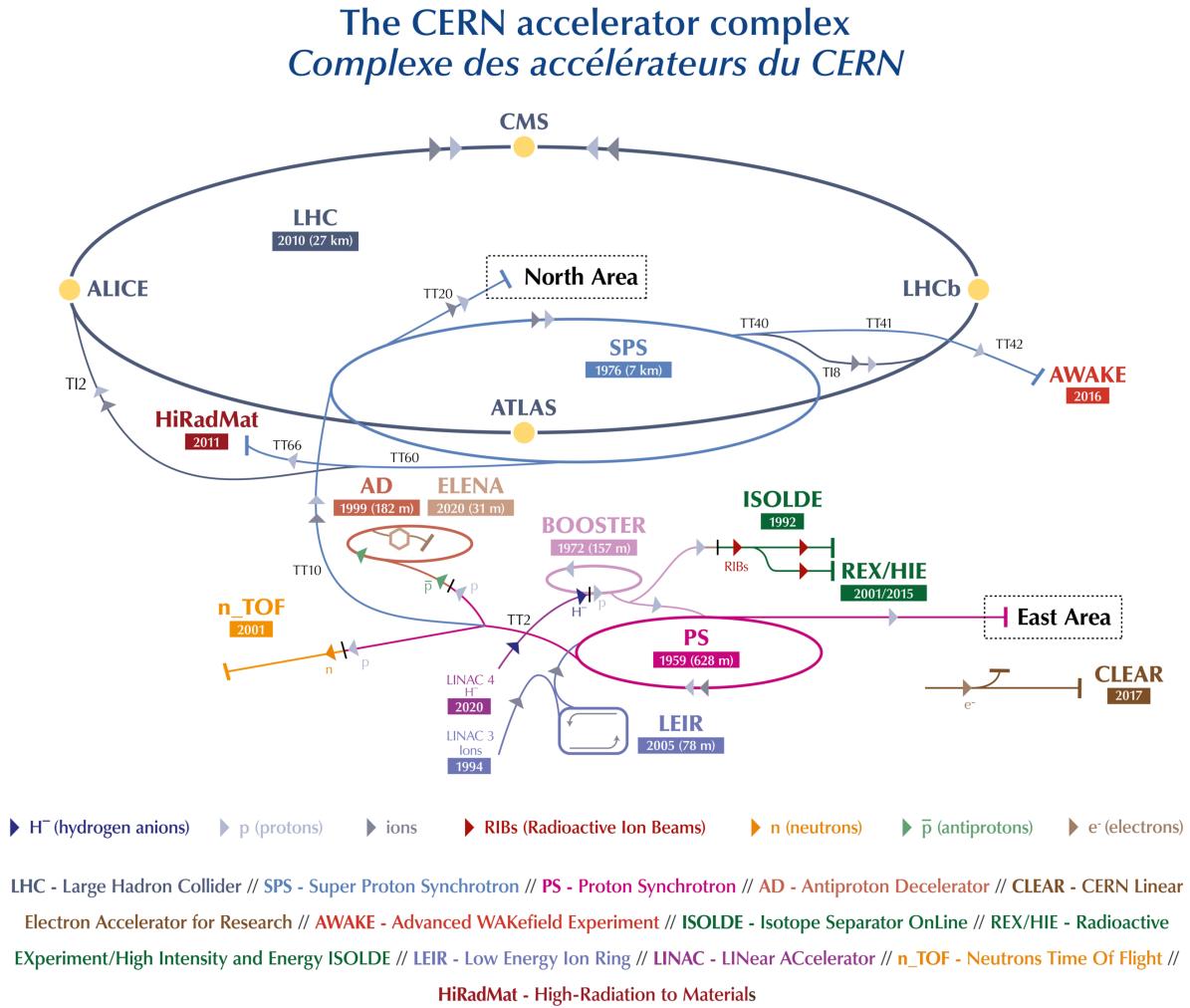


Figure 2.1: A schematic of the CERN accelerator complex [13]

285

2.1.1 Integrated Luminosity

The number of events generated in a collisions for a given process is,

$$N = L\sigma \quad (2.1)$$

286

where σ is cross-section of the process and L is the luminosity of the LHC.

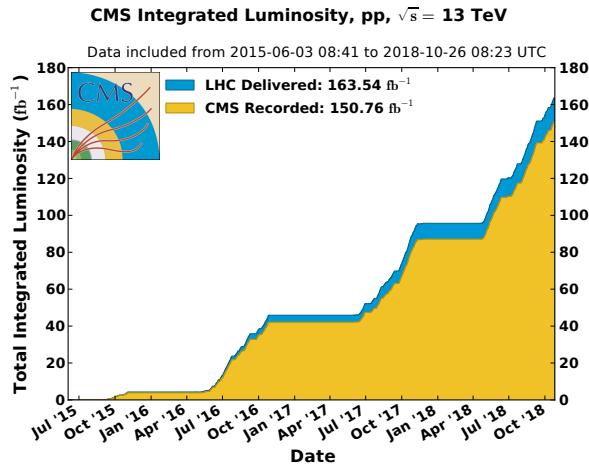


Figure 2.2: Cumulative delivered and recorded luminosity versus time for 2015–2018 proton-proton collisions [14]

287

Cumulative luminosity delivered and recorded by CMS during run-2 operation is shown in Figure 2.2. For run-2 standard physics analysis luminosity recorded during 2016–2018 is considered, and only runs certified as “golden” by CMS Luminosity physics object group (POG) for analysis are used. The total luminosity for run-2 standard physics is 137.19 fb^{-1} and separately for years in Table 2.1 [15–17].

Table 2.1: Standard physics luminosity for run-2

2016	2017	2018	run-2
35.92 fb^{-1}	41.53 fb^{-1}	59.74 fb^{-1}	137.19 fb^{-1}

292

2.2 The CMS Detector

293 The CMS detector is a general purpose detector. A cutaway view of the detector is
 294 shown in Figure 2.3. The detector is cylindrical with dimensions 21 meters long, and 15
 295 meters in diameter, and the whole detector weighs about 14000 tonnes. The detector is built
 296 in slices with central region called “barrel”, and two closing end sides called “endcap”. A
 297 superconducting solenoid generates magnetic field of 3.8 T inside and 2 T outside, and to
 298 contain the magnetic field outside of solenoid and support structure of the detector massive
 299 steel yokes are used.

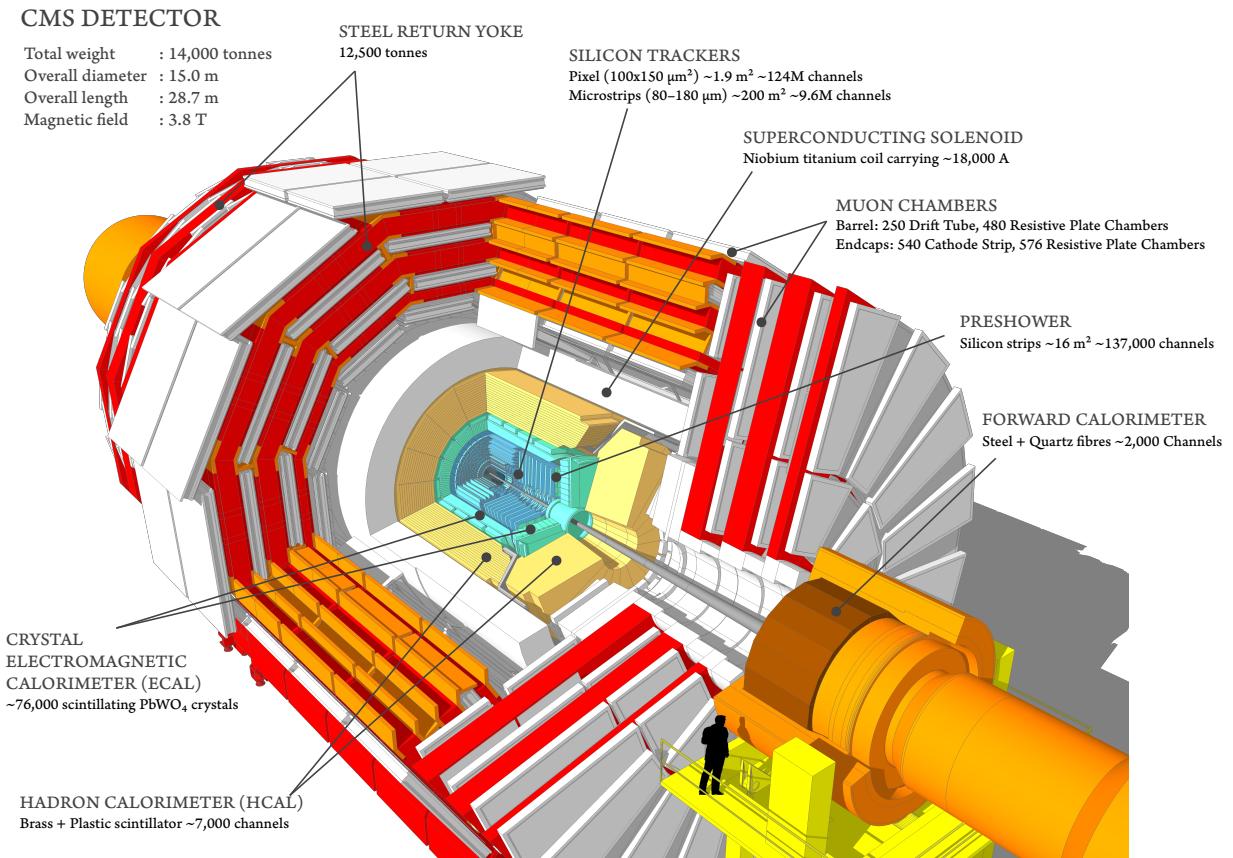


Figure 2.3: The CMS detector cutaway view [18]

300 The slice view of CMS in Figure 2.4 shows how different particles leave signature in CMS
 301 detector. Neutral particles such photons, neutrinos, and hadrons will leave no track in Silicon
 302 Tracker (ST), and are identified by only energy deposited or missing energy. Electrons are
 303 identified from the track in ST and energy deposit in Electromagnetic Calorimeter (ECAL),
 304 hadrons are heavier and they pass through ECAL and deposit their energy completely in
 305 Hadronic Calorimeter (HCAL), leaving only small fraction of energy in ECAL. Since muons
 306 are minimum ionizing particle (MIP), they pass through whole detector with very small
 307 fraction of energy deposit in ECAL and HCAL.

308 This section describes the subsystems of CMS detector. For detailed technical description
 309 refer to [19].

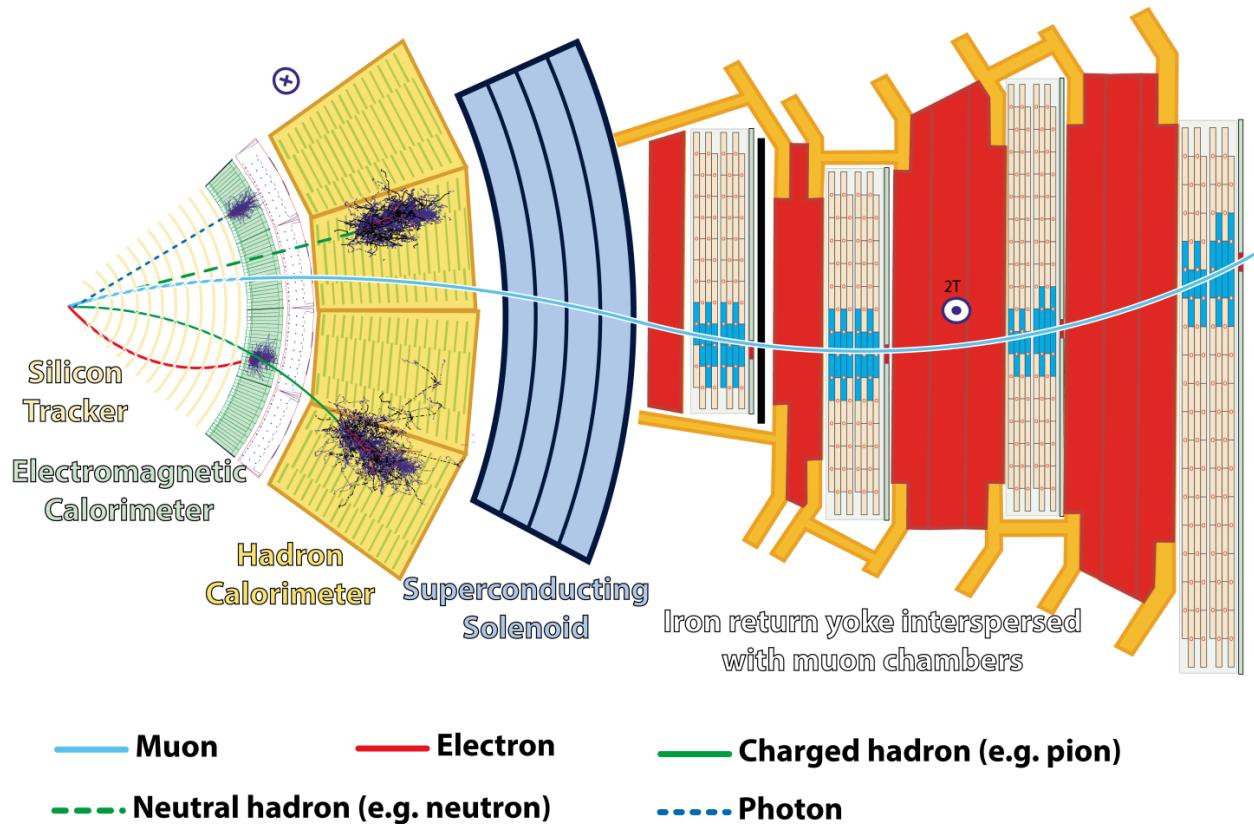


Figure 2.4: The CMS detector slice view [20]

310

2.2.1 The CMS Coordinate System

311 CMS uses interaction point (IP) of collisions as origin to define right-handed coordinate
 312 system. The z -axis is along the beamline, the x -axis points toward the center of the LHC,
 313 and the y -axis points upwards, toward Earth's surface. The transverse plane $x - y$ is used as
 314 to calculate most commonly used quantities like transverse momentum p_T and energy E_T .

To describe the direction of particles leaving the IP, azimuthal ϕ and polar θ angles are used. ϕ is measured around the beam axis, and θ is measured from the beam axis. In collider physics, pseudorapidity η (Lorentz invariant) is used to describe direction from beam pipe instead of θ as,

$$\eta = -\ln[\tan \theta/2] \quad (2.2)$$

and sometimes in terms of rapidity y as,

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \quad (2.3)$$

Particles kinematics can be completely described in terms of p_T , η , ϕ , and E_T or mass. The distance between the two particles ΔR in $\eta - \phi$ plane is described as,

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad (2.4)$$

315

2.2.2 The Superconducting Magnet

316 The superconducting magnet is the main part of the CMS detector, it is 12.5 meters long
 317 and 6.3 meters in diameter. The magnet is cooled to 4.5 K and 20 kA current flows through
 318 it to generate 3.8 T of magnetic field with stored energy of 2.6 GJ.

319 The Figure 2.5 shows visible superconducting magnet and iron yoke when part of CMS
320 detector was lowered in the underground cavern during installation in 2007.

321 The key purpose of magnet is to determine the momentum and the sign of charged
322 particles by bending them. The momentum resolution of the particles will decrease with
323 increase in p_T , with constant 3.8 T magnetic field inside and it has momentum resolution
324 of $\Delta p/p \approx 10\%$, which is enough to determine unambiguously the sign of muons with
325 momentum of $\approx 1 \text{ TeV}/c$.

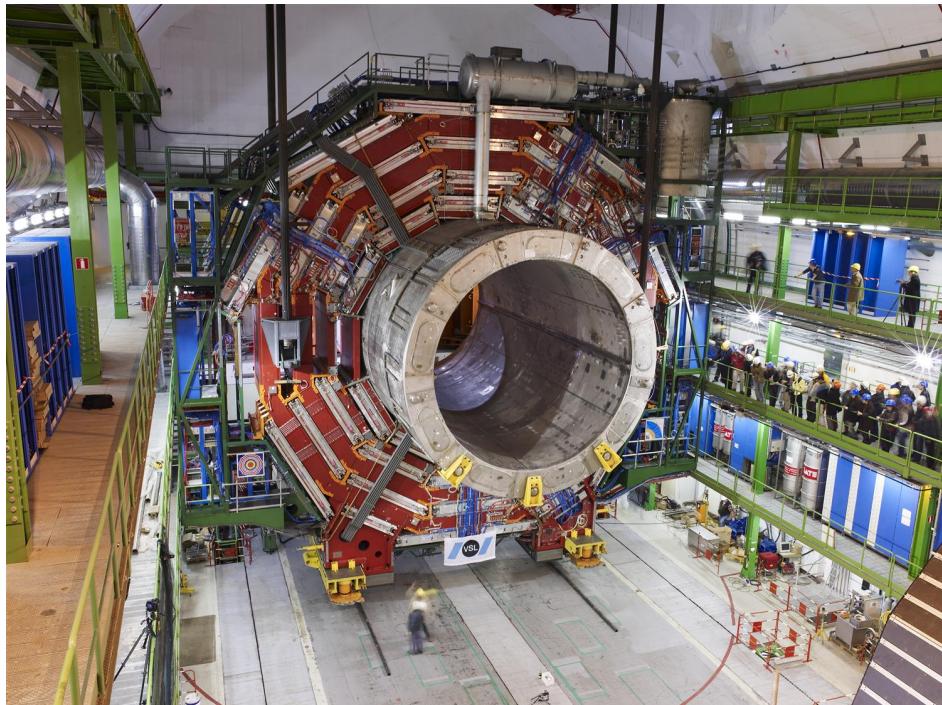


Figure 2.5: The picture of the CMS detector central part when lowered in underground cavern with superconducting magnet and iron yoke visible [21].

326

2.2.3 The Tracking System

327 The CMS tracking system ST is the innermost part of the detector, it is made up of pixel
 328 and strip detectors. The main goal of ST is to reconstruct the tracks of the charged particles
 329 with high precision in high pileup environment.

330 Silicon is most commonly used material for making tracking systems because of it's
 331 semiconductor properties, and high radiation hardness which is essential for the innermost
 332 detector. When a p-n junction is built on silicon substrate it creates a depletion zone with
 333 no charge carriers at the junction, and whenever a charged particle pass through the deple-
 334 tion zone it creates a electron-hole pair, and under reverse bias this electron-hole generates
 335 electrical signal. The CMS tracking consists of about 124 million channels of such junctions
 336 in pixel detector and 10 million in strip detector.

337 The pixel detector was upgraded in 2017 and the comparison of layers before and after
 338 the upgrade is shown in Figure 2.6. It is made up of four barrel layers and three endcaps,
 339 with nearest barrel layer being 3 cm away from beamline for precise measurement of IP.
 340 Because of the large number of pixel channels, the readout is done by Application-specific
 341 integrated circuits (ASICs).

342 The outermost part of ST detector is made of silicon strips. It allows large coverage by
 343 reducing number of readout channels. It has 10 layers in barrel region and 12 discs in endcap
 344 region. For better signal-to-noise ratio and radiation tolerance both pixel and strip operates
 345 at -20 °C.

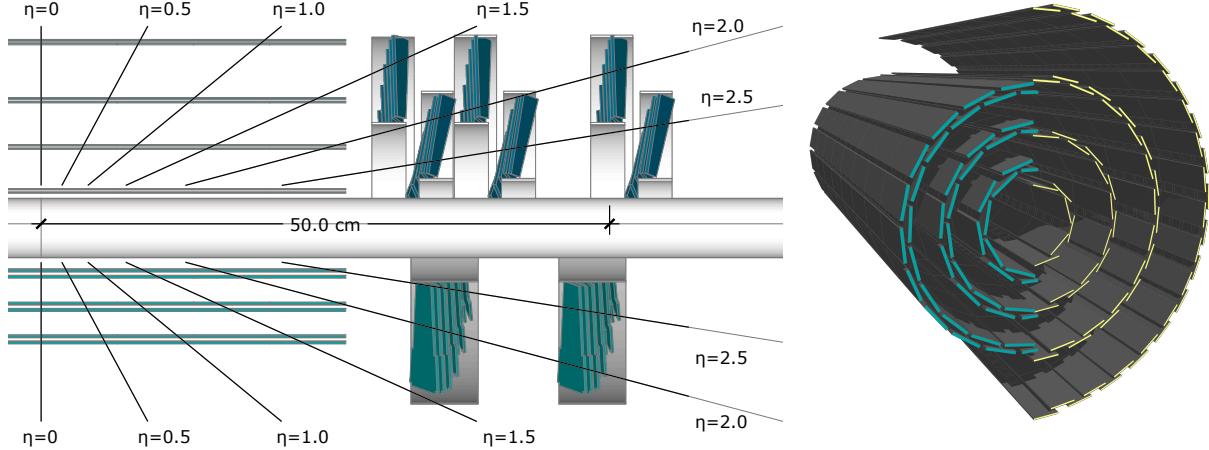


Figure 2.6: The CMS pixel upgrade. The left is cross sectional view of pixel detector layers before upgrade (bottom) and after Phase 1 upgrade (top). The right is view pixel barrel before upgrade (left) and after upgrade (right) [22].

2.2.4 The Electromagnetic Calorimeter

The ECAL active material is made of lead tungstate (PbWO_4) scintillating crystals and two layers silicon strip for preshower in front of the endcaps. The crystals in central barrel section are mounted in quasi-projective geometry pointing towards IP and covers $|\eta| < 1.48$, and two endcaps extends the coverage to $|\eta| = 3.0$. The schematic layout of ECAL is shown in Figure 2.7 and the picture of endcap quadrant when assembled in Figure 2.8

The main purpose of ECAL is to determine energy and positions of electromagnetically interacting particles. To determine particle need to completely deposit their energy, except electron and photons all other particles pass through ECAL crystals with only small fraction of energy signature in crystals. When electron and photon interacts with PbWO_4 it starts the process of electromagnetic shower and continues until the energy the energy of the incident particle is below threshold, which is about 1 MeV.

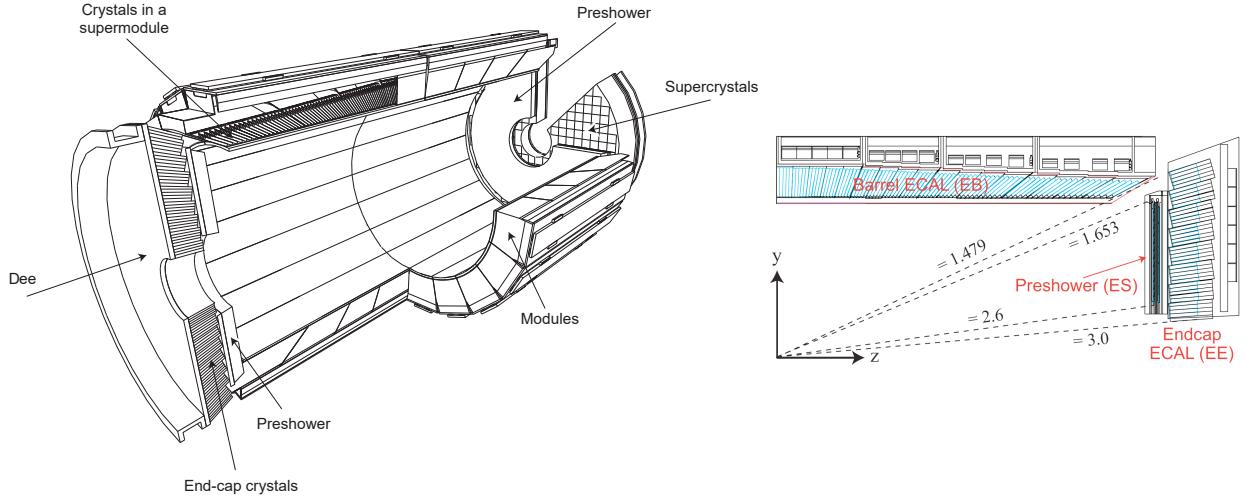


Figure 2.7: The CMS ECAL schematic layout. The left is schematic showing arrangement of superclusters in barrel and endcap (with preshower layers). The right is $y - z$ plane quarter view of ECAL layout [23].

The resolution of the ECAL energy measurements can be described as,

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \quad (2.5)$$

358 where S is the stochastic term, N is related to the noise, and C is a constant offset.

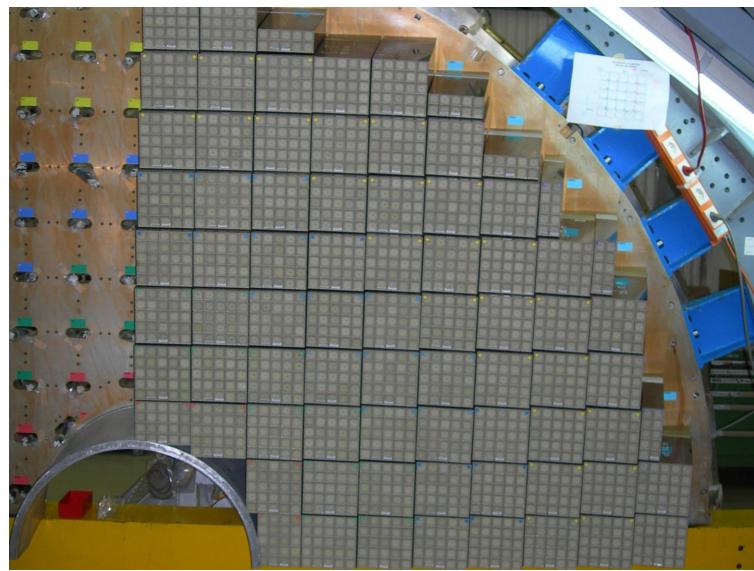


Figure 2.8: The ECAL endcap quadrant assembled view [24].

359

2.2.5 The Hadronic Calorimeter

360 HCAL is last subdetector inside solenoid after ECAL and the first half of barrel HCAL
 361 inserted is shown in Figure 2.9. Similar to ECAL the purpose of HCAL is to shower hadrons,
 362 and measure their energy and position. HCAL is made up of towers pointing towards IP
 363 and each tower is made up of sampling layers with alternating layers of plastic scintillator
 364 and brass. Brass acts as absorber in HCAL and causes hadrons to shower, then from the
 365 light output of scintillator receiving secondary shower particles gives the amount of energy
 366 deposit in each layer. In phase 1 upgrade the HCAL was upgraded to give energy deposit
 367 as function of depths and the depth segmentation schematic is show in Figure 2.10 and the
 368 details of upgrade are in technical design report [25].



Figure 2.9: The first half of the barrel HCAL inserted into the superconducting solenoid (April 2006) [26].

369 HCAL consists barrel (HB) and two endcaps (HE) located inside solenoid. These two
 370 subsystems combined cover region $|\eta| = 3.0$, which is most of physics analysis done in CMS.
 371 There are two other subsystems of HCAL outside solenoid, a forward HCAL (HF) and
 372 outer barrel HCAL (HO). HO was added to ensure there is no leak from the particles that
 373 make past the solenoid. HF extends the coverage to $|\eta| = 5.0$ and is based Cherenkov
 374 radiation principle unlike other subsystems of HCAL, and it uses quartz fiber as active
 375 material with steel absorbers. HF is used most commonly used by heavy ion analysis.

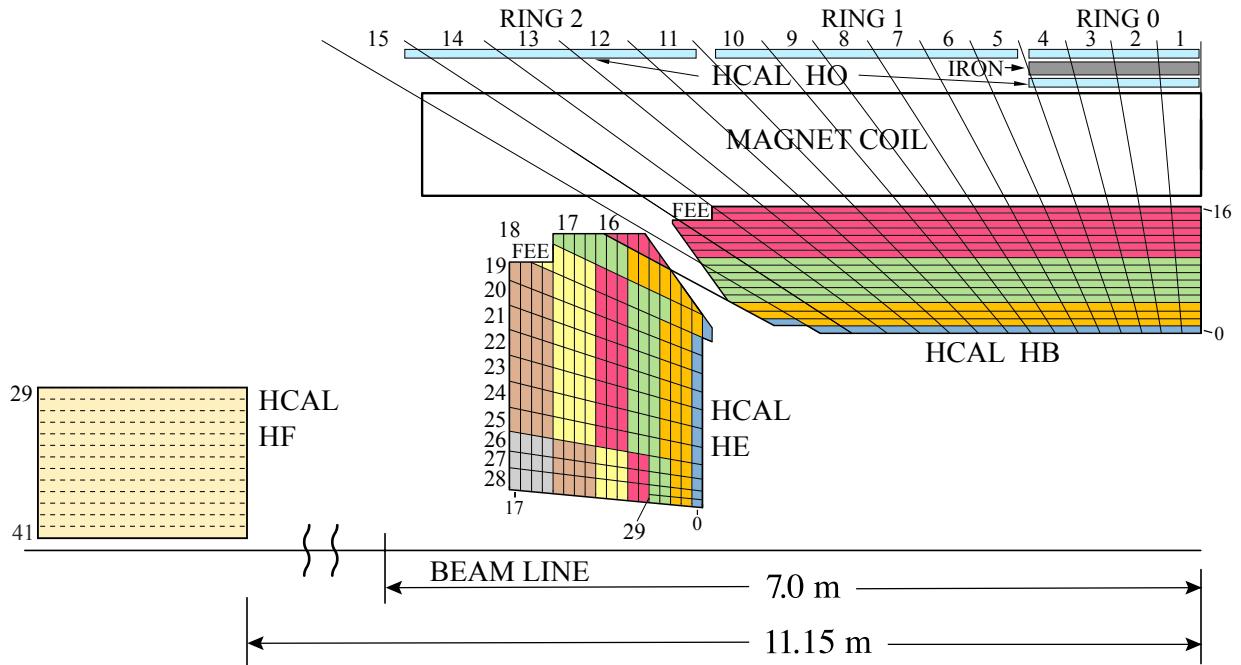


Figure 2.10: The HCAL depth segmentation after phase 1 upgrade [27].

376

2.2.6 Muon Detector

377 The outermost subsystem in the CMS detector is the muon detector. Unlike electrons,
 378 muons are MIPs, they do not lose much of their energy while passing through tracker,
 379 calorimeter and solenoid. Muon detector is build to identify, measure momentum and trigger

380 the events with muons. Like other subsystems, muon detector consists of barrel and endcap
 381 detector and schematic layout is highlighted in Figure 2.11.

382 The muon detector consists of three subsystems drift tubes (DTs), cathode strip chambers
 383 (CSCs) and resistive plate chambers (RPCs).

384 The DTs are wire gas detectors filled with Argon and composed of many tube cells of
 385 about 4 cm. Muon passing through these tubes ionizes Argon and free electron is detected
 386 with wires as cathode. Each DT is about 2 meters by 2.5 meters in size, and there are four
 387 layers of the DTs interleaved with iron yoke parallel to the beam pipe in barrel region. The
 388 drift time is the order of about 380 ns.

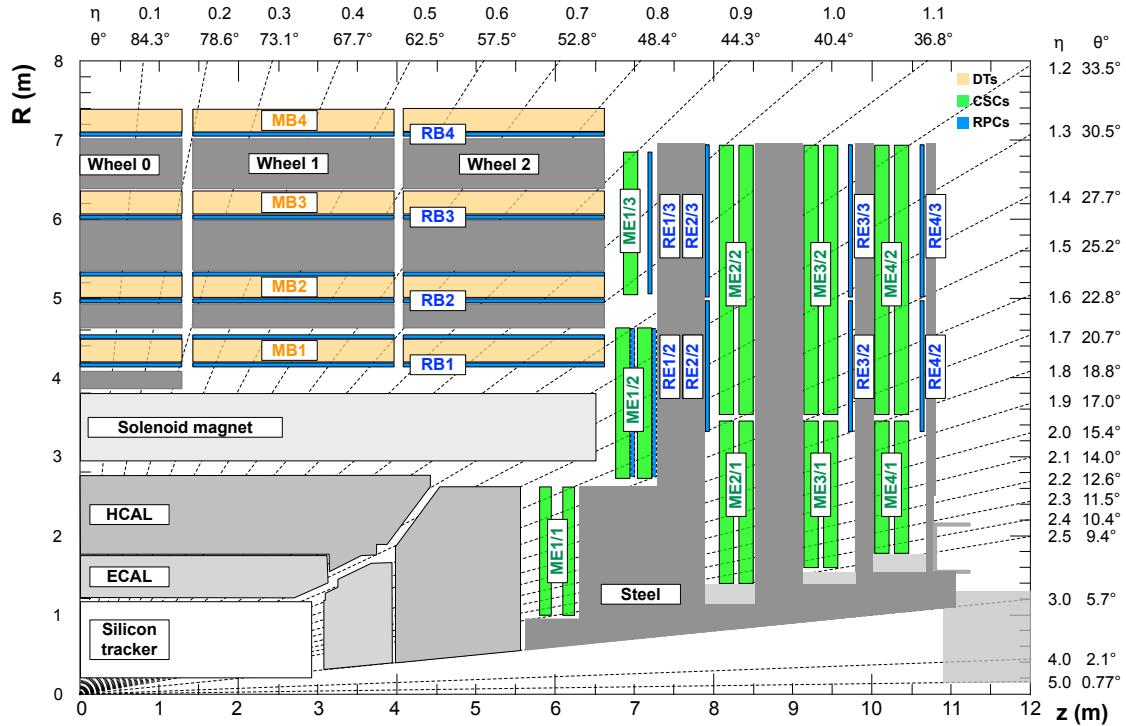


Figure 2.11: The quadrant view of CMS subdetectors layout, and the coverage of the muon detector DTs, CSCs, and RPCs highlighted [28].

389 The CSCs are based on same principle as DTs, and are made of multi-wire proportional
 390 chambers consisting of 6 anode planes interleaved with 7 cathode planes. They have time

³⁹¹ resolution smaller than 5 ns. The CSCs are used in endcap region, where radiation hardness
³⁹² is required, and non uniform magnetic field does neutrinos effects the the measurement.

³⁹³ The RPCs are made up of two high resistive parallel plates, with oppositely charged plates
³⁹⁴ and gas volume between them. When a charged particles passes through it and ionizes the
³⁹⁵ gas, it creates an avalanche and charge is collected by metallic readout strips. RPCs have
³⁹⁶ poor position resolution but fast readout of the order of 1 ns, which is fast compared to DTs,
³⁹⁷ this is the reason there are 1 or 2 RPCs attached to both DTs, and CSCs.

³⁹⁸ **2.2.7 Level 1 Trigger**

³⁹⁹ Since proton-proton collisions happens every bunch crossing which is 25 ns apart, which
⁴⁰⁰ is equivalent to 40 MHz collisions rate. At this collisions rate, the data storage required will
⁴⁰¹ be enormous and CMS can only record up to 1000 events per second. Since most of events
⁴⁰² does not contain interesting physics events, they can be thrown away. To do this CMS has
⁴⁰³ two tier trigger system Level-1 Trigger (L1T), and High Level Trigger (HLT).

⁴⁰⁴ The L1T is the foremost electronic processing system through which event information is
⁴⁰⁵ processed before it is passed to second trigger system HLT. The L1T is designed to make fast
⁴⁰⁶ decisions in about $3.8\ \mu s$, and only uses ECAL, HCAL and muon system to make decision.
⁴⁰⁷ L1T cut downs the data rate from 40 MHz to 100 kHz. The L1T electronics is placed next
⁴⁰⁸ to the detector in underground cavern for fast transfer of data.

⁴⁰⁹ The HLT further reduces the data rate from 100 kHz to about 1 kHz using a computer
⁴¹⁰ farm with nearly 26000 cores. HLT uses all the available information from the event to make
⁴¹¹ decision in about 300ms. HLT is modular by design to allow the use of information from
⁴¹² different systems to construct multiple paths called HLT paths, for example the single muon

⁴¹³ HLT path will save event with at least one muon passing the selection criteria set in HLT
⁴¹⁴ path. Events passing at least on HLT path are save for offline physics analysis.

CHAPTER 3

EVENT SIMULATION AND RECONSTRUCTION

417 The proton-proton collision at LHC produces shower of particles, before the event information can be easily used in an analysis, the data collected goes through iterative process
 418 of reconstruct particles produced in collision. CMS uses particle flow (PF) algorithm to
 419 reconstruct 4-vectors of muons, electrons, photons, hadrons, jets and missing transverse
 420 momentum [29].

422 To analyze the data collected and compare it with theoretical model, events are simulated
 423 using Monte Carlo (MC) event generators and are passed through detector simulation and
 424 PF so that MC events can be treated same as real events.

425 This chapter describes the basic ingredients for object reconstruction, PF candidates and
 426 MC event generators used in this analysis.

3.1 Track Reconstruction and Calorimeter Clustering

428 For complete particle reconstruction two main ingredients are tracks left by particle in the
 429 detector, and energy deposit in calorimeter. This section describes track reconstruction from
 430 the hits in Tracker and Muon Detector, and energy deposit measurement from calorimeter
 431 clustering.

432 Track reconstruction requires reconstructed hits, and seed generation which are described
 433 in [30], then the track reconstruction is done using pattern recognition which is based on
 434 combinatorial Kalman Filtering (KF) method [31]. It is an iterative process starting from

435 seed layer the track is estimated and then proceeds to next layer one by one, at each successive
 436 layer the track trajectory is better known. There can be multiple hits in each new layer, for
 437 this multiple trajectory candidates are created. All the trajectory candidates are grown in
 438 parallel to avoid bias, and truncated at each layer to prevent exponential increase in number
 439 of candidates. Then finally the track is fitted to compute momentum and vertex information.

440 The main purpose of calorimeter clustering is to determine position and energy deposit
 441 of the particle. A cluster in a calorimeter is a local group of energy deposits that are
 442 spatially consistent with a electromagnetic or hadronic shower. First the topological clusters
 443 are identified, a topological cluster is a contiguous region of energy deposit, then a seed
 444 is identified in topological cluster with certain energy threshold, and highest among the 8
 445 neighbors for ECAL and 4 neighbors for HCAL. Now starting with seed energy and position,
 446 the neighbors energy are added, and new position is calculated. For the case when we have
 447 just one seed, there is only iteration until all the neighbors are added, in case when we
 448 have more than one seeds in a cluster, the energy from neighbors is shared, the fraction of
 449 energy shared depends on the energy and position of the cluster, after the first iteration of
 450 calculation of the energy and position the process is repeated with new values of cluster's
 451 energy and position until either the maximum iteration is reached or cluster's energy and
 452 position values are converged.

453 3.2 Reconstructed Particles

454 After tracks and calorimeter clusters are formed, PF links this information from the
 455 detectors together to form objects as broadly discussed in Section 2.2 and shown in Figure 2.4.
 456 This section describes the properties of those reconstructed particle candidates.

457

3.2.1 Muons

458 Reconstructing muon with best precision is the key ingredient for many physics searches.
 459 Muons reconstruction and identification uses all the information from tracker, calorimeters
 460 and muon detector. There are two types of reconstruction performed “Global” and “Tracker”
 461 for muon candidates. Global muons are formed combining and refitting muon hits in the
 462 muon detector with compatible track from ST, and the tracker muons are formed by extrap-
 463 olating tracks from ST to segments in muon detector.

464 Once the muon candidates are found, the kinematics properties (p_T, η, ϕ) are calculated
 465 from track fitting, and other properties such as distance form primary vertex (PV) d_{xy}, d_z ,
 466 number of hits in the tracker and muon system, tracker based relative isolation (3.1) in a
 467 cone of $\Delta R = 0.3$, and PF relative based isolation (3.2) in a cone of $\Delta R = 0.4$ are stored for
 468 cleaning and isolating muons for physics analysis.

The tracker and PF based relative isolation are defined as,

$$\text{TkIso03} = \left(\sum p_T^{\text{Tracks (PV)}} \right) / (p_T^\mu) \quad (3.1)$$

$$\text{PFRelIso04} = \left(\sum p_T^{\text{CH (PV)}} + \min \left[0, \sum E_T^{\text{NH}} + \sum E_T^\gamma - 0.5 \sum p_T^{\text{CH (PU)}} \right] \right) / (p_T^\mu) \quad (3.2)$$

469 where “Tracks (PV)” refers to all the tracks in tracker and coming from PV, “CH (PV)”
 470 and “CH (PU)” refers to charged hadrons coming from PV and pile up (PU) respectively,
 471 “NH” refers to neutral hadrons, μ refers to muon, and γ refers to photon.

472 There are multiple source of muons whenever collision event happens, they can be real
 473 muons or hadrons which are misidentified as muons, these hadrons are able “punch” through
 474 HCAL and leaves hit in muon detector. The real muons of interest are called “prompt” muons
 475 and others are either usually referred as “fake” or “non-prompt”. Fake muons can originate

476 from decay of pions and kaons in flight usually identified with a “kink” in track or from
 477 heavy flavor decay of b or c-quarks which are identified with tracks not originating from PV.
 478 The prompt muons are the ones coming from decay of H, W, Z bosons and τ leptons, and
 479 have small impact parameter from PV, have hits in both tracker and muon detector, and
 480 are typically well isolated.

481 In addition to muons from collision events, there can be cosmic muons from pion decay
 482 in upper atmosphere. Cosmic muons are generally not in-time with collision and far from
 483 interaction points.

484 **3.2.2 Electrons and Photons**

485 Since there is large amount of material in tracker, electrons often emit bremsstrahlung
 486 photons when passing through tracker volume, and photons can further decay to e^-e^+ pair
 487 which complicates the tracking algorithm. The energy deposit of such electrons emitting
 488 bremsstrahlung will have large spread in ϕ direction because the magnetic field will bend
 489 electrons in ϕ whereas photons are unaffected. For this reason electron and photon recon-
 490 struction are done together, and the Gaussian-sum filter (GSF) algorithm is used for electron
 491 track reconstruction which takes care of kinks in electrons track because of hard emission [32].

492 An electron is reconstructed when an ECAL cluster matches a GSF track, and a photon
 493 is reconstructed when an ECAL cluster with E_T more than 10 GeV is found and have no
 494 matching GSF track. To prevent electron and photon from being misidentified as jets certain
 495 conditions are applied, for electron the number of GSF track matching with ECAL cluster
 496 is limited to maximum of two, and energy deposit in a cone of $\Delta R = 0.15$ in HCAL around
 497 the position of electrons and photons is required to be less than 10%.

498 Similar to muons, after electron and photon reconstruction is done, their kinematics
 499 properties are calculated and various other properties required for cut based and Multivariate
 500 analysis (MVA) based identification are stored. The detailed description of electrons and
 501 photons identification technique and properties used in this dissertation can be found in
 502 Reference [33].

503 **3.2.3 Hadrons and Jets**

504 Quarks and gluons produced in a collision event are not detected directly, because of color
 505 confinement, they go through fragmentation and hadronization making a collimated spray of
 506 particles mostly made of hadrons and are called “jets”. Charged hadrons are reconstructed
 507 when a HCAL cluster can be associated with one or more tracks, if the track association
 508 fails the cluster is reconstructed as neutral hadron.

509 Jet in CMS are reconstructed using FASTJET package [34], which takes input of all PF
 510 candidates and associated tracks. The clustering basically combines 4-vectors of particles
 511 iteratively and stop when distance between two particles (d_{ij}) is higher than stopping distance
 512 (d_{iB}).

513 d_{ij} and d_{iB} are defined as,

$$d_{ij} = \min(p_{Ti}^{2p}, p_{Tj}^{2p}) \frac{\Delta R_{ij}^2}{R^2} \quad (3.3)$$

$$d_{iB} = p_{Ti}^{2p} \quad (3.4)$$

513 where p is the parameter for different clustering algorithms, R is the cone size, and ΔR_{ij}
 514 is distance between two particles in iteration.

515 Anti- k_T (AK) is the most use jet algorithm in physics analysis, this corresponds to $p = -1$,
 516 this means the hard particles will be clustered first in this clustering algorithm. The cone
 517 size used for standard jets (AK4) is $R = 0.4$, and for large jets (AK8) often called “fatjet”
 518 is $R = 0.8$.

519 To mitigate the effect of PU contamination in jets two most commonly used techniques
 520 are charged hadron subtraction (CHS) and Pileup Per Particle Identification (PUPPI) [35].
 521 CHS as the name suggests removes all the PF in the jet clustering which are originating from
 522 PU vertices, and it is a standard technique for AK4 jets. PUPPI works by identifying PU
 523 in an event from charged PU information, then it assigns a weight to all the other particles
 524 inside jet, such as neutral particles, the weight is then used to rescale momentum of those
 525 particles. The main limitation of CHS is that it only removes charged PU contribution,
 526 for larger jets it can be issue, since it is clustering larger number of particles and can have
 527 significant contribution from neutral hadrons, for this reason PUPPI technique is used for
 528 AK8 jets.

529 To improve the jet selection and reject jets originating purely from PU two methods
 530 are used in this dissertation, jet identification based on multiplicities and energy fraction
 531 of particles contained in the jet, and MVA based PU identification which uses jets shape
 532 variables to discriminate prompt jet from pileup jets. The details of PU mitigation and
 533 identification used in CMS are described in Reference [36].

534 After jets reconstruction is complete the in addition to calculating kinematics properties
 535 (p_T , η , ϕ and mass), various other properties such as b-quark tagging and quark-gluon
 536 likelihood are also calculated and stored.

⁵³⁷ **3.2.3.1 N-Subjetiness and Deep Taggers**

⁵³⁸ The origin of fatjets are usually when heavy energetic particle often referred to as
⁵³⁹ “boosted” decays hadronically, for example boosted W or Z bosons decaying to a pair of
⁵⁴⁰ quarks. To find and discriminate the fatjet of interest based on its substructure the two
⁵⁴¹ technique studied and used are N-Subjetiness [37] and “deep tagger” [38].

N-Subjetiness is defined as,

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k}\} \quad (3.5)$$

where k runs over constituent particles in a jet, $\Delta R_{J,k}$ is the distance between subjet J and k constituent, and d_0 is the normalization constant defined as,

$$d_0 = \sum_k p_{T,k} R_0 \quad (3.6)$$

⁵⁴² τ_N quantifies to what degree a jet can be regarded as made of N jets. The small values of
⁵⁴³ the τ_N means a jet is more likely to have N or less subjets, and higher value means it will at
⁵⁴⁴ least have $N + 1$ subjets. Rather than using τ_N alone, ratio of different τ_N variables is used,
⁵⁴⁵ which more discriminating for cases like W vs QCD jets. Figure 3.1 shows distribution of τ_{21}
⁵⁴⁶ and τ_{32} shapes in signal and background. τ_{21} is used to discriminate fatjets with 2-prongs
⁵⁴⁷ (W/Z/H) and τ_{32} with 3-prongs (t-quark) substructure against QCD jets.

⁵⁴⁸ Deep Tagger for AK8 are Machine learning (ML) based tagger developed to determine
⁵⁴⁹ origin of a fatjet. These taggers are trained on particle level information from PF and
⁵⁵⁰ provide multi class tagging probabilities. In addition to there are version of these taggers is
⁵⁵¹ also which is de-correlated from the mass of jet, this is important for analysis including this

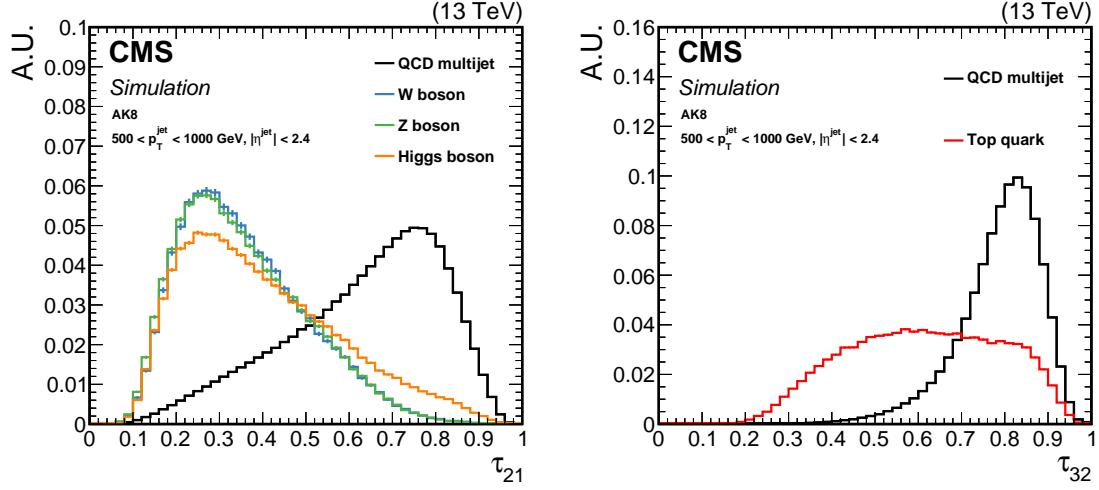


Figure 3.1: Comparison of τ_{21} and τ_{32} shapes for signal and background in AK8 jets. The left is τ_{21} distribution showing discrimination W/Z/H jets vs QCD jet, and the right is τ_{32} distribution for t-quark vs QCD jets [38].

- 552 dissertation where we utilize mass regions of fatjet to normalize background contribution.
 553 Figure 3.2 describes the architecture of “DeepAK8” tagging.

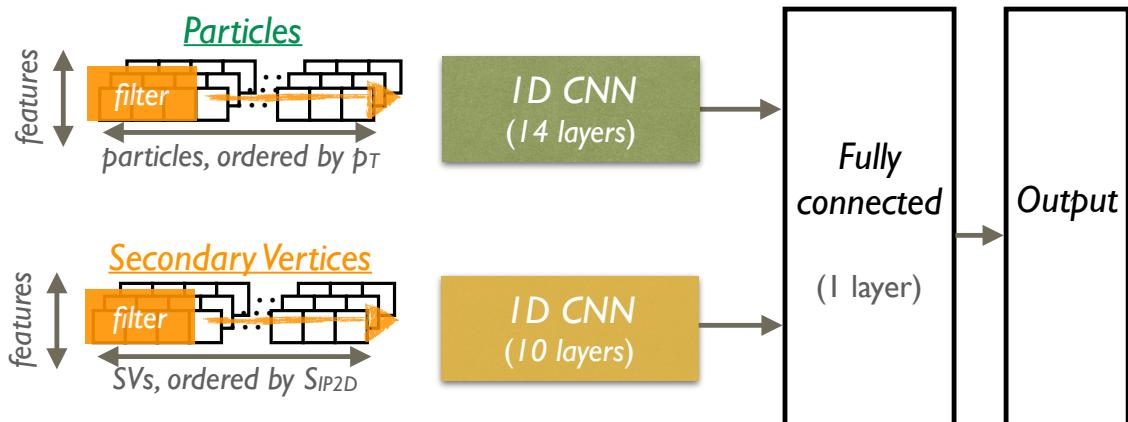


Figure 3.2: The network architecture of DeepAK8 [38]

⁵⁵⁴ **3.2.3.2 Softdrop Mass**

⁵⁵⁵ Fatjets can also have contamination coming from wide angle soft initial state radiation
⁵⁵⁶ (ISR) and multiple hadron scattering, which affects the mass calculation of the jet, to remove
⁵⁵⁷ such contamination and have better mass reconstruction, the “softdrop” mass algorithm [39]
⁵⁵⁸ is used.

Softdrop is a declustering algorithm which removes the particle from the jet with radius R_0 , when the following condition between two particles is satisfied,

$$\frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}} > z_{cut} \left(\frac{\Delta R_{12}}{R_0} \right)^\beta \quad (3.7)$$

⁵⁵⁹ where ΔR_{12} is the distance between the two particles, z_{cut} and β are the parameters for
⁵⁶⁰ tuning softdrop declustering. For fatjets used in CMS, they have softdrop applied with
⁵⁶¹ $\beta = 0$ and $z_{cut} = 0.1$ which vetoes both soft and soft-collinear emissions in a jet.

⁵⁶² **3.2.4 Missing transverse momentum**

⁵⁶³ Invisible particles like neutrinos cannot be detected at CMS directly. Kinematics of
⁵⁶⁴ such particles can be determined using laws of conservation of total momentum. In case
⁵⁶⁵ of proton-proton collision, the actual collision happens between quarks contained in proton
⁵⁶⁶ and quarks carry fraction of proton momentum and can not be determined exactly, for this
⁵⁶⁷ reasons kinematic determination of invisible particles is limited to transverse plane only.

After all the particles are reconstructed in an event, their p_T 's can be used to determine missing transverse momentum as,

$$\vec{p}_T^{miss} = - \sum \vec{p}_T \quad (3.8)$$

568 It's usually neutrinos which contributes to missing transverse momentum, and they have
 569 very small mass, the missing transverse momentum is then equivalent to Missing Transverse
 570 Energy (MET), which is most often used term in physics analysis.

571 3.3 TBD Monte Carlo Simulation

572 3.3.1 TBD Generators

573 3.3.2 TBD Hadronization

CHAPTER 4

VBS MEASUREMENT IN ZVJJ FINAL STATE

576 As discussed in Section 1.2 this analysis targets VBS of ZV with two jets in final state.
 577 The goal of the analysis is to reduce contribution from background processes as much as
 578 possible without loosing much of signal, and measure signal strength and significance.

579 Since Z decays leptonically and V is decaying hadronically, the phase-space of this analysis
 580 can be either l^+l^-jjjj or l^+l^-Jjj , where l are leptons, j are narrow jets and J is a boosted
 581 (wider) jet. The phase-space is divided into two broad regions signal and controlled, signal
 582 region is constructed based on theory such that it is mostly signal process and controlled
 583 region is basically orthogonal to the signal region where we expect contributions mostly from
 584 background processes.

585 The analysis is performed “blind” to avoid intrinsic bias i.e. until the analysis procedure
 586 is finalized, the collision data is only used in controlled regions. Once the analysis technique
 587 is optimized using MC samples and validated against collision data in controlled regions.
 588 Once the analysis technique is satisfactory and approved by CMS Physics Group then the
 589 results are “un-blinded” i.e. measurements are done using collision data in signal region.

4.1 Dataset and Simulation

591 As discussed in Section 2.2.7 only events those pass Level-1 Trigger and HLT paths
 592 are saved for further processing, MC simulation also have these identical step during event
 593 generation to mimic Level-1 Trigger and HLT paths.

594 CMS collaboration processes the datasets centrally and provides various tiers of datasets
 595 such as “RECO” datasets, which contains reconstructed objects and no skimming. Average
 596 size of an event saved at “RECO” tier is 480 kB per event and on average an analysis will
 597 process 3 billion events, which makes this tier not practical in terms of computer processing
 598 time and storage if each analysis starts from “RECO”. CMS centrally processes these datasets
 599 further and removes certain objects or features to reduce the average event size but still
 600 covering majority of the analysis to make use of the reduced datasets.

601 This analysis uses NANOAOD tier with version “v7” of datasets which has average event
 602 size of 2–3 kB.

603 **4.1.1 Data**

604 The collision data events used in this analysis are all certified by CMS Data Quality
 605 Monitoring (DQM) and Data Certification (DC) group. The primary trigger object in HLT
 606 paths are leptons p_T threshold, and since in our final state we are looking for Z boson decaying
 607 into two leptons, we require single and double lepton trigger for our analysis. Depending
 608 on the detector conditions and LHC storage capacity we have slightly different threshold in
 609 triggers across different years. The Table 4.1 contains the list of HLT paths used in this
 610 analysis.

611 **4.1.2 MC Simulations**

612 The EW VBS process which is our signal is generated with MADGRAPH5+PYTHIA8 at
 613 leading order (LO) with α_{EW}^6 order i.e. all vertices in tree level Feynman diagram are EW ver-
 614 tices. The QCD induced VBS background process which is very similar to our signal is gener-

Table 4.1: Trigger paths used to select events in CMS collision data

Dataset	Year	HLT Path
Single Muon	2016	HLT_IsoMu24 HLT_IsoTkMu24
	2017	HLT_IsoMu27
	2018	HLT_IsoMu24
Single Electron	2016	HLT_Ele27_WPTight_Gsf HLT_Ele25_eta2p1_WPTight_Gsf
	2017	HLT_Ele35_WPTight_Gsf
	2018	HLT_Ele32_WPTight_Gsf
Double Muon	2016	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ HLT_TkMu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ
	2017	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8 HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8
	2018	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8
	2016	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL
Double Electron	2017	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL
	2018	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL

615 ated with same configuration but with $\alpha_{EW}^4 \alpha_{QCD}^2$ order i.e. two of the six vertices are QCD.
 616 The dominant background to analysis Drell-Yan (DY) + Jets and, top-quark based processes
 617 consisting of single top-quark (t and s-channel), single top-quark in association with W boson
 618 (tW) and top-quark pair (tt) production. The DY + Jets are generated at LO with MAD-
 619 GRAPH5+PYTHIA8 in bins of HT i.e. scalar sum of all the jets p_T in the event, to have more
 620 statistics for higher HT bins. Top-quark background process are generated at next-to-next-
 621 to-leading order (NNLO), s-channel is generated with MADGRAPH_MC@NLO5+PYTHIA8
 622 and others t-channel, tW, tt are generated with POWHEG+PYTHIA8. The complete list of

₆₂₃ Table 4.2, 4.3, 4.4 contains the list of Signal and Background MC samples used for modeling
₆₂₄ in this analysis.

Table 4.2: List of MC samples for Signal and Background modeling

Process	Year	Cross Section (pb)	
VBS_EWK (Signal)	2016	WminusTo2JZTo2LJJ_EWK_LO_SM_MJJ100PTJ10_TuneCUETP8M1_13TeV-madgraph-pythia8	0.02982
	2016	WplusTo2JZTo2LJJ_EWK_LO_SM_MJJ100PTJ10_TuneCUETP8M1_13TeV-madgraph-pythia8	0.05401
	2016	ZTo2LZTo2JJ_EWK_LO_SM_MJJ100PTJ10_TuneCUETP8M1_13TeV-madgraph-pythia8	0.01589
	2017, 2018	WminusTo2JZTo2LJJ_EWK_LO_SM_MJJ100PTJ10_TuneCP5_13TeV-madgraph-pythia8	0.02982
	2017, 2018	WplusTo2JZTo2LJJ_EWK_LO_SM_MJJ100PTJ10_TuneCP5_13TeV-madgraph-pythia8	0.05401
	2017, 2018	ZTo2LZTo2JJ_EWK_LO_SM_MJJ100PTJ10_TuneCP5_13TeV-madgraph-pythia8	0.01589
	2016	WminusTo2JZTo2LJJ_QCD_LO_SM_MJJ100PTJ10_TuneCUETP8M1_13TeV-madgraph-pythia8	0.3488
	2016	WplusTo2JZTo2LJJ_QCD_LO_SM_MJJ100PTJ10_TuneCUETP8M1_13TeV-madgraph-pythia8	0.575
	2016	ZTo2LZTo2JJ_QCD_LO_SM_MJJ100PTJ10_TuneCUETP8M1_13TeV-madgraph-pythia8	0.3449
	2017, 2018	WminusTo2JZTo2LJJ_QCD_LO_SM_MJJ100PTJ10_TuneCP5_13TeV-madgraph-pythia8	0.3488
VBS_QCD (Background)	2017, 2018	WplusTo2JZTo2LJJ_QCD_LO_SM_MJJ100PTJ10_TuneCP5_13TeV-madgraph-pythia8	0.3488
	2017, 2018	ZTo2LZTo2JJ_QCD_LO_SM_MJJ100PTJ10_TuneCP5_13TeV-madgraph-pythia8	0.575
	2017, 2018	ZTo2LZTo2JJ_QCD_LO_SM_MJJ100PTJ10_TuneCP5_13TeV-madgraph-pythia8	0.3449
	2016	DYJetsToLL_M-50_HT-70to100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	169.9
	2016	DYJetsToLL_M-50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	147.4
DY + Jets LO (Background)	2016	DYJetsToLL_M-50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	41.04
	2016	DYJetsToLL_M-50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	5.674
	2016	DYJetsToLL_M-50_HT-600to800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.358
	2016	DYJetsToLL_M-50_HT-800to1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.6229
	2016	DYJetsToLL_M-50_HT-1200to2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.1512
	2016	DYJetsToLL_M-50_HT-2500toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.003659
	2017	DYJetsToLL_M-50_HT-70to100_TuneCP5_13TeV-madgraphMLM-pythia8	167.33
	2017	DYJetsToLL_M-50_HT-100to200_TuneCP5_13TeV-madgraphMLM-pythia8	161.1
	2017	DYJetsToLL_M-50_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8	48.66
	2017	DYJetsToLL_M-50_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8	6.968
	2017	DYJetsToLL_M-50_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8	1.743
	2017	DYJetsToLL_M-50_HT-800to1200_TuneCP5_13TeV-madgraphMLM-pythia8	0.8052
	2017	DYJetsToLL_M-50_HT-1200to2500_TuneCP5_13TeV-madgraphMLM-pythia8	0.1933
	2017	DYJetsToLL_M-50_HT-2500toInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.003468

Table 4.3: List of MC samples for Signal and Background modeling

Process	Year	Dataset Name	Cross Section (pb)
DY + Jets LO (Background)	2018	DYJetsToLL_M-50_HT-70to100_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	167.33
	2018	DYJetsToLL_M-50_HT-100to200_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	161.1
	2018	DYJetsToLL_M-50_HT-200to400_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	48.66
	2018	DYJetsToLL_M-50_HT-400to600_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	6.968
	2018	DYJetsToLL_M-50_HT-600to800_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	1.743
	2018	DYJetsToLL_M-50_HT-800to1200_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.8052
	2018	DYJetsToLL_M-50_HT-1200to2500_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.1933
	2018	DYJetsToLL_M-50_HT-2500toInf_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.003468
	2016	ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV_PSweights-powheg-pythia8	38.06
	2016	ST_tW_antitop_5f_inclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8	34.97
Top (Background)	2016	ST_tW_top_5f_NoFullyHadronicDecays_13TeV_PSweights-powheg-pythia8	38.09
	2016	ST_tW_top_5f_inclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8	34.91
	2016	ST_t-channel1_antitop_4f_InclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8	67.91
	2016	ST_t-channel1_top_4f_InclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8	113.3
	2016	ST_s-channel1_4f_leptonDecays_13TeV_PSweights-amcatnlo-pythia8	3.365
	2016	ST_s-channel1_4f_hadronicDecays_TuneCP5_PSweights_13TeV-amcatnlo-pythia8	11.24
	2016	ST_s-channel1_4f_InclusiveDecays_13TeV-amcatnlo-pythia8	10.12
	2016	TTToHadronic_TuneCP5_PSweights_13TeV-powheg-pythia8	377.96
	2016	TTToSemiLeptonic_TuneCP5_PSweights_13TeV-powheg-pythia8	365.34
	2016	TTTo2L2Nu_TuneCP5_PSweights_13TeV-powheg-pythia8	88.29
	2017	TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8	365.34
	2017	TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	86.99
	2017	TTToHadronic_TuneCP5_13TeV-powheg-pythia8	377.96
	2017	ST_s-channel1_antitop_leptonDecays_13TeV-PSweights_powheg-pythia	1.33
	2017	ST_s-channel1_top_leptonDecays_13TeV-PSweights_powheg-pythia	2.13
	2017	ST_t-channel1_antitop_5f_TuneCP5_PSweights_13TeV-powheg-pythia8	27.19
	2017	ST_t-channel1_top_5f_TuneCP5_13TeV-powheg-pythia8	45.7
	2017	ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	12.04
	2017	ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	12.04

Table 4.4: List of MC samples for Signal and Background modeling

Process	Year	Dataset Name	Cross Section (pb)
Top (Background)	2018	TTToSemileptonic_TuneCP5_13TeV-powheg-pythia8	365.34
	2018	TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	86.99
	2018	TTToHadronic_TuneCP5_13TeV-powheg-pythia8	377.96
	2018	ST_s-channel_antitop_leptonDecays_13TeV-PSweights_powheg-pythia	1.33
	2018	ST_s-channel_top_leptonDecays_13TeV-PSweights_powheg-pythia	2.13
	2018	ST_t-channel_antitop_5f_TuneCP5_13TeV-powheg-pythia8	27.19
	2018	ST_t-channel_top_5f_TuneCP5_13TeV-powheg-pythia8	45.7
	2018	ST_tW_DS_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	12.04
	2018	ST_tW_DS_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	12.04

625

4.2 Event Selection

626 In first stage of selection events are selected if an events has minimum number of objects
 627 for analysis categories i.e. at least two leptons of same flavor ($p_T > 10 \text{ GeV}$), and either
 628 four narrow jets (AK4) (Resolved ZV category) or two narrow jets (AK4) plus one wider jet
 629 (AK8) (Boosted ZV category).

630 After initial skimming step, leptons are selected with OSSF for Z candidate with following
 631 selections for muon and electron channels:

632 • **Muons:** Muons with $p_T < 10 \text{ GeV}$, $|\eta| > 2.4$, failing loose ID, and PF relative isolation
 633 in cone $R = 0.4$ (“pfRelIso04”) more than 0.40 are vetoed [40]. Then exactly two
 634 tight muons with opposite charge are selected with $p_T > 20 \text{ GeV}$, passing tight ID,
 635 “pfRelIso04” less than 0.15 and impact parameters $d_{xy} < 0.01, d_z < 0.1$, for all the
 636 years.

637 • **Electrons:** Electrons with $p_T < 10 \text{ GeV}$, $|\eta| > 2.5$, and failing “cutBased_HLTPreSel”
 638 2016 or loose “cutBased” for 2017, 2018 [33]. Tight selection of electron is differ-
 639 ent 2016 and 2017–2018, with $p_T > 20 \text{ GeV}$ same for all years, for 2016 year elec-
 640 trons with passing “mvaSpring16GP_WP90” and, “pfRelIso03_all” less than 0.0571 for
 641 $|\eta| > 1.479$ and less than 0.0588 otherwise, for 2017 and 2018 years, electrons passing
 642 “mvaFall17V2Iso_WP90” and “pfRelIso03_all” less than 0.06 is required.

643 After lepton selection, Z candidates kinematics are calculated using p_T, η, ϕ , mass of the
 644 leptons and events with Z mass in range $[76, 106] \text{ GeV}$ are kept.

645 • VBS Tagged Jets

646 • V Jet Candidate

4.2.1 Boosted ZV DY+Jets Control Region

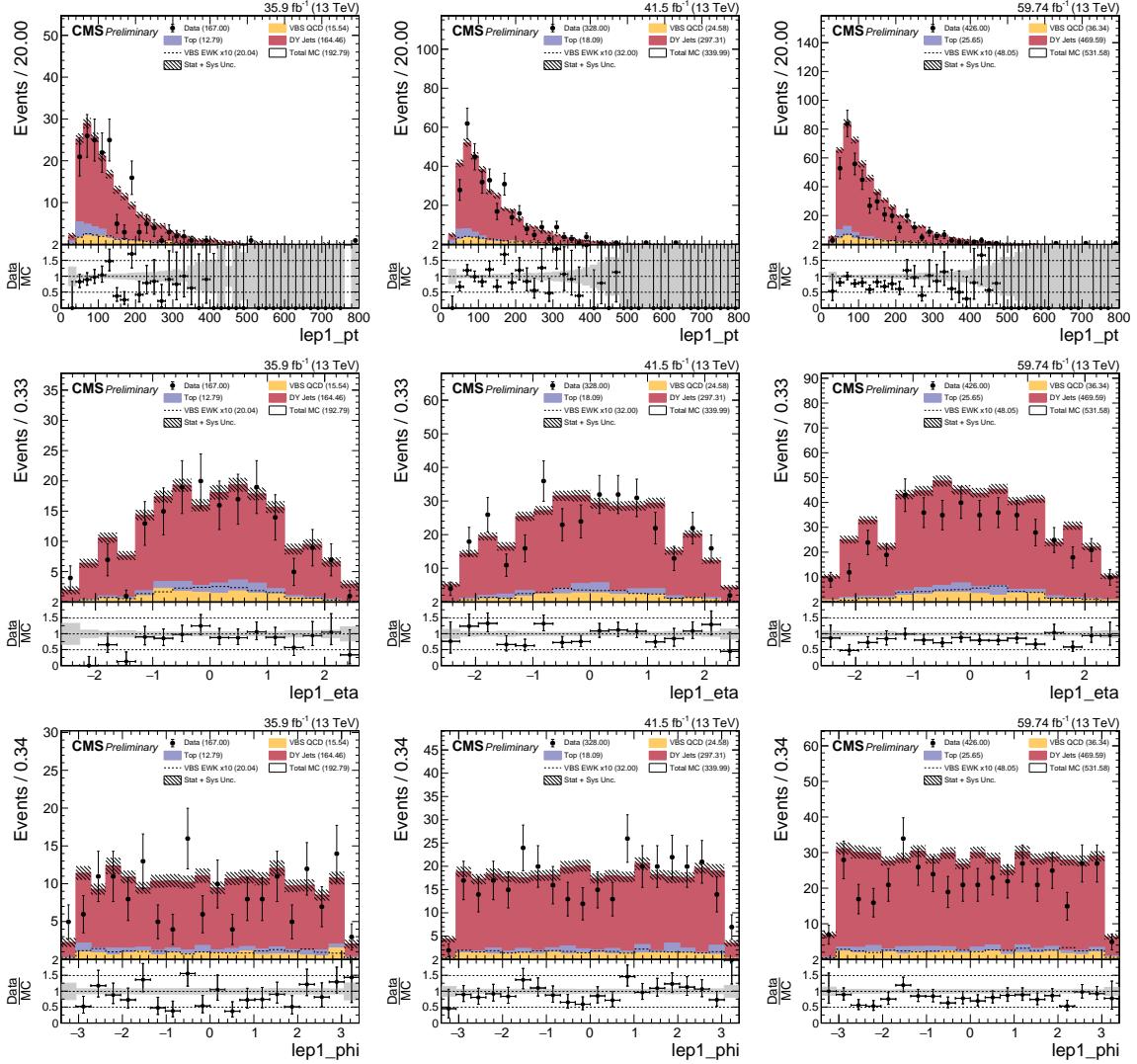


Figure 4.1: DY+Jets Control Region: Leading electron kinematics in Boosted ZV Channel. From Left to Right: 2016, 2017, and 2018. From Top to Bottom: p_T , η , ϕ .

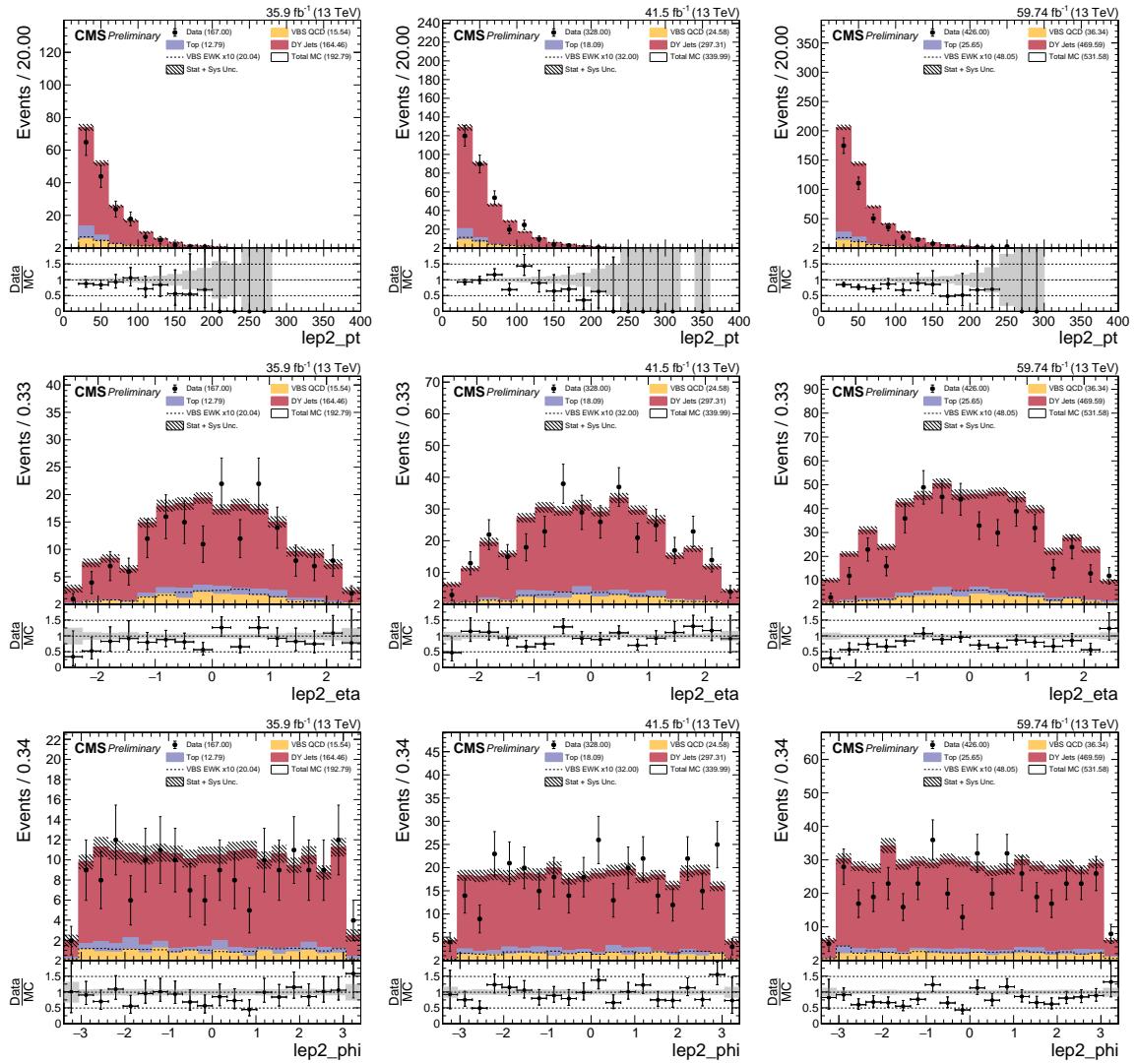


Figure 4.2: DY+Jets Control Region: Trailing electron kinematics in Boosted ZV Channel. From Left to Right: 2016, 2017, and 2018. From Top to Bottom: p_T , η , ϕ .

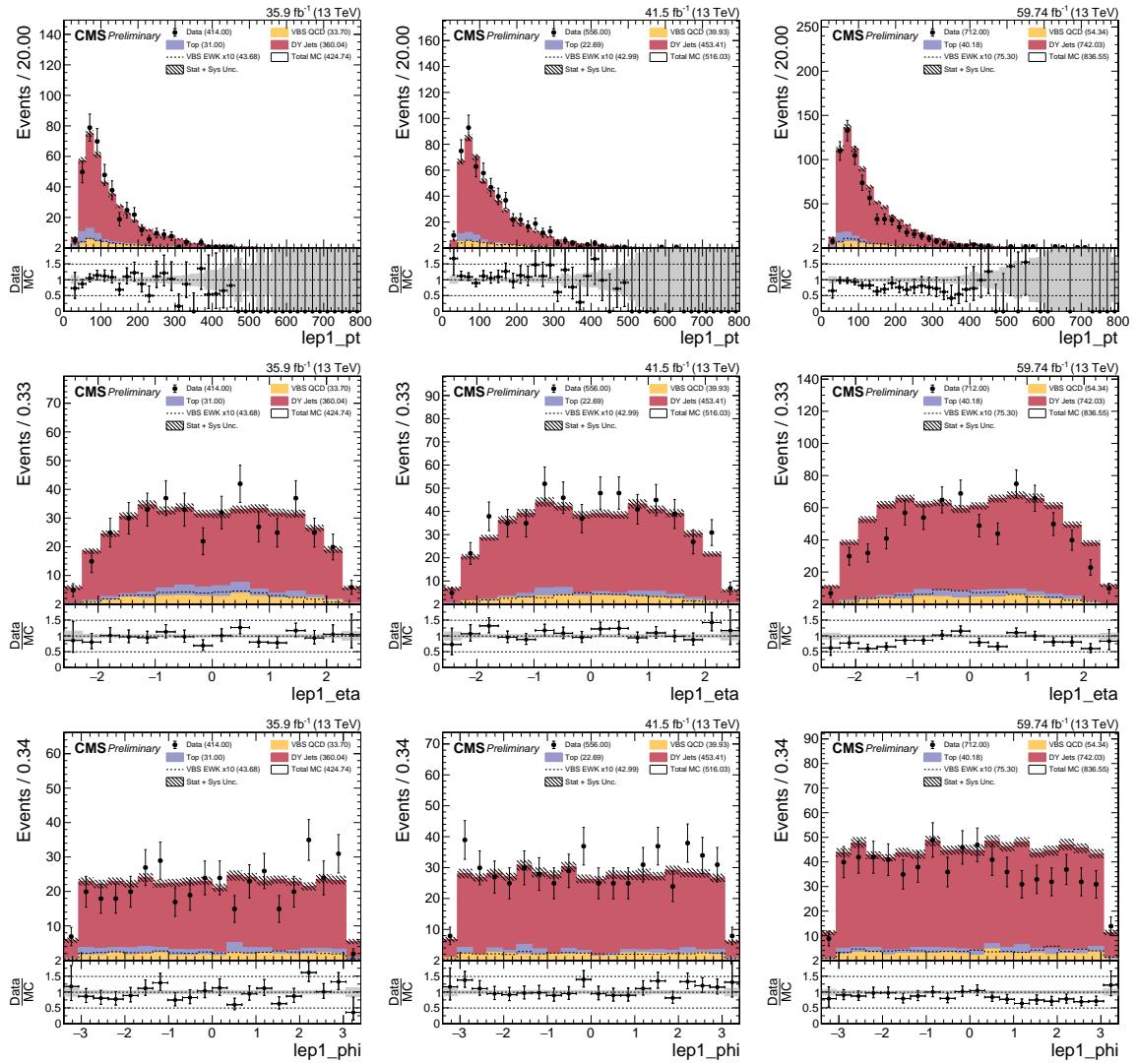


Figure 4.3: DY+Jets Control Region: Leading muon kinematics in Boosted ZV Channel. From Left to Right: 2016, 2017, and 2018. From Top to Bottom: p_T , η , ϕ .

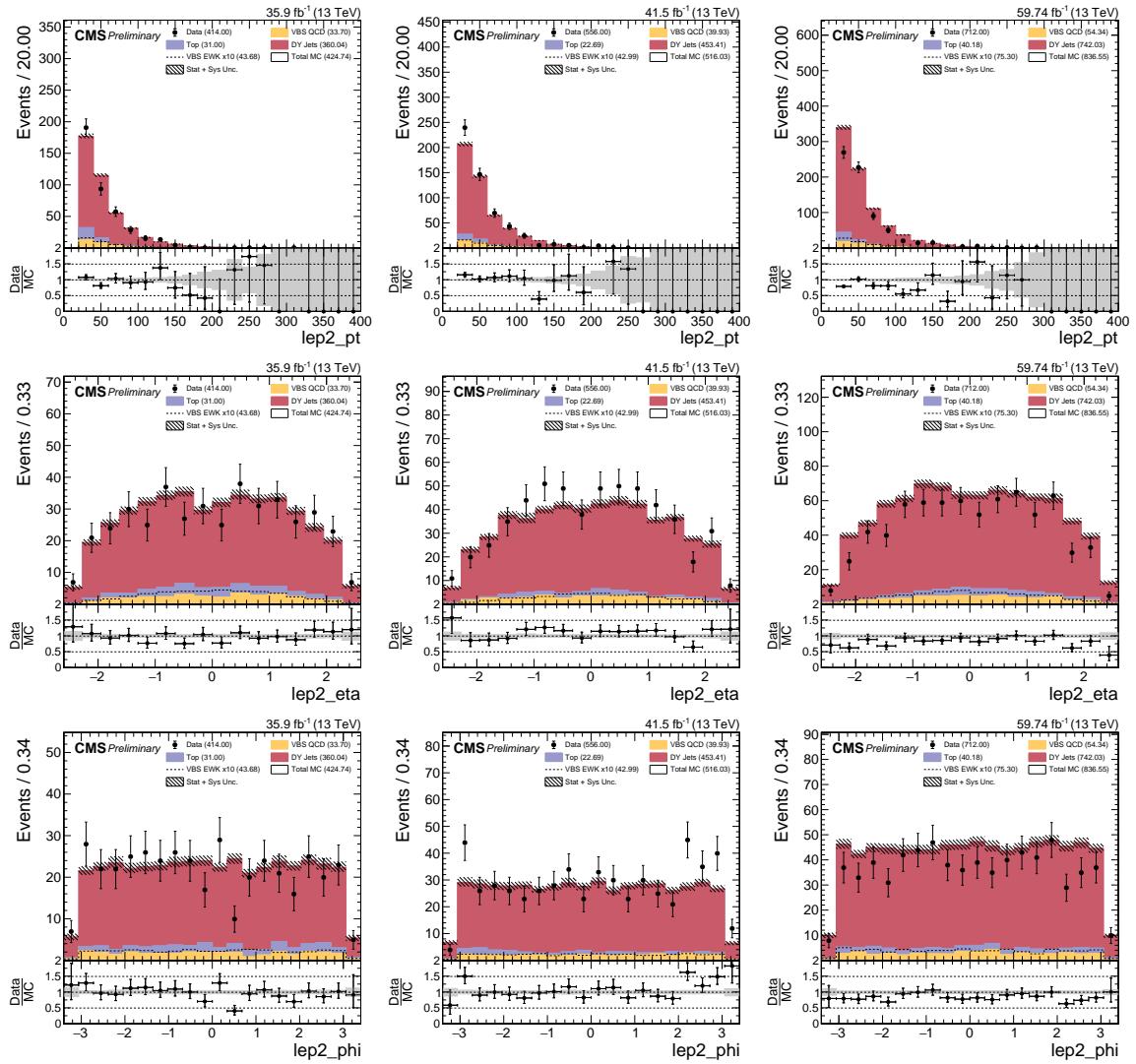


Figure 4.4: DY+Jets Control Region: Trailing muon kinematics in Boosted ZV Channel. From Left to Right: 2016, 2017, and 2018. From Top to Bottom: p_T , η , ϕ .

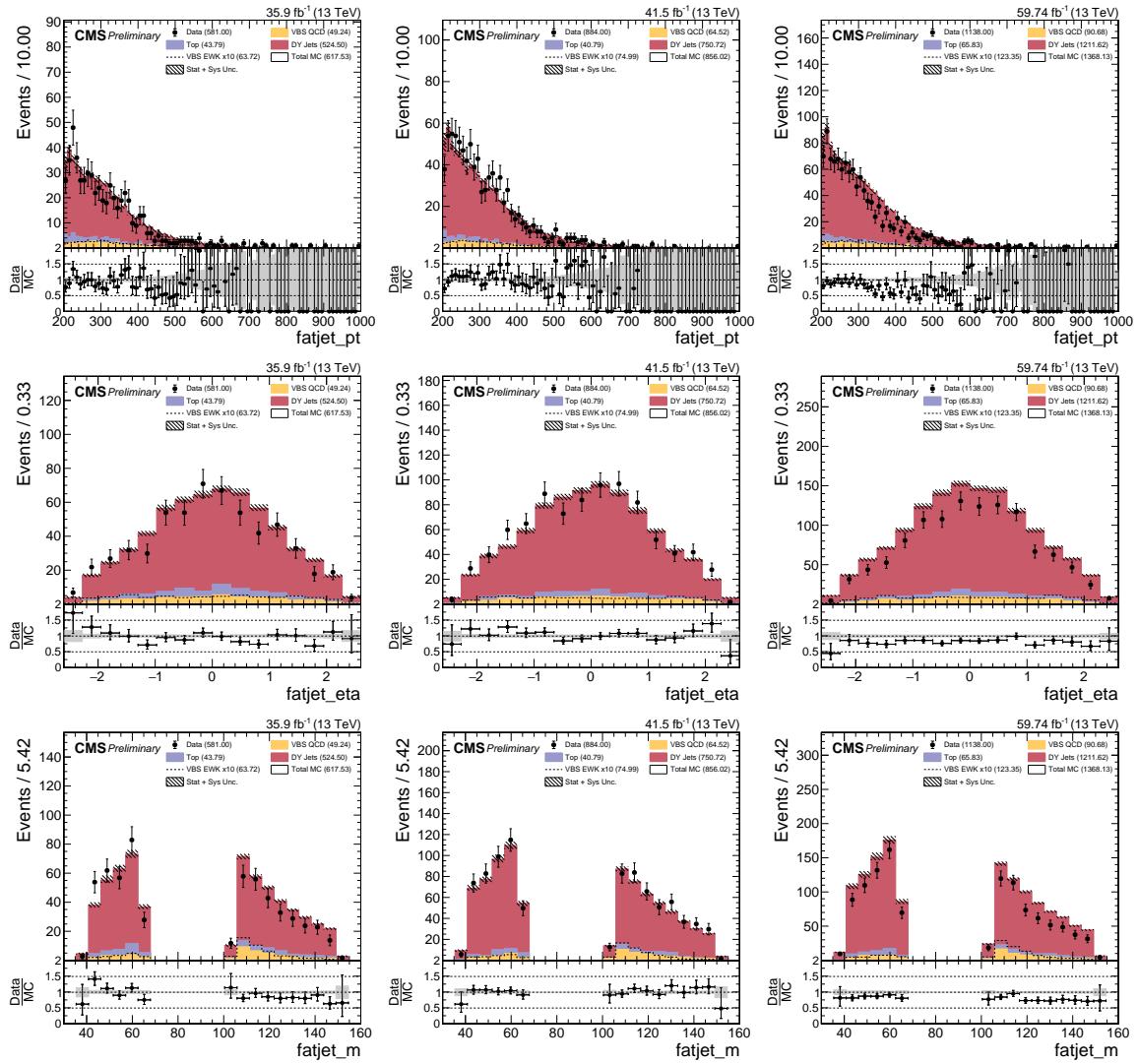


Figure 4.5: DY+Jets Control Region: Hadronic boson kinematics in Boosted ZV Channel. From Left to Right: 2016, 2017, and 2018. From Top to Bottom: p_T , η , mass m .

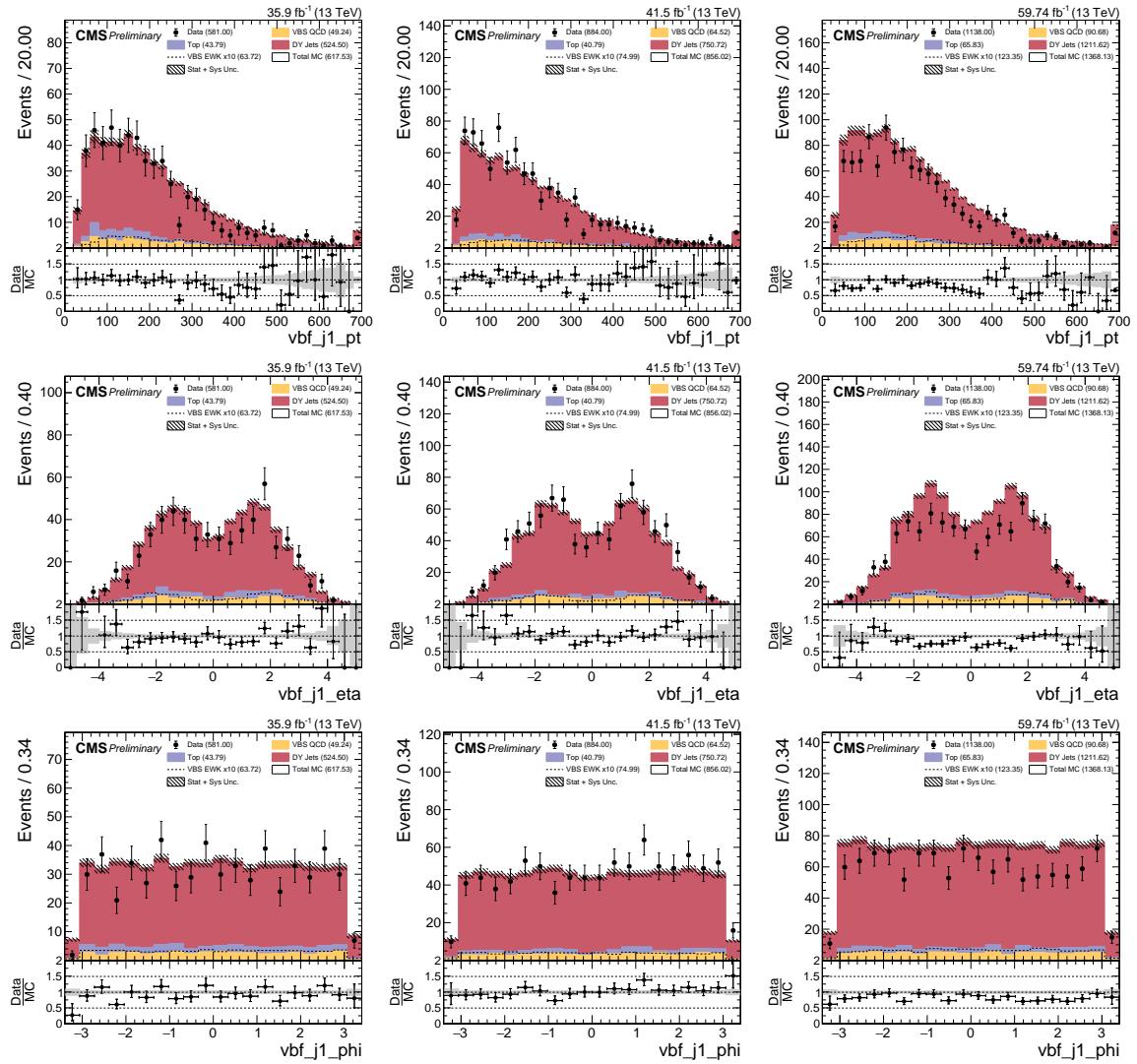


Figure 4.6: DY+Jets Control Region: Leading VBS tagged jet kinematics in Boosted ZV Channel. From Left to Right: 2016, 2017, and 2018. From Top to Bottom: p_T , η , ϕ .

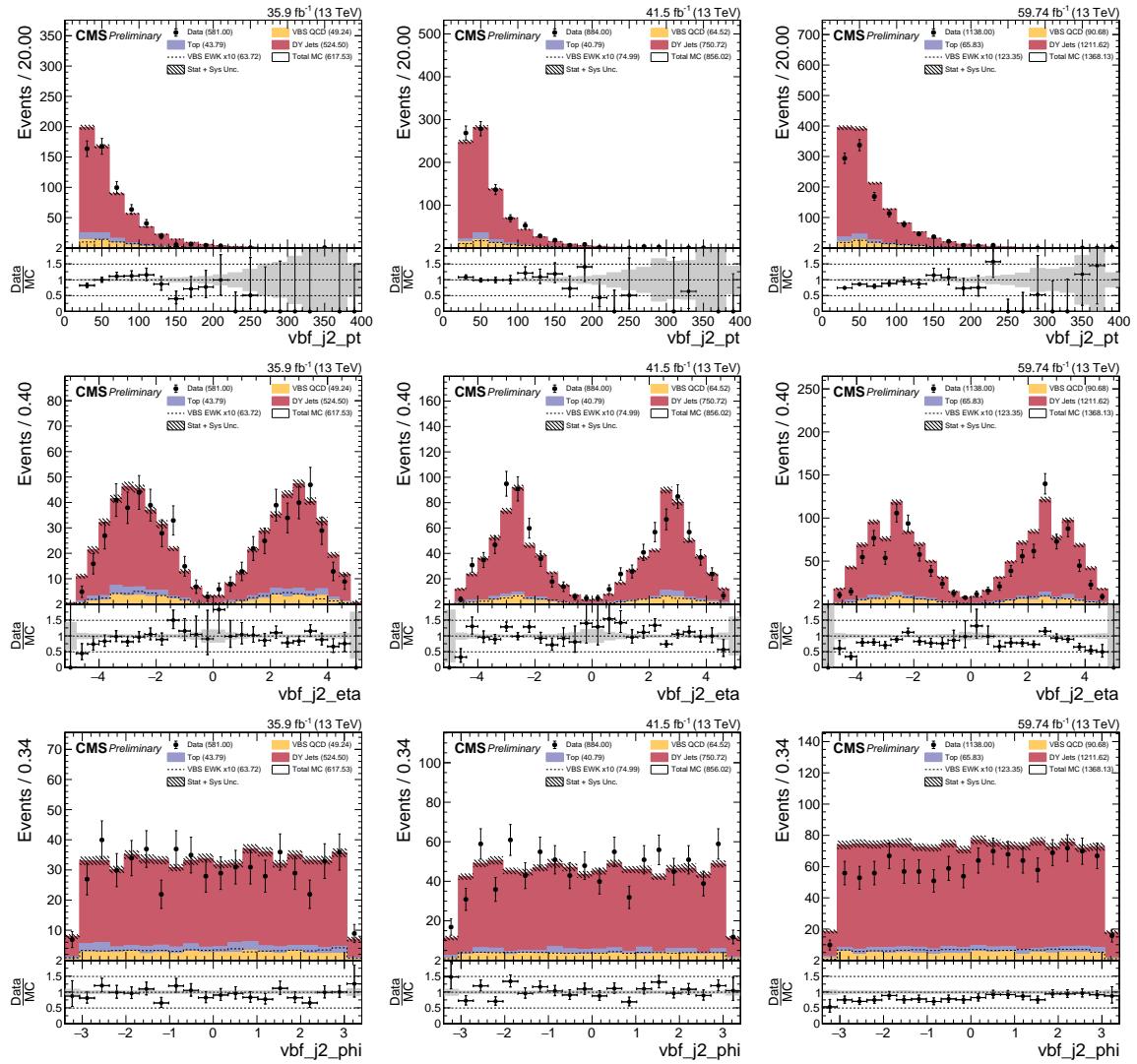


Figure 4.7: DY+Jets Control Region: Trailing VBS tagged jet kinematics in Boosted ZV Channel. From Left to Right: 2016, 2017, and 2018. From Top to Bottom: p_T , η , ϕ .

4.2.2 Resolved ZV DY+Jets Control Region

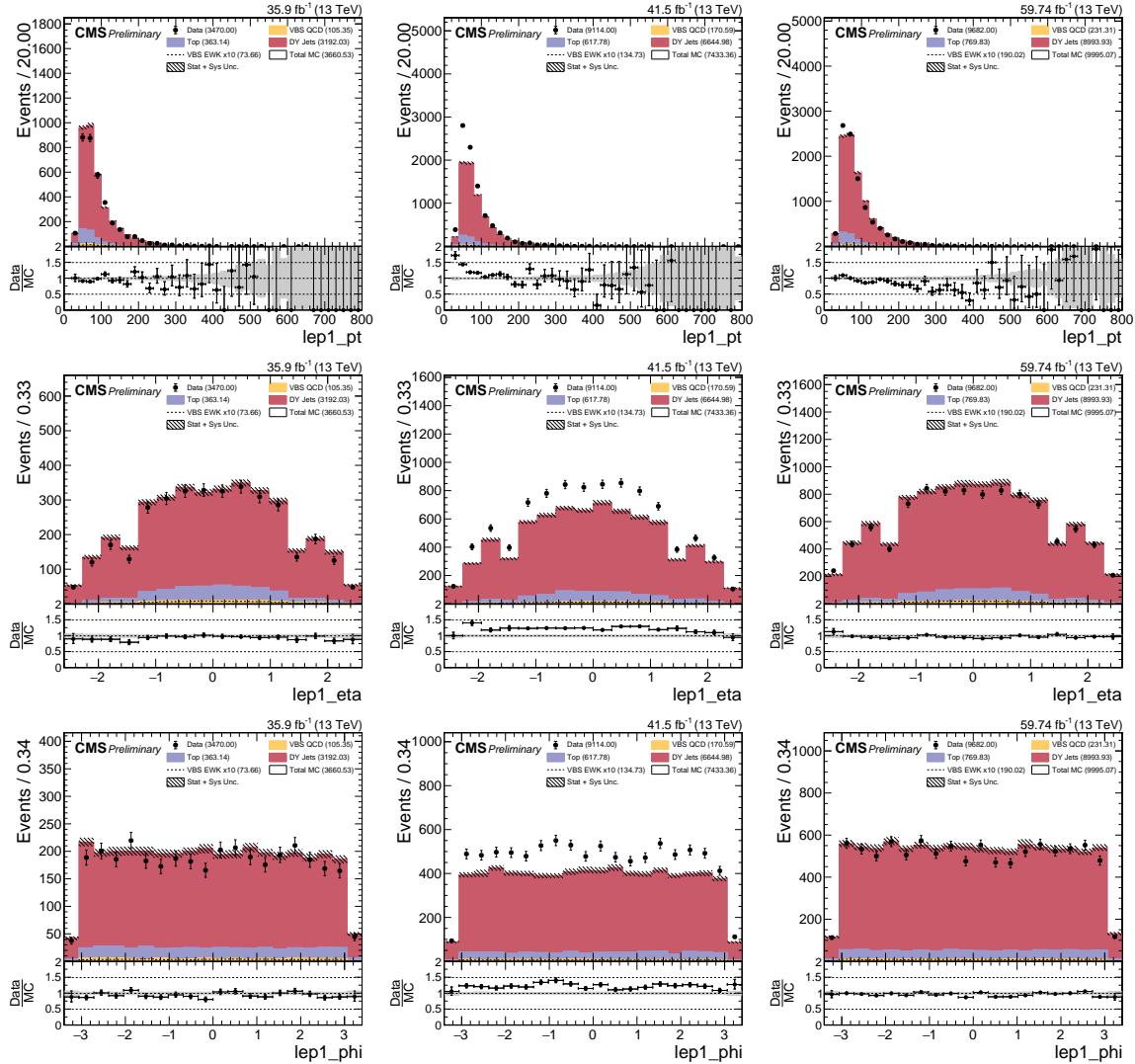


Figure 4.8: DY+Jets Control Region: Leading electron kinematics in Resolved ZV Channel. From Left to Right: 2016, 2017, and 2018. From Top to Bottom: p_T , η , ϕ .

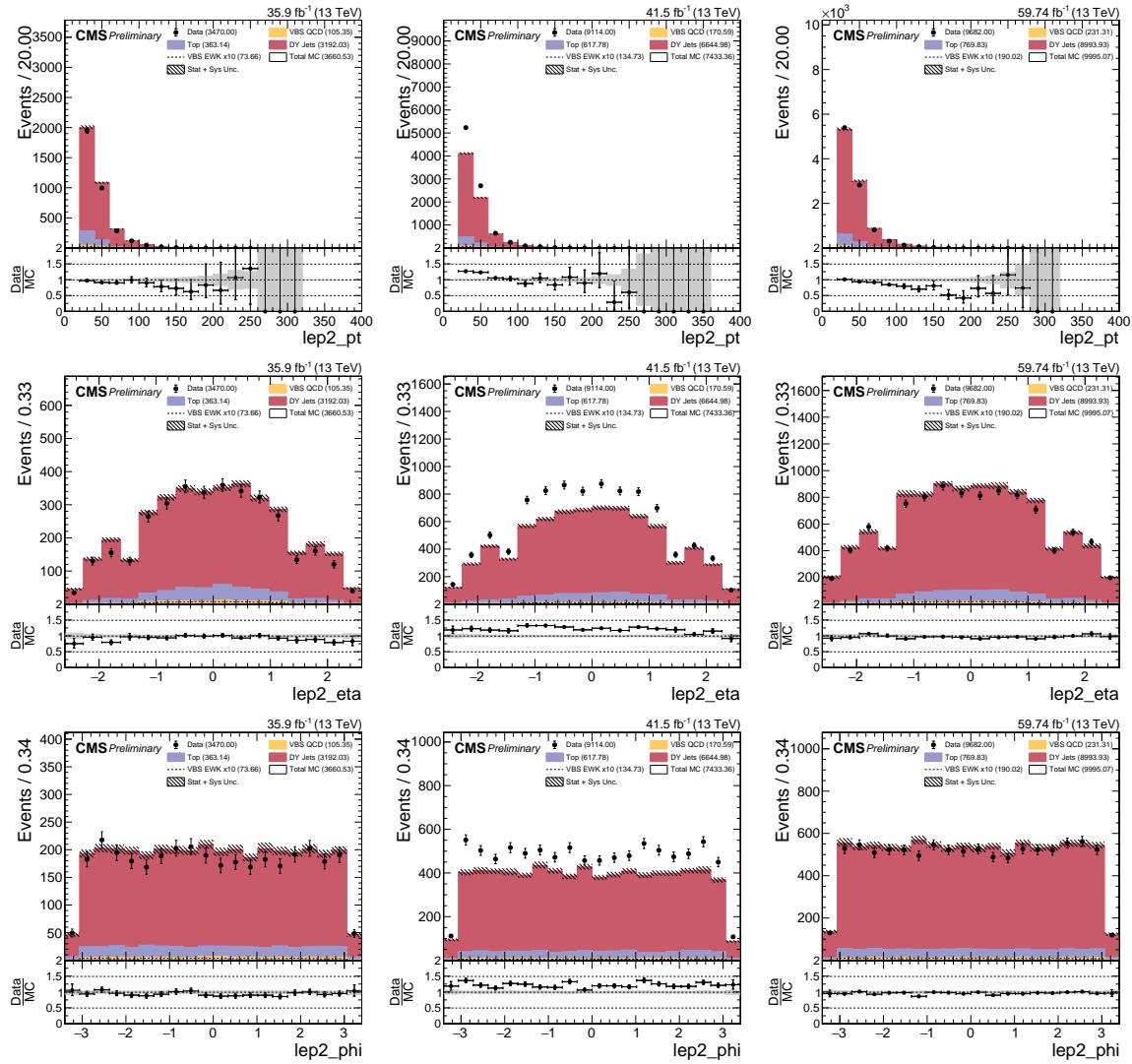


Figure 4.9: DY+Jets Control Region: Trailing electron kinematics in Resolved ZV Channel. From Left to Right: 2016, 2017, and 2018. From Top to Bottom: p_T , η , ϕ .

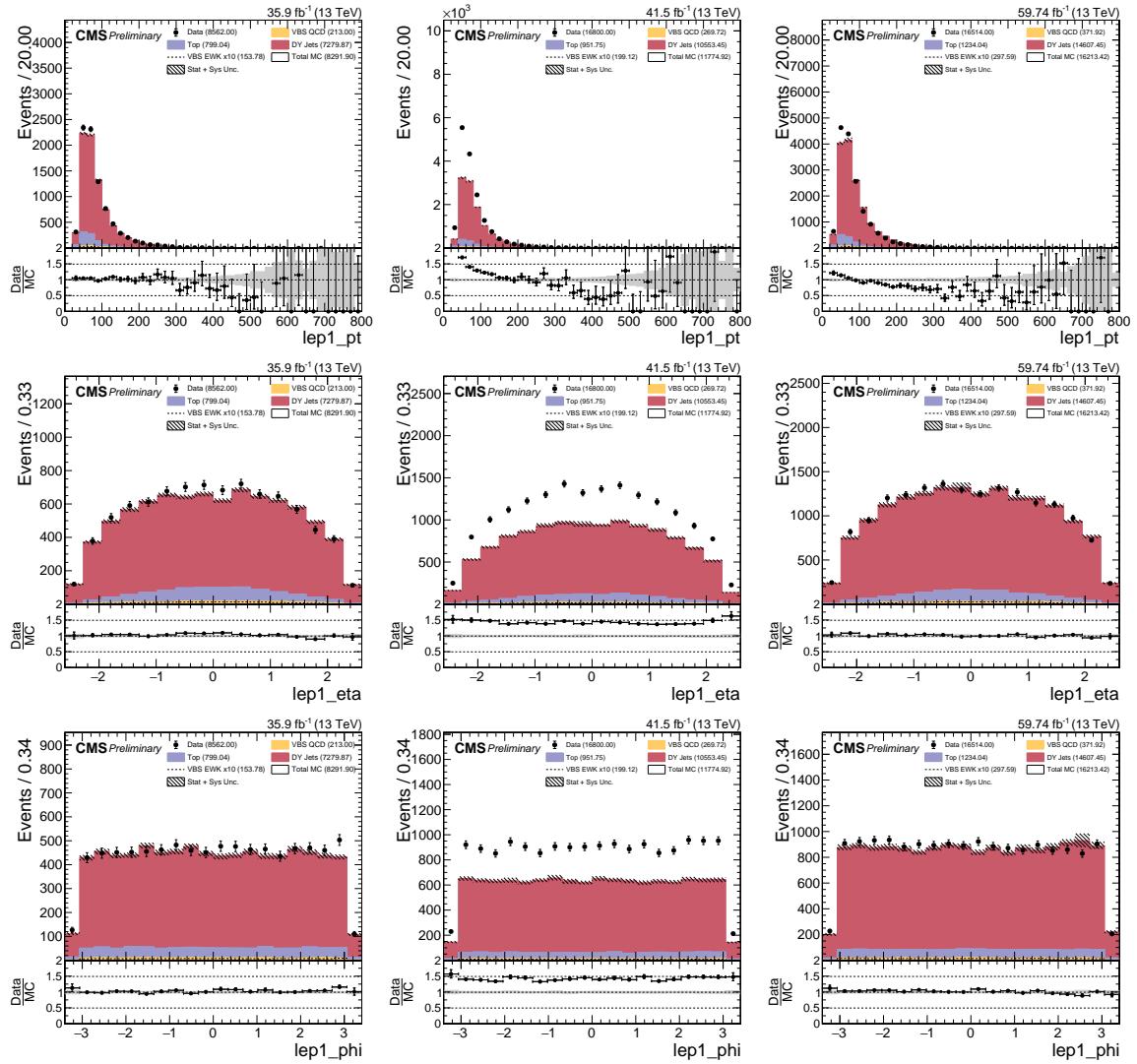


Figure 4.10: DY+Jets Control Region: Leading muon kinematics in Resolved ZV Channel. From Left to Right: 2016, 2017, and 2018. From Top to Bottom: p_T , η , ϕ .

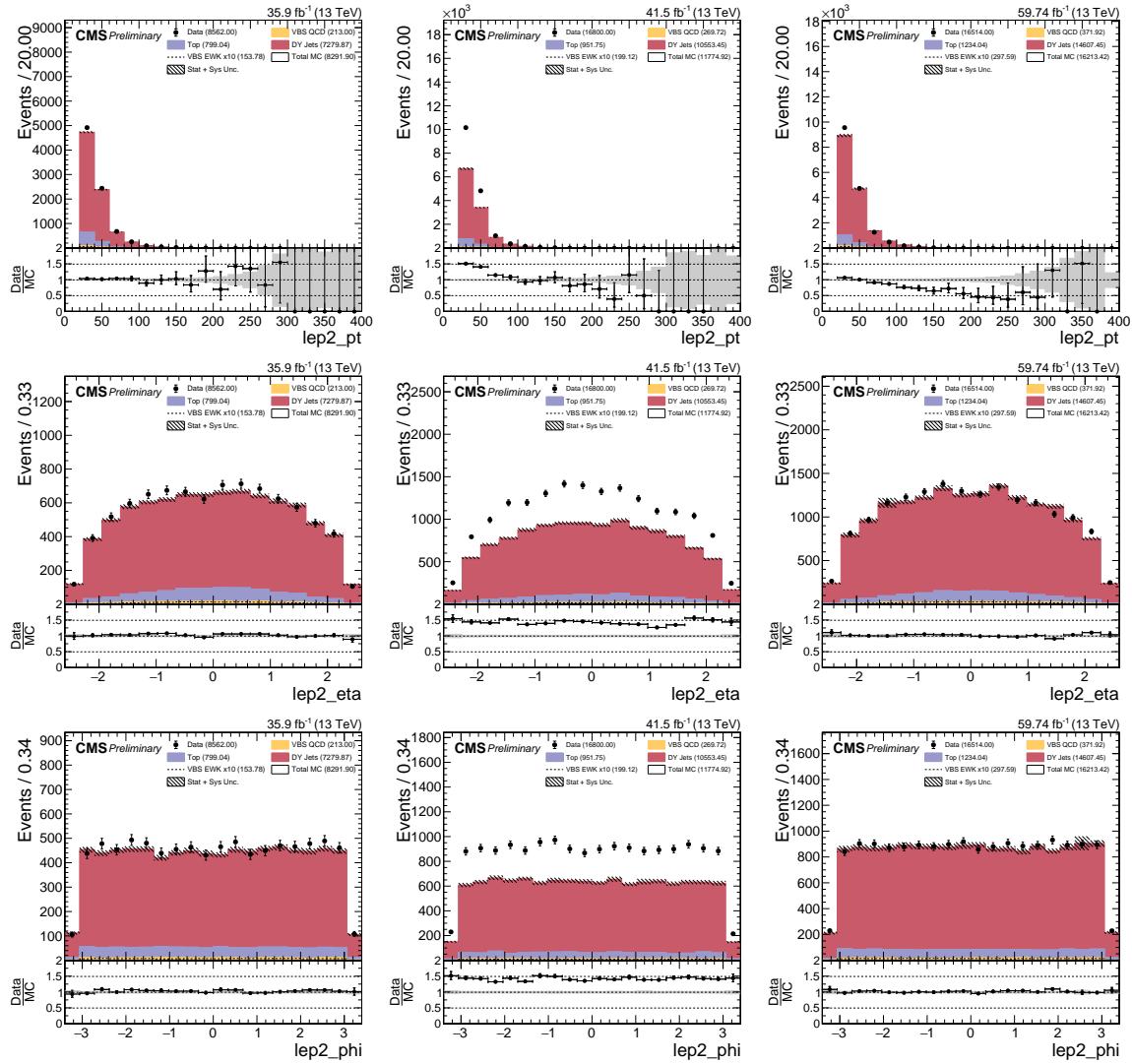


Figure 4.11: DY+Jets Control Region: Trailing muon kinematics in Resolved ZV Channel. From Left to Right: 2016, 2017, and 2018. From Top to Bottom: p_T , η , ϕ .

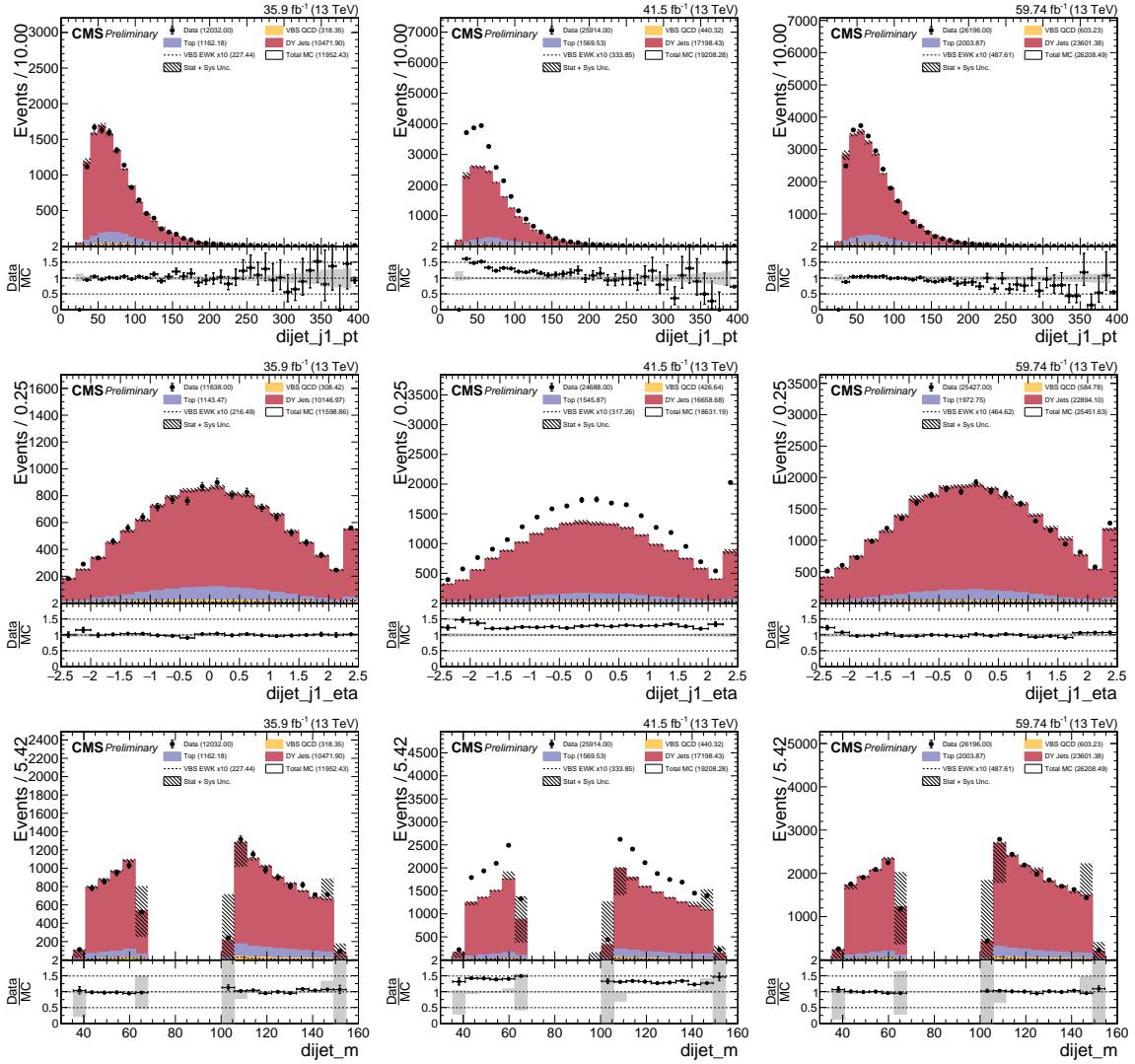


Figure 4.12: DY+Jets Control Region: Hadronic boson leading jet kinematics in Resolved ZV Channel. From Left to Right: 2016, 2017, and 2018. From Top to Bottom: p_T , η , invariant mass m_{jj} .

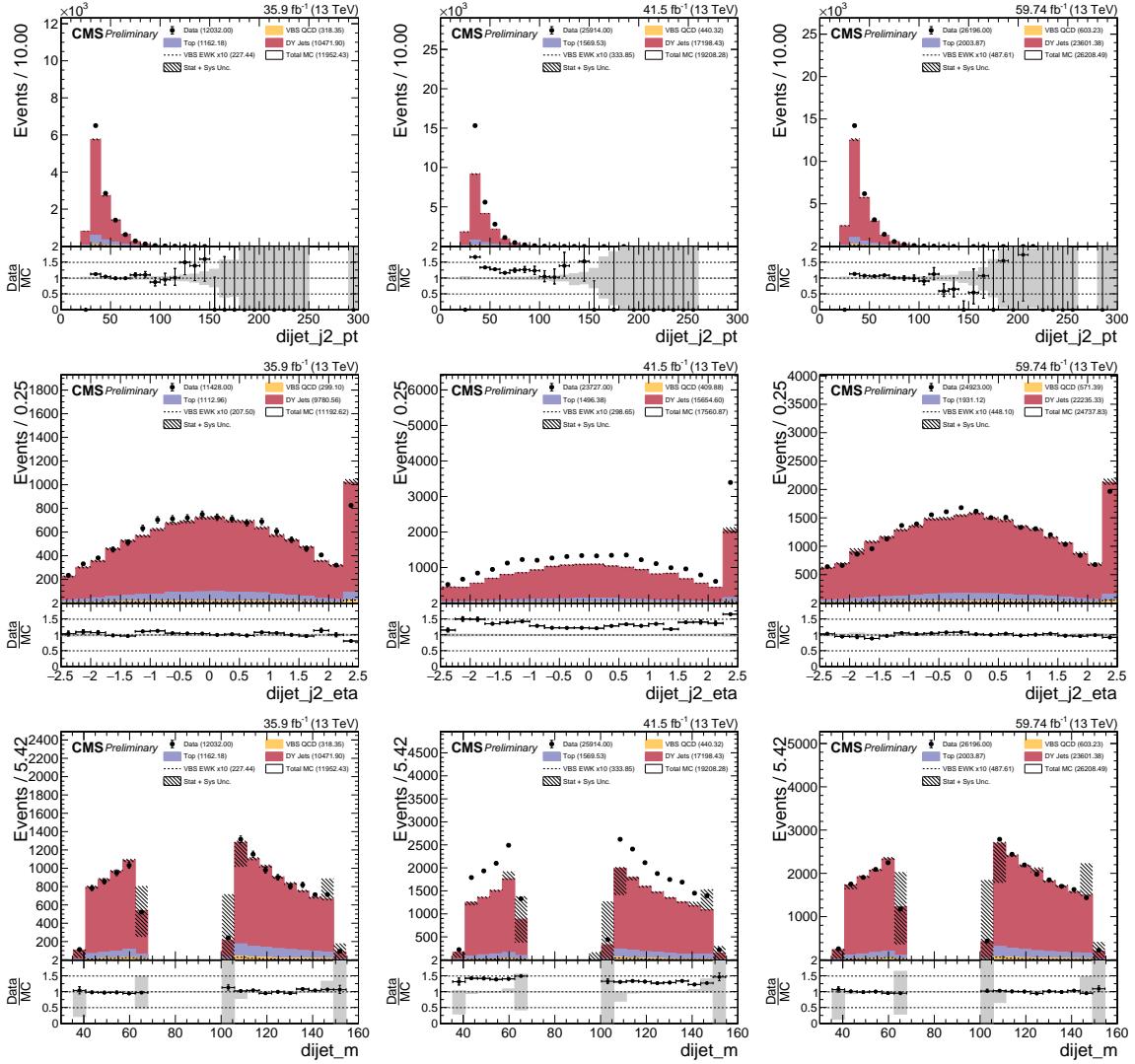


Figure 4.13: DY+Jets Control Region: Hadronic boson trailing jet kinematics in Resolved ZV Channel. From Left to Right: 2016, 2017, and 2018. From Top to Bottom: p_T , η , invariant mass m_{jj} .

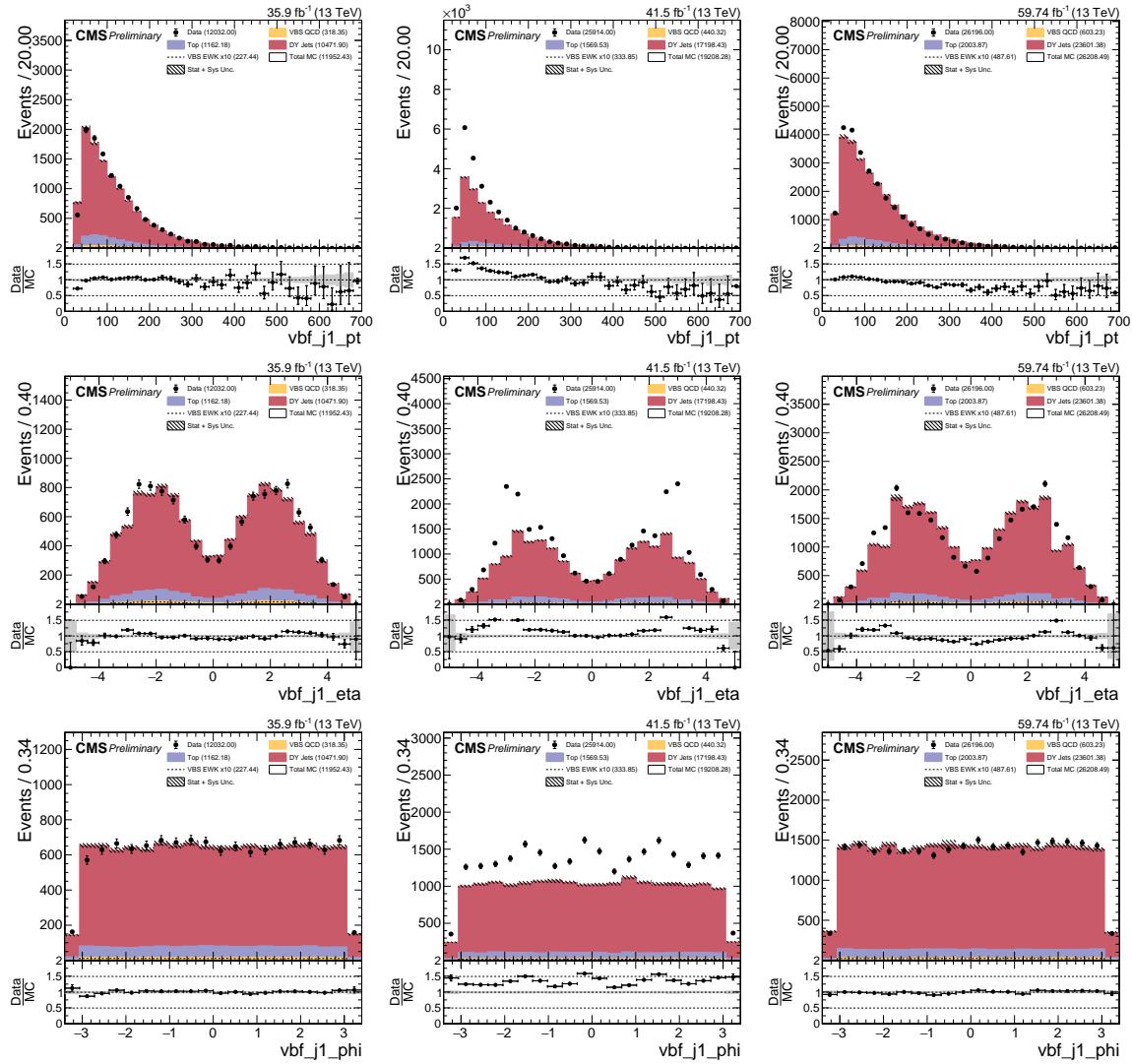


Figure 4.14: DY+Jets Control Region: Leading VBS tagged jet kinematics in Resolved ZV Channel. From Left to Right: 2016, 2017, and 2018. From Top to Bottom: p_T , η , ϕ .

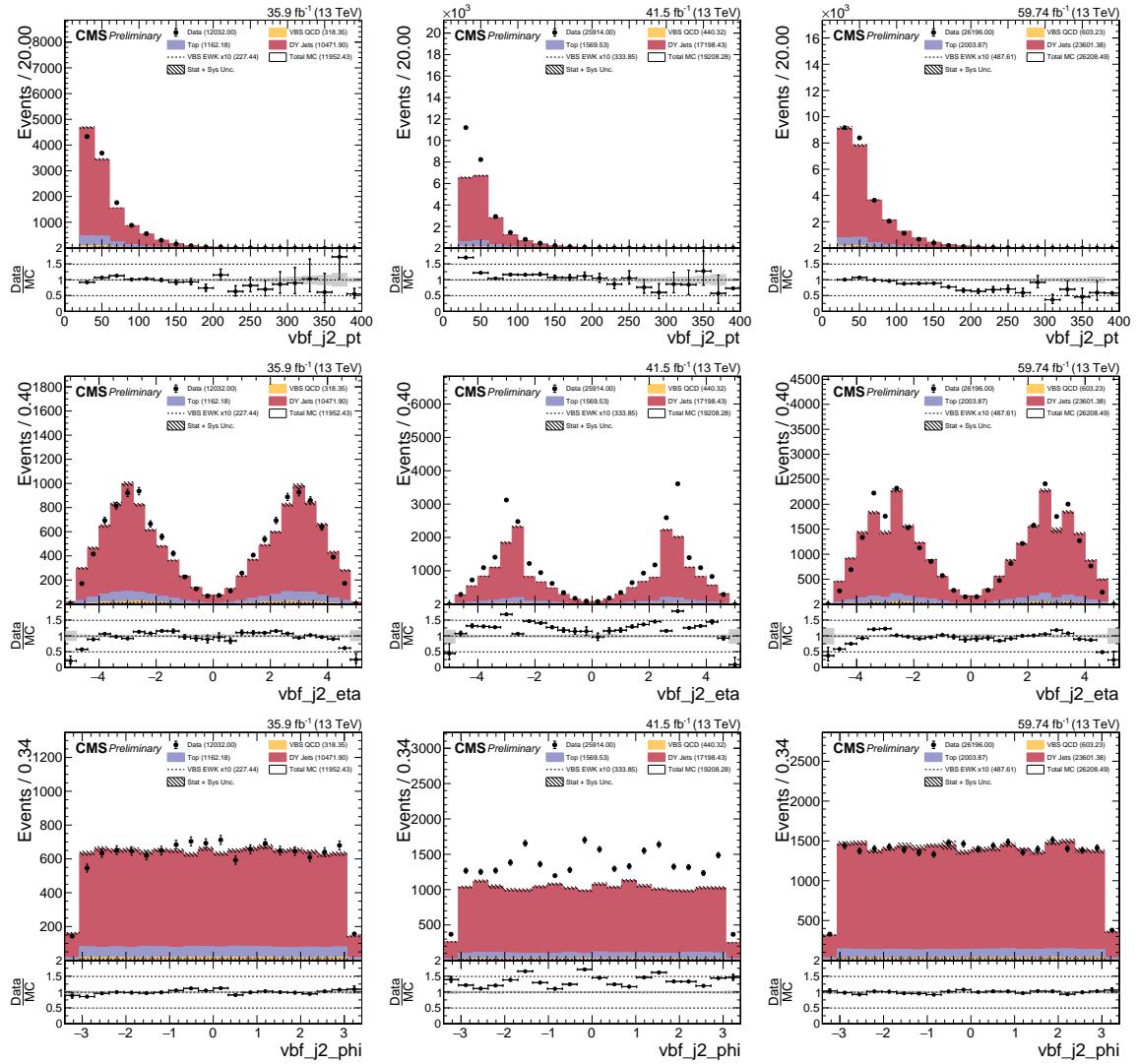


Figure 4.15: DY+Jets Control Region: Trailing VBS tagged jet kinematics in Resolved ZV Channel. From Left to Right: 2016, 2017, and 2018. From Top to Bottom: p_T , η , ϕ .

649

4.3 Machine Learning Modeling

650 Instead of traditional cut-based analysis, we decided to use MVA a.k.a. ML technique to
 651 build a signal vs background classifier. The main reasoning behind using a MVA technique
 652 is so that we can build a model which can learn our analysis topology from looser selection
 653 regions and still let us keep higher statistics for final measurement.

654

4.3.1 Algorithm: Gradient Boosted Decision Tree

655 Tools used Toolkit for Multivariate Analysis (TMVA) part of ROOT package (FIXME
 656 reference). FIXME some details mathematical details about BDT and specifically gradient
 657 BDT.

658

4.3.2 Training and Results

659 Two models were trained for boosted and resolved topology and the training was done
 660 using combined MC from all years 2016, 2017, 2018 to benefit from larger statistics see
 661 Table 4.5, signal MC “VBS_EWK” was trained against background “DY + Jets LO” since
 662 that is dominant background in our analysis. Each models Boosted decision tree (BDT)
 663 hyper-parameters were tunned to prevent under and over-fitting, and input variables used
 664 were also pruned in order of importance and keeping model metrics area under curve (AUC)
 665 of Receiver Operating Characteristic (ROC) relatively unaffected. Distribution of final input
 666 variables used in training are shown in Figure 4.16 and Figure 4.17.

Table 4.5: Training and Testing Statistics

		Number of Events		
Channel	Dataset	Training	Testing	Total
Boosted	Signal	7404	7405	14809
	Background	46991	46991	93982
Resolved	Signal	23425	23425	46850
	Background	209368	209368	418736

667 After training TMVA BDT evaluates input variables and ranks them in terms of impor-
 668 tance and separation they provide in classification is listed in Table 4.6, and the correlation
 669 matrix of variable is show in Figure 4.18.

Table 4.6: Training Input Variable Ranking

Channel	Variable	Variable Name	Importance	Separation
Boosted	M_{JJ}^{VBS}	vbf_m	0.2496	0.1348
	Zeppenfeld Z_V^*	zeppLep_deta	0.2396	0.1116
	QGL j_1^{VBS}	vbf2_AK4_qgid	0.1889	0.02413
	QGL j_2^{VBS}	vbf1_AK4_qgid	0.1780	0.02330
	M_{VV}	dibos_m	0.1439	0.005308
Resolved	Zeppenfeld Z_V^*	zeppLep_deta	0.1955	0.1219
	M_{JJ}^{VBS}	vbf_m	0.1822	0.07998
	HT^*	ht_resolved	0.1693	0.04201
	QGL j_1^{VBS}	vbf2_AK4_qgid	0.1403	0.02159
	QGL j_2^{VBS}	vbf1_AK4_qgid	0.1341	0.03235
	M_{VV}	dibos_m	0.09098	0.01112
	η_{lep2}	lep2_eta	0.08760	0.01755

670 The under and over-fitting of trained model is checked by Kolmogorov—Smirnov (K-S)
 671 test and ROC curves comparison between training and testing datasets. If the tests are not
 672 acceptable, then the training is redone with adjusted parameters. The Figure 4.19 show
 673 MVA score and ROC curves of the BDT models. Models final metrics are listed in Table 4.7.

Table 4.7: Models Metrics

Channel	AUC (%)
Boosted	78
Resolved	79

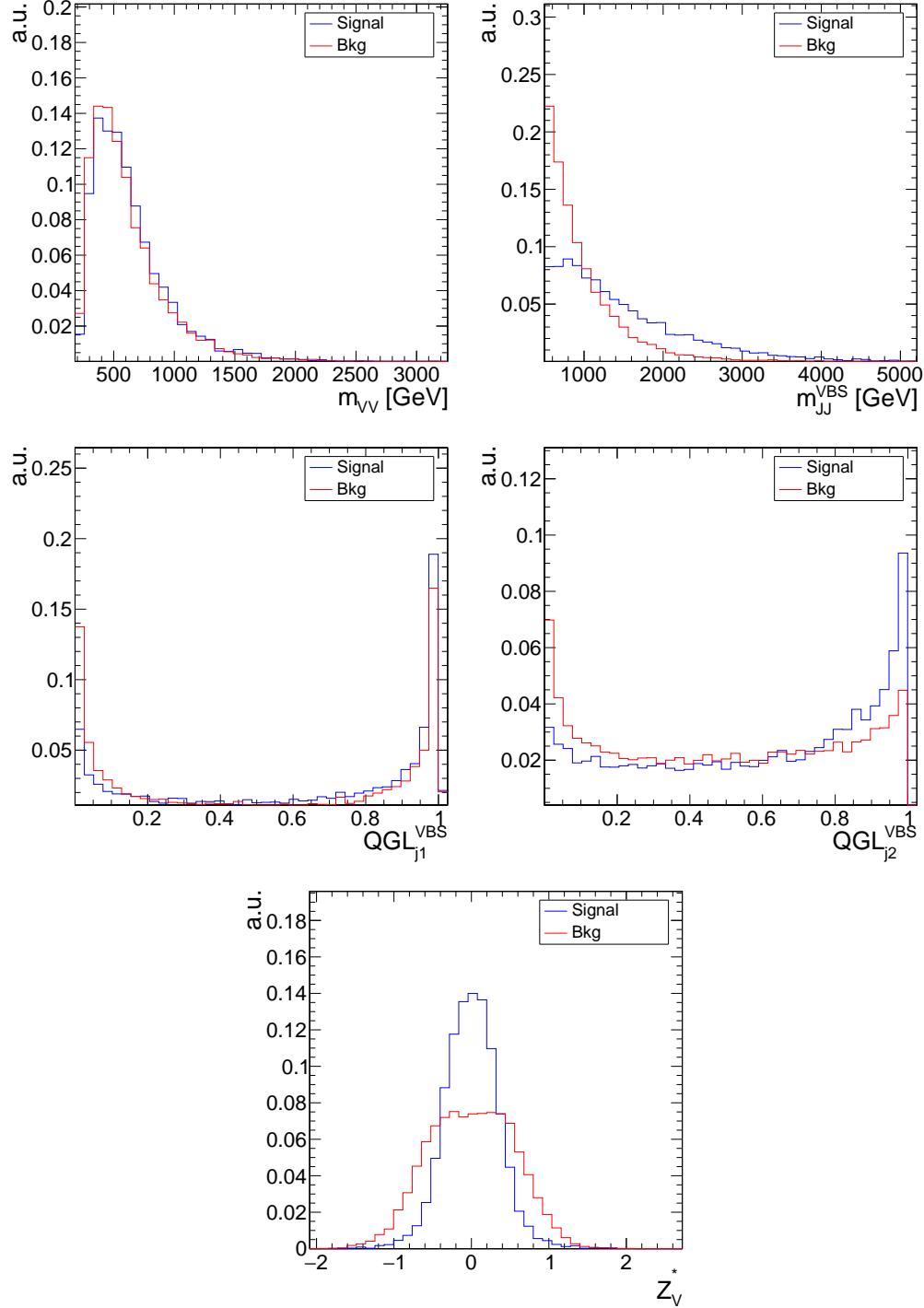


Figure 4.16: Inputs Variables (combined for Run2 MC) for training Boosted ZV BDT Classifier. From top: Diboson invariant mass, VBS tagged jets invariant mass, Quark Gluon Likelihood (QGL) of leading VBS tagged jet, QGL of trailing VBS tagged jet, Zeppenfeld variable of leptonic boson.

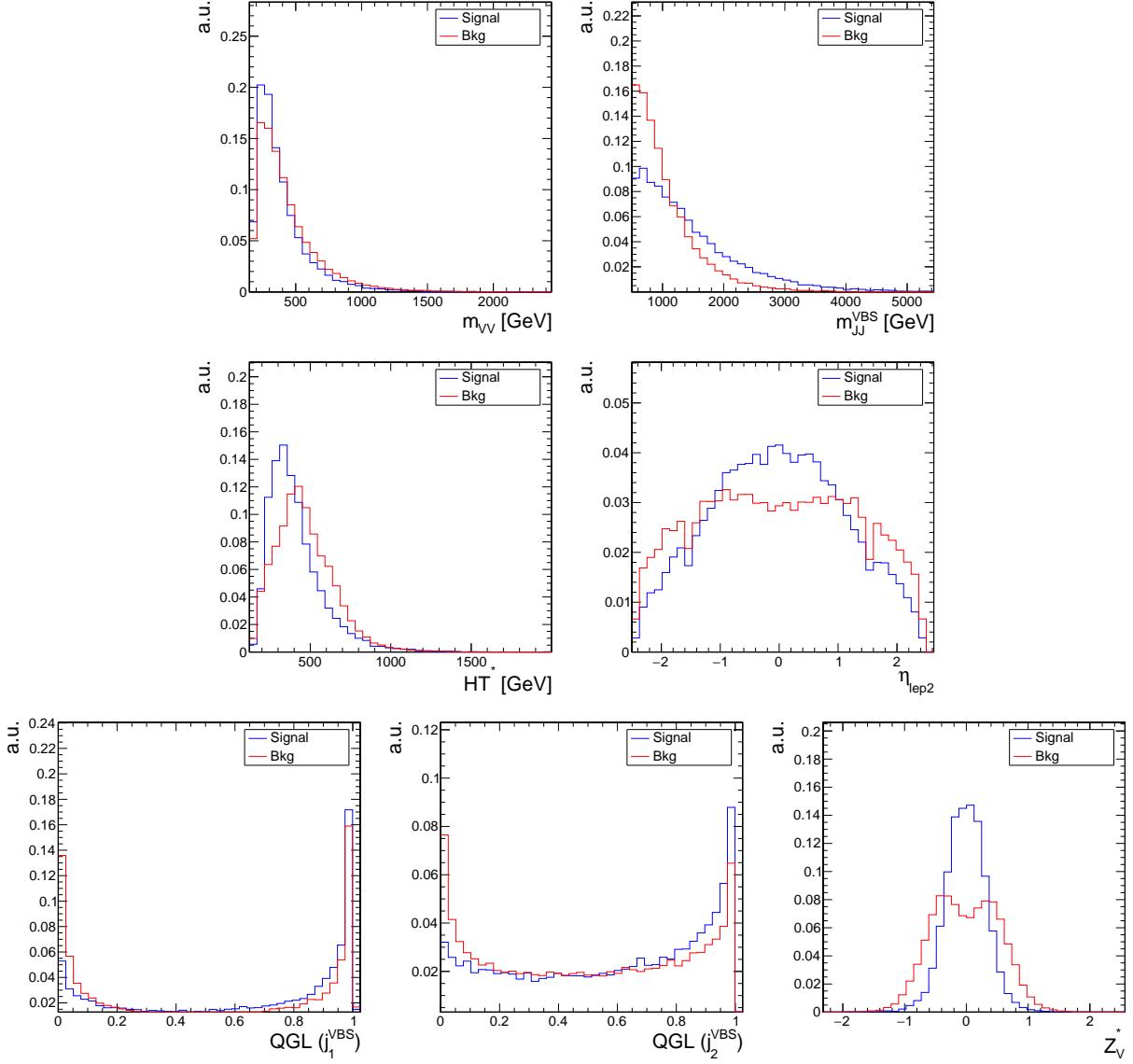


Figure 4.17: Inputs Variables (combined for Run2 MC) for training Resolved ZV BDT Classifier. From top: Diboson invariant mass, VBS tagged jets invariant mass, HT^* (p_T sum of jets), trailing lepton η , QGL of leading VBS tagged jet, QGL of trailing VBS tagged jet, Zeppenfeld variable of leptonic boson.

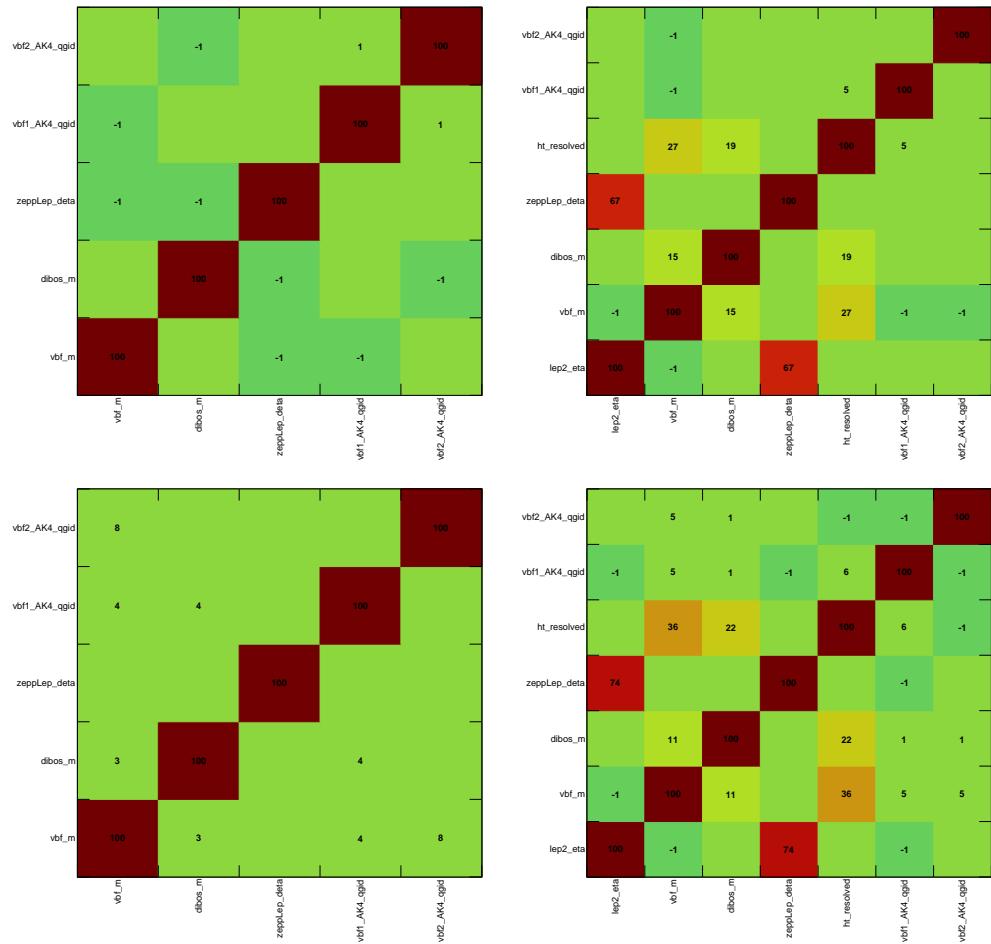


Figure 4.18: Correlation Matrix for Signal and Background. From Left to Right: Boosted, Resolved. From Top to Bottom: Signal, Background

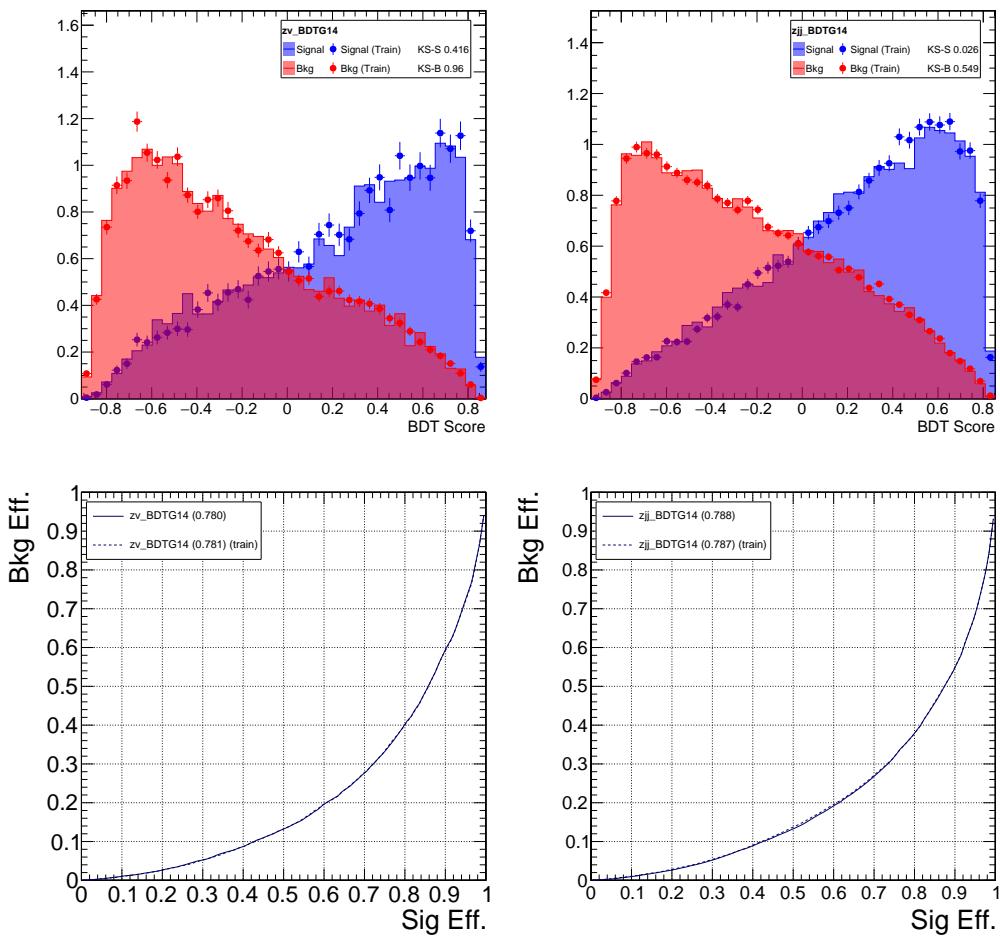


Figure 4.19: From Left to Right: Boosted, Resolved. Top to Bottom: MVA Score of BDT models, ROC Curves.

674

4.3.3 MVA Score inference in Signal Region

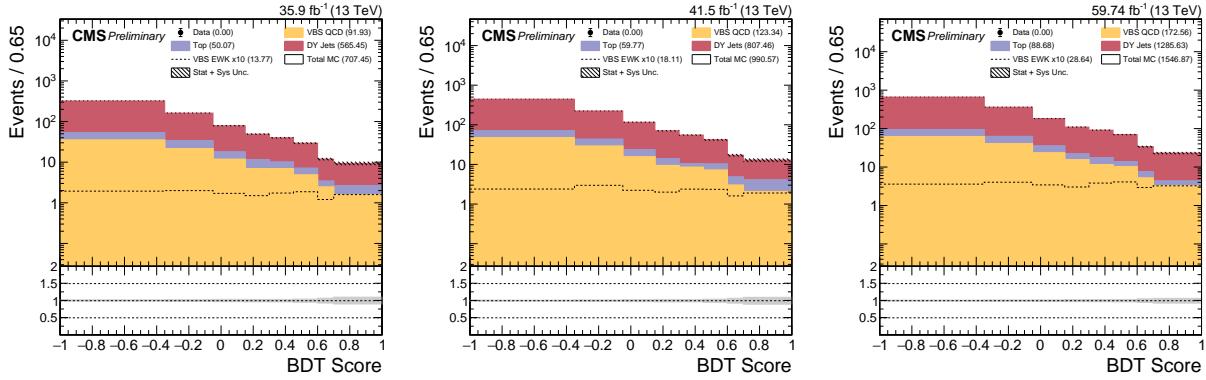


Figure 4.20: MVA Score in Signal Region for Boosted ZV Channel.

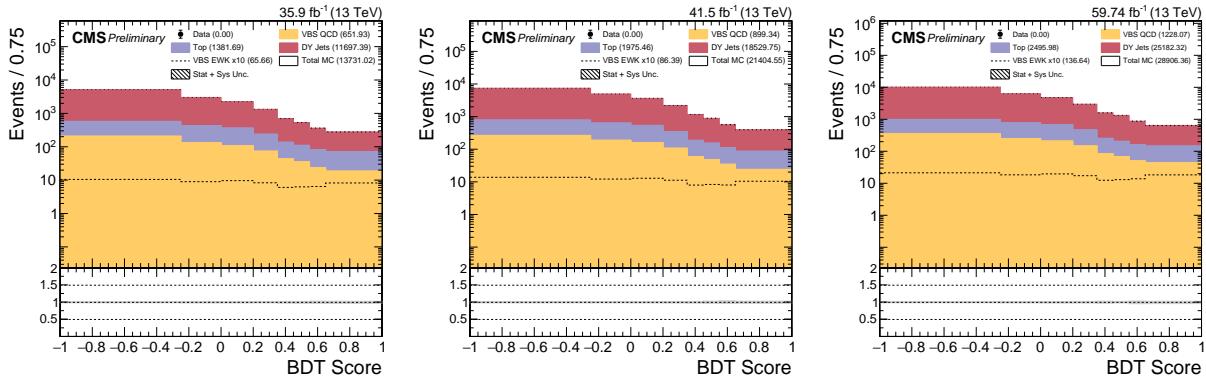


Figure 4.21: MVA Score in Signal Region for Resolved ZV Channel.

675

4.4 Measurement

676

About COMBINELIMIT used for calculating significance.

677

4.4.1 Statistical and Systematic Uncertainties

678

4.4.2 Significance

Table 4.8: Significance

Channel	2016	2017	2018	Combined
Boosted	0.66	0.68	0.88	1.29
Resolved	0.51	0.46	0.62	0.92
Combined	1.59			

4.4.3 Impact Plots

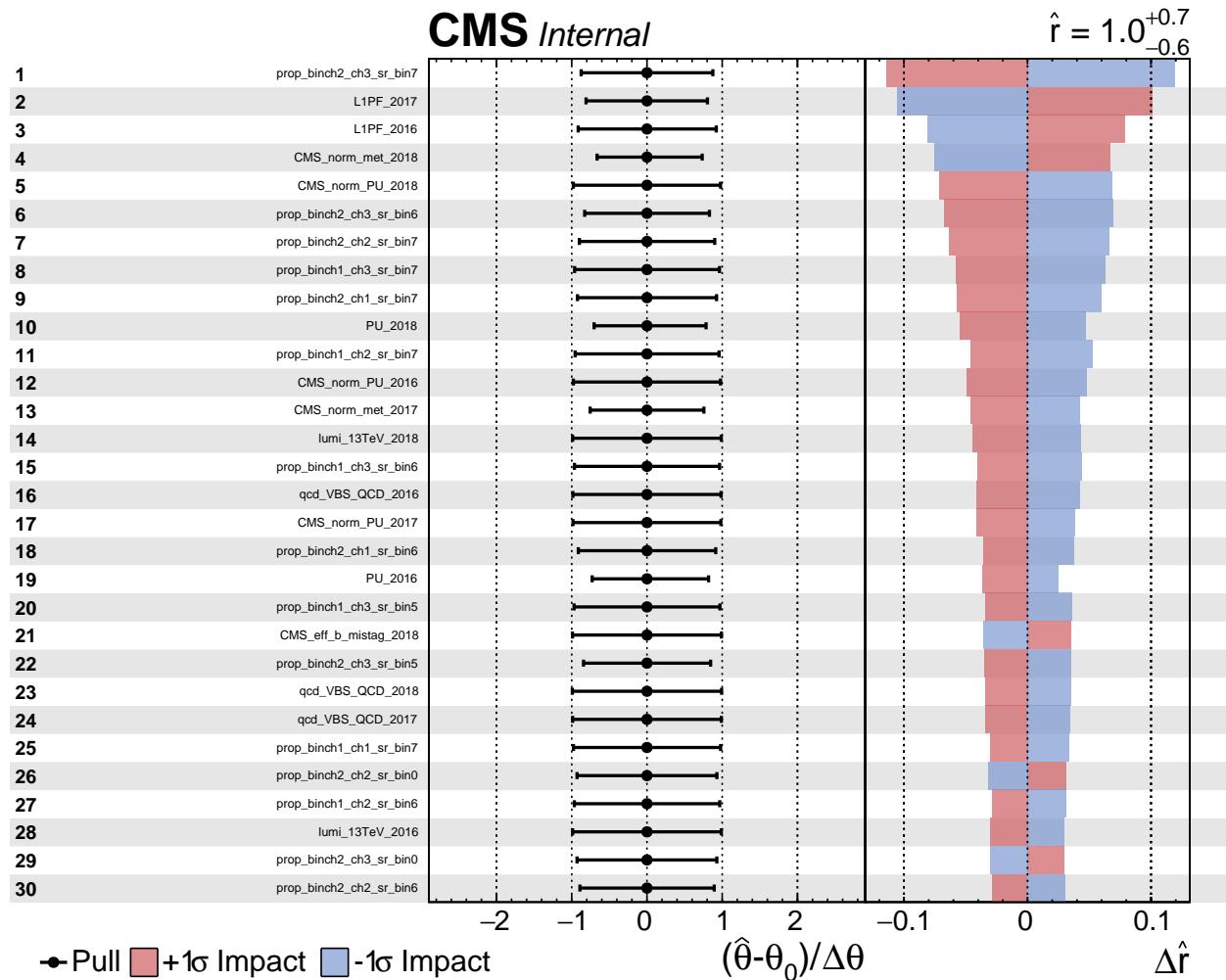


Figure 4.22: Impact Plots

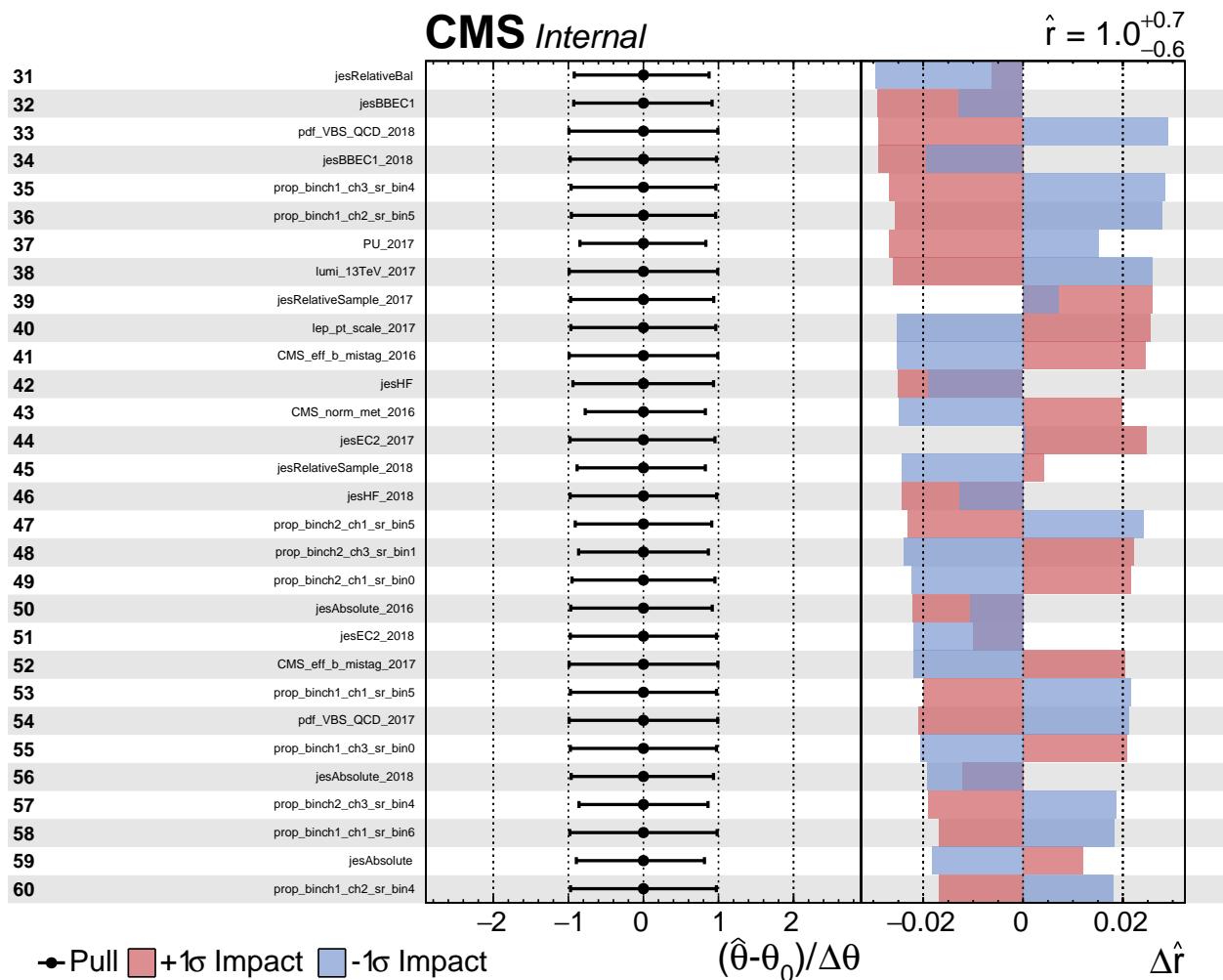


Figure 4.23: Impact Plots

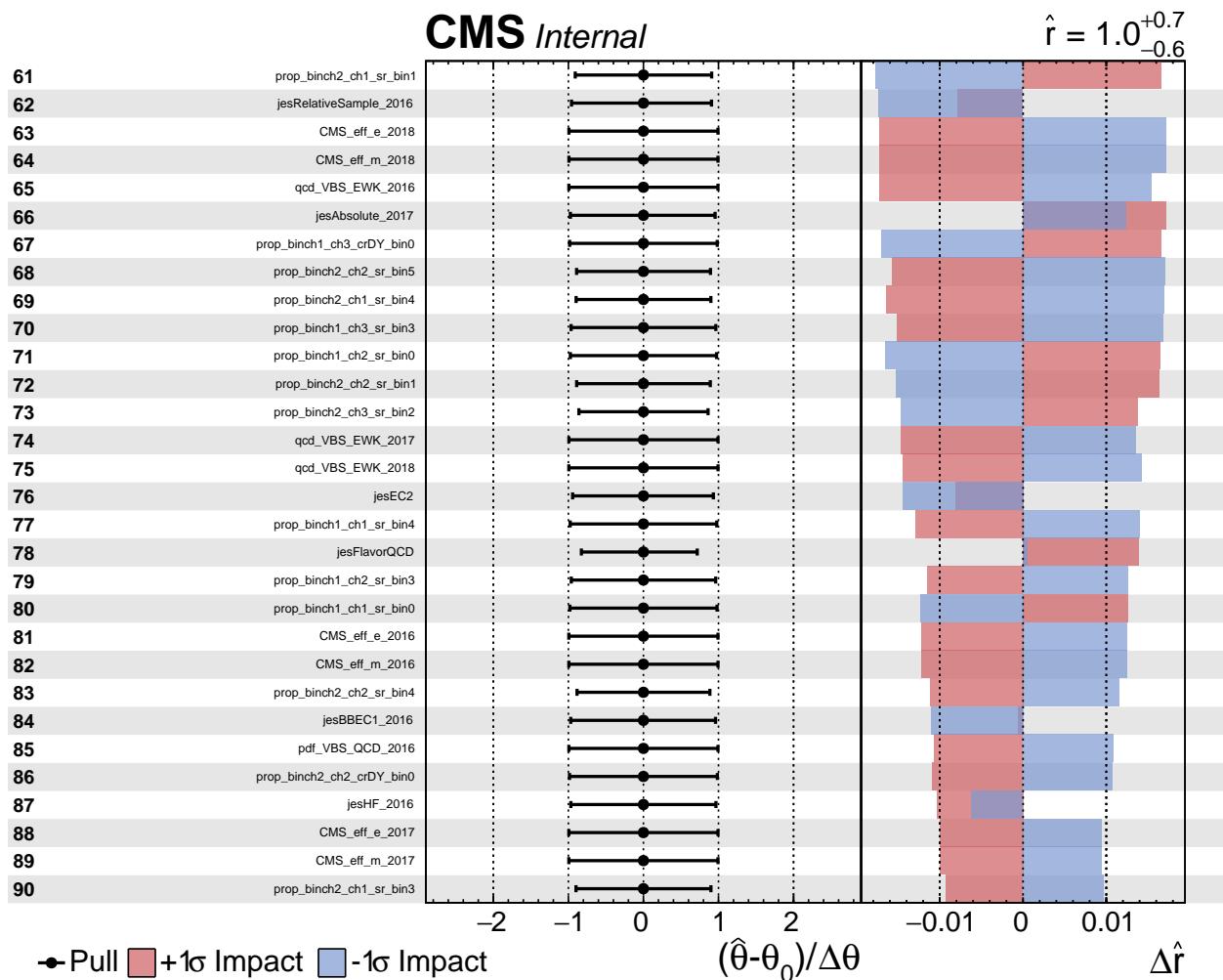


Figure 4.24: Impact Plots

4.4.4 Postfit Plots

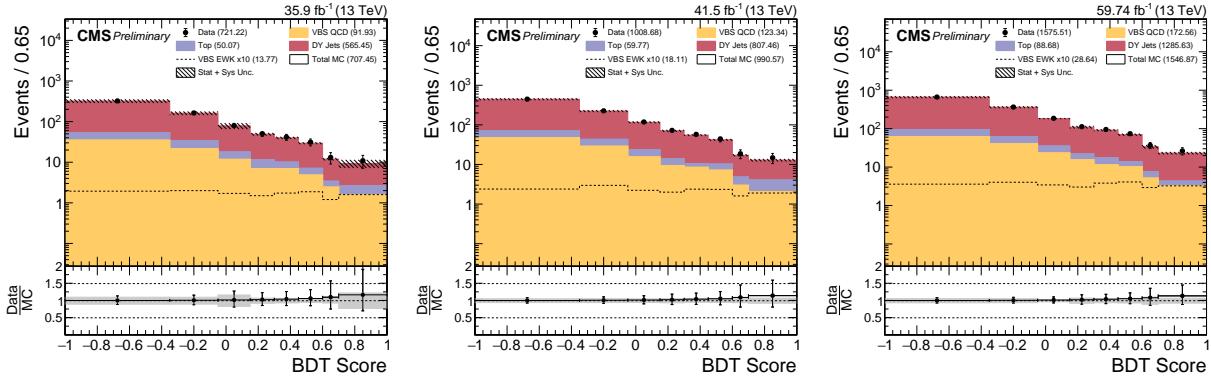


Figure 4.25: (Asimov Data) MVA Score postfit in Signal Region for Boosted ZV Channel.

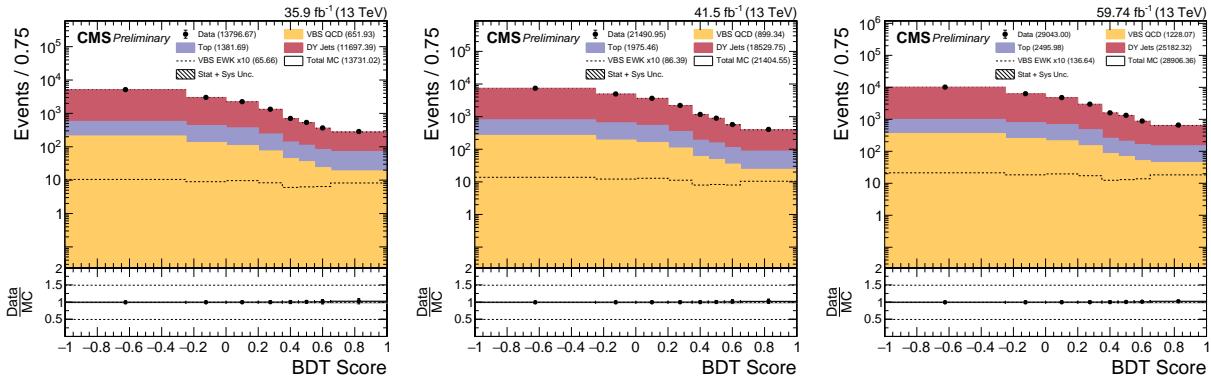


Figure 4.26: (Asimov Data) MVA Score postfit in Signal Region for Resolved ZV Channel.
(Asimov Data)

681

CHAPTER 5

682

HIGH GRANULARITY CALORIMETER UPGRADE

683

About High Luminosity LHC (HL-LHC) High Granularity Calorimeter (HGCAL) Up-

684 grade

685

5.1 Technical Design

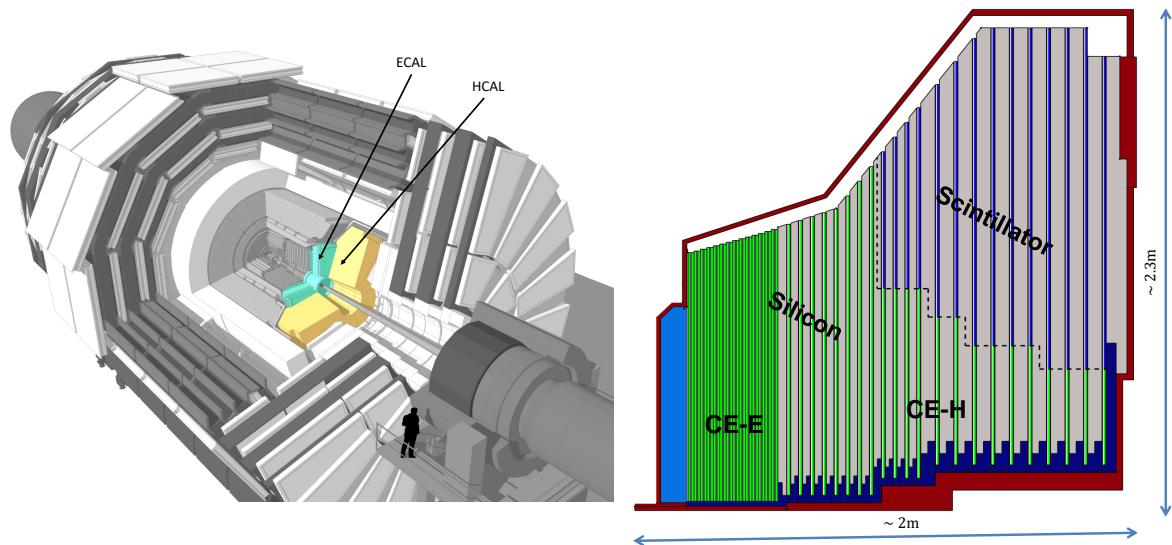


Figure 5.1: Overview of CE [41, 42]

686

5.2 Scintillator Tiles

687

5.2.1 End of Life Scenario

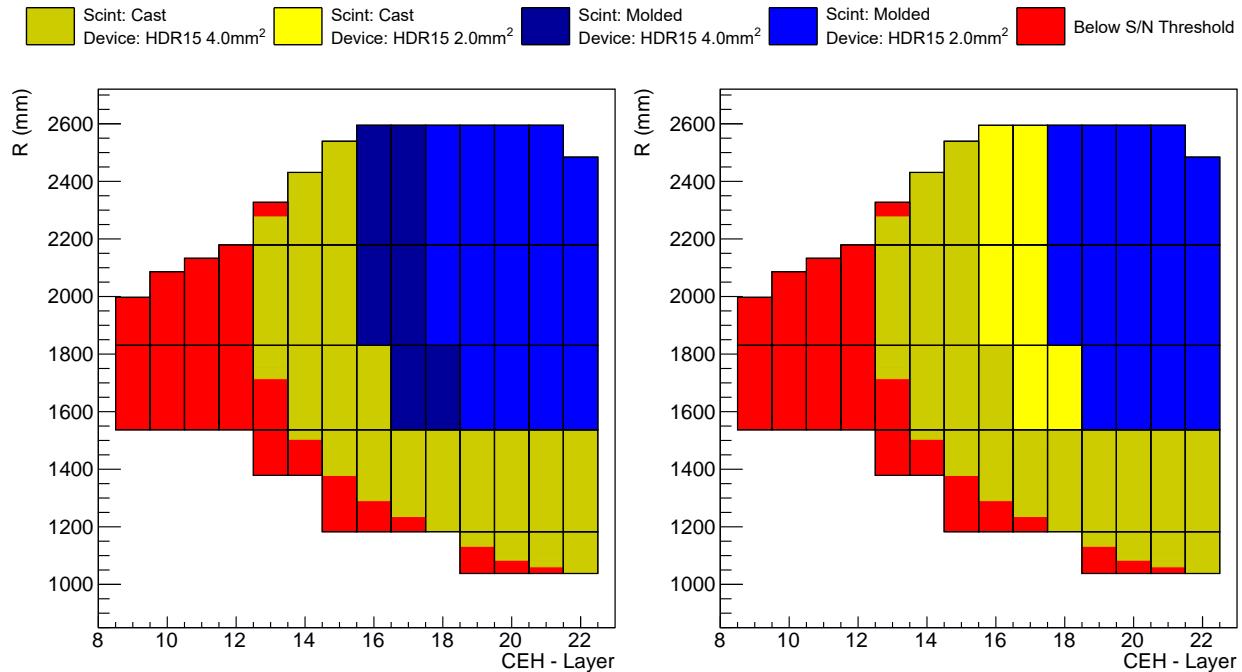


Figure 5.2: HGCAL scenarios

Table 5.1: HGCAL scenarios comparison

		Scene A	Scene B
Cast Scintillator	Cell Count	148, 608	176, 256
	Total Area	185.13 m ²	239.45 m ²
	Percentage	50.6 %	65.5 %
Injection Molded Scintillator	Cell Count	91, 008	63, 360
	Total Area	180.5 m ²	126.18 m ²
	Percentage	49.4 %	34.5 %
SiPMs Count	2 mm ²	63, 360	91, 008
	4 mm ²	100, 224	148, 608

688

CHAPTER 6

689

CONCLUSIONS

REFERENCES

- [1] Wikipedia contributors. “Standard Model of Elementary Particles” (July 2021). Accessed September 07, 2021. URL: https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.
- [2] C. N. Yang and R. L. Mills. “Conservation of Isotopic Spin and Isotopic Gauge Invariance”. *Phys. Rev.* 96 (1 Oct. 1954), pp. 191–195. DOI: [10.1103/PhysRev.96.191](https://doi.org/10.1103/PhysRev.96.191). URL: <https://link.aps.org/doi/10.1103/PhysRev.96.191>.
- [3] S. L. Glashow. “The renormalizability of vector meson interactions”. *Nuclear Physics* 10 (Feb. 1959), pp. 107–117. DOI: [10.1016/0029-5582\(59\)90196-8](https://doi.org/10.1016/0029-5582(59)90196-8).
- [4] S. Weinberg. “A Model of Leptons”. *Physical Review Letters* 19 (Nov. 1967), pp. 1264–1266. DOI: [10.1103/physrevlett.19.1264](https://doi.org/10.1103/physrevlett.19.1264).
- [5] A. Salam and J. C. Ward. “Weak and electromagnetic interactions”. *Il Nuovo Cimento* 11 (Feb. 1959), pp. 568–577. DOI: [10.1007/bf02726525](https://doi.org/10.1007/bf02726525).
- [6] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. *Physical Review Letters* 13 (Aug. 1964), pp. 321–323. DOI: [10.1103/physrevlett.13.321](https://doi.org/10.1103/physrevlett.13.321).
- [7] P. W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. *Physical Review Letters* 13 (Oct. 1964), pp. 508–509. DOI: [10.1103/physrevlett.13.508](https://doi.org/10.1103/physrevlett.13.508).
- [8] J. Ellis. “Higgs Physics”. en (2015). DOI: [10.5170/CERN-2015-004.117](https://doi.org/10.5170/CERN-2015-004.117).
- [9] B. W. Lee, C. Quigg, and H. B. Thacker. “Weak interactions at very high energies: The role of the Higgs-boson mass”. *Physical Review D* 16 (Sept. 1977), pp. 1519–1531. DOI: [10.1103/physrevd.16.1519](https://doi.org/10.1103/physrevd.16.1519).

- [712] [10] B. W. Lee, C. Quigg, and H. B. Thacker. “Strength of Weak Interactions at Very
 [713] High Energies and the Higgs Boson Mass”. *Physical Review Letters* 38 (Apr. 1977),
 [714] pp. 883–885. DOI: 10.1103/physrevlett.38.883.
- [715] [11] A. Denner and T. Hahn. “Radiative Corrections to $W^+W^- \rightarrow W^+W^-$ in the Elec-
 [716] troweak Standard Model”. *Nucl.Phys.B* 525:27-50, 1998 (Nov. 12, 1997). DOI: 10.1016/
 [717] S0550-3213(98)00287-9.
- [718] [12] L. Evans and P. Bryant. “LHC Machine”. *Journal of Instrumentation* 3.08 (Aug. 2008),
 [719] S08001–S08001. DOI: 10.1088/1748-0221/3/08/s08001.
- [720] [13] E. Mobs. “The CERN accelerator complex - 2019” (July 2019). Accessed August 16,
 [721] 2021. URL: <http://cds.cern.ch/record/2684277/files/>.
- [722] [14] Public CMS Luminosity Information. “Cumulative delivered and recorded luminosity
 [723] versus time for run-2 pp data” (2018). Accessed August 17, 2021. URL: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- [725] [15] The CMS Collaboration. “CMS Luminosity Measurements for the 2016 Data Taking
 [726] Period” (2017). URL: <https://cds.cern.ch/record/2257069>.
- [727] [16] The CMS Collaboration. “CMS luminosity measurement for the 2017 data-taking pe-
 [728] riod at $\sqrt{s} = 13$ TeV” (2018). URL: <https://cds.cern.ch/record/2621960>.
- [729] [17] The CMS Collaboration. “CMS luminosity measurement for the 2018 data-taking pe-
 [730] riod at $\sqrt{s} = 13$ TeV” (2019). URL: <https://cds.cern.ch/record/2676164>.
- [731] [18] T. Sakuma. “Cutaway diagrams of CMS detector” (May 2019). Accessed August 17,
 [732] 2021. URL: <https://cds.cern.ch/record/2665537>.
- [733] [19] The CMS Collaboration. “The CMS experiment at the CERN LHC”. *Journal of Instru-
 [734] mentation* 3.08 (Aug. 2008), S08004–S08004. DOI: 10.1088/1748-0221/3/08/s08004.
 [735] URL: <https://doi.org/10.1088/1748-0221/3/08/s08004>.

- 736 [20] The CMS Collaboration. “CMS Detector Slice” (Jan. 2016). Accessed August 18, 2021.
737 URL: <https://cds.cern.ch/record/2120661>.
- 738 [21] M. Brice and C. Marcelloni. “The central part of CMS is lowered” (Feb. 2007). Accessed
739 August 22, 2021. URL: <https://cds.cern.ch/record/1020311>.
- 740 [22] T. Sakuma. “SketchUp Drawings of the Phase 1 pixel detector alongside the current
741 pixel detector” (June 2012). Accessed August 17, 2021. URL: <https://cms-docdb.cern.ch/cgi-bin/PublicDocDB>ShowDocument?docid=6473>.
- 742
- 743 [23] The CMS Collaboration. “The CMS ECAL performance with examples”. *Journal of
744 Instrumentation* 9.02 (Feb. 2014), pp. C02008–C02008. DOI: [10.1088/1748-0221/9/
745 02/c02008](https://doi.org/10.1088/1748-0221/9/02/c02008).
- 746 [24] The CMS Collaboration. “Images of CMS ECAL Endcap (EE)” (Nov. 2008). Accessed
747 August 22, 2021. URL: <https://cds.cern.ch/record/1431479>.
- 748 [25] J. Mans et al. “CMS Technical Design Report for the Phase 1 Upgrade of the Hadron
749 Calorimeter” (Sept. 2012). URL: <https://cds.cern.ch/record/1481837>.
- 750 [26] The CMS Collaboration. “Images of the CMS HCAL Barrel (HB)” (2008). Accessed
751 August 22, 2021. URL: <https://cds.cern.ch/record/1431485>.
- 752 [27] The CMS Collaboration. “HCAL Depth Segmentation Phase 1 Upgrade” (2012). Ac-
753 cessed August 22, 2021. URL: <https://home.fnal.gov/~chlebana/CMS/Phase1>.
- 754 [28] The CMS Collaboration. “Performance of the CMS Muon Detectors in early 2016
755 collision runs” (2016). Accessed August 22, 2021. URL: [https://twiki.cern.ch/
756 twiki/pub/CMSPublic/MuonDPGPublic160729](https://twiki.cern.ch/twiki/pub/CMSPublic/MuonDPGPublic160729).
- 757 [29] The CMS Collaboration. “Particle-flow reconstruction and global event description
758 with the CMS detector”. *Journal of Instrumentation* 12.10 (Oct. 2017), P10003–
759 P10003. DOI: [10.1088/1748-0221/12/10/p10003](https://doi.org/10.1088/1748-0221/12/10/p10003).

- [30] S. Cucciarelli et al. “Track reconstruction, primary vertex finding and seed generation with the Pixel Detector”. CMS-NOTE-2006-026 (Jan. 2006). URL: <https://cds.cern.ch/record/927384>.
- [31] W. Adam et al. “Track Reconstruction in the CMS tracker”. CMS-NOTE-2006-041 (Dec. 2006). URL: <https://cds.cern.ch/record/934067>.
- [32] W. Adam et al. “Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC”. *Journal of Physics G: Nuclear and Particle Physics* 31.9 (July 2005), N9–N20. DOI: [10.1088/0954-3899/31/9/n01](https://doi.org/10.1088/0954-3899/31/9/n01).
- The CMS Collaboration. “Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC”. *JINST* 16 (2021) P05014 (Dec. 12, 2020). DOI: [10.1088/1748-0221/16/05/P05014](https://doi.org/10.1088/1748-0221/16/05/P05014).
- M. Cacciari, G. P. Salam, and G. Soyez. “FastJet user manual” (Nov. 25, 2011). DOI: [10.1140/epjc/s10052-012-1896-2](https://doi.org/10.1140/epjc/s10052-012-1896-2).
- D. Bertolini et al. “Pileup Per Particle Identification”. *JHEP* 1410 (2014) 59 (July 22, 2014). DOI: [10.1007/JHEP10\(2014\)059](https://doi.org/10.1007/JHEP10(2014)059).
- The CMS Collaboration. “Pileup mitigation at CMS in 13 TeV data”. *Journal of Instrumentation* 15.09 (Sept. 2020), P09018–P09018. DOI: [10.1088/1748-0221/15/09/p09018](https://doi.org/10.1088/1748-0221/15/09/p09018).
- J. Thaler and K. V. Tilburg. “Identifying Boosted Objects with N-subjettiness”. *JHEP* 1103:015, 2011 (Nov. 10, 2010). DOI: [10.1007/JHEP03\(2011\)015](https://doi.org/10.1007/JHEP03(2011)015).
- The CMS Collaboration. “Identification of heavy, energetic, hadronically decaying particles using machine-learning techniques”. *JINST* 15 (2020) P06005 (Apr. 17, 2020). DOI: [10.1088/1748-0221/15/06/P06005](https://doi.org/10.1088/1748-0221/15/06/P06005).

- 783 [39] A. J. Larkoski et al. “Soft Drop”. *JHEP* **1405** (2014) 146 (Feb. 11, 2014). DOI: 10.
784 1007/JHEP05(2014)146.
- 785 [40] CMS Collaboration. “Performance of the CMS muon detector and muon reconstruction
786 with proton-proton collisions at $\sqrt{s} = 13$ TeV”. *JINST* **13** (2018) P06015 (Apr. 12,
787 2018). DOI: 10.1088/1748-0221/13/06/P06015. arXiv: 1804.04528.
- 788 [41] D. Barney. “Overview slide of CE with main parameters” (Oct. 2019). Accessed August
789 24, 2021. URL: <https://cms-docdb.cern.ch/cgi-bin/PublicDocDB>ShowDocument?docid=13251>.
- 790 [42] D. Barney and T. Sakuma. “Sketchup images highlighting the sub-detectors” (Sept.
791 2017). Accessed August 24, 2021. URL: <https://cds.cern.ch/record/2628519>.

793

APPENDIX A

794

MACHINE LEARNING

795 Write about ML, MVA, BDT, etc.

796

APPENDIX B

797

DATA ANALYSIS

798 Data Analysis Code and Stuff