

OBSERVATION OF HIGGS BOSON DECAY TO BOTTOM QUARKS  
WHEN PRODUCED IN ASSOCIATION WITH A WEAK VECTOR BOSON

By

SEAN-JIUN WANG

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2019

© 2019 Sean-Jiun Wang

For my parents, who abandoned the familiar comforts of Taiwan and came to the United States because they wanted the best for their future children.

## ACKNOWLEDGMENTS

I want to first and foremost thank Jacobo Konigsberg for giving me the opportunity to work and learn under his guidance. Because of this, I was able to contribute to a long-awaited and historic result at the cutting edge of the field. I could not have asked for a better advisor, and I will cherish his invaluable support and instruction. I also want to thank Ivan Furic and Darin Acosta for providing undergraduate research opportunities that inspired my continued work on the CMS experiment and awakened my interest in machine learning. As our research is made possible by the University of Florida and the U.S. Department of Energy, I thank them for their support.

I also want to recognize and thank the CMS collaboration and those involved in LHC operations, whose tireless work continues to produce datasets of epic size that are a joy to explore. I also thank the colleagues and office mates who enriched my time at Fermilab, specifically Souvik Das, Jia Fu Low, and Caterina Vernieri. Such a challenging analysis wouldn't have been a success without the shared vision and expertise of the  $VH(b\bar{b})$  analysis team, so I thank and congratulate my closest colleagues on the job well done. Special thanks go to Michele de Gruttola and Christopher Palmer for their mentorship. I thank the Cookmans and the Paulsons, who showed me tremendous hospitality during my travels.

I thank my family for their endless love and encouragement and for giving me the freedom to choose how I might make the world a better place. I also thank my wife, Gwynndolyn, who has shown amazing patience and understanding while I travelled around the world chasing particles. I am forever grateful for her love and sacrifice which has allowed me to pursue my research. And finally, I thank God, who after creating the whirling cosmos saw fit to fashion humanity with the means to not only perceive its beauty but also fathom its design.

## TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS . . . . .	4
LIST OF TABLES . . . . .	8
LIST OF FIGURES . . . . .	9
ABSTRACT . . . . .	12
CHAPTER	
1 INTRODUCTION . . . . .	13
1.1 The Standard Model . . . . .	13
1.1.1 Fermions . . . . .	14
1.1.2 Bosons . . . . .	15
1.1.3 The Higgs Mechanism . . . . .	16
1.1.3.1 Gauge boson masses . . . . .	18
1.1.3.2 Fermion masses . . . . .	20
1.2 The Higgs Boson . . . . .	23
1.2.1 Phenomenology . . . . .	23
1.2.1.1 Production modes . . . . .	24
1.2.1.2 Decay modes . . . . .	27
1.2.2 Discovery . . . . .	29
1.3 Searches for $VH(b\bar{b})$ . . . . .	29
1.3.1 Motivation for $VH(b\bar{b})$ . . . . .	30
1.3.2 Previous Results . . . . .	33
2 EXPERIMENTAL APPARATUS . . . . .	35
2.1 The Large Hadron Collider . . . . .	36
2.2 The Compact Muon Solenoid . . . . .	40
2.2.1 Detector Subsystems . . . . .	44
2.2.1.1 Silicon tracker . . . . .	44
2.2.1.2 Electromagnetic calorimeter . . . . .	46
2.2.1.3 Hadronic calorimeter . . . . .	48
2.2.1.4 Muon subsystems . . . . .	49
2.2.2 Data Acquisition and Triggers . . . . .	51
2.2.3 Computing Resources . . . . .	53
3 PHYSICS OBJECT RECONSTRUCTION . . . . .	54
3.1 Particle Flow Algorithm . . . . .	54
3.1.1 Detector-level Objects . . . . .	55
3.1.1.1 Tracks . . . . .	56
3.1.1.2 Calorimeter clusters . . . . .	58

3.1.1.3	Standalone muon tracks . . . . .	60
3.1.2	Particle Candidates . . . . .	61
3.1.2.1	Electrons . . . . .	61
3.1.2.2	Photons . . . . .	62
3.1.2.3	Muons . . . . .	63
3.1.2.4	Charged and neutral hadrons . . . . .	64
3.2	Primary Vertices . . . . .	65
3.3	Tau Leptons . . . . .	66
3.4	Lepton Isolation . . . . .	66
3.5	Jets . . . . .	67
3.5.1	Reconstruction . . . . .	67
3.5.2	Energy Scale and Resolution Corrections . . . . .	70
3.5.3	<i>b</i> -Tagging . . . . .	72
3.6	Missing Transverse Energy . . . . .	75
4	ANALYSIS STRATEGY . . . . .	78
4.1	Data and Simulation . . . . .	78
4.1.1	Data . . . . .	78
4.1.2	Simulation . . . . .	79
4.1.3	Triggers . . . . .	84
4.1.4	Residual Monte-Carlo Corrections . . . . .	88
4.2	Physics Objects . . . . .	90
4.2.1	Vector Boson Candidate . . . . .	95
4.2.2	Higgs Boson Candidate . . . . .	96
4.3	Event Selection . . . . .	99
4.3.1	Signal Regions . . . . .	100
4.3.2	Background Control Regions . . . . .	102
4.4	Binned Multivariate Shape Analysis . . . . .	104
4.4.1	Multivariate Discriminator . . . . .	106
4.4.2	Binned Shape Analysis . . . . .	110
4.4.3	$VZ(b\bar{b})$ Cross-check Analysis . . . . .	113
4.4.4	Dijet Invariant Mass Cross-check Analysis . . . . .	114
4.5	Systematic Uncertainties . . . . .	116
5	RESULTS . . . . .	119
5.1	Statistical Treatment . . . . .	119
5.2	$VZ(b\bar{b})$ Analysis . . . . .	121
5.3	Dijet Invariant Mass Analysis . . . . .	121
5.4	$VH(b\bar{b})$ Analysis . . . . .	124
5.5	Combinations with Past CMS Results . . . . .	128
6	CONCLUSIONS . . . . .	130
	APPENDIX	

A	EVENT DISPLAYS . . . . .	132
B	CONTROL REGION DISTRIBUTIONS . . . . .	138
REFERENCES . . . . .		148
BIOGRAPHICAL SKETCH . . . . .		158

## LIST OF TABLES

<u>Table</u>	<u>page</u>
1-1 Production Cross Sections for the Higgs Boson and SM Background Process at the LHC . . . . .	34
3-1 Iterative Tracking Summary . . . . .	58
4-1 Signal Samples for $VH(b\bar{b})$ 2017 . . . . .	80
4-2 Diboson Samples for $VH(b\bar{b})$ 2017 . . . . .	80
4-3 $W+$ jets ( $W \rightarrow \ell\bar{\nu}$ ) Samples for $VH(b\bar{b})$ 2017 . . . . .	81
4-4 $Z+$ jets ( $Z \rightarrow \ell\bar{\ell}$ ) Samples for $VH(b\bar{b})$ 2017 . . . . .	82
4-5 $Z+$ jets ( $Z \rightarrow \nu\bar{\nu}$ ) Samples for $VH(b\bar{b})$ 2017 . . . . .	83
4-6 $t\bar{t}$ Samples for $VH(b\bar{b})$ 2017 . . . . .	83
4-7 Single Top Samples for $VH(b\bar{b})$ 2017 . . . . .	84
4-8 QCD Samples for $VH(b\bar{b})$ 2017 . . . . .	84
4-9 L1 and HLT Triggers for $VH(b\bar{b})$ 2017 . . . . .	85
4-10 2017 DeepCSV Working Points . . . . .	94
4-11 Signal Region Definitions for the 2017 $VH(b\bar{b})$ Analysis . . . . .	101
4-12 $Z(\nu\bar{\nu})H$ Control Region Definitions for the 2017 $VH(b\bar{b})$ Analysis . . . . .	104
4-13 $W(\ell\nu)H$ Control Region Definitions for the 2017 $VH(b\bar{b})$ Analysis . . . . .	105
4-14 $Z(\ell\bar{\ell})H$ Control Region Definitions for the 2017 $VH(b\bar{b})$ Analysis . . . . .	105
4-15 Event Classification Variables . . . . .	107
4-16 Mass-correlated Input Features . . . . .	115
5-1 $VZ(b\bar{b})$ Analysis Scale Factors . . . . .	121
5-2 Dijet Invariant Mass Analysis Categories . . . . .	124
5-3 Dijet Invariant Mass Analysis Scale Factors . . . . .	124
5-4 $VH(b\bar{b})$ Analysis Scale Factors . . . . .	125
5-5 $VH(b\bar{b})$ Main Uncertainty Sources . . . . .	128
A-1 Event Display Candidate Information . . . . .	132

## LIST OF FIGURES

<u>Figure</u>		<u>page</u>
1-1 Standard Model Infographic . . . . .		14
1-2 Shape of the Higgs Field Potential . . . . .		17
1-3 Higgs Boson Interaction Vertices . . . . .		24
1-4 Higgs Boson Production Cross Sections . . . . .		25
1-5 Higgs Boson Production Diagrams . . . . .		26
1-6 Higgs Boson Branching Ratios . . . . .		28
1-7 Production Cross Sections at the LHC . . . . .		31
1-8 CMS Measurements of SM Production Cross Sections . . . . .		32
1-9 $VH(b\bar{b})$ Background Process Diagrams . . . . .		32
2-1 Overall View of the LHC . . . . .		36
2-2 CERN Accelerator Complex . . . . .		38
2-3 LHC Dipole and Quadrupole Magnets . . . . .		38
2-4 Cutaway View of the CMS Detector . . . . .		41
2-5 Sagitta of a Charged Particle Track . . . . .		42
2-6 Coordinate System of the CMS Detector . . . . .		44
2-7 Cross Sectional View of the CMS Detector . . . . .		45
2-8 Schematic for the CMS Silicon Tracker . . . . .		45
2-9 Schematic for the CMS ECAL . . . . .		47
2-10 Schematic for the CMS HCAL . . . . .		48
2-11 Schematic for the CMS Muon Subsystems . . . . .		49
2-12 Overview of the Level 1 Trigger System . . . . .		52
3-1 Particle Flow Infographic . . . . .		55
3-2 Event Display of High-Pileup Collision Event . . . . .		56
3-3 Charged Particle Trajectory Parameterization . . . . .		57
3-4 Track Reconstruction Performance . . . . .		59

3-5	Event Display of Calorimeter Clusters . . . . .	60
3-6	Electron Reconstruction Efficiency . . . . .	62
3-7	Photon Reconstruction Efficiency . . . . .	63
3-8	Muon Reconstruction Efficiency . . . . .	65
3-9	Infrared and Collinear Unsafety . . . . .	68
3-10	Jet Energy Scale Corrections Infographic . . . . .	70
3-11	Illustration of a Heavy-Flavor Jet . . . . .	72
3-12	CMS $b$ -Tagging Performance . . . . .	74
3-13	Illustration of Missing Transverse Momentum . . . . .	76
4-1	2017 Signal MC Rescaling Weights for $W^+(\ell\nu)H \rightarrow b\bar{b}$ . . . . .	80
4-2	Trigger Efficiency for $Z(\nu\bar{\nu})H(b\bar{b})$ in 2017 Data . . . . .	87
4-3	Trigger Efficiency for $Z(\nu\bar{\nu})H(b\bar{b})$ in 2017 MC . . . . .	87
4-4	2017 MET Trigger Efficiency Correction for $Z(\nu\bar{\nu})H(b\bar{b})$ . . . . .	88
4-5	2017 $V+jets$ MC Electroweak Correction . . . . .	89
4-6	NLO to LO Ratio of $\Delta\eta(jj)$ for $Z+jets$ ( $Z \rightarrow \ell\bar{\ell}$ ) Samples . . . . .	90
4-7	Average Jet Energy Density in 2017 Data . . . . .	92
4-8	Performance of $b$ -jet Energy Regression . . . . .	98
4-9	Performance of $Z(\ell\bar{\ell})H$ Kinematic Fit . . . . .	98
4-10	$VH(b\bar{b})$ Control Region Definitions . . . . .	103
4-11	Input Feature Distributions for the $Z(\nu\bar{\nu})H$ Channel . . . . .	108
4-12	Normalized DNN Discriminant Distributions . . . . .	111
4-13	Multi-background DNN Discriminant Distributions . . . . .	112
4-14	Background Sculpting Mitigation . . . . .	116
5-1	$VZ(b\bar{b})$ Signal Region Distributions . . . . .	122
5-2	$VZ(b\bar{b})$ Combined Signal Region Distribution . . . . .	123
5-3	$VZ(b\bar{b})$ Analysis Signal Strengths . . . . .	123
5-4	Combined Dijet Invariant Mass Distribution . . . . .	125

5-5	$VH(b\bar{b})$ Signal Region Distributions . . . . .	126
5-6	$VH(b\bar{b})$ Combined Signal Region Distribution . . . . .	127
5-7	$VH(b\bar{b})$ Analysis Signal Strengths . . . . .	127
5-8	$H(b\bar{b})$ Combination Signal Strengths . . . . .	129
A-1	Event Display for $Z(\nu\bar{\nu})H(b\bar{b})$ Candidate . . . . .	133
A-2	Event Display for $W(e\nu)H(b\bar{b})$ Candidate . . . . .	134
A-3	Event Display for $W(\mu\nu)H(b\bar{b})$ Candidate . . . . .	135
A-4	Event Display for $Z(e\bar{e})H(b\bar{b})$ Candidate . . . . .	136
A-5	Event Display for $Z(\mu\bar{\mu})H(b\bar{b})$ Candidate . . . . .	137
B-1	$t\bar{t}$ Control Region Distributions for the $Z(\nu\bar{\nu})H$ Channel . . . . .	139
B-2	Additional $t\bar{t}$ Control Region Distributions for the $Z(\nu\bar{\nu})H$ Channel . . . . .	140
B-3	$Z+$ light Control Region Distributions for the $Z(\nu\bar{\nu})H$ Channel . . . . .	141
B-4	Additional $Z+$ light Control Region Distributions for the $Z(\nu\bar{\nu})H$ Channel . . .	142
B-5	$Z+$ heavy Control Region Distributions for the $Z(\nu\bar{\nu})H$ Channel . . . . .	143
B-6	Additional $Z+$ heavy Control Region Distributions for the $Z(\nu\bar{\nu})H$ Channel . .	144
B-7	Control Region Distributions for the $W(\ell\nu)H$ Channels . . . . .	145
B-8	Low $p_T(V)$ Control Region Distributions for the $Z(\ell\bar{\ell})H$ Channels . . . . .	146
B-9	High $p_T(V)$ Control Region Distributions for the $Z(\ell\bar{\ell})H$ Channels . . . . .	147

Abstract of Dissertation Presented to the Graduate School  
of the University of Florida in Partial Fulfillment of the  
Requirements for the Degree of Doctor of Philosophy

OBSERVATION OF HIGGS BOSON DECAY TO BOTTOM QUARKS  
WHEN PRODUCED IN ASSOCIATION WITH A WEAK VECTOR BOSON

By

Sean-Jiun Wang

May 2019

Chair: Jacobo Konigsberg

Cochair: Ivan K. Furic

Major: Physics

A search for the Standard Model Higgs boson ( $H$ ) produced in association with a weak vector boson ( $VH$ ) and decaying into a bottom-antibottom quark pair ( $b\bar{b}$ ) is reported for the decay channels  $Z(\nu\bar{\nu})H(b\bar{b})$ ,  $W(e\nu)H(b\bar{b})$ ,  $W(\mu\nu)H(b\bar{b})$ ,  $Z(e\bar{e})H(b\bar{b})$ , and  $Z(\mu\bar{\mu})H(b\bar{b})$ , with a focus on  $Z(\nu\bar{\nu})H(b\bar{b})$  in particular. The search is performed with a dataset corresponding to an integrated luminosity of  $41.3 \text{ fb}^{-1}$  at a center-of-mass energy of  $\sqrt{s} = 13 \text{ TeV}$  recorded by the CMS experiment at the LHC during Run 2 in 2017. An excess of events is observed above the expected background with a significance of 3.3 standard deviations, which is compatible with the Standard Model expectation of 3.1 standard deviations for a Higgs boson of mass  $m_H = 125.09 \text{ GeV}$ . When combined with the results of previous  $VH$  measurements and  $H \rightarrow b\bar{b}$  searches using other Higgs production modes, the observed (expected) significance is 5.6 (5.5) standard deviations. This represents the first observation of  $H \rightarrow b\bar{b}$  by the CMS experiment.

## CHAPTER 1

### INTRODUCTION

The fundamental structure of the universe has remained an enduring fascination of mankind. The philosophers of antiquity speculated that the world could be decomposed into the base elements of water, earth, fire, and air. The alchemists of the Middle Ages, regarding the classical elements to be expressions of hidden substance, sought to unveil its nature with crude experiments. The scholars of the Renaissance rejected the Aristotelian description of physical phenomena for its empirical failures and pondered if atoms were the indivisible units of matter. The chemists of the 19th century advanced and refined the atomic theory in their pursuit to catalog the pure elements. The physicists of the 20th century realized the necessity of a quantum description of the universe while delving into the subatomic realm. The scientific paradigm of the current era now rests on the theoretical framework known as the Standard Model of particle physics which, though understood to be incomplete, has had its predictions verified with remarkable accuracy.

#### 1.1 The Standard Model

The fundamental forces<sup>1</sup> and forms of matter within the universe are explained by the Standard Model in terms of elementary particles, their antiparticles, and their interactions. The following subsections highlight the features of the model necessary to motivate the treatise of this paper. A formal and rigorous treatment of the theory is beyond the scope of this paper but may be found in such texts as Ref. [1] and [2].

The Standard Model is a relativistic quantum field theory (QFT) in which fundamental particles are understood to be excitations of their associated fields that permeate all of space-time. The fundamental forces are generated by the internal gauge symmetry of the model  $SU(3) \times SU(2) \times U(1)$ , which gives rise to the strong ( $SU(3)$ ), weak ( $SU(2)$ ),

---

<sup>1</sup> Gravity, which does not have a satisfactory quantum description, falls outside of the Standard Model.

and electromagnetic ( $U(1)$ ) interactions. Its current formulation is thus a combination of electroweak theory with quantum chromodynamics (QCD). The particles manifested by the model may be broadly categorized into *fermions* and *bosons* and a visual summary is shown in Figure 1-1.

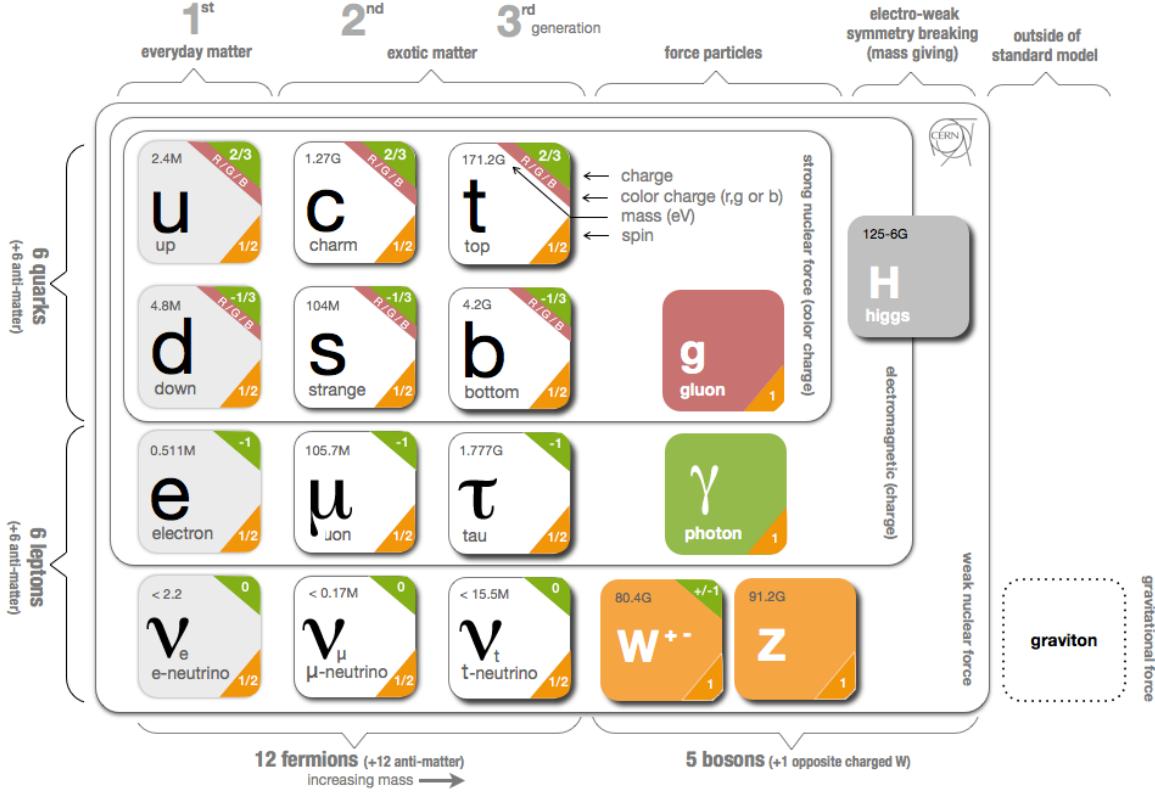


Figure 1-1. An infographic of the Standard Model [3]. The elementary particles are arranged in their usual generational pairs and the surrounding solid lines divide them into sections based on the fundamental forces they experience.

### 1.1.1 Fermions

The fermions encompass all forms of ordinary and exotic matter and obey Fermi-Dirac statistics because of their half-integer spin, namely, spin- $\frac{1}{2}$ . They are divided into *leptons* and *quarks*. The electrically charged leptons interact via the electromagnetic and weak nuclear forces and include the familiar *electron* ( $e$ ), the *muon* ( $\mu$ ), and the *tau* ( $\tau$ ). The charged leptons also have neutral counterparts ( $\nu_e$ ,  $\nu_\mu$ ,  $\nu_\tau$ ) known as *neutrinos*, which interact solely via the weak nuclear force. The quarks, having both electric and color charge, interact via the electromagnetic and strong and weak nuclear forces and include

the up ( $u$ ), down ( $d$ ), charm ( $c$ ), strange ( $s$ ), bottom ( $b$ ), and top ( $t$ ). Although the leptons can exist freely, the nature of the strong interaction gives rise to color confinement, whereby quarks only appear within composite particles called hadrons. The bound states of quark doublets are known as mesons while those of quark triplets, of which the familiar proton is an example, are known as baryons.

The leptons and quarks are each arranged into pairs by flavour quantum number and sorted into three generations of increasing mass<sup>2</sup>, as depicted in Figure 1-1. The more massive particles of the higher generations decay into the stable particles of the first generation, which corroborates the observation that ordinary matter consists of electrons and up and down quarks which are bound in protons and neutrons. There are no known constraints on the number of fermion generations, although experimental results suggest there are only three.

### 1.1.2 Bosons

The bosons, which obey Bose-Einstein statistics because of their integer spin, are divided into vector and scalar bosons. The spin-1 vector, or gauge, bosons mediate the fundamental forces and are exchanged between elementary particles during their interactions. The photon ( $\gamma$ ) mediates the electromagnetic force, which is responsible for the phenomenon of intermolecular repulsion.<sup>3</sup> The  $W$  and  $Z$  bosons mediate the weak nuclear force that causes nuclear decay. And finally, the gluons ( $g$ ) mediate the strong nuclear force that binds quarks into hadrons and even protons and neutrons in nuclei. The only spin-0 scalar boson in the theory is the Higgs boson ( $H$ ), which is responsible for the masses of the heavy gauge bosons and fermions.

---

<sup>2</sup> The observation of neutrino oscillations disproved the prediction that they are massless but its implications for the lepton generations are unclear.

<sup>3</sup> It is this repulsion that prevents us from passing through objects.

### 1.1.3 The Higgs Mechanism

The requirement of gauge invariance under SU(2) leads to a prediction of massless gauge bosons, which stands in contrast to the observation that the weak vector bosons are massive. In order to reconcile this observation while preserving the gauge invariance of the interaction, a mechanism of *spontaneous symmetry breaking* was incorporated into the unified electroweak theory proposed by Sheldon Glashow [4], Abdus Salam [5], and Steven Weinberg [6]. This mechanism, which came to be known as the *Higgs mechanism*, was independently developed by Robert Brout and François Englert [7], its namesake Peter Higgs [8], and Gerald Guralnik, Richard Hagen, and Sir Thomas Kibble [9].

The mechanism introduces into the Standard Model a scalar field  $\phi$  represented as a complex doublet

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (1-1)$$

and its potential  $V(\phi)$  of the form

$$V(\phi) = \mu^2 \phi^\dagger \phi + \frac{\lambda}{2} (\phi^\dagger \phi)^2, \quad (1-2)$$

where the real parameters  $\mu^2$  and  $\lambda$  represent the mass and self-coupling, respectively. The self-coupling  $\lambda$  is taken to be positive by convention such that the potential is bounded from below. When  $\mu^2 > 0$ , the potential is parabolic in shape and the ground state of the vacuum is at  $\phi = 0$ , keeping its symmetries intact. However, when  $\mu^2 < 0$ , the minimum of the potential is no longer at  $\phi = 0$  but rather<sup>4</sup>

$$v = \left( \frac{-\mu^2}{\lambda} \right)^{1/2}, \quad (1-3)$$

---

<sup>4</sup> The solution is unique up to a phase  $e^{i\theta}$ , but it is customary to take  $\theta = 0$ .

where the vacuum now attains an expectation value  $v$  in its ground state. The shape of such a potential is illustrated in Figure 1-2. The Higgs mechanism thus provides the means to spontaneously break the  $SU(2) \times U(1)$  symmetry within the model.

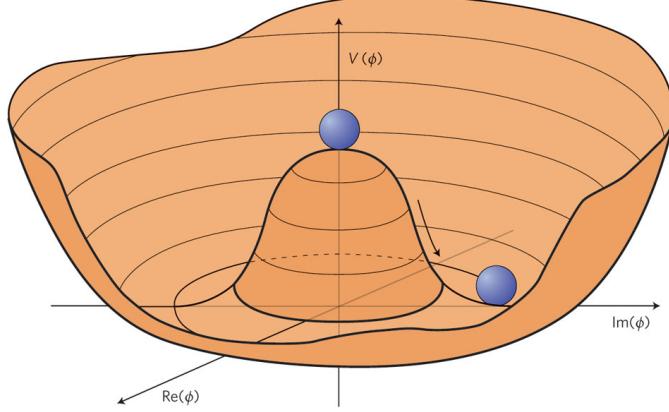


Figure 1-2. The shape of the Higgs field potential for  $\mu^2 < 0$  [10]. The transition of the field from its unstable state at the origin to its true ground state spontaneously breaks the rotational symmetry of the system.

Because particles are excitations of their fields, the presence of the Higgs field also suggests the existence of the Higgs boson. An expression for the mass of this new boson can be determined by expanding the potential around the vacuum ground state  $\phi_0$ . By choosing the unitary gauge, the components of  $\phi_0$  become

$$(\phi_1)_0 = 0, (\phi_2)_0 = 0, (\phi_3)_0 = v, (\phi_4)_0 = 0 \quad (1-4)$$

and  $\phi_0$  becomes simply

$$\phi_0 = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (1-5)$$

Then for small fluctuations  $\phi(x) = \phi_0 + H(x)$ , the potential to second order in  $H$  becomes

$$\begin{aligned} V(\phi) &= \frac{\mu^2}{2}(v^2 + 2vH + H^2) + \frac{\lambda}{4}(v^2 + 2vH + H^2)^2 \\ &= \frac{1}{2}(\mu^2 v^2 + \frac{\lambda}{2}v^4) + (\mu^2 v + \lambda v^3)H + \frac{1}{2}(\mu^2 + 3\lambda v^2)H^2 + \mathcal{O}(H^3) \\ &= V_0 - \frac{1}{2}(2\mu^2)H^2 + \mathcal{O}(H^3), \end{aligned} \quad (1-6)$$

where  $V_0$  collects the constant terms and Equation 1-3 has been used to vanish the first order term and simplify the second order term. The mass can be read from the coefficient of the second order term in  $H$ :

$$m_H^2 = -2\mu^2 \implies m_H = \sqrt{2}\mu = \sqrt{2\lambda}v, \quad (1-7)$$

where the sign of the mass was dropped because  $\mu^2$  was chosen to be negative. Although the Standard Model predicts a massive Higgs boson, its dependence on the value of the unknown self-coupling parameter  $\lambda$  and the vacuum expectation value  $v$  means it must be determined experimentally.

### 1.1.3.1 Gauge boson masses

The necessity of the Higgs mechanism is realized by how it generates the masses of the gauge bosons. The Lagrangian for the Higgs may be written as

$$\mathcal{L}_\phi = (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi) = (D_\mu\phi)^\dagger(D^\mu\phi) - \mu^2\phi^\dagger\phi - \frac{\lambda}{2}(\phi^\dagger\phi)^2, \quad (1-8)$$

By definition, the Higgs field has weak hypercharge  $Y = +1$  and weak isospin  $T_3 = -1/2$ . The weak hypercharge  $Y$  and weak isospin  $T$  are the generators of the SU(2) and U(1) symmetries, respectively, with the choice of their values made to satisfy the relation

$$Q = T_3 + \frac{1}{2}Y, \quad (1-9)$$

where  $Q$  is the electric charge.<sup>5</sup> These choices require the covariant derivative of the kinetic term to be of the form

$$D_\mu\phi = (\partial_\mu - \frac{i}{2}g\boldsymbol{\sigma} \cdot \mathbf{W}_\mu - \frac{i}{2}g'B_\mu)\phi, \quad (1-10)$$

where  $\mathbf{W}_\mu$  and  $B_\mu$  are, respectively, the SU(2) and U(1) gauge bosons,  $g$  and  $g'$  are the corresponding gauge couplings, and  $\boldsymbol{\sigma} = \sigma^a, a = 1, 2, 3$  are the Pauli matrices.

---

<sup>5</sup> Hence, the Higgs boson is a neutral particle.

The gauge boson masses are contained in the kinetic term of the Lagrangian and can be determined by evaluating its expression for the vacuum ground state. The covariant derivative becomes

$$\begin{aligned} D_\mu \phi &\rightarrow \left[ \partial_\mu - \frac{i}{2}g \begin{pmatrix} W_\mu^3 & W_\mu^1 - iW_\mu^2 \\ W_\mu^1 + iW_\mu^2 & -W_\mu^3 \end{pmatrix} - \frac{i}{2}g' B_\mu \right] \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix} \\ &= -\frac{i}{2} \begin{pmatrix} \frac{g}{\sqrt{2}}(W_\mu^1 - iW_\mu^2)(v + H) \\ i\sqrt{2}\partial_\mu H + \frac{1}{\sqrt{2}}(-gW_\mu^3 + g'B_\mu)(v + H) \end{pmatrix} \end{aligned} \quad (1-11)$$

and substituting this expression into Equation 1-8 while ignoring the potential term for clarity yields

$$\begin{aligned} \mathcal{L}_\phi &= \frac{1}{4} \left| \begin{pmatrix} \frac{g}{\sqrt{2}}(W_\mu^1 - iW_\mu^2)(v + H) \\ i\sqrt{2}\partial_\mu H + \frac{1}{\sqrt{2}}(-gW_\mu^3 + g'B_\mu)(v + H) \end{pmatrix} \right|^2 \\ &= \frac{1}{4} \left[ 2(\partial_\mu H)^2 + \left( \frac{g^2}{2} ((W_\mu^1)^2 + (W_\mu^2)^2) + \frac{1}{2} (g'B_\mu - gW_\mu^3)^2 \right) (v + H)^2 \right] \\ &= \frac{1}{2}(\partial_\mu H)^2 + \frac{g^2}{8}(W_\mu^1)^2 v^2 + \frac{g^2}{8}(W_\mu^2)^2 v^2 + \frac{1}{8} (g'B_\mu - gW_\mu^3)^2 v^2 + \mathcal{O}(H), \end{aligned} \quad (1-12)$$

where higher order terms in  $H$  have been discarded. The result of Equation 1-12 reveals that the broken symmetry of the Higgs field causes the  $W_\mu^1$  and  $W_\mu^2$  vector fields to both acquire a mass of

$$m_W^2 = \frac{g^2 v^2}{4} \implies m_W = \frac{gv}{2}. \quad (1-13)$$

The linear combination of vector fields  $(g'B_\mu - gW_\mu^3)$  also acquires a mass that depends on the couplings  $g$  and  $g'$ , and is typically expressed as

$$m_Z^2 = \frac{(g^2 + g'^2)v^2}{4} \implies m_Z = \frac{\sqrt{g^2 + g'^2}v}{2}. \quad (1-14)$$

The excitations of these massive vector fields are the weak gauge bosons  $W^+$ ,  $W^-$ , and  $Z^0$  which have been observed in nature and whose masses have been measured[11] to be

$$m_{W^+} = m_{W^-} = 80.379 \pm 0.012 \text{ GeV} \text{ and } m_{Z^0} = 91.1876 \pm 0.0021 \text{ GeV}. \quad (1-15)$$

The derivation would be complete but for a missing gauge boson. Recall our earlier transformation to the unitary gauge which fixed three of the four degrees of freedom of the Higgs field. According to Goldstone's Theorem[12], these degrees of freedom become the longitudinal polarizations of the now massive weak gauge bosons. The untouched degree of freedom thus corresponds to the massless photon  $\gamma$ . By taking the electromagnetic field  $A_\mu$  to be proportional to the missing linear combination of fields ( $g'W_\mu^3 + gB_\mu$ ), it can be explicitly introduced into the kinetic term of the Lagrangian while remaining orthogonal to the other fields.

### 1.1.3.2 Fermion masses

Although the Dirac Lagrangian describes the dynamics of the fermions, it is inadequate for explaining their observed masses. This is because a Dirac mass term of the form

$$m\bar{\psi}\psi = m(\bar{\psi}_L + \bar{\psi}_R)(\psi_L + \psi_R) = m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L) \quad (1-16)$$

violates the gauge invariance of the model by coupling together left-handed and right-handed fermions which transform differently under both SU(2) and U(1). The masses of the fermions are instead another consequence of spontaneous symmetry breaking, and their coupling to the scalar Higgs field through Yukawa interactions gives rise to their mass terms in the Standard Model Lagrangian.

The Yukawa interaction of the first generation leptons to the Higgs field  $\phi$  is given by

$$\mathcal{L}_{\text{Yukawa}} = -\lambda_e (\bar{L}\phi e_R + \bar{e}_R\phi^\dagger L), \quad (1-17)$$

where  $\lambda_e$  is the coupling constant, the SU(2) singlet  $e_R$  represents the right-handed electron, and the SU(2) doublet  $L$  represents the left-handed neutrino<sup>6</sup> and electron, i.e.

$$L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}. \quad (1-18)$$

Evaluating this expression for the vacuum ground state yields the electron mass term

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}} &= -\lambda_e \left( \bar{L} \phi_0 e_R + \bar{e}_R \phi_0^\dagger L \right) \\ &= -\frac{\lambda_e}{\sqrt{2}} \left[ \begin{pmatrix} \bar{\nu}_L & \bar{e}_L \end{pmatrix} \begin{pmatrix} 0 \\ v \end{pmatrix}_0 e_R + \bar{e}_R \begin{pmatrix} 0 & v \end{pmatrix}_0 \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \right] \\ &= -\frac{\lambda_e v}{\sqrt{2}} (\bar{e}_L e_R + \bar{e}_R e_L), \end{aligned} \quad (1-19)$$

which gives the electron mass

$$m_e = \frac{\lambda_e v}{\sqrt{2}}. \quad (1-20)$$

Therefore the electron, and similarly for the charged leptons of the other generations, is a mixture of left-handed and right-handed fields which acquire a mass that is proportional to the vacuum expectation value  $v$  and their coupling to the Higgs field.

The Yukawa interaction of the quarks, taking into account the existence of the right-handed SU(2) singlet  $u_R$ , is similarly given by

$$\mathcal{L}_{\text{Yukawa}} = -\Lambda_d^{ij} (\bar{Q}_L^i \phi d_R^j + \bar{d}_R^j \phi^\dagger Q_L^i) - \Lambda_u^{ij} (\bar{Q}_L^i \tilde{\phi} u_R^j + \bar{u}_R^j \tilde{\phi}^\dagger Q_L^i), \quad (1-21)$$

where  $\Lambda_d^{ij}$  and  $\Lambda_u^{ij}$  are  $3 \times 3$  complex matrices of couplings for the down-type and up-type quarks, respectively, whose indices  $i$  and  $j$  run over the quark generations,  $d_R$  and  $u_R$  are

<sup>6</sup> The neutrinos have been experimentally measured, within uncertainties, to have left-handed helicity only.

the right-handed singlets,  $Q_L$  is the left-handed doublet

$$Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \quad (1-22)$$

and  $\tilde{\phi}$  is the charge conjugate of the Higgs doublet

$$\tilde{\phi} = \begin{pmatrix} \phi^{0*} \\ -\phi^{+*} \end{pmatrix} \implies \tilde{\phi}_0 = \begin{pmatrix} v \\ 0 \end{pmatrix}. \quad (1-23)$$

Proceeding as before, evaluating this expression for the vacuum ground state yields

$$\mathcal{L}_{\text{Yukawa}} = -\frac{v}{\sqrt{2}} \Lambda_d^{ij} (\bar{d}_L^i d_R^j + \bar{d}_R^j d_L^i) - \frac{v}{\sqrt{2}} \Lambda_u^{ij} (\bar{u}_L^i u_R^j + \bar{u}_R^j u_L^i), \quad (1-24)$$

from which the quark masses are found to be

$$M_d^{ij} = \frac{\Lambda_d^{ij} v}{\sqrt{2}} \text{ and } M_u^{ij} = \frac{\Lambda_u^{ij} v}{\sqrt{2}}. \quad (1-25)$$

These mass terms resemble those of leptons except that they are matrices which depend on the Yukawa couplings of the mixed quark states to the Higgs field. The physical quark fields correspond to the mass eigenstates obtained by diagonalizing the mass matrices  $M_d^{ij}$  and  $M_u^{ij}$ . A general complex matrix may be diagonalized by a biunitary transformation, so there exist unitary matrices  $\mathbf{D}_L$ ,  $\mathbf{D}_R$ ,  $\mathbf{U}_L$ , and  $\mathbf{U}_R$  such that

$$\mathbf{D}_L^\dagger \mathbf{M}_d \mathbf{D}_R = \begin{pmatrix} m_d & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_b \end{pmatrix}, \quad \mathbf{U}_L^\dagger \mathbf{M}_u \mathbf{U}_R = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_t \end{pmatrix}, \quad (1-26)$$

where the mass eigenvalues are real and positive. But this transformation also amounts to diagonalizing the Yukawa coupling matrices, thus revealing that the quark masses also depend on their couplings and the vacuum expectation value in the same fashion as the charged leptons.

## 1.2 The Higgs Boson

While the Standard Model predicts the existence of the Higgs boson, it does not predict the value of its self-coupling  $\lambda$  and therefore its mass  $m_H$  is an unknown parameter of the model. Theoretical constraints based on unitarity bounds, the stability of the vacuum, and the energy scale of Standard Model physics only place the value of  $m_H$  within a wide range from 100 GeV up to 1 TeV. Stronger limits on  $m_H$ , not to mention an observation of the elusive boson, would have to be obtained from experimental results. During the 1980s, experimental searches for the Higgs boson began in earnest as particle accelerators finally reached energies that enabled the predicted mass range to be studied. The current era of high-energy physics experiment is now dominated by experiments at the Large Hadron Collider (LHC) and so the following qualitative review of Standard Model Higgs boson phenomenology is framed in its context.

### 1.2.1 Phenomenology

The Standard Model Higgs boson is predicted to be a spinless particle with zero electric or color charge that is even under the combined symmetry of charge conjugation and parity (CP-symmetry). It is also predicted to couple to the gauge bosons, fermions, and itself<sup>7</sup> in proportion to their masses according to

$$\begin{aligned} g_{Hf\bar{f}} &= \frac{m_f}{v}, & g_{HVV} &= \frac{2m_V^2}{v}, & g_{HHVV} &= \frac{2m_V^2}{v^2}, \\ g_{HHH} &= \frac{3m_H^2}{v}, & g_{HHHH} &= \frac{3m_H^2}{v^2}, \end{aligned} \tag{1-27}$$

where  $V = W$  or  $Z$  and  $v$  is the vacuum expectation value.<sup>[11]</sup> These interactions may be expressed as the Feynman diagram vertices shown in Figure 1-3. Because the Higgs boson might be coaxed into existence by these interactions, the production modes of the Higgs boson have informed the development of modern hadron colliders. By the same token,

---

<sup>7</sup> The Higgs self-couplings come from the higher order terms of Equation 1-6.

those interactions provide the means for the massive Higgs boson to decay into lighter and more stable particles. The decay modes of the Higgs boson thus influence the design of particle detectors capable of accurately measuring the properties of its decays products.

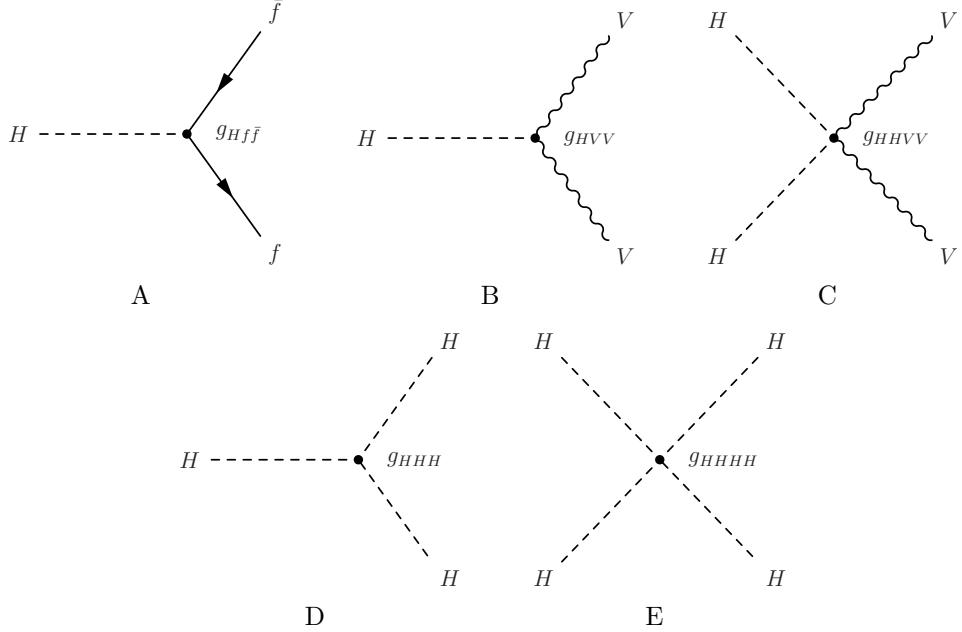


Figure 1-3. The Feynman diagrams for the vertices of the Standard Model Higgs boson representing the A) fermion-Higgs interaction; B) trilinear gauge-Higgs interaction; C) quartic gauge-Higgs interaction; D) trilinear self-interaction; E) quartic self-interaction.

### 1.2.1.1 Production modes

The production of massive particles such as the Higgs boson proceeds from the interactions of hadrons made to collide at high energies. The probability that such a collision results in a specific interaction is known as a cross section  $\sigma$ . The production cross sections of the Higgs boson depend on the center-of-mass energy of the collision  $\sqrt{s}$  and its mass  $m_H$ , and have been determined by theoretical calculations to behave as shown in Figure 1-4.

The dominant production mode is gluon fusion ( $pp \rightarrow H$ ), in which a gluon from each of the colliding hadrons form a virtual quark loop. Because the Higgs boson couples to quarks in proportion to their masses, it can be radiated from the resulting virtual

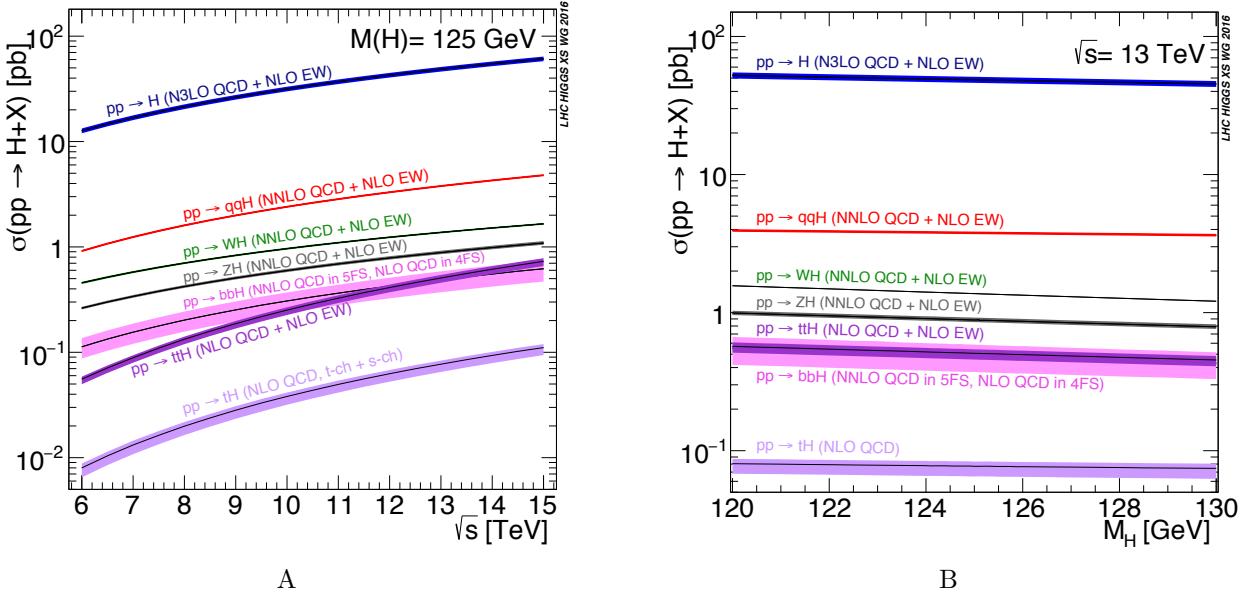


Figure 1-4. The production cross section for the Higgs boson as A) a function of the center of mass energy  $\sqrt{s}$  for  $m_H = 125 \text{ GeV}$ ; B) a function of  $m_H$  for  $\sqrt{s} = 13 \text{ TeV}$ . The shaded bands show the combined parametric and theoretical uncertainties and the labels indicate the production mode and any radiative corrections considered.[13]

top quark or, to a lesser extent, bottom quark loop. The production mode with the second-largest cross section is vector boson fusion ( $pp \rightarrow q\bar{q}H$ ), which proceeds from the parton scattering of two quarks, or anti-quarks, via their exchange of weak vector bosons  $V$ . This produces a Higgs boson because of the allowed trilinear gauge-Higgs interaction. The vector boson associated production mode ( $pp \rightarrow VH$ ), or *Higgstrahlung*, has the third-largest cross section. This production mechanism proceeds from the weak interaction of a quark and an anti-quark from which a Higgs boson is radiated by the virtual  $W$  or  $Z$  weak vector boson. Because the  $Z$  associated production can also be induced via a virtual top quark loop, contributions of nearly 15% to the total production cross section of this mode come from the process  $gg \rightarrow ZH$ . The fourth relevant production mode is top quark pair associated production ( $pp \rightarrow t\bar{t}H$ ). Rather than forming a virtual top quark loop, a gluon from each of the colliding hadrons decays into a top quark-antiquark pair and a Higgs boson is produced by the annihilation of a top quark of one pair and the

top anti-quark of the other. The smallest cross sections are contributed by the single top quark associated production and bottom quark pair associated production modes, which are suppressed relative to the main production modes. The Feynman diagrams for the main production modes are shown in Figure 1-5.

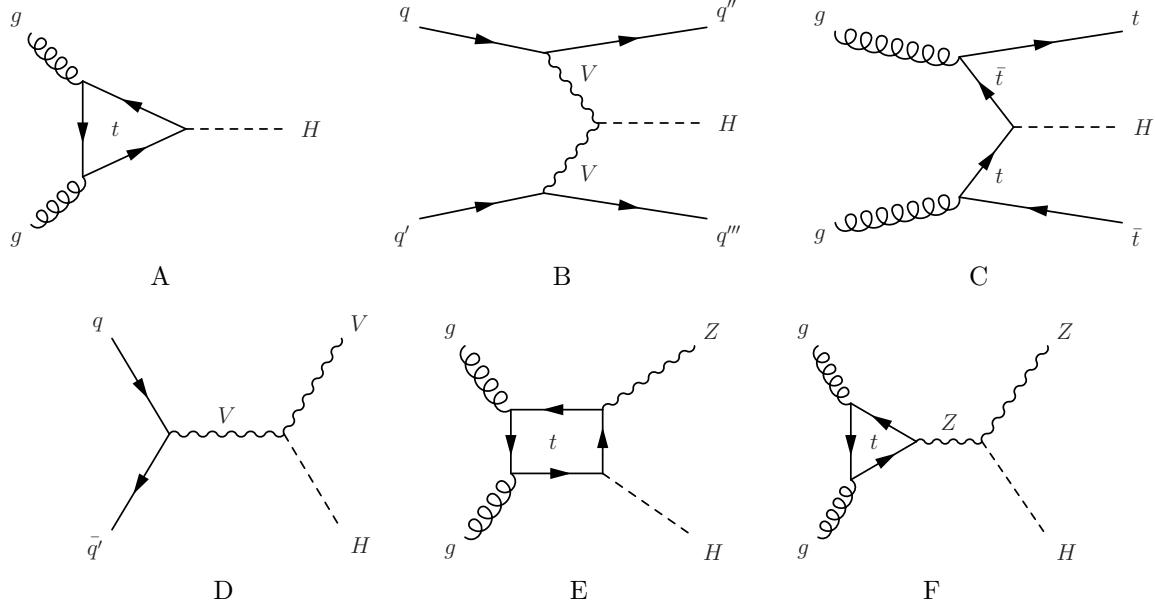


Figure 1-5. The Feynman diagrams for the Higgs boson production processes of A) gluon fusion; B) vector boson fusion; C) top quark pair associated production; D) vector boson associated production; E), F) top-loop induced contributions to the  $Z$  boson associated production.

For the current center-of-mass energy of proton collisions at the LHC,  $\sqrt{s} = 13$  TeV, and a Higgs boson mass of  $m_H = 125$  GeV, the state-of-the-art values computed for the production cross sections<sup>8</sup> are

$$\sigma(pp \rightarrow H) = 48.58^{+4.6\%}_{-6.7\%} \text{ (theory)} \pm 3.2\% \text{ (PDF + } \alpha_s \text{) pb},$$

$$\sigma(pp \rightarrow qqH) = 3.782^{+0.4\%}_{-0.3\%} \text{ (QCD scale)} \pm 2.1\% \text{ (PDF + } \alpha_s \text{) pb},$$

$$\sigma(pp \rightarrow WH) = 1.373^{+0.5\%}_{-0.7\%} \text{ (QCD scale)} \pm 1.9\% \text{ (PDF + } \alpha_s \text{) pb},$$

---

<sup>8</sup> The cross sections are presented in units of picobarns (pb), where 1 barn =  $10^{-24}$  cm<sup>2</sup>.

$$\begin{aligned}\sigma_{\text{total}}(pp \rightarrow ZH) &= 0.8839^{+3.8\%}_{-3.1\%} (\text{QCD scale}) \pm 1.6\% (\text{PDF} + \alpha_s) \text{ pb}, \\ \sigma(gg \rightarrow ZH) &= 0.1227^{+25.1\%}_{-18.9\%} (\text{QCD scale}) \pm 2.4\% (\text{PDF} + \alpha_s) \text{ pb}, \\ \sigma(pp \rightarrow ttH) &= 0.5071^{+5.8\%}_{-9.2\%} (\text{QCD scale}) \pm 3.6\% (\text{PDF} + \alpha_s) \text{ pb}.\end{aligned}$$

The uncertainties are due to variations on the QCD renormalization scale (QCD scale), the parton distribution function (PDF) that determines the momenta of the interacting partons, and the strong coupling constant ( $\alpha_s$ ). The accuracy of these cross sections are determined by the inclusion of higher-order terms during their calculation, which is generally next-to-next-to-leading (NNLO) order in QCD and next-to-leading order (NLO) in electroweak corrections. The exceptions are that the gluon fusion cross section is calculated further to N3LO in QCD, hence acquiring additional theoretical uncertainties, and the top quark pair associated production cross section which is calculated only to NLO in QCD. The methods and theoretical treatment of these cross section calculations are thoroughly documented in Ref. [13].

### 1.2.1.2 Decay modes

With a predicted lifetime<sup>9</sup> of  $1.6 \times 10^{-22}$  s, the Higgs boson decays almost immediately and its presence can only be inferred from its decay products. Based on its predicted properties, the final state of its decay must be electrically and color neutral. By conservation of mass, it must also decay into lighter particles so, for example, a decay into a top quark-antiquark pair  $t\bar{t}$  is forbidden for low mass Higgs bosons. However, because the Higgs boson couples to particles in proportion to their masses, the heaviest allowed decays will be the most favored.

The probability of a particular decay mode occurring is given by its *branching ratio*, and is shown in Figure 1-6 as a function of the Higgs boson mass  $m_H$ . For a Higgs boson

---

<sup>9</sup> The total decay width of the Higgs boson is predicted to be  $\Gamma_H = 4.088$  MeV and the lifetime of a particle is given by  $\tau = \hbar / \Gamma$ .

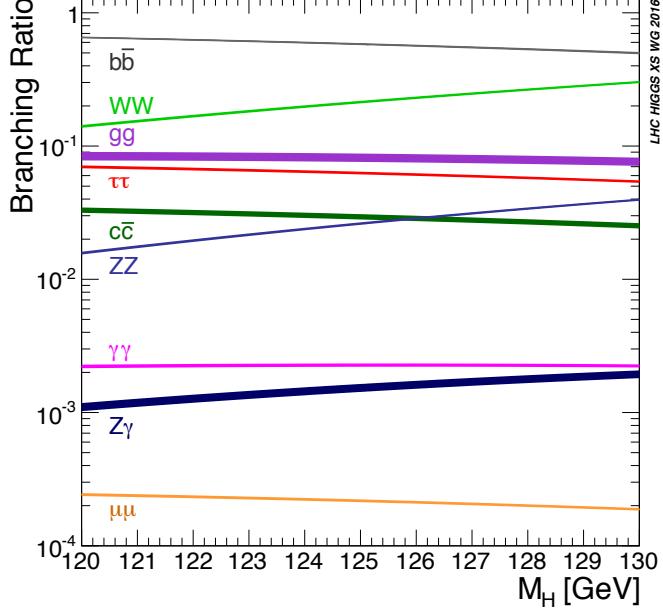


Figure 1-6. The branching ratios of the Higgs boson as a function of its mass  $m_H$ . [13].  
The shaded bands indicate the theoretical uncertainty of the calculation while the labels indicate the decay mode.

mass of  $m_H = 125$  GeV, the state-of-the-art values computed for the branching ratios of its allowed decays are, in descending order,

$$\begin{aligned}
\text{BR}(H \rightarrow b\bar{b}) &= 0.5824^{+0.65\%}_{-0.65\%} (\text{theory})^{+0.72\%}_{-0.74\%} (m_q)^{+0.78\%}_{-0.80\%} (\alpha_s), \\
\text{BR}(H \rightarrow WW) &= 0.2137^{+0.99\%}_{-0.99\%} (\text{theory})^{+0.99\%}_{-0.98\%} (m_q)^{+0.66\%}_{-0.63\%} (\alpha_s), \\
\text{BR}(H \rightarrow gg) &= 0.08187^{+3.40\%}_{-3.41\%} (\text{theory})^{+1.12\%}_{-1.13\%} (m_q)^{+3.69\%}_{-3.61\%} (\alpha_s), \\
\text{BR}(H \rightarrow \tau\bar{\tau}) &= 0.06272^{+1.17\%}_{-1.16\%} (\text{theory})^{+0.98\%}_{-0.98\%} (m_q)^{+0.62\%}_{-0.62\%} (\alpha_s), \\
\text{BR}(H \rightarrow c\bar{c}) &= 0.02891^{+1.20\%}_{-1.20\%} (\text{theory})^{+5.26\%}_{-0.98\%} (m_q)^{+1.25\%}_{-1.25\%} (\alpha_s), \\
\text{BR}(H \rightarrow ZZ) &= 0.02619^{+0.99\%}_{-0.99\%} (\text{theory})^{+0.99\%}_{-0.98\%} (m_q)^{+0.66\%}_{-0.63\%} (\alpha_s), \\
\text{BR}(H \rightarrow \gamma\gamma) &= 0.002270^{+1.73\%}_{-1.72\%} (\text{theory})^{+0.93\%}_{-0.99\%} (m_q)^{+0.61\%}_{-0.62\%} (\alpha_s), \\
\text{BR}(H \rightarrow Z\gamma) &= 0.001533^{+5.71\%}_{-5.71\%} (\text{theory})^{+0.98\%}_{-1.01\%} (m_q)^{+0.58\%}_{-0.65\%} (\alpha_s), \\
\text{BR}(H \rightarrow \mu\bar{\mu}) &= 0.0002176^{+1.23\%}_{-1.23\%} (\text{theory})^{+0.97\%}_{-0.99\%} (m_q)^{+0.59\%}_{-0.64\%} (\alpha_s).
\end{aligned}$$

The theoretical uncertainties account for missing higher-order QCD and electroweak corrections, while the parametric uncertainties cover the variations of the quark masses  $m_q$ , where  $q = c, b, t$ , and the strong coupling constant  $\alpha_s$  which are input parameters to the calculation. The methods and theoretical treatment of the computation of these branching ratios is also documented in Ref. [13].

### 1.2.2 Discovery

The discovery of a new boson with a mass close to 125 GeV was announced on July 4, 2012 at CERN, with independent observations achieved by the ATLAS[14] and CMS[15] collaborations. Within a year’s time, the new particle would be verified by both experiments to have zero spin and positive parity[16, 17] and its observed couplings would remain consistent with a Standard Model Higgs boson.[18] At this point, the new boson was recognized as a Higgs boson, and in 2013 the Nobel Prize in Physics would be awarded to François Englert and Peter Higgs for their discovery of the Higgs mechanism and prediction of a Higgs boson almost half a century before its discovery.

The decay channels with the highest sensitivities were  $H \rightarrow ZZ$ , with each  $Z$  boson subsequently decaying into a pair of charged leptons, and  $H \rightarrow \gamma\gamma$ , for which ATLAS and CMS both had observed local significances above the expected background of  $3\sigma$  and  $4\sigma$ , respectively. While those channels individually passed the  $3\sigma$  threshold for establishing evidence, it would take their combination with other channels for ATLAS to surpass and CMS to reach  $5\sigma$ , the threshold for declaring a discovery. By the beginning of 2015, when the LHC would resume operations since its first planned maintainence period, the Higgs production modes of gluon fusion and vector boson fusion had been observed, as well as its decays to  $WW$ ,  $ZZ$ , and  $\gamma\gamma$ .

## 1.3 Searches for $VH(b\bar{b})$

The preferred decay mode of the Higgs boson, as determined by its measured mass of  $m_H = 125.26 \pm 0.16$  GeV[11], is to a bottom quark-antiquark pair, henceforth  $H(b\bar{b})$ , with a branching ratio of nearly 59%. An observation would establish clearly the coupling

of the Higgs boson to bottom quarks, and to down-type quarks in general. Moreover, as the dominant decay mode, a precise measurement of its branching ratio has direct ramifications for improving the constraints on the total decay width of the Higgs boson and provides an opportunity to check for anomalous Yukawa couplings that may present a case for new physics beyond the Standard Model. The observation of  $H(b\bar{b})$  is therefore of great scientific interest and experiments must be dedicated towards searching for this decay.

### 1.3.1 Motivation for $VH(b\bar{b})$

The experimental challenge of searching for  $H(b\bar{b})$  is in distinguishing its decay signature, or signal, from the immense background of Standard Model processes. The bottom and antibottom quarks produced by the decay will both hadronize due to color confinement, forming two unstable B-hadrons that subsequently decay into cones of lighter particles known as jets. Such a final state is not unique, as QCD processes which produce many jets are a common occurrence. However, even if the pair of jets can be correctly identified as originating from bottom quarks and not lighter quarks or gluons, the multijet production of a bottom quark-antiquark pair occurs at a rate that is seven orders of magnitude greater than the gluon fusion production of a Higgs boson that decays into a bottom quark-antiquark pair, as shown in Figure 1-7. While advances in the field of jet substructure may eventually enable such a strategy[19], a direct search for final states containing only two  $b$ -jets faces a large and irreducible multijet background.

One solution for mitigating the multijet background is searching for  $H(b\bar{b})$  using the remaining Higgs boson production modes and exploiting their final state topologies. The vector boson fusion final state is fully hadronic, as the two scattering quarks will form their own jets, and is challenging in its own right. The top-quark associated production final state potentially contains leptons from the dominant decay of the top quark,  $t \rightarrow Wb$ , but a combinatorial complexity arises with the presence of additional  $b$ -jets. The key lies in using the weak vector boson associated production mode, henceforth  $VH(b\bar{b})$ , which

offers a final state that can be distinguished by the leptonic decays of either the  $W$  or  $Z$ , thereby enabling the reduction of the multijet background to negligible levels.

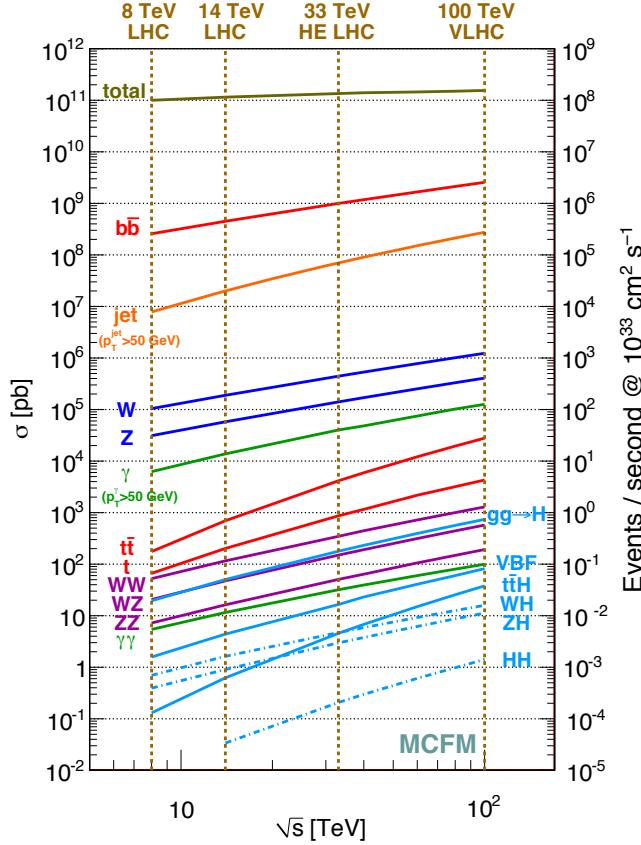


Figure 1-7. The predicted production cross sections of various processes for the range of center-of-mass energies achieved and within reach at the LHC.[20]

Searches based on  $VH(b\bar{b})$  must also contend with known Standard Model background processes, such as those measured in Figure 1-8, which mimic its final states. The main irreducible backgrounds come from  $W$  and  $Z$  bosons produced in association with jets, or  $V + \text{jets}$ , and the production of top quark-antiquark pairs, or  $t\bar{t}$ , which have cross sections that are three to four orders of magnitude larger than that of  $VH(b\bar{b})$ . The single top and diboson, or  $VV$ , processes are also important backgrounds, but with cross sections that are only one to two orders of magnitude larger than that of  $VH(b\bar{b})$ . The Feynman diagrams for examples of these background processes are shown in Figure 1-9.

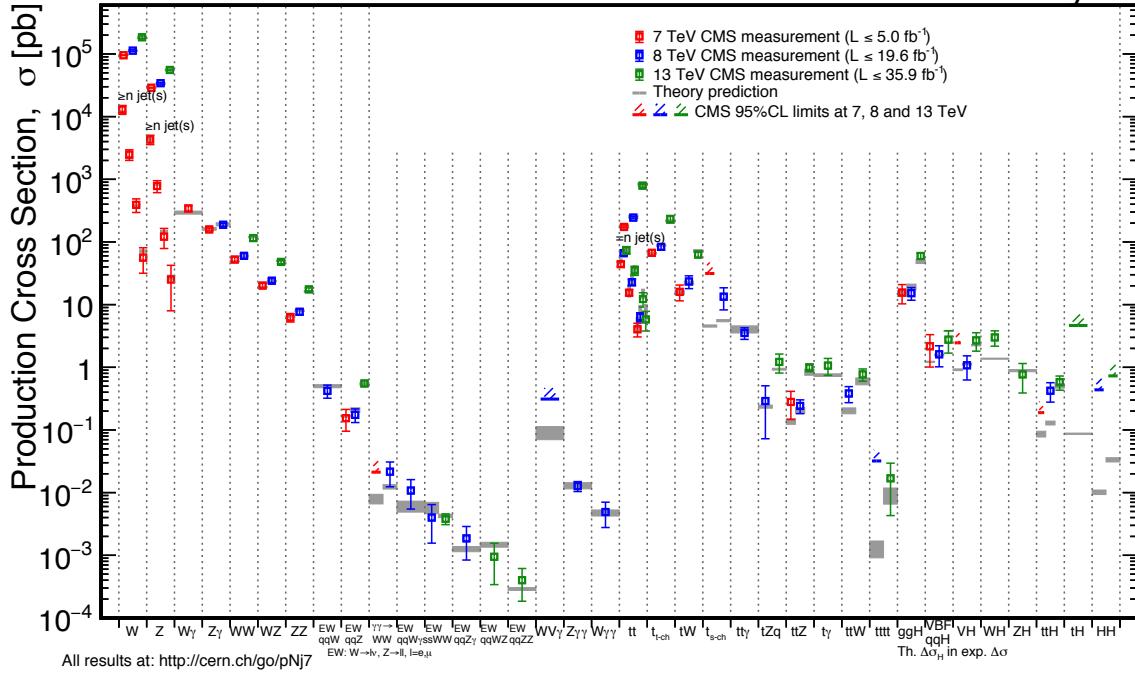


Figure 1-8. Measurements of the production cross sections of the Standard Model processes by the CMS experiment.[21]. The agreement with prediction holds remarkably well across the different center-of-mass energies of collisions at the LHC.

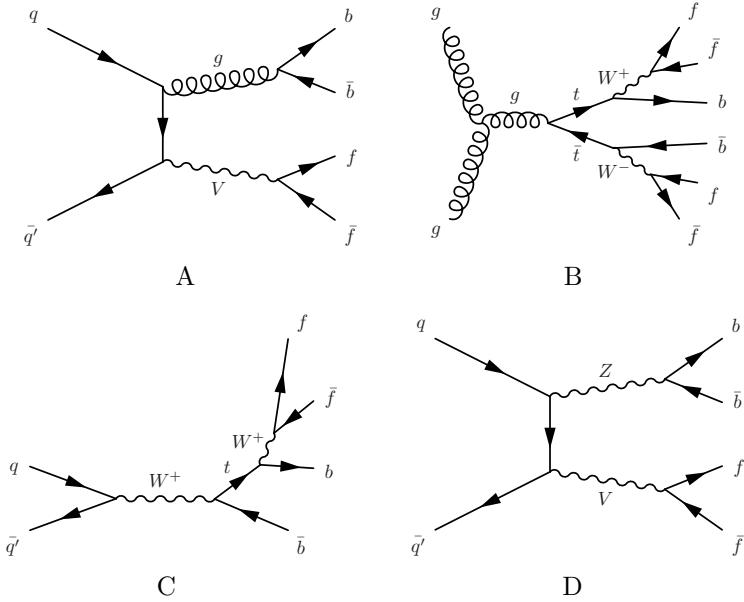


Figure 1-9. The Feynman diagrams for the Standard Model background processes to  $VH(b\bar{b})$ . A) vector boson production in association with jets ( $V + \text{jets}$ ); B) top quark-antiquark pair production ( $t\bar{t}$ ); C) single top D) diboson ( $VV$ ).

### 1.3.2 Previous Results

The first searches for  $VH(b\bar{b})$  coincided with the first searches for the Higgs boson and began at the Large Electron-Positron (LEP) collider which operated from 1989 to 2000 at CERN in Geneva, Switzerland. At LEP, a Higgs boson would have been produced primarily through  $Z$  boson associated production initiated by the annihilation of an electron-positron pair. Although the Higgs boson was not found, the four LEP experiments ALEPH, DELPHI, OPAL, and L3 analyzed the full dataset collected at center-of-mass energies above  $\sqrt{s} = 189$  GeV and established a lower bound of 114.4 GeV for the mass of the Higgs boson at the 95% confidence level.[22]

The first searches for  $VH(b\bar{b})$  using hadron collisions were done at the Tevatron at Fermi National Accelerator Laboratory (FNAL) in Batavia, Illinois. The Tevatron collided protons and antiprotons at center-of-mass energies up to  $\sqrt{s} = 1.96$  TeV and operated from 1985 until 2011. Although the Higgs boson was not observed, a combined analysis of the full Tevatron datasets collected by the CDF and D0 experiments yielded an observed significance of  $2.8\sigma$ , just shy of establishing evidence for the  $VH(b\bar{b})$  decay.[23]

The most recent searches for  $VH(b\bar{b})$  have taken place at the Large Hadron Collider (LHC), also located at CERN, using collisions between protons. During Run 1, which lasted from 2010 to 2013, the LHC collided protons at center-of-mass energies of  $\sqrt{s} = 7$  and 8 TeV. Analysis of the Run 1 datasets were unable to obtain evidence for  $VH(b\bar{b})$  with observed significances of  $1.4\sigma$  and  $2.5\sigma$  by the ATLAS and CMS experiments, respectively.[24, 25] A combination of these results only brought the observed significance up to  $2.6\sigma$ .[26]

After its upgrades during Long Shutdown 1 (LS1), the LHC resumed operations and Run 2 commenced with the center-of-mass energy of collisions increased to  $\sqrt{s} = 13$  TeV. Although the increase in energy meant a larger cross section for  $VH(b\bar{b})$ , it also led to relatively larger increases in the cross sections of the Standard Model background processes as demonstrated in Table 1-1, further complicating the search for  $VH(b\bar{b})$ .

However, this pessimism was unwarranted, as ATLAS and CMS both established evidence for the  $VH(b\bar{b})$  decay by observing significances of  $3.6\sigma$  and  $3.8\sigma$ , respectively, after combining the results of their initial Run 2 and Run 1 analyses.[\[27, 28\]](#)

Table 1-1. A list of Higgs boson and Standard Model background production cross sections at center-of-mass energies of  $\sqrt{s} = 8$  and 13 TeV, along with the relative increase at  $\sqrt{s} = 13$  TeV.

Cross Section	$\sqrt{s} = 8$ TeV	$\sqrt{s} = 13$ TeV	Relative Increase at $\sqrt{s} = 13$ TeV
$\sigma(pp \rightarrow H)$	19.4 pb	44.1 pb	2.27
$\sigma(pp \rightarrow qqH)$	1.6 pb	3.8 pb	2.38
$\sigma(pp \rightarrow VH)$	1.23 pb	2.26 pb	1.84
$\sigma(pp \rightarrow ttH)$	133 fb	507 fb	3.81
$\sigma(pp \rightarrow t\bar{t})$	253 pb	832 pb	3.29

Naive projections assuming increased statistics and similar systematic uncertainties suggested that an observation of  $VH(b\bar{b})$  may be possible if the experimental sensitivity could be improved by about 20%. With Run 2 on track to produce more data in 2017 than 2016, such a historic result appeared to be on the horizon. The analysis of the 2017 dataset and the results obtained by the CMS experiment are the subject of this dissertation.

## CHAPTER 2

### EXPERIMENTAL APPARATUS

It is nature's irony, that the study of subatomic particles requires the largest and most intricate machines known to date in human history. Because the physics of such particles are described by their interactions, high energy physics experiments typically accelerate particles to extreme kinetic energies, collide them against each other or dense targets, and infer the details of their interaction from the resulting final state particles. Such collisions can be *elastic*, where the interacting particles merely exchange energy and remain unchanged, or *inelastic*, where the interaction changes the incoming particles and has a chance of creating new, massive particles because of the time-energy uncertainty relation. Collisions of the latter kind are thus of great interest to experimentalists in the search for new particles and their decays.

Under the consideration of providing as much energy as possible towards the interaction, modern experiments favor colliding beams over fixed-target experiments. By momentum conservation, some of the energy of a fixed-target collision must be converted into the kinetic energies of final state particles, while there is no such constraint for colliding beams. Thus, at high energies, the dependence of the center-of-mass energy  $\sqrt{s}$  on the incoming beam energy  $E_{\text{beam}}$  in the laboratory frame scales as  $\sqrt{s} \propto E_{\text{beam}}^{1/2}$  for fixed-targets, whereas it scales linearly for colliding beams  $\sqrt{s} \propto E_{\text{beam}}$ . Because the center-of-mass energy of the collision fixes the energy available for the production of new particles, colliding beams hold the advantage.

Modern experiments also favor circular rather than linear particle accelerators. Because the length of a linear accelerator is correlated with its maximum beam energy, the size required to probe the current energy frontier is prohibitively large, and even then particles can only be accelerated once through the machine. Instead, circular colliders can accelerate particles for multiple orbits until they reach the desired energy scale, and this circulation also allows for the same beams to be collided multiple times to achieve

a *luminosity*, a measure of particle flux, that is higher than linear accelerators. The problem faced by circular colliders is synchrotron radiation, where charged particles accelerated along a circular trajectory lose energy. Because a particle's energy loss due to synchrotron radiation is inversely proportional to the fourth power of its mass, electron beams experience incredible energy losses compared to proton beams. For this reason, modern experiments prefer the acceleration and collision of hadron beams.

## 2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) at CERN is the culmination of contemporary high energy experiment sensibilities. The world's largest machine, it is a circular accelerator with a circumference of 27 km, spanning across the Franco-Swiss border, and is housed within an underground tunnel at an average depth of 100 m. An illustration is provided in Figure 2-1. It features two parallel evacuated beam pipes in which proton beams circulate in opposite directions. Although the design energy of each proton beam is 7 TeV, the beams are currently circulating at an energy of 6.5 TeV to give a center-of-mass energy of 13 TeV for the one billion proton collisions it produces per second.

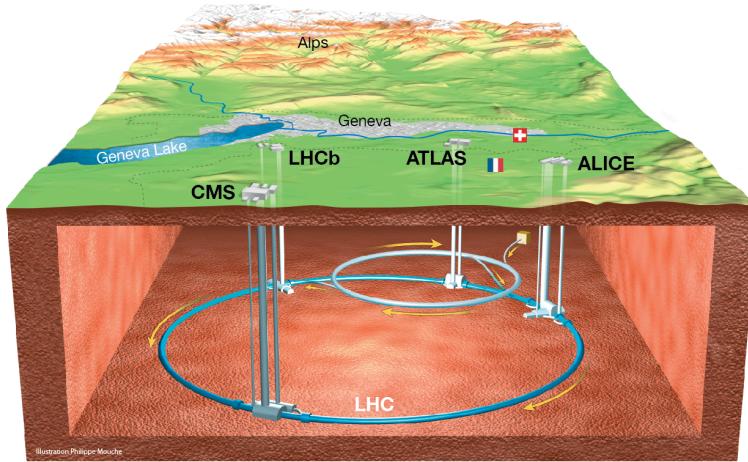


Figure 2-1. An overall view of the LHC showing the underground tunnel and its experiments in relation to its geography.[29]

Because the LHC reuses the tunnels of the Large Electron-Positron (LEP) Collider, additional excavation was avoided and it was integrated as the final stage of CERN's accelerator complex, shown in Figure 2-2. The source of protons is a canister of hydrogen gas to which a strong electric field is applied to strip the electrons from the hydrogen nuclei. The protons then pass through the linear accelerator LINAC2 and are accelerated to an energy of 50 MeV. They are then injected into successive circular accelerators: the Proton Synchrotron Booster which accelerates the beams to 1.4 GeV, the Proton Synchrotron (PS) which accelerates the beams to 25 GeV, the Super Proton Synchrotron (SPS) which further accelerates the beams to 450 GeV, and finally the LHC itself. The oscillating radio frequency (RF) fields used to accelerate the protons results in beams composed of proton *bunches* as opposed to a continuous stream. The LHC currently circulates up to 2556 bunches per beam, which is not far from its design value of 2808 bunches per beam.

The circular orbit of the proton beams are maintained by the 1,232 bending magnets lining the LHC. The bending magnets are superconducting dipole magnets which generate 8.3 T magnetic fields powerful enough to bend the high energy beams. In order to minimize the oscillation of the bunches around their trajectories and achieve high luminosity, 392 focusing magnets are also placed along the LHC ring. The focusing magnets are superconducting quadrupole magnets with a magnetic field gradient of 223 T/m. They are placed in alternating pairs to account for their nature of focusing along one direction while defocusing the orthogonal direction. Schematics for these two main types of magnets are shown in Figure 2-3. In order to maintain the superconductivity of these magnets, the cryogenic systems cool them down to a temperature of 1.9 K.

The beams are directed to cross through each other at four sites around the LHC known as interactions points. At these points, fine control of the beam is crucial because it has direct ramifications on the number of possible interactions. The interaction rate of any

## CERN's Accelerator Complex

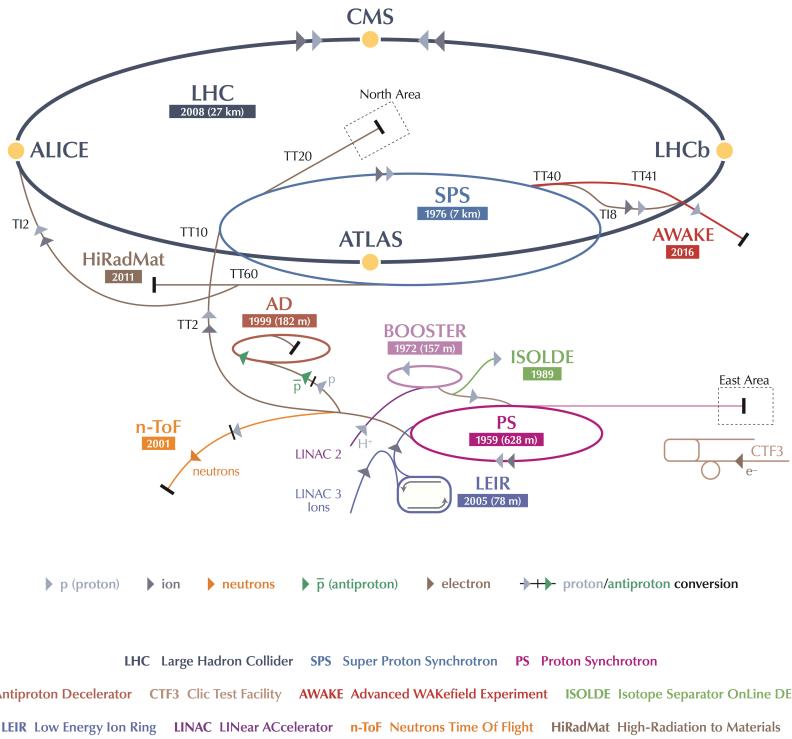


Figure 2-2. A complete schematic of the CERN acerclator complex and experiments.[30]

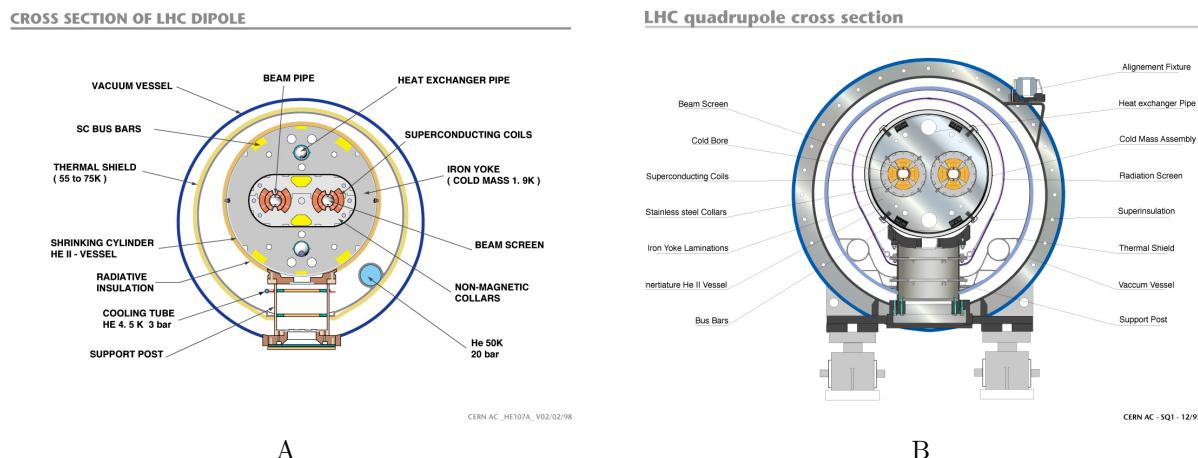


Figure 2-3. The cross sectional views for the main types of LHC magnets: A) dipole magnet used for bending the beams; B) quadrupole magnet used for focusing the beams.[31, 32]

given process is proportional to its cross section  $\sigma$  and the luminosity  $\mathcal{L}$  of the collider

$$\frac{dN}{dt} = \sigma \mathcal{L}, \quad (2-1)$$

where  $N$  refers to the number of interactions. The cross section, or interaction probability, is fixed by physics while the luminosity is determined by the collider's design, and therefore the highest possible luminosity is desired in order to observe rare interactions such as those that produce Higgs bosons. For equal, bunched beams with a rounded profile, the luminosity can be expressed as

$$\mathcal{L} = \frac{N^2 k_b f \gamma}{4\pi \epsilon_n \beta^*} F, \quad (2-2)$$

where the various parameters are determined by the design of the collider:  $N$  is the number of protons per bunch,  $k_b$  is the number of bunches,  $f$  is the revolution frequency,  $\gamma$  is the typical Lorentz factor,  $\epsilon_n$  is the normalized emittance,  $\beta^*$  is the value of the betatron function at the interaction point, and  $F$  is the geometrical reduction factor due to the crossing angle.[33]

The LHC has a 25 ns bunch spacing, or time between bunch crossings at the interaction point, which places a constraint on how large  $k_b$  and  $f$  can be. The spacing cannot be lowered due to operational safety concerns, such as providing kicker magnets enough time to ramp up and divert the bunches toward a beam dump for absorption. The normalized emittance  $\epsilon_n$  is a measure of the kinematic phase space occupied by the particles in the bunch and can be improved by reducing the size of the bunch and the spread in momentum of its particles. The beta function is a measure of the beam's transverse size and so its value at the interaction point  $\beta^*$  is determined by how well the magnets can focus the beam at the interaction point. The geometrical factor  $F$  corrects the amount by which the colliding bunches overlap as they are made to cross at a shallow angle. While a head-on collision would result in the greatest overlap, it also increases

the chances of long range electromagnetic interactions between bunches, so there is some interplay in determining the crossing angle.

Although the design luminosity of the LHC is  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , it has exceeded expectations by achieving a luminosity of twice the designed value. As a measure of particle flux, the luminosity that has been discussed so far is a dynamic quantity and is better called *instantaneous luminosity*. The expected number of interactions  $N$  for a specific process is thus given by

$$N = \sigma \int \mathcal{L} dt = \sigma \mathcal{L}_{int}, \quad (2-3)$$

where  $\mathcal{L}_{int}$  is the *integrated luminosity*, and so the total integrated luminosity is quoted in reference to the size of particle physics datasets.

Surrounding each of the interaction points of the LHC are its four main experiments: A Large Ion Collider Experiment (ALICE), A Toroidal LHC ApparatuS (ATLAS), the Compact Muon Solenoid (CMS), and the Large Hadron Collider beauty (LHCb). Both ATLAS and CMS are general-purpose detectors designed to search for the Higgs boson and perform precise measurements of its properties, while also exploring the new energy frontier in search of physics beyond the Standard Model. The LHC can also accelerate lead ions, which is used by ALICE to study heavy-ion collisions and the dynamics of quark-gluon plasma. As general-purpose detectors, ATLAS and CMS also study heavy-ion collisions. The LHCb experiment specializes in studying the electroweak and QCD physics of heavy flavor quarks and measuring the properties of B mesons.

## 2.2 The Compact Muon Solenoid

The Compact Muon Solenoid (CMS), illustrated in Figure 2-4, is a general-purpose particle detector and one of the two main experiments at the LHC. Its cylindrical design is impressively compact for its purpose, with a diameter of 15 m and a length of 28.7 m. Although heavy, weighing in at 14,000 tonnes, the bulk of this weight comes from its steel return yoke and structural supports which together weigh 12,500 tonnes. The return

yoke guides and contains the 3.8 T magnetic field generated by the detector's namesake superconducting solenoid, which is cooled by its cryostat to a temperature of 4.5 K during operation.

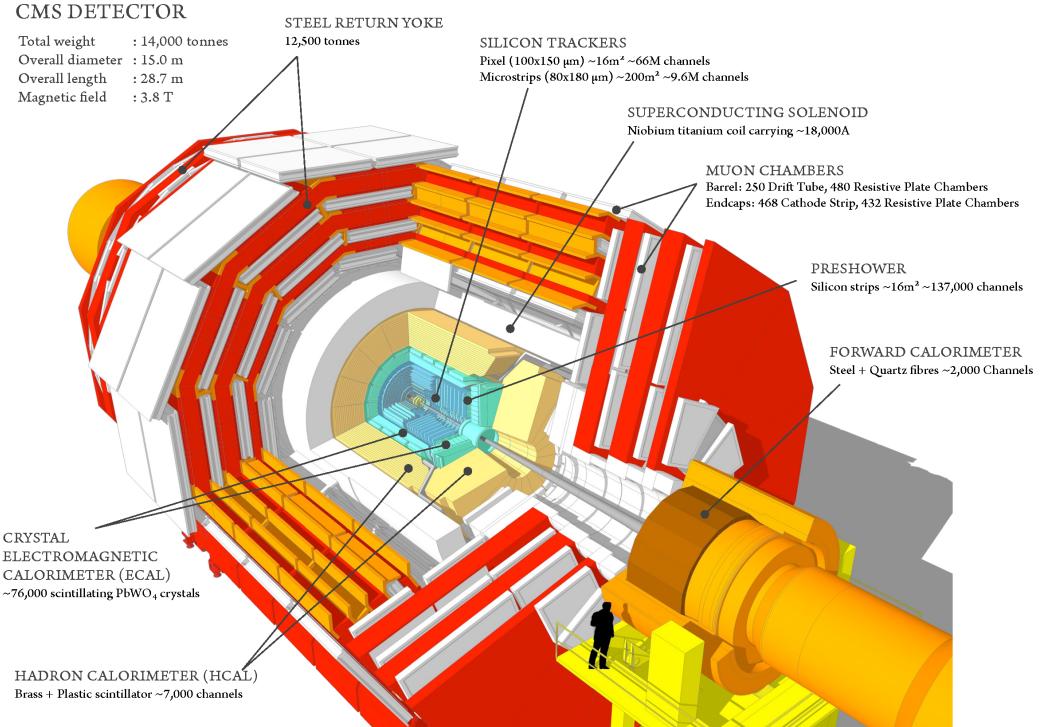


Figure 2-4. A cutaway view of the CMS detector with its main subsystems and components labeled.[34]

The solenoid is central to the design of the CMS detector because it provides a uniform magnetic field capable of bending the trajectories of charged particles as they travel through the detector, thereby enabling the measurement of their electric charge and transverse momentum  $p_T$ . This follows because a particle with electric charge  $q$  and velocity  $\mathbf{v}$  moving through a uniform magnetic field  $\mathbf{B}$  experiences a Lorentz force given by

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B}. \quad (2-4)$$

Since the force is transverse to the direction of motion, the particle travels along a helical path of radius  $R$  with a handedness determined by its electric charge. An application of Newton's 2<sup>nd</sup> Law upon the circular motion of such a particle with mass  $m$  and velocity

transverse to the direction of motion  $v$  further shows that its transverse momentum is determined by the radius of curvature of its path

$$\frac{mv^2}{R} = qvB \implies p_T = qRB, \quad (2-5)$$

where  $p_T = mv$  is the transverse momentum of the particle.

While the helical paths of charged particles travelling with low momentum are fully contained within the detector, the trajectories of charged particles with high momentum are recorded as incomplete arcs, either because they decay mid-flight or, in the case of muons, travel out of the detector. While the radius of curvature of such highly energetic particles cannot be directly measured, it can be obtained by measuring the sagitta of their *track*, the arc they trace through the detector.

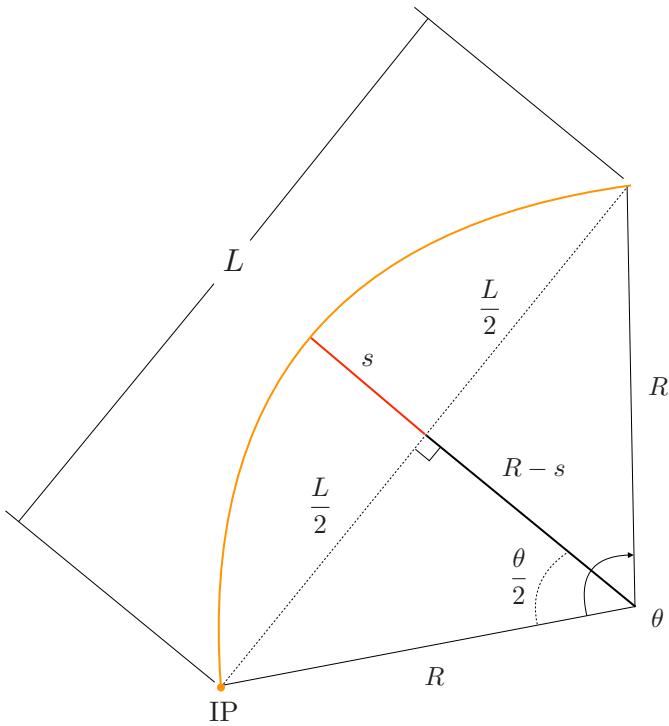


Figure 2-5. A diagram defining the sagitta of an arc representing a charged particle track (orange) travelling away from the interaction point (IP). The track has a chord of length  $L$ , an opening angle of  $\theta$ , a radius of curvature  $R$ , and a sagitta (red) of length  $s$ .

From Figure 2-5, a relationship between the chord length  $L$ , radius of curvature  $R$ , and sagitta  $s$  of the track can be constructed using the Pythagorean Theorem such that

$$R^2 = (R - s)^2 + \left(\frac{L}{2}\right)^2. \quad (2-6)$$

Solving for  $R$  with the reasonable assumption that  $s \ll L$  yields

$$R = \frac{L^2}{8s} + \frac{s}{2} \approx \frac{L^2}{8s}, \quad (2-7)$$

and therefore the  $p_T$  of a charged particle can be determined from the chord length and sagitta of its track

$$p_T = \frac{qBL^2}{8s}. \quad (2-8)$$

The magnetic field strength  $B$  and, in part, the chord length  $L$  are dictated by the design of the detector, and so the  $p_T$  resolution is limited by the resolution of the sagitta. As their momentum increases, charged particles leave straighter tracks which make it difficult to accurately measure the sagitta, thereby degrading the resolution at high momentum.

Because the magnetic field within the solenoid is oriented along the direction of the beam, it motivates the coordinate system adopted for the CMS detector that is shown in Figure 2-6. The geometrical center of the detector defines the origin of the right-handed Cartesian coordinate system with the  $z$ -axis oriented along the direction of the anticlockwise proton beam from the LHC, the  $x$ -axis pointing horizontally towards the center of the LHC ring, and the  $y$ -axis pointing vertically upwards out of the plane of the LHC ring. The shape of the detector also permits a cylindrical description about the same origin. With the  $x$ -axis and  $z$ -axis taken to be the polar and longitudinal axes, respectively, a cylindrical coordinate system can be defined with the azimuthal angle  $\phi$  taken with respect to the positive  $x$ -axis ( $\phi = 0$ ) and the polar angle  $\theta$  taken with respect to the positive  $z$ -axis ( $\theta = 0$ ). Because the polar angle  $\theta$  is not Lorentz invariant under

boosts along the direction of the beam, it is typically transformed into the *pseudorapidity*

$$\eta = -\ln \left[ \tan \left( \frac{\theta}{2} \right) \right] \quad (2-9)$$

such that the positive  $z$ -axis has  $\eta = +\infty$  and the positive  $y$ -axis has  $\eta = 0$ .

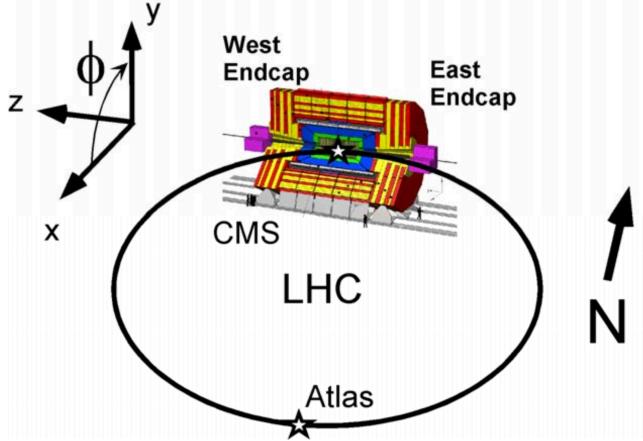


Figure 2-6. The coordinate system convention used for the CMS detector.[\[35\]](#)

### 2.2.1 Detector Subsystems

As with all modern general-purpose detectors, the guiding principle in the design of the CMS detector is hermeticity. By maximizing the angular coverage of the space surrounding the interaction point, the final state particles produced by proton-proton collisions are more likely to be recorded by the detector's subsystems. The CMS detector is composed of layered subsystems as illustrated in Figure 2-7, each designed to detect and measure the properties of different types of particles. Each of the detector subsystems are described in the following sections, starting from those closest to the interaction point. A thorough discussion of the detector's design is available in Ref. [\[33\]](#), the first volume of the official CMS Technical Design Report (TDR).

#### 2.2.1.1 Silicon tracker

The tracker is designed to detect charged particles and measure their trajectories or “tracks”. As the innermost layer surrounding the interaction point, its components must have a fast response time to record the tracks of short-lived particles, good spatial

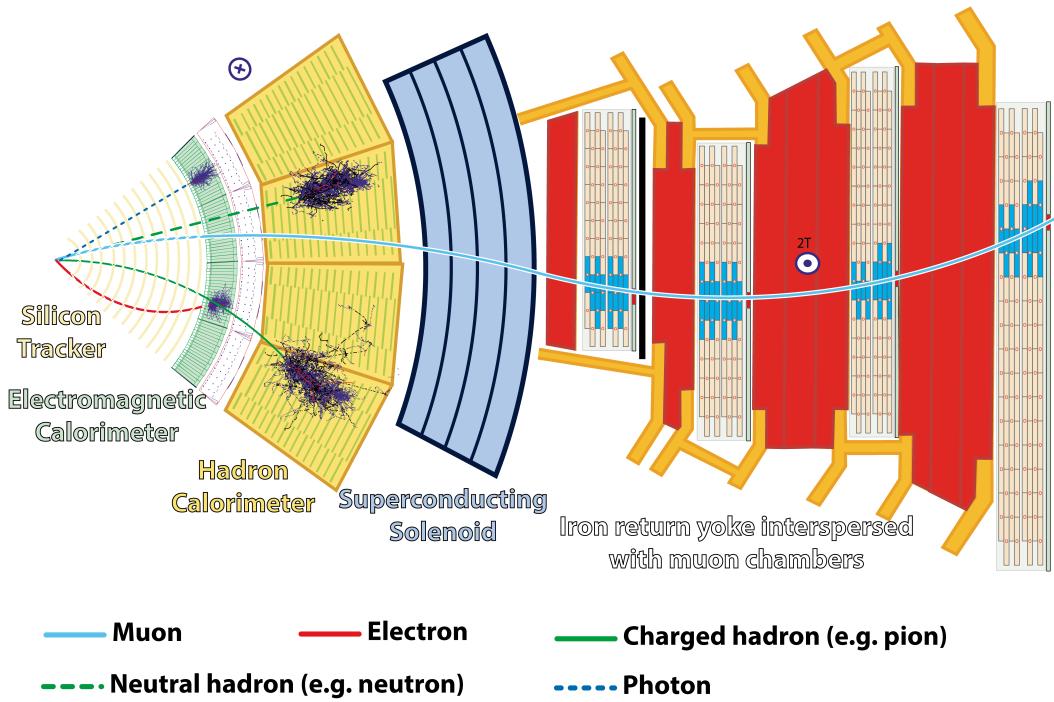


Figure 2-7. A cross sectional view of the CMS detector in the  $r\phi$ -plane revealing the different subsystems and the particles they are used to detect.[36]

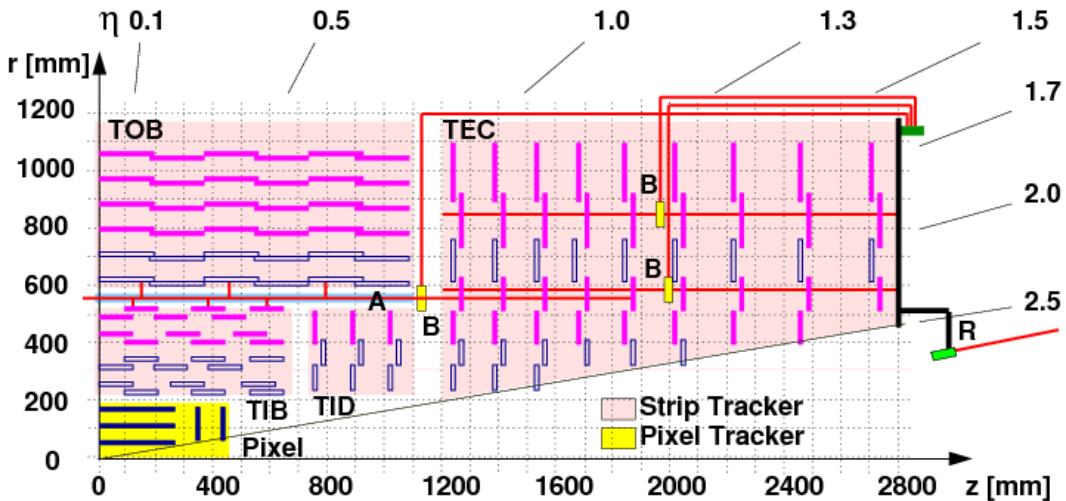


Figure 2-8. A schematic diagram showing a cross sectional view of the CMS detector's silicon tracker in the positive  $yz$ -plane.[37]

resolution to distinguish between tracks of different particles, and excellent radiation tolerance to withstand the high intensity of the proton beams. These concerns motivated the choice of semiconductor detectors made from silicon, which offer fast response times and excellent energy resolution and acceptable radiation hardness, although degradation is expected over time. When a charged particle passes through a semiconductor module, valence band electrons are promoted to the conduction band and leave behind “holes”. By applying an electric field, the electrons and holes are collected by electrodes to generate a signal indicating the passage of a charged particle.

The tracker is composed of 16,588 silicon modules distributed between the pixel tracker, which segments its sensors into 66 million pixels of size  $100 \times 150 \mu\text{m}^2$ <sup>21</sup>, and strip tracker, which segments its sensors into 9.6 million strips that are 80 to 180  $\mu\text{m}$  wide. The tracker is divided into four regions as shown in Figure 2-8, with the inner barrel (TIB) and outer barrel (TOB) arranged in cylindrical layers and the inner disks (TID) and endcaps (TEC) arranged in layered disks. This offers full angular coverage in  $\phi$  and  $|\eta| < 2.5$ , and the granularity of the pixels and strips achieves a  $p_T$  resolution of approximately 1.5% for charged particles with  $1 \text{ GeV} < p_T < 10 \text{ GeV}$  in the central region of the tracker.[38]

### 2.2.1.2 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) is designed to detect electrons and photons based on their energy deposition. When highly energetic electrons or photons interact with matter, they initiate an electromagnetic shower via bremsstrahlung or electron-positron pair production, respectively. Because these secondary particles may interact further with the material, the electromagnetic shower proceeds until the energies of the produced particles fall below a critical energy and energy losses proceed through ionization. If the material is a scintillator, it can absorb the low energy particles produced by the shower

---

<sup>21</sup> For reference, the average width of a human hair is around 100  $\mu\text{m}$ .

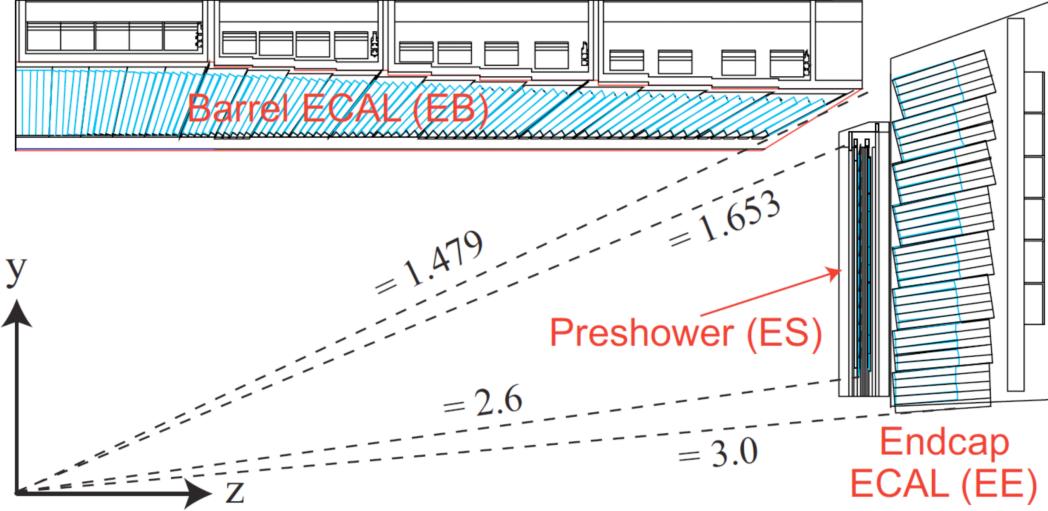


Figure 2-9. A schematic diagram showing a cross sectional view of the CMS detector’s electromagnetic calorimeter in the positive  $yz$ -plane.[39]

and produce an amount of light proportional to their energies that can be measured by photodetectors.

To ensure that the energies of incident electrons and photons are collected by the ECAL, lead tungstate ( $\text{PbWO}_4$ ) crystals were chosen for their low radiation length and Molière radius, fast response, and radiation hardness. The radiation length is the mean distance over which a particle radiates approximately 63% of their energy, while the Molière radius specifies the radius of the cylindrical volume that contains approximately 90% of the energy of the incident electron or photon. By selecting a material that minimizes both of these properties, the full absorption of electrons and photons can be achieved within a limited material budget.

The ECAL is composed of 75,848 lead tungstate crystals distributed between the ECAL barrel (EB) which uses 61,200 crystals and the ECAL endcap (EE) which uses 14,648 crystals, as shown in Figure 2-9. The additional ECAL preshower (EC) layer is placed before the EE and consists of alternating lead and silicon modules instead of lead tungstate crystals. The finer granularity of the EC is used to reject closely-spaced photon pairs produced by the decay of neutral pions based on their energy deposits, which could otherwise be identified as a high energy photon by the ECAL. On the whole, the ECAL

offers full angular coverage in  $\phi$  and  $|\eta| < 3.0$  and achieves an energy resolution between 2-5% for electrons and 1.1-5% for photons.[40]

### 2.2.1.3 Hadronic calorimeter

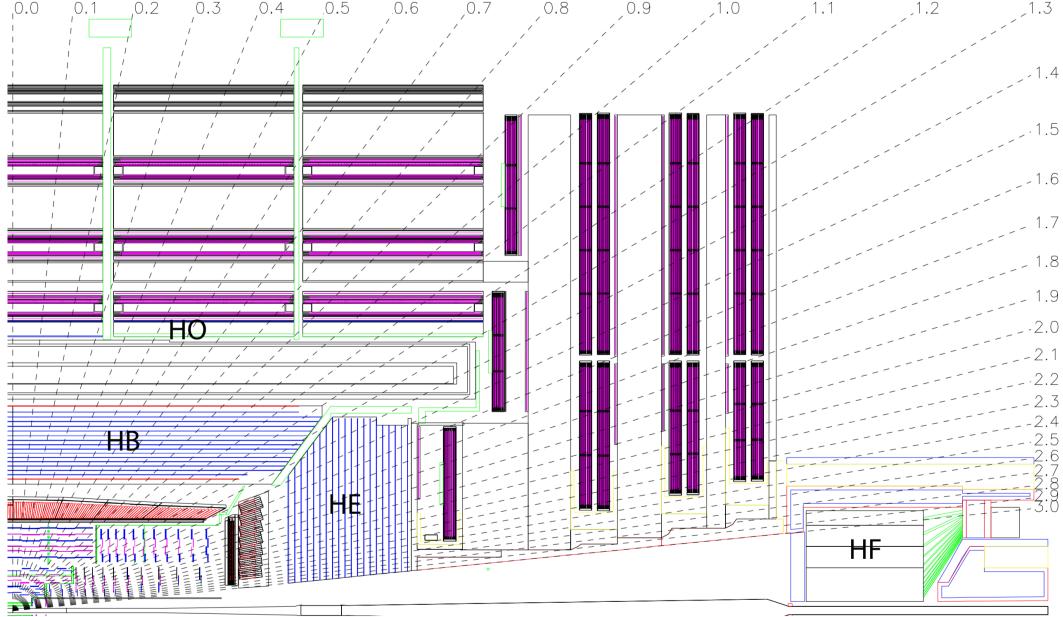


Figure 2-10. A schematic diagram showing a cross sectional view of the CMS detector's hadronic calorimeter in the positive  $yz$ -plane.[41]

The hadronic calorimeter (HCAL) is designed to detect charged and neutral hadrons based on their energy deposition. When hadrons interact with matter, they initiate a hadronic shower which produces a cascade of particles through the production of secondary hadrons, nuclear deexcitation, and pion and muon decays. Unlike electromagnetic showers, the lengths of hadronic showers are determined by the nuclear interaction length, or the mean distance travelled by a hadron within a material before experiencing an inelastic nuclear interaction, which is larger than the radiation length for the same materials. The HCAL was therefore chosen to be made from brass, which has a short interaction length that maximizes the amount of material to absorb hadrons. Unlike the ECAL, which is a *homogeneous calorimeter* because its crystals are both absorber and scintillator, the HCAL is a *sampling calorimeter* composed of alternating layers of brass absorber and plastic scintillator tiles which necessarily means the hadronic shower is

only sampled at discrete times during its evolution. The scintillation light is directed by wave-length shifting fibers embedded within the tiles to hybrid photodiodes which produce an amplified electronic signal.

The HCAL is divided into the inner barrel (HB), outer barrel (HO), and endcap (HE) regions as shown in Figure 2-10. This offers full angular coverage in  $\phi$  and  $|\eta| < 3.0$ . In order to increase the coverage at high  $\eta$  to recover particles produced at shallow angles to the beam line, an additional hadronic forward (HF) calorimeters are placed outside both endcaps. The HF is composed of steel and quartz fibers which collect light in the form of Cerenkov radiation and extends coverage down to  $|\eta| < 5.0$ . As the HCAL is concerned with the measurement of hadrons, its performance is typically assessed in the context of jet and missing transverse energy  $p_T^{\text{miss}}$  resolution measurements and is covered in Chapter 3.

#### 2.2.1.4 Muon subsystems

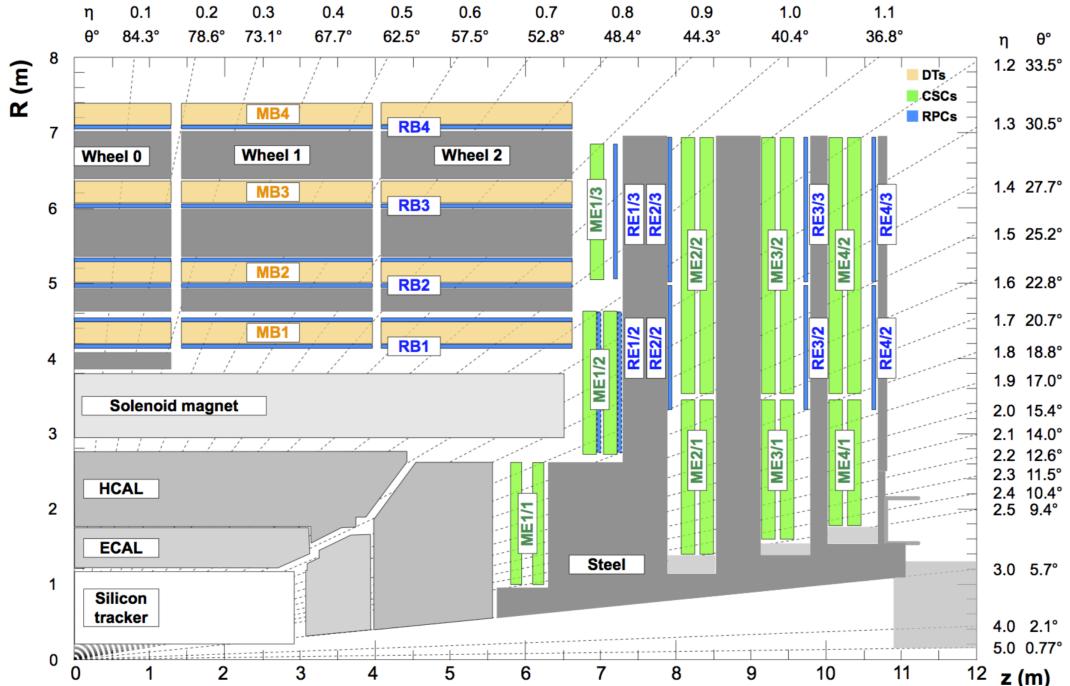


Figure 2-11. A schematic diagram showing a cross sectional view of the CMS detector's muon subsystems in the positive  $yz$ -plane.[42]

The muon subsystems form the outermost layer of the CMS detector and are designed to detect muons and measure their tracks. It is located far from the interaction point because the muon is a weakly interacting particle that easily penetrates through all preceding layers. While its function is similar to that of the silicon tracker, the muon subsystems eschew solid state for gas ionization chambers to accommodate the amount of surface area required to be covered. The muon subsystems are divided as shown in Figure 2-11 into barrel ( $|\eta| < 0.9$ , overlap ( $0.9 < |\eta| < 1.2$ ), and endcap ( $1.2 < |\eta| < 2.5$ ) regions which deploy three different implementations of gas ionization chambers.

Drift tubes (DTs) are used in the barrel region because of the low muon rate and magnetic field. They are 4 cm wide tubes containing an anode wire and a gaseous mixture of argon and carbon dioxide with a ratio of 85% to 15%, respectively. The passage of a muon through the tube ionizes the gas molecules, whose electrons “drift” towards the anode wiring under the force of an applied electric field. The position of the muon is determined by the location of the signal along the wire and the travel time based on the drift velocity of electrons within the tube.

Cathode strip chambers (CSCs) are used in the endcap region because of the high muon rate and non-uniform magnetic field. These trapezoidal chambers are composed of six gaps filled with a gas mixture of 40% argon, 50% carbon dioxide, and 10% carbon tetrafluoride ( $CF_4$ ) that house planes of cathode strips oriented radially and anode wires oriented perpendicular to the strips. The passage of a muon through the chamber ionizes the gas molecules, with the electrons travelling to the anode wire and creating an avalanche of electrons and the positive ions travelling to the cathode strips and creating charge pulses. Because the wires and strips are perpendicular, the  $r\phi$  position of the muon can be determined at six points and combined to form a track local to the chamber.

Resistive plate chambers (RPCs) are used both in the barrel and endcap regions, offering good spatial resolution and an excellent time resolution of approximately 3 ns. They consist of two parallel plates made of a resin with high resistivity, one of which is an

anode and the other a cathode. The plates surround a 2 mm wide gap filled with a gas mixture of 95% tetrafluoroethane ( $C_2H_2F_4$ ) and 5% isobutane ( $i - C_4H_{10}$ ). The passage of a muon through the chamber ionizes the gas molecules, producing an avalanche of electrons which are collected by metallic strips attached to one of the plates.

### 2.2.2 Data Acquisition and Triggers

At the LHC, approximately one billion proton-proton collisions occur each second and each hard scattering event takes on average one megabyte of storage, generating one petabyte of data every second. With contemporary particle detectors having data channels numbering in the millions, the volume of information is impossible to preserve. However, the vast majority of collision events are produced by common and well understood physics mechanisms rather than the rare processes which motivated the LHC's physics program. This observation informs the design of *trigger* systems at the LHC experiments which are used to reduce the event rate by filtering for interesting signatures.

At the trigger level, low latencies are required to accomodate the 25 ns bunch crossing rate, which results in the overlap of particles from the current collision with those from the previous and next collisions. In order to maintain the trigger rates at acceptable levels, a hierarchical system of triggers is employed by the CMS experiment to filter events at the hardware and software levels. By implementing robust algorithms and taking advantage of data parallelism, the triggers are optimized and run extremely quickly. A full review of the data acquisition and trigger systems of the CMS experiment may be found in Ref. [43].

The *Level 1 Trigger* (L1T) is the first stage of the trigger system used by the CMS experiment to filter events and is implemented in hardware processors with a fixed latency of 4  $\mu s$ . It consists of custom electronics, such as field-programmable gate arrays (FPGAs) and application-specific integrated circuits (ASICs), that are integrated with the calorimetry and muon subsystems. The decisions made by the L1T are based on the presence of *trigger primitives* proposed by the detector subsystems, representing candidate particles such as muons or jets, which are processed individually and then evaluated as a

whole by the global trigger (GT) for a final decision to accept or reject the event based on  $p_T$  and  $E_T$  thresholds. If an event is accepted, the buffered information from the detector subsystems are read for offline storage. The L1T reduces the event rate from 40 MHz to a maximum of 100 kHz. An overview of the L1T is show in Figure 2-12.

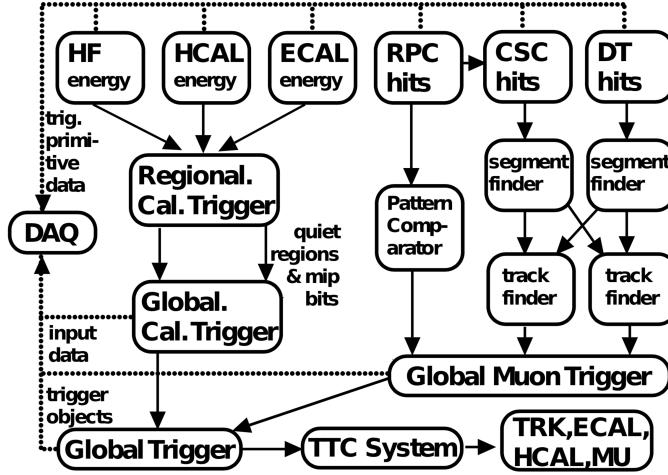


Figure 2-12. An overview of the Level 1 Trigger System used by the CMS detector. The detector subsystems send trigger primitives to the DAQ and regional triggers. Global triggers for the calorimeter and muon subsystems then send data to the global trigger for the final decision to reject an event or accept it and release information buffered in the DAQ systems.[43]

The High Level Trigger (HLT), the second and last stage of the trigger system, is implemented in software that is executed on a computing cluster consisting of 26,000 cores. Because the algorithms use the full precision information recorded by the detector subsystems, the HLT has an average processing time of 260 ms per event.[44] There are several hundred algorithms, or *HLT paths*, each designed for a particular physics goal and defining a sequence of modules that reconstruct objects of interest and performs selections based on their properties. If any modules reject the event, the remainder of the HLT path is short-circuited and the event is rejected. If an event is accepted, it is sent to the CMS Tier-0 computing center for storage and full event reconstruction. The HLT reduces the event rate from 100 kHz at the L1T level to around 100 Hz.

### 2.2.3 Computing Resources

In order to handle the computational requirements of the CMS experiment for tasks ranging from event reconstruction and storage to end-user data analysis, a distributed grid computing model is adopted. This computing infrastructure, known as the Worldwide LHC Computing Grid (WLCG), is a tiered system. The Tier-0 data center at CERN handles the reconstruction and long term storage of events. The data is transferred upon request to national laboratories and universities which host Tier-1 data centers located across Western Europe and in the United States and Taiwan and the smaller Tier-2 data centers used primarily for data analysis. A thorough review of the computing strategy is available in Ref. [45].

## CHAPTER 3

### PHYSICS OBJECT RECONSTRUCTION

In the early days of particle physics experimentation, charged particles were visually identified by analyzing photographs of the ionization tracks left behind in cloud chambers and bubble chambers. Given the higher collision energies and instantaneous luminosities demanded by modern experiments, the amount of information recorded for a collision event renders such visual analyses intractable. The final state particles produced by a proton collision at the LHC are recorded as electronic signals by the CMS detector, and the accurate interpretation of these signals as physics objects is what enables the full reconstruction of the collision’s aftermath. The definition and reconstruction of the standard physics objects, with an emphasis on those used by the  $VH(b\bar{b})$  analysis, are described in this chapter.

#### 3.1 Particle Flow Algorithm

The hermetic design of the CMS detector and the granularity of its subsystems enabled the first successful deployment of a *particle flow (PF)* based reconstruction algorithm at a hadron collider experiment.[46] Although the individual subsystems are capable of reconstructing the particles for which they were designed, a more accurate and global event description can be achieved by combining the measurements obtained by the subsystems as a whole. Since its commissioning, the PF algorithm has been used online to improve the efficiency of the High-Level Trigger (HLT) and offline to improve the quality of the reconstructed particle candidates considered by physics analyses.

The PF algorithm starts by collecting the detector-level objects, or *elements*, produced by each detector subsystem: the silicon tracker and muon chambers both provide charged particle *tracks*, while the electromagnetic calorimeter (ECAL) and hadronic calorimeter (HCAL) both provide *clusters* of absorbed energy. A *link algorithm* which tests the compatibility of pairs of elements from different subsystems is used to generate *blocks* of elements that are directly linked or indirectly linked through common

elements. Individual particles are subsequently identified and reconstructed within each block, starting with muons then proceeding to electrons, photons, and charged and neutral hadrons. As particles are reconstructed, the elements associated with that particle are removed from the block such that each particle is reconstructed from a set of unique elements. Once all the blocks have been processed and all particles in the event have been identified and reconstructed, a post-processing algorithm is used to correct misidentified or misreconstructed muons with large transverse momentum  $p_T$  which can artificially increase the reconstructed missing transverse momentum  $p_T^{\text{miss}}$  in the event.

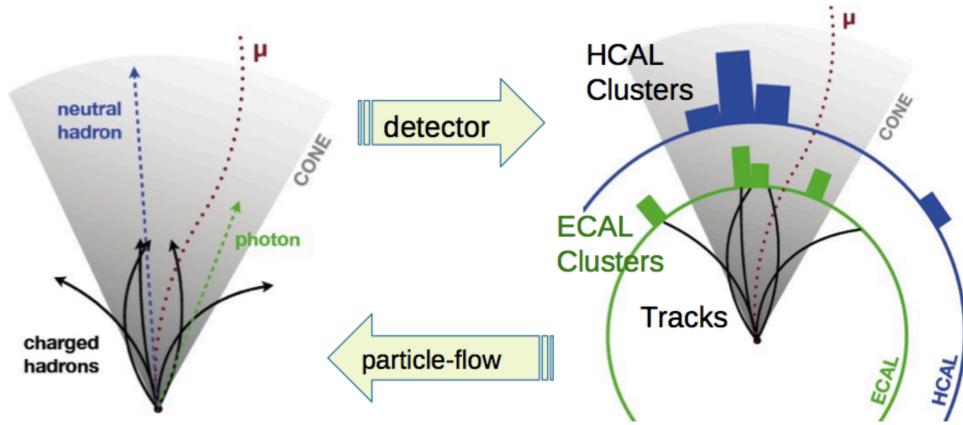


Figure 3-1. An infographic illustrating the particle flow reconstruction paradigm.[47]

At this stage, the particle candidates proposed by the PF algorithm are ready to be used in physics analyses. In practice, the particle candidates are processed further by passing them to algorithms which employ different clustering strategies to reconstruct jets. Finally, The candidate particles and jets which satisfy the additional criteria recommended by the various physics object groups (POGs) within the CMS collaboration become the standard physics objects considered by the physics analyses.

### 3.1.1 Detector-level Objects

Each detector subsystem produces objects which become the fundamental inputs to the PF algorithm. The two main types of detector-level objects are charged particle tracks, with a distinction between tracks from the silicon tracker and standalone muon

tracks from the muon chambers, and calorimeter clusters. By considering these objects together, the full information gathered by the detector may be brought to bear for the task of identifying and reconstructing particles.

### 3.1.1.1 Tracks

The silicon tracker, the inner-most detector subsystem of the CMS detector, reconstructs tracks from *hits*, which are clusters of signals gathered by the pixel and strip trackers. This task quickly becomes a combinatorial challenge because of its dense operating environment due to *pileup*. The bunched nature of the LHC’s proton beams results in multiple proton-proton collisions each time they cross the interaction point as illustrated in Figure 3-2, and those collisions which overlap the one of interest are deemed pileup interactions. Although the number of hits in the tracker increases linearly with pileup, the number of possible combinations of hits to form tracks grows exponentially.

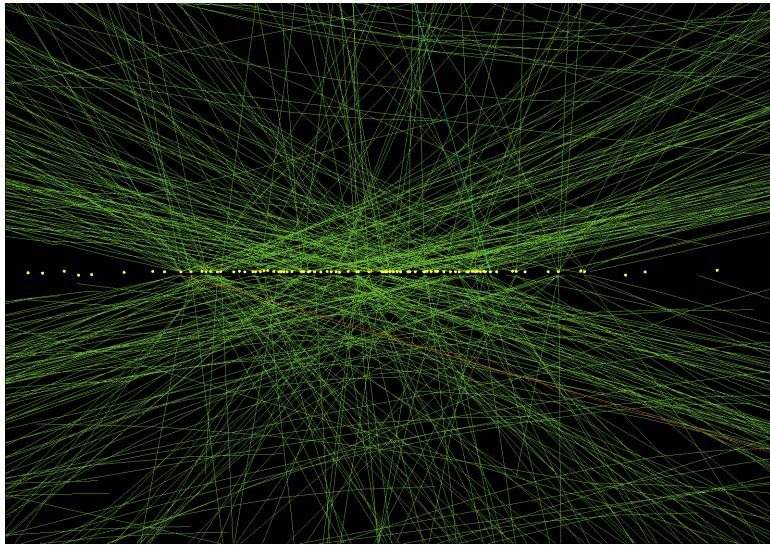


Figure 3-2. An event display for a collision from the high-pileup run 198609 which shows 78 reconstructed vertices and their associated charged particle tracks.[48]

In order to accurately and quickly reconstruct tracks, an iterative tracking algorithm[49] is employed by the CMS experiment which applies several iterations of a combinatorial track finder (CTF) based on Kalman filtering[50] to a collection of hits. Each iteration of the CTF proceeds through the following stages:

- **Seed Generation:** Initial trajectory candidates composed of three hits or two hits and a beamspot or vertex constraint are proposed as seeds.
- **Trajectory Building:** The seeds are extrapolated towards compatible hits in subsequent layers and also towards a single “invalid” or fake hit to account for the case where the corresponding hit was not recorded. A Kalman filter then updates the trajectory based on the compatible hit to form track candidates. This extrapolation procedure terminates upon reaching the outermost layer of the tracker or a when “stop condition” is satisfied, such as surpassing a maximum number of invalid hits. Because a single seed may produce multiple track candidates and different seeds may produce the same track candidate, an ambiguity resolution based on the fraction of hits shared between pairs of track candidates is applied to prevent double counting.
- **Track Fitting:** The collection of hits for each track candidate is refitted by a Kalman filter and smoother to obtain optimal estimates of the track parameters illustrated in Figure 3-3: the signed transverse curvature  $qp_T$ , the polar angle  $\cot \theta_0$ , the azimuthal angle  $\phi_0$ , the longitudinal impact parameter  $d_z$ , and the signed transverse impact parameter  $d_0$ , all of which are defined at the point of closest approach to the beam axis.

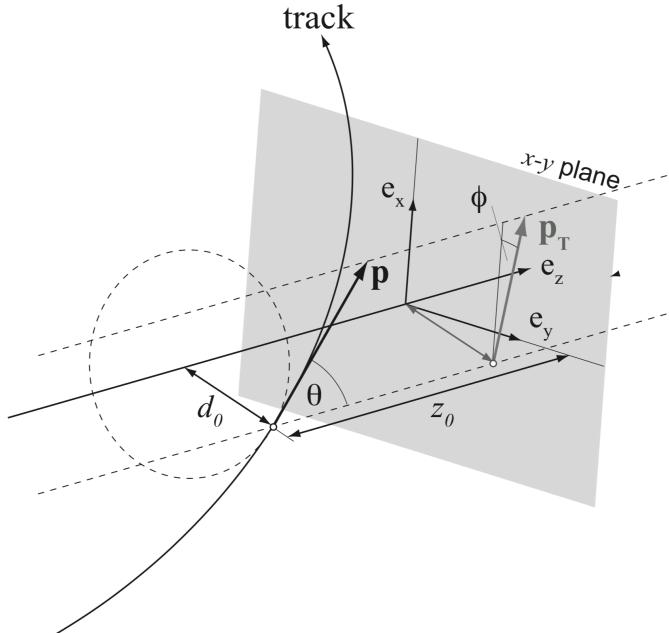


Figure 3-3. The parameterization of a charged particle’s trajectory through a magnetic field.[51]

The CTF is run for a total of ten iterations. The first iteration seeks to reconstruct the easiest tracks which originate close to the interaction point, also called *prompt*

tracks, and have high  $p_T$ . The last two iterations attempt to reconstruct muons and involves muon tracks which are described in further detail in section 3.1.1.3. Because successive iterations face increasingly difficult tasks, the hits associated with selected track candidates are masked in subsequent iterations to reduce the combinatorial complexity. The seeds and reconstruction targets of the ten iterations are summarized by Table 3-1.

Table 3-1. A summary of the ten tracking iterations used by the CMS tracker. The variable  $R$  in the final column denotes the distance from the beam axis to the production position of the targeted track.[46]

Iteration	Seed Type	Reconstructed Tracks
1	pixel triplets	prompt, high $p_T$
2	pixel triplets	from b hadron decays, $R \lesssim 5$ cm
3	pixel triplets	prompt, low $p_T$
4	pixel pairs	recover high $p_T$
5	pixel+strip triplets	displaced, $R \lesssim 7$ cm
6	strip triplets/pairs	very displaced, $R \lesssim 25$ cm
7	strip triplets/pairs	very displaced, $R \lesssim 60$ cm
8	pixel+strip pairs	inside high $p_T$ jets
9	muon-tagged tracks	muons
10	muon detectors	muons

The tracking efficiency and misreconstruction rate are shown in Figure 3-4. The full iterative tracking method achieves a 90% reconstruction efficiency for tracks with  $p_T$  between 1 GeV and 10 GeV, to be compared with the 70-80% efficiency of the global CTF, and maintains an approximately 8-10% higher efficiency than the global CTF for tracks with  $p_T > 100$  GeV. Although the iterative tracking method incurs a slightly higher misreconstruction rate than the global CTF, it is able to reconstruct over half of the tracks missed by the global CTF while running twice as fast.

### 3.1.1.2 Calorimeter clusters

The ECAL and HCAL, which are the intermediate detector subsystems of the CMS detector, reconstruct energy clusters left behind by traversing particles. The electromagnetic interactions of particles with the ECAL materials induce electromagnetic showers which deposit their energy in lead tungstate crystals. The strong interactions of particles with the HCAL materials induce hadronic showers and jets which deposit their

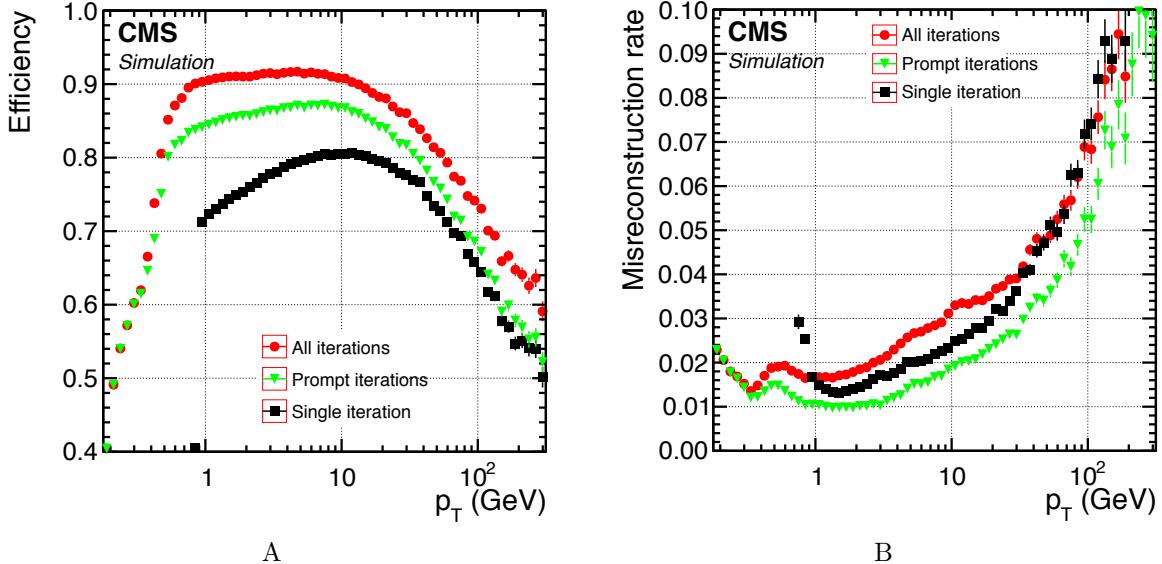


Figure 3-4. The A) tracking efficiency and B) misreconstruction rate of charged hadron tracks in simulated multijet events for a single, global iteration of the combinatorial track finder (black squares), the prompt iterations (iterations 1, 2, 3, 4, 5, and 7) of the iterative tracking method (green triangles), and the full iterative tracking method (red circles) as a function of track  $p_T$ .[46]

energies in alternating layers of brass or steel absorbing plates and plastic scintillator. The electronic signals produced by these showers are measured by the cells of their respective calorimeter.

The ECAL clusters are seeded by cells which contain a local energy maxima above a detector-dependent threshold. The *Island algorithm* is used in both the ECAL barrel and endcaps to cluster the eight cells directly adjacent to the seed in the  $\eta\phi$ -plane with measured energies above twice the noise level, while the *hybrid algorithm* is used in the ECAL barrel only and forms bars that are 3 or 5 crystals wide in *eta* before searching along the *phi*-direction. Both of these algorithms propose *topological clusters* which are themselves clustered into *superclusters*. The most energetic topological cluster is used as the supercluster seed, and nearby clusters are gathered into a supercluster with some spread in the *phi*-direction to recover the energy from photon conversions and bremsstrahlung which are spread by the solenoid's magnetic field. Although 94% of the energy of a photon or electron is contained by a  $3 \times 3$  cluster of crystals, and up to 97%

for  $5 \times 5$  clusters, detector materials before the ECAL result in energy losses which are accounted for by applying corrections to the supercluster energies. Clusters are similarly constructed using information from the HCAL, although the algorithms proceed by considering only those four cells which share an edge with the seed cell. A visualization of ECAL and HCAL clusters are shown in Figure 3-5.

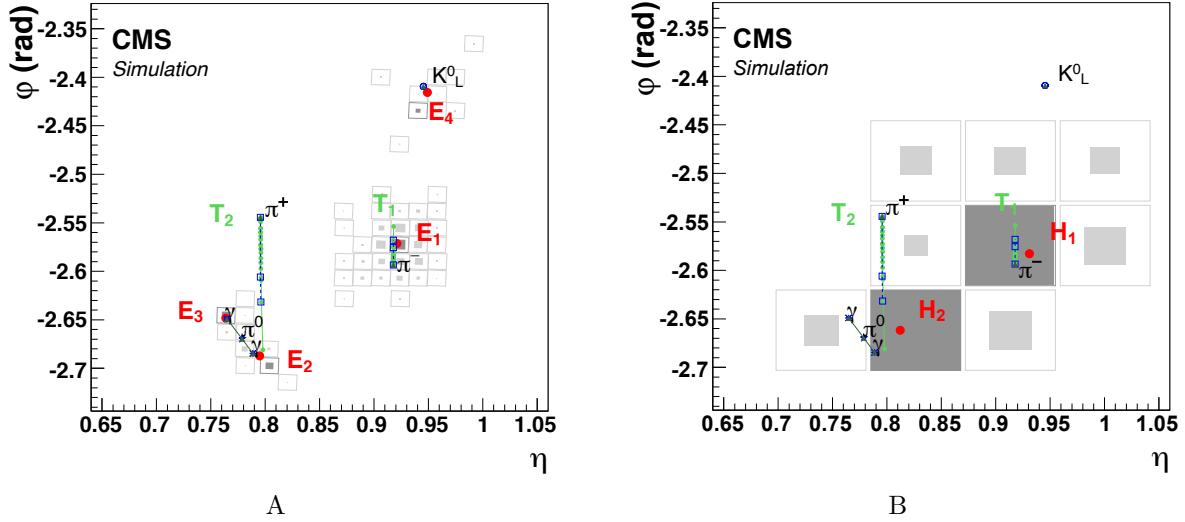


Figure 3-5. An event display of a simulated jet of five particles projected onto A) the ECAL surface and B) the HCAL surface. The green lines represent reconstructed tracks while the red dots represent reconstructed cluster centroids. Four ECAL clusters and two HCAL clusters are shown.[46]

### 3.1.1.3 Standalone muon tracks

The muon chambers also reconstruct tracks from hits and, being the outer-most detector subsystems of the CMS detector, it is expected that these tracks represent muon trajectories. For drift tubes (DTs), hits are reconstructed along anode wires based on the measured arrival time of electrons produced when a traversing muon ionizes the gas in the drift cell. For cathode strip chambers (CSCs), hits are reconstructed from the intersection of clustered signals from cathode strips and anode wire groups. For resistive plate chambers (RPCs), the hits are reconstructed as the centroid of the clustered strip signals. Because the CSCs and DTs have multiple layers, as opposed to the single layer design of the RPCs, their hits are also locally reconstructed as *track segments*.

Groups of DT and CSC track segments from the innermost chambers are used as seeds to generate trajectory candidates. A Kalman filter[52] is used to form track candidates by combining the seed with track segments and hits found in the outer layers. This process is also repeated in reverse, from the outer layers towards the track seed, and the resulting pair is combined into a single track candidate which is then extrapolated to the beam axis and refit with a vertex constraint. This optimal track candidate, based solely on the information obtained by the muon chambers, is what is referred to as a *standalone muon track*.

### 3.1.2 Particle Candidates

The PF algorithm identifies and reconstructs candidate electrons, photons, muons, and charged and neutral hadrons using the elements contributed by the various detector subsystems. Although not an exhaustive list, these five basic particle types are sufficient for the further reconstruction of more complex objects such as jets or  $\tau$  leptons. As discussed in the following sections, the benefits of this holistic approach towards particle reconstruction are realized as more robust particle candidates and improved estimates of their properties.

#### 3.1.2.1 Electrons

Electron candidates are formed from seed tracks and their associated ECAL clusters which are compatible with the trajectory extrapolated from the track. The emission of photons by electrons as they travel through the silicon tracker poses a difficulty for the tracking algorithm which relies on pattern recognition, so a Gaussian-sum filter (GSF)[53] is used to recover hits for the seed track which may have been missed during the trajectory building step. The direction of the seed track is taken to be the direction of the electron candidate and a combination of the seed track momentum and corrected ECAL cluster energies are used to assign the energy of the electron candidate.

This method of electron reconstruction achieves a high identification efficiency for genuine electrons but also tends to misidentify charged hadrons as electrons. The electron

identification can be improved by training boosted decision tree (BDT) classifiers to discriminate against fake electrons using fourteen electron candidate features.[54] Those candidates with classifier scores above an analysis-dependent threshold are identified as electrons. The electron reconstruction efficiency as measured with a data sample of proton-proton collisions collected at  $\sqrt{s} = 8$  TeV and corresponding to an integrated luminosity of  $19.7 \text{ fb}^{-1}$  is shown in Figure 3-6. The reconstruction efficiency is at least 85% for electrons with  $p_T > 15$  GeV and plateaus at approximately 97% for electrons with  $p_T > 30$  GeV.

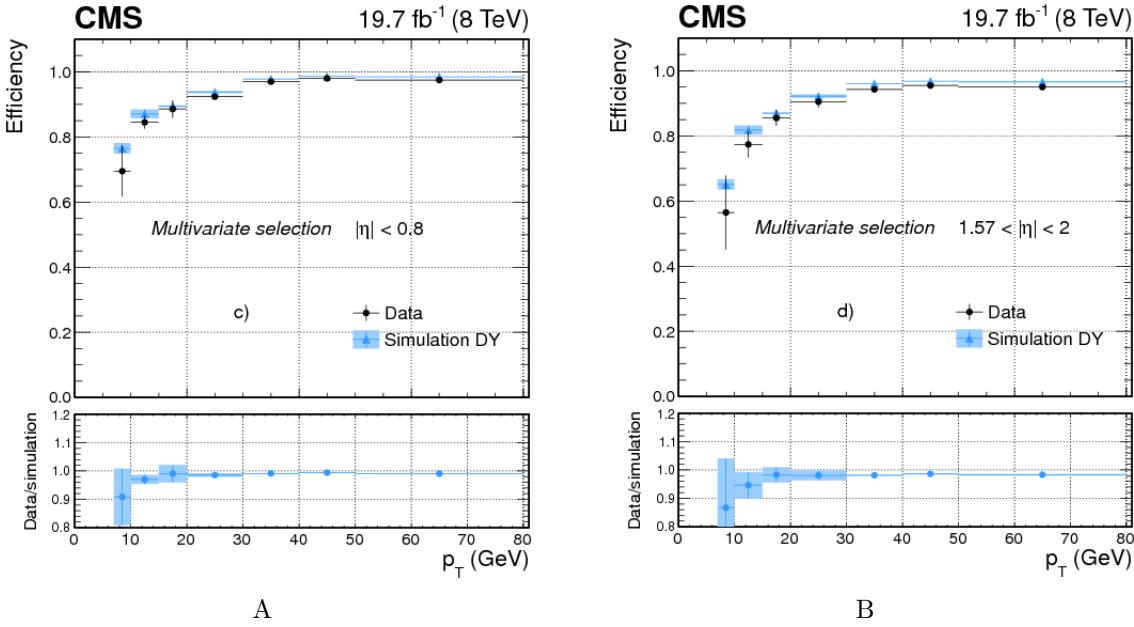


Figure 3-6. The electron reconstruction efficiency as a function of electron  $p_T$  for dielectron data events (black points) and Drell-Yan simulation (blue triangles) in regions of A) low  $\eta$  and B) high  $\eta$ . The medium working point derived from the BDT classifier was used as the multivariate selection.[54]

### 3.1.2.2 Photons

Photon candidates are seeded by ECAL superclusters which are not linked to a track. However, photons travelling through the silicon tracker are likely to interact with the material and convert into  $e^+e^-$  pairs. To ensure that *converted* photons are correctly

reconstructed, pairs of oppositely-charged tracks identified as electrons are associated with the supercluster.

The identification of photons is also improved by training a BDT to discriminate prompt photons from background using features which describe the shape of the shower in the ECAL and aspects of the photon candidates supercluster.[55] Those candidates with classifier scores above an analysis-dependent threshold are identified as photons. The photon reconstruction efficiency measured with a data sample of proton-proton collisions collected at  $\sqrt{s} = 8$  TeV and corresponding to an integrated luminosity of  $19.7 \text{ fb}^{-1}$  is shown in Figure 3-7. For the BDT threshold that achieves 80% signal efficiency and approximately 93% background rejection, the photon reconstruction efficiency typically falls between 80-90% for photons with  $p_T > 20$  GeV.

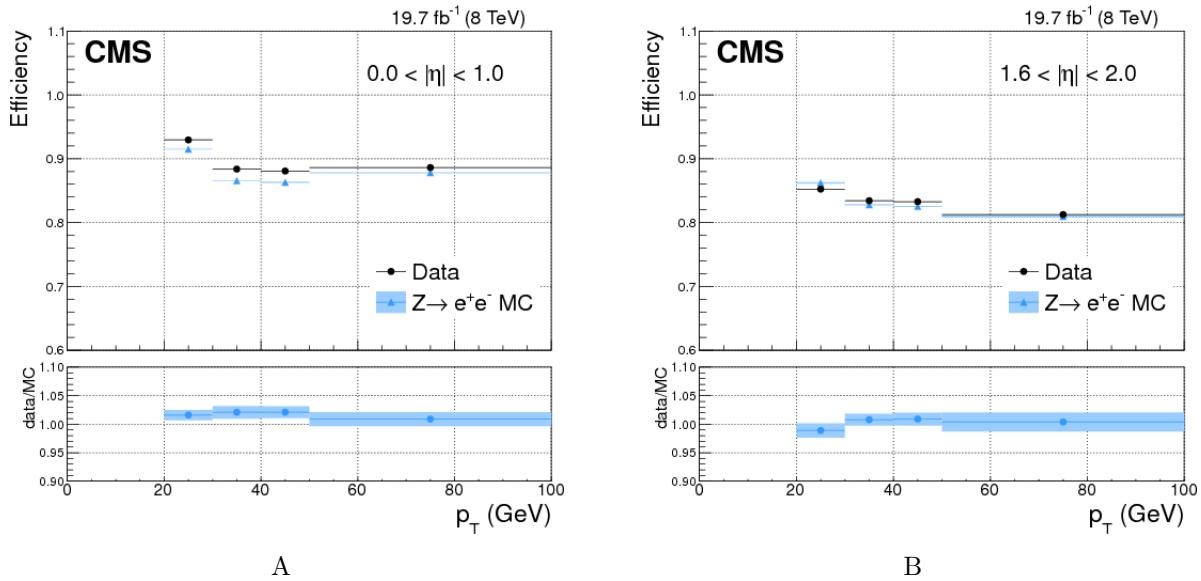


Figure 3-7. The photon reconstruction efficiency as a function of photon  $p_T$  for dielectron data events (black points) and Drell-Yan simulation (blue triangles) in regions of A) low  $\eta$  and B) high  $\eta$ . An example working point for the BDT classifier which achieves 80% signal efficiency was used as the multivariate selection.[55]

### 3.1.2.3 Muons

Muon candidates are seeded by standalone muon tracks, as well as *tracker muon tracks* and *global muon tracks* which are formed by combining information from the

silicon tracker and muon chambers during the last two iterations of the iterative tracking algorithm. The tracker muon tracks are generated by an “inside-out” algorithm which extrapolates tracks from the silicon tracker which pass a momentum threshold to track segments identified in the DT and CSC chambers which are compatible with the hypothesized trajectory. Global muon tracks are similarly generated using an “outside-in” algorithm which extrapolates standalone muon tracks to tracks from the silicon tracker, followed by a combined Kalman filter fit. Those tracker muon tracks and global muon tracks which share silicon tracker tracks are identified as the same muon candidate. If any calorimeter clusters are associated with the muon tracks, they must be compatible with the energy deposition signature of a minimum ionizing particle.

A variety of muon identification algorithms are defined by the CMS Muon POG to serve the needs of different physics analyses. These algorithms combine low-level variables, such as the number of hits per track, and high-level variables, such as the compatibility between the tracks, track segments, and primary vertex, to identify specific types of muons under different conditions. For example, the common *loose muon ID* is designed to efficiently identify tracker muons or global muons as prompt muons which originate from the primary vertex or light and heavy flavor decays.[\[56\]](#) The reconstruction efficiency when using the loose muon ID, which exceeds 99% over the full  $\eta$  range, is measured using a 2015 data sample of proton-proton collisions collected at  $\sqrt{s} = 13$  TeV that corresponds to an integrated luminosity of approximately  $2 \text{ fb}^{-1}$  and is shown in Figure 3-8.

### 3.1.2.4 Charged and neutral hadrons

Charged hadron candidates are seeded by HCAL clusters which are linked to at least one track and possibly ECAL clusters. Neutral hadron candidates within the tracker’s acceptance, or  $|\eta| < 2.5$ , are instead seeded by HCAL clusters which are not linked to a track while candidates outside of the tracker’s acceptance are seeded by linked ECAL and HCAL clusters which present an energy excesses inconsistent with the signature of a charged hadron’s energy deposition. Because such hadrons are created by the

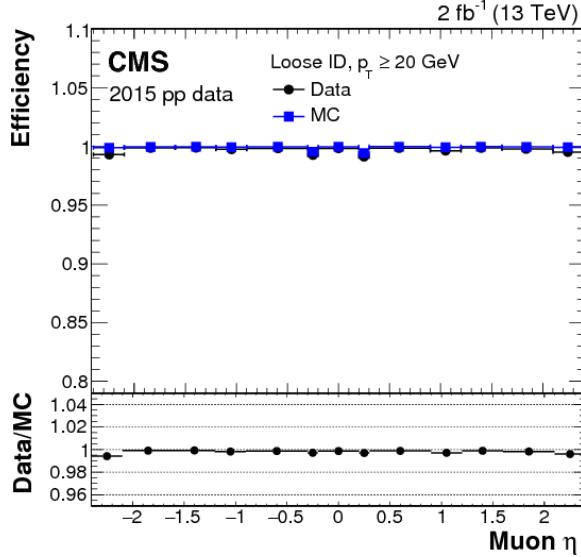


Figure 3-8. The muon reconstruction efficiency as a function of  $\eta$  for 2015 data events with a pair of reconstructed muons (black points) and Drell-Yan simulation (blue squares). The muon candidates were required to have  $p_T > 20$  GeV and pass the loose muon ID.[56]

hadronization of final state particles, their performance is better discussed in the context of jet reconstruction and is reserved for section 3.5.

### 3.2 Primary Vertices

Although multiple proton-proton interactions occur when the beams collide, only those interactions which cause the event to trigger are of interest. The position where such an interaction occurs is known as a *primary vertex* and its reconstruction is vital because it is the point of origin of the particles initially produced by the hard-scattering process.

The reconstruction of primary vertices begins with the clustering of tracks using a deterministic annealing algorithm.[49] The tracks of each primary vertex candidate are then passed to an adaptive vertex fit to improve the estimate of the spatial position of the vertex.[57] The fit also assigns weights to each track in the cluster based on their compatibility with the primary vertex position, the sum of which is proportional to the number of degrees of freedom of the fit. Based on their fit results, only those candidates which have a  $z$  position within 24 cm of the nominal detector center, a radial distance

within 2 cm of the beamspot axis, and a vertex fit exceeding four degrees of freedom are considered primary vertices.

### 3.3 Tau Leptons

Besides the electron and muon candidates proposed by the PF algorithm, a third lepton remains to be reconstructed: the tau lepton. With a mass of  $m_\tau \approx 1.78$  GeV[11], the tau is the only lepton able to decay hadronically through weak interactions, which accounts for almost two-thirds of its decays. The remaining decay channels are purely leptonic, with the tau decaying into either an electron or muon. The hadrons-plus-strips (HPS) algorithm[58, 59] is used to reconstruct the hadronically decaying taus by analyzing the constituents within reconstructed jets. The photons and electrons within the seed jets are clustered together into *strips* because neutral pions, being the most common hadron produced by tau decays, predominantly decay into photon pairs. Based on the number of strips and charged particles found within the jet, a tau candidate is assigned to one of three possible decay topologies:

- **One prong:** a single charged hadron and no strips.
- **Two prong:** a single charged hadron and one strip.
- **Three prong:** a single charged hadron and two strips, or three charged hadrons.

Finally, to reduce the misidentification rate of jets as taus, a BDT is trained to classify taus using jet features, such as the impact parameter of the leading track and the flight distance, as well as jet constituent features, such as the multiplicity of electrons and photons within the jet.

### 3.4 Lepton Isolation

Electrons and muons can either be prompt, produced directly by electroweak decays of massive particles suchs as the  $W$  or  $Z$  bosons, or non-prompt, produced by the subsequent decays of taus or jets. Lepton *isolation*, a measure of the lack of activity around the charged particle track, is used to reconstruct prompt electrons and muons while rejecting leptons which are jet constituents. The isolation is defined as the ratio of

the total transverse momentum of the particles within a cone around the lepton to the transverse momentum of the lepton itself

$$I_{\text{PF}} = \frac{1}{p_T} \left( \sum_{\gamma} p_T^{\gamma} + \sum_{h^\pm} p_T^{h^\pm} + \sum_{h^0} p_T^{h^0} \right), \quad (3-1)$$

where the superscripts indicate photons ( $\gamma$ ), charged hadrons ( $h^\pm$ ), and neutral hadrons ( $h^0$ ) and the lack of a superscript denotes the lepton of interest. The identification of prompt leptons proceeds by placing thresholds directly on the lepton isolation, with lower thresholds indicating a tighter criterion, or passing the lepton isolation as one of several inputs to a classifier.

### 3.5 Jets

The experimental signature of quarks and gluons produced by a hard scattering process are collimated sprays of particles known as *jets*. Due to color confinement, quarks and gluons are not observed directly but through the hadrons they immediately form, which are themselves unstable and inferred from the decay of their partons. This process of hadronization and fragmentation continues until the hadrons are absorbed or the remaining decay products are stable particles. Because jet constituents leave signatures all throughout the detector volume, the holistic approach of the PF algorithm is well-suited for jet reconstruction.

#### 3.5.1 Reconstruction

The primary concerns for jet reconstruction are the identification and association of a jet's constituent particles. The identification of possible constituent particles is undertaken by the PF algorithm, where prompt and isolated electrons, muons, and photons are removed from consideration. The association of constituents to form a jet is handled by specialized jet finding algorithms.

Jet finding algorithms broadly fall under two categories, cone algorithms and sequential recombination algorithms. Cone algorithms use a “top-down” approach based on the assumption that jets can be contained within conical envelopea. Cones of

radius  $R$  are first seeded along regions of energy flow and stable cones, which have axes compatible with the direction of the vectorial sum of the momenta of particles within the cone, are kept. Overlapping stable cones are further split or merged under the algorithm converges. Sequential recombination algorithms instead use a “bottom-up” approach based on the assumption that jet constituents form well-defined clusters due to their collimation. Distance metrics in coordinate space and in transverse momentum space are defined and calculated between pairs of constituent candidates, with the minimum distance pair of either space used to seed the jet cluster. The distances are then recalculated and candidate particles combined with the jet cluster iteratively until reaching a stop condition.

Because the theoretically motivated observable quantities describing parton decays are manifested experimentally as the properties of reconstructed jets, the implementation of jet finding algorithms must be infrared and collinear (IRC) safe. Jet finding algorithms are IRC unsafe if they are sensitive to soft emissions, such as gluon radiation, and collinear splittings of particles as illustrated in Figure 3-9. By satisfying the requirement of IRC safety, the jets reconstructed by the algorithm may be considered valid approximations of the final state partons.

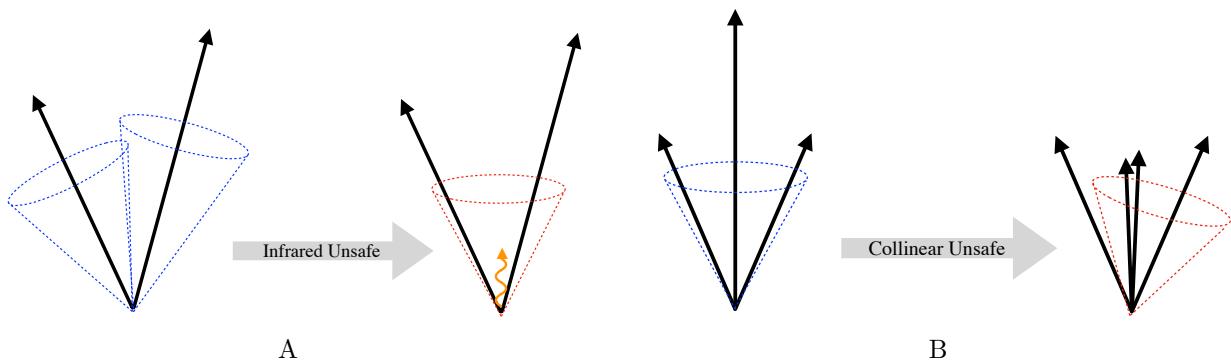


Figure 3-9. An illustration of the consequences of infrared and collinear unsafety for jet finding: A) an example of infrared unsafety, where a soft emission (orange) causes the algorithm to merge two overlapping but otherwise distinct jets; B) an example of collinear unsafety, where the collinear splitting of a constituent alters the jet returned by the algorithm.

Although cone algorithms were previously favored for their ease of implementation, simple jet geometries, and faster execution compared to sequential recombination algorithms, the vast majority of cone algorithms are not IRC safe. With the release of the **FastJet**[60] software package, efficient implementations of sequential recombination algorithms with  $\mathcal{O}(n \log n)$  time complexity were readily adopted at the LHC. The CMS experiment uses the *anti- $k_T$*  jet clustering algorithm[61] to reconstruct jets from PF candidates. The *anti- $k_T$*  algorithm defines the distance in transverse momentum space as

$$d_{ij} = \min \left( \frac{1}{p_{Ti}^2}, \frac{1}{p_{Tj}^2} \right) \times \frac{R_{ij}^2}{R}, \quad (3-2)$$

where  $i$  and  $j$  are the indices of the PF candidates,  $R$  is the radius parameter which governs the size of the reconstructed jets, and  $R_{ij}^2$  is the Euclidean distance between the PF candidates in the  $\eta\phi$ -plane

$$R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2. \quad (3-3)$$

The *anti- $k_T$*  algorithm also defines an additional distance between the beam axis and the  $i$ -th particle to be

$$d_{iB} = \frac{1}{p_{Ti}^2}, \quad (3-4)$$

which is used to classify a finished cluster as a jet.

The algorithm proceeds by finding the minimum over the set of distances  $\{d_{ij}, d_{iB}\}$ . If the minimum distance is between a pair of particles  $i$  and  $j$ , their four-vectors are summed to form a new particle and particles  $i$  and  $j$  are removed from the set of particles considered by the algorithm. If the minimum distance is between a particle and the beam axis, the particle is considered a jet and removed from the set of particles considered by the algorithm. The algorithm continues until all initial particles have been combined into jets. Because the distance measure given by Equation 3-2 is dominated by high  $p_T$  the *anti- $k_T$*  algorithm tends to cluster soft particles with the hardest particles first, producing perfectly conical jets.

### 3.5.2 Energy Scale and Resolution Corrections

Disagreements between the measured energies of reconstructed jets and the true energies of their corresponding partons are expected. The jet's constituent particles encounter unavoidable energy losses as they travel through the detector material and their measured energies are biased by the non-linear response of the detector subsystems to their signals. Because the four-momentum of a jet is taken to be the vectorial sum of the momenta of its clustered PF candidates, these issues and more are compounded together and require a dedicated calibration.

In order to calibrate jet energies, the recommended procedure by the CMS experiment applies a series of *jet energy scale* (JES) corrections which scale the four-momentum of jets. The various detector effects are addressed using a factorized approach shown in Figure 3-10, where each of the corrections are designed to target specific effects and are applied in a fixed-order sequence. The first correction aims to remove energy contributions due to pileup particles and also spurious detector noise, and is parameterized as a function of the energy density  $\rho$ , jet area  $A$ , and the jet  $p_T$  and  $\eta$ . The second correction addresses the effects of the calorimeter response, which is non-uniform in  $\eta$  and non-linear in  $p_T$ . The next two corrections address percent level residual differences observed between the truth level simulation and data. A final, optional correction can also be applied to correct the flavor, or type of parton, which initiated the jet.

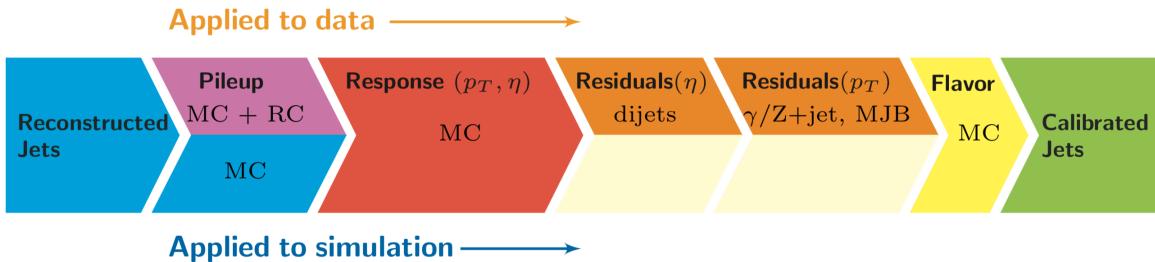


Figure 3-10. An infographic of the jet energy scale corrections used by the CMS experiment for data and Monte-Carlo simulated samples. Corrections derived from simulation are marked with MC, while RC stands for random cone and MJB stands for multijet background.[62]

After the jet energies are calibrated, the jet energy resolution (JER) in Monte-Carlo (MC) simulation is observed to be better than for data. Because the effects of detector acceptance and performance are not completely modelled in simulation, a “smearing” procedure is adopted which modifies the four-momenta of reconstructed jets in order to match the energy resolution observed in data. The smearing can be performed using the scaling method which scales the four-momentum with a correction of the form

$$c_{\text{JER}} = 1 + (s_{\text{JER}} - 1) \frac{p_{\text{T}} - p_{\text{T}}^{\text{ptcl}}}{p_{\text{T}}}, \quad (3-5)$$

where  $s_{\text{JER}}$  is the simulation-to-data resolution scale factor,  $p_{\text{T}}$  is the transverse momentum of the reconstructed jet, and  $p_{\text{T}}^{\text{ptcl}}$  is the transverse momentum of the corresponding jet but clustered from generator-level particles. The scaling method assumes that the particle-level jet matches well the generator-level jet, otherwise it introduces a large shift in the response. The alternative method of stochastic smearing relaxes this assumption, with a randomly sampled correction of the form

$$c_{\text{JER}} = 1 + \mathcal{N}(0, \sigma_{\text{JER}}) \sqrt{\max(s_{\text{JER}}^2 - 1, 0)}, \quad (3-6)$$

where  $\mathcal{N}(0, \sigma_{\text{JER}})$  denotes a normal distribution with zero mean and variance determined by the relative  $p_{\text{T}}$  resolution in simulation  $\sigma_{\text{JER}}$  and  $s_{\text{JER}}$  is the simulation-to-data resolution scale factor. For either method, negative correction factors are set to zero. Typically, a hybrid approach is recommended where the stochastic smearing is applied by default and the scaling smearing is applied if a jet has a generator-level match.

In physics analyses, both the JES corrections and JER smearings are applied to the reconstructed jets and systematic uncertainties are assessed for the different systematic sources based on their  $1\sigma$  shifts that is fully correlated in jet  $p_{\text{T}}$  and  $\eta$ . This amounts to at most a 3% uncertainty in the phase space considered by most analyses. A thorough review of the derivations of the JES corrections, JER smearing, and their uncertainties and are available in Ref. [62].

### 3.5.3 *b*-Tagging

The long lifetime of *b*-hadrons, quark bound states containing a bottom quark, can be attributed to the mass of the *b*-quark, which is heavier than the common quarks such as the up or down quarks, and the suppression of its decays to lighter quarks by the Cabibbo-Kobayashi-Maskawa (CKM) matrix. Because of this longer lifetime, *b*-hadrons travel a distance on the order a few millimeters before decaying, resulting in a characteristic displaced vertex relative to the primary vertex of the hard scattering process, as shown in Figure 3-11. This property enables the jets formed by the hadronization of *b*-quarks to be distinguished from those of other quarks by using *b*-tagging methods which determine a jet's quark flavor.

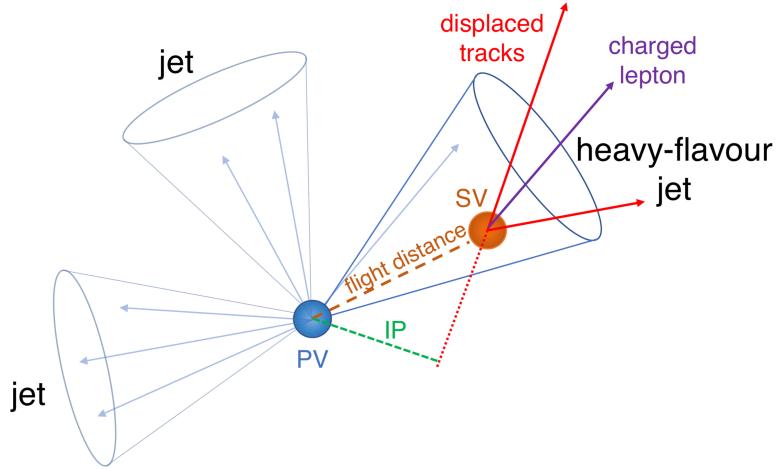


Figure 3-11. An illustration of a heavy-flavor (*c* or *b*-hadron) jet emphasizing its characteristic property of a secondary vertex (SV) that is displaced from the primary vertex (PV).[63]

The latest *b*-tagging algorithm developed and used by the CMS experiment is the Deep Combined Secondary Vertex (DeepCSV)[63] tagger. The DeepCSV tagger is a classifier based on a fully-connected neural network architecture with four hidden layers, each with 100 rectified linear (ReLU) units, implemented using the Keras[64] neural network framework with Tensorflow[65] as the graph compilation and automatic differentiation backend. The final layer of the neural network is composed of five linear units passed through a softmax activation function. A softmax output layer was chosen

because the output values are interpreted to be the probability that a jet belongs to one of five possible classes based on its particle content: exactly one  $b$ -hadron, at least two  $b$ -hadrons, exactly one  $c$ -hadron and no  $b$ -hadrons, at least two  $c$ -hadrons and no  $b$ -hadrons, and no  $c$  or  $b$ -hadrons.

The DeepCSV tagger is trained using information from the secondary vertex candidates proposed by the inclusive vertex finding (IVF) algorithm, the information of the tracks associated with the secondary vertices such as their impact parameter, as well as the jet  $p_T$  and  $\eta$ . The combination of low-level and high-level variables enables the DeepCSV tagger to not only perform better than previous approaches in terms of  $b$ -jet identification efficiency but also crucially reduces the misidentification probability of light jets as  $b$ -jets as shown in Figure 3-12. For physics analyses, three standard working points are defined for use which trade off  $b$ -jet identification efficiency ( $\varepsilon_b$ ) for misidentification probability ( $\varepsilon_{uds_g}$ ): the loose (L) working point achieves  $\varepsilon_b = 84\%$  and  $\varepsilon_{uds_g} = 11\%$ , the medium (M) working point achieves  $\varepsilon_b = 68\%$  and  $\varepsilon_{uds_g} = 1.1\%$ , and the tight (T) working point achieves  $\varepsilon_b = 50\%$  and  $\varepsilon_{uds_g} = 0.1\%$ .

Discrepancies are observed in the distributions of some of the input variables to the DeepCSV tagger between data and simulation which arise due to imperfect modeling of the detector and the modelling of parton showering and hadronization by the MC generators. These discrepancies translate to differences in  $b$ -tagging efficiency between data and simulation. In order to correct for these effects, a scale factor is applied on a per jet basis as a function of jet  $p_T$  and  $\eta$  defined by

$$SF_f = \frac{\varepsilon_f^{\text{data}}(p_T, \eta)}{\varepsilon_f^{\text{MC}}(p_T, \eta)}, \quad (3-7)$$

where  $f$  is the jet flavor,  $\varepsilon_f^{\text{data}}(p_T, \eta)$  is the flavor tagging efficiency in data, and  $\varepsilon_f^{\text{mc}}(p_T, \eta)$  is the flavor tagging efficiency in MC. One method of determining the tagging efficiency in data is by applying the tag-and-probe (TnP) technique to data events passing a selection for single-lepton  $t\bar{t}$  decays. Because top quarks decay into a  $W$  and a bottom quark nearly

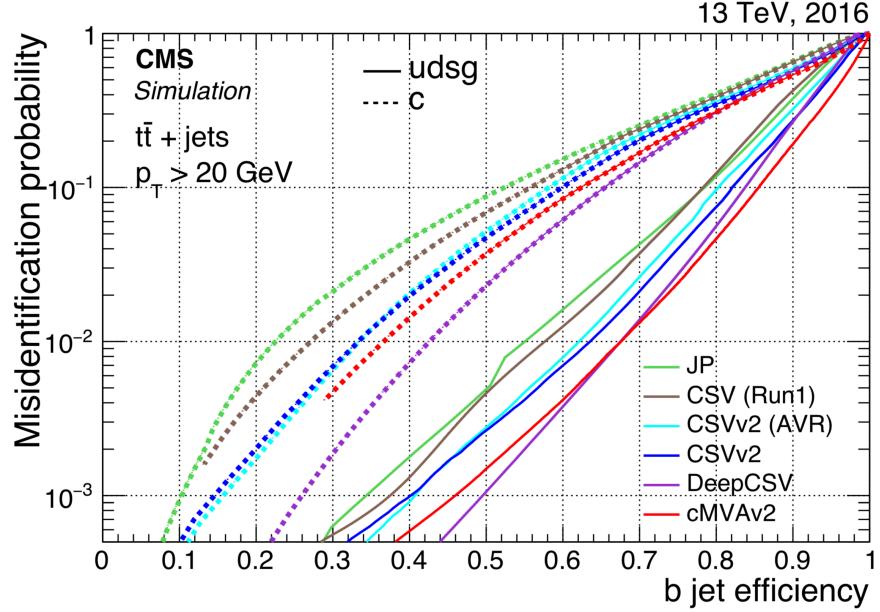


Figure 3-12. The receiver operating characteristic (ROC) curves which characterize the performance of various  $b$ -tagging algorithms developed by the CMS experient and evaluated using a  $t\bar{t} + \text{jets}$  Monte-Carlo simulated sample. The DeepCSV algorithm's  $b$ -tagging performance against light flavored jets is given by the solid purple line.[63]

96% of the time and the  $W$  boson decays hadronically around 67% of the time, such a selection of events where one of the  $W$  bosons decays leptonically and the other hadronically offers at least two  $b$ -jets and multiple light jets. If a data event contains a  $b$ -jet which satisfies a tagging requirement, the remaining “probe” jets are used to provide an unbiased estimate of the  $b$ -tagging efficiency from the fraction of probe jets passing the tagging requirement with respect to all probe jets according to the definition

$$\varepsilon_b = \frac{N_b^{\text{tagged}}}{N_b^{\text{vetoed}} + N_b^{\text{tagged}}}, \quad (3-8)$$

where  $N_b^{\text{tagged}}$  is the number of probe jets passing the tagging requirements and  $N_b^{\text{vetoed}}$  is the number of probe jets failing the tagging requirements.

The  $b$ -tagging scale factors incur systematic effects through the measurement of the  $b$ -tagging efficiency. These systematic sources can arise from theory and simulation, such as uncertainties for the factorization and renormalization scales, the branching fractions of

$b$ -hadron decays, and gluon splitting, or data, such as uncertainties for pileup and the jet energy scale and resolution. The  $b$ -tagging algorithms and their performance, as well as the determination of their efficiency scale factors and uncertainty, are described in detail in Ref. [63].

### 3.6 Missing Transverse Energy

Although the CMS detector has a near hermetic design, neutrinos and perhaps other theorized particles which are neutral and weakly interacting escape detection. Their presence is indicated by an imbalance in the transverse momentum sum of all “visible” particles as illustrated in Figure 3-13, from which it is inferred that the energy missing from the final state must have been carried away by “invisible” particles to satisfy energy conservation. This imbalance motivates the definition of missing transverse momentum  $\text{vec}\mathbf{p}_\text{T}^\text{miss}$  in an event as

$$\vec{p}_\text{T}^\text{miss} = - \sum \vec{p}_\text{T}, \quad (3-9)$$

or the negative vectorial sum over the transverse momenta of all visible final state particles. The missing transverse energy  $p_\text{T}^\text{miss}$  is taken to be the magnitude of  $\vec{p}_\text{T}^\text{miss}$ . A complete treatment of the reconstruction of  $p_\text{T}^\text{miss}$  including different methods and their performance, may be found in Ref. [66], but the following discussion focuses on the standard approach.

The reconstruction of  $p_\text{T}^\text{miss}$  is challenging because it relies on the accurate reconstruction of all other particles, the precise measurements of their properties, and inherits all of their corrections and the effects of their systematic sources. Given its robust performance, the standard PF algorithm is applied towards the reconstruction of  $p_\text{T}^\text{miss}$ , where the vectorial transverse momentum sum is performed over all PF candidate particles, to produce PF  $p_\text{T}^\text{miss}$ . However, the quantity at this stage is considered “raw”  $p_\text{T}^\text{miss}$  because it is systematically different from true  $p_\text{T}^\text{miss}$  due to non-linear detector responses and misreconstruction of the PF candidates which take on the role of the visible particles. The quantity typically used is *Type-I corrected*  $p_\text{T}^\text{miss}$ , which propagates the jet energy scale

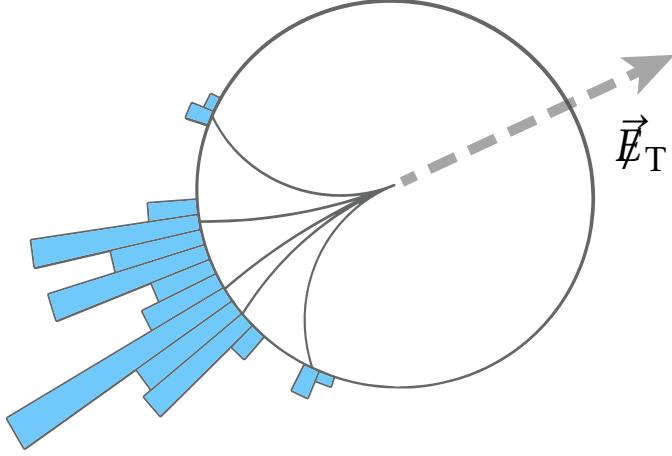


Figure 3-13. An illustration of missing transverse momentum  $p_T^{\text{miss}}$  (gray dashed line), which is inferred from the imbalanced transverse momentum sum of visible particles which were detected by their tracks (black solid lines) and calorimeter energy deposits (blue rectangles). The symbol  $\vec{E}_T$  which used to denote missing transverse momentum has been deprecated.[67]

(JES) corrections to the determination of  $p_T^{\text{miss}}$ . By distinguishing those PF candidates which have been clustered together as jets, we can reformulate Equation 3-9 to naturally incorporate the JES corrections

$$\vec{p}_T^{\text{miss}} = - \sum_{\text{jets}} \vec{p}_T^{\text{JEC}} - \sum_{\text{unclustered}} \vec{p}_T, \quad (3-10)$$

where  $\vec{p}_T^{\text{JEC}}$  denotes the  $\vec{p}_T$  of the corrected jets.

In addition to these corrections, *filters* are designed by the JetMET POG to help physics analyses reject events with artificially large  $p_T^{\text{miss}}$  due to specific noise sources. The following filters were recommended to be applied for 2017 data, as well as MC simulation under 2017 data-taking conditions:

- The **GoodVerticesFilter** rejects events with large  $p_T^{\text{miss}}$  due to noisy vertex reconstruction caused by pileup.
- The **GlobalTightHaloFilter** rejects events with large  $p_T^{\text{miss}}$  due to beam-halo particles, which are the result of beam protons interacting with the instrumentation used to collimate the beam or residual gas particles in the beam pipe vacuum. Information from the CSCs in the muon endcap are used to identify the passage of

halo muons, which have a non-negligible chance of interacting with the calorimeters to produce clusters with energy on the order of several hundred GeV.

- The **HBHENoiseFilter** rejects events with large  $p_T^{\text{miss}}$  due to sporadic HCAL noise from the hybrid photodiodes (HPD) and readout box (RBX) electronics and ion feedback noise that can affect several HPD pixels. These spurious signals can be identified by their high HPD occupancy and pulse shape information from the RBX. An additional **HBHENoiseIsoFilter** performs the same task but adds additional isolation requirements to potential HCAL noise clusters by applying a topological filter that searches for neighboring activity in the HCAL as well as the ECAL and tracker.
- The **EcalDeadCellTriggerPrimitiveFilter** rejects events with large  $p_T^{\text{miss}}$  due to large energy losses from particles travelling through dead ECAL cells which have non-functioning data links.
- The **BadPFMuonFilter** rejects events with large  $p_T^{\text{miss}}$  due to the inclusion of poor quality muon candidates which barely passed the PF algorithm's thresholds and have large mismeasured  $p_T$ .
- The **BadChargedCandidateFilter** rejects events with large  $p_T^{\text{miss}}$  due to the misreconstruction of poor quality muon candidates, where the PF algorithm incorrectly associates them with an energetic HCAL cluster to form a charged hadron candidate.
- The **EcalBadCalibFilter** rejects events with large  $p_T^{\text{miss}}$  due to sporadic high energy noise found in endcap crystals in the very forward region of the ECAL.
- The **eeBadScFilter**, which is applicable for data and not MC simulation, rejects events with large  $p_T^{\text{miss}}$  due to high-amplitude anomalous pulses produced by data channels from the ECAL endcap which are identified by the total energy and number of low-quality hits within an ECAL supercluster.

## CHAPTER 4

### ANALYSIS STRATEGY

The first  $VH(b\bar{b})$  analysis conducted by the CMS collaboration in 2012[68] demonstrated that an observation of the  $H(b\bar{b})$  decay was feasible by combining searches for the  $Z(\nu\bar{\nu})H(b\bar{b})$ ,  $W(e\nu)H(b\bar{b})$ ,  $W(\mu\nu)H(b\bar{b})$ ,  $Z(e\bar{e})H(b\bar{b})$ , and  $Z(\mu\bar{\mu})H(b\bar{b})$  decay channels. The overall strategy that it pioneered has remained largely intact, even through the transition to LHC Run 2. The search for  $VH(b\bar{b})$  using the 2016 dataset would establish evidence for the  $H(b\bar{b})$  decay in 2017[28] with only incremental improvements. The search for  $VH(b\bar{b})$  using the 2017 dataset would build on the physics insights and experiences obtained the year prior while bringing to bear additional event reconstruction techniques and deep learning to finally observe the  $H(b\bar{b})$  decay in 2018.[69] This chapter describes the analysis strategy used by CMS to achieve its state of the art result, with a focus on the  $Z(\nu\bar{\nu})H(b\bar{b})$  decay channel in particular.<sup>1</sup>

#### 4.1 Data and Simulation

##### 4.1.1 Data

The datasets used for this analysis were collected by the CMS detector throughout 2017 and produced by proton-proton collisions with a center-of-mass energy  $\sqrt{s} = 13$  TeV and a bunch spacing of 25 ns. The  $Z(\nu\bar{\nu})H$  channel uses the MET dataset, the  $W(e\nu)H$  channel uses the SingleElectron dataset, the  $W(\mu\nu)H$  channel uses the SingleMuon dataset, the  $Z(e\bar{e})H$  channel uses the DoubleEG dataset, and the  $Z(\mu\bar{\mu})H$  channel uses the DoubleMuon dataset. All of these primary datasets have an approximate total luminosity of  $41.3 \text{ fb}^{-1}$  when only considering *golden* events, which were certified by data quality monitoring teams to have been collected during periods of normal function for all detector subsystems. The JSON certification file

---

<sup>1</sup> This chapter includes content adapted from Ref. [28] and [69], as the author's work directly contributed to those publications.

`Cert_294927-306462_13TeV_EOY2017ReReco_Collisions17_JSON_v1.txt` was used to filter for golden events.

#### 4.1.2 Simulation

Simulated samples for each of the relevant signal and background processes are used to design and validate the analysis strategy in an unbiased manner. These samples are produced using different Monte-Carlo (MC) generators and processed by the GEANT4[70–72] software which realistically models the detector’s response to the passage of particles through its materials. The reconstruction of these MC events was performed using CMS SoftWare (CMSSW) version 94X releases set to match the 2017 data taking conditions.

The signal samples, generated using the POWHEG v2[73–75] event generator, are listed in Table 4-1. The quark-induced signal processes are generated to next-to-leading order (NLO) accuracy in QCD using the MiNLO procedure [76, 77], while the gluon-induced contributions to the  $ZH$  signal processes are generated only to leading order (LO) accuracy in QCD. Their production cross sections and branching fractions are taken to be the state of the art values calculated for a Higgs boson mass of  $m_H = 125$  GeV from Ref. [13], and specifically those listed in sections 1.2.1.1 and 1.2.1.2. These production cross sections are further rescaled to next-to-next-to-leading order (NNLO) accuracy in QCD and NLO electroweak accuracy as a function of the transverse momentum of the vector boson  $p_T(V)$ , according to the combined calculations of the VHNNLO[78–81], VH@NNLO[82, 83], and HAWK v2.0[84] event generators as described in Ref. [13]. The event weights used to rescale the samples are specific to each signal process and an example of the weights applied to the  $W^+(\ell\nu)H \rightarrow b\bar{b}$  channel are shown in Figure 4-1.

The diboson samples listed in Table 4-2 are generated using the MADGRAPH5\_AMC@NLO v2.4.2[85] event generator at NLO accuracy with the FxFx[86] merging scheme and up to two additional partons. The exception is the inclusive  $ZZ$  sample which is only simulated to LO accuracy and generated using PYTHIA v8.230[87] event generator. Although the cross section of the ZZTo2L2Q sample has been obtained from an NLO calculation,

Table 4-1. The Monte-Carlo samples and their cross sections for the signal processes considered by the 2017  $VH(b\bar{b})$  analysis.

Sample	$\sigma(\text{pb})$
WminusH_HToBB_WToLNu_M125_13TeV_powheg_pythia8	$0.5824 \times 0.533 \times 0.108535$
WplusH_HToBB_WToLNu_M125_13TeV_powheg_pythia8	$0.5824 \times 0.840 \times 0.108535$
ZH_HToBB_ZToLL_M125_13TeV_powheg_pythia8	$0.5824 \times (0.8839 - 0.1227) \times 0.10974$
ZH_HToBB_ZToNuNu_M125_13TeV_powheg_pythia8	$0.5824 \times (0.8839 - 0.1227) \times 0.20103$
ggZH_HToBB_ZToLL_M125_13TeV_powheg_pythia8	$0.5824 \times 0.1227 \times 0.10974$
ggZH_HToBB_ZToNuNu_M125_13TeV_powheg_pythia8	$0.5824 \times 0.1227 \times 0.20103$

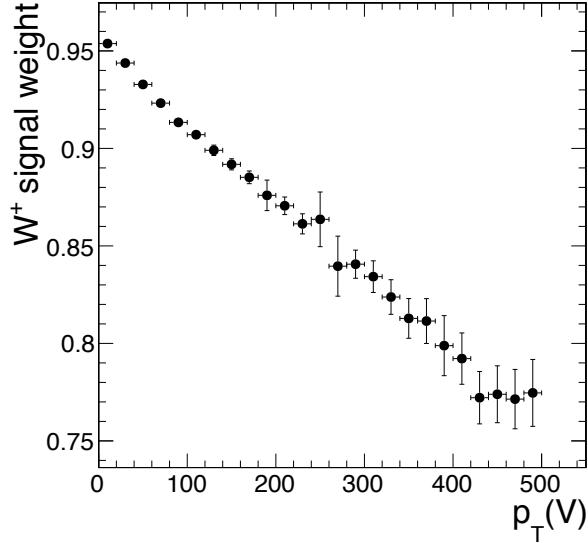


Figure 4-1. The event weights used to rescale the production cross section of the  $W^+(\ell\nu)H \rightarrow b\bar{b}$  signal Monte-Carlo sample to NNLO QCD and NLO electroweak accuracy.

the other cross sections are taken to be the product of their measured inclusive cross sections[88–90] and the appropriate branching ratio as computed by the Particle Data Group (PDG) in Ref. [11].

Table 4-2. The Monte-Carlo samples and their cross sections for the diboson processes considered by the 2017  $VH(b\bar{b})$  analysis.

Sample	$\sigma(\text{pb})$
WWTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	50.86
WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	10.88
ZZ_TuneCP5_13TeV-pythia8	14.60
ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.69

The  $V + \text{jets}$  samples are also produced using the `MADGRAPH5_AMC@NLO v2.4.2` event generator but at LO accuracy with the MLM matching scheme[91]. Besides

their inclusive and HT-binned configurations,  $b$ -quark enriched versions with up to four additional partons are also generated to increase the statistical power of these samples in the phase space most relevant for signal events because the  $V + \text{jets}$  processes are the primary irreducible backgrounds. The samples for  $W + \text{jets}$  with  $W \rightarrow \ell\nu$  are listed in Table 4-3, while the samples for  $Z + \text{jets}$  with  $Z \rightarrow \ell\bar{\ell}$  are listed in Table 4-4 and with  $Z \rightarrow \nu\bar{\nu}$  are listed in Table 4-5. The cross sections of the  $V + \text{jets}$  samples are multiplied by  $k$ -factors of 1.21 and 1.23 for the  $W + \text{jets}$  and  $Z + \text{jets}$  samples, respectively, which rescale them to their NNLO cross sections as calculated by the FEWZ 3.1[92–94] software. The cross sections of the  $b$ -quark enriched samples are further multiplied by a stitching factor which allows them to be used in conjunction with the inclusive and HT-binned versions by appropriately scaling their cross sections.

Table 4-3. The Monte-Carlo samples and their cross sections for the  $W + \text{jets}$  ( $W \rightarrow \ell\nu$ ) processes considered by the 2017  $VH(b\bar{b})$  analysis. Stitching factors of 1.5 and 1.12 are applied to the `WBJets` and `BGenFilter`  $b$ -quark enriched samples, respectively.

Sample	$\sigma(\text{pb})$
<code>WJetsToLNu_TuneCP5_13TeV-madgraphMLM-pythia8</code>	$1.21 \times 52940.0$
<code>WJetsToLNu_HT-100To200_TuneCP5_13TeV-madgraphMLM-pythia8</code>	$1.21 \times 1395.0$
<code>WJetsToLNu_HT-200To400_TuneCP5_13TeV-madgraphMLM-pythia8</code>	$1.21 \times 407.9$
<code>WJetsToLNu_HT-400To600_TuneCP5_13TeV-madgraphMLM-pythia8</code>	$1.21 \times 57.48$
<code>WJetsToLNu_HT-600To800_TuneCP5_13TeV-madgraphMLM-pythia8</code>	$1.21 \times 12.87$
<code>WJetsToLNu_HT-800To1200_TuneCP5_13TeV-madgraphMLM-pythia8</code>	$1.21 \times 5.366$
<code>WJetsToLNu_HT-1200To2500_TuneCP5_13TeV-madgraphMLM-pythia8</code>	$1.21 \times 1.074$
<code>WJetsToLNu_HT-2500ToInf_TuneCP5_13TeV-madgraphMLM-pythia8</code>	$1.21 \times 0.03216$
<code>WBJetsToLNu_Wpt-100to200_TuneCP5_13TeV-madgraphMLM-pythia8</code>	$1.21 \times 1.5 \times 7.35$
<code>WBJetsToLNu_Wpt-200toInf_TuneCP5_13TeV-madgraphMLM-pythia8</code>	$1.21 \times 1.5 \times 1.1$
<code>WJetsToLNu_BGenFilter_Wpt-100to200_TuneCP5_13TeV-madgraphMLM-pythia8</code>	$1.21 \times 1.12 \times 26.6$
<code>WJetsToLNu_BGenFilter_Wpt-200toInf_TuneCP5_13TeV-madgraphMLM-pythia8</code>	$1.21 \times 1.12 \times 3.9$

The  $t\bar{t}$ [95] samples listed in Table 4-6 are generated to NLO accuracy using the POWHEG v2 event generator and their cross sections are rescaled to NNLO accuracy using the next-to-next-to-leading-logarithm (NNLL) values calculated using the `TOP++` v2.0[96] software. The single top production samples listed in Table 4-7 are also generated

Table 4-4. The Monte-Carlo samples and their cross sections for the  $Z + \text{jets}$  ( $Z \rightarrow \ell\bar{\ell}$ ) processes considered by the 2017  $VH(b\bar{b})$  analysis. Stitching factors of 1.085 and 1.15 are applied to the DYJetsToLL and BGenFilter  $b$ -quark enriched samples, respectively.

Sample	$\sigma(\text{pb})$
DYJetsToLL_M-4to50_HT-100to200_TuneCP5_13TeV	$1.23 \times 204.0$
-madgraphMLM-pythia8	
DYJetsToLL_M-4to50_HT-200to400_TuneCP5_13TeV	$1.23 \times 54.4$
-madgraphMLM-pythia8	
DYJetsToLL_M-4to50_HT-400to600_TuneCP5_13TeV	$1.23 \times 5.70$
-madgraphMLM-pythia8	
DYJetsToLL_M-4to50_HT-600toInf_TuneCP5_13TeV	$1.23 \times 1.85$
-madgraphMLM-pythia8	
DYJetsToLL_M-50_TuneCP5_13TeV	$1.23 \times 5343.0$
-madgraphMLM-pythia8	
DYJetsToLL_M-50_HT-100to200_TuneCP5_13TeV	$1.23 \times 161.1$
-madgraphMLM-pythia8	
DYJetsToLL_M-50_HT-200to400_TuneCP5_13TeV	$1.23 \times 48.66$
-madgraphMLM-pythia8	
DYJetsToLL_M-50_HT-400to600_TuneCP5_13TeV	$1.23 \times 6.97$
-madgraphMLM-pythia8	
DYJetsToLL_M-50_HT-600to800_TuneCP5_13TeV	$1.23 \times 1.743$
-madgraphMLM-pythia8	
DYJetsToLL_M-50_HT-800to1200_TuneCP5_13TeV	$1.23 \times 0.805$
-madgraphMLM-pythia8	
DYJetsToLL_M-50_HT-1200to2500_TuneCP5_13TeV	$1.23 \times 0.193$
-madgraphMLM-pythia8	
DYJetsToLL_M-50_HT-2500toInf_TuneCP5_13TeV	$1.23 \times 0.00347$
-madgraphMLM-pythia8	
DYBJetsToLL_M-50_Zpt-100to200_TuneCP5_13TeV	$1.23 \times 1.085 \times 4.042$
-madgraphMLM-pythia8	
DYBJetsToLL_M-50_Zpt-200toInf_TuneCP5_13TeV	$1.23 \times 1.085 \times 0.4286$
-madgraphMLM-pythia8	
DYJetsToLL_BGenFilter_Zpt-100to200_M-50_TuneCP5_13TeV	$1.23 \times 1.15 \times 3.384$
-madgraphMLM-pythia8	
DYJetsToLL_BGenFilter_Zpt-200toInf_M-50_TuneCP5_13TeV	$1.23 \times 1.15 \times 0.5327$
-madgraphMLM-pythia8	

Table 4-5. The Monte-Carlo samples and their cross sections for the  $Z + \text{jets}$  ( $Z \rightarrow \nu\bar{\nu}$ ) processes considered by the 2017  $VH(b\bar{b})$  analysis. Stitching factors of 1.085 and 1.11 are applied to the `ZBJetsToNuNu` and `BGenFilter`  $b$ -quark enriched samples, respectively.

Sample	$\sigma(\text{pb})$
<code>ZJetsToNuNu_HT-100To200_13TeV-madgraph</code>	$1.23 \times 304.2$
<code>ZJetsToNuNu_HT-200To400_13TeV-madgraph</code>	$1.23 \times 91.92$
<code>ZJetsToNuNu_HT-400To600_13TeV-madgraph</code>	$1.23 \times 13.18$
<code>ZJetsToNuNu_HT-600To800_13TeV-madgraph</code>	$1.23 \times 3.258$
<code>ZJetsToNuNu_HT-800To1200_13TeV-madgraph</code>	$1.23 \times 1.496$
<code>ZJetsToNuNu_HT-1200To2500_13TeV-madgraph</code>	$1.23 \times 0.3419$
<code>ZJetsToNuNu_HT-2500ToInf_13TeV-madgraph</code>	$1.23 \times 0.005112$
<code>ZBJetsToNuNu_M-50_Zpt-100to200_TuneCP5_13TeV-madgraphMLM-pythia8</code>	$1.23 \times 1.085 \times 7.7$
<code>ZBJetsToNuNu_M-50_Zpt-200toInf_TuneCP5_13TeV-madgraphMLM-pythia8</code>	$1.23 \times 1.085 \times 0.8131$
<code>ZJetsToNuNu_BGenFilter_Zpt-100to200_M-50_TuneCP5_13TeV-madgraphMLM-pythia8</code>	$1.23 \times 1.11 \times 3 \times 2.139$
<code>ZJetsToNuNu_BGenFilter_Zpt-200toInf_M-50_TuneCP5_13TeV-madgraphMLM-pythia8</code>	$1.23 \times 1.11 \times 3 \times 0.3287$

to NLO accuracy, with the t-channel[97] and tW-channel[98] processes generated using the POWHEG v2 event generator and the s-channel[99] process generated using the `MADGRAPH5_AMC@NLO` v2.4.2 event generator. Their cross sections are also rescaled to their corresponding values obtained from NNLO calculations.[100, 101] Finally, the QCD or multi-jet samples listed in Table 4-8 are generated to LO accuracy using the `MADGRAPH5_AMC@NLO` v2.4.2 event generator with the MLM matching scheme.

Table 4-6. The Monte-Carlo samples and their cross sections for the  $t\bar{t}$  processes considered by the 2017  $VH(b\bar{b})$  analysis.

Sample	$\sigma(\text{pb})$
<code>TTTo2L2Nu_TuneCP5_PSweights_13TeV-powheg-pythia8</code>	88.29
<code>TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8</code>	365.34
<code>TTToHadronic_TuneCP5_PSweights_13TeV-powheg-pythia8</code>	377.96

The set of parton distribution functions, which define the distribution of a hadron's momentum among its partons, used to generate all of the MC samples was chosen to be the NNPDF3.1[102] set. The parton showering and hadronization was handled by interfacing the matrix element generators with PYTHIA v8.230. Finally, additional

Table 4-7. The Monte-Carlo samples and their cross sections for the single top processes considered by the 2017  $VH(b\bar{b})$  analysis.

Sample	$\sigma(\text{pb})$
ST_s-channel_4f_leptonDecays_TuneCP5_PSweights_13TeV -amcatnlo-pythia8	$0.325 \times 10.32$
ST_t-channel_antitop_4f_inclusiveDecays_TuneCP5_13TeV -powhegV2-madspin-pythia8	$0.325 \times 80.95$
ST_t-channel_top_4f_inclusiveDecays_TuneCP5_13TeV -powhegV2-madspin-pythia8	$0.325 \times 136.02$
ST_tW_antitop_5f_inclusiveDecays_TuneCP5_PSweights_13TeV -powheg-pythia8	35.85
ST_tW_top_5f_inclusiveDecays_TuneCP5_PSweights_13TeV -powheg-pythia8	35.85

Table 4-8. The Monte-Carlo samples and their cross sections for the QCD processes considered by the 2017  $VH(b\bar{b})$  analysis.

Sample	$\sigma(\text{pb})$
QCD_HT100to200_TuneCP5_13TeV-madgraphMLM-pythia8	27990000
QCD_HT200to300_TuneCP5_13TeV-madgraphMLM-pythia8	1547000
QCD_HT300to500_TuneCP5_13TeV-madgraphMLM-pythia8	322600
QCD_HT500to700_TuneCP5_13TeV-madgraphMLM-pythia8	29980
QCD_HT700to1000_TuneCP5_13TeV-madgraphMLM-pythia8	6334
QCD_HT1000to1500_TuneCP5_13TeV-madgraphMLM-pythia8	1088
QCD_HT1500to2000_TuneCP5_13TeV-madgraphMLM-pythia8	99.11
QCD_HT2000toInf_TuneCP5_13TeV-madgraphMLM-pythia8	20.23

$pp$  interactions are added to the hard-scattering process to simulate pileup, with a multiplicity distribution matched to the 2017 data taking conditions.

#### 4.1.3 Triggers

Each channel employs a subset of the available triggers to select data events which are consistent with their signal hypothesis. The  $Z(\nu\bar{\nu})H$  channel uses triggers which place thresholds on the missing transverse energy (MET), missing transverse hadronic energy (MHT), and transverse hadronic energy (HT) in an event. The  $W(e\nu)H$  and  $W(\mu\nu)H$  channels both use single lepton triggers, while the  $Z(e\bar{e})H$  and  $Z(\mu\bar{\mu})H$  channels both use di-lepton triggers, which place thresholds on the transverse momentum of isolated leptons. The specific triggers used by each channel are detailed in Table 4-9. Because the triggers

are also emulated during simulation, the MC events are also required to satisfy the same triggers as used in data.

Table 4-9. The L1 and HLT triggers used by the 2017  $VH(b\bar{b})$  analysis, organized by decay channel.

Channel	L1 Seeds	HLT Paths
$Z(\nu\bar{\nu})H$	(L1_ETM110 OR L1_ETMHF120) OR L1_ETMHF110_HTT60er	HLT_PFMET120_PFMHT120_IDTight OR HLT_PFMET120_PFMHT120_IDTight_PFHT60
$W(e\nu)H$	L1_SingleEG38 OR L1_SingleIsoEG30 OR L1_SingleIsoEG28er2p1 OR L1_DoubleEG_25_12	HLT_Ele32_WPTight_Gsf_L1DoubleEG
$W(\mu\nu)H$	L1_SingleMu22	HLT_IsoMu27
$Z(e\bar{e})H$	L1_SingleEG30 OR L1_SingleIsoEG22er OR L1_SingleIsoEG24 OR L1_DoubleEG_15_10	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL
$Z(\mu\bar{\mu})H$	L1_DoubleMu_12_5	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8

The primary trigger for the  $Z(\nu\bar{\nu})H$  channel is `HLT_PFMET120_PFMHT120_IDTight`, which requires both particle-flow (PF) MET and MHT to be above 120 GeV. The corresponding triggers at L1 are seeded by ETM with thresholds ranging from 100 GeV to 120 GeV which varies based on the instantaneous luminosity of the LHC in order to maintain reasonable trigger rates. The online PF MET, though similar to the offline version, uses a simplified version of tracking and is evaluated as the transverse momentum ( $p_T$ ) imbalance of all PF objects reconstructed at the HLT which includes photons, electrons, muons, and corrected jets from neutral and charged hadrons. The online PF MHT considers corrected PF jets with  $p_T > 20\text{GeV}$  and  $|\eta| < 5.2$  which have neutral hadronic fraction  $< 0.9$ , neutral electromagnetic fraction  $< 0.99$ , and at least one constituent. The jets contributing to the PF MHT which lie within the tracker acceptance, or  $|\eta| < 2.4$ , are also required to have charged hadronic fraction  $> 0$ , charged electromagnetic fraction  $< 0.99$ , and charged multiplicity  $> 0$ .

The secondary trigger for the  $Z(\nu\bar{\nu})H$  channel, `HLT_PFMET120_PFMHT120_IDTight_PFHT60`, additionally requires the PF HT to be above 60 GeV. This additional requirement was verified to pose no significant impact, as there are at least two highly-energetic jets present

in the events considered by the analysis. Because the primary trigger was sometimes inactive during run period F, this secondary trigger is used to guarantee full coverage throughout 2017 by taking the logical **OR** of the two triggers.

The trigger efficiency of the logical **OR** of the primary and secondary MET triggers is measured using the SingleElectron primary dataset. Because the single-electron triggers are orthogonal to the MET triggers, this dataset provides an unbiased sample with which the trigger efficiency can be measured. In addition to passing the **OR** of the primary and secondary MET triggers, events are also required to have two jets within the tracker acceptance and the electron is required to have  $|\Delta\phi(e, \text{MET})| < 2.5$ . This separation in azimuthal angle rejects events which have an electron that is back-to-back with the reconstructed MET in order to avoid bias from the L1 MET. As the triggers are parameterized by both the online PF MET and PF MHT, the efficiency is measured as a function of the minimum of the offline MET and MHT, i.e.  $\min(\text{MET}, \text{MHT})$ . The trigger efficiency curve is derived by fitting the data points shown in Figure 4-2 with the convolution of a crystal ball function and a step function.

The MET trigger efficiency can also be measured using MC samples to assess the trigger emulation performance. The trigger efficiency curve of the emulated MET triggers is obtained using the same procedure as for data and is shown in Figure 4-3. An efficiency correction is obtained by taking the ratio of the fitted trigger efficiency curve for data to that of simulation and is shown in Figure 4-4. The correction reaches up to 9% in the “turn-on” region of the efficiency curve, but then remains close to unity upon reaching the plateau. The uncertainty of this efficiency correction were determined by the eigenvector decomposition of the covariance matrices of the fitted functions for data and simulation, and those which have non-negligible effects on the shape of the final discriminant or process normalizations are assessed as systematic uncertainties.

The trigger efficiencies for the other decays channels are similarly handled. The triggers used by the  $W(\ell\nu)H$  channel reach efficiencies of approximately 90% for electrons

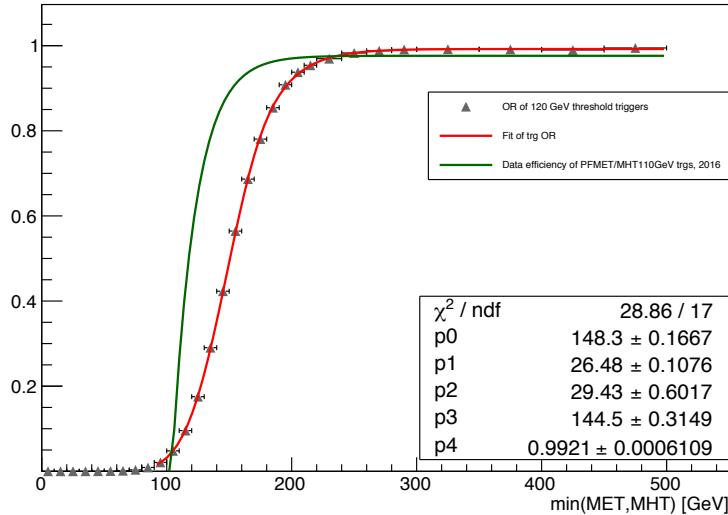


Figure 4-2. The trigger efficiency for the  $Z(\nu\bar{\nu})H(b\bar{b})$  channel as a function of the minimum of the offline MET and MHT measured using the full 2017 SingleElectron primary dataset. The green curve represents the efficiency in 2016 data of the logical OR of analogous triggers used by the 2016 analysis. The red curve represents the efficiency of the logical OR of the triggers used for the 2017 analysis.

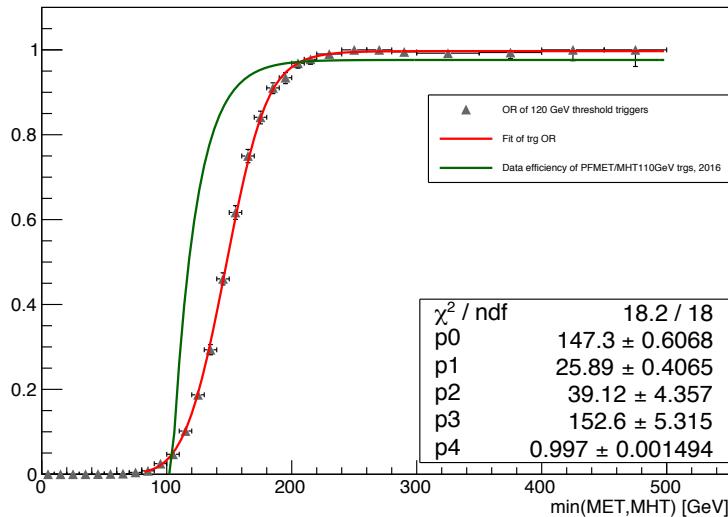


Figure 4-3. The emulated trigger efficiency for the  $Z(\nu\bar{\nu})H(b\bar{b})$  channel as a function of the minimum of the offline MET and MHT measured using a  $W + 3$ -jets Monte-Carlo sample. The green curve represents the efficiency in 2016 data of the logical OR of analogous triggers used by the 2016 analysis. The red curve represents the efficiency of the emulated logical OR of the triggers used for the 2017 analysis.

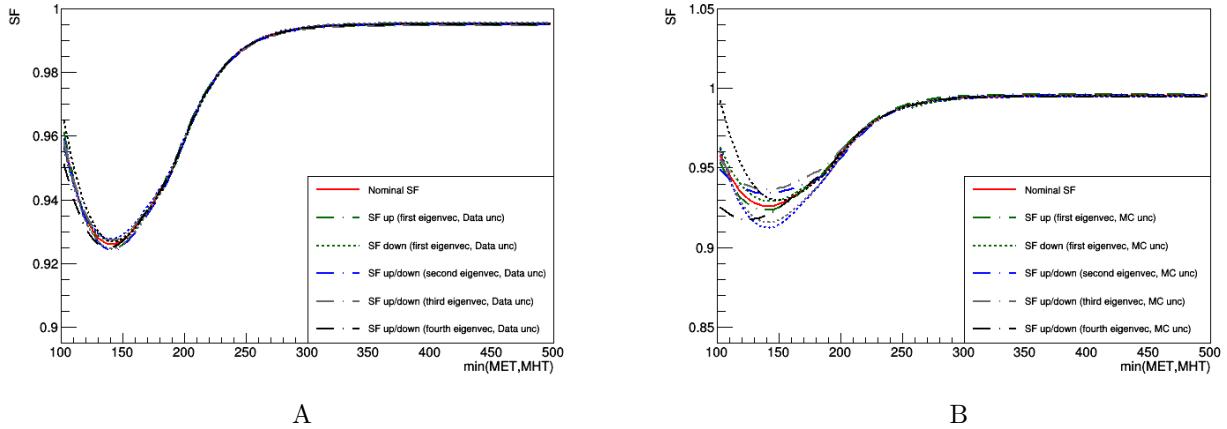


Figure 4-4. The MET trigger efficiency correction for the  $Z(\nu\bar{\nu})H(b\bar{b})$  channel as a function of the minimum of the offline MET and MHT. The red curve represents the nominal correction, while the dotted and dash-dotted curves represent the variations in the correction due to the four leading uncertainties in A) data and B) Monte-Carlo.

and 95% for muons. The triggers used by the  $Z(\ell\bar{\ell})H$  channel reach efficiencies of approximately 96% for electrons and 91% for muons.

#### 4.1.4 Residual Monte-Carlo Corrections

Although the MC samples provide realistic simulations of physics processes, discrepancies between the shapes of kinematic distributions in data and MC remain a concern. The presence of such observed differences is not unexpected, given the fixed-order accuracy of and the assumptions made by the event generators. Residual corrections therefore need to be applied to the MC samples to improve the agreement between data and MC for key distributions.

The first such correction addresses the difference between the  $p_T(V)$  spectrum in data and the  $V+\text{jets}$  MC samples. The  $p_T(V)$  distribution is harder for simulation than for data because the simulation does not include higher-order electroweak corrections.[103] The  $V+\text{jets}$  samples are therefore reweighted as a function of the  $p_T(V)$  to apply the NLO electroweak correction shown in Figure 4-5 which accounts for discrepancies of up to 10% for  $p_T(V)$  near 400 GeV.

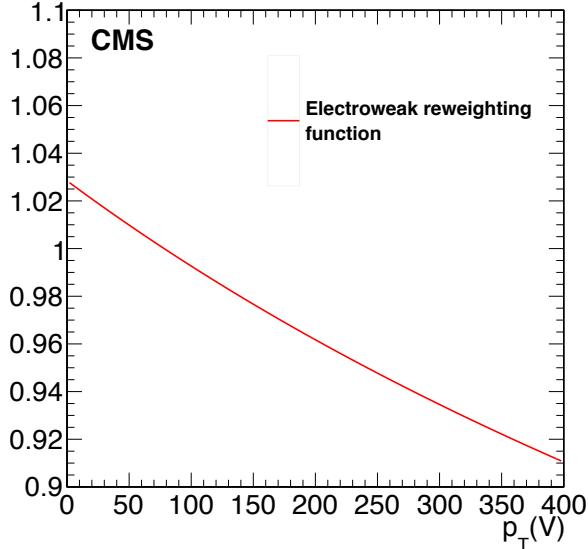


Figure 4-5. The NLO electroweak correction applied to the  $V + \text{jets}$  Monte-Carlo samples.

The second such correction addresses the difference between the  $p_T$  spectra of the top quarks in data and the  $t\bar{t}$  MC samples. The  $p_T$  spectra of top quarks in data are observed to be softer than predicted by the event generators. The  $t\bar{t}$  samples are therefore reweighted as a function of the top quark  $p_T$  according to the official recommendation by the CMS experiment.[104] This correction is only applicable for the  $Z(\nu\bar{\nu})H$  and  $Z(\ell\bar{\ell})H$  channels, and is superceded by an equivalent reweighting specific to the  $W(\ell\nu)H$  channel.

A third correction addresses the discrepancy between the di-jet invariant mass  $m(jj)$  distribution in data and the LO  $V + \text{jets}$  MC samples. Although NLO  $V + \text{jets}$  MC samples are readily available and show good agreement with data for the  $m(jj)$  distribution, they are not used by the analysis because their limited statistical power results in an over 10% decrease in expected sensitivity with respect to the LO  $V + \text{jets}$  samples. A differential LO-to-NLO correction based on the separation in  $\eta$  between the two  $b$ -quark jets from the Higgs boson decay  $\Delta\eta(jj)$  is derived as a ratio of the NLO to LO  $V + \text{jets}$  samples following the procedure outlined in Ref. [28]. An example of this ratio is shown in Figure 4-6. These reweighting functions improve the agreement between data and the LO  $V + \text{jets}$

samples for the  $m(jj)$  and  $p_T(V)$  distributions while demonstrating a negligible effect on the remaining distributions. The full reweighting is assessed as a systematic uncertainty.

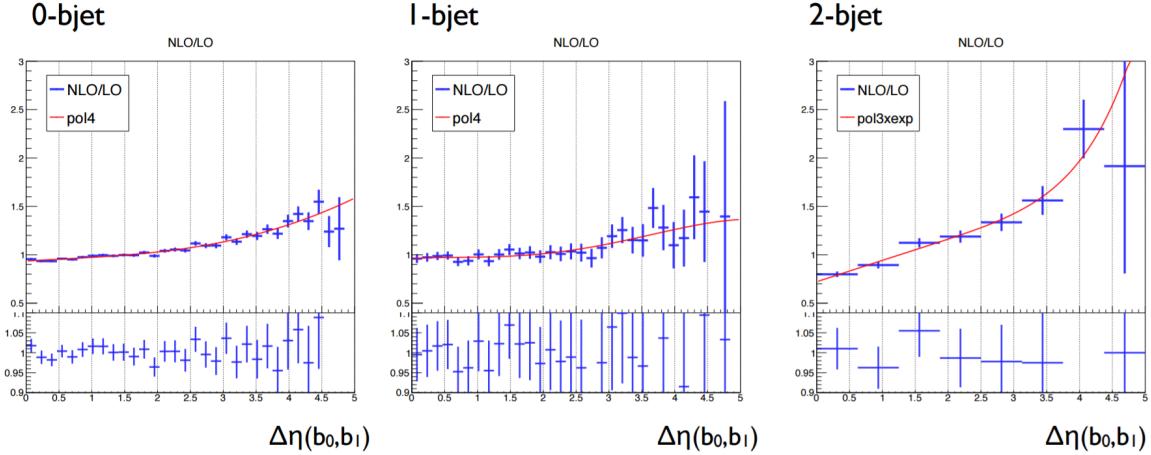


Figure 4-6. The ratio of  $\Delta\eta(jj)$  between the NLO to LO  $Z + \text{jets}$  ( $Z \rightarrow \ell\bar{\ell}$ ) Monte-Carlo samples divided into categories based on the number of true  $b$ -jets present.

A final correction addresses the downward trend in the data to MC ratio for the  $p_T(V)$  distribution that is observed in the control regions of the  $W(\ell\nu)H$  channel. A simultaneous fit of the  $p_T(V)$  distribution to data in the control regions of the  $W(\ell\nu)H$  channel are used to derive independent linear reweighting functions for the  $t\bar{t}$ ,  $W$  with light-flavored jets ( $W + udcsg$ ), and combination of  $W$  with  $b$ -jets ( $W + b\bar{b}$ ) and single top processes. The relative compositions of the background processes were fixed during the fit and the reweighting preserves the overall normalization. The uncertainties on the fitted slopes of the linear reweighting functions are taken to be the systematic uncertainties of this  $p_T(V)$  correction. An uncertainty of 13% is assessed for the  $t\bar{t}$  reweighting, while a 6% uncertainty is assessed for both the  $W + udcsg$  and  $W + b\bar{b}$  and single top combined reweightings. These uncertainties sufficiently cover any perceived discrepancies after all corrections have been applied.

## 4.2 Physics Objects

The standard physics object candidates reconstructed by the PF algorithm and described in Chapter 3 are used by the analysis. The identification and selection of

electrons, muons, jets,  $b$ -jets, and MET proceeds according to recipes approved and provided by their corresponding Physics Object Group (POG). Recommended working points and corrections are also applied, as well as additional reconstruction techniques developed for the  $VH(b\bar{b})$  analysis.

The primary vertex candidates are reconstructed according to the deterministic annealing algorithm.[49] The signal vertex, the primary vertex corresponding to the primary hard interaction which triggered the event, is chosen to be the primary vertex with the largest  $\sum p_T^2$ , where the sum runs over all elementary particles identified at the tracking level, namely track-jets, track-MET, and charged leptons. In addition, displaced tracks originating from the decay of  $b$ -hadrons are associated with their proper primary vertex to increase the probability of choosing the correct signal vertex, especially for the  $Z(\nu\bar{\nu})H$  channel which searches for a decay final state with two  $b$ -jets and large MET.

Because of the high instantaneous luminosity and tight bunch crossing timing, the number of reconstructed primary vertices increases with the amount of pileup, or additional proton-proton interactions that may be *in-time*, occurring within the same bunch crossing considered by the event, or *out-of-time*, occurring within the previous or next bunch crossings. The presence of pileup interactions impacts the choice of signal vertex as well as the reconstruction of all other objects by worsening the momentum resolution of jets and, by extension, MET and affecting the calculation of lepton isolation and the performance of  $b$ -taggers. The analysis addresses pileup effects by employing the following approaches:

- **Charged hadronic subtraction (CHS)** is an algorithm integrated with the reconstruction algorithm for PF jets which filters all charged hadrons whose trajectories are incompatible with the primary interaction.
- **FastJet subtraction** is an algorithm provided by the **FastJet**[60] software package which calculates the average momentum density per unit area due to pileup for each event which is used to decontaminate jet and lepton isolation cones.

Finally, to account for observed differences between the pileup distribution in data and MC simulation, the simulated events are reweighted to match the amount of generated pileup. After the effects of pileup have been treated, the agreement between data and simulation is much better, as demonstrated in Figure 4-7.

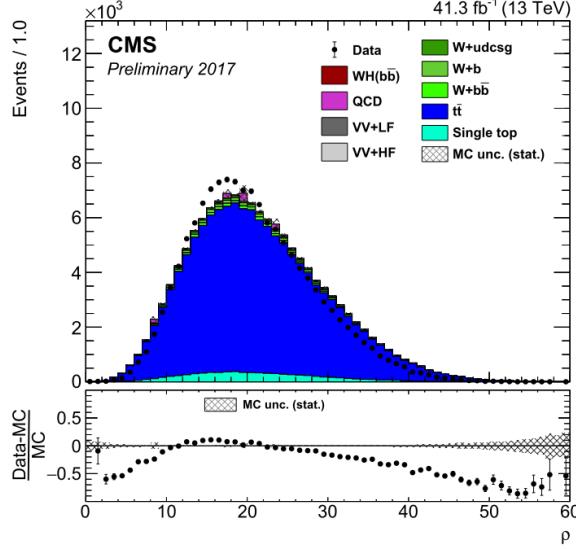


Figure 4-7. The distribution of the average jet energy density  $\rho$  for 2017 data and simulation in the  $t\bar{t}$  control region of the  $W(\ell\nu)H$  channel which is defined in section 4.3.2.

Electrons are reconstructed using the Gaussian-sum filter (GSF) algorithm.[53] The electron candidates are then selected by first applying a loose relative isolation cut of 0.4, after which they must pass tighter identification requirements based on multivariate discriminator thresholds determined by the Electron-Gamma (EGM) POG. Two of the prescribed working points are used:

- The loose working point (WP90) achieves an electron selection efficiency of 90% and is used by the analysis to perform an initial categorization of event candidates for every decay channel as well as to identify the electrons in the  $Z(e\bar{e})H$  channel.
- The tight working point (WP80) achieves an electron selection efficiency of 80% and is used to identify electrons in the  $W(e\nu)H$  channel and suppress background.

In order to reduce the multijet background, the electron isolation cut is specifically tightened to 0.06 for the  $W(e\nu)H$  channel. The  $Z(e\bar{e})H$  channel loosened the electron

isolation cut to 0.15 because its requirement of two leptons in the final state virtually eliminates all multijet background for that channel, allowing it to increase its signal efficiency at little cost.

Muons are reconstructed using tracker muon tracks and global muon tracks identified by the PF algorithm. The muon candidates are also selected by first applying a loose relative isolation cut of 0.4, after which they are required to pass tighter criteria proposed by the Muon POG which are based on the results of the global muon track fit and information from the tracker which are used to determine if it the muon is prompt. The  $W(\mu\nu)H$  channel tightened the muon isolation cut to 0.06, which by chance is the same value used by the  $W(e\nu)H$  channel for the electron isolation cut, in order to remove data excesses in the poorly modelled tails of the isolation distribution. The  $Z(\mu\bar{\mu})H$  channel loosened the muon isolation cut to 0.12 for same reasons the  $Z(e\bar{e})H$  channel loosened its electron isolation cut.

Jets are reconstructed using the anti- $k_T$  jet clustering algorithm[61] to cluster together PF candidates with a radius parameter of  $R = 0.4$  to form so-called “AK4” jets. The jet energy scale (JES) corrections recommended by the Jet/MET (JME) POG and derived for 2017 data and simulation were applied to calibrate the energy of the jets. The jet energies are also smeared in simulation in order for the resolution in simulation to match the resolution in data. A loose jet identification is applied that rejects misreconstructed candidates due to detector noise or contamination from pileup energy. Jets overlapping a preselected electron or muon are discarded, and jets which are outside of the tracker acceptance of  $|\eta| < 2.5$  are not considered by the analysis.

The identification of jets initiated by the hadronization of  $b$ -quarks is performed using the Deep Combined Secondary Vertex (DeepCSV)[63]  $b$ -tagging algorithm. The DeepCSV tagger uses as its discriminator a neural network classifier which takes as input information from secondary vertices and track impact parameters associated with a jet in an event and outputs five continuous values that are interpreted to be the probabilities

that the jet belongs to one of five different possible hadronization scenarios described in section 3.5.3. In the context of this analysis, the DeepCSV score refers to the summed probability that a jet originated from a  $b$ -quark or two  $b$ -quarks, as the Higgs boson will be reconstructed from a compatible pair of  $b$ -jet candidates. The  $b$ -Tagging and Vertexing (BTV) POG recommends three working points, shown in Table 4-10, which trade off between  $b$ -tagging efficiency and misidentification rate.

Table 4-10. The 2017  $b$ -tagging working points defined for the DeepCSV algorithm, along with their associated misidentification rates.[63]

Working Point	DeepCSV Score	Misidentification Rate
Loose	0.1522	10.0%
Medium	0.4941	1.0%
Tight	0.8001	0.1%

Observed discrepancies in the  $b$ -tagging efficiency between data and simulation are addressed by calibrating the DeepCSV discriminator values using a tag-and-probe (TnP) approach such that the the DeepCSV distribution for jets in simulation matches the distribution observed in data control regions. The binned distribution of the probe jets in data is compared to the expected distribution from simulation and an iterative procedure scales the content of each bin to match while preserving the expected proportions of light and heavy flavor jets. The procedure itself is carried out in bins of  $p_T$  and  $\eta$ . The ratio of the scaled distribution in simulation to the original distribution in simulation defines a per jet weight  $w_j$  which is function of DeepCSV and parameterized by the jet  $p_T$ ,  $\eta$ , and flavor. For an event with  $N_{\text{jets}}$ , the product of the weights for each jet  $j$  defines an overall  $b$ -tag calibration weight

$$w_{b\text{-tag}} = \prod_{j=1}^{N_{\text{jets}}} w_j (\text{DeepCSV}_j | p_{Tj}, \eta_j, \text{flavor}_j). \quad (4-1)$$

The missing transverse momentum  $\text{vec}p_T^{\text{miss}}$  is reconstructed as the negative vectorial sum of the transverse momentum of all PF objects identified in the event. The magnitude of  $\text{vec}p_T^{\text{miss}}$  is the missing transverse energy  $p_T^{\text{miss}}$ , or colloquially MET. A

Type-I correction, where the JES corrections are applied to the PF jets which enter the vectorial sum, is applied to calibrate the MET. The jet smearing is also propagated to the MET calculation to more realistically match the resolution of the MET in data. Finally, events with large MET caused by spurious detector noise are rejected by applying the standard set of MET filters recommended by the JetMET POG. In addition to the MET, the missing hadronic transverse energy MHT is also used by the  $Z(\nu\bar{\nu})H$  channel, which is defined as the magnitude of the missing hadronic transverse momentum, or the negative vectorial sum of the transverse momenta of only jets with  $p_T > 30$  GeV and  $|\eta| < 2.4$ .

Aside from the vector boson and Higgs boson decays, additional hadronic activity is expected to be low for signal events. The remaining activity is expected to be soft, and so any low  $p_T$  tracks originating from the signal vertex but not associated with the vector boson or Higgs boson candidates are used to reconstruct the “soft activity”. These tracks are then clustered into AK4 “soft-track jets” which have been demonstrated to capably reconstruct the hadronization of very low energy partons.[\[105\]](#) To further separate signal from background, the  $VH(b\bar{b})$  analysis defines and uses the quantity SA5, which is the number of soft-track jets in the event with  $p_T > 5$  GeV.

#### 4.2.1 Vector Boson Candidate

For the  $Z(\nu\bar{\nu})H$  channel, the vector boson is simply reconstructed from large MET because the  $Z$  decays invisibly to a pair of neutrinos. Candidate  $Z(\nu\bar{\nu})H$  events are required to have  $\text{MET} > 150$  GeV, and the transverse momentum of the reconstructed  $Z$  candidate is defined as  $p_T(Z) = \min(\text{MET}, \text{MHT})$ . In order to remove observed discrepancies between the data and simulation in the turn-on region of the MET trigger efficiency curve, the analysis of the  $Z(\nu\bar{\nu})H$  channel only considers events with  $p_T(Z) > 170$  GeV.

For the remaining channels, the vector boson reconstruction involves the selection and identification of isolated leptons. For the  $Z(\ell\bar{\ell})H$  channels, the  $Z$  boson is reconstructed from pairs of isolated electrons or muons which have a dilepton invariant mass  $m(\ell\ell)$

within the range  $75 \text{ GeV} < m(\ell\ell) < 105 \text{ GeV}$  which covers the on and off-shell decays of the  $Z$  boson. Separate analyses of the  $Z(\ell\bar{\ell})H$  channels are then performed for  $Z$  boson candidates in a low  $p_T$  category ( $50 \text{ GeV} < p_T(Z) < 150 \text{ GeV}$ ) and high  $p_T$  category ( $p_T(Z) > 150 \text{ GeV}$ ). For the  $W(\ell\nu)H$  channels, the  $W$  boson is reconstructed from a single isolated electron or muon and the MET and the transverse momentum and transverse mass of the  $W$  boson candidate is defined as

$$p_T(W) = \sqrt{(\text{MET}_x + p_x^\ell)^2 + (\text{MET}_y + p_y^\ell)^2} \quad (4-2)$$

and

$$m_T(W) = \sqrt{(\text{MET}_x + p_x^\ell)^2 - p_T(W)^2}, \quad (4-3)$$

where  $p^\ell$  is the four-momentum of the isolated lepton.

#### 4.2.2 Higgs Boson Candidate

For all decay channels, the Higgs boson is most efficiently reconstructed using the pair of  $b$ -jets with the largest DeepCSV score per event, where the leading  $b$ -jet is required to pass the tight working point cut for all channels except  $Z(\ell\bar{\ell})H$  and the subleading  $b$ -jet is required to pass the loose working point. Because the identification of the Higgs boson candidate is crucial to the analysis, advanced techniques are used to improve the dijet invariant mass resolution. The application of a  $b$ -jet energy regression, a kinematic fit in the  $Z(\ell\bar{\ell})H$  channel, and a recovery of final state radiation (FSR) results in a channel-dependent improvement of 10-23% for the mass resolution.

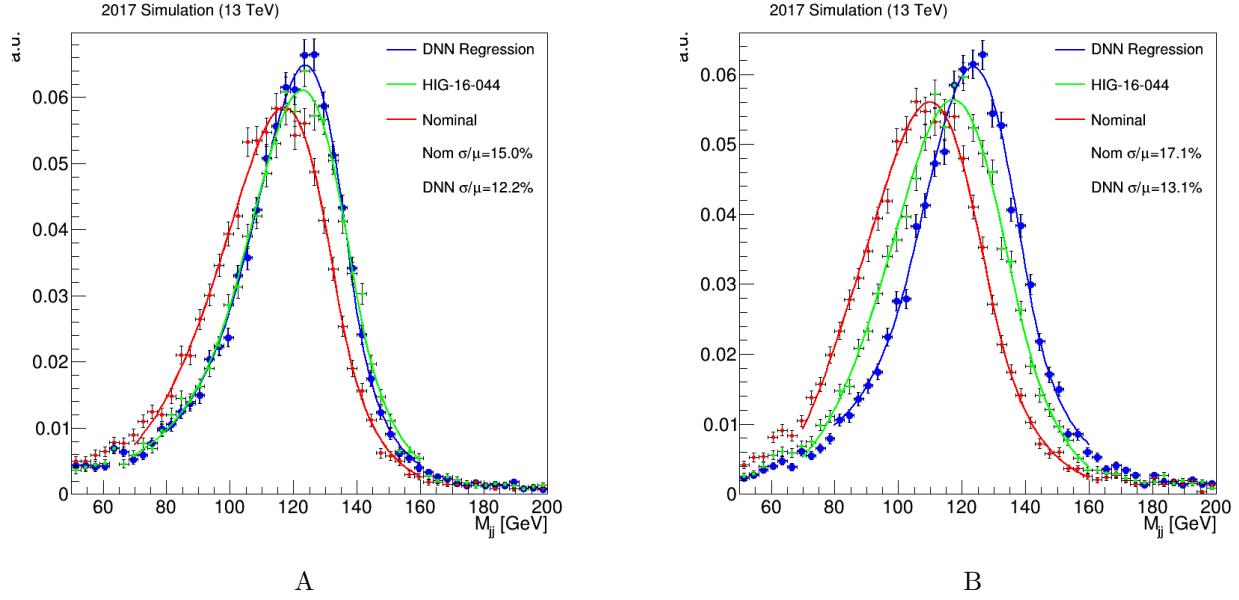
The  $b$ -jet energy regression is used to correct the energies of  $b$ -jets by providing an accurate estimate of the true  $b$ -jet energy and resolution. The multitarget regression is modelled using a neural network implemented using the Keras[64] neural network framework with Tensorflow[65] as the backend. There are 43 input features based on the  $b$ -jet kinematics, pileup in the event, leading tracks in the  $b$ -jet, soft-lepton tracks in the  $b$ -jet, secondary vertex candidates, and  $b$ -jet energy fractions carried by electromagnetic, charged, and neutral constituents in five rings of  $\Delta R$  within the jet cone. The architecture

of the neural network model uses six hidden layers with the number of nodes in each layer being 1024, 1024, 1024, 512, 256, and 128, respectively. The hidden layer nodes are leaky rectified linear units (Leaky ReLU)[106] with a slope of  $\beta = 0.2$  for the negative part of its argument:

$$f(x) = \begin{cases} x, & \text{if } x > 0 \\ \beta x, & \text{otherwise} \end{cases} = \begin{cases} x, & \text{if } x > 0 \\ 0.2x, & \text{otherwise} \end{cases}. \quad (4-4)$$

Batch-normalization[107] and dropout[108] with a probability of 0.1 are applied to the outputs of each hidden layer to increase the training speed and regularize the model. The output layer consists of three linear nodes, one of which is trained to estimate the jet energy correction while the others are trained to estimate the jet energy resolution at 25% and 75% quantiles. A composite loss function is used, which is taken to be the sum of the Huber loss[109] for the jet energy correction task and two asymmetrically-weighted quantile losses for the jet energy resolution tasks. The training was performed with stochastic gradient descent using adaptive moment estimation, or the Adam optimizer.[110] The model hyperparameters were optimized by taking the best performing random model out of 50 trials sampled from a discrete grid. By calibrating the  $b$ -jets energies using the corrections obtained by the regression, channel-dependent improvements of 10-23% for the mass resolution are observed, with the improvements to the  $Z(\ell\bar{\ell})H$  channel shown in Figure 4-8.

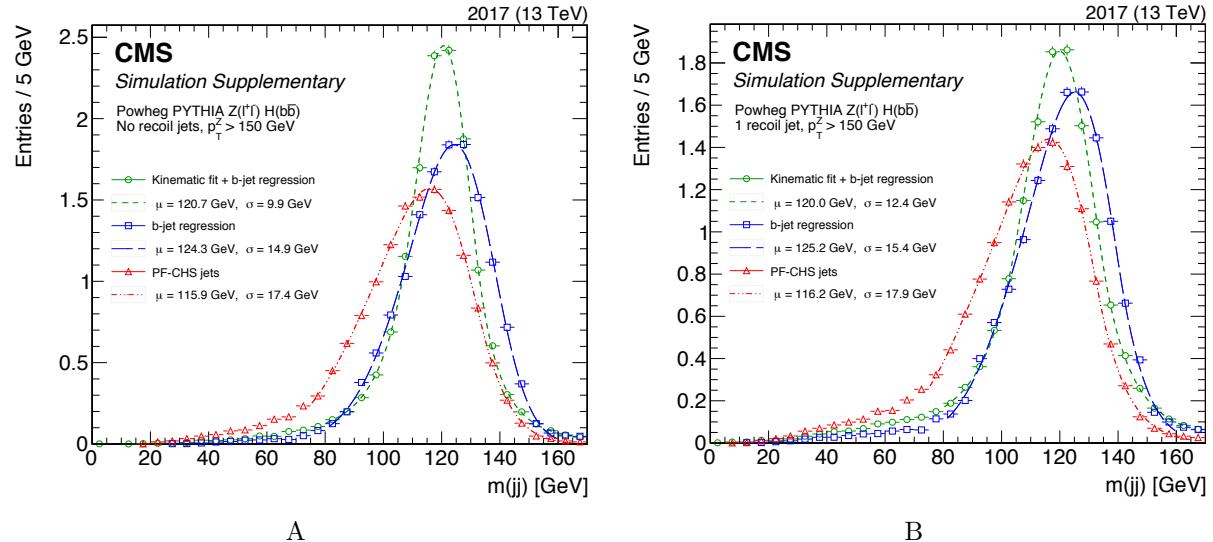
The kinematic fit is specific to the  $Z(\ell\bar{\ell})H$  channel because its events have no genuine MET, thereby admitting a unique kinematic constraint which may be exploited to improve the dijet invariant mass resolution. The **PhysicsTools/KinFitter** package distributed with CMSSW is used to perform the fit, which consists of first constraining the dilepton system to the  $Z$  boson mass and then constraining the dilepton-dijet system in the transverse plane, all within uncertainties. This balancing of the dilepton and dijet systems also considers a recoil vector to account for jets caused by initial state radiation (ISR).



A

B

Figure 4-8. The dijet invariant mass distribution of the  $Z(\ell\bar{\ell})H$  channel before (red) and after the  $b$ -jet energy regression is applied for A)  $b$ -jets which do not contain leptons and B)  $b$ -jets which contain leptons from semi-leptonic decays of  $b$ -hadrons. The artificial neural network based regression (blue) used by the 2017 analysis outperforms the boosted decision tree based regression (green) used by the 2016 analysis.



A

B

Figure 4-9. The dijet invariant mass distribution of the  $Z(\ell\bar{\ell})H$  channel before (red) and after the  $b$ -jet energy regression corrections (blue) and the kinematic fit (green) are applied for A) events without a recoiling jet and B) events with a single recoiling jet.

The kinematic fit further improves the dijet invariant mass resolution in the  $Z(\ell\bar{\ell})H$  channel by at least 20% on top of the  $b$ -jet energy regression, as shown in Figure 4-9.

Although a small effect, the final state partons can emit FSR radiation that is collinear to the direction of the  $b$ -jet. In order to recover FSR belonging to the Higgs boson candidates, the four-vectors of soft jets which lie within a cone of  $\Delta R < 0.8$  around either of the candidate  $b$ -jets used to reconstruct the Higgs boson are added vectorially to the Higgs boson candidate. The FSR recovery results in a channel-dependent improvement of 1-5% for the dijet invariant mass resolution.

### 4.3 Event Selection

The analysis applies a channel-dependent event selection in order to maximize the signal purity while reducing the level of background. These selections are primarily motivated by the final state topology of signal process. The  $VH(b\bar{b})$  decay considered by the analysis features a leptonically decaying vector boson that recoils away from a pair of  $b$ -jets in the final state. The vector boson candidate and dijet system are therefore expected to be back-to-back in the transverse plane, such that the difference in azimuthal angle between them peaks at  $\pi$ . Because the invariant mass of the dijet system is expected to be compatible with the mass of the Higgs boson in the range  $100 \text{ GeV} < m_H < 150 \text{ GeV}$ , the  $p_T$  distributions of the Higgs jet candidates are expected to peak at roughly  $m_H/2$ . Additional activity in the form of isolated leptons which do not originate from the vector boson candidate and soft hadrons are expected to be low.

The event preselections are further refined by considering the final state topologies of the background processes. The  $V+\text{jets}$  process is the dominant background for all decay channels because it features a true  $W$  or  $Z$  boson and one or more jets. Due to its large cross section and similarity to the signal topology, this is the dominant background for all decay channels. It is not entirely irreducible, as detailed comparisons of the final state reveal differences from the signal in terms of the  $p_T$  distribution of the vector boson, which is typically softer than for signal, and the dijet invariant mass distribution, which falls

more sharply. In addition, the majority of these jets are light jets which can be rejected by selecting Higgs jet candidates from those jets which pass the DeepCSV working points.

The  $t\bar{t}$ +jets background process is a dominant background for the  $W(\ell\nu)H$  channel. The preferential decay of the top quark to a  $W$  boson and a  $b$ -quark results in a final state with up to two genuine  $W$  bosons, either of which may decay leptonically, and at least two  $b$ -jets. However,  $t\bar{t}$  background events typically present a higher jet multiplicity, given that one of the  $W$  bosons may decay hadronically, and a broader distribution for the difference in azimuthal angle between the vector boson candidate and dijet system. The single top background process is irreducible but only represents approximately 10-20% of the background composition in the  $W(\ell\nu)H$  channel, and even less for the other decay channels, due to its smaller cross section.

The Standard Model (SM) diboson production ( $WW$ ,  $WZ$ ,  $ZZ$ ) processes are also important to consider. Because one of the vector bosons can decay leptonically and the other hadronically, their final state topologies mimic that of the signal processes and produces a peak in the dijet invariant mass distribution that is within a few standard deviations of the Higgs mass peak. As this is another irreducible background, diboson events can only be rejected by having a good dijet invariant mass resolution that enables the separation of the diboson and Higgs mass peaks. Incidentally, the properties that make the diboson background difficult to handle also enable it to serve as a “standard candle” for the analysis by providing a well-understood SM process with which to validate the analysis strategy.

#### 4.3.1 Signal Regions

For each decay channel, a signal region is defined based on a selection that reduces the background contamination. The simulated events in the signal region are reserved for the training and validation of the classifier model used to discriminate signal from background, as well as the final statistical analysis, while the data events in the signal

region remain blinded until all validation checks have been satisfied. The cuts used to define the signal regions of each decay channel are listed in Table 4-11.

Table 4-11. The cuts defining the signal regions of the decay channels for the  $VH(b\bar{b})$  analysis. The values of the kinematic quantities are quoted in GeV and the values of the differences in azimuthal angle are quoted in radians. For the  $W(\ell\nu)H$  column, the two values for  $p_T^\ell$  and Tightened Lepton Isolation are applicable for the  $W(\mu\nu)H$  and  $W(e\nu)H$  channels, respectively. For the  $Z(\ell\bar{\ell})H$  column, the values for  $p_T(V)$  denote the low and high  $p_T(V)$  analysis categories, respective.

Variable	$Z(\nu\bar{\nu})H$	$W(\ell\nu)H$	$Z(\ell\bar{\ell})H$
$p_T(V)$	$> 170$	$> 150$	$[50, 150], > 150$
$m(\ell\ell)$	—	—	$[75, 105]$
$p_T^\ell$	—	$> 25, > 30$	$> 20$
$p_T(jj)$	$> 120$	$> 100$	—
$m(jj)$	$[60, 160]$	$[90, 150]$	$[90, 150]$
$p_T(j)_{\max}$	$> 60$	$> 25$	$> 20$
$p_T(j)_{\min}$	$> 35$	$> 25$	$> 20$
$b\text{-tag}_{\max}$	$>$ Tight	$>$ Tight	$>$ Loose
$b\text{-tag}_{\min}$	$>$ Loose	$>$ Loose	$>$ Loose
$N_j^{\text{add}}$	$< 2$	$< 2$	—
$N_\ell^{\text{add}}$	$= 0$	$= 0$	—
$p_T^{\text{miss}}$	$> 170$	—	—
Anti-QCD	Yes	—	—
$ \Delta\phi(V, H) $	$> 2.0$	$> 2.5$	$> 2.5$
$ \Delta\phi(p_T^{\text{miss}}, p_T^{\text{miss,track}}) $	$< 0.5$	—	—
$ \Delta\phi(p_T^{\text{miss}}, \ell) $	—	$< 2.0$	—
Lepton Isolation	—	0.06, 0.06	—

The analysis of the  $Z(\nu\bar{\nu})H$  channel is performed in the region of phase space where the transverse momentum of the vector boson  $p_T(V)$ , which is equivalent to the  $p_T^{\text{miss}}$  for this signal process, is greater than 170 GeV due to concerns of data to simulation disagreement in the regions in and below the region of the trigger turn-on. This high MET requirement also helps to “boost” the events considered by the analysis, which offers a better signal-to-background ratio because the  $p_T(V)$  distribution is harder for signal than it is for background. The boosted topology is further reinforced by requiring the transverse momentum of the dijet system  $p_T(jj)$  to be greater than 120 GeV and the difference in azimuthal angle between the MET and the dijet system  $|\Delta\phi(V, H)|$  to be greater than 2.

The dijet invariant mass  $m(jj)$  window is chosen to range from 60 GeV to 160 GeV, which is wider than for the other decay channels in order to recover background statistics lost by the aggressive thresholds placed on the transverse momenta of the Higgs jet candidates. For the Higgs jet candidate pairs, the  $b$ -jet with the greater DeepCSV score is denoted by  $b\text{-tag}_{\max}$  and the lesser by  $b\text{-tag}_{\min}$ , while the larger  $b$ -jet transverse momentum is denoted by  $p_{\mathrm{T}}(j)_{\max}$  and the smaller by  $p_{\mathrm{T}}(j)_{\min}$ . The  $Z(\nu\bar{\nu})H$  analysis places large, asymmetric cuts on  $p_{\mathrm{T}}(j)_{\max}$  and  $p_{\mathrm{T}}(j)_{\min}$  because of an observed reduction of multijet background.

The number of additional jets  $N_j^{\mathrm{add}}$  is at most one and the number of additional leptons  $N_{\ell}^{\mathrm{add}}$  must be zero in order to reject  $t\bar{t}$  and multijet backgrounds which have large jet multiplicities and leptons in the final state. The difference in azimuthal angle between the MET and tracker MET  $|\Delta\phi(p_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss,track}})|$  is chosen to be less than 0.5 as the direction of the tracker MET is expected to be close to the direction of true MET, given that the  $Z$  decay does not produce charged leptons. Finally, the Anti-QCD cut is used to strongly reject the multijet background in the  $Z(\nu\bar{\nu})H$  channel while maintaining high signal efficiency. The cut is placed on the minimum of the set of differences in azimuthal angle between each jet in an event with  $p_{\mathrm{T}} > 30$  GeV and the MET. By requiring this minimum difference to be greater than 0.5, events with fake MET produced by the mismeasurement of jet  $p_{\mathrm{T}}$  are removed from the analysis, with approximately 93% multijet background rejection at 96% signal efficiency in simulation. The remaining variables which involve leptons are not applicable to the  $Z(\nu\bar{\nu})H$  channel.

### 4.3.2 Background Control Regions

For each decay channel, several control regions are defined by inverting and modifying signal region cuts to produce mutually exclusive samples that are enriched in specific background processes and orthogonal to the signal region. The control regions are used by the analysis to validate the agreement of MC simulation to data and to determine the normalization of key background processes with minimal extrapolation uncertainties

during the final statistical analysis. A diagram summarizing the design of the control regions is show in Figure 4-10.

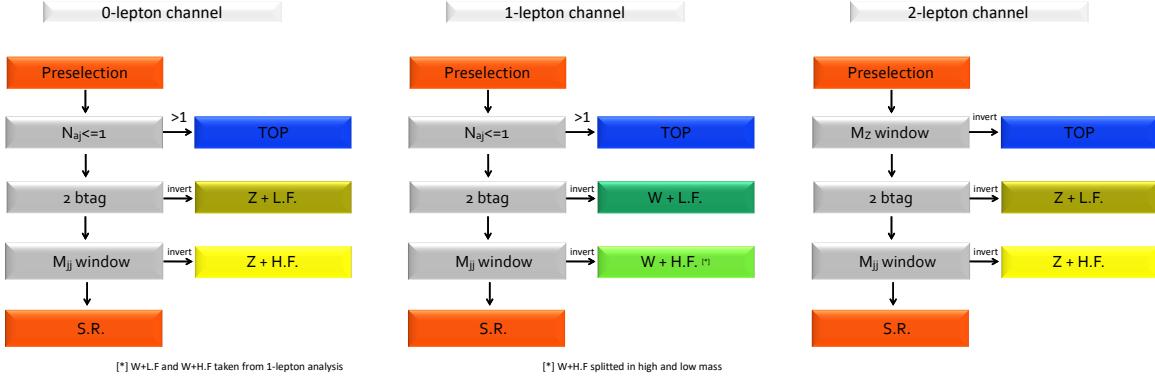


Figure 4-10. An infographic summarizing the inverted cuts used by each decay channel to define their respective control regions.

The detailed cuts which define the control regions of the  $Z(\nu\bar{\nu})H$  channel are listed in Table 4-12. The  $Z+$ heavy process, where  $Z+$ jets events produce at least one  $b$ -jet, is the background process most similar to the  $Z(\nu\bar{\nu})H$  signal. In order to create a  $Z+$ heavy enriched control region, the dijet invariant mass window required for the signal region is simply inverted such that the control region focuses on the mass sidebands. Because  $t\bar{t}$  is an important background process that is also present in the  $Z+$ heavy control region with relatively large contributions, a dedicated control region enriched in the  $t\bar{t}$  process is defined by loosening the signal region's Anti-QCD cut threshold to  $\pi/2$ , loosening the DeepCSV working point to Medium for  $b\text{-tag}_{\max}$ , and requiring at least two additional jets with  $p_T > 30$  GeV and at least one additional lepton. The requirement of at least one additional lepton can result in overlap with the  $t\bar{t}$  control region for the  $W(\ell\nu)H$  channel, so care is taken to ensure that these two control regions remain distinct. A control region enriched in the  $Z+$ light process, where  $Z+$ jets events produce only light flavored jets, is defined by removing the dijet invariant mass window and inverting the DeepCSV working point cut placed on  $b\text{-tag}_{\max}$ , creating a fairly pure control region that helps to constrain the normalization of the background process.

The detailed cuts which define the control regions of the  $W(\ell\nu)H$  and  $Z(\ell\bar{\ell})H$  channels are listed in Table 4-13 and Table 4-14, respectively. Distributions of important variables in all of the control regions of the different decay channels are found in Appendix B. Because the  $W+$ jets processes can have large relative contributions in the  $Z+$ light and  $Z+$ heavy control regions of the  $Z(\nu\bar{\nu})H$  channel, the  $W+$ light and  $W+$ heavy control regions of the  $W(\ell\nu)H$  channel are adopted by the  $Z(\nu\bar{\nu})H$  channel to constrain the normalization of those processes.

Table 4-12. The cuts defining the control regions of the  $Z(\nu\bar{\nu})H$  decay channel for the  $VH(b\bar{b})$  analysis. The values of the kinematic quantities are quoted in GeV and the values of the differences in azimuthal angle are quoted in radians.

Variable	$t\bar{t}$	$Z+\text{light}$	$Z+\text{heavy}$
VDecay Type	$W(\mu\nu)$	$Z(\nu\bar{\nu})$	$Z(\nu\bar{\nu})$
$p_T^{\text{miss}}$	$> 170$	$> 170$	$> 170$
$m(jj)$	—	—	$\notin [60, 160]$
$p_T(jj)$	$> 120$	$> 120$	$> 120$
$b\text{-tag}_{\max}$	$>$ Medium	$<$ Medium	$>$ Tight
$b\text{-tag}_{\min}$	$>$ Loose	$>$ Loose	$>$ Loose
$p_T(j)_{\max}$	$> 60$	$> 60$	$> 60$
$p_T(j)_{\min}$	$> 35$	$> 35$	$> 35$
$ \Delta\phi(V, H) $	$> 2.0$	$> 2.0$	$> 2.0$
$N_{\ell}^{\text{add}}$	$\geq 1$	$= 0$	$= 0$
$N_j^{\text{add}}$	$\geq 2$	$\leq 1$	$\leq 1$
$ \Delta\phi(p_T^{\text{miss}}, p_T^{\text{miss,track}}) $	—	$< 0.5$	$< 0.5$
Anti-QCD	Modified	Yes	Yes

#### 4.4 Binned Multivariate Shape Analysis

In order to observe the  $VH(b\bar{b})$  decay, the distributions considered by the signal extraction fit must enable the hypothesized  $VH(b\bar{b})$  signal to be distinguished from the SM background. Although the dijet invariant mass distribution is a powerful physical observable on its own, statistical learning methods can leverage the information spread across multiple key variables to further separate the signal from background. Because multivariate discriminants can achieve signal-to-background ratios which exceed those of individual physical observables, a multivariate approach is employed by this analysis to maximize the sensitivity to the signal process.

Table 4-13. The cuts defining the control regions of the  $W(e\nu)H$  and  $W(\mu\nu)H$  decay channels for the  $VH(b\bar{b})$  analysis. The values of the kinematic quantities are quoted in GeV and the values of the differences in azimuthal angle are quoted in radians. The  $p_T^{\text{miss}}$ Significance is defined as the ratio of  $p_T^{\text{miss}}$  to the scalar sum of the transverse energy of all particle-flow objects in the event.

Variable	$t\bar{t}$	$W+\text{light}$	$W+\text{heavy}$
VDecay Type	$W(\mu\nu)$	$W(\mu\nu)$	$W(\mu\nu)$
$p_T(V)$	$> 150$	$> 150$	$> 150$
$ \Delta\phi(p_T^{\text{miss}}, \ell) $	$< 2.0$	$< 2.0$	$< 2.0$
$m(jj)$	$< 250$	$< 250$	$\notin [90, 150]$
$p_T(jj)$	$> 100$	$> 100$	$> 100$
$b\text{-tag}_{\text{max}}$	$> \text{Tight}$	$< \text{Medium}$	$> \text{Tight}$
$b\text{-tag}_{\text{min}}$	$> \text{Loose}$	$> \text{Loose}$	$> \text{Loose}$
$p_T(j)_{\text{max}}$	$> 25$	$> 25$	$> 25$
$p_T(j)_{\text{min}}$	$> 25$	$> 25$	$> 25$
$ \Delta\phi(V, H) $	$> 2.5$	$> 2.5$	$> 2.5$
$N_{\ell}^{\text{add}}$	$\geq 1$	—	$= 0$
$N_j^{\text{add}}$	—	$\geq 1$	$\leq 1$
$ \Delta\phi(p_T^{\text{miss}}, p_T^{\text{miss,track}}) $	—	$< 0.5$	$< 0.5$
$p_T^{\text{miss}}$ Significance	—	$> 2.0$	$> 2.0$

Table 4-14. The cuts defining the control regions of the  $Z(e\bar{e})H$  and  $Z(\mu\bar{\mu})H$  decay channels for the  $VH(b\bar{b})$  analysis. The values of the kinematic quantities are quoted in GeV and the values of the differences in azimuthal angle are quoted in radians.

Variable	$t\bar{t}$	$Z+\text{light}$	$Z+\text{heavy}$
VDecay Type	$Z(\ell\ell)$	$Z(\ell\ell)$	$Z(\ell\ell)$
$p_T(V)$	$[50, 150], > 150$	$[50, 150], > 150$	$[50, 150], > 150$
$m(\ell\ell)$	$\notin [0, 10], \notin [75, 120]$	$[75, 105]$	$[85, 97]$
$m(jj)$	—	$[90, 150]$	$\notin [90, 150]$
$b\text{-tag}_{\text{max}}$	$> \text{Tight}$	$< \text{Loose}$	$> \text{Tight}$
$b\text{-tag}_{\text{min}}$	$> \text{Loose}$	$< \text{Loose}$	$> \text{Loose}$
$p_T(j)_{\text{max}}$	$> 20$	$> 20$	$> 20$
$p_T(j)_{\text{min}}$	$> 20$	$> 20$	$> 20$
$ \Delta\phi(V, H) $	—	$> 2.5$	$> 2.5$
$p_T^{\text{miss}}$	—	—	$< 60$

#### 4.4.1 Multivariate Discriminator

For this analysis, the multivariate discriminator is a classifier with a single, continuous valued output that scores events based on how “signal-like” they appear, with the convention that more “signal-like” events receive higher scores. Because the final state topologies and background contributions are specific to each decay channel, separate instances of the classifier are trained for the  $Z(\nu\bar{\nu})H$ ,  $W(\ell\nu)H$ , and the low and high  $p_T(V)$   $Z(\ell\bar{\ell})H$  categories, where the training may be inclusive or exclusive for the  $e$  and  $\mu$  final states. The discriminant distributions of these distinct models are used for the final signal extraction.

The classifiers are trained in a supervised manner using the MC simulated events which pass the signal region selection of their respective decay channel. The simulated events for the  $VH(b\bar{b})$  processes are labelled as the signal class and all other processes are labelled as the background class. For each decay channel, the simulated events are further divided into a training set and a validation set based on the parity of their generator level ID values to obtain two statistically independent samples which are roughly equal in size, typically around 200,000 events. The training set is used solely to train the classifier while the validation set is used to evaluate the generalization performance of the model and tune the hyperparameters of the machine learning algorithm. Because of the limited statistics available for the 2017 MC samples, the classifiers are trained and validated using the 2016 MC samples while the 2017 MC samples are used to provide the signal region discriminant distribution considered by the signal extraction fit.

The variables which are used as the classifier input features are listed in Table 4-15. The final set of input features were selected using a “leave-one-out” strategy, where an initial subset of variables chosen based on their physics efficacy for each channel are further refined by iteratively holding out one variable while training the classifier with the rest. By evaluating the performance of each of these classifiers on the validation set, an optimal set of input features is determined. Because the training and validation sets

are derived from 2016 MC samples but the 2017 MC samples and data are used for the signal extraction fit, a lack of agreement between the data and simulated samples in the distributions of the input variables could jeopardize the generalizability of the classifier. It is found that the distributions of the input features remain fairly consistent between the 2016 and 2017 MC samples.

Table 4-15. The variables which define the input features used to train the event classifier models for the multivariate discriminant analysis. The values of the kinematic quantities are quoted in GeV and the values of the differences in azimuthal angle are quoted in radians.

Variable	Description	$Z(\nu\bar{\nu})H$	$W(\ell\nu)H$	$Z(\ell\ell)H$
$m(jj)$	dijet invariant mass	✓	✓	✓
$p_T(jj)$	dijet transverse momentum	✓	✓	✓
$p_T(j)_{\max}, p_T(j)_{\min}$	transverse momentum of the two $b$ -jets	✓		✓
$\Delta R(jj)$	distance between the two $b$ -jets in $\eta\phi$ space			✓
$\Delta\eta(jj)$	difference in $\eta$ between the two $b$ -jets	✓		✓
$\Delta\phi(jj)$	difference in $\phi$ between the two $b$ -jets	✓		
$p_T(V)$	vector boson transverse momentum		✓	✓
$\Delta\phi(V, H)$	difference in $\phi$ between the dijet system and vector boson	✓	✓	✓
$p_T(jj)/p_T(V)$	$p_T$ ratio of the dijet system to the vector boson			✓
$m(Z)$	reconstructed mass of the $Z$ boson			✓
$b\text{-tag}_{\max}$	leading DeepCSV score of the two $b$ -jets	✓		✓
$b\text{-tag}_{\min}$	subleading DeepCSV score of the two $b$ -jets	✓	✓	✓
$b\text{-tag}_{\text{add}}$	leading DeepCSV score of all additional jets	✓		
$p_T^{\text{miss}}$	missing transverse energy	✓	✓	✓
$ \Delta\phi(p_T^{\text{miss}}, j) $	difference in $\phi$ between $p_T^{\text{miss}}$ and the nearest jet with $p_T > 30$ GeV	✓		
$ \Delta\phi(p_T^{\text{miss}}, \ell) $	difference in $\phi$ between $p_T^{\text{miss}}$ and the isolated lepton		✓	
$m_T$	mass of the isolated lepton $\vec{p}_T + \vec{p}_T^{\text{miss}}$			✓
$m_t$	reconstructed mass of the top quark			✓
$N_j^{\text{add}}$	number of additional jets		✓	✓
$p_T(j)_{\text{add}}$	highest $p_T$ of all additional jets	✓		
SA5	number of soft-track jets with $p_T > 5$ GeV	✓	✓	✓

Two different machine learning algorithms were considered for the final multivariate discriminant. Gradient boosted decision tree (BDT)[111] classifiers implemented using the

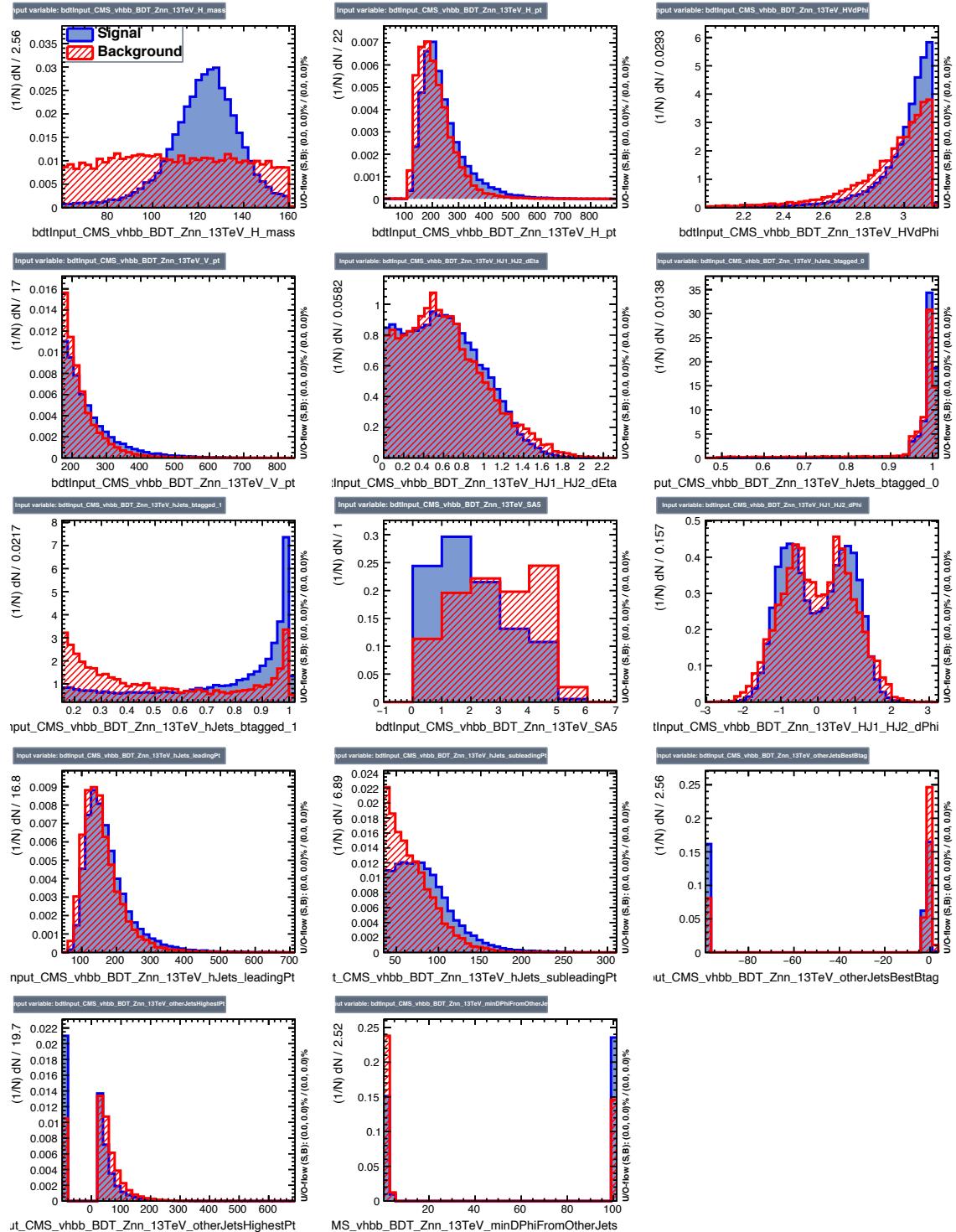


Figure 4-11. The distributions of select input features for the  $Z(\nu\bar{\nu})H$  decay channel. In order to emphasize the separation between signal (blue) and background (red), the distributions are normalized.

`TMVA` v4.2.1[112] software package were used to establish a performance baseline, given their broad acceptance by the high energy physics community because of the relative interpretability of their results and good “out-of-the-box” performance with little to no input feature preprocessing or hyperparameter tuning. Multilayer perceptron[113] or artificial neural network (ANN) classifiers implemented using the `Keras` neural network framework with `Tensorflow` as the computational backend were compared to the incumbent BDT classifiers to determine if recent advances in the implementation and training of ANN classifiers resulted in superior performance.

Optimal BDT and ANN classification models were selected using a grid-based hyperparameter search with the area under the receiver operating characteristic (ROC) curve evaluated on the validation set as the performance metric. Because overfitting can result in models which fail to generalize because they are sensitive to fine variations in the training set, the Kolmogorov-Smirnov test statistic was used to estimate the compatibility of the discriminant distribution of the training set to the validation set. The compatibility of the normalized shapes of the discriminant distributions for the training and validations sets were also visually inspected.

The performance of the optimized BDT and ANN classifiers were found to be nearly equivalent, with both models achieving similar area under the ROC curve values near 0.80 when evaluated on the validation set. The ANN classifier was chosen as the final discriminant for the analysis. The chosen neural network architecture, dubbed “DNN” for deep neural network, takes as input the training features of a specific decay channel, each scaled to zero mean and unit variance, and passes them through five hidden layers with 32 nodes in each layer. The hidden layer nodes are Leaky ReLU units with a slope of  $\beta = 0.2$  for the negative part of its argument. Residual connections[114], which have been shown to benefit the training of deep neural networks by mitigating the problem of vanishing or exploding gradients during training, were made between the first, second, and third hidden layers and the last hidden layer and between the second and fourth

hidden layers. Batch-normalization and dropout with a probability of 0.1 are applied to the outputs of each hidden layer to increase the training speed and regularize the model. The output layer consists of two linear nodes passed through a softmax activation function. A softmax output layer was chosen because the output values are interpreted to be the probability that an event belongs to one of the two possible classes of signal or background. The DNN was trained to minimize the cross entropy loss using stochastic gradient descent with adaptive moment estimation, or the Adam optimizer. The output node of the DNN chosen to represent the probability that an event belongs to the signal class defines a discriminant whose values fall within the inclusive range  $[0, 1]$ , where events with lower scores are more background-like and events with higher scores are more signal-like. The normalized discriminant distributions of each of the trained DNN classifiers are shown in Figure 4-12.

Because the  $Z$ +heavy control region of the  $Z(\nu\bar{\nu})H$  channel and the  $W$ +heavy control region of the  $W(\ell\nu)H$  channel are not pure in the enriched background, the normalization of the  $V+b$  and  $V+b\bar{b}$  processes are difficult to constrain. A variation of the DNN is therefore used to enhance the separation of the contributing background processes in the  $V$ +heavy flavor control regions of the  $Z(\nu\bar{\nu})H$  and  $W(\ell\nu)H$  channels. The input features are the same as those used for the signal classification task, except that the training and validation events pass the control region selection of their respective decay channel. The training is carried out as before, with the neural network architecture modified to use a five-way softmax output layer whose values are interpreted to be the probability that an event belongs to one of five possible background classes:  $V+b\bar{b}$ ,  $t\bar{t}$ ,  $V$ +light,  $V+b$ , and single top. The discriminant distributions of these multi-background DNN classifiers are shown in Figure 4-13.

#### 4.4.2 Binned Shape Analysis

The final signal extraction is performed using a simultaneous binned maximum likelihood fit of the signal and background control regions of all decay channels and

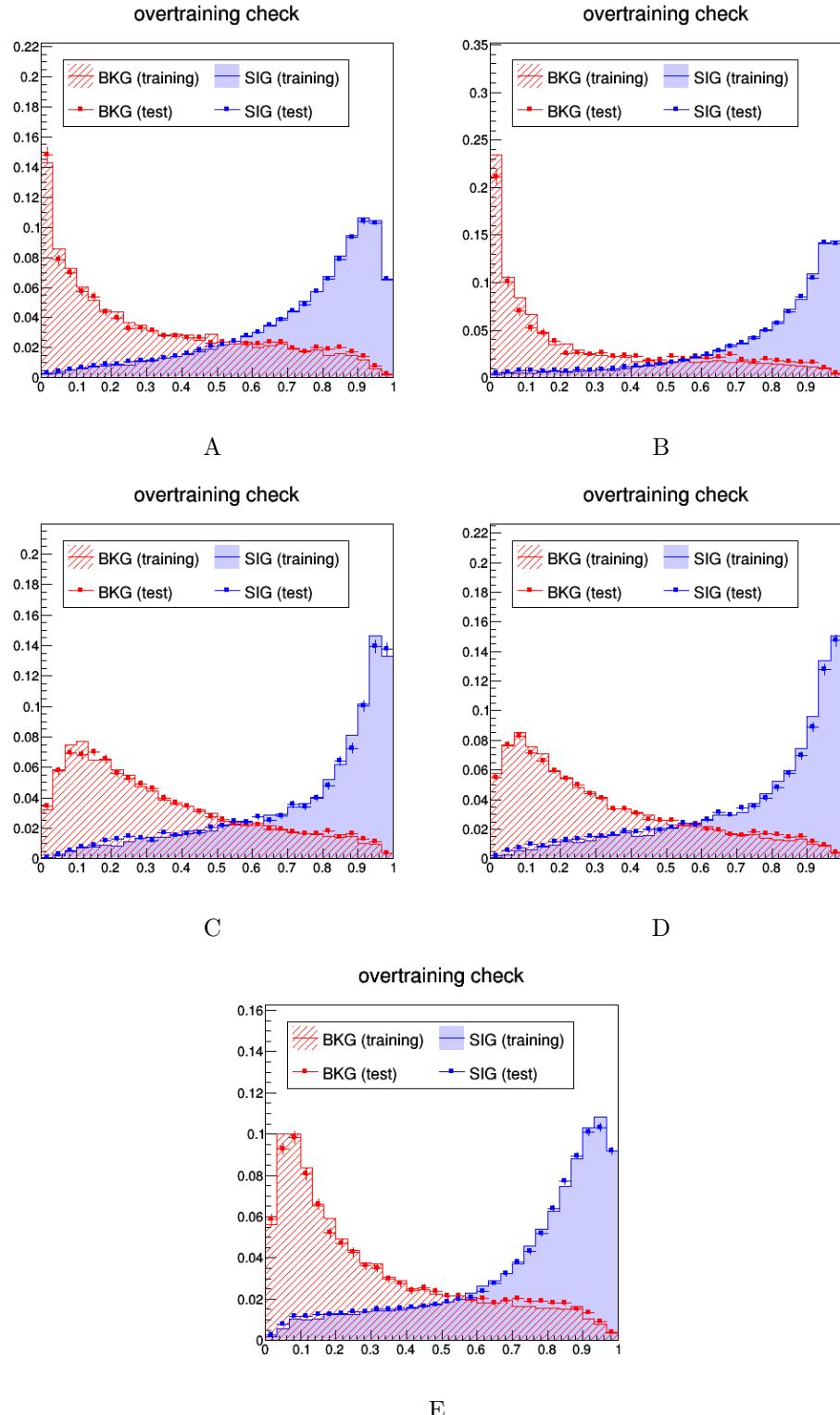


Figure 4-12. The normalized DNN discriminant distributions for the A) low  $p_T(V)$   $Z(\ell\bar{\ell})H$ ; B) high  $p_T(V)$   $Z(\ell\bar{\ell})H$ ; C)  $W(e\nu)H$ ; D)  $W(\mu\nu)H$ ; and E)  $Z(\nu\bar{\nu})H$  channels. The signal distributions are shown for the training (solid blue) and validation (blue points) sets. The background distributions are also shown for the training (hashed red) and validation (red points) sets.

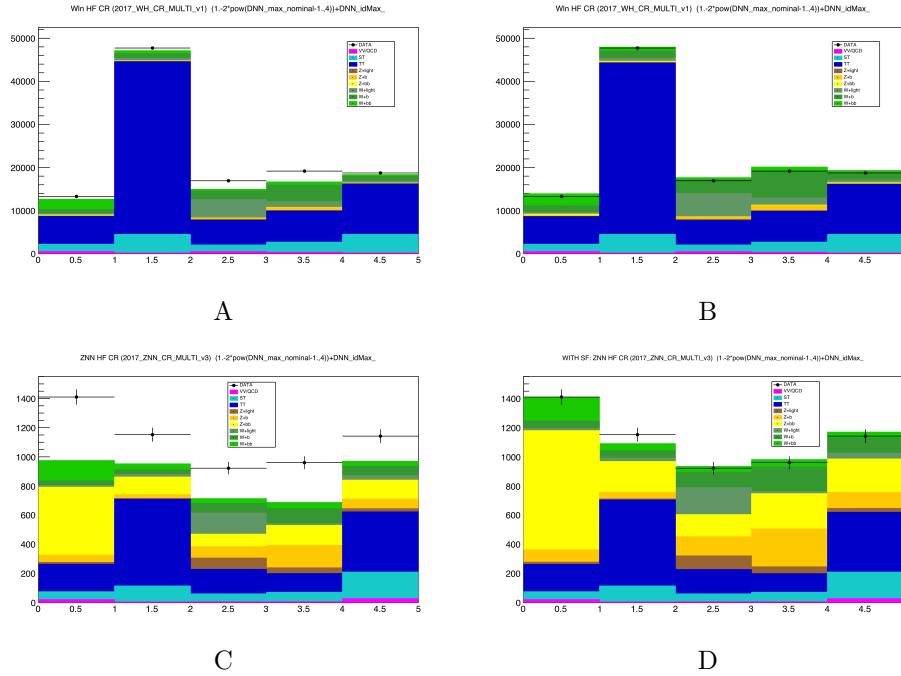


Figure 4-13. The multi-background DNN discriminant distributions for the  $W+\text{heavy}$  control region of the  $W(\ell\nu)H$  channel (top row) and  $Z+\text{heavy}$  control region of the  $Z(\nu\bar{\nu})H$  channel (bottom row) both A), C) before and B), D) after the background process normalizations have been adjusted.

analysis regions. The fitted distributions for the signal regions are the discriminant distributions derived from the DNN signal classifiers. The fitted distributions for the  $W+\text{heavy}$  control region of the  $W(\ell\nu)H$  channel and  $Z+\text{heavy}$  control region of the  $Z(\nu\bar{\nu})H$  channel are the discriminant distributions derived from their dedicated multi-background DNN classifiers. The fitted distributions for the remaining background control regions are the  $b\text{-tag}_{\min}$  distributions derived from the subleading DeepCSV score of the pair of  $b$ -jet candidates.

Whereas the signal region distributions are chosen so as to maximize the signal-to-background ratio to improve the sensitivity to signal, the background control region distributions are chosen based on their separation of the individual background components to improve the estimation of the background normalizations. The simultaneous binned fit treats the normalizations of the  $t\bar{t}$ ,  $W+\text{light}$ ,  $W+\text{heavy}$ ,  $Z+\text{light}$ , and  $Z+\text{heavy}$  processes as free

parameters, with the  $V$ +heavy processes further distinguished between its  $V+b$  and  $V+b\bar{b}$  components. The fit determines the necessary *scale factors* which are used to scale the MC simulated yield of those background components while taking into account the variations in template shape due to systematic sources.

The simultaneous binned fit extracts the signal strength  $\mu$ , which is defined as the ratio of the measured production cross section multiplied by the branching fraction to the expected SM value for the  $VH(b\bar{b})$  process. The discovery significance of any observed excess of events over the expected SM background in the fit is computed using the profile likelihood asymptotic approximation.[115–118]

#### 4.4.3 $VZ(b\bar{b})$ Cross-check Analysis

In order to validate the analysis strategy, a cross-check analysis is performed by applying the same analysis strategy to search for the  $VZ(b\bar{b})$  decay, a well-understood SM process, with minimal modifications. The first necessary change to the analysis strategy is the widening of the dijet invariant mass windows in the signal regions, along with the corresponding changes for the  $V$ +heavy background control regions, of the  $W(\ell\nu)H$  and  $Z(\ell\bar{\ell})H$  decay channels by changing the lower boundary from 90 GeV to 60 GeV, such that the diboson mass peak can be fully resolved given that the mass of the  $Z$  boson is approximately 91.2 GeV. The second necessary change is the retraining of the multivariate discriminators on redefined signal and background classes, where the diboson processes  $WZ(b\bar{b})$  and  $ZZ(b\bar{b})$  are now treated as signal and the  $VH(b\bar{b})$  processes are treated as background. Once the multivariate discriminators have been retrained, a simultaneous binned maximum likelihood fit of the signal and background control regions of all decay channels and analysis regions is performed to extract the signal strength of the  $VZ(b\bar{b})$  process and measure the discovery significance of any observed excess of events over the expected SM background.

#### 4.4.4 Dijet Invariant Mass Cross-check Analysis

The dijet invariant mass distribution is the most powerful individual observable because it features a  $VH(b\bar{b})$  signal mass peak which is easily discernable from a flat background distribution. Although a dijet invariant mass analysis is typically 5-20% less sensitive to the signal than a multivariate analysis depending on the decay channel, it remains an invaluable cross-check analysis because it validates the robustness of the multivariate approach and provides a physically interpretable visual fit result in which the  $Z \rightarrow b\bar{b}$  and  $H \rightarrow b\bar{b}$  mass peaks are clearly resolved in a widened dijet invariant mass window.

The main difficulty in designing a mass analysis is the definition of the signal region, as both cut-based and multivariate approaches can achieve high signal purity at the expense of introducing distortions of the background mass shapes. Because the choice of cuts or multivariate discriminator scores are highly correlated with the mass, the background mass distributions can become artificially “sculpted” into a peak which reduces the discovery significance of the signal events. This suboptimal behavior poses no issue for the multivariate analysis, which fits directly the discriminant distribution based on a classifier that has been trained with the dijet invariant mass, at the expense of a physically interpretable result.

An alternative strategy implemented by this analysis is the mitigation of background sculpting by decorrelating the discriminator score from the mass. The decorrelation strategy employed by this analysis consisted of reevaluating the DNN classifiers over the training and test sets while marginalizing over the mass information of the input features and without retraining. For the set of input features for each decay channel, those features most correlated with the mass were determined to be those which exceeded a threshold of 0.10 for the magnitude of their Spearman rank correlation coefficient evaluated with respect to the mass. The values of those features, along with the mass, were then fixed to the mean values of their corresponding distribution over the background events contained

in the training set of their respective decay channel, which are listed in Table 4-16. The output of this procedure produces a “massless” DNN score for each event. While this strategy was not extensively optimized, it was proven to reduce the sculpting of the background mass shapes to negligible levels, as illustrated in Figure 4-14 for the  $Z(\nu\bar{\nu})H$  channel.

Table 4-16. The input features used by the multivariate discriminant analysis which are most correlated with the dijet invariant mass. For each decay channel, the mean background values evaluated over the training set are reported for the relevant features. The values of the kinematic quantities are quoted in GeV.

Variable	$Z(\nu\bar{\nu})H$	$W(e\nu)H$	$W(\mu\nu)H$	$Z(\ell\ell)H$ , Low $p_T(V)$	$Z(\ell\ell)H$ , High $p_T(V)$
$m(jj)$	110.80	121.00	120.10	117.50	118.40
$\sigma_m(jj)$	-	-	-	7.00	21.30
$\Delta\eta(jj)$	0.62	0.75	0.76	1.36	0.93
$p_T(jj)$	206.30	-	-	-	-
$p_T(j)_{\text{max}}$	154.30	142.40	140.50	85.70	172.90
$p_T(j)_{\text{min}}$	68.30	59.50	59.00	41.30	55.30
$p_T(V)$	220.60	-	-	-	-

Given that the massless DNN score is by construction decorrelated from the mass, it may be used to define event categories for which the dijet invariant mass distributions have increased signal-to-background ratios without biasing the distribution. The AUTOCATEGORIZER[119] software package was used to define the event categories. The AUTOCATEGORIZER implements a greedy tree-based optimization algorithm that seeks to maximize the expected discovery significance by splitting events into distinct categories based on the provided categorization variables, evaluating the approximate median discovery significance[120] for a probability density function (PDF) of interest in each category, and choosing the category with the highest significance at each step. The final number of categories is determined by the maximum number of terminal leaf nodes allowed for the categorization tree. In the context of this analysis, the categorization variable was the massless DNN score and the PDF of interest was the dijet invariant mass distribution. The optimal number of mass categories was found to be four based on separate grid search optimizations for each decay channel.

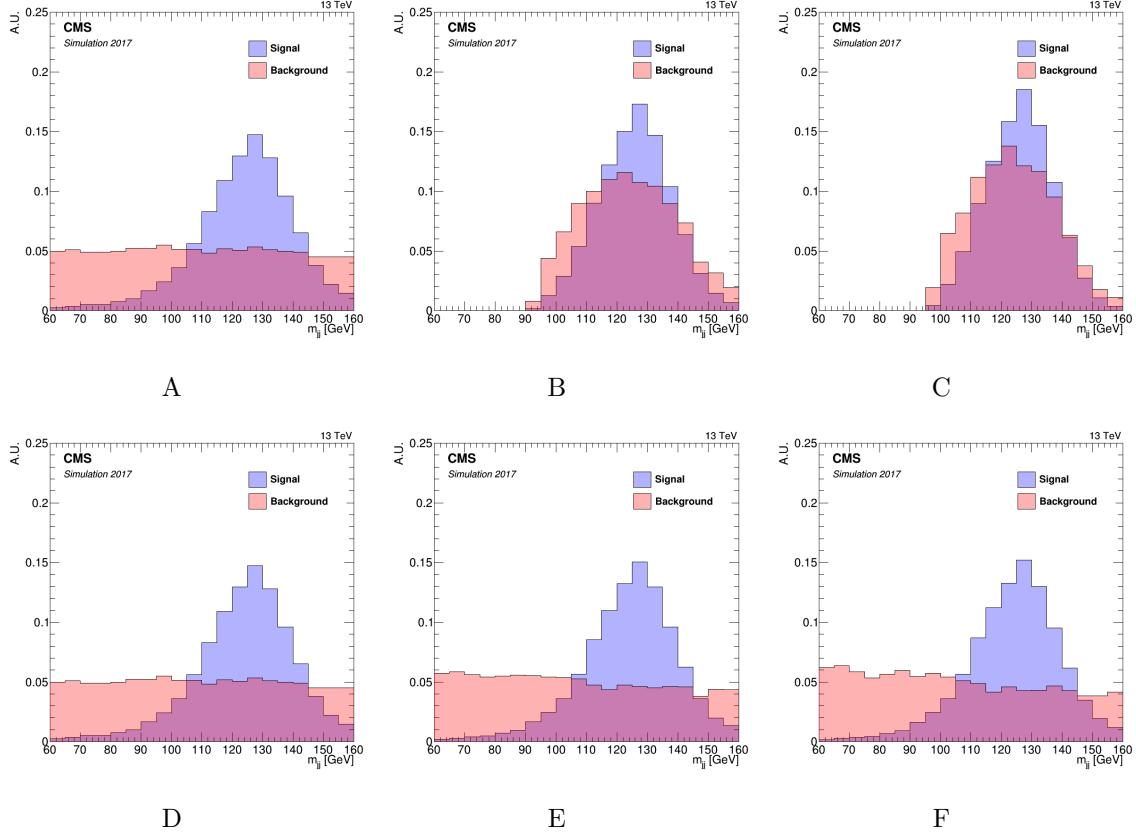


Figure 4-14. The normalized signal (blue) and background (red) dijet invariant mass distributions in the signal region of the  $Z(\nu\bar{\nu})H$  channel. The top row shows the dijet invariant mass distribution for A) all events; B) events with DNN score  $> 0.5$ ; C) events with DNN score  $> 0.65$ . The bottom row shows the dijet invariant mass distribution for D) all events; E) events with massless DNN score  $> 0.5$ ; F) events with massless DNN score  $> 0.65$ .

#### 4.5 Systematic Uncertainties

The results of the analysis depend on the expected yields of the signal and background processes and the shapes of the distributions participating in the maximum likelihood fit. The final fit must therefore account for those systematic uncertainties which affect the normalizations and shapes of these processes. The following experimental and theoretical sources of systematic uncertainty are included in the fit as independent nuisance parameters.

- **Luminosity:** the total systematic uncertainty of the 2017 luminosity measurement by the CMS experiment is 2.3%.[\[121\]](#)

- **MET+jets Trigger:** the variation of the parameters describing the efficiency curve of the MET triggers used by the  $Z(\nu\bar{\nu})H$  channel according to their statistical uncertainties results in an estimated systematic uncertainty of 1%.
- **Lepton Efficiencies:** the electron and muon triggers and their reconstruction and identification efficiencies are determined in data using the tag-and-probe technique with  $Z$  bosons and the statistical uncertainties of these efficiencies are used to estimate the systematic uncertainty which is found to be at most 2%.
- **Jet Energy Scale:** the energy scale of each jet is varied up and down within one standard deviation for each of 27 independent sources of systematic uncertainty prescribed by the JetMET POG. Because the corresponding nuisance parameters of the jet energy scale uncertainties were prone to over-constraining in the signal extraction fit, the assumption that the systematics are fully correlated over the entire phase space was relaxed and the systematics decorrelated across five bins of  $p_T$  and three bins of  $\eta$ .
- **Jet Energy Resolution:** the energy resolution of each  $b$ -jet, after the  $b$ -jet energy regression has been applied, is smeared by the 10% systematic uncertainty prescribed by the JetMET POG. Because these systematics affect the shape of the multivariate discriminant, the DNN is reevaluated after the relevant inputs have been smeared to define the shape variations.
- **$b$ -Tagging:** an average systematic uncertainty of 6% per  $b$ -jet, 12% per  $c$ -jet, and 15% per light-jet is applied to account for variations in normalization. Because these systematics affect the shape of the multivariate discriminant, the DNN is reevaluated after the  $b$ -tagging inputs have been varied according to the systematic uncertainties prescribed by the BTV POG to define the shape variations. For the same reason as the jet energy scale uncertainties, the  $b$ -tagging systematics are also decorrelated in bins of  $p_T$  and  $\eta$  in the same manner.
- **Unclustered MET:** a systematic uncertainty of 3% for the calibration of unclustered MET, the missing energy associated with particles which were not clustered into jets, is estimated according to the procedure recommended by the JetMET POG.
- **Signal Cross Section:** the total signal cross section calculated to NNLO accuracy has a total uncertainty of 4%.[\[122\]](#)
- **$H \rightarrow b\bar{b}$  Branching Ratio:** a theoretical uncertainty of 0.5% is estimated by the LHC Higgs Cross Section Working Group.[\[13\]](#)
- **Signal  $p_T$  Spectrum:** the NLO electroweak[\[123–125\]](#) and NNLO QCD corrections[\[126\]](#) applied to the Monte-Carlo samples for the signal processes have systematic uncertainties of 2% and 5%, respectively.

- **$W(\ell\nu)H$   $p_T(V)$  Reweighting:** the  $p_T(V)$  reweightings derived by the  $W(\ell\nu)H$  channel, which are described in section 4.1.4, result in systematic uncertainties of 13% for the  $t\bar{t}$  process and 6% for the  $W+udcsg$  and  $W+b\bar{b}$  and single top combined processes.
- **$\Delta\eta(jj)$  Reweighting:** the  $\Delta\eta(jj)$  reweightings, described in section 4.1.4, affect the shape of the multivariate discriminant. The DNN is reevaluated after the  $\Delta\eta(jj)$  reweighting has been varied within its statistical uncertainties to define the shape variations.
- **Background Estimation:** the total uncertainty of the background estimates include contributions from uncertainty sources due to data-driven methods, simulation, and theory. An uncertainty of 15% is assumed for the single top and diboson processes based on the uncertainty of their cross section measurements, while the uncertainty of the remaining background processes are determined by the control region distributions participating in the fit.
- **Monte-Carlo Statistics:** the normalization uncertainties of the signal and background processes account for the finite size of their corresponding MC samples. The shape of the multivariate discriminant is also allowed to vary within the bin-by-bin statistical uncertainties of the MC samples.
- **Parton Distribution Functions:** the uncertainty in the knowledge of the proton quark content is encoded by a set of 101 NNPDF MC replicas. For each process, the root mean square (RMS) of all variations is evaluated for each bin of the DNN distribution and the largest variation is taken to provide the normalization uncertainty. The resulting yield uncertainties are 1% for the  $VH$  signal processes, 5% for the  $V + udcsg$  process, 3% for the  $V + b$  process, 2% for the  $V + b\bar{b}$  process, 3% for the  $VV + udcsg$  process, 2% for the  $VV + b$  and  $VV + b\bar{b}$  processes combined, and 0.5% for the  $t\bar{t}$  process.
- **QCD Renormalization and Factorization Scales:** the perturbative QCD renormalization and factorization scale variations are estimated by halving and doubling their nominal values. These resulting uncertainties are considered separately for each process and taken to be uncorrelated.

## CHAPTER 5

### RESULTS

The principle results reported by this analysis are the discovery significance and an estimate of the signal strength  $\mu$  of the  $VH(b\bar{b})$  decay. The discovery significance is the statistical significance of any observed excess of events in data over the expected SM background. The signal strength is defined as the measured production cross section times the  $H \rightarrow b\bar{b}$  branching fraction divided by the expected SM value or  $\sigma/\sigma_{\text{SM}}$ . The statistical analysis was performed using the COMBINE[127] software package developed by the CMS collaboration which provides a command line interface for the statistical tools implemented in RooFit[128] and RooStats[129].

#### 5.1 Statistical Treatment

The discovery significance is computed using the profile likelihood asymptotic approximation.[116–118, 130] In this approach, the significance is calculated from a profile likelihood ratio where the signal strength is set to zero in the numerator and is allowed to float freely in the denominator. The test statistic defined by the profile likelihood ratio is thus

$$q_0 = -2 \ln \frac{\mathcal{L}(\text{data}|b, \hat{\theta}_0)}{\mathcal{L}(\text{data}|\hat{\mu} \times s + b, \hat{\theta})} \text{ for } \hat{\mu} > 0, \quad (5-1)$$

where data comes from either the observed dataset when calculating the observed significance or a generated toy dataset when calculating the expected significance,  $s$  and  $b$  are respectively the expected number and distribution of signal and background events,  $\hat{\theta}_0$  is the estimated value of the nuisance parameter that maximizes the likelihood of the numerator, and  $\hat{\mu}$  and  $\hat{\theta}$  are the estimated values of the signal strength and nuisance parameter that maximizes the likelihood of the denominator. The value of  $\hat{\mu}$  is restricted to be positive to remain physically meaningful. Assuming that the number of events is large, the distribution of this test statistic may be determined asymptotically using Wilke’s theorem.

The local  $p$ -value is defined as the probability of obtaining a value of  $q_0$  greater than or equal to the value of  $q_0$  observed in data under the background-only hypothesis such that

$$p_0 = P(q_0 \geq q_0^{\text{data}} | b). \quad (5-2)$$

The value of  $p_0$  is therefore a measure of the probability that an upward fluctuation of the background is responsible for an observed excess of events in the absence of signal. Each value of  $p_0$  corresponds to a unique level of significance  $z$  that is computed from the right-sided tail of a standard normal distribution

$$p_0 = \frac{1}{\sqrt{2\pi}} \int_z^{+\infty} e^{-\frac{1}{2}x^2} dx. \quad (5-3)$$

This discovery significance is typically quoted in terms of the number of standard deviations  $\sigma$ , with a threshold of  $3\sigma$  ( $z = 3$ ) required to establish evidence for a decay and a threshold of  $5\sigma$  ( $z = 5$ ) required to establish an observation of a decay.

In the presence of an excess of events, the best fit value of the estimated signal strength  $\hat{\mu}$  is determined by a simultaneous binned maximum likelihood fit over the signal and control regions of all decay channels during which the signal strength  $\hat{\mu}$  is allowed to vary freely and the nuisance parameters  $\hat{\theta}$  vary within their uncertainties. For the  $VH(b\bar{b})$  analysis, the signal extraction fit is performed using the DNN-based multivariate discriminants in the signal regions and the distributions of the subleading DeepCSV score of the two  $b$ -jets  $b\text{-tag}_{\min}$  in the control regions. The signal extraction fit for the  $VZ(b\bar{b})$  cross-check analysis proceeds in the same manner, with the only differences being a wider dijet invariant mass window in the signal region and a DNN classifier trained to identify diboson events as signal. For the dijet invariant mass cross-check analysis, the fit is performed using the dijet invariant mass distributions of the four distinct categories in the signal region of each decay channel while the control regions and their distributions remain the same as in the  $VH(b\bar{b})$  analysis.

## 5.2 $VZ(b\bar{b})$ Analysis

The data to MC scale factors obtained by the maximum likelihood fit are shown in Table 5-1. The post-fit multivariate discriminant distributions of each channel are shown individually in Figure 5-1 and combined in Figure 5-2. The expected post-fit significance of the excess over the background prediction is  $5.0\sigma$ , while the observed significance is  $5.2\sigma$ . The best fit signal strength is  $\mu = 1.05^{+0.22}_{-0.21}$  for all channels combined and the signal strengths of the individual channels, shown in Figure 5-3, are compatible with a  $p$ -value of 64%. These results are in good agreement with past measurements of the  $VZ(b\bar{b})$  process and the observation of this known Standard Model process serves as an important validation of the analysis strategy.

Table 5-1. The data to Monte-Carlo scale factors obtained by the simultaneous maximum likelihood fit over all  $VZ(b\bar{b})$  channels. The errors include both statistical and systematic uncertainties.

Process	$Z(\nu\bar{\nu})H$	$W(\ell\nu)H$	$Z(\ell\ell)H$ , Low $p_T(V)$	$Z(\ell\ell)H$ , High $p_T(V)$
$W + udcsg$	$1.04 \pm 0.01$	$1.04 \pm 0.01$	-	-
$W + b$	$2.02 \pm 0.09$	$2.02 \pm 0.09$	-	-
$W + b\bar{b}$	$2.02 \pm 0.13$	$2.02 \pm 0.13$	-	-
$Z + udcsg$	$0.86 \pm 0.05$	-	$0.88 \pm 0.01$	$0.80 \pm 0.01$
$Z + b$	$1.07 \pm 0.14$	-	$0.89 \pm 0.05$	$1.13 \pm 0.08$
$Z + b\bar{b}$	$1.20 \pm 0.07$	-	$0.84 \pm 0.03$	$0.95 \pm 0.05$
$t\bar{t}$	$0.97 \pm 0.02$	$0.93 \pm 0.01$	$0.88 \pm 0.01$	$0.90 \pm 0.02$

## 5.3 Dijet Invariant Mass Analysis

The optimal massless DNN score values that divide the signal regions of each of the decay channels into four separate mass categories are listed in Table 5-2. The data to MC scale factors obtained by the maximum likelihood fit are shown in Table 5-3. The expected post-fit significance of the excess over the background prediction is  $2.19\sigma$ , while the observed significance is  $1.33\sigma$ . The best fit signal strength is  $\mu = 0.59^{+0.45}_{-0.45}$  for all channels combined. By combining the dijet invariant mass distributions of all channels with weight  $S/(S + B)$ , where  $S$  is the post-fit  $VH(b\bar{b})$  signal yield and  $B$  is the post-fit total background yield specific to each channel, and then subtracting the contributions of

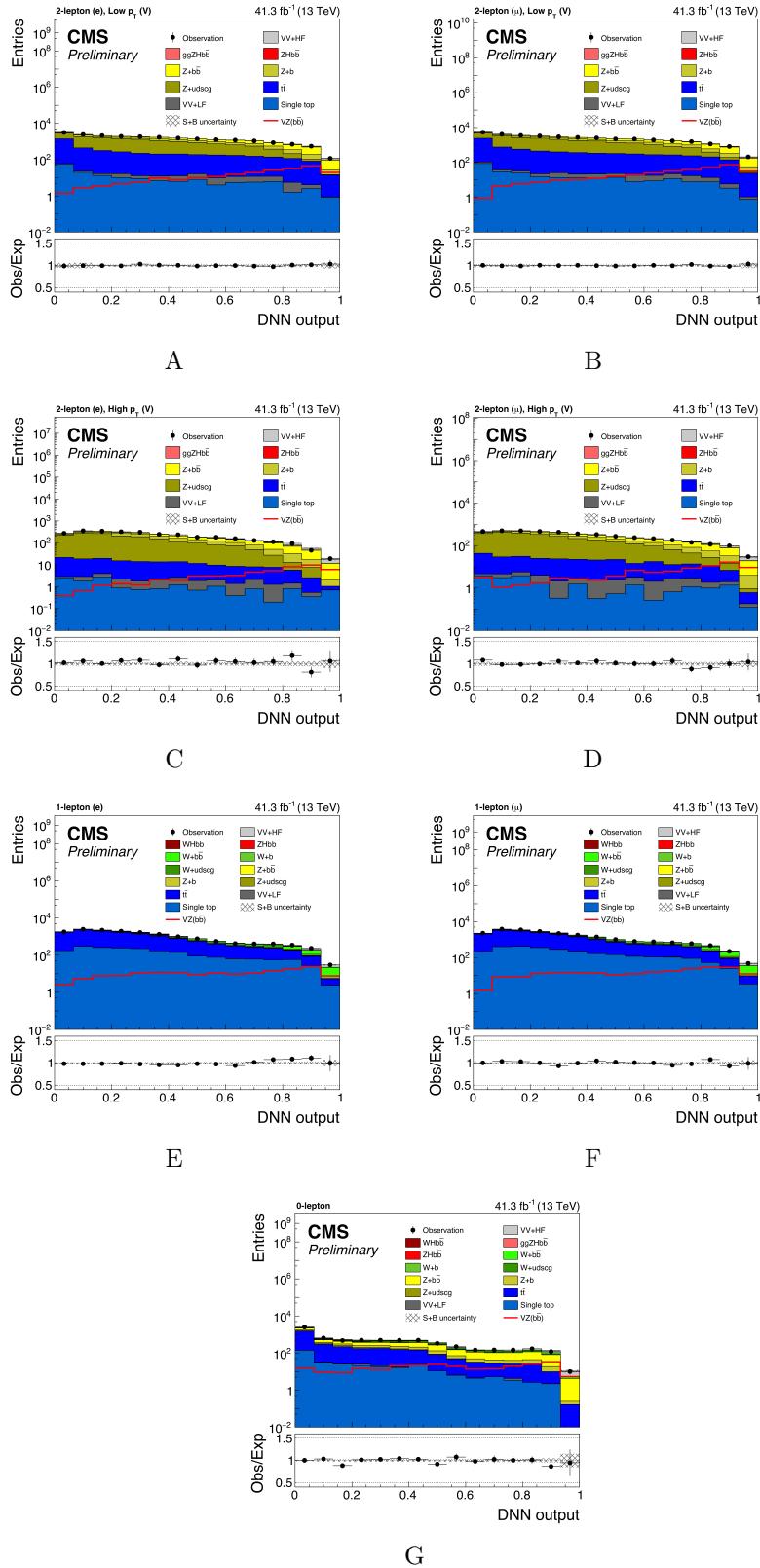


Figure 5-1. The post-fit  $VZ(b\bar{b})$  multivariate discriminant distributions of the A) low  $p_T(V)$   $Z(e\bar{e})H$ ; B) low  $p_T(V)$   $Z(\mu\bar{\mu})H$ ; C) high  $p_T(V)$   $Z(e\bar{e})H$ ; D) high  $p_T(V)$   $Z(\mu\bar{\mu})H$ ; E)  $W(e\nu)H$ ; F)  $W(\mu\nu)H$ ; and G)  $Z(\nu\bar{\nu})H$  channels.

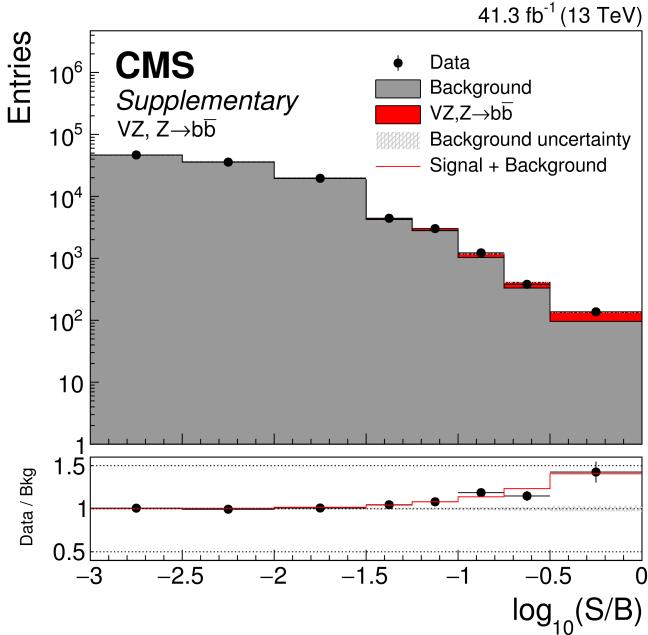


Figure 5-2. The post-fit multivariate discriminant distributions of signal, background, and data event yields combined for all  $VZ(b\bar{b})$  channels and sorted into bins of similar signal-to-background ratio. The  $VZ(b\bar{b})$  signal contribution (red) has been stacked above the sum of all background yields (gray). The bottom panel shows the data to background ratio (black points) and the sum of signal and background contributions divided by the background yield (red line).

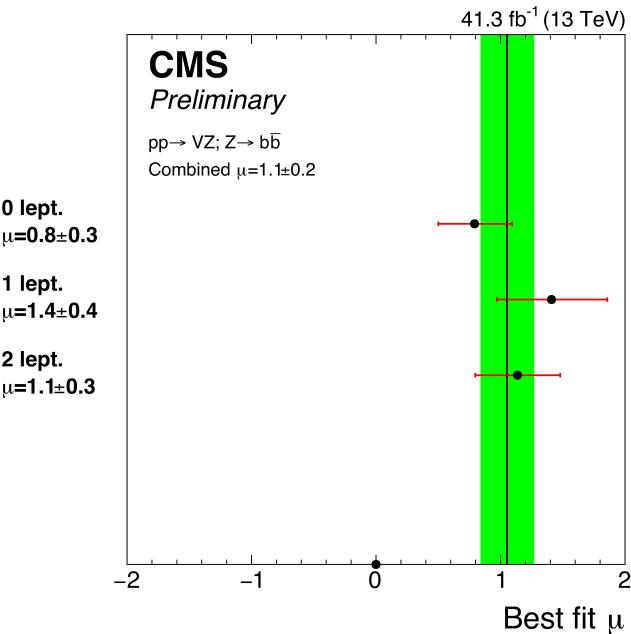


Figure 5-3. The best fit signal strengths and their uncertainties for each of the channels of the  $VZ(b\bar{b})$  cross-check analysis.

all background processes except for  $VZ(b\bar{b})$ , the capability of resolving the  $VH(b\bar{b})$  mass peak against the  $VZ(b\bar{b})$  mass peak without a multivariate analysis is shown in Figure 5-4.

Table 5-2. The massless DNN scores which define the boundaries of the four dijet invariant mass categories of each  $VH(b\bar{b})$  channel.

Channel	Mass Category Boundaries
$Z(\nu\bar{\nu})H$	[0.00, 0.558, 0.768, 0.838, 1.00]
$W(\ell\nu)H$	[0.00, 0.550, 0.856, 0.941, 1.00]
$Z(\ell\bar{\ell})H$ , Low $p_T(V)$	[0.00, 0.662, 0.875, 0.927, 1.00]
$Z(\ell\bar{\ell})H$ , High $p_T(V)$	[0.00, 0.486, 0.836, 0.899, 1.00]

Table 5-3. The data to Monte-Carlo scale factors obtained by the simultaneous maximum likelihood fit over all  $VH(b\bar{b})$  channels for the dijet invariant mass analysis. The errors include both statistical and systematic uncertainties.

Process	$Z(\nu\bar{\nu})H$	$W(\ell\nu)H$	$Z(\ell\bar{\ell})H$ , Low $p_T(V)$	$Z(\ell\bar{\ell})H$ , High $p_T(V)$
$W + udcsg$	$1.02 \pm 0.07$	$1.02 \pm 0.07$	-	-
$W + b$	$1.82 \pm 0.14$	$1.82 \pm 0.14$	-	-
$W + b\bar{b}$	$2.05 \pm 0.22$	$2.05 \pm 0.22$	-	-
$Z + udcsg$	$0.95 \pm 0.08$	-	$0.88 \pm 0.06$	$0.81 \pm 0.05$
$Z + b$	$1.16 \pm 0.16$	-	$0.99 \pm 0.13$	$1.12 \pm 0.11$
$Z + b\bar{b}$	$1.00 \pm 0.08$	-	$0.72 \pm 0.06$	$0.82 \pm 0.07$
$t\bar{t}$	$0.97 \pm 0.08$	$0.90 \pm 0.07$	$0.89 \pm 0.07$	$0.88 \pm 0.07$

## 5.4 $VH(b\bar{b})$ Analysis

The data to MC scale factors obtained by the maximum likelihood fit are shown in Table 5-4 and show good agreement with those obtained by the  $VZ(b\bar{b})$  cross-check analysis. The post-fit multivariate discriminant distributions of each channel are shown individually in Figure 5-5 and combined in Figure 5-6. The expected post-fit significance of the excess over the background prediction is  $3.1\sigma$ , while the observed significance is  $3.3\sigma$ . The best fit signal strength is  $\mu = 1.08^{+0.35}_{-0.33}$  for all channels combined and the signal strengths of the individual channels, shown in Figure 5-7, are compatible with a  $p$ -value of 96%. These results reestablish evidence for the  $VH(b\bar{b})$  decay. The contributions of the major uncertainty sources are listed in Table 5-5 and decomposed into four groups: theory, size of simulated samples, experimental, and statistical. For each major source of uncertainty, the largest individual subcomponents are also listed.

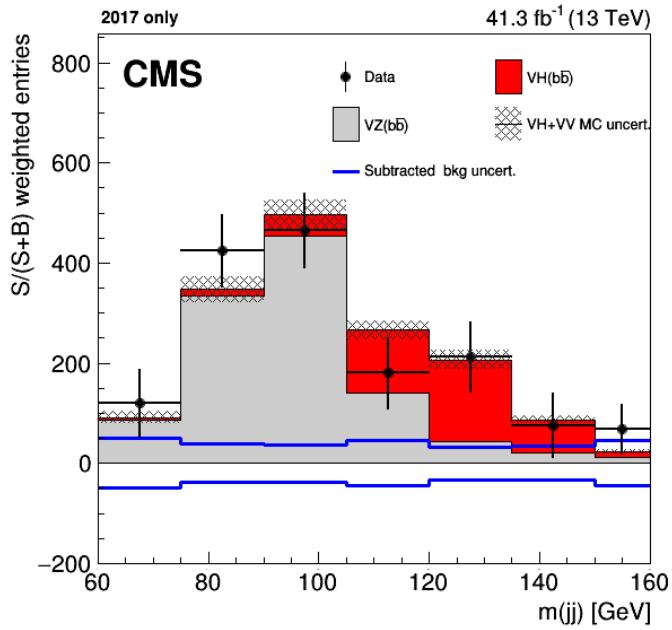


Figure 5-4. The dijet invariant mass distribution for events weighted by  $S/(S + B)$  in all channels combined. The data (black) and the fitted  $VH(b\bar{b})$  signal (red) and  $VZ(b\bar{b})$  background (gray) distributions are shown with all other fitted backgrounds subtracted. The vertical error bars represent the  $1\sigma$  statistical uncertainty on the data before background subtraction, while the blue error band represents the  $1\sigma$  total uncertainty on the signal and all background processes.

Table 5-4. The data to Monte-Carlo scale factors obtained by the simultaneous maximum likelihood fit over all  $VH(b\bar{b})$  channels. The errors include both statistical and systematic uncertainties.

Process	$Z(\nu\bar{\nu})H$	$W(\ell\nu)H$	$Z(\ell\ell)H$ , Low $p_T(V)$	$Z(\ell\ell)H$ , High $p_T(V)$
$W + udcsg$	$1.04 \pm 0.07$	$1.04 \pm 0.07$	-	-
$W + b$	$2.09 \pm 0.16$	$2.09 \pm 0.16$	-	-
$W + b\bar{b}$	$1.74 \pm 0.21$	$1.74 \pm 0.21$	-	-
$Z + udcsg$	$0.95 \pm 0.09$	-	$0.89 \pm 0.06$	$0.81 \pm 0.05$
$Z + b$	$1.02 \pm 0.17$	-	$0.94 \pm 0.12$	$1.17 \pm 0.10$
$Z + b\bar{b}$	$1.20 \pm 0.11$	-	$0.81 \pm 0.07$	$0.88 \pm 0.08$
$t\bar{t}$	$0.99 \pm 0.07$	$0.93 \pm 0.07$	$0.89 \pm 0.07$	$0.91 \pm 0.07$

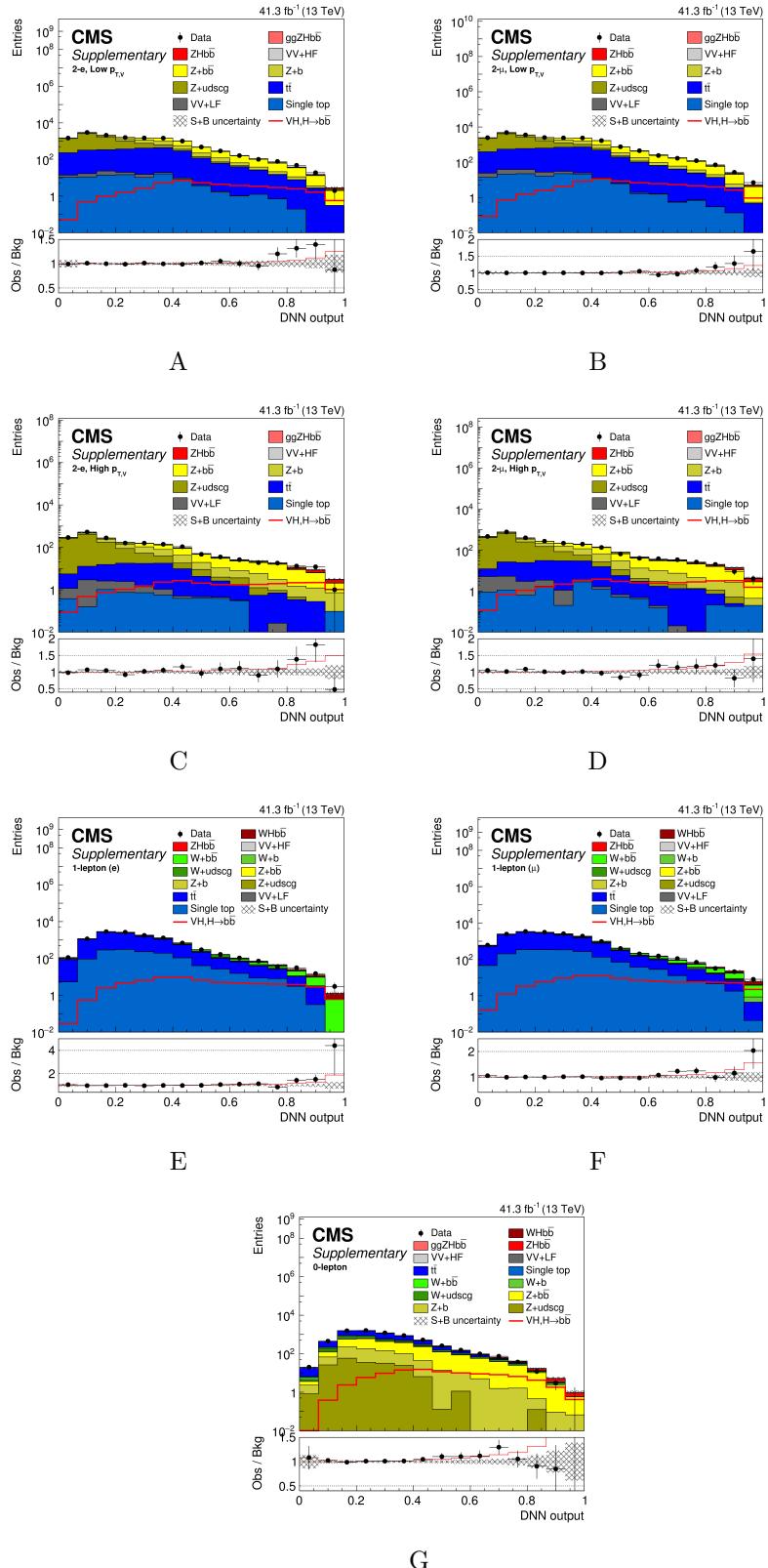


Figure 5-5. The post-fit  $VH(\bar{b}b)$  multivariate discriminant distributions of the A) low  $p_T(V)$   $Z(e\bar{e})H$ ; B) low  $p_T(V)$   $Z(\mu\bar{\mu})H$ ; C) high  $p_T(V)$   $Z(e\bar{e})H$ ; D) high  $p_T(V)$   $Z(\mu\bar{\mu})H$ ; E)  $W(e\nu)H$ ; F)  $W(\mu\nu)H$ ; and G)  $Z(\nu\bar{\nu})H$  channels.

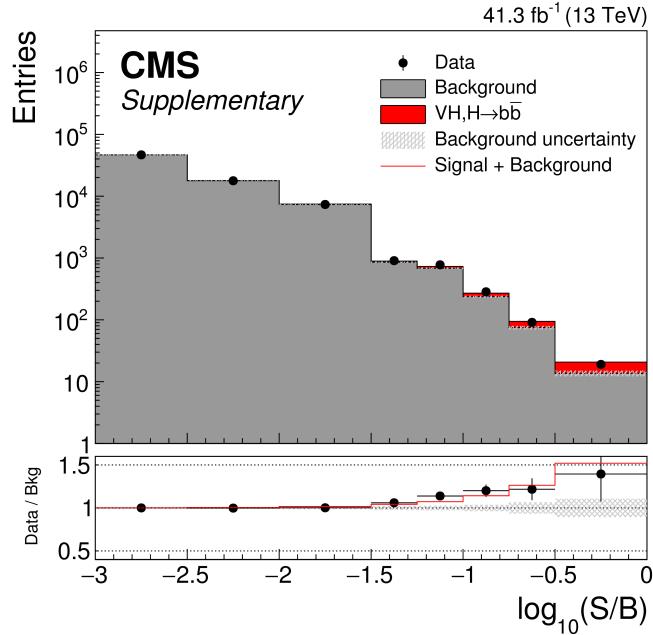


Figure 5-6. The post-fit multivariate discriminant distributions of signal, background, and data event yields combined for all  $VH(b\bar{b})$  channels and sorted into bins of similar signal-to-background ratio. The  $VH(b\bar{b})$  signal contribution (red) has been stacked above the sum of all background yields (gray). The bottom panel shows the data to background ratio (black points) and the sum of signal and background contributions divided by the background yield (red line).

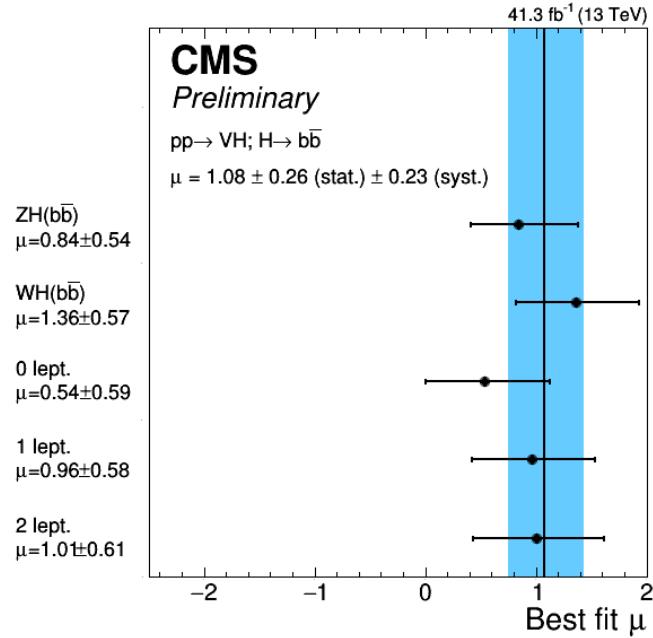


Figure 5-7. The best fit signal strengths and their uncertainties for each of the channels of the  $VH(b\bar{b})$  analysis.

Table 5-5. The contributions of the main uncertainty sources in the measurement of the  $VH(b\bar{b})$  signal strength. Due to correlations between the nuisance parameters of different sources in the maximum likelihood fit, the sum in quadrature of each source is not necessarily the total uncertainty of each major component.

Uncertainty Source	$\Delta\mu$	
Theory	+0.11	-0.09
- Signal theory	+0.07	-0.04
- Background theory	+0.08	-0.08
Size of simulated samples	+0.12	-0.12
Experimental	+0.16	-0.15
- Luminosity	+0.03	-0.03
- Lepton efficiency	+0.02	-0.01
- Jet energy scale and resolution	+0.05	-0.05
- $b$ -Tagging efficiency	+0.09	-0.08
- Other experimental uncertainties	+0.10	-0.09
Statistical	+0.26	-0.26
- Background control region normalization	+0.12	-0.12
Total	+0.35	-0.33

## 5.5 Combinations with Past CMS Results

A combination of the 2017  $VH(b\bar{b})$  analysis results with those of the 2016  $VH(b\bar{b})$  analysis published in Ref. [28] increases the expected and observed significance to  $4.16\sigma$  and  $4.36\sigma$ , respectively, with a best fit signal strength of  $\mu = 1.06^{+0.25}_{-0.26}$ . A further combination of these analyses with the Run 1 analysis published in Ref. [25] further increases the expected and observed significance to  $4.88\sigma$  and  $4.82\sigma$ , respectively, with a best fit signal strength of  $\mu = 1.01^{+0.22}_{-0.23}$ . The results of this combination fall short of establishing an observation of the  $VH(b\bar{b})$  decay.

A global fit is performed that combines all available CMS analyses of different  $H \rightarrow b\bar{b}$  production processes, including this analysis of  $VH(b\bar{b})$ , gluon fusion[19], vector boson fusion[131], and  $t\bar{t}$  associated production[132–134]. Because the results of these analyses span multiple datasets collected at 7, 8, and 13 TeV and different detector operating conditions, most systematic uncertainty sources are treated as uncorrelated. For analyses using the same collision energies, the jet energy scale uncertainties are treated as correlated between those processes. The theoretical uncertainties are correlated between

all processes and datasets. This grand combination achieves an expected and observed significance of  $5.5\sigma$  and  $5.6\sigma$ , respectively. The best fit signal strength is  $\mu = 1.04 \pm 0.20$ , and the signal strengths of the different  $H \rightarrow b\bar{b}$  production modes are shown in Figure 5-8. This marks the first observation of the  $H \rightarrow b\bar{b}$  decay at the CMS experiment and the measured signal strength indicates agreement with the Standard Model expectation.

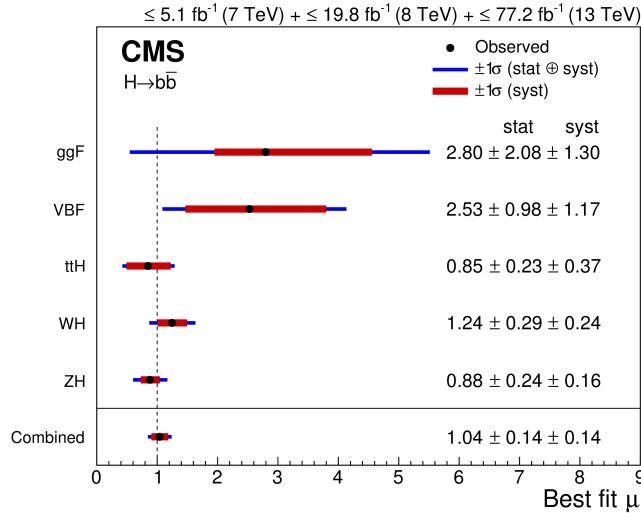


Figure 5-8. The best fit signal strengths and their uncertainties for each of the analyses of the different  $H \rightarrow b\bar{b}$  production modes that participated in the grand combination.

## CHAPTER 6 CONCLUSIONS

A search for the Standard Model Higgs boson produced in association with a weak vector boson and decaying into a bottom-antibottom quark pair ( $b\bar{b}$ ) has been presented for the  $VH(b\bar{b})$  decay channels  $Z(\nu\bar{\nu})H(b\bar{b})$ ,  $W(e\nu)H(b\bar{b})$ ,  $W(\mu\nu)H(b\bar{b})$ ,  $Z(e\bar{e})H(b\bar{b})$ , and  $Z(\mu\bar{\mu})H(b\bar{b})$ . The search is performed with a dataset corresponding to an integrated luminosity of  $41.3 \text{ fb}^{-1}$  at a center-of-mass energy of  $\sqrt{s} = 13 \text{ TeV}$  recorded by the CMS experiment at the LHC during Run 2 in 2017. An excess of events in data over the expected background is observed with a significance of 3.3 standard deviations, as compared to an expected significance of 3.1 standard deviations under the Standard Model prediction for a Higgs boson of mass  $m_H = 125.09 \text{ GeV}$  decaying to bottom quarks. The corresponding signal strength of this excess relative to the Standard Model Higgs boson is measured to be  $\mu = 1.08^{+0.35}_{-0.33}$ , which suggests compatibility with the Standard Model.

A grand combination of this result is performed with previous measurements made by the CMS collaboration of the associated vector boson, gluon fusion, vector boson fusion, and associated top quark production modes for  $H \rightarrow b\bar{b}$ . This combination achieves an observed significance of 5.6 standard deviations, with an expected significance of 5.5 standard deviations, and measures an overall signal strength of  $\mu = 1.04 \pm 0.20$  for the  $H \rightarrow b\bar{b}$  decay. This constitutes the observation of the  $H \rightarrow b\bar{b}$  decay by the CMS collaboration, with no clear indication of physics beyond the Standard Model.

Although the  $H \rightarrow b\bar{b}$  decay has been observed, the continued study of this decay presents the opportunity to check for anomalous Higgs' Yukawa couplings and, given it has the largest branching fraction of all Higgs' decays, to improve the constraints placed on the total decay width of the Higgs boson. The focus of future analyses will then prioritize improving the precision of the measurement of the  $VH(b\bar{b})$  production cross section and the  $H \rightarrow b\bar{b}$  branching fraction. Given that current jet flavor tagging algorithms are capable of identifying  $c$ -jets, this analysis may be readily adapted towards a

search for the  $H \rightarrow c\bar{c}$  decay via the vector boson associated production channels  $VH(c\bar{c})$  in order to study the Higgs boson coupling to a second generation quark.

By refining the strategy adopted for the dijet invariant mass analysis, where a multivariate classifier is evaluated in such a way that its output is decorrelated from the fitted distribution, the effects of shape systematics could be reduced as the dijet invariant mass distributions of the background processes are typically flat. Similar techniques that warrant further investigation include a method of rescaling input features to train a multivariate classifier is decorrelated from a target variable that has been successfully used by the  $H \rightarrow \mu\bar{\mu}$  analysis[135] and an adversarial training method that has been proposed to decorrelate jet tagging algorithms from the jet mass[136]. By careful retraining of the multivariate classifier instead of marginalizing over the dijet invariant mass distribution, these alternative approaches may enable a dijet invariant mass analysis achieve a sensitivity that is competitive with a multivariate analysis.

Future analyses of the  $VH(b\bar{b})$  decay into Run 3 and beyond will face increased luminosities which may cause the dijet invariant mass resolution and  $b$ -tagging performance to degrade due to increased pileup. More robust methods of pileup removal such as the pileup per particle identification (PUPPI)[137] algorithm may be required under these harsher conditions. Future analyses performed at higher center-of-mass energies may also benefit from a dedicated jet substructure[138] analysis, given that the bottom quarks from the Higgs decay become highly collimated at very high boosts. While typical jet algorithms may be unable to resolve the merged  $b$ -jets, jet substructure techniques such as those used by the search for  $H \rightarrow b\bar{b}$  using the gluon fusion production mode[19] are capable of identifying jet constituents and their flavor. Novel applications of machine learning in the field of high energy physics towards jet flavor tagging, such as the DeepJet[139] algorithm, and discovery significance optimization[140] may improve the sensitivity of the analysis.

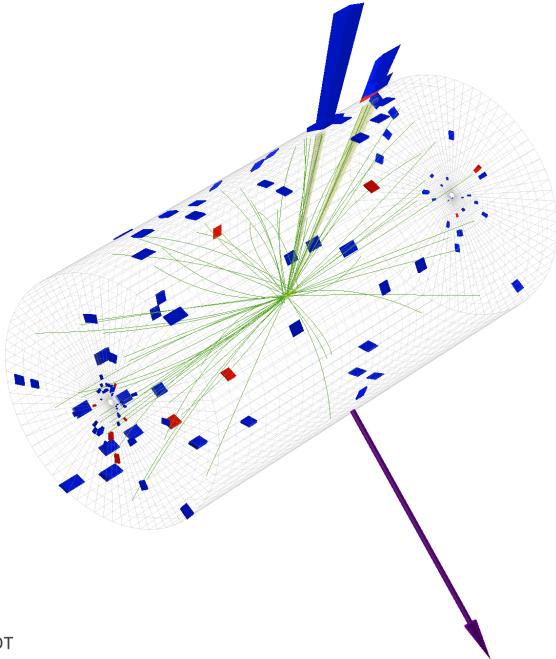
## APPENDIX A EVENT DISPLAYS

Collision events recorded by the CMS detector can be visualized as event displays using the Fireworks application developed by the CMS collaboration.[141] Event displays are used to visually summarize particle interactions with the detector’s subsystems by rendering reconstructed objects of interest such as particle tracks and jets over a 3D model of the detector. Figures A-1 through A-5 each show multiple different perspectives of an event display for a candidate signal region event from each decay channel. The candidate events were chosen based on their high DNN scores, which indicate that the DNN of their corresponding decay channel classified them as a signal event with high confidence. Although the event metadata are provided on the event displays, other event-level quantities of interest are shown in Table A-1.

Table A-1. Information for the candidate signal region events visualized by the event display. The values of the kinematic quantities are reported in units of GeV and the  $b$ -tag algorithm is DeepCSV.

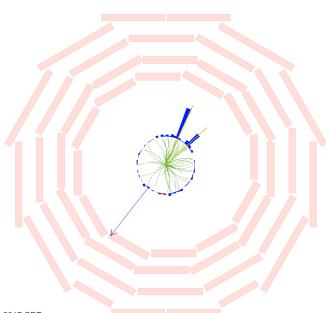
Channel	$m(jj)$	$p_T(jj)$	$p_T(j)_{\max}$	$p_T(j)_{\min}$	$b$ -tag <sub>max</sub>	$b$ -tag <sub>min</sub>	$p_T(V)$	DNN score
$Z(\nu\bar{\nu})H$	121.56	319.43	195.99	129.81	0.999	0.993	319.10	0.975
$W(e\nu)H$	126.49	259.66	179.31	87.31	0.999	0.995	295.75	0.989
$W(\mu\nu)H$	117.69	250.55	177.33	85.28	0.996	0.995	259.80	0.993
$Z(e\bar{e})H$	114.14	286.17	216.54	97.82	0.990	0.957	311.33	1.00
$Z(\mu\bar{\mu})H$	111.39	100.12	78.66	69.53	0.984	0.975	466.83	1.00

For each figure, the perspective of the largest subfigure is a projection of the event display onto the  $xz$ -plane, unless otherwise specified. The perspective of both of the smaller subfigures is a projection of the event display onto the  $xy$ -plane, with the subfigure on the right zoomed in on the interaction point. This level of zoom is necessary to show the secondary vertices and their associated tracks, shown in black, and to appreciate that those tracks do not extrapolate well to any of the primary vertices, shown in yellow.



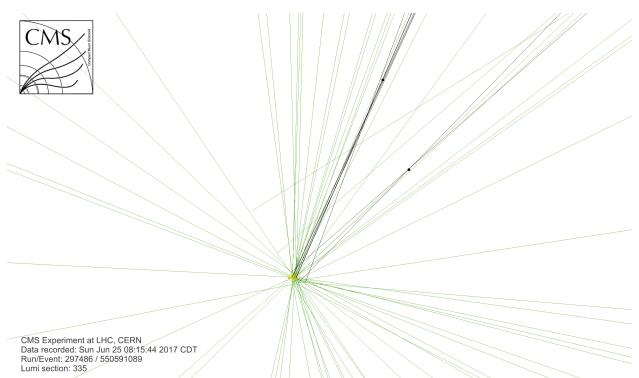
CMS Experiment at LHC, CERN  
Data recorded: Sun Jun 25 08:15:44 2017 CDT  
Run/Event: 297486 / 550591089  
Lumi section: 335

A



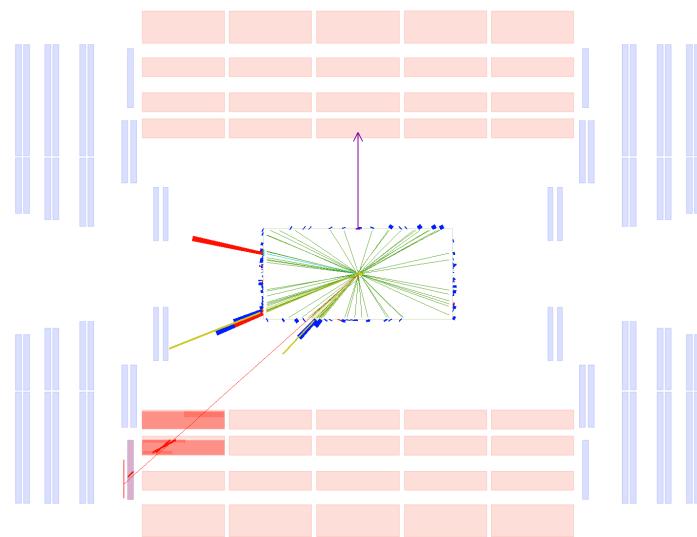
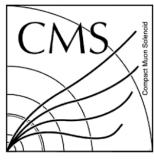
CMS Experiment at LHC, CERN  
Data recorded: Sun Jun 25 08:15:44 2017 CDT  
Run/Event: 297486 / 550591089  
Lumi section: 335

B



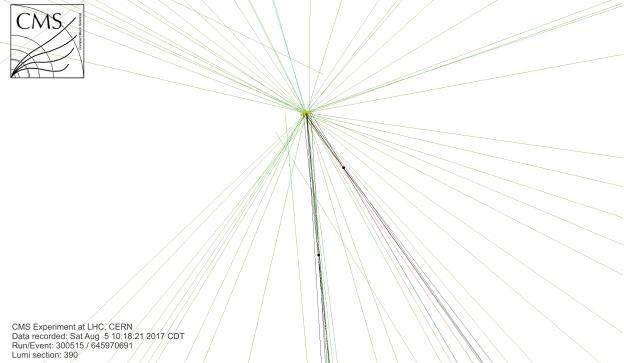
C

Figure A-1. The event display for a candidate  $Z(\nu\bar{\nu})H(b\bar{b})$  event A) from a 3D perspective view; B) projected onto the  $xy$ -plane; and C) projected onto the  $xy$ -plane and zoomed in on the interaction point. The purple arrow represents the missing transverse momentum from the invisible leptonic decay of the  $Z$  boson while the two yellow cones with their blue and red calorimeter towers represent the two  $b$ -jets from the decay of the Higgs boson. The secondary vertices of the  $b$ -jets and their associated tracks (black) are visible in C.



CMS Experiment at LHC, CERN  
Data recorded: Sat Aug 5 10:18:21 2017 CDT  
Run/Event: 300515 / 645970691  
Lumi section: 390

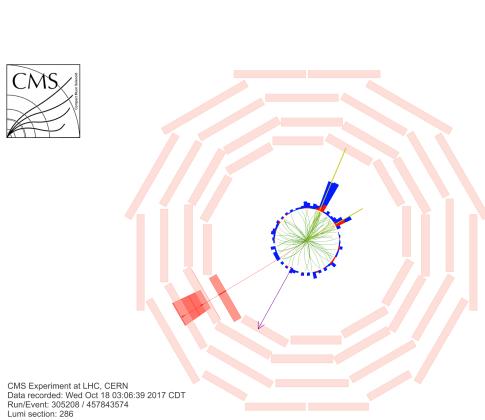
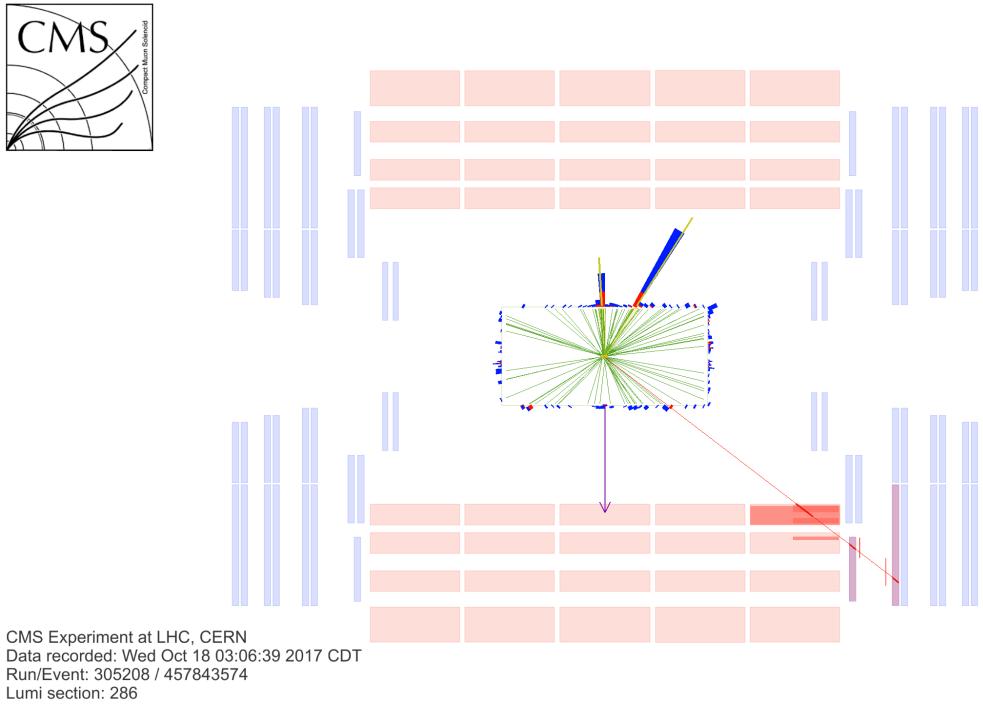
A



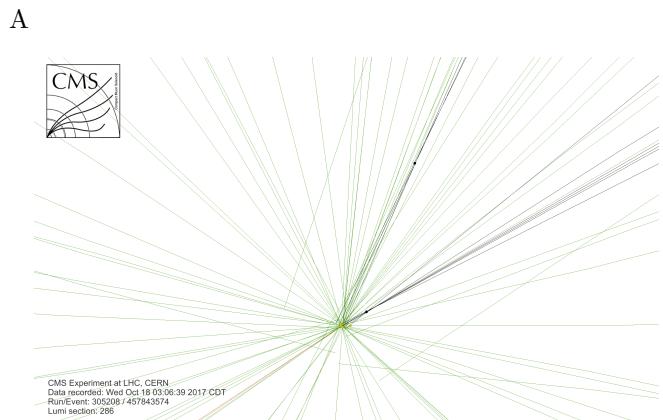
B

C

Figure A-2. The event display for a candidate  $W(e\nu)H(b\bar{b})$  event A) projected onto the  $xz$ -plane; B) projected onto the  $xy$ -plane; and C) projected onto the  $xy$ -plane and zoomed in on the interaction point. The cyan track leading to a red calorimeter tower and the purple arrow represent the electron and the missing transverse momentum from the leptonic decay of the  $W$  boson, respectively, while the two yellow cones with their blue and red calorimeter towers represent the two  $b$ -jets from the decay of the Higgs boson. The secondary vertices of the  $b$ -jets and their associated tracks (black) are visible in C. The red track associated with one of the  $b$ -jets is likely a muon resulting from a semi-leptonic decay of the  $b$ -hadron.



B



C

Figure A-3. The event display for a candidate  $W(\mu\nu)H(b\bar{b})$  event A) projected onto the  $xz$ -plane; B) projected onto the  $xy$ -plane; and C) projected onto the  $xy$ -plane and zoomed onto the interaction point. The red track and the purple arrow represent the muon and the missing transverse momentum from the leptonic decay of the  $W$  boson, respectively, while the two yellow cones with their blue and red calorimeter towers represent the two  $b$ -jets from the decay of the Higgs boson. The secondary vertices of the  $b$ -jets and their associated tracks (black) are visible in C.

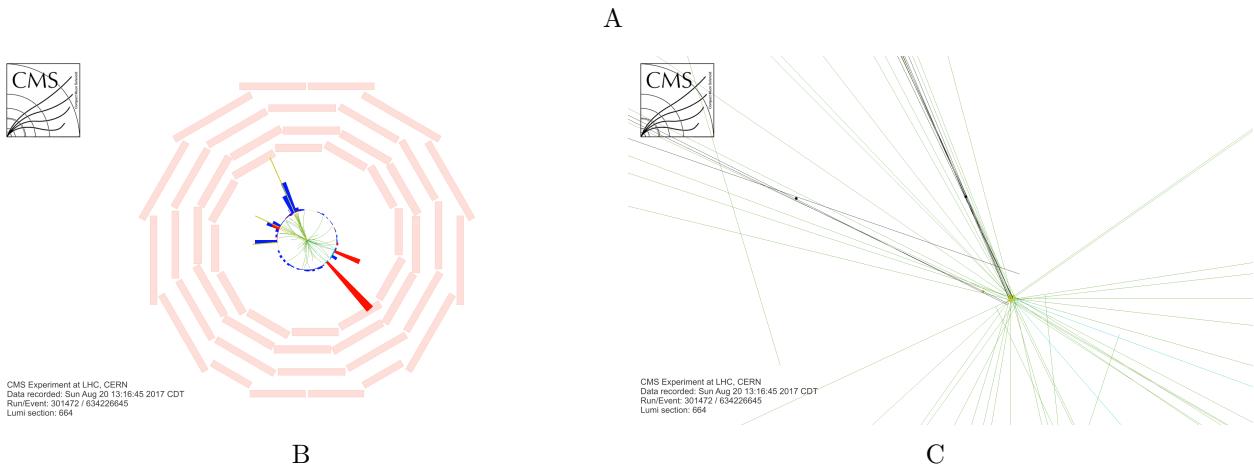
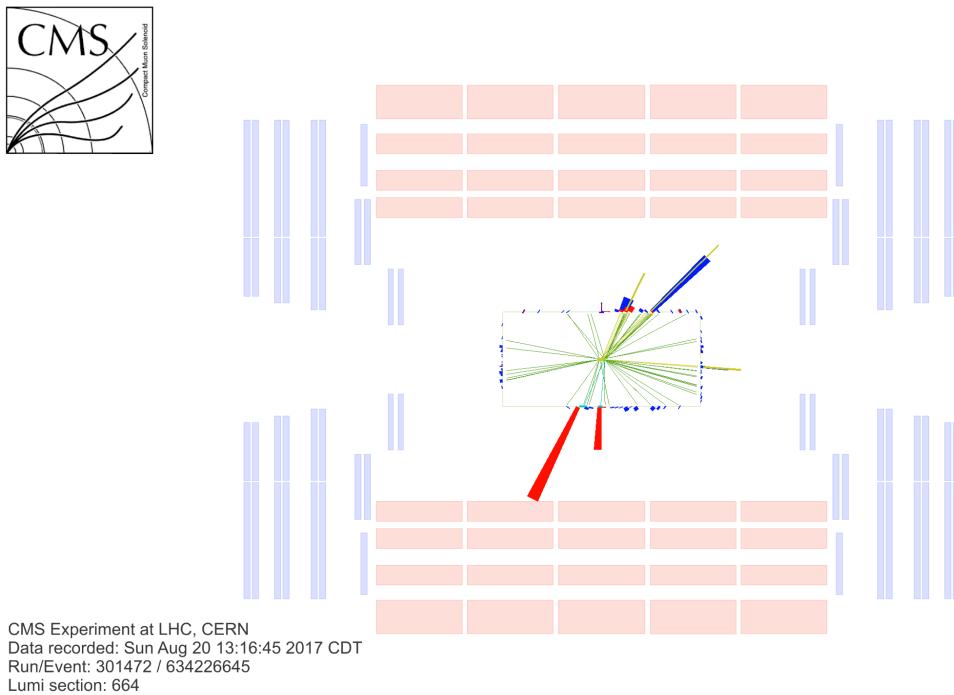


Figure A-4. The event display for a candidate  $Z(e\bar{e})H(b\bar{b})$  event A) projected onto the  $xz$ -plane; B) projected onto the  $xy$ -plane; and C) projected onto the  $xy$ -plane and zoomed onto the interaction point. The cyan tracks leading to red calorimeter towers represent the electrons from the leptonic decay of the  $Z$  boson, while the two yellow cones with their blue and red calorimeter towers represent the two  $b$ -jets from the decay of the Higgs boson. The secondary vertices of the  $b$ -jets and their associated tracks (black) are visible in C. The low missing transverse momentum close in azimuthal angle to one of the  $b$ -jets is suggestive of jet energy mismeasurement. This may be related to the blue calorimeter tower without obvious associated tracks, which may have resulted from the decay of a neutral hadron.

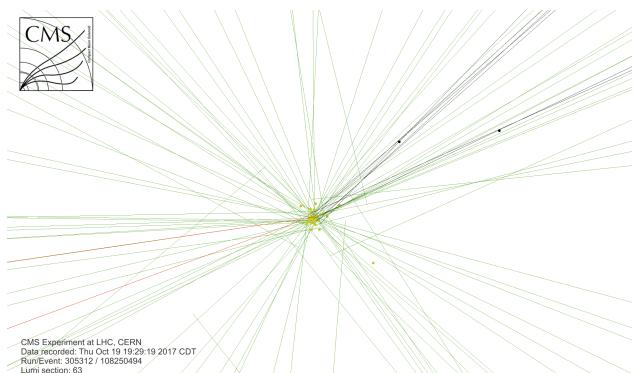
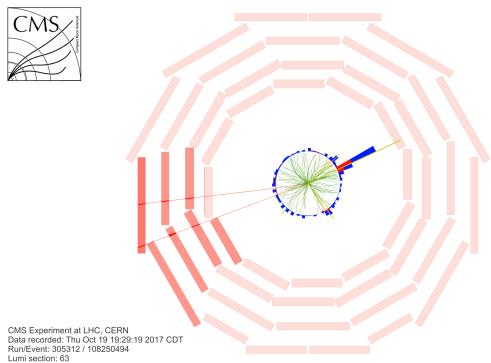
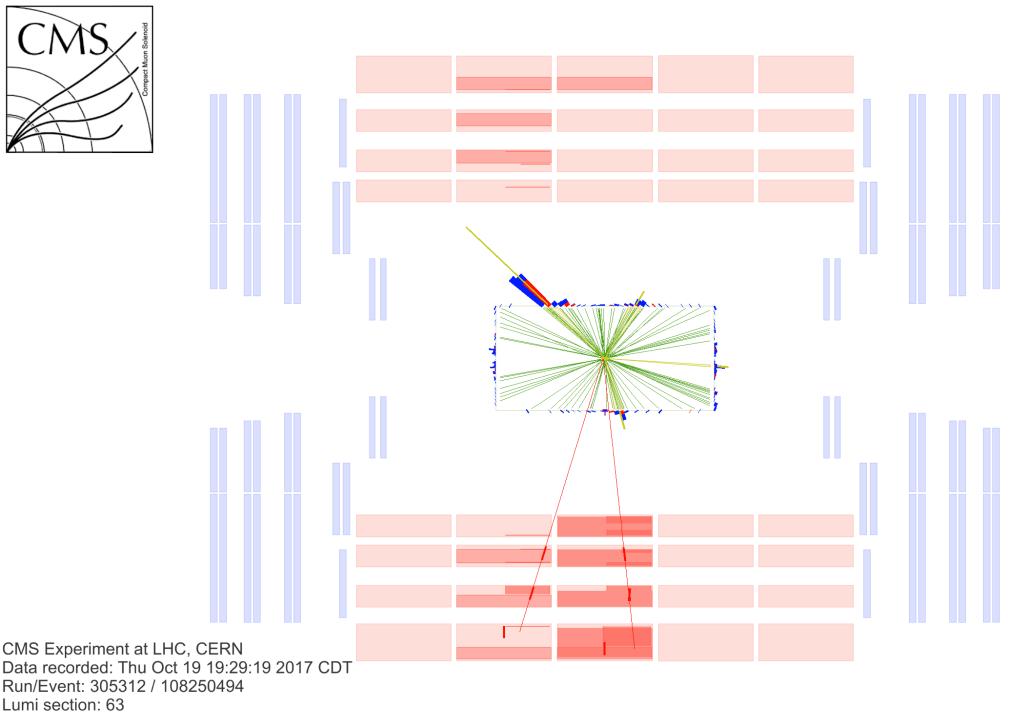


Figure A-5. The event display for a candidate  $Z(\mu\bar{\mu})H(b\bar{b})$  event A) projected onto the  $xz$ -plane; B) projected onto the  $xy$ -plane; and C) projected onto the  $xy$ -plane and zoomed onto the interaction point. The red tracks represent the muons from the leptonic decay of the  $Z$  boson, while the two yellow cones with their blue and red calorimeter towers located opposite to the red tracks represent the two  $b$ -jets from the decay of the Higgs boson. The secondary vertices of the  $b$ -jets and their associated tracks (black) are visible in C. The two other yellow cones could represent jets caused by initial or final state radiation, neither of which were  $b$ -tagged.

## APPENDIX B CONTROL REGION DISTRIBUTIONS

The control regions defined by the  $VH(b\bar{b})$  analysis are used to validate the agreement between data and Monte-Carlo (MC) simulation and to determine the normalizations of the dominant background processes during the signal extraction. The distributions of variables for the  $Z(\nu\bar{\nu})H$  channel are shown in Figures B-1 and B-2 for the  $t\bar{t}$  control region, Figures B-3 and B-4 for the  $Z+$ light control region, and Figures B-5 and B-6 for the  $Z+$ heavy control region. Select control region distributions of variables are shown in Figure B-7 for the  $W(\ell\nu)H$  channels and B-8 and B-9 for the low and high  $p_T(V)$  regions of the  $Z(\ell\bar{\ell})H$  channels. The distributions are visualized as stacked histograms to better convey the relative contributions of the individual background processes. All residual MC corrections have been applied, as well as the appropriate scale factors which adjust the dominant background normalizations. The residual discrepancies between data and simulation are attributed to adjustments that account for the various systematic sources, which are determined during the signal extraction fit but absent for the purposes of these visualizations.

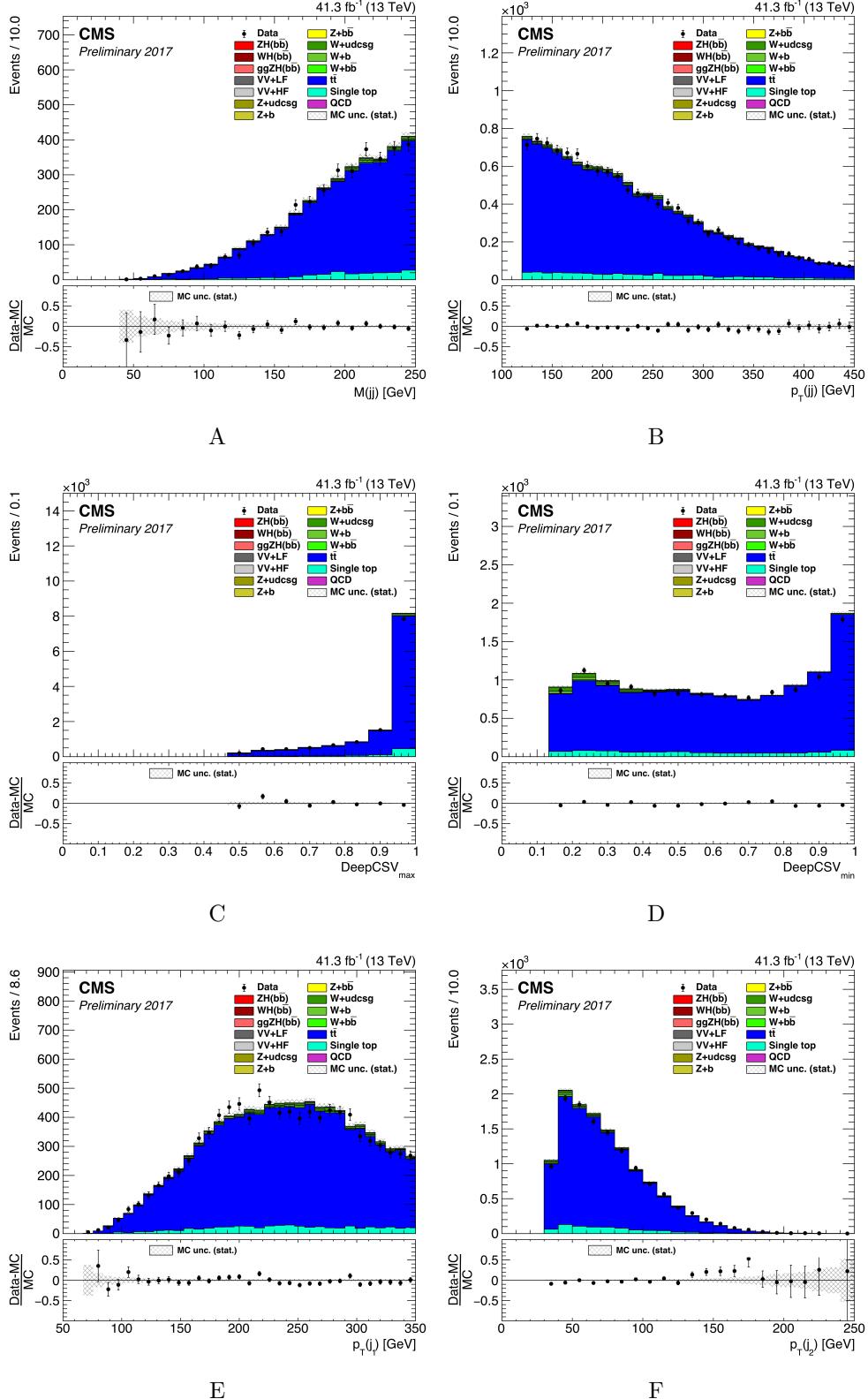


Figure B-1. The distributions of variables in the  $t\bar{t}$  control region of the  $Z(\nu\bar{\nu})H$  channel:  
 A)  $m(jj)$ , B)  $p_T(jj)$ , C)  $b$ -tag<sub>max</sub>, D)  $b$ -tag<sub>min</sub>, E)  $p_T(j_1)$ max, F)  $p_T(j_2)$ min.

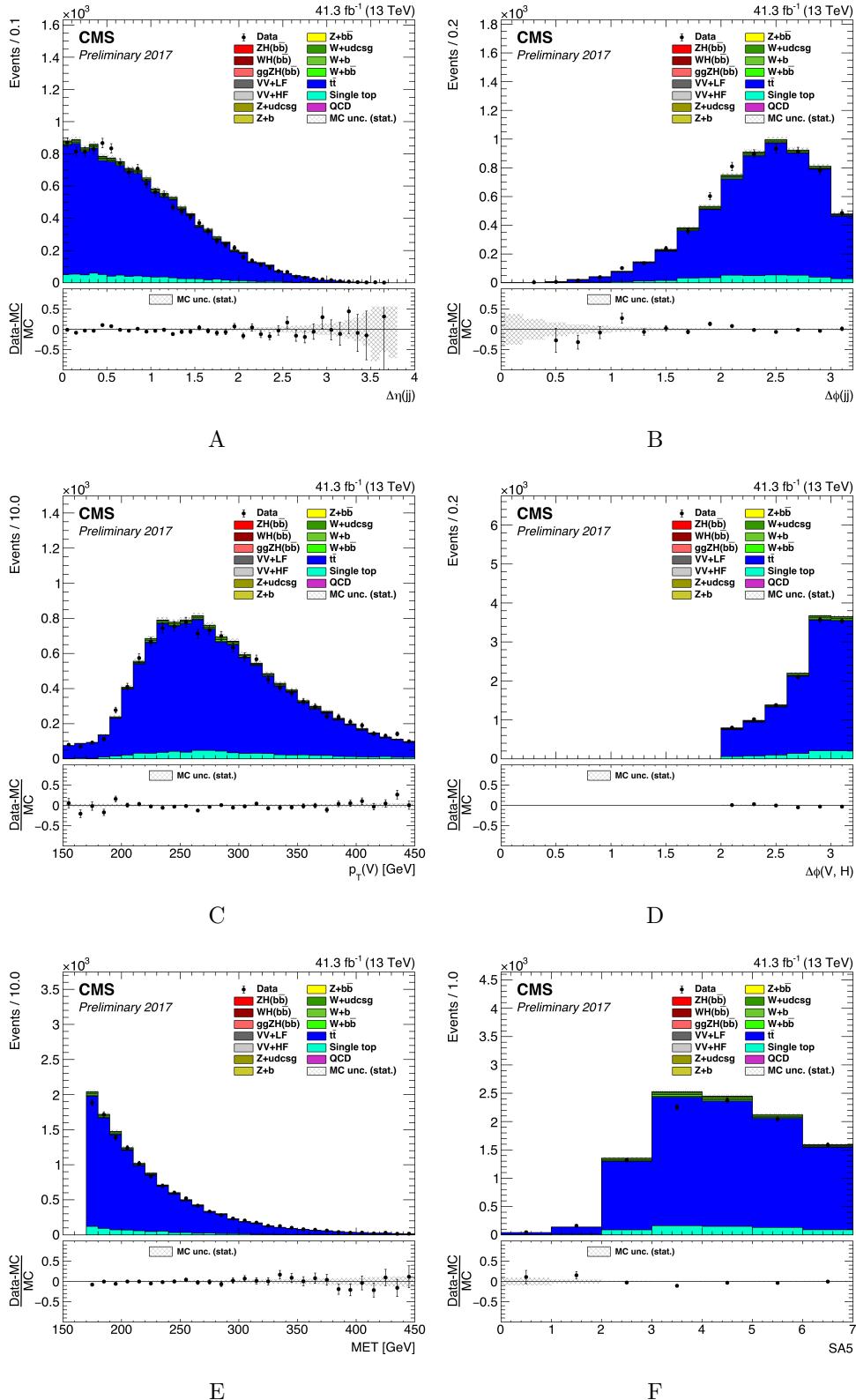


Figure B-2. The distributions of variables in the  $t\bar{t}$  control region of the  $Z(\nu\bar{\nu})H$  channel:  
 A)  $|\Delta\eta(j_1, j_2)|$ , B)  $|\Delta\phi(j_1, j_2)|$ , C)  $p_T(V)$ , D)  $|\Delta\phi(V, H)|$ , E)  $p_T^{\text{miss}}$ , F) SA5.

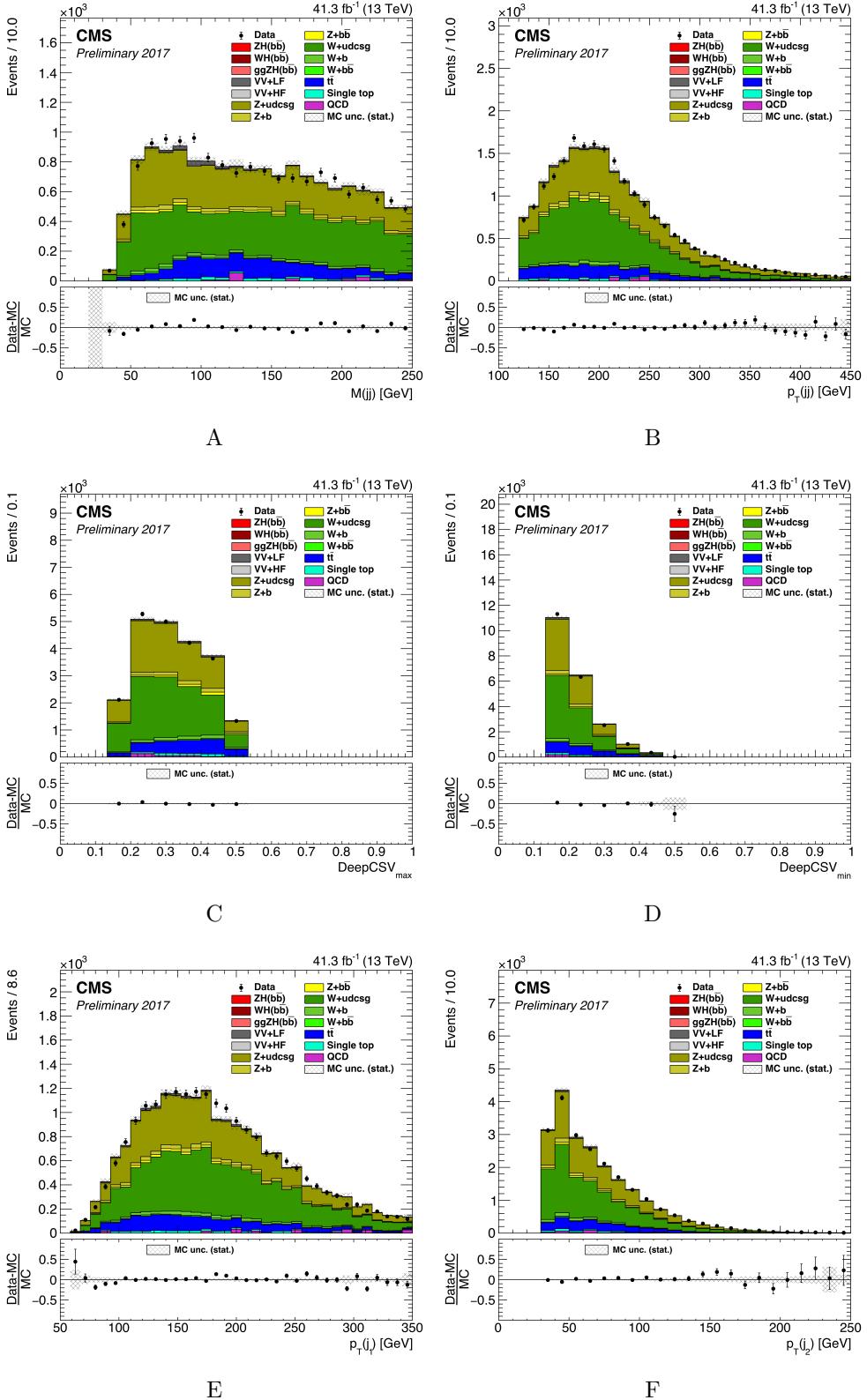


Figure B-3. The distributions of variables in the  $Z + \text{light}$  control region of the  $Z(\nu\bar{\nu})H$  channel: A)  $m(jj)$ , B)  $p_T(jj)$ , C)  $b$ -tag<sub>max</sub>, D)  $b$ -tag<sub>min</sub>, E)  $p_T(j_1)$ <sub>max</sub>, F)  $p_T(j_1)$ <sub>min</sub>.

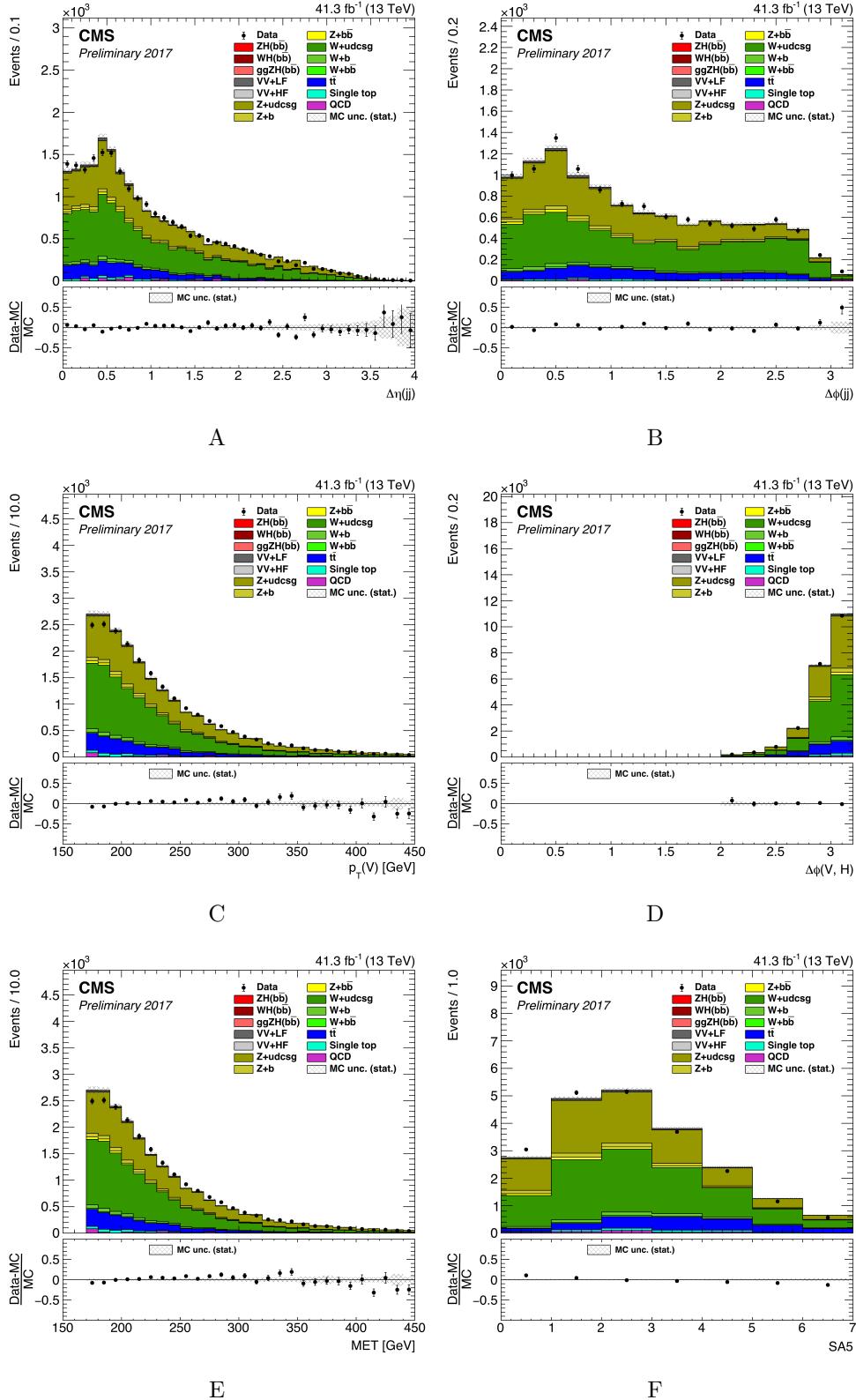


Figure B-4. The distributions of variables in the  $Z + \text{light}$  control region of the  $Z(\nu\bar{\nu})H$  channel: A)  $|\Delta\eta(j_1, j_2)|$ , B)  $|\Delta\phi(j_1, j_2)|$ , C)  $p_T(V)$ , D)  $|\Delta\phi(V, H)|$ , E)  $p_T^{\text{miss}}$ , F)  $\text{SA5}$ .

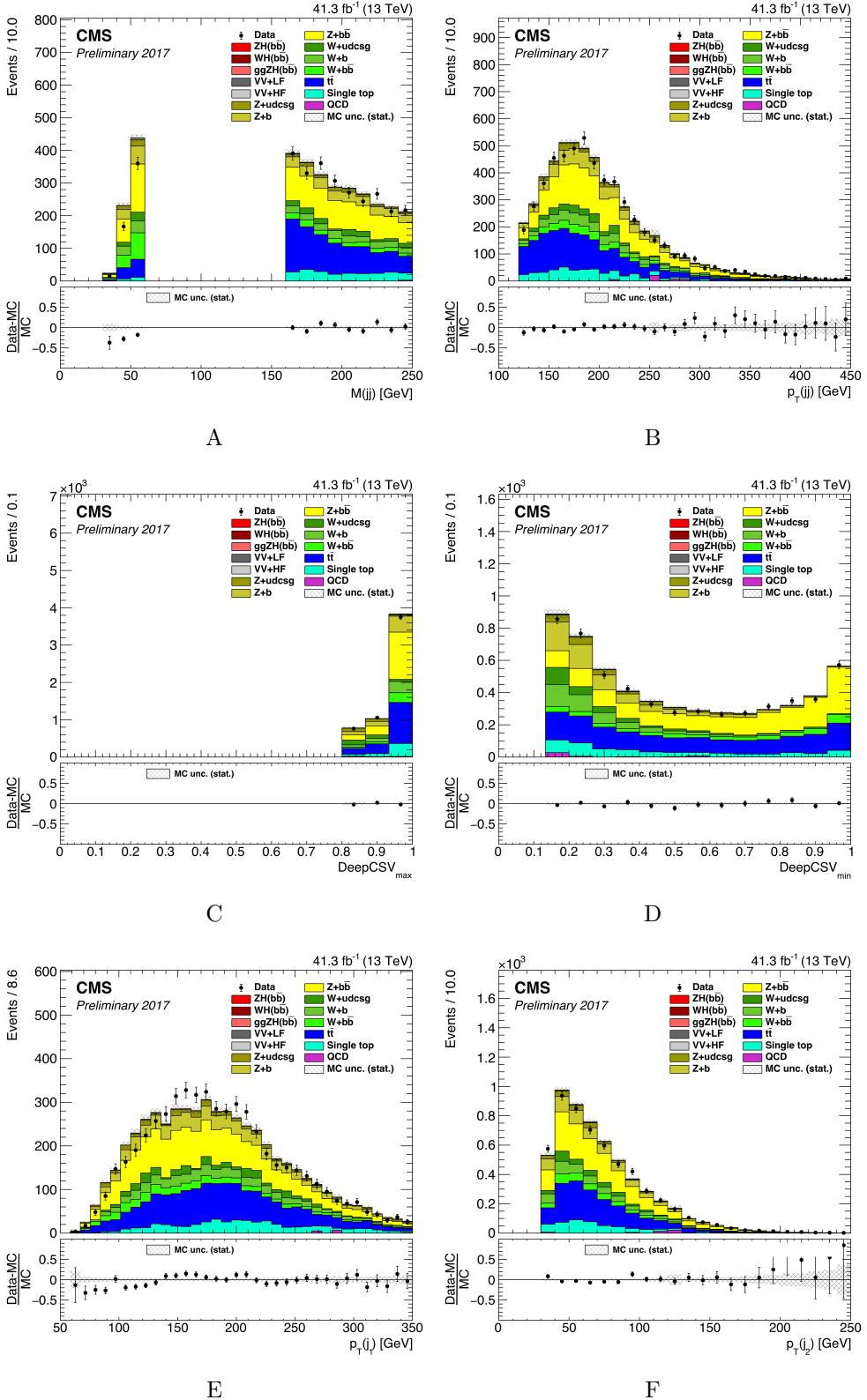


Figure B-5. The distributions of variables in the  $Z + \text{heavy}$  control region of the  $Z(\nu\bar{\nu})H$  channel: A)  $m(jj)$ , B)  $p_T(jj)$ , C)  $b$ -tag<sub>max</sub>, D)  $b$ -tag<sub>min</sub>, E)  $p_T(j)_\text{max}$ , F)  $p_T(j)_\text{min}$ .

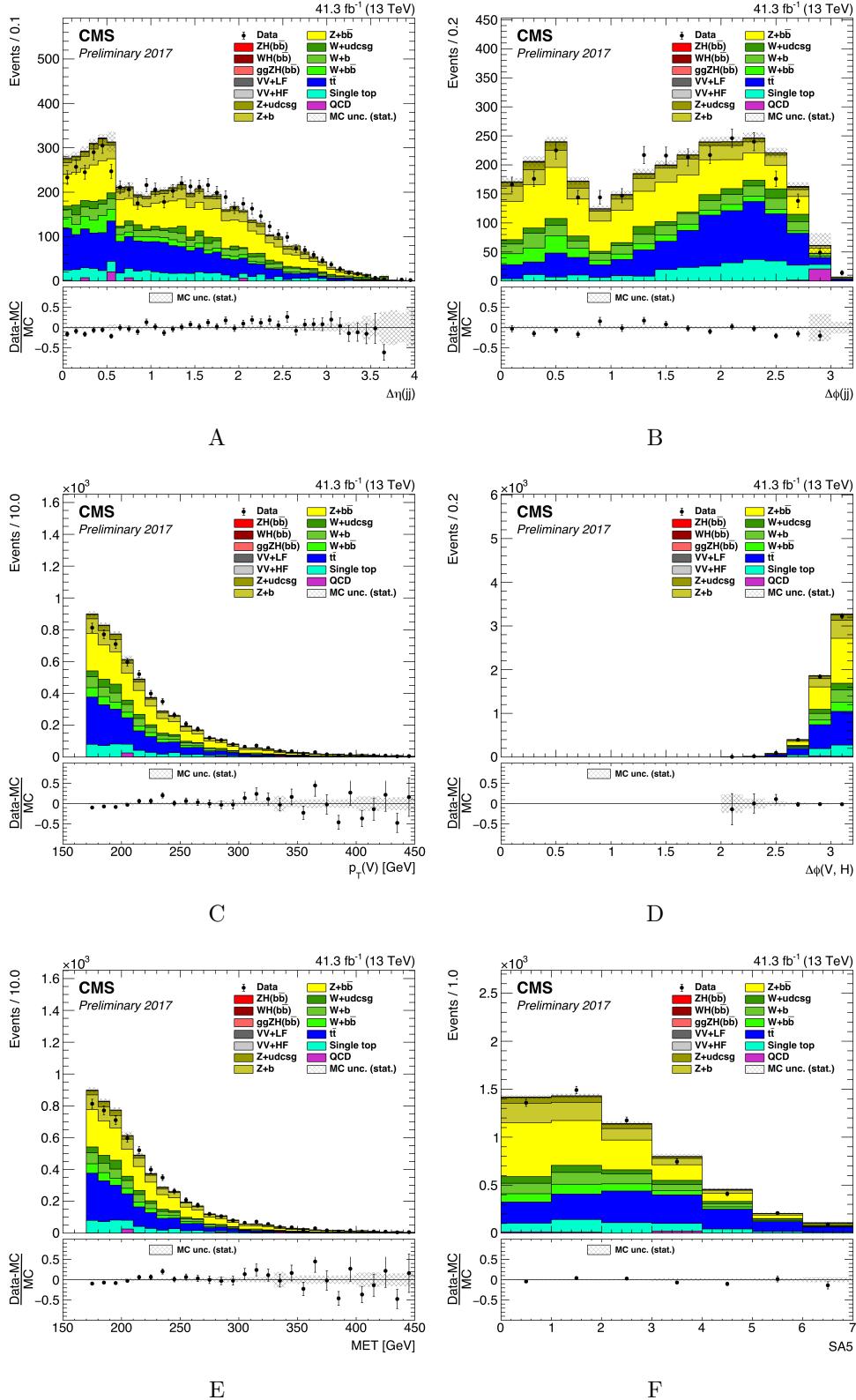


Figure B-6. The distributions of variables in the  $Z + \text{heavy}$  control region of the  $Z(\nu\bar{\nu})H$  channel: A)  $|\Delta\eta(j_1, j_2)|$ , B)  $|\Delta\phi(j_1, j_2)|$ , C)  $p_T(V)$ , D)  $|\Delta\phi(V, H)|$ , E)  $p_T^{\text{miss}}$ , F) SA5.

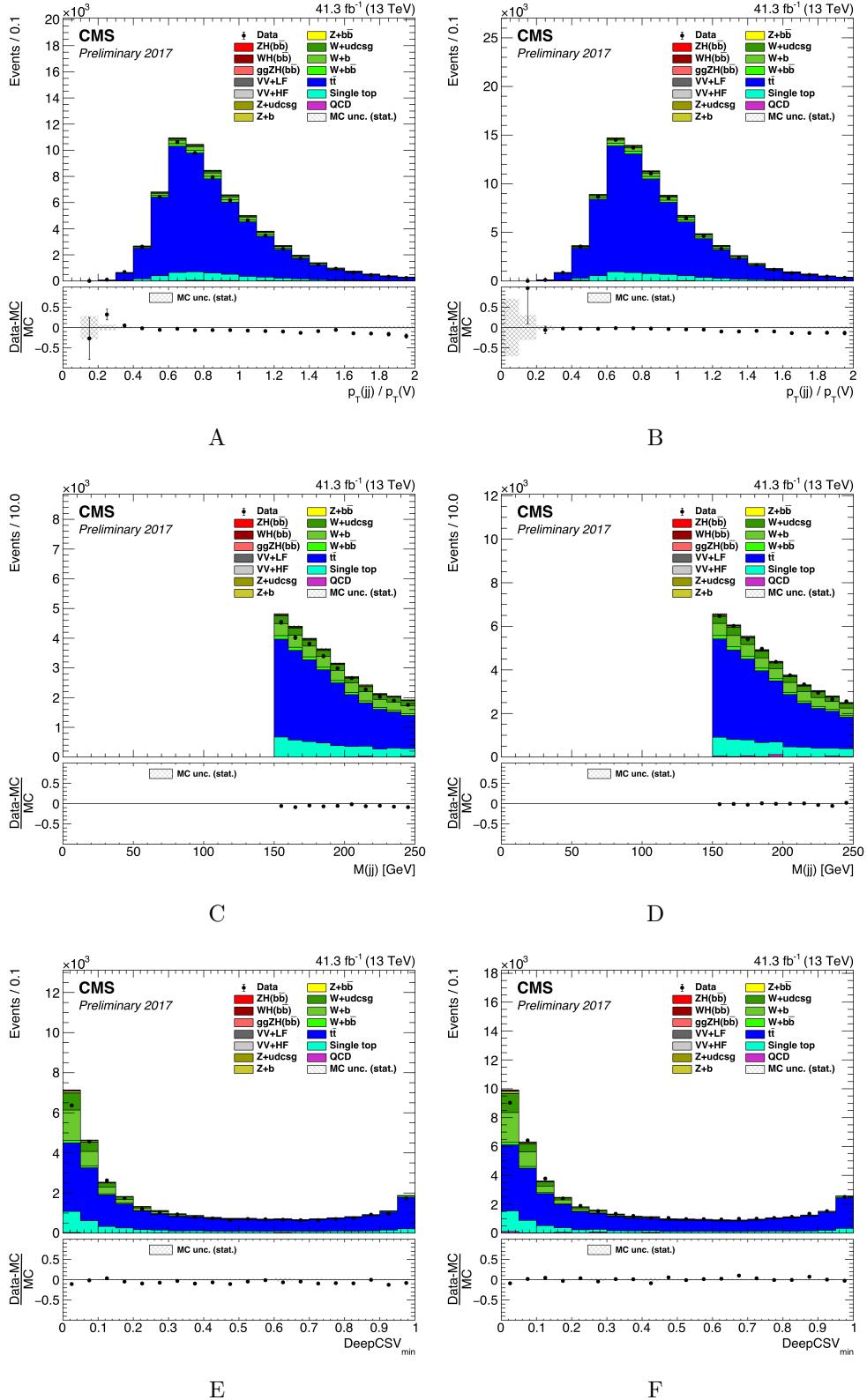


Figure B-7. The distributions of variables of the  $W(e\nu)H$  channel (left column) and the  $W(\mu\nu)H$  channel (right column): A), B)  $p_T(jj)/p_T(V)$  for the  $t\bar{t}$  control region; C), D)  $m(jj)$  for the  $W+H$  control region; E), F)  $b$ -tag<sub>min</sub> for the high mass  $W+H$  control region.

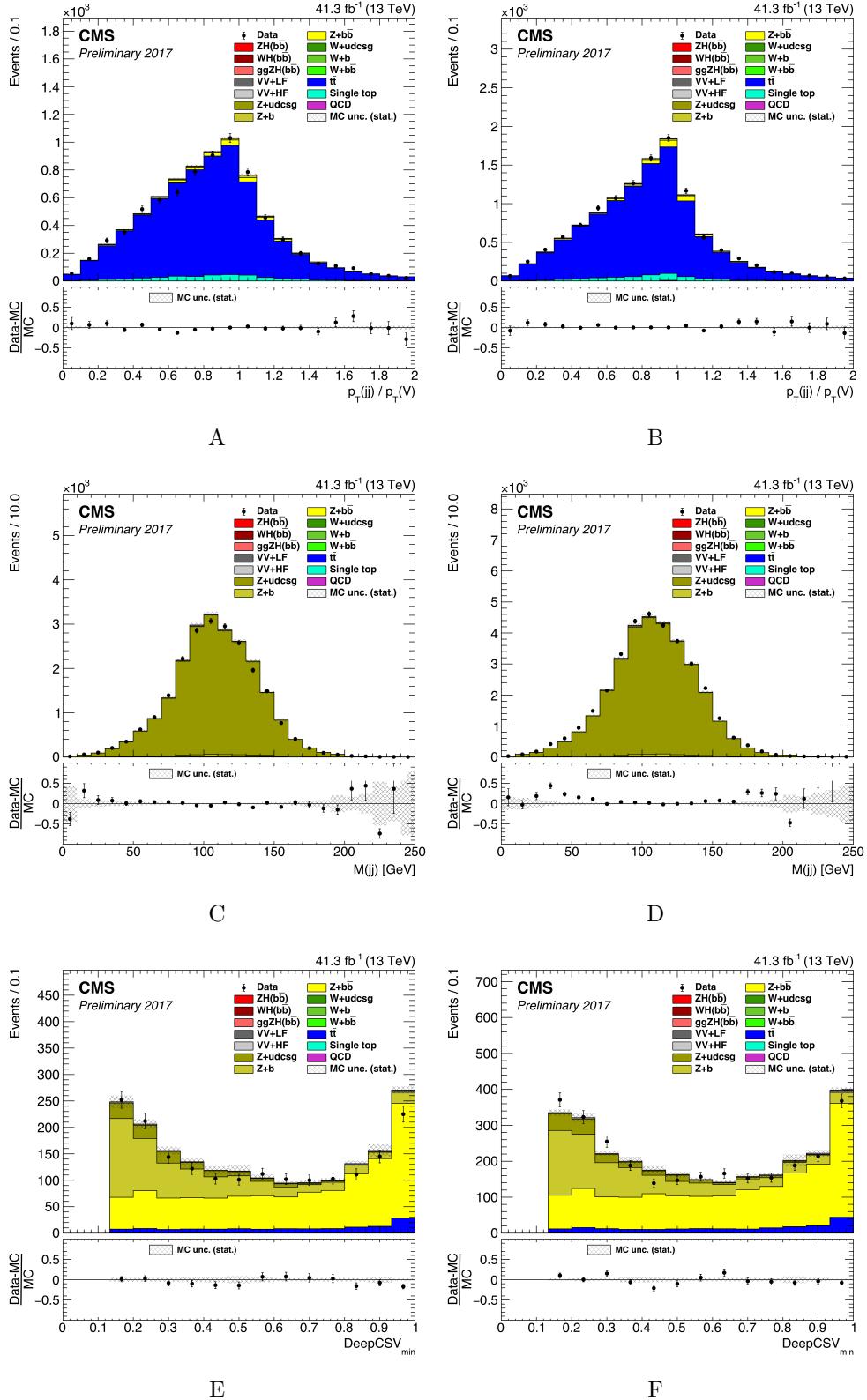


Figure B-8. The distributions of variables of the  $Z(e\bar{e})H$  channel (left column) and the  $Z(\mu\bar{\mu})H$  channel (right column) in the low  $p_T(V)$  region: A), B)  $p_T(jj)/p_T(V)$  for the  $t\bar{t}$  control region; C), D)  $m(jj)$  for the  $Z+light$  control region; E), F)  $b$ -tag<sub>min</sub> for the  $Z+heavy$  control region.

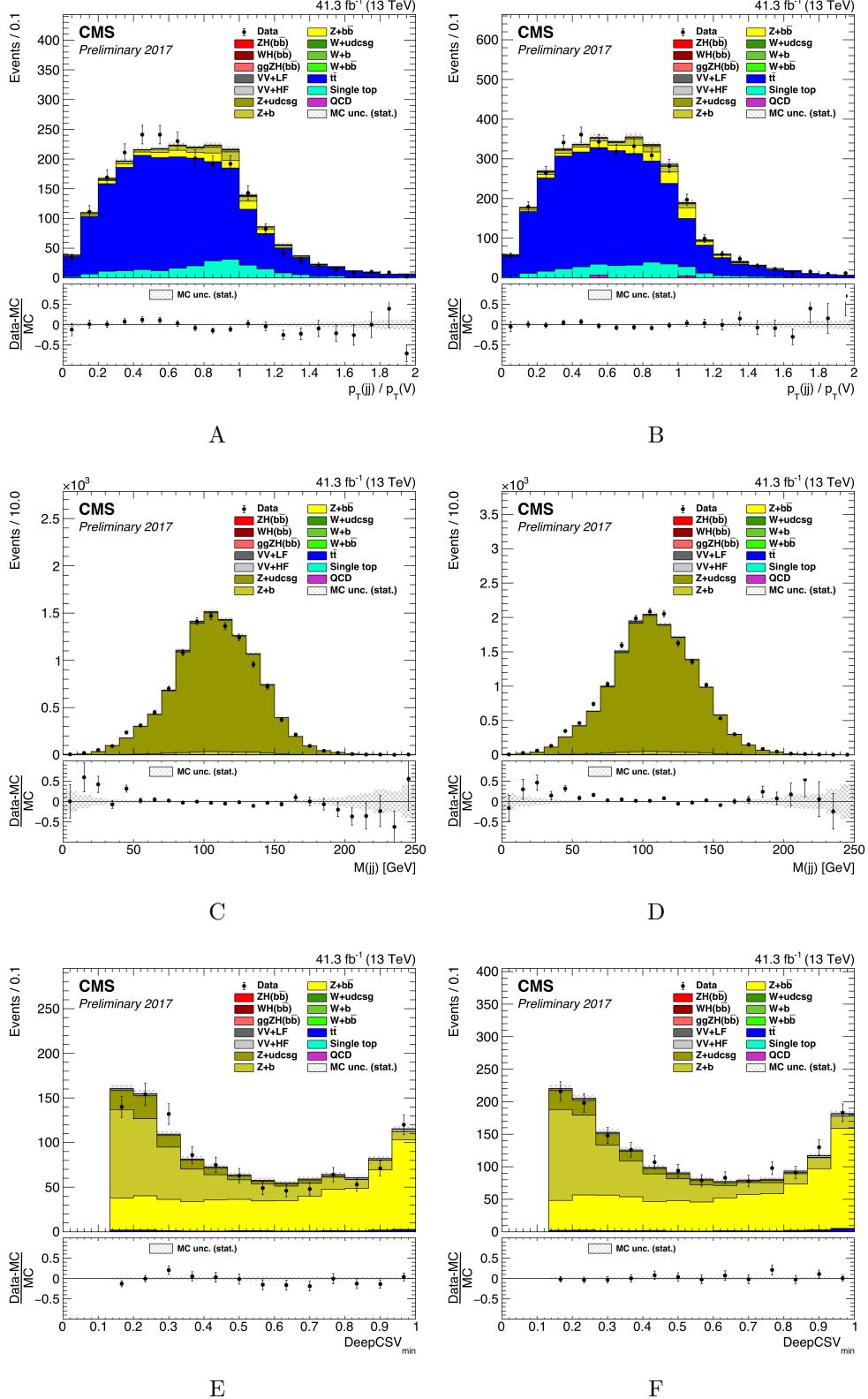


Figure B-9. The distributions of variables of the  $Z(e\bar{e})H$  channel (left column) and the  $Z(\mu\bar{\mu})H$  channel (right column) in the high  $p_T(V)$  region: A), B)  $p_T(jj)/p_T(V)$  for the  $t\bar{t}$  control region; C), D)  $m(jj)$  for the  $Z+$ light control region; E), F)  $b$ -tag<sub>min</sub> for the  $Z+$ heavy control region.

## REFERENCES

- [1] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory* (Westview Press, 1995).
- [2] M. Schwartz, *Quantum Field Theory and the Standard Model* (Cambridge University Press, 2014).
- [3] A. Purcell, Go on a particle quest at the first CERN webfest, [CERN Bulletin](#) **35**, 10 (2012).
- [4] S. Glashow, Partial-symmetries of weak interactions, [Nuclear Physics](#) **22**, 579 (1961).
- [5] A. Salam, Weak and electromagnetic interactions, *8th Nobel Symposium Lerum, Sweden, May 19-25, 1968*, Conf. Proc. **C680519**, 367 (1968).
- [6] S. Weinberg, A model of leptons, [Phys. Rev. Lett.](#) **19**, 1264 (1967).
- [7] F. Englert and R. Brout, Broken symmetry and the mass of gauge vector mesons, [Phys. Rev. Lett.](#) **13**, 321 (1964).
- [8] P. Higgs, Broken symmetries and the masses of gauge bosons, [Phys. Rev. Lett.](#) **13**, 508 (1964).
- [9] G. Guralnik, C. Hagen, and T. Kibble, Global conservation laws and massless particles, [Phys. Rev. Lett.](#) **13**, 585 (1964).
- [10] W. de Boer (CMS), The discovery of the Higgs boson with the CMS detector and its implications for supersymmetry and cosmology, in *Time and Matter 2013 (TAM2013) Venice, Italy* (2013) [arxiv:1309.0721 \[hep-ph\]](#).
- [11] M. Tanabashi *et al.* (Particle Data Group), Review of particle physics, [Phys. Rev. D](#) **98**, 030001 (2018).
- [12] J. Goldstone, Field theories with ⟨⟨ superconductor ⟩⟩ solutions, [Il Nuovo Cimento \(1955-1965\)](#) **19**, 154 (1961).
- [13] D. de Florian *et al.* (LHC Higgs Cross Section Working Group), *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, Tech. Rep. (2016) [arxiv:1610.07922 \[hep-ph\]](#).
- [14] ATLAS Collaboration, Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, [Phys. Lett. B](#) **716**, 1 (2012), [arxiv:1207.7214](#).
- [15] CMS Collaboration, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, [Phys. Lett. B](#) **716**, 30 (2012), [arxiv:1207.7235](#).

- [16] ATLAS Collaboration, Evidence for the spin-0 nature of the Higgs boson using ATLAS data, *Phys. Lett. B* **726**, 120 (2013).
- [17] CMS Collaboration, Study of the mass and spin-parity of the Higgs boson candidate via its decays to  $Z$  boson pairs, *Phys. Rev. Lett.* **110**, 081803 (2013).
- [18] ATLAS and CMS Collaborations, Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at  $\sqrt{s} = 7$  and 8 TeV, *Journal of High Energy Physics* **2016**, 45 (2016).
- [19] CMS Collaboration, Inclusive search for a highly boosted Higgs boson decaying to a bottom quark-antiquark pair, *Phys. Rev. Lett.* **120**, 071802 (2018).
- [20] J. Campbell, K. Ellis, W. Giele, T. Neumann, and C. Williams, MCFM - Monte Carlo for FeMtobarn processes (2018).
- [21] CMS Collaboration, Summaries of CMS cross section measurements (2018).
- [22] ALEPH, DELPHI, L3, and OPAL Collaborations, Search for the Standard Model Higgs boson at LEP, *Phys. Lett. B* **565**, 61 (2003).
- [23] CDF and D0 Collaborations, Evidence for a particle produced in association with weak bosons and decaying to a bottom-antibottom quark pair in Higgs boson searches at the Tevatron, *Phys. Rev. Lett.* **109**, 071804 (2012).
- [24] ATLAS Collaboration, Search for the  $b\bar{b}$  decay of the Standard Model Higgs boson in associated ( $W/Z$ ) $H$  production with the ATLAS detector, *Journal of High Energy Physics* **2015**, 69 (2015).
- [25] CMS Collaboration, Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV, *The European Physical Journal C* **75**, 212 (2015).
- [26] ATLAS and CMS Collaborations, Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at  $\sqrt{s} = 7$  and 8 TeV, *Journal of High Energy Physics* **2016**, 45 (2016).
- [27] ATLAS Collaboration, Evidence for the  $H \rightarrow b\bar{b}$  decay with the ATLAS detector, *Journal of High Energy Physics* **2017**, 24 (2017).
- [28] CMS Collaboration, Evidence for the Higgs boson decay to a bottom quark-antiquark pair, *Physics Letters B* **780**, 501 (2018).
- [29] P. Mouche, *Overall view of the LHC. Vue d'ensemble du LHC* (2014).
- [30] J. Haffner, *The CERN accelerator complex. Complexe des accélérateurs du CERN* (2013).

- [31] J. Caron, Cross section of LHC dipole. Dipole LHC: coupe transversale. (1998).
- [32] J. Caron, LHC quadrupole cross section. (1998).
- [33] CMS Collaboration, *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*, Technical Design Report CMS (CERN, Geneva, 2006).
- [34] T. Sakuma and T. McCauley, *Detector and event visualization with SketchUp at the CMS experiment*, Tech. Rep. CMS-CR-2013-379. arxiv:1311.4942 (CERN, Geneva, 2013).
- [35] R. Breedon *et al.*, CSC strip, wire, chamber, and electronics conventions (2007).
- [36] D. Barney, CMS detector slice. (2016).
- [37] CMS Collaboration, Alignment of the CMS silicon tracker during commissioning with cosmic rays, *Journal of Instrumentation* **5** (03), T03009.
- [38] CMS Collaboration, Description and performance of track and primary-vertex reconstruction with the CMS tracker, *Journal of Instrumentation* **9** (10), P10009.
- [39] A. Benaglia, The CMS ECAL performance with examples, *Journal of Instrumentation* **9** (02), C02008.
- [40] CMS Collaboration, Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at  $\sqrt{s} = 7$  TeV, *Journal of Instrumentation* **8** (09), P09009.
- [41] CMS Collaboration, The CMS experiment at the CERN LHC, *Journal of Instrumentation* **3** (08), S08004.
- [42] CMS Collaboration, Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at  $\sqrt{s} = 13$  TeV, *Journal of Instrumentation* **13** (06), P06015.
- [43] CMS Collaboration, The CMS trigger system, *Journal of Instrumentation* **12** (01), P01020.
- [44] S. Donato, CMS trigger performance, *EPJ Web Conf.* **182**, 02037 (2018).
- [45] CMS Collaboration, *CMS computing: Technical Design Report*, Technical Design Report CMS (CERN, Geneva, 2005).
- [46] CMS Collaboration, Particle-flow reconstruction and global event description with the CMS detector, *Journal of Instrumentation* **12** (10), P10003.
- [47] B. Mangano, Performance of particle flow in CMS (2013).
- [48] A. Holzner, 78 reconstructed vertices in event from high-pileup run 198609 (2012).

- [49] CMS Collaboration, Description and performance of track and primary-vertex reconstruction with the CMS tracker, *Journal of Instrumentation* **9** (10), P10009.
- [50] W. Adam, B. Mangano, T. Speer, and T. Todorov, *Track Reconstruction in the CMS tracker*, Tech. Rep. CMS-NOTE-2006-041 (CERN, Geneva, 2006).
- [51] A. Salzburger, Inner detector reconstruction part I - tracking (2008).
- [52] R. Frühwirth, Application of Kalman filtering to track and vertex fitting, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **262**, 444 (1987).
- [53] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov, Reconstruction of electrons with the gaussian-sum filter in the CMS tracker at the LHC, *Journal of Physics G: Nuclear and Particle Physics* **31**, N9 (2005).
- [54] CMS Collaboration, Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV, *Journal of Instrumentation* **10** (06), P06005.
- [55] CMS Collaboration, Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV, *Journal of Instrumentation* **10** (08), P08010.
- [56] CMS Collaboration, Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at  $\sqrt{s} = 13$  TeV, *Journal of Instrumentation* **13** (06), P06015.
- [57] W. Waltenberger, R. Frühwirth, and P. Vanlaer, Adaptive vertex fitting, *Journal of Physics G: Nuclear and Particle Physics* **34**, N343 (2007).
- [58] CMS Collaboration, Performance of  $\tau$ -lepton reconstruction and identification in CMS, *Journal of Instrumentation* **7** (01), P01001.
- [59] CMS Collaboration, Reconstruction and identification of  $\tau$  lepton decays to hadrons and  $\nu_\tau$  at CMS, *Journal of Instrumentation* **11** (01), P01019.
- [60] M. Cacciari, G. Salam, and G. Soyez, FastJet user manual, *The European Physical Journal C* **72**, 1896 (2012).
- [61] M. Cacciari, G. Salam, and G. Soyez, The anti-kt jet clustering algorithm, *Journal of High Energy Physics* **2008**, 063 (2008).
- [62] CMS Collaboration, Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV, *Journal of Instrumentation* **12** (02), P02014.
- [63] CMS Collaboration, Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV, *Journal of Instrumentation* **13** (05), P05011.
- [64] F. Chollet *et al.*, Keras, <https://keras.io> (2015).

- [65] M. Abadi *et al.*, TensorFlow: Large-scale machine learning on heterogeneous systems (2015), software available from tensorflow.org.
- [66] CMS Collaboration, Performance of the CMS missing transverse momentum reconstruction in pp data at  $\sqrt{s} = 8$  TeV, *Journal of Instrumentation* **10** (02), P02006.
- [67] T. Sakuma, Missing ET schematic diagram, <https://cms-docdb.cern.ch/cgi-bin/PublicDocDB>ShowDocument?docid=12312> (2014).
- [68] CMS Collaboration, Search for the standard model Higgs boson produced in association with a  $W$  or a  $Z$  boson and decaying to bottom quarks, *Phys. Rev. D* **89**, 012003 (2014).
- [69] CMS Collaboration, Observation of Higgs boson decay to bottom quarks, *Phys. Rev. Lett.* **121**, 121801 (2018).
- [70] S. Agostinelli *et al.*, Geant4 — a simulation toolkit, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506**, 250 (2003).
- [71] J. Allison *et al.*, Geant4 developments and applications, *IEEE Transactions on Nuclear Science* **53**, 270 (2006).
- [72] J. Allison *et al.*, Recent developments in Geant4, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **835**, 186 (2016).
- [73] P. Nason, A new method for combining NLO QCD with shower Monte Carlo algorithms, *Journal of High Energy Physics* **2004**, 040 (2004).
- [74] S. Frixione, P. Nason, and C. Oleari, Matching NLO QCD computations with parton shower simulations: the POWHEG method, *Journal of High Energy Physics* **2007**, 070 (2007).
- [75] S. Alioli, P. Nason, C. Oleari, and E. Re, A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX, *Journal of High Energy Physics* **2010**, 43 (2010).
- [76] K. Hamilton, P. Nason, and G. Zanderighi, MINLO: multi-scale improved NLO, *Journal of High Energy Physics* **2012**, 155 (2012).
- [77] G. Luisini, P. Nason, C. Oleari, and F. Tramontano,  $HW^\pm/HZ + 0$  and 1 jet at nlo with the POWHEG BOX interfaced to GoSam and their merging within MiNLO, *Journal of High Energy Physics* **2013**, 83 (2013).
- [78] G. Ferrera, M. Grazzini, and F. Tramontano, Higher-order QCD effects for associated WH production and decay at the LHC, *Journal of High Energy Physics* **2014**, 39 (2014).

- [79] G. Ferrera, M. Grazzini, and F. Tramontano, Associated ZH production at hadron colliders: The fully differential NNLO QCD calculation, *Physics Letters B* **740**, 51 (2015).
- [80] G. Ferrera, M. Grazzini, and F. Tramontano, Associated Higgs- $W$ -boson production at hadron colliders: A fully exclusive QCD calculation at NNLO, *Phys. Rev. Lett.* **107**, 152003 (2011).
- [81] G. Ferrera, G. Somogyi, and F. Tramontano, Associated production of a Higgs boson decaying into bottom quarks at the LHC in full NNLO QCD, *Physics Letters B* **780**, 346 (2018).
- [82] O. Brein, R. Harlander, and T. Zirke, vh@nnlo—Higgs strahlung at hadron colliders, *Computer Physics Communications* **184**, 998 (2013).
- [83] R. Harlander, S. Liebler, and T. Zirke, Higgs strahlung at the Large Hadron Collider in the 2-Higgs-doublet model, *Journal of High Energy Physics* **2014**, 23 (2014).
- [84] A. Denner, S. Dittmaier, S. Kallweit, and A. Mück, HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders, *Computer Physics Communications* **195**, 161 (2015).
- [85] J. Alwall *et al.*, The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, *Journal of High Energy Physics* **2014**, 79 (2014).
- [86] R. Frederix and S. Frixione, Merging meets matching in MC@NLO, *Journal of High Energy Physics* **2012**, 61 (2012).
- [87] T. Sjöstrand *et al.*, An introduction to PYTHIA 8.2, *Computer Physics Communications* **191**, 159 (2015).
- [88] CMS Collaboration (CMS Collaboration), *Measurement of the WW cross section pp collisions at  $\sqrt{s} = 13$  TeV*, Tech. Rep. CMS-PAS-SMP-16-006 (CERN, Geneva, 2016).
- [89] CMS Collaboration (CMS Collaboration), *Measurements of the  $pp \rightarrow WZ$  inclusive and differential production cross section and constraints on charged anomalous triple gauge couplings at  $\sqrt{s} = 13$  TeV*, Tech. Rep. CMS-PAS-SMP-18-002 (CERN, Geneva, 2018).
- [90] CMS Collaboration, Measurement of the ZZ production cross section and  $Z \rightarrow \ell^+ \ell^- \ell'^+ \ell'^-$  branching fraction in pp collisions at  $\sqrt{s} = 13$  TeV, *Physics Letters B* **763**, 280 (2016).
- [91] J. Alwall *et al.*, Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions, *The European Physical Journal C* **53**, 473 (2008).

- [92] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order, *Computer Physics Communications* **182**, 2388 (2011).
- [93] Y. Li and F. Petriello, Combining QCD and electroweak corrections to dilepton production in the framework of the FEWZ simulation code, *Phys. Rev. D* **86**, 094034 (2012).
- [94] S. Quackenbush, R. Gavin, Y. Li, and F. Petriello, W physics at the LHC with FEWZ 2.1, *Computer Physics Communications* **184**, 209 (2013).
- [95] S. Frixione, G. Ridolfi, and P. Nason, A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction, *Journal of High Energy Physics* **2007**, 126 (2007).
- [96] M. Czakon and A. Mitov, Top++: A program for the calculation of the top-pair cross-section at hadron colliders, *Computer Physics Communications* **185**, 2930 (2014).
- [97] R. Frederix, E. Re, and P. Torrielli, Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO, *Journal of High Energy Physics* **2012**, 130 (2012).
- [98] E. Re, Single-top Wt-channel production matched with parton showers using the POWHEG method, *The European Physical Journal C* **71**, 1547 (2011).
- [99] S. Alioli, P. Nason, C. Oleari, and E. Re, NLO single-top production matched with shower in POWHEG: s - and t -channel contributions, *Journal of High Energy Physics* **2009**, 111 (2009).
- [100] M. Czakon, P. Fiedler, and A. Mitov, Total top-quark pair-production cross section at hadron colliders through  $\mathcal{O}(\alpha_S^4)$ , *Phys. Rev. Lett.* **110**, 252004 (2013).
- [101] N. Kidonakis, Differential and total cross sections for top pair and single top production, *20th International Workshop on Deep-Inelastic Scattering and Related Subjects, DIS 2012* (2012).
- [102] R. Ball *et al.*, Parton distributions from high-precision collider data, *The European Physical Journal C* **77**, 663 (2017).
- [103] S. Kallweit, J. Lindert, P. Maierhöfer, S. Pozzorini, and M. Schönherr, NLO QCD+EW predictions for V + jets including off-shell vector-boson decays and multijet merging, *Journal of High Energy Physics* **2016**, 21 (2016).
- [104] CMS Collaboration, Measurement of differential cross sections for top quark pair production using the lepton + jets final state in proton-proton collisions at 13 TeV, *Phys. Rev. D* **95**, 092001 (2017).

- [105] CMS Collaboration, *Commissioning of TrackJets in pp Collisions at 7 TeV*, Tech. Rep. CMS-PAS-JME-10-006 (CERN, Geneva, 2010).
- [106] A. Maas, A. Hannun, and A. Ng, Rectifier nonlinearities improve neural network acoustic models, in *Proc. icml*, Vol. 30 (2013) p. 3.
- [107] S. Ioffe and C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift (2015), [arxiv:1502.03167 \[cs.LG\]](https://arxiv.org/abs/1502.03167) .
- [108] G. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, Improving neural networks by preventing co-adaptation of feature detectors (2012), [arxiv:1207.0580 \[cs.NE\]](https://arxiv.org/abs/1207.0580) .
- [109] P. Huber, Robust estimation of a location parameter, *Ann. Math. Statist.* **35**, 73 (1964).
- [110] D. Kingma and J. Ba, Adam: A method for stochastic optimization (2014), [arxiv:1412.6980 \[cs.LG\]](https://arxiv.org/abs/1412.6980) .
- [111] J. Friedman, Greedy function approximation: A gradient boosting machine., *Ann. Statist.* **29**, 1189 (2001).
- [112] P. Speckmayer, A. Höcker, J. Stelzer, and H. Voss, The toolkit for multivariate data analysis, TMVA 4, *Journal of Physics: Conference Series* **219**, 032057 (2010).
- [113] F. Rosenblatt, The perceptron: A probabilistic model for information storage and organization in the brain (1958).
- [114] K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition, *CoRR* **abs/1512.03385** (2015), arXiv:1512.03385 .
- [115] J. Thomas, Confidence level computation for combining searches with small statistics, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **434**, 435 (1999).
- [116] A. Read, Presentation of search results: the  $CL_s$  technique, *Journal of Physics G: Nuclear and Particle Physics* **28**, 2693 (2002).
- [117] ATLAS and CMS Collaborations, LHC Higgs Combination Group, *Procedure for the LHC Higgs boson search combination in summer 2011*, Tech. Rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11 (CERN, Geneva, 2011).
- [118] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Asymptotic formulae for likelihood-based tests of new physics, *The European Physical Journal C* **71**, 1554 (2011).
- [119] A. Carnes, Autocategorizer, <https://github.com/acarnes/bdt> (2018).

- [120] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Asymptotic formulae for likelihood-based tests of new physics, *The European Physical Journal C* **71**, 1554 (2011).
- [121] CMS Collaboration, *CMS luminosity measurement for the 2017 data-taking period at  $\sqrt{s} = 13$  TeV*, Tech. Rep. CMS-PAS-LUM-17-004 (CERN, Geneva, 2018).
- [122] S. Dittmaier *et al.* (LHC Higgs Cross Section Working Group), *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, CERN Yellow Reports: Monographs (CERN, Geneva, 2011).
- [123] M. Ciccolini, A. Denner, and S. Dittmaier, Strong and electroweak corrections to the production of a Higgs Boson + 2 Jets via weak interactions at the Large Hadron Collider, *Phys. Rev. Lett.* **99**, 161803 (2007).
- [124] M. Ciccolini, A. Denner, and S. Dittmaier, Electroweak and QCD corrections to Higgs production via vector-boson fusion at the CERN LHC, *Phys. Rev. D* **77**, 013002 (2008).
- [125] *NLO QCD and electroweak corrections to Higgs strahlung off W/Z bosons at Tevatron and the LHC with HAWK*, Vol. 134 (Proceedings of Science, 2012).
- [126] G. Ferrera, M. Grazzini, and F. Tramontano, Associated Higgs-W-Boson production at hadron colliders: A fully exclusive QCD calculation at NNLO, *Phys. Rev. Lett.* **107**, 152003 (2011).
- [127] CMS Collaboration, Cms higgs combination toolkit, <https://github.com/cms-analysis/HiggsAnalysis-CombinedLimit> (2018).
- [128] W. Verkerke and D. Kirkby, Rooftit users manual v2.91, [https://root.cern.ch/download/doc/RooFit\\_Users\\_Manual\\_2.91-33.pdf](https://root.cern.ch/download/doc/RooFit_Users_Manual_2.91-33.pdf) (2008).
- [129] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke (ROOT Collaboration), *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, Tech. Rep. CERN-OPEN-2012-016 (New York U., New York, 2012).
- [130] T. Junk, Confidence level computation for combining searches with small statistics, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **434**, 435 (1999).
- [131] CMS Collaboration, Search for the standard model higgs boson produced through vector boson fusion and decaying to  $b\bar{b}$ , *Phys. Rev. D* **92**, 032008 (2015).
- [132] CMS Collaboration, Observation of  $t\bar{t}H$  production, *Phys. Rev. Lett.* **120**, 231801 (2018).

- [133] CMS Collaboration, Search for the standard model Higgs boson produced in association with a top-quark pair in pp collisions at the LHC, *Journal of High Energy Physics* **2013**, 145 (2013).
- [134] CMS Collaboration, Search for the associated production of the Higgs boson with a top-quark pair, *Journal of High Energy Physics* **2014**, 87 (2014).
- [135] CMS Collaboration, Search for the Higgs boson decaying to two muons in proton-proton collisions at  $\sqrt{s} = 13$  TeV, *Phys. Rev. Lett.* **122**, 021801 (2019).
- [136] C. Shimmin, P. Sadowski, P. Baldi, E. Weik, D. Whiteson, E. Goul, and A. Søgaard, Decorrelated jet substructure tagging using adversarial neural networks, *Phys. Rev. D* **96**, 074034 (2017).
- [137] D. Bertolini, P. Harris, M. Low, and N. Tran, Pileup per particle identification, *Journal of High Energy Physics* **2014**, 59 (2014).
- [138] J. Butterworth, A. Davison, M. Rubin, and G. Salam, Jet substructure as a new Higgs-search channel at the Large Hadron Collider, *Phys. Rev. Lett.* **100**, 242001 (2008).
- [139] M. Stoye, J. Kieseler, M. Verzetti, H. Qu, L. Gouskos, and A. Stakia, DeepJet: Generic physics object based jet multiclass classification for LHC experiments (2017).
- [140] A. Elwood and D. Krücker, Direct optimisation of the discovery significance when training neural networks to search for new physics in particle colliders (2018), arxiv:1806.00322 [hep-ex].
- [141] D. Kovalskyi *et al.*, Fireworks: A physics event display for CMS, *Journal of Physics: Conference Series* **219**, 032014 (2010)

## BIOGRAPHICAL SKETCH

Sean-Jiun Wang was born in West Palm Beach, Florida. After graduating from Suncoast High School with an interest in physics, he attended the University of Florida where he obtained a Bachelor of Science in physics and mathematics. For his graduate education, he remained at the University of Florida to pursue experimental high energy physics. Upon completing his coursework, he moved to the Franco-Swiss border where he began his graduate research with the CMS experiment at CERN. A year later, he was awarded the 2016 LPC Graduate Scholar fellowship and moved to Fermi National Accelerator Laboratory in Illinois where he would work closely with his advisor Dr. Jacobo Konigsberg until the end of his graduate research. He received a Doctor of Philosophy in the spring of 2019.