

NORTHWESTERN UNIVERSITY

Search for Higgs Boson Decays to a Z Boson and a Photon

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Physics

By

James Bueghly

EVANSTON, ILLINOIS

June 2022

© Copyright by James Bueghly 2022

All Rights Reserved

## ABSTRACT

Search for Higgs Boson Decays to a Z Boson and a Photon

James Bueghly

Since its discovery in 2012 at the Large Hadron Collider (LHC), efforts have been made to measure and characterize the properties of the Higgs boson. Among these efforts have been searches for rare decays of the Higgs predicted by the standard model (SM) of particle physics. One such decay is the process  $H \rightarrow Z\gamma$ , which has an expected branching fraction of  $\mathcal{B}(H \rightarrow Z\gamma) = (1.57 \pm 0.09) \times 10^{-3}$  in the SM, assuming a Higgs boson mass of  $m_H = 125.38$  GeV. This decay mode has not yet been experimentally observed, and its observation and measurement remains an important goal of Higgs physics research at the LHC. In addition, a measurement of this decay mode at a rate deviating from the SM prediction would provide indirect evidence of new physics beyond the SM.

This thesis presents a search for  $H \rightarrow Z\gamma$ , where  $Z \rightarrow \ell^+\ell^-$  with  $\ell = e$  or  $\mu$ . The search is performed using a sample of proton-proton (pp) collision data at a center-of-mass energy of 13 TeV, recorded by the Compact Muon Solenoid experiment at the LHC, corresponding to an integrated luminosity of  $138 \text{ fb}^{-1}$ . Events are assigned to mutually exclusive categories, which exploit differences in both event topology and kinematics of

distinct Higgs production mechanisms to enhance signal sensitivity. To detect a potential signal, fits are performed to the distributions of  $\ell^+\ell^-\gamma$  invariant mass in each of these categories simultaneously. The signal strength  $\mu$ , defined as the product of the cross section and the branching fraction [ $\sigma(pp \rightarrow H)\mathcal{B}(H \rightarrow Z\gamma)$ ] relative to the SM expectation, is found to be  $\mu = 2.4 \pm 0.9$  for  $m_H = 125.38$  GeV. This measurement corresponds to  $\sigma(pp \rightarrow H)\mathcal{B}(H \rightarrow Z\gamma) = 0.21 \pm 0.08$  pb. The statistical significance of the observed excess of events is 2.7 standard deviations. The observed (expected) upper limit at 95% confidence level on  $\mu$  is 4.1 (1.8). The ratio of branching fractions  $\mathcal{B}(H \rightarrow Z\gamma)/\mathcal{B}(H \rightarrow \gamma\gamma)$  is measured to be  $1.5^{+0.7}_{-0.6}$ , which agrees with the SM prediction at the 1.5 standard deviation level.

## Acknowledgements

I would like to thank my adviser, Mayda Velasco, for the opportunity to pursue a goal I dreamed up years ago and for her support along the way. I'd also like to thank my many colleagues at Northwestern University, National Central University in Taiwan, CERN, and Fermilab who helped me grow both as a scientist and a person over the course of my Ph.D.

Special thanks are due to those who significantly impacted the work contained in this thesis: My coanalyzer Ming-Yan Lee, whose discipline and tireless work ethic kept us all going through the many challenges we faced; postdoc Andrew Gilbert, whose expertise and guidance were critical; and the CMS leaders who guided and reviewed our work, including the Higgs conveners, Analysis Review Committee, and our colleague and language editor Jeff Richman. Without all of you, this thesis would not exist.

To my Northwestern colleagues: I will always be grateful for your friendship over these years. The moments we shared away from the work kept me sane, focused, and balanced. It is impossible to name all of them here, but the squash games with Nate Odell, the lunches with Nate, Brian Pollack, Thoth Gunter, and Joseph Cordero Mercado, and our adventures at COFI in San Juan were all very meaningful to me.

Last, but not least, I want to thank my parents, sisters, and my partner Anna. You all have shown me so much love, encouragement, and patience throughout this process. For that, I feel extremely fortunate. This is for all of you.

## Table of Contents

ABSTRACT	3
Acknowledgements	5
Table of Contents	6
List of Tables	9
List of Figures	11
Chapter 1. Introduction	22
Chapter 2. Theory	25
2.1. The Standard Model	25
2.2. The Elementary Particles	25
2.3. Electroweak Theory	27
2.4. Spontaneous Symmetry Breaking (Higgs Mechanism)	28
2.5. Higgs Production	34
2.6. Higgs Decay	35
2.7. Physics Beyond the Standard Model	36
Chapter 3. Experiment Description	40
3.1. The Large Hadron Collider	40

3.2. The Compact Muon Solenoid	42
3.3. Trigger System	48
3.4. Object Reconstruction	50
Chapter 4. Analysis Overview	57
4.1. History of the $H \rightarrow Z\gamma$ Search	57
4.2. The CMS Run 2 Strategy	60
Chapter 5. Data and Simulated Samples	62
5.1. $Z+jets$ Overlap Removal	65
5.2. Photon Internal Conversion	65
5.3. Transverse Momentum Reweighting	66
5.4. Pileup Reweighting	67
5.5. Level-1 Trigger Prefiring Reweighting	68
Chapter 6. Physics Object and Event Selection	70
6.1. Primary Vertex	70
6.2. Triggers	70
6.3. Photon Selection	72
6.4. Electron Selection	74
6.5. Muon Selection	75
6.6. Jet Selection	76
6.7. Object Corrections	77
6.8. Event Selection	78
6.9. Kinematic Fit	79

Chapter 7. Event Categorization	81
7.1. Kinematic Boosted Decision Tree	82
7.2. Vector Boson Fusion Boosted Decision Tree	83
7.3. Categorization Procedure	84
7.4. Dropping of Boosted Category	88
Chapter 8. Signal and Background Modeling	90
8.1. Signal Modeling	90
8.2. Resonant Background Modeling	91
8.3. Nonresonant Background Modeling	91
Chapter 9. Systematic Uncertainties	109
Chapter 10. Statistical Analysis	113
Chapter 11. Results and Interpretation	117
11.1. Ratio Measurement	120
Chapter 12. Conclusion	126
References	127
Appendix A. Signal and Resonant Background Fits	136
A.1. 2017 Signal Fits	136
A.2. 2018 Signal Fits	136
A.3. 2017 Resonant Background Fits	136
A.4. 2018 Resonant Background Fits	136

## List of Tables

5.1	Summary of MINIAOD data samples used.	63
5.2	Simulated event samples used in the analysis. For Higgs samples, the listed cross sections and branching fractions correspond to the values for $m_H = 125$ GeV.	64
6.1	Percent improvement in dilepton and three-body mass resolution (measured by $\sigma_{eff}$ ) after the kinematic fit.	80
7.1	Summary of the category definitions. The lepton-tagged category requires at least one additional electron or muon. Dijet categories are defined by regions of $\mathcal{D}_{VBF}$ and untagged categories are defined by regions of $\mathcal{D}_{kin}$ .	86
7.2	Yields and approximate significance ( $S/\sqrt{B}$ ) for each category, where $S$ and $B$ are the expected number of signal and background events in the narrowest $m_{\ell^+\ell^-\gamma}$ interval containing 95% of the expected signal distribution. Also shown is the $m_{\ell^+\ell^-\gamma}$ resolution, computed using the narrowest interval containing 68% of the expected signal distribution.	87
8.1	Bias and coverage results for each category.	107

- 9.1 Sources of systematic uncertainty affecting the simulated signal. The  
normalization effect on the expected yield, or the effect on the signal  
shape parameters, is given as indicated, with the values averaged over  
all event categories. The third column shows the uncertainties that  
have a correlated effect across the three data-taking periods. 112
- 11.1 Correlated uncertainties between  $H \rightarrow Z\gamma$  and  $H \rightarrow \gamma\gamma$  for the  
 $\mathcal{B}(H \rightarrow Z\gamma)/\mathcal{B}(H \rightarrow \gamma\gamma)$  measurement. 123

## List of Figures

1.1	Feynman diagrams for $H \rightarrow Z\gamma$ decay.	23
2.1	Table of elementary particles in the SM. The leftmost three columns correspond to three generations of fermions, the fourth column from the left shows the gauge bosons which mediate the elementary forces, and the rightmost column shows the Higgs boson. For each particle, mass, charge, and spin are labeled, with mass values given as of 2019.	26
2.2	Comparison of the shapes of complex scalar potentials without symmetry breaking (left) and with spontaneous symmetry breaking (right). The transverse axes represent the real-imaginary $\Phi$ plane, and the vertical axis represents the magnitude of the potential $V(\Phi)$ . The diagram at right corresponds to the shape of the Higgs potential in the SM.	30
2.3	Cross sections for different Higgs boson production mechanisms as a function of center of mass energy for $m_H = 125$ GeV.	34
2.4	Branching fractions for different Higgs decay modes as a function of $m_H$ .	35
2.5	Contours of constant branching fraction enhancement relative to the SM for the $W'$ (left), charged scalar (center), and charged lepton	

(right) scenarios. The solid lines represent contours of enhancement in  $\mathcal{B}(H \rightarrow Z\gamma)$ , while the dashed lines correspond to  $H \rightarrow \gamma\gamma$ . The yellow boxed numbers label the enhancement for each  $H \rightarrow Z\gamma$  contour, and the unboxed numbers label the enhancement for each  $H \rightarrow \gamma\gamma$  contour.

37

2.6	Ratio of decay widths $\Gamma(H \rightarrow Z\gamma)/\Gamma(H \rightarrow \gamma\gamma)$ as a function of the mass coefficient of the additional scalar field(s) for seven BSM models with extended Higgs sectors. The models plotted on the left correspond to one additional singly-charged boson, and those on the right to one additional singly-charged and one additional doubly-charged boson.	38
2.7	Examples of two-loop diagrams contributing to $H \rightarrow Z\gamma$ decay.	39
3.1	CERN accelerator complex.	41
3.2	CMS detector apparatus.	43
3.3	Cross section of the CMS inner tracking system in the r-z plane.	44
3.4	Response of the CMS ECAL crystals to injected laser light as a function of time, broken into regions of $ \eta $ .	46
3.5	Diagram of the CMS detector in the r-z plane showing the components of the muon system.	48
3.6	Visualization of different types of particles interacting with the CMS detector.	51

3.7	Energy resolution for simulated electrons in the barrel (left) and endcaps (right) as a function of generated $p_{\text{T}}$ . The green markers represent tracker measurement only, the orange markers represent ECAL measurement only, and the purple markers represent the combination of tracker and ECAL information.	52
3.8	Left: dimuon mass resolution as a function of leading muon $\eta$ for $Z \rightarrow \mu^+ \mu^-$ events in 2018 data and simulation. Right: reconstructed dimuon mass.	54
3.9	An example of the first few steps of the anti- $k_{\text{T}}$ jet clustering algorithm.	55
3.10	Jet energy resolution as a function of simulated jet $p_{\text{T}}$ in the barrel (left) and in the endcap (right) regions. Calo refers to the sum of ECAL and HCAL energy deposits, while PF corresponds to jets clustered from PF particles using the anti- $k_{\text{T}}$ algorithm.	56
4.1	ATLAS (left) and CMS (right) Run 1 limit results as a function of $m_{\text{H}}$ .	59
4.2	Left: CMS 2016 limit results as a function of $m_{\text{H}}$ . Right: ATLAS Run 2 signal plus background fit, for $m_{\text{H}} = 125.09 \text{ GeV}$ , weighted by $\ln(1+S_{68}/B_{68})$ , where $S_{68}$ and $B_{68}$ are the expected signal and background events in an $m_{\ell^+\ell^-\gamma}$ window containing 68% of the expected signal yield.	60

- 5.1 Left: Background simulation modeling of  $p_T^{\ell^+\ell^-\gamma}$  in each data-taking year. There is a shape discrepancy between 2016 and 2017–2018 due to the updated underlying event tuning. Right: Final comparison between data and simulation after the  $p_T^{\ell^+\ell^-\gamma}$  reweighting is performed on 2017 and 2018 background simulation. 67
- 5.2 Distributions of the number of primary vertices in data and simulation for  $e^+e^-\gamma$  events in 2016 (left), 2017 (center), and 2018 (right) after pileup reweighting. 68
- 7.1 The  $\mathcal{D}_{\text{VBF}}$  (left) and  $\mathcal{D}_{\text{kin}}$  (right) distributions for signal, simulated background, and data. The  $\mathcal{D}_{\text{VBF}}$  distribution includes only dijet-tagged events, and the  $\mathcal{D}_{\text{kin}}$  distribution includes only untagged events. The sum of contributions from all signal production mechanisms is shown by the blue line, while the contribution from only the VBF mechanism is shown by the red line. The uncertainty band incorporates all statistical and systematic uncertainties in the expected background. The dashed lines indicate the boundaries for the dijet and untagged categories. The gray shaded region in the  $\mathcal{D}_{\text{kin}}$  distribution is excluded from the analysis. 86
- 7.2 Left:  $\ell\ell\gamma p_T$  for different signal MC processes. Right: optimization of the boosted category by scanning the expected limit as function of  $\ell\ell\gamma p_T$ . 89

7.3	Left: $\ell\ell\gamma p_T$ for untagged events and dijet events. Right: The $m_{\ell\ell\gamma}$ spectrum of untagged events and dijet events.	89
8.1	Fits to simulated $m_{\ell^+\ell^-\gamma}$ signal distributions in the electron channel for $m_H = 125$ GeV for the 2016 data-taking period. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component. The top four plots correspond to the untagged categories, and the bottom three plots correspond to the dijet categories.	92
8.2	Fits to simulated $m_{\ell^+\ell^-\gamma}$ signal distributions in the muon channel for $m_H = 125$ GeV for the 2016 data-taking period. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component. The top four plots correspond to the untagged categories, and the bottom three plots correspond to the dijet categories.	93
8.3	Fits to simulated $m_{\ell^+\ell^-\gamma}$ signal distributions in the electron and muon channels combined in the lepton-tagged category for $m_H = 125$ GeV for the 2016 data-taking period. The left plot shows the fit to simulated ZH production events, and the right plot shows the fit to simulated WH production events. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component.	94

- 8.4 Fits to simulated  $m_{\mu^+\mu^-\gamma}$  resonant background distributions from  $H \rightarrow \mu^+\mu^-$  for  $m_H = 125$  GeV for the 2016 data-taking period. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component. The top four plots correspond to the untagged categories, and the bottom three plots correspond to the dijet categories. 95
- 8.5 Fits to simulated  $m_{\ell^+\ell^-\gamma}$  resonant background distributions from  $H \rightarrow \mu^+\mu^-$  in the electron and muon channels combined in the lepton-tagged category for  $m_H = 125$  GeV for the 2016 data-taking period. The left plot shows the fit to simulated ZH production events, and the right plot shows the fit to simulated WH production events. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component. 96
- 8.6 Background functions determined by the discrete profiling method. The top four plots correspond to the untagged categories, and the bottom four plots correspond to the dijet categories and lepton tag category. 103
- 8.7 Comparison of background fits with and without modeling the turn-on. The top four plots correspond to the untagged categories, and the bottom four plots correspond to the dijet categories and lepton tag category. 104

11.1 Fits to the  $m_{\ell^+\ell^-\gamma}$  data distribution in the lepton-tagged (upper left), dijet 1 (upper right), dijet 2 (lower left), and dijet 3 (lower right) categories. In the upper panel, the red solid line shows the result of a signal-plus-background fit to the given category. The red dashed line shows the background component of the fit. The green and yellow bands represent the 68 and 95% CL uncertainties in the fit. Also plotted is the expected SM signal, scaled by a factor of 10. In the lower panel, the data minus the background component of the fit is shown.

118

11.2 Fits to the  $m_{\ell^+\ell^-\gamma}$  data distribution in the untagged 1 (upper left), untagged 2 (upper right), untagged 3 (lower left), and untagged 4 (lower right) categories. In the upper panel, the red solid line shows the result of a signal-plus-background fit to the given category. The red dashed line shows the background component of the fit. The green and yellow bands represent the 68 and 95% CL uncertainties in the fit. Also plotted is the expected SM signal, scaled by a factor of 10. In the lower panel, the data minus the background component of the fit is shown.

119

11.3 Sum over all categories of the data points and signal-plus-background model after the simultaneous fit to each  $m_{\ell^+\ell^-\gamma}$  distribution. The contribution from each category is weighted by  $S/(S + B)$ , as defined in the text. In the upper panel, the red solid line shows the signal-plus-background fit. The red dashed line shows the background

component of the fit. The green and yellow bands represent the 68 and 95% CL uncertainties in the fit. Also plotted is the expected SM signal weighted by  $S/(S + B)$  and scaled by a factor of 10. In the lower panel, the data minus the background component of the fit is shown.

120

- 11.4 Observed signal strength ( $\mu$ ) for a SM Higgs boson with  $m_H = 125.38 \text{ GeV}$ . The labels “untagged combined,” “dijet combined,” and “combined” represent the results obtained from simultaneous fits of the untagged categories, dijet categories, and full set of categories, respectively. The black solid line shows  $\mu = 1$ , and the red dashed line shows the best fit value  $\hat{\mu} = 2.4 \pm 0.9$  of all categories combined. 121

- 11.5 Upper limit (95% CL) on the signal strength ( $\mu$ ) relative to the SM prediction, as a function of the assumed value of the Higgs boson mass used in the fit. 122

- 11.6 Profile likelihood scans of  $\mu_{\gamma\gamma}$  and  $\mu_{Z\gamma}/\mu_{\gamma\gamma}$  with fixed  $m_H = 125.38 \text{ GeV}$ . 124

- 11.6 Nuisance parameter impacts on  $\mu_{Z\gamma}/\mu_{\gamma\gamma}$ . 125

- A.1 Fits to simulated  $m_{\ell^+\ell^-\gamma}$  signal distributions in the electron channel for  $m_H = 125 \text{ GeV}$  for the 2017 data-taking period. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function

component. The top four plots correspond to the untagged categories,  
and the bottom three plots correspond to the dijet categories. 137

A.2 Fits to simulated  $m_{\ell^+\ell^-\gamma}$  signal distributions in the muon channel for  
 $m_H = 125$  GeV for the 2017 data-taking period. The blue line shows  
the total fit function, the green line shows the Crystal Ball function  
component, and the red line shows the Gaussian function component.

The top four plots correspond to the untagged categories, and the  
bottom three plots correspond to the dijet categories. 138

A.3 Fits to simulated  $m_{\ell^+\ell^-\gamma}$  signal distributions in the electron and muon  
channels combined in the lepton-tagged category for  $m_H = 125$  GeV  
for the 2017 data-taking period. The left plot shows the fit to  
simulated ZH production events, and the right plot shows the fit to  
simulated WH production events. The blue line shows the total fit  
function, the green line shows the Crystal Ball function component,  
and the red line shows the Gaussian function component. 139

A.4 Fits to simulated  $m_{\ell^+\ell^-\gamma}$  signal distributions in the electron channel  
for  $m_H = 125$  GeV for the 2018 data-taking period. The blue line  
shows the total fit function, the green line shows the Crystal Ball  
function component, and the red line shows the Gaussian function  
component. The top four plots correspond to the untagged categories,  
and the bottom three plots correspond to the dijet categories. 140

A.5 Fits to simulated  $m_{\ell^+\ell^-\gamma}$  signal distributions in the muon channel for  
 $m_H = 125$  GeV for the 2018 data-taking period. The blue line shows

the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component.

The top four plots correspond to the untagged categories, and the bottom three plots correspond to the dijet categories. 141

A.6 Fits to simulated  $m_{\ell^+\ell^-\gamma}$  signal distributions in the electron and muon channels combined in the lepton-tagged category for  $m_H = 125$  GeV for the 2018 data-taking period. The left plot shows the fit to simulated ZH production events, and the right plot shows the fit to simulated WH production events. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component. 142

A.7 Fits to simulated  $m_{\mu^+\mu^-\gamma}$  resonant background distributions from  $H \rightarrow \mu^+\mu^-$  for  $m_H = 125$  GeV for the 2017 data-taking period. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component. The top four plots correspond to the untagged categories, and the bottom three plots correspond to the dijet categories. 143

A.8 Fits to simulated  $m_{\ell^+\ell^-\gamma}$  resonant background distributions from  $H \rightarrow \mu^+\mu^-$  in the electron and muon channels combined in the lepton-tagged category for  $m_H = 125$  GeV for the 2017 data-taking period. The left plot shows the fit to simulated ZH production events, and the right plot shows the fit to simulated WH production events. The blue line shows the total fit function, the green line shows the

Crystal Ball function component, and the red line shows the Gaussian function component.

144

A.9 Fits to simulated  $m_{\mu^+\mu^-\gamma}$  resonant background distributions from  $H \rightarrow \mu^+\mu^-$  for  $m_H = 125$  GeV for the 2018 data-taking period. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component. The top four plots correspond to the untagged categories, and the bottom three plots correspond to the dijet categories.

145

A.10 Fits to simulated  $m_{\ell^+\ell^-\gamma}$  resonant background distributions from  $H \rightarrow \mu^+\mu^-$  in the electron and muon channels combined in the lepton-tagged category for  $m_H = 125$  GeV for the 2018 data-taking period. The left plot shows the fit to simulated ZH production events, and the right plot shows the fit to simulated WH production events. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component.

146

## CHAPTER 1

### Introduction

The standard model (SM) of particle physics is one of the most successful physics theories in history. It describes all of the fundamental particles we currently know in our universe, including the ways in which those particles interact with one another. Since its development in the second half of the twentieth century, the SM has been extensively tested through many measurements by numerous experiments. The Large Hadron Collider (LHC) was built with the goal of further testing the SM, while also searching for possible beyond the SM (BSM) physics at groundbreaking energy scales. One of the main goals of the LHC was to search for and observe the Higgs boson, a critical piece of the SM associated with the origin of particle mass.

Since the discovery of the Higgs boson [1, 2, 3] at the LHC, an extensive program of measurements [4] has been undertaken to determine its properties and couplings to different types of particles and to assess whether these properties are consistent with those predicted by the SM. With the successful running of the LHC, large data samples of proton-proton (pp) collisions at  $\sqrt{s} = 13$  TeV have been accumulated, increasing the sensitivity to rare decays of the Higgs boson. Such decays also provide probes for possible contributions arising from BSM physics and include the process  $H \rightarrow Z\gamma$  [5, 6, 7, 8, 9, 10, 11].

Figure 1.1 shows Feynman diagrams for the key SM contributions to the  $H \rightarrow Z\gamma$  decay process. Experimentally, the final state resulting from  $Z \rightarrow \ell^+\ell^-$  ( $\ell = e$  or  $\mu$ ) is the most

accessible, since the leptons are highly distinctive, well-measured, and provide a means to trigger the recording of the events. In the SM, the expected branching fraction for  $H \rightarrow Z\gamma$  is  $\mathcal{B}(H \rightarrow Z\gamma) = (1.57 \pm 0.09) \times 10^{-3}$ , assuming a Higgs boson mass of  $m_H = 125.38$  GeV. This branching fraction is comparable to  $\mathcal{B}(H \rightarrow \gamma\gamma) = (2.27 \pm 0.04) \times 10^{-3}$  [12, 13]. The value  $m_H = 125.38 \pm 0.14$  GeV is taken from the most recent Compact Muon Solenoid (CMS) Higgs boson mass measurement [14], which uses the combination of  $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ^* \rightarrow 4\ell$  results from the 2011–2012 and 2016 data samples. The ratio  $\mathcal{B}(H \rightarrow Z\gamma)/\mathcal{B}(H \rightarrow \gamma\gamma) = 0.69 \pm 0.04$  is potentially sensitive to BSM physics, such as supersymmetry (SUSY) and extended Higgs sectors [15, 16, 17, 18]. The effects from these models can shift the  $H \rightarrow Z\gamma$  and  $H \rightarrow \gamma\gamma$  branching fractions by different amounts, making the ratio the most sensitive observable. The impact on the ratio varies by model, but can be up to 20% for two Higgs doublet or minimal SUSY models.

The ATLAS (A Toroidal LHC ApparatuS) and CMS Collaborations have performed searches for the decay  $H \rightarrow Z\gamma \rightarrow \ell^+\ell^-\gamma$  [19, 20, 21, 22] at  $\sqrt{s} = 7, 8$ , and 13 TeV in the  $e^+e^-\gamma$  and  $\mu^+\mu^-\gamma$  final states. The most stringent bound has been set by the

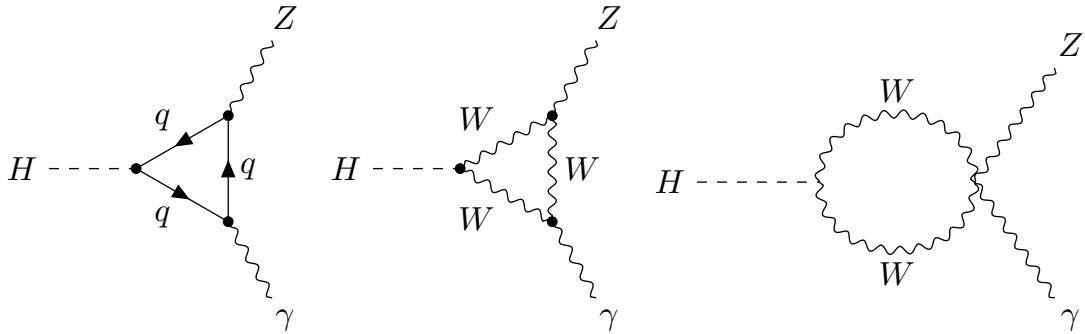


Figure 1.1. Feynman diagrams for  $H \rightarrow Z\gamma$  decay.

ATLAS Collaboration using  $\sqrt{s} = 13$  TeV data corresponding to an integrated luminosity of  $139 \text{ fb}^{-1}$ . The observed (expected) upper limit at 95% confidence level (CL) on  $\sigma(\text{pp} \rightarrow \text{H})\mathcal{B}(\text{H} \rightarrow Z\gamma)$  relative to the SM is 3.6 (2.6), assuming  $m_{\text{H}} = 125.09$  GeV. The ATLAS experiment has reported evidence at the 3.2 standard deviation level for the decay  $\text{H} \rightarrow \ell^+\ell^-\gamma$  with  $m_{\ell^+\ell^-} < 30$  GeV using both of the dilepton channels [23]. The CMS Collaboration has also searched for the  $\text{H} \rightarrow \ell^+\ell^-\gamma$  process with  $m_{\ell^+\ell^-} < 50$  GeV in the dimuon channel at  $\sqrt{s} = 8$  [24] and 13 [21] TeV.

This thesis describes a search for the decay  $\text{H} \rightarrow Z\gamma$ , where  $Z \rightarrow \ell^+\ell^-$  and  $m_{\ell^+\ell^-} > 50$  GeV. This phase space is chosen in order to target on-shell  $Z$  bosons, since the region  $m_{\ell^+\ell^-} < 50$  GeV contains a contribution from an additional process,  $\text{H} \rightarrow \gamma^*\gamma \rightarrow \ell^+\ell^-\gamma$  [7]. The data sample corresponds to an integrated luminosity of  $138 \text{ fb}^{-1}$  of pp collisions at  $\sqrt{s} = 13$  TeV accumulated between 2016 and 2018 by the CMS detector at the LHC. This thesis is organized as follows. The relevant theoretical physics background is provided in Chapter 2, and the CMS detector and event reconstruction are described in Chapter 3. Chapter 4 provides an overview of the  $\text{H} \rightarrow Z\gamma$  search strategy, and the data and simulated event samples are described in Chapter 5. Chapter 6 outlines the object and event selection, and Chapter 7 discusses event categorization. The signal and background modeling procedure is presented in Chapter 8, with systematic uncertainties discussed in Chapter 9. Chapter 10 describes the statistical procedure. The final results of the search are discussed in Chapter 11, followed by a conclusion in Chapter 12.

## CHAPTER 2

# Theory

### 2.1. The Standard Model

The SM is currently our best theoretical framework for understanding the nature of fundamental particles. It is rooted in the idea that particles exist as excitations of quantum fields. These fields are constructed so as to obey fundamental symmetries of nature, and quantum field theory describes both the particles of our universe and their interactions. The SM is not only elegant and extensive, but provides a wide variety of measurable observables for the experimentalist to probe. So far, many measurements have been made of the known elementary particles and their interactions, and the predictions of the SM have held up in each case. In this respect, it is a wildly successful theory. In other respects, it is obviously incomplete. It does not account for the gravitational force, dark matter, or dark energy, among other phenomena. Thus, there is value in testing the SM even more carefully with experiments like those at the LHC, in hopes of refining our understanding and potentially discovering new physics.

### 2.2. The Elementary Particles

Figure 2.1 catalogs the elementary particles of the SM. These particles come in two broad types: fermions and bosons. Fermions are particles of half-integer spin that obey the Pauli exclusion principle. They are, therefore, responsible for the structure of matter. Each fermion has a corresponding antiparticle that has opposite electrical charge, but is

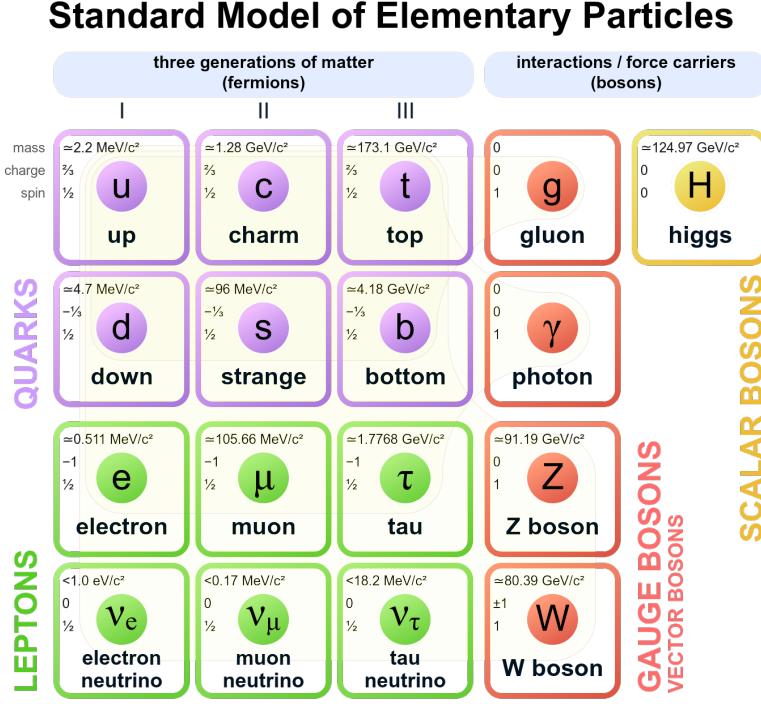


Figure 2.1. Table [25] of elementary particles in the SM. The leftmost three columns correspond to three generations of fermions, the fourth column from the left shows the gauge bosons which mediate the fundamental forces, and the rightmost column shows the Higgs boson. For each particle, mass, charge, and spin are labeled, with mass values given as of 2019.

otherwise identical. The fermions are subdivided into two types in three generations, the generations corresponding to mass. Quarks are fermions that carry both fractional electric charge and color charge, so they participate in the electroweak and strong interactions. Leptons are fermions that carry integer electric charge and participate in the electroweak interaction. In contrast to fermions, bosons have integer spin. The vector gauge bosons function as mediators of the fundamental forces of nature. The gluons mediate the strong force, the photon the electromagnetic force, and the W and Z bosons the weak force. Finally, the electrically neutral, scalar Higgs boson emerges due to electroweak symmetry

breaking, described later in this section. The interaction of the Higgs with the other particles in the SM is responsible for particle masses. In the context of this thesis, which is a search for  $H \rightarrow Z\gamma$ , the most relevant elementary particles are the Higgs boson,  $Z$  boson, leptons, and photon. Therefore, the remainder of this section will describe electroweak theory and the Higgs mechanism in more detail. Other aspects of the SM, such as quantum chromodynamics, are of importance in collider physics, but of less relevance for this specific search. Such topics are omitted in order to narrow the scope of the discussion.

### 2.3. Electroweak Theory

To motivate the Higgs mechanism and Higgs interactions, it is useful to first consider the electroweak interaction absent the Higgs. The electroweak piece of the SM Lagrangian can be written as

$$(2.1) \quad \mathcal{L}_{EW} = -\frac{1}{4}A_a^{\mu\nu}A_{\mu\nu}^a - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} + i \sum_L \bar{L}\gamma^\mu D_{\mu,L}L + i \sum_R \bar{R}\gamma^\mu D_{\mu,R}R$$

$$(2.2) \quad A_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu + g\varepsilon_{abc}W_\mu^b W_\nu^c$$

$$(2.3) \quad B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$$

$$(2.4) \quad D_{\mu,L} = \partial_\mu + \frac{ig}{2}\tau \cdot A_\mu + \frac{ig'}{2}B_\mu Y$$

$$(2.5) \quad D_{\mu,R} = \partial_\mu + \frac{ig'}{2}B_\mu Y.$$

In the above equations,  $A_{\mu\nu}^a$  represents an  $SU(2)_L$  triplet of gauge fields and  $B_{\mu\nu}$  represents a  $U(1)$  gauge field. The third and fourth terms in equation 2.1 describe the interactions of these gauge fields with the fermions, both quarks and leptons. Note that the interaction

depends on the chirality of the fermions, where the left-handed fermions are denoted as  $L$  and the right-handed as  $R$ . The right-handed fermions interact only with the  $U(1)$  field, while the left-handed fermions interact with the  $U(1)$  and  $SU(2)_L$  fields. This is encoded by the derivative operators  $D_{\mu,L}$  and  $D_{\mu,R}$  defined in equations 2.4 and 2.5 in terms of the fields, Pauli matrices ( $\tau$ ), hypercharge operator  $Y$ , and the couplings  $g$  and  $g'$ .

## 2.4. Spontaneous Symmetry Breaking (Higgs Mechanism)

One important property of the electroweak Lagrangian in equation 2.1 is that the bosons associated with the gauge fields are all massless. However, the direct observation of charged and neutral current interactions [26, 27] at CERN in 1973 implied that the  $W$  and  $Z$  boson must have relatively large masses in the range of 50–100 GeV. Indeed, the massive  $W$  and  $Z$  bosons were later discovered and measured [28, 29] at the CERN Super Proton Synchrotron in 1983. A solution to this shortcoming of the electroweak theory came in the form of the Higgs mechanism [30, 31, 32], which spontaneously breaks the  $SU(2) \times U(1)$  gauge symmetry. One consequence of this is the addition of a real Higgs field accompanied by a massive Higgs boson. The  $W$  and  $Z$  boson masses arise naturally due to the electroweak symmetry breaking, while fermion masses are explained via additional Yukawa couplings to the Higgs boson. The discovery of the Higgs boson and subsequent measurements of its properties have provided experimental verification that the Higgs mechanism is a central piece of the Standard Model.

To understand how the Higgs mechanism works, consider the introduction of a complex scalar field  $\Phi$ , which transforms as a doublet under  $SU(2)_L$ .

$$(2.6) \quad \Phi = \begin{bmatrix} \phi^+ \\ \phi^0 \end{bmatrix}$$

Its contribution to the Lagrangian is given by:

$$(2.7) \quad \mathcal{L}_H = (D_L^\mu \Phi)^\dagger (D_{\mu,L} \Phi) - V(\Phi)$$

where the Higgs potential takes the form

$$(2.8) \quad V(\Phi) = \mu^2 |\Phi^\dagger \Phi| + \lambda \left( |\Phi^\dagger \Phi| \right)^2.$$

Consider the case in which the parameters of the Higgs potential  $\lambda$  and  $\mu$  satisfy the conditions  $\lambda > 0$  and  $\mu^2 < 0$ . Then the shape of the potential is shown in Fig. 2.2 (right). There is no minimum of the potential at  $\Phi = 0$ . Rather, an infinite set of minima lie around a circle in the complex plane. Hence, it is said that  $\Phi$  has a nonzero vacuum expectation value (VEV). The value of the VEV in terms of  $\mu$  and  $\lambda$  can be determined by explicitly minimizing the potential:

$$(2.9) \quad \begin{aligned} \frac{\partial}{\partial(\Phi^\dagger \Phi)} V(\Phi) &= 0 \\ \mu^2 + 2\lambda \left( |\Phi^\dagger \Phi| \right) &= 0 \\ \mu^2 + 2\lambda \left[ (\phi^+)^2 + (\phi^0)^2 \right] &= 0 \end{aligned}$$

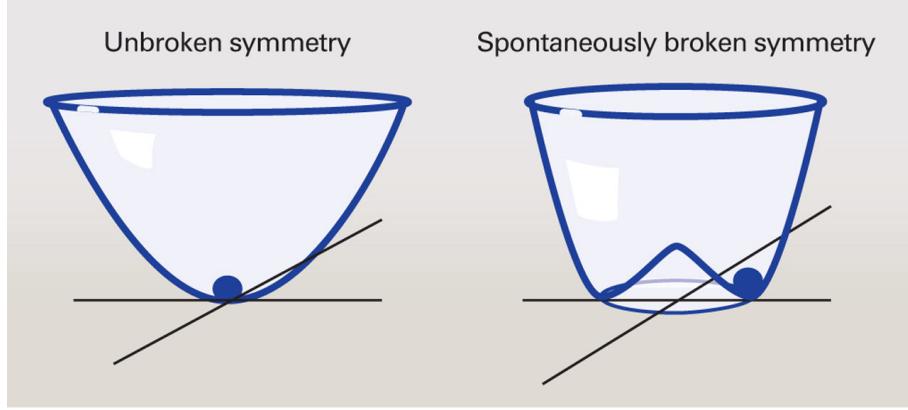


Figure 2.2. Comparison of the shapes of complex scalar potentials without symmetry breaking (left) and with spontaneous symmetry breaking (right) [33]. The transverse axes represent the real-imaginary  $\Phi$  plane, and the vertical axis represents the magnitude of the potential  $V(\Phi)$ . The diagram at right corresponds to the shape of the Higgs potential in the SM.

This can be minimized in many ways depending on individual values of  $\phi^+$  and  $\phi^0$  in the vacuum. By convention, and without loss of generality, we choose the case in which  $\phi^+ = 0$ . In this case we obtain the equation

$$(2.10) \quad \phi^0 = \sqrt{\frac{-\mu^2}{2\lambda}} = \frac{1}{\sqrt{2}}v$$

where we have defined  $v \equiv \sqrt{-\mu^2/\lambda}$ .

The existence of the Higgs VEV has profound implications. To see this, it is helpful to reparameterize the scalar doublet field  $\Phi$  as follows:

$$(2.11) \quad \Phi = \frac{1}{\sqrt{2}} e^{i \frac{\tau^a}{2} \theta_a(x)} \begin{bmatrix} 0 \\ v + h(x) \end{bmatrix}$$

As  $\Phi$  is invariant under local  $SU(2)_L$  gauge transformations, the prefactor may be rotated away. This is equivalent to setting  $\theta(x) = 0$  in equation 2.11. This choice of gauge is

known as the unitary gauge, and leads to

$$(2.12) \quad \Phi = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ v + h(x) \end{bmatrix}$$

Given the above form of  $\Phi$ , we can now evaluate the Higgs Lagrangian (equation 2.7), starting with the kinetic term  $(D_L^\mu \Phi)^\dagger (D_{\mu,L} \Phi)$ .

$$\begin{aligned} (D_L^\mu \Phi)^\dagger (D_{\mu,L} \Phi) &= \left| \left( \partial_\mu - \frac{ig}{2} \tau \cdot A_\mu - \frac{ig'}{2} B_\mu Y \right) \Phi \right|^2 \\ &= \frac{1}{2} \left| \begin{bmatrix} \partial_\mu - \frac{i}{2}(gA_\mu^3 + g'B_\mu) & -\frac{ig}{2}(A_\mu^1 - iA_\mu^2) \\ -\frac{ig}{2}(A_\mu^1 + iA_\mu^2) & \partial_\mu + \frac{i}{2}(gA_\mu^3 - g'B_\mu) \end{bmatrix} \begin{bmatrix} 0 \\ v + h(x) \end{bmatrix} \right|^2 \\ &= \frac{1}{2} \left| \begin{bmatrix} -\frac{ig}{2}(A_\mu^1 - iA_\mu^2)(v + h(x)) \\ \partial_\mu h(x) + \frac{i}{2}(gA_\mu^3 - g'B_\mu)(v + h(x)) \end{bmatrix} \right|^2 \\ &= \frac{1}{2} \partial_\mu h(x) \partial^\mu h(x) + \frac{1}{8} (gA_\mu^3 - g'B_\mu)(gA_\mu^\mu - g'B^\mu)(v + h(x))^2 \\ &\quad + \frac{g^2}{8} (A_\mu^1 - iA_\mu^2)(A_\mu^1 + iA_\mu^2)(v + h(x))^2 \end{aligned} \quad (2.13)$$

With some foreknowledge of the result, we define the physical gauge fields and their masses.

$$(2.14) \quad W_\mu^\pm = \frac{1}{\sqrt{2}} (A_\mu^1 \mp iA_\mu^2) \quad m_W = \frac{gv}{2}$$

$$(2.15) \quad Z_\mu = \frac{1}{\sqrt{g^2 + g'^2}} (gA_\mu^3 - g'B_\mu) \quad m_Z = \sqrt{g^2 + g'^2} \frac{v}{2}$$

$$(2.16) \quad A_\mu = \frac{1}{\sqrt{g^2 + g'^2}} (g'A_\mu^3 + gB_\mu) \quad m_A = 0$$

Then the kinetic term of the Higgs Lagrangian can be recast as

$$\begin{aligned}
 (D_L^\mu \Phi)^\dagger (D_{\mu,L} \Phi) = & \frac{1}{2} \partial_\mu h(x) \partial^\mu h(x) \\
 & + \frac{1}{2} m_Z^2 Z_\mu Z^\mu + m_W^2 W_\mu^+ W^{-\mu} \\
 & + \frac{v}{4} (g^2 + g'^2) Z_\mu Z^\mu h + \frac{1}{8} (g^2 + g'^2) Z_\mu Z^\mu h^2 \\
 (2.17) \quad & + \frac{v}{4} g^2 W_\mu^+ W^{-\mu} h + \frac{1}{8} g^2 W_\mu^+ W^{-\mu} h^2
 \end{aligned}$$

Equation 2.17 provides a great deal of information on the physical ramifications of the Higgs mechanism. The first term is the kinetic term of the physical Higgs boson field. The second and third terms are the mass terms of the Z and W bosons, respectively. The fourth and fifth terms show the linear and quadratic couplings of the Z boson to the Higgs boson, respectively. Finally, the sixth and seventh terms show the linear and quadratic couplings of the W boson to the Higgs boson, respectively. Given the form of equation 2.12, a similar expansion can be carried out on the Higgs potential (equation 2.8). Here, the terms involving the physical Higgs boson are most interesting, so constant terms are dropped.

$$\begin{aligned}
 V(\Phi) = & \frac{\mu^2}{2} (v + h)^2 + \frac{\lambda}{4} ((v + h)^2)^2 \\
 (2.18) \quad & \rightarrow \lambda v^2 h^2 + \lambda v h^3 + \frac{\lambda}{4} h^4
 \end{aligned}$$

The first term is a Higgs mass term, with  $m_H = \sqrt{2\lambda v^2}$ . The second and third terms describe the Higgs trilinear and quartic self-couplings, respectively.

The Lagrangian including the Higgs doublet  $\Phi$  can be further extended to incorporate interactions between the Higgs and fermion fields. These interactions, along with the nonzero Higgs VEV, provide a mechanism to generate the fermion masses. The interactions take the form of Yukawa couplings:

$$(2.19) \quad \mathcal{L}_{Yukawa} = Y_{ij}^d \bar{Q}_L^i \Phi d_R^j + Y_{ij}^u \bar{Q}_L^i \tilde{\Phi} u_R^j + Y_{ij}^e \bar{L}_L^i \Phi e_R^j + h.c.,$$

where  $\tilde{\Phi} \equiv i\tau_2 \Phi^*$ , and  $u$ ,  $d$ , and  $e$  represent up-type quarks, down-type quarks, and leptons, respectively. The constants  $Y_{ij}^a$  denote the Yukawa couplings in each case,  $Q^i$  is the set of SU(2) quark doublets, and  $L^i$  is the set of SU(2) lepton doublets.

$$(2.20) \quad \tilde{\Phi} \equiv i\tau_2 \Phi^*$$

Plugging in the unitary gauge parameterization of equation 2.12, this evaluates to

$$(2.21) \quad \mathcal{L}_{Yukawa} = \frac{Y_{ij}^d}{\sqrt{2}} \bar{d}_L^i (v + h) d_R^j + \frac{Y_{ij}^u}{\sqrt{2}} \bar{u}_L^i (v + h) u_R^j + \frac{Y_{ij}^e}{\sqrt{2}} \bar{e}_L^i (v + h) e_R^j$$

It is worth looking closely at the terms in equation 2.21. For a given fermion type, the first term in the parentheses is a fermion mass term. The second term in the parentheses gives the coupling of the fermion to the Higgs boson. With this in mind, we observe that the fermion mass is given in terms of the couplings and the Higgs vev by

$$(2.22) \quad m_f = \frac{y_f v}{\sqrt{2}}$$

where  $y_f$  is the relevant value taken from the Yukawa coupling matrix. In addition, we see that the strength of the fermion coupling to the Higgs boson is given by  $\frac{m_f}{v}$ . Thus, fermions

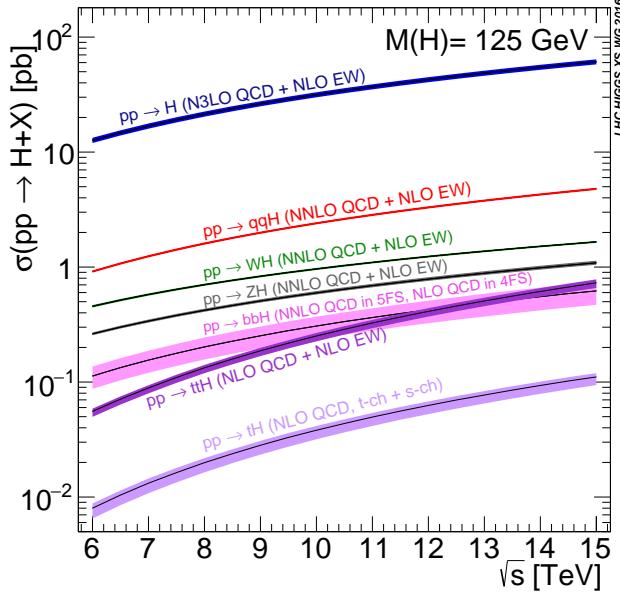


Figure 2.3. Cross sections for different Higgs boson production mechanisms as a function of center of mass energy for  $m_H = 125 \text{ GeV}$  [12].

couple to the Higgs boson with strength directly proportional to their masses. This result has important consequences in the context of collider experiments, as it regulates the rates of Higgs production and decay related to each fermion-Higgs interaction.

## 2.5. Higgs Production

At a hadron collider like the LHC, the Higgs boson can be produced via several different mechanisms, with each mechanism occurring at a certain rate. The dominant production mechanism is gluon-gluon fusion ( $ggH$ ), followed by vector boson fusion (VBF) production, which is about an order of magnitude rarer. Less common production mechanisms, but still relevant for a search at the LHC, are the associated production mechanisms  $WH$ ,  $ZH$ , and  $t\bar{t}H$ ). The cross sections for each Higgs production mechanism as a function of center of mass energy are shown in Fig. 2.3.

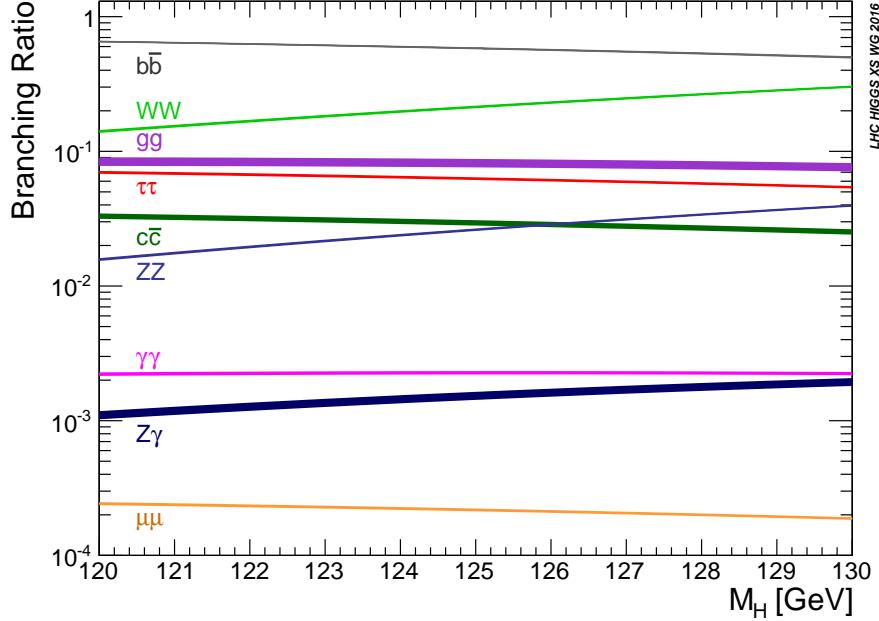


Figure 2.4. Branching fractions for different Higgs decay modes as a function of  $m_H$  [12].

## 2.6. Higgs Decay

The Higgs boson decays rapidly after it is produced at the LHC. It can decay into a variety of final states, each occurring at a certain rate. Tree-level Higgs decays occur at rates proportional to the square of the mass of the decay products. Decays to  $b\bar{b}$ ,  $W^+W^-$ ,  $\tau^+\tau^-$ , and  $ZZ^*$  are among the most common and experimentally accessible at the LHC. Other Higgs decays occur at loop-level, including  $H \rightarrow \gamma\gamma$  and  $H \rightarrow Z\gamma$ . Figure 2.4 shows the branching fractions at  $\sqrt{s} = 13$  TeV for different decay modes as a function of Higgs boson mass. The branching fraction of  $H \rightarrow Z\gamma$  is suppressed relative to several other production modes, so a search for  $H \rightarrow Z\gamma$  must rely on the clean  $\ell^+\ell^-\gamma$  signature.

## 2.7. Physics Beyond the Standard Model

One of the goals of the LHC is to search for BSM physics. While the decay  $H \rightarrow Z\gamma$  is predicted in the SM, searching for it and measuring its branching fraction can indirectly probe potential new physics phenomena. A few unique features of  $H \rightarrow Z\gamma$  make it well-suited as a BSM probe. First, the decay is loop-induced, which means that the introduction of new particles will contribute to the loop, interfering and altering the final branching fraction. Secondly, the decay  $H \rightarrow \gamma\gamma$  is a very similar loop-induced process. This means any new particle content entering into the  $H \rightarrow Z\gamma$  loop will also enter into the  $H \rightarrow \gamma\gamma$  loop. However, in the case of  $H \rightarrow Z\gamma$ , the  $Z$  boson couples to loop particles according to the  $SU(2) \times U(1)$  quantum number rather than just the electric charge, as is the case for  $H \rightarrow \gamma\gamma$ . It follows that potential BSM physics scenarios can shift the branching fractions of  $H \rightarrow Z\gamma$  and  $H \rightarrow \gamma\gamma$  by differing amounts, making the ratio  $\mathcal{B}(H \rightarrow Z\gamma)/\mathcal{B}(H \rightarrow \gamma\gamma)$  a sensitive observable. In the SM,  $\mathcal{B}(H \rightarrow Z\gamma)/\mathcal{B}(H \rightarrow \gamma\gamma) = 0.69 \pm 0.04$ . A measurement of a significant deviation from this value would be an indicator of possible BSM physics. Below, we describe a few examples of BSM scenarios that can affect this ratio.

One way to consider BSM effects in the context of  $H \rightarrow Z\gamma$  is to introduce new particles, like a  $W'$  boson, charged scalar, or a pair of charged leptons. This is the approach taken by Carena, Low, and Wagner [16]. The  $W'$  possibility is motivated by the fact that the  $W$  boson loop is the dominant contribution in the SM. The  $W'$  can be defined by an  $SU(2)$  triplet and characterized by mass and Higgs coupling parameters. Figure 2.5 (left) shows contours of constant branching fraction enhancement in the  $H \rightarrow Z\gamma$  and  $H \rightarrow \gamma\gamma$  decay channels in the mass-coupling plane for the  $W'$  model. Depending on

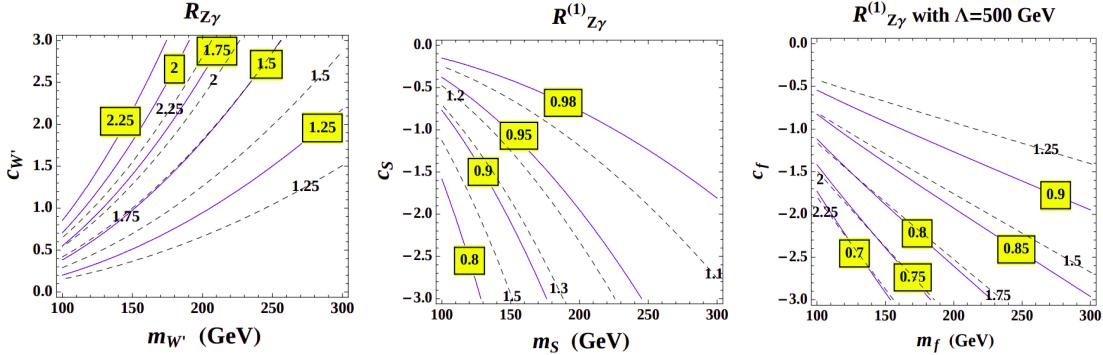


Figure 2.5. Contours of constant branching fraction enhancement relative to the SM for the  $W'$  (left), charged scalar (center), and charged lepton (right) scenarios [16]. The solid lines represent contours of enhancement in  $\mathcal{B}(H \rightarrow Z\gamma)$ , while the dashed lines correspond to  $H \rightarrow \gamma\gamma$ . The yellow boxed numbers label the enhancement for each  $H \rightarrow Z\gamma$  contour, and the unboxed numbers label the enhancement for each  $H \rightarrow \gamma\gamma$  contour.

the parameter values, the level of enhancement can be up to double the SM value, and it differs significantly in the two channels. Similarly, we can imagine a new charged scalar or a pair of new charged leptons. The branching fraction enhancement contour plots in Fig. 2.5 show that for similar regions in the mass-coupling plane,  $H \rightarrow Z\gamma$  can be enhanced, while  $H \rightarrow \gamma\gamma$  is simultaneously diminished.

Other relevant BSM physics models include extended Higgs sectors. Such models often contain charged Higgs bosons, which enter into the  $H \rightarrow Z\gamma$  and  $H \rightarrow \gamma\gamma$  loops and impact the branching fractions. In Ref. [17], Chiang and Yagyu catalog extended Higgs models of varying types: models with one singly-charged boson, those with one singly-charged and one doubly-charged boson, and those with two singly-charged bosons. In total, the authors consider thirteen models, including the two Higgs doublet model, the minimal SUSY model, and the Higgs triplet model. Figure 2.6 gives the ratio of decay widths  $\Gamma(H \rightarrow Z\gamma)/\Gamma(H \rightarrow \gamma\gamma)$  as a function of the mass coefficient of the additional

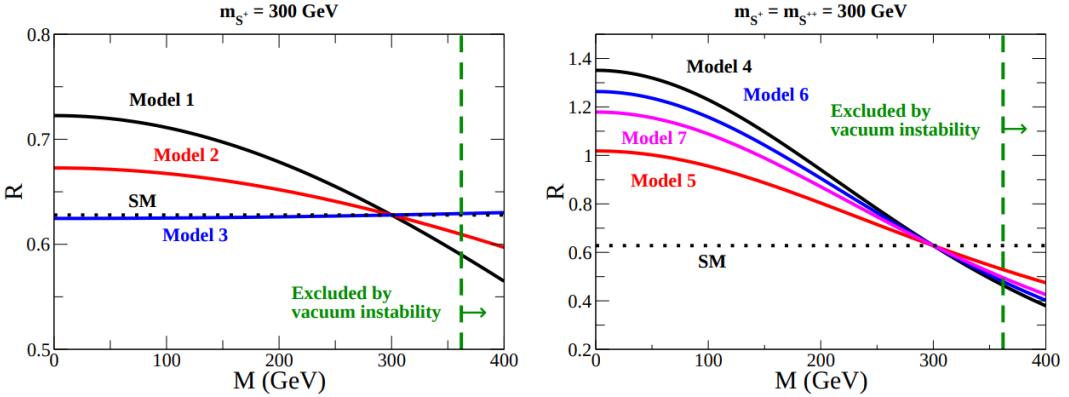


Figure 2.6. Ratio of decay widths  $\Gamma(H \rightarrow Z\gamma)/\Gamma(H \rightarrow \gamma\gamma)$  as a function of the mass coefficient of the additional scalar field(s) for seven BSM models with extended Higgs sectors [17]. The models plotted on the left correspond to one additional singly-charged boson, and those on the right to one additional singly-charged and one additional doubly-charged boson.

scalar field(s) for seven of the models. For mass scales accessible at the LHC, the ratio can be significantly shifted relative to the SM prediction.

Another approach, discussed by Degrassi and Vitti in Ref. [10], is to consider how an anomalous Higgs trilinear self-coupling would modify the  $H \rightarrow Z\gamma$  branching fraction. In the SM, the value of  $\lambda$  in equation 2.18 is the same for each of the three terms. However, one can consider a case in which the value of the trilinear term is different than the value in the SM. This would enter into amplitudes for two-loop processes, such as those shown in Fig. 2.7. In turn, this would modify the  $H \rightarrow Z\gamma$  branching fraction relative to the SM prediction. An advantage of this approach is that it does not depend on the specifics of the BSM model giving rise to the anomalous trilinear Higgs self-coupling, so it can serve as a more general probe for new physics. Degrassi and Vitti show that  $H \rightarrow Z\gamma$  is among the most sensitive Higgs decay modes to an anomalous trilinear Higgs self-coupling, more sensitive than  $H \rightarrow \gamma\gamma$  and only slightly less sensitive than  $H \rightarrow ZZ^*$ .

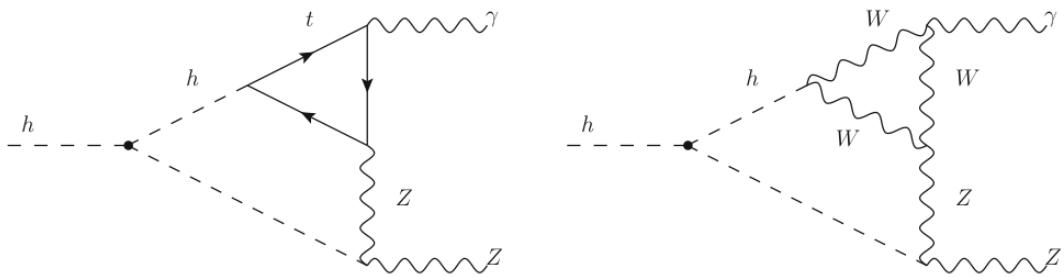


Figure 2.7. Examples of two-loop diagrams [10] contributing to  $H \rightarrow Z\gamma$  decay. The presence of the three-Higgs vertex makes these diagrams sensitive to an anomalous trilinear Higgs self-coupling.

## CHAPTER 3

# Experiment Description

### 3.1. The Large Hadron Collider

The Large Hadron Collider (LHC) [34] is a high energy pp collider that serves several modern particle physics detector experiments. Built at CERN, straddling the French-Swiss border, and turned on for operation in 2008, it is the most powerful particle collider in history. The LHC was designed to collide protons at a maximum center of mass energy of  $\sqrt{s} = 14 \text{ TeV}$  with a maximum instantaneous luminosity of  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ . To achieve this performance, the 27 km tunnel originally constructed for the Large Electron-Positron Collider (LEP) [35] was repurposed to separately accelerate two counterrotating proton beams. The separate acceleration of the beams is made possible by oppositely oriented magnetic dipole fields in the two rings. When the beams achieve the desired energy, they are collided at one of four interaction points.

The protons collided in the LHC are first gathered, bunched, and accelerated in other parts of the CERN accelerator complex before injection into the LHC ring. A detailed description of the injection chain can be found in Ref. [36], and a short summary is provided below. First, protons are stripped off of hydrogen atoms by a duoplasmatron and accelerated to 50 MeV by a linear accelerator (LINAC). Next, the protons are further accelerated in a series of synchrotron rings of increasing size, the Proton Synchrotron Booster (PSB), Proton Synchrotron (PS), and Super Proton Synchrotron (SPS). The SPS brings

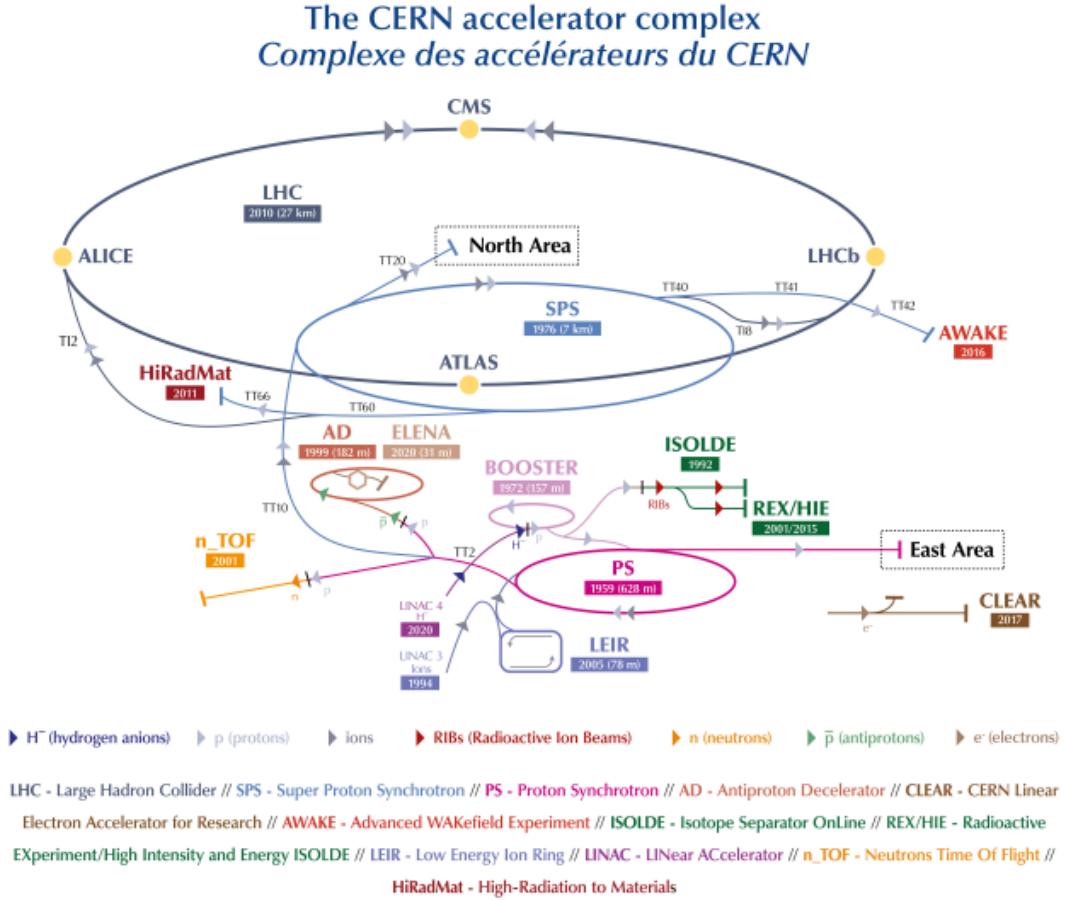


Figure 3.1. CERN accelerator complex [37].

the proton beam energy up to 450 GeV. At this point, the beam is injected to the two LHC rings to be accelerated up to the final collision energy. Acceleration in the main ring is achieved by a system of thousands of superconducting dipole magnets, along with hundreds of correcting quadrupole magnets. A liquid helium cooling system is used to maintain the magnets at cryogenic temperatures. A full diagram of the CERN accelerator complex [37], including the parts relevant for the LHC, is shown in Fig. 3.1.

### 3.2. The Compact Muon Solenoid

The CMS apparatus [38] is a multipurpose, nearly hermetic detector, designed to trigger on [39, 40] and identify photons, electrons, muons, and (charged and neutral) hadrons [41, 42, 43, 44]. The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. The ECAL consists of 75,848 lead tungstate crystals, which provide coverage in pseudorapidity  $|\eta| < 1.48$  in a barrel region (EB) and  $1.48 < |\eta| < 3.0$  in two endcap regions (EE). Preshower detectors consisting of two planes of silicon sensors interleaved with a total of 3 radiation lengths of lead are located in front of each EE detector. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. Figure 3.2 shows a schematic diagram of the CMS detector, and a more detailed description of the individual detector components and detector operation is provided below.

#### 3.2.1. Superconducting Magnet

The function of the superconducting magnet within the CMS detector is to bend the trajectories of charged particles. This is crucial for accurate particle identification and momentum measurement. Designed to provide a maximum magnetic field of 4 T, its operating strength during pp collision runs is set to 3.8 T. The bulk of the magnet is composed of NbTi, cooled by liquid helium to 4.5 K, which is below the critical temperature

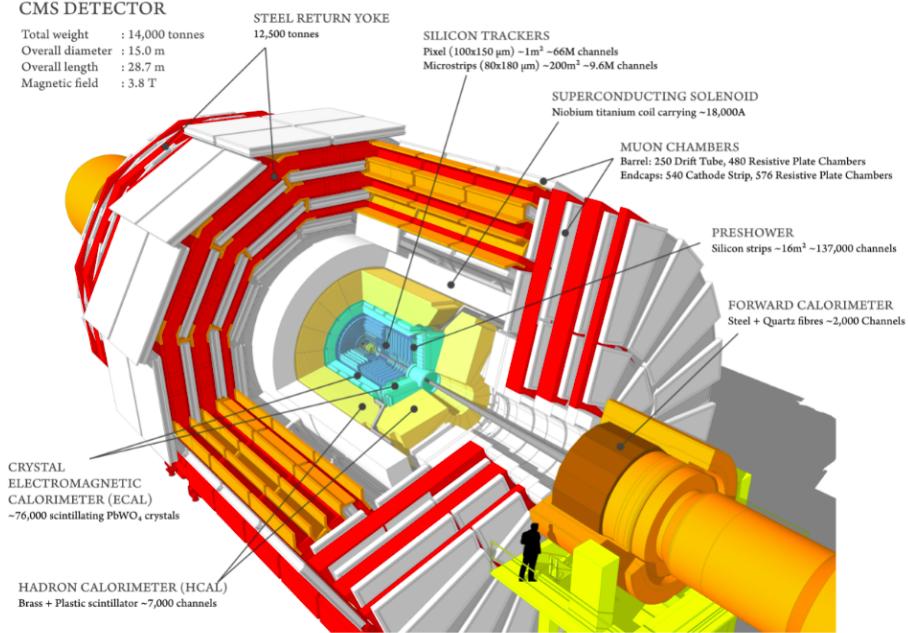


Figure 3.2. CMS detector apparatus. [45]

for superconductivity. The magnet has a length of 12.5m, diameter of 6.3m, and mass of 220 tons. The solenoid encloses several detector components, including the tracker and the majority of the calorimeters. A steel flux-return yoke is built around the solenoid, and is composed of 5 wheels and two endcaps weighing a total of roughly 10,000 tons.

### 3.2.2. Inner Tracking System

The inner tracking system is designed to reconstruct charged particle trajectories and vertices arising from particle decays. Comprised of a silicon pixel detector and silicon strip tracker, it covers the pseudorapidity region  $|\eta| < 2.5$  and is designed for efficient and precise measurement of charged particles with transverse momentum ( $p_T$ ) above about 1 GeV. At the LHC design luminosity for pp collisions, each bunch crossing leads to about

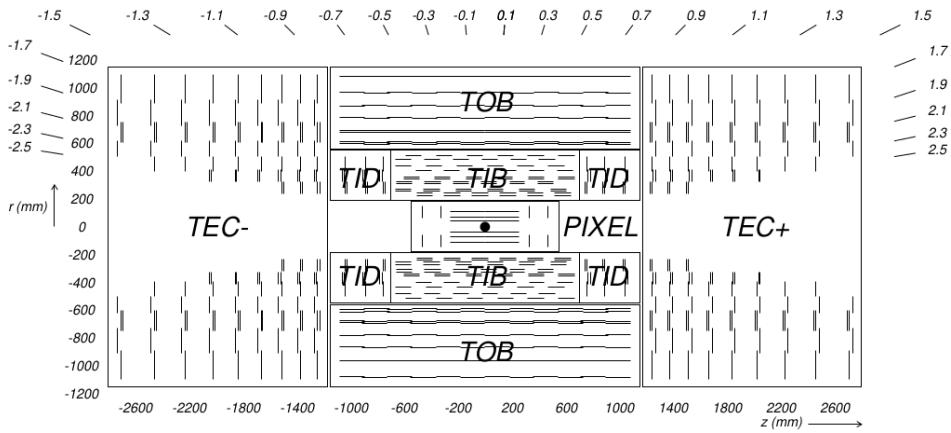


Figure 3.3. Cross section of the CMS inner tracking system in the r-z plane. [38]

1,000 hits in the inner tracking system. As such, the inner tracking system was designed to be maximally radiation tolerant.

The silicon pixel detector covers the inner region of  $r < 10\text{ cm}$ . It is composed of  $100 \times 150 \mu\text{m}^2$  pixels, arranged in three barrel layers and two endcap disks. Due to the extreme radiation environment, the innermost layer was designed to be replaced after about two years of LHC operation. In response to LHC running conditions and detector degradation, a replacement and upgrade of the full pixel detector was made during the LHC extended year-end technical stop in 2016–2017.

The silicon strip tracker covers the region  $20 < r < 116\text{ cm}$ . It is divided into three subsystems: the Tracker Inner Barrel (TIB), Tracker Inner Disks (TID), and Tracker Outer Barrel (TOB). The strip thickness is  $320 \mu\text{m}$  ( $500 \mu\text{m}$ ) in the TIB/TID (TOB). A schematic diagram of the CMS inner tracking system in the r-z plane is shown in Fig. 3.3.

### 3.2.3. Electromagnetic Calorimeter (ECAL)

The electromagnetic calorimeter (ECAL) is designed to identify and measure electromagnetic particles by inducing and characterizing electromagnetic showers. It is made up of 61,200 lead tungstate ( $\text{PbWO}_4$ ) crystals in the barrel (EB), which covers  $|\eta| < 1.479$ , plus 7,324 crystals in each endcap (EE), which cover the range  $1.479 < |\eta| < 3.0$ . A preshower system lies in front of the EE. The preshower is a lead and silicon sampling calorimeter designed to aid in the identification of neutral pions and to improve the spatial resolution in the endcap region. The crystal length corresponds to 25.8 (24.7) radiation lengths in the EB (EE). When electromagnetic particles, such as electrons and photons, encounter the ECAL and interact with the crystal material, electromagnetic showers are induced. Light is emitted by the ECAL crystals proportional to the energy of constituent particles in the shower, allowing for reconstruction of the shower energy. An array of fast and radiation-tolerant photodetectors detects the emitted light, and this information is saved off detector. For this purpose, the EB uses avalanche photodiodes, while the EE uses vacuum phototriodes. As shown in Fig. 3.4, the ECAL has maintained good performance with minimal degradation throughout the duration of Run 1 and Run 2 of the LHC.

### 3.2.4. Hadronic Calorimeter (HCAL)

Measurements of hadrons, jets, and missing transverse momentum are crucial at a hadron collider like the LHC. The CMS hadronic calorimeter (HCAL) is instrumental in such measurements. The HCAL is primarily a sampling calorimeter broken into several sub-detectors. The barrel (HB) covers the range  $|\eta| < 1.3$ , the endcaps (HE) cover the range

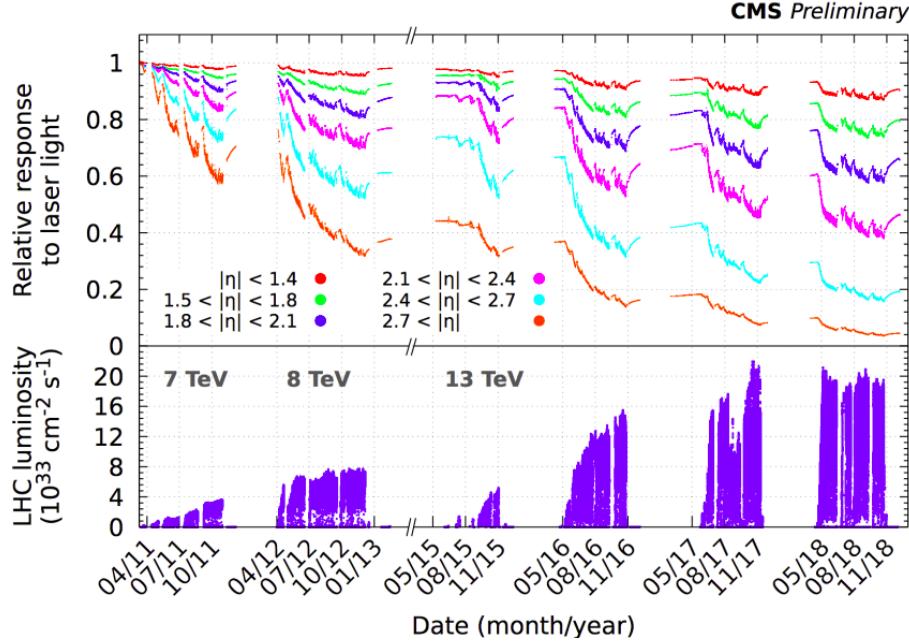


Figure 3.4. Response of the CMS ECAL crystals to injected laser light as a function of time, broken into regions of  $|\eta|$  [46].

$1.3 < |\eta| < 3$ , the forward calorimeter (HF) covers the range  $3 < |\eta| < 5$ , and an outer barrel calorimeter (HO) covers  $|\eta| < 1.3$  outside of the solenoid magnet.

Layers of brass absorber are interleaved with plastic scintillator in the HB and HE, with a total absorber thickness of 5.82–10.6 interaction lengths in the HB and about 10 interaction lengths in the HE. Scintillation light emitted by the active material is measured and transmitted off detector using hybrid photodiodes. The design of the HF was motivated by the extreme particle fluxes in the very forward region. The HF is made up of steel plate absorbers, in which quartz fibers are inserted to serve as the active material. Cherenkov light is produced when shower particles moving at speeds above the Cherenkov threshold pass through the fibers, and a calculable fraction of the light is captured by the fibers. The function of the HO is to recover the energy in showers that

leak through the HB, which is needed for the accurate measurement of missing transverse momentum. In the HO, the solenoid coil, itself, as well as a thick iron tail catcher, function as absorbers. This extends the effective HCAL depth to 11.8 interaction lengths everywhere except at the barrel-endcap boundary.

### 3.2.5. Muon Detectors

The CMS muon detector system is comprised of a set of gas ionization detectors embedded in the flux-return yoke outside the solenoid magnet. These detectors cover the range  $|\eta| < 2.4$ . The positioning of the muon detectors outside the magnet takes advantage of the fact that muons deposit minimal energy in the detector materials within the solenoid, such as the calorimeters. Most other types of particles will have deposited their energy before reaching the muon system, so using the combination of muon detector and tracker information, CMS is able to reconstruct muons with good momentum resolution. The gas ionization detectors in the muon system come in three types: drift tubes (DTs), cathode strip chambers (CSCs), and resistive plate chambers (RPCs). The DTs cover the barrel region ( $|\eta| < 1.2$ ) and are arranged in four stations, each with 60–70 drift chambers containing a gas mixture of 85% Ar and 15% CO<sub>2</sub>. The CSCs are multiwire proportional chambers composed of anode wire planes interleaved with cathode panels. Muons passing through the CSCs ionize a gas mixture of 40% Ar, 50% CO<sub>2</sub>, and 10% CF<sub>4</sub>, and the electrons flow to the anodes, yielding a detectable avalanche of charge. Covering the region  $|\eta| < 2.1$ , the RPCs are composed of oppositely charged parallel plates enclosing a gas mixture of primarily C<sub>2</sub>H<sub>2</sub>F<sub>4</sub>. One advantage of the RPCs is their good time resolution, which is less than the LHC bunch spacing time of 25ns. This makes

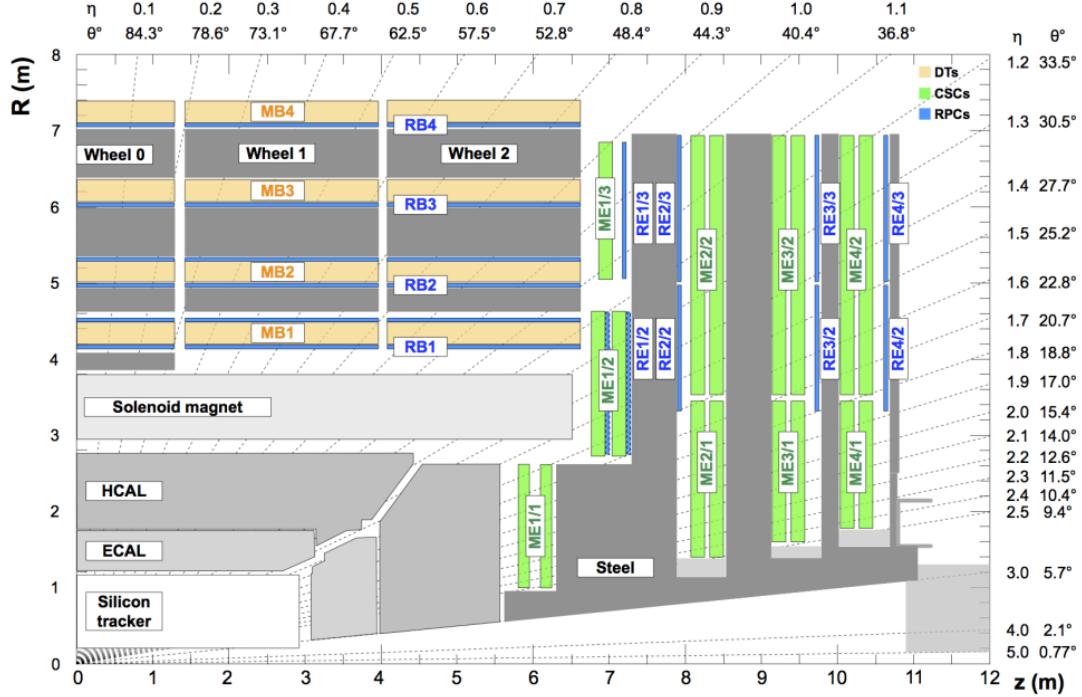


Figure 3.5. Diagram of the CMS detector in the  $r$ - $z$  plane showing the components of the muon system [38].

them useful for fast muon triggering that matches muon tracks to the relevant bunch crossing. Figure 3.5 shows a cross-sectional view of the CMS detector highlighting the muon system components.

### 3.3. Trigger System

At the LHC, the proton beam crossing interval is 25 ns, which corresponds to a pp collision rate of 40 MHz. Recording the full set of detector information for each collision is unfeasible, as there is far too much data to save to disk. As a result, the CMS experiment employs a two-tiered trigger system, designed to preserve information of physics interest while reducing the stored event rate to a more manageable 100 Hz. The first level of the trigger (L1) is a hardware trigger, which reduces the event rate from 40 MHz to about

100 kHz. The second layer is the high level trigger (HLT), a processor farm using software optimized for fast processing.

The L1 trigger is implemented using field-programmable gate arrays and application-specific integrated circuits, and has local, regional, and global components. First, local patterns in the calorimeters, track segments, and hit patterns in the muon chambers are processed by the Trigger Primitive Generators (TPGs). The TPGs rank and sort primitive objects corresponding to particle candidates based on energy, momentum, and reconstruction quality. Information from the TPGs is taken as input by regional triggers for the calorimeters and muon systems. These regional triggers further determine candidate physics objects, as well as energy sums and isolation information. Finally, this information is fed to a set of global triggers, which rank trigger objects across the entire detector. The final global trigger must decide whether to accept or reject an event based on information input from the global calorimeter and muon trigger systems. Algorithms used for this decision can be based on single object  $p_T$  thresholds and multiplicity-related requirements, such as the presence of multiple jets, among other criteria. Events passing the L1 trigger are read out and passed to the HLT for further processing.

In contrast to the L1 trigger, the HLT incorporates the full physics object reconstruction based on the full precision of the detector. This allows the HLT to accept and reject events with algorithms of similar quality to those used by offline analyses. Over 13,000 central processing unit cores are dedicated to HLT processing. For speed and efficiency, reconstruction and filtering algorithms are applied in increasing order of complexity. If a filter sequence fails, the rest of the reconstruction is skipped. Additionally, processing is done regionally based on the L1 candidates and relevant detector components passed as

input to the HLT. Events that pass the HLT are assigned to relevant data streams based on their physics content. For the  $H \rightarrow Z\gamma$  search, the relevant streams are the Double Muon and Double Electron streams, triggered by events with two muons or two electrons passing a set of minimum  $p_T$  requirements and other cuts.

### 3.4. Object Reconstruction

The global event reconstruction (also called particle-flow (PF) event reconstruction [47]) aims to reconstruct and identify each individual particle in an event, with an optimized combination of all subdetector information. In this process, the identification of the particle type (photon, electron, muon, charged hadron, neutral hadron) plays an important role in the determination of the particle direction and energy. Figure 3.6 provides a simple visualization of how each type of particle types interacts with each of the CMS subdetectors. An overview of the reconstruction procedure for each particle type is provided below.

#### 3.4.1. Photon and Electron Reconstruction

Photons and electrons interact with the material in the ECAL, depositing the majority of their energy before reaching the HCAL. As they interact, photons convert into electron-positron pairs and electrons radiate bremsstrahlung photons. This leads to the formation of an electromagnetic shower. Because of this, the original particle energy is split into multiple energy deposits, which must be combined to reconstruct the original energy. Deposits in individual ECAL crystals are combined into clusters, and these clusters are in turn combined into superclusters. Two algorithms are used to define superclusters: the

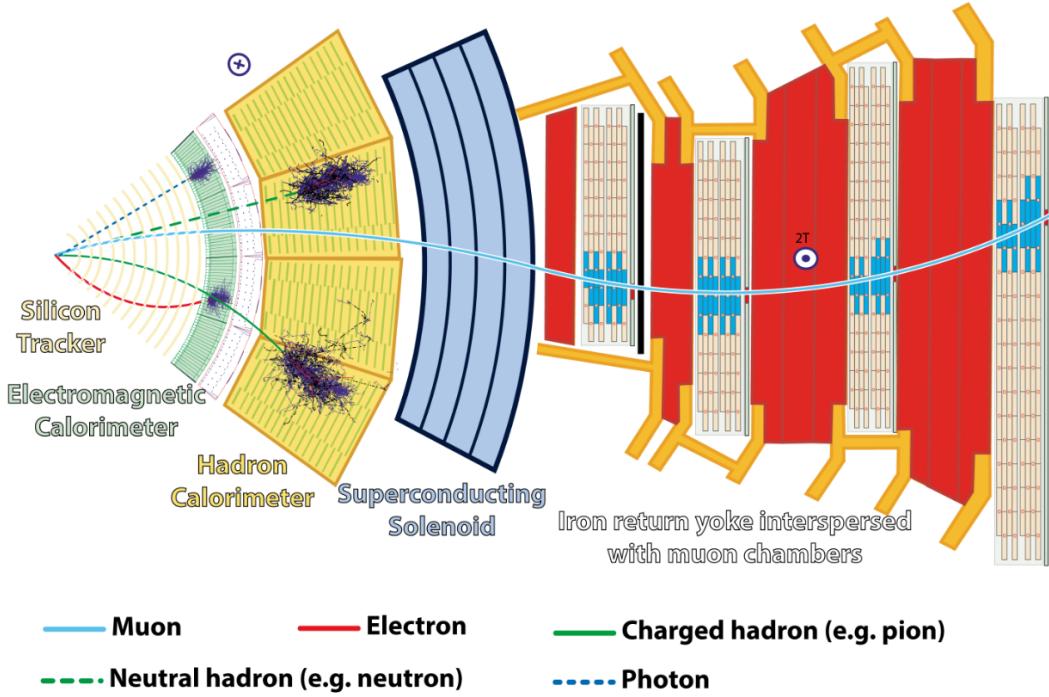


Figure 3.6. Visualization of different types of particles interacting with the CMS detector.

“mustache” algorithm, and the “refined” algorithm. The mustache algorithm defines a seed cluster with energy above a certain threshold, and then combines it with other clusters within a region of the  $\eta\text{-}\phi$  plane centered at the seed location. The refined algorithm uses the mustache superclusters and tracking information to extrapolate bremsstrahlung and conversion tracks, in order to decide whether a cluster should belong to the supercluster.

The distinction between photons and electrons is made using tracking information, where photons are associated with no tracks and electrons with tracks. The Gaussian Sum Filter (GSF) [48] track fitting algorithm is used to identify and characterize tracks that might be associated with an electron. It first begins with a hit pattern in the tracker, which is used as a seed. This seed can either be tracker-driven, coming from the collection

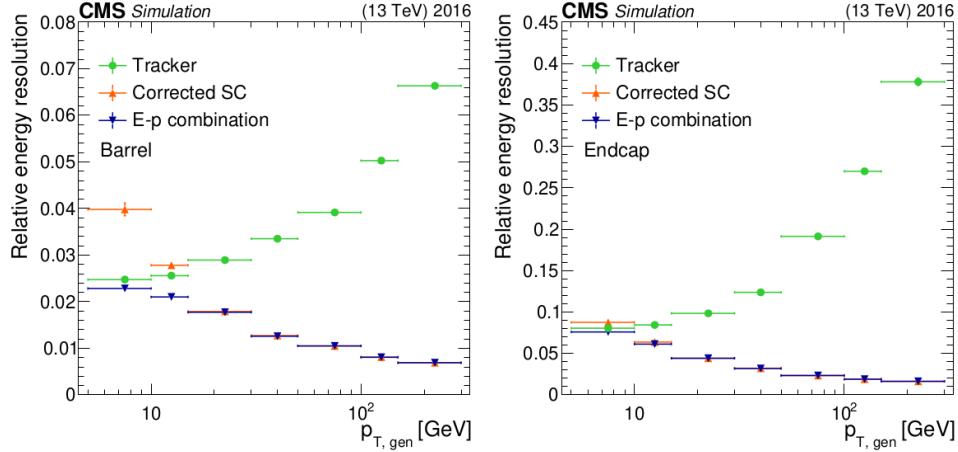


Figure 3.7. Energy resolution for simulated electrons in the barrel (left) and endcaps (right) as a function of generated  $p_T$  [49]. The green markers represent tracker measurement only, the orange markers represent ECAL measurement only, and the purple markers represent the combination of tracker and ECAL information.

of generic tracks tested for mutual compatibility, or it can be ECAL-driven, where a mustache supercluster is compared in location with a collection of tracker hit patterns to determine if the supercluster is consistent with the trajectory indicated by the track. Electron seeds are then converted into reconstructed electron tracks. In the absence of any GSF electron tracks, a photon candidate is obtained. Additional distinction between photons and electrons is obtained through further selection requirements. Fig. 3.7 shows the energy resolution for simulated electrons in the barrel and endcaps as a function of generated  $p_T$ . The measured energy resolution for electrons produced in Z boson decays in pp collision data ranges from 2–5%, depending on electron pseudorapidity and energy loss through bremsstrahlung in the detector material [49].

### 3.4.2. Muon Reconstruction

Muon reconstruction utilizes information from the muon detectors and the tracker. First, detector hits in the CSCs, DTs, and RPCs are used to build standalone tracks using a Kalman filter [50] technique. Subsequently, these standalone muon tracks are combined with tracker information via two algorithms. So-called “tracker muons” are reconstructed using an “inside-out” algorithm, which starts from tracker tracks and matches them to DT or CSC segments. So-called “global muons” are reconstructed with an “outside-in” approach, which starts from standalone muon tracks and matches them to tracker tracks using a Kalman filter technique. If both algorithms reconstruct a muon sharing the same tracker track, the two outputs are merged into a single muon candidate. In general, the tracker muon algorithm is more efficient in the region of low muon  $p_T$ , while the global algorithm is more efficient at high  $p_T$ .

The momentum of muons is obtained from the corresponding track momentum. Matching muons to tracks measured in the silicon tracker results in a  $p_T$  resolution, for muons with  $p_T$  up to 100 GeV, of 1% in the barrel and 3% in the endcaps. The  $p_T$  resolution in the barrel is better than 7% for muons with  $p_T$  up to 1 TeV [43]. As shown in Fig. 3.8, this corresponds to a dimuon mass resolution in  $Z \rightarrow \mu^+ \mu^-$  events of roughly 1–2 GeV, depending on the pseudorapidity of the muons [51].

### 3.4.3. Hadron Reconstruction

All particles within an event that are not reconstructed as photons, electrons, or muons are reconstructed as hadrons. Note that within the tracker acceptance ( $|\eta| < 2.5$ ), all ECAL clusters not linked to a track are considered photons. All of the remaining ECAL and

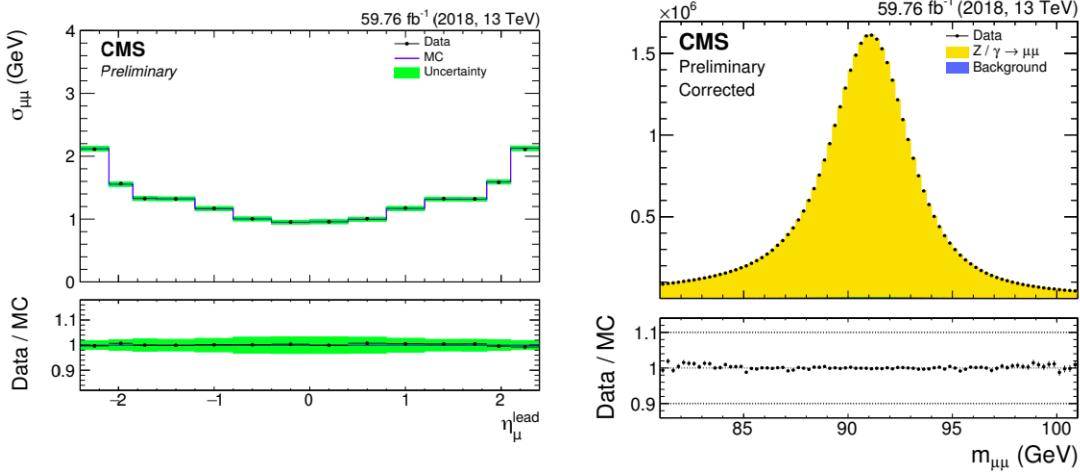


Figure 3.8. Left: dimuon mass resolution as a function of leading muon  $\eta$  for  $Z \rightarrow \mu^+ \mu^-$  events in 2018 data and simulation. Right: reconstructed dimuon mass. [51]

HCAL clusters are processed further by the PF algorithm. The HCAL clusters are first linked to any associated tracks and ECAL clusters. Then, a comparison is made between the calorimetric energy sum and the sum of the track momenta. If the calorimetric energy sum is compatible with the sum of track momenta, a charged hadron is reconstructed. If the calorimetric energy sum is greater than the sum of track momenta, the excess prompts the reconstruction of photons and neutral hadrons, depending on the size of the excess.

#### 3.4.4. Jet Reconstruction

For each event, hadronic jets are reconstructed as clusters of PF particles, including photons, electrons, muons, and hadrons. For clustering, the infrared and collinear-safe anti- $k_T$  algorithm [52, 53] is used. The goal of jet clustering algorithms is to combine PF particles together into a single jet object by merging them together based on their relative geometric distances and transverse momenta ( $k_T$ ). In the case of the anti- $k_T$  algorithm, clustering

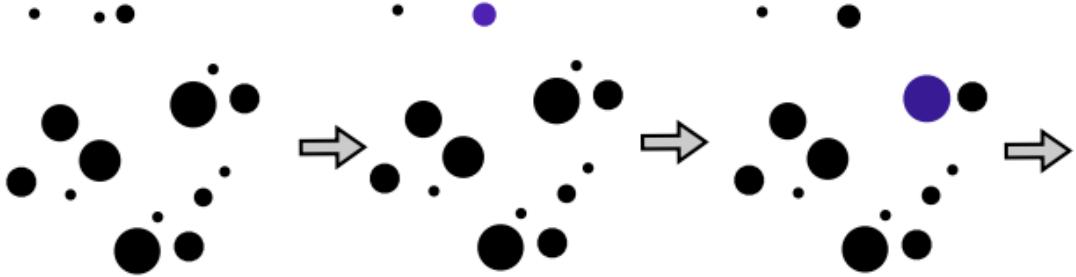


Figure 3.9. An example of the first few steps of the anti- $k_{\mathrm{T}}$  jet clustering algorithm.

proceeds from smallest geometric distance to largest, weighted by squared inverse  $k_{\mathrm{T}}$ . This inverse weighting preferentially combines PF particles with higher momenta before lower momenta, for a given distance. The advantage of this approach is that low momentum particles are more likely to be clustered with high momentum counterparts, rather than with other low momentum particles. It follows that these low momentum particles will not significantly modify the jet shape, whereas the more important, high-momentum particles, will. A precise definition of the anti- $k_{\mathrm{T}}$  algorithm is given in equation 3.1.

$$(3.1) \quad d_{ij} = \min(k_{T,i}^{2p}, k_{T,j}^{2p}) \frac{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}{R^2}, \quad d_i = k_{T,i}^{2p}$$

The indices  $i$  and  $j$  refer to two PF particles, the parameter  $p$  is taken to be -1, and the radius parameter  $R$  can be tuned. For standard jets in CMS, as used in this analysis,  $R$  is set to 0.4. Particles are clustered from smallest  $d_{ij}$  to largest, until there is no  $d_{ij}$  smaller than  $k_{ti}^{-2}$ , at which point the final jet is defined. A visual representation of the anti- $k_{\mathrm{T}}$  algorithm is shown in Fig. 3.9.

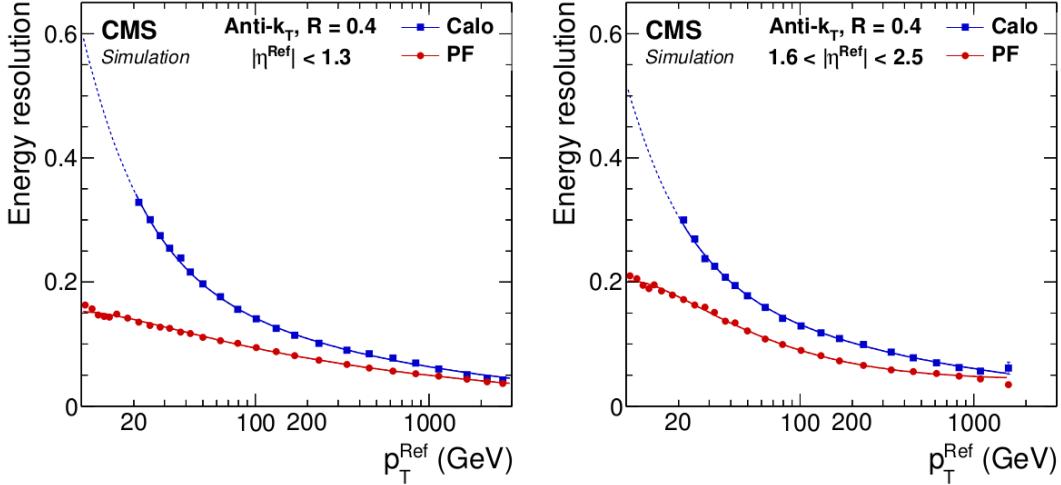


Figure 3.10. Jet energy resolution as a function of simulated jet  $p_T$  in the barrel (left) and in the endcap (right) regions [47]. Calo refers to the sum of ECAL and HCAL energy deposits, while PF corresponds to jets clustered from PF particles using the anti- $k_T$  algorithm.

Jet momentum is determined as the vectorial sum of all particle momenta in the jet, and is found from simulation to be, on average, within 5–10% of the true momentum over the entire  $p_T$  spectrum and detector acceptance. Additional pp interactions within the same or nearby bunch crossings (pileup) can contribute additional tracks and calorimetric energy depositions to the jet momentum. To mitigate this effect, charged particles identified as originating from pileup vertices are discarded, and an offset correction is applied to correct for the remaining contributions [54]. The jet energy resolution, as shown in Fig. 3.10, typically amounts to 15–20% at 30 GeV, 10% at 100 GeV, and 5% at 1 TeV [55].

## CHAPTER 4

### Analysis Overview

#### 4.1. History of the $H \rightarrow Z\gamma$ Search

Both the CMS and ATLAS collaborations have undertaken the search for  $H \rightarrow Z\gamma$  since Run 1 of the LHC. Before the CMS Run 2 analysis, which is the focus of this thesis, each collaboration published results using Run 1 data at  $\sqrt{s} = 7$  and 8 TeV [20, 19], CMS published a result with 2016 data at  $\sqrt{s} = 13$  TeV [21], and ATLAS published a full Run 2 result at  $\sqrt{s} = 13$  TeV [22]. As such, the present analysis represents the continuation of a broader effort, so we will give a brief history of these prior searches to put our work in context. While some of the analysis strategies of past searches inform our current strategy, there are also many differences and innovations in our analysis. We will emphasize the ways in which our analysis overlaps and differs with previous approaches, and will show that our innovations have significantly improved the expected sensitivity and statistical robustness of the search.

##### 4.1.1. The Run 1 Searches

In 2014, ATLAS published its Run 1 results [19] for the search for  $H \rightarrow Z\gamma$  where  $Z \rightarrow \ell^+ \ell^-$  with  $\ell = e$  or  $\mu$ . The baseline strategy was a standard dilepton plus photon selection, with  $m_{\ell^+ \ell^-}$  near the  $Z$  boson mass and  $\ell^+ \ell^- \gamma$  invariant mass ( $m_{\ell^+ \ell^- \gamma}$ ) between 115 and 170 GeV. The resolution of  $m_{\ell^+ \ell^- \gamma}$  was improved using a kinematic fit procedure,

accounting for the true Z boson line shape. Events were categorized based on lepton flavor; the pseudorapidity difference between the photon and Z boson; and the variable  $p_T^t$ , defined as  $|\vec{p}_T^{Z\gamma} \times \hat{t}|$ , where  $\hat{t} = (\vec{p}_T^Z - \vec{p}_T^\gamma)/|\vec{p}_T^Z - \vec{p}_T^\gamma|$  [56, 57], the  $p_T$  of the  $Z\gamma$  system that is perpendicular to the difference of the three-momenta of the Z boson and the photon, a quantity that is strongly correlated with the  $p_T$  of the  $\ell^+\ell^-\gamma$  system. Limits on the signal strength relative to the SM prediction, shown in Fig. 4.1 (left), were obtained from a simultaneous fit to the  $m_{\ell^+\ell^-\gamma}$  distributions in all categories.

The CMS Run 1 analysis [20] was published in 2015 for the  $\mu^+\mu^-\gamma$  and  $e^+e^-\gamma$  final states. It employed a standard dilepton plus photon selection, but the  $m_{\ell^+\ell^-\gamma}$  range was 100–190 GeV, which includes a kinematic turn-on around 110–115 GeV, mainly driven by the photon  $p_T$  selection. The choice to model this turn-on increased the complexity of the fit relative to the ATLAS analysis above, but was made in order to minimize potential bias. Events were categorized based on flavor; presence of a dijet system; lepton and photon pseudorapidity; and the photon  $R_9$ , which is the energy sum of the 3x3 ECAL crystal array centered on the most energetic crystal divided by the supercluster energy, and is associated with the quality of the photon. The CMS Run 1 limits on the signal strength are shown in Fig. 4.1 (right).

#### 4.1.2. The Previous Run 2 Searches

The first CMS  $H \rightarrow Z\gamma$  results at  $\sqrt{s} = 13$  TeV [21] were published in 2018 using the 2016 data set, corresponding to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The search strategy was largely similar to the CMS Run 1 analysis, with a few main differences. First, a boosted category was introduced to the analysis, defined by  $p_T^{\ell^+\ell^-\gamma} > 60 \text{ GeV}$ . This category

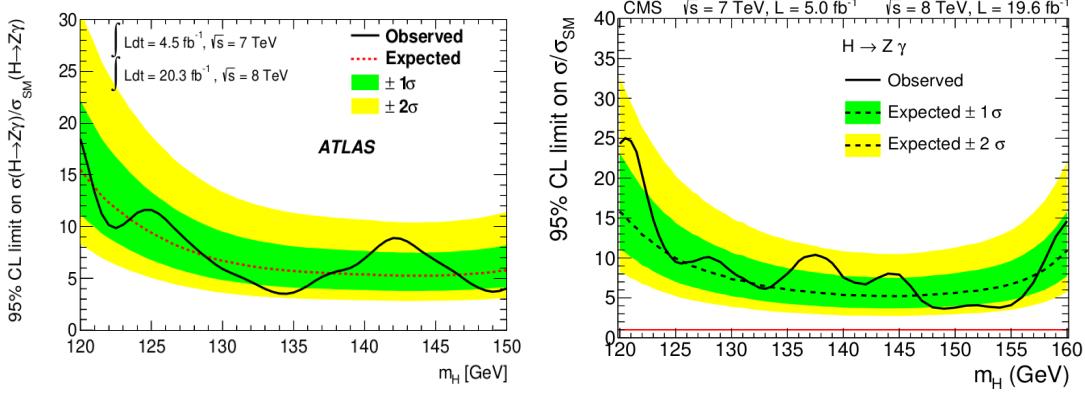


Figure 4.1. ATLAS (left [19]) and CMS (right [20]) Run 1 limit results as a function of  $m_H$ .

targeted events with a Higgs boson recoiling against a jet. Second, the range of  $m_{\ell^+\ell^-\gamma}$  was restricted to 115–170 GeV in order to avoid fitting the kinematic turn-on. The decision of whether or not to fit the turn-on became critically important in the full CMS Run 2 analysis, and will be discussed in more detail later in this thesis. Limits from the CMS 2016 search are shown in Fig. 4.2 (left).

The ATLAS full Run 2 search [22] was published in 2020, with a largely similar strategy to the Run 1 analysis. As in Run 1, a kinematic fit was used to improve the  $m_{\ell^+\ell^-\gamma}$  resolution and the variable  $p_T^t$  was used for categorization. However, there were a few major differences introduced in the Run 2 analysis. The fit range was changed to  $105 \text{ GeV} < m_{\ell^+\ell^-\gamma} < 160 \text{ GeV}$  and some aspects of the categorization strategy were updated. For categorization, the most significant change was the use of a boosted decision tree (BDT) trained to separate VBF Higgs production events from other Higgs production mechanisms and backgrounds. As will be discussed later in this thesis, the use of BDTs for categorization can significantly improve the sensitivity of searches for  $H \rightarrow Z\gamma$ . Figure

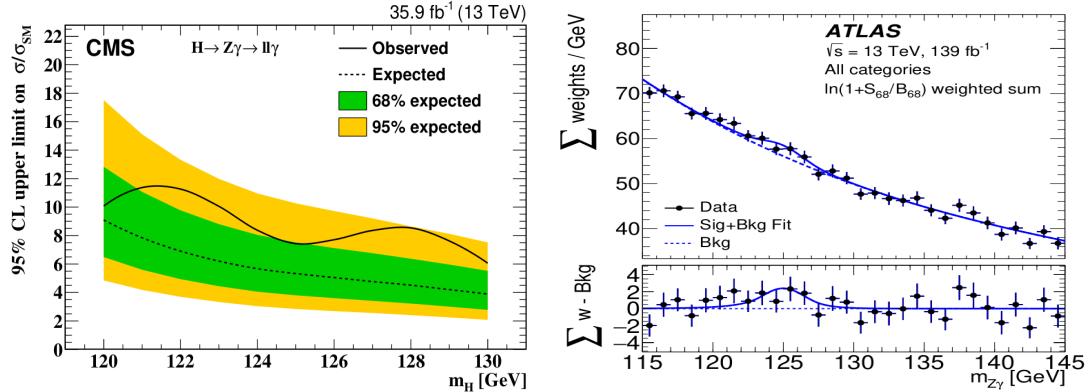


Figure 4.2. Left: CMS 2016 limit results [21] as a function of  $m_H$ . Right: ATLAS Run 2 signal plus background fit [22], for  $m_H = 125.09 \text{ GeV}$ , weighted by  $\ln(1+S_{68}/B_{68})$ , where  $S_{68}$  and  $B_{68}$  are the expected signal and background events in an  $m_{\ell^+\ell^-\gamma}$  window containing 68% of the expected signal yield.

4.2 (right) shows a summary of the fit results for the ATLAS Run 2 search. An excess in data was observed, corresponding to 2.2 standard deviations for  $m_H = 125.09 \text{ GeV}$ .

## 4.2. The CMS Run 2 Strategy

The remainder of this thesis provides a thorough description of the CMS Run 2 search for  $H \rightarrow Z\gamma$ . First, a brief overview is provided below. The data sample corresponds to an integrated luminosity of  $138 \text{ fb}^{-1}$  of pp collisions at  $\sqrt{s} = 13 \text{ TeV}$  accumulated between 2016 and 2018. Much of the analysis strategy has been inspired by the prior CMS and ATLAS  $H \rightarrow Z\gamma$  searches, as well as other CMS Higgs analyses. The baseline event selection is similar to the previous CMS searches. One crucial difference with respect to the CMS 2016 analysis is that the  $m_{\ell^+\ell^-\gamma}$  range has been extended to  $105\text{--}170 \text{ GeV}$ , including the kinematic turn-on, as was done in the CMS Run 1 analysis. This choice will be discussed in detail later in this thesis. Additionally, a kinematic fit procedure has been

added to the analysis to improve the  $m_{\ell^+\ell^-\gamma}$  resolution, similar to what has been done in the previous ATLAS  $H \rightarrow Z\gamma$  analyses and in other CMS Higgs analyses, including  $H \rightarrow ZZ^* \rightarrow 4\ell$  [58].

The sensitivity of the analysis is enhanced by searching for Higgs boson production in a variety of mechanisms, including gluon-gluon fusion ( $ggH$ ); vector boson fusion (VBF); and the associated production of a Higgs boson with either a weak vector boson ( $VH$ , where  $V = Z$  or  $W$ ) or a top quark pair ( $t\bar{t}H$ ). The dominant backgrounds arise from Drell–Yan production in association with an initial-state photon ( $Z/\gamma^* + \gamma$ ) and Drell–Yan production in association with jets, where a jet or additional lepton is misidentified as a photon ( $Z/\gamma^* + \text{jets}$ ). After using a set of discriminating variables to suppress background in the different production mechanisms, the signal is identified as a narrow resonant peak around  $m_H$  in the distribution of  $m_{\ell^+\ell^-\gamma}$ .

The data sample is divided into mutually exclusive categories according to (i) the presence of an additional lepton produced by  $Z(\rightarrow \ell^+\ell^-)$  or  $W(\rightarrow \ell\nu)$  decay, indicating the possible associated production of a Higgs boson with a  $W$  or  $Z$  boson, or  $t\bar{t}H$  production with a leptonic top quark decay; (ii) the value of a multivariate analysis (MVA) discriminant characterizing the kinematic properties of a dijet system together with the  $\ell^+\ell^-\gamma$  candidate, indicating possible VBF production; and (iii) the value of an MVA discriminant characterizing the kinematic properties of the  $\ell^+\ell^-\gamma$  system. This MVA categorization procedure is an innovation of the CMS Run 2 analysis. While a VBF MVA was used in the ATLAS full Run 2 analysis, we are the first to also leverage a kinematic MVA for maximal discrimination between signal and background.

## CHAPTER 5

### Data and Simulated Samples

The data sample used for the CMS Run 2  $H \rightarrow Z\gamma$  search corresponds to a total integrated luminosity of  $138\text{ fb}^{-1}$  and was collected over a data-taking period spanning three years:  $36.3\text{ fb}^{-1}$  in 2016,  $41.5\text{ fb}^{-1}$  in 2017, and  $59.8\text{ fb}^{-1}$  in 2018 [59, 60, 61]. To be considered in the analysis, events must satisfy the HLT requirements for at least one of the dielectron or dimuon triggers. The dielectron trigger requires a leading (subleading) electron with  $p_T > 23$  (12) GeV, while the dimuon trigger requires a leading (subleading) muon with  $p_T > 17$  (8) GeV. The corresponding CMS data streams are listed in Table 5.1. All data samples use the CMS MINIAOD data format. To ensure data quality, luminosity masks are applied based on recommendations from the Physics Performance and Dataset group.

Signal samples for ggH, VBF, VH, and  $t\bar{t}H$  production, with  $H \rightarrow Z\gamma$  and  $Z \rightarrow \ell^+\ell^-$  ( $\ell = e, \mu$ , or  $\tau$ ), are generated at next-to-leading order (NLO) using POWHEG v2.0 [62, 63]. These samples are produced for  $m_H$  of 120, 125, and 130 GeV. The dominant backgrounds,  $Z/\gamma^*(\rightarrow \ell^+\ell^-)+\gamma$  and  $Z/\gamma^*(\rightarrow \ell^+\ell^-)+\text{jets}$ , are generated at NLO using the MADGRAPH5\_aMC@NLO v2.6.0 (v2.6.1) generator [64] for 2016 (2017 and 2018) samples. Events arising from  $t\bar{t}$  production [65] are a relatively minor background and are generated at NLO with POWHEG v2.0 [62, 63]. The background from vector boson scattering (VBS) production of  $Z/\gamma^*+\gamma$  pairs, with the Z boson decaying to a pair of leptons, is simulated at leading order using the MADGRAPH5\_aMC@NLO generator. The decay  $H \rightarrow \mu^+\mu^-$

Table 5.1. Summary of MINIAOD data samples used.

$e^+e^-\gamma$ final state	$\mu^+\mu^+\gamma$ final state
/DoubleEG/Run2016B-17Jul2018_ver2-v1	/DoubleMuon/Run2016B-17Jul2018_ver2-v1
/DoubleEG/Run2016C-17Jul2018-v1	/DoubleMuon/Run2016C-17Jul2018-v1
/DoubleEG/Run2016D-17Jul2018-v1	/DoubleMuon/Run2016D-17Jul2018-v1
/DoubleEG/Run2016E-17Jul2018-v1	/DoubleMuon/Run2016E-17Jul2018-v1
/DoubleEG/Run2016F-17Jul2018-v1	/DoubleMuon/Run2016F-17Jul2018-v1
/DoubleEG/Run2016G-17Jul2018-v1	/DoubleMuon/Run2016G-17Jul2018-v1
/DoubleEG/Run2016H-17Jul2018-v1	/DoubleMuon/Run2016H-17Jul2018-v1
/DoubleEG/Run2016H-17Jul2018-v1	/DoubleMuon/Run2016H-17Jul2018-v1
/DoubleEG/Run2017B-31Mar2018-v1	/DoubleMuon/Run2017B-31Mar2018-v1
/DoubleEG/Run2017C-31Mar2018-v1	/DoubleMuon/Run2017C-31Mar2018-v1
/DoubleEG/Run2017D-31Mar2018-v1	/DoubleMuon/Run2017D-31Mar2018-v1
/DoubleEG/Run2017E-31Mar2018-v1	/DoubleMuon/Run2017E-31Mar2018-v1
/DoubleEG/Run2017F-31Mar2018-v1	/DoubleMuon/Run2017F-31Mar2018-v1
/EGamma/Run2018A-17Sep2018-v2	/DoubleMuon/Run2018A-17Sep2018-v2
/EGamma/Run2018B-17Sep2018-v1	/DoubleMuon/Run2018B-17Sep2018-v1
/EGamma/Run2018C-17Sep2018-v1	/DoubleMuon/Run2018C-17Sep2018-v1
/EGamma/Run2018D-22Jan2019-v2	/DoubleMuon/Run2018D-PromptReco-v2

is considered as a resonant background and is generated for the ggH, VBF, VH, and  $t\bar{t}H$  production mechanisms. The ggH production cross section is computed at next-to-next-to-NLO precision in QCD and at NLO in electroweak (EWK) theory [66]. The cross sections for Higgs boson production in the VBF [67] and VH [68] mechanisms are calculated at next-to-NLO in QCD, including NLO EWK corrections, while the  $t\bar{t}H$  cross section is computed at NLO in QCD and EWK theory [69].

A full list of simulated physics processes used in the analysis is provided in Table 5.2, along with cross section and branching fraction information. For the  $H \rightarrow Z\gamma$  and  $H \rightarrow \mu^+\mu^-$  samples, the SM Higgs boson production cross sections and branching fractions recommended by the LHC Higgs Working Group [12] are used for each mass point. The values for the Higgs boson processes listed in Table 5.2 correspond to  $m_H = 125$  GeV.

Table 5.2. Simulated event samples used in the analysis. For Higgs samples, the listed cross sections and branching fractions correspond to the values for  $m_H = 125$  GeV.

process	cross section (pb)	branching fraction
ggH $\rightarrow Z\gamma \rightarrow \ell^+\ell^-\gamma$	48.58	$1.548 \times 10^{-3}$
q $\bar{q}$ H $\rightarrow Z\gamma \rightarrow \ell^+\ell^-\gamma$	3.782	$1.548 \times 10^{-3}$
ZH $\rightarrow Z\gamma \rightarrow \ell^+\ell^-\gamma$	0.8839	$1.548 \times 10^{-3}$
W $^+H \rightarrow Z\gamma \rightarrow \ell^+\ell^-\gamma$	0.84	$1.548 \times 10^{-3}$
W $^-H \rightarrow Z\gamma \rightarrow \ell^+\ell^-\gamma$	0.532	$1.548 \times 10^{-3}$
t $\bar{t}H \rightarrow Z\gamma \rightarrow \ell^+\ell^-\gamma$	0.5071	$1.548 \times 10^{-3}$
ggH $\rightarrow \mu^+\mu^-$	48.58	$2.176 \times 10^{-4}$
q $\bar{q}$ H $\rightarrow \mu^+\mu^-$	3.782	$2.176 \times 10^{-4}$
ZH $\rightarrow \mu^+\mu^-$	0.8839	$2.176 \times 10^{-4}$
W $^+H \rightarrow \mu^+\mu^-$	0.84	$2.176 \times 10^{-4}$
W $^-H \rightarrow \mu^+\mu^-$	0.532	$2.176 \times 10^{-4}$
t $\bar{t}H \rightarrow \mu^+\mu^-$	0.5071	$2.176 \times 10^{-4}$
Z/ $\gamma^* \rightarrow \ell^+\ell^-\gamma$	117.864	—
Z/ $\gamma^* \rightarrow \ell^+\ell^- + \text{jets}$	6225.42	—
t $\bar{t}$ (inclusive)	831.76	—
VBF Z/ $\gamma^* + \gamma$	0.1097	—

All simulated events are interfaced with PYTHIA v8.226 (v8.230) [70] with the CUETP8M1 [71] (CP5 [72]) underlying event tune for 2016 (2017–2018) for the fragmentation and hadronization of partons and the internal bremsstrahlung of the leptons. The NLO parton distribution function (PDF) set, NNPDF v3.0 [73] (NNPDF v3.1) [74], is used to produce these samples in 2016 (2017–2018). The response of the CMS detector is modeled using the GEANT4 program [75]. The simulated events are reweighted to correct for differences between data and simulation in the number of additional pp interactions, trigger efficiencies, selection efficiencies, and efficiencies of isolation requirements for photons, electrons, and muons. These corrections and their associated uncertainties will be described in more

detail throughout the remainder of this thesis. The rest of this chapter will discuss general procedures and corrections related to the data and simulated samples.

### 5.1. Z+jets Overlap Removal

The two dominant backgrounds in this analysis are  $Z/\gamma^*(\rightarrow \ell^+\ell^-) + \gamma$ , or  $Z\gamma$  for short, and  $Z/\gamma^*(\rightarrow \ell^+\ell^-) + \text{jets}$ , or  $Z+\text{jets}$  for short. Both processes are simulated independently in order to obtain the best model of the kinematics of the  $\ell^+\ell^-\gamma$  system. The  $Z\gamma$  sample simulates photons from initial-state and final-state radiation at the matrix element level and includes a better computation of the interference between the corresponding Feynman diagrams than the  $Z+\text{jets}$  sample. However, the  $Z+\text{jets}$  simulated sample also includes the contribution from  $Z\gamma$ . Therefore, this contribution must be removed from the  $Z+\text{jets}$  sample in order to avoid double counting events and to utilize the better kinematic modeling of the  $Z\gamma$  sample. To handle this,  $Z+\text{jets}$  events are removed from the analysis if they contain prompt, final state photons within  $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.1$  of the photon selected to reconstruct the  $\ell^+\ell^-\gamma$  system. The events with prompt, final state photons are tagged using generator-level truth information.

### 5.2. Photon Internal Conversion

The PYTHIA 8 showering algorithm automatically includes a component in which  $H \rightarrow \gamma\gamma$  photons are internally converted via the process  $H \rightarrow \gamma\gamma^* \rightarrow f\bar{f}\gamma\gamma$ , with  $f\bar{f}$  mostly  $e^+e^-$ . This also affects the  $H \rightarrow Z\gamma$  signal and  $Z\gamma$  background samples. Since the normalization of these samples does not include the internal conversion component, the cross section must be adjusted accordingly. To determine the amount by which the cross section is adjusted, we measure the internal conversion rate in simulation. We define a

ratio where the numerator is the number of events with hard process final state photons and the denominator is total number of events in the simulation. Subtracting this ratio from unity gives the internal conversion rate. The hard process final state photon flagging is done using generator-level truth information. In our analysis, the internal conversion rate is about 3.1% (2016) and 3.0% (2017 and 2018). We, therefore, divide the affected sample cross sections by 0.969 (0.97) for 2016 (2017 and 2018).

### 5.3. Transverse Momentum Reweighting

The underlying event tuning was updated from CUETP8M1 (2016) to CP5 (2017–2018). Unfortunately, it was discovered that this update caused significant mismodeling of the  $\ell^+\ell^-\gamma p_T$  spectrum in the  $Z\gamma$  and  $Z+jets$  background samples in 2017 and 2018. Figure 5.1 (left) shows the clear difference between simulated  $p_T^{\ell^+\ell^-\gamma}$  in 2016 and 2017–2018. As will be described later in this thesis, the variable  $p_T^{\ell^+\ell^-\gamma}/m_{\ell^+\ell^-\gamma}$  has good discriminating power between signal and background and is a useful feature for MVA training. Therefore, it is important to correct this discrepancy. To do this, a simple reweighting of the distribution is performed using two sideband regions in  $m_{\ell^+\ell^-\gamma}$ . The sidebands are defined as 115–120 and 130–135 GeV. The resulting weights are then applied to the  $Z\gamma$  and  $Z+jets$  simulated events in 2017 and 2018. Figure 5.1 (right) shows the final distribution of  $p_T^{\ell^+\ell^-\gamma}/m_{\ell^+\ell^-\gamma}$  for all data-taking years combined after the  $p_T^{\ell^+\ell^-\gamma}$  reweighting is applied. The overall agreement between data and simulation is good over the full  $p_T$  spectrum.

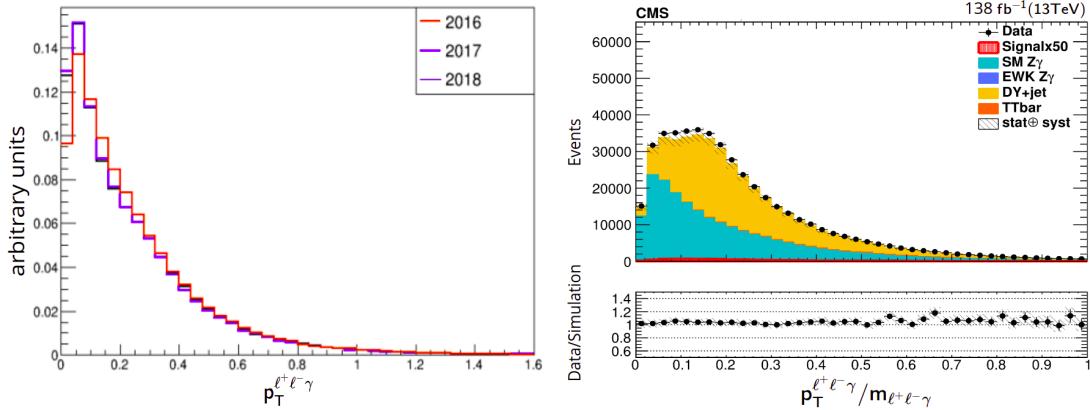


Figure 5.1. Left: Background simulation modeling of  $p_T^{\ell^+\ell^-\gamma}$  in each data-taking year. There is a shape discrepancy between 2016 and 2017–2018 due to the updated underlying event tuning. Right: Final comparison between data and simulation after the  $p_T^{\ell^+\ell^-\gamma}$  reweighting is performed on 2017 and 2018 background simulation.

#### 5.4. Pileup Reweighting

In order to correctly model the pileup distribution in data, a pileup reweighting procedure must be applied to the simulation. Although the Deterministic Annealing primary vertex reconstruction [76] has been shown to be efficient and well-behaved up to the observed levels of pileup, the final distribution for the number of reconstructed primary vertices is still sensitive to differences between data and simulation in the primary vertex reconstruction and underlying event. Additionally, the offline event selection criteria and trigger can potentially impact the distribution for the number of reconstructed vertices. To handle this, the pileup distribution for data is first derived by using the per-bunch-crossing, per-luminosity-section instantaneous luminosity, together with the total pp inelastic cross section. The total pp inelastic cross section is taken to be 66 mb, 72.4 mb,

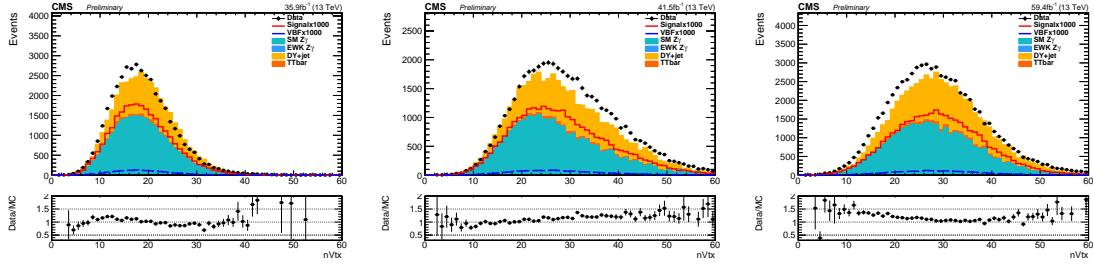


Figure 5.2. Distributions of the number of primary vertices in data and simulation for  $e^+e^-\gamma$  events in 2016 (left), 2017 (center), and 2018 (right) after pileup reweighting.

and 80 mb for 2016, 2017, and 2018 respectively. This yields an expected pileup distribution, correctly weighted over the entire data-taking period. The simulation is then reweighted to match this expected pileup distribution. Figure 5.2 shows the number of primary vertices after pileup reweighting for each data-taking year. The uncertainties associated with pileup reweighting are accounted for in the analysis and are discussed further in Chapter 9.

### 5.5. Level-1 Trigger Prefiring Reweighting

In 2016 and 2017, a gradual timing shift of the ECAL was not properly propagated to L1 trigger primitives, resulting in a significant fraction of high-pseudorapidity trigger primitives mistakenly associated to the previous bunch crossing [77]. Since L1 trigger rules forbid two consecutive bunch crossings from triggering, a consequence is that events can self-veto if a significant amount of ECAL energy is found in the region  $2 < |\eta| < 3$ . This effect is not described in simulation. To handle this, a weight is applied to each simulated event based on the probability of an event not to prefire and self-veto. The prefire probability is determined by measuring the timing distributions of physics objects in the subset of data events for which prefire is impossible due to the lack of a L1 trigger accept

in the previous bunch crossing. The measurement is parameterized in  $\eta$  and  $p_T$ . The final event weight applied to the simulation is the product of the non-prefiring probability of all objects, as shown in equation 5.1.

$$(5.1) \quad \omega = 1 - P(\text{Prefiring}) = \prod_{i=\text{photons, jets, muons}} (1 - \epsilon_i^{\text{pref}}(\eta, p_T^{\text{EM}, \mu}))$$

The impact of the L1 prefiring for electrons (muons) is a 1.4 (0.8)% loss in signal yield for 2016 and 2.4 (1.1)% loss in signal yield for 2017. The uncertainties associated with the L1 prefiring reweighting are accounted for in the analysis and are discussed further in Chapter 9.

## CHAPTER 6

### Physics Object and Event Selection

This chapter outlines the basic  $H \rightarrow Z\gamma$  physics object and event selection before any event categorization is done. Along the way, we will also point out the most relevant corrections applied, including trigger scale factors (SFs), physics object energy and momentum corrections, and identification and isolation SFs.

#### 6.1. Primary Vertex

Events are required to have at least one good primary vertex (PV) with a reconstructed longitudinal position within 24 cm of the geometric center of the detector and a transverse position within 2 cm of the nominal beam collision point. To avoid fake vertices, the PV reconstruction is required to have more than four degrees of freedom. In the case of multiple satisfactory vertices, the vertex with the highest scalar sum of squared  $p_T$  is chosen.

#### 6.2. Triggers

The topology and basic kinematics of the  $H \rightarrow Z\gamma$  process guide the choice of triggers used in this analysis. As this is a three body decay, the  $p_T$  of the photon tends to be less than in other channels like  $H \rightarrow \gamma\gamma$ . Consequently, the CMS photon trigger  $p_T$  thresholds are too great to make them viable options. Instead, we trigger on the leptons arising from the decay of the Z boson, which tend to have larger  $p_T$ . The best approach to maximize

signal efficiency is to use the double lepton triggers. The double muon trigger has  $p_T$  thresholds of 17 and 8 GeV, and the double electron trigger has thresholds of 23 and 12 GeV. These are the lowest unprescaled double lepton triggers generally available for CMS Run 2 analysis. Events passing both the double muon and double electron triggers are treated as double muon events.

The triggers are applied to both data and simulation. Trigger efficiencies and scale factors are measured using the simulation samples corresponding to each data-taking year. These measurements use a tag and probe method [CITE]. The tag and probe method takes advantage of the high purity of  $Z \rightarrow \ell^+\ell^-$  events near the Z mass peak. One lepton functions as the tag, and satisfies a set of tight trigger, identification, isolation,  $p_T$  requirements. The second lepton, the probe, must pass a looser selection and is used to measure the efficiency in question. Using this approach, trigger efficiencies for each leg of a given double lepton trigger are measured in both data and simulation. Then a corrective scale factor, defined as the ratio of data efficiency to simulation efficiency, is applied to the simulation. Scale factors are measured and applied in bins of  $p_T$  and  $|\eta|$ .

For the double electron trigger efficiency measurements, the tag electron must satisfy the following requirements. It must pass the single electron trigger, pass a tight cut-based identification, have  $p_T > 30$  (35) in 2016 (2017/2018), and have  $|\eta| < 2.5$ . The probe electron must pass the loose electron MVA identification (with isolation) requirement. The efficiencies for each leg of the trigger are measured separately, so in each case, the probe electron must match the trigger leg being measured. The efficiencies and scale factors of each double electron trigger leg for 2016, 2017, and 2018 are shown in Figures [FIGS].

For the double muon trigger efficiency measurements, the tag muon must satisfy the following requirements. It must pass the single muon trigger, pass a tight cut-based identification, have  $p_T > 26$  (29) GeV for 2016 (2017/2018) and satisfy  $|\eta| < 2.4$ . The probe muon must pass the  $H \rightarrow ZZ$  identification and isolation cuts. The details of the  $H \rightarrow ZZ$  muon identification will be described later. The efficiencies for each leg of the trigger are measured separately, so in each case, the probe muon must match the trigger leg being measured. The efficiencies and scale factors of each double muon trigger leg for 2016, 2017, and 2018 are shown in Figures [FIGS].

### 6.3. Photon Selection

Photons are identified using a boosted decision tree MVA discriminator trained on photon plus jets simulation. The features used in the training include supercluster kinematics, isolation variables, and shower shape variables [REF]. To reject electrons faking photons, a conversion-safe electron veto is applied. In order to improve agreement between simulation and data in the  $H \rightarrow Z\gamma$  analysis, shower shape corrections are taken from the Higgs to diphoton analysis [REF] and applied to simulated events. The original MVA provided by the EGamma POG is then reevaluated after these corrections. The following features are corrected:  $R_9$ , defined as the ratio of energy in the 5x5 array of ECAL crystals to the supercluster energy;  $S_4$ , defined as the ratio of the maximum energy 2x2 array to the energy of the 5x5 array; the energy weighted shower widths  $\sigma_\eta$  and  $\sigma_\phi$ ; the energy weighted widths by crystal index  $\sigma_{i\eta i\eta}$  and  $\sigma_{i\eta i\phi}$ ; photon isolation, charged isolation with respect to the primary vertex, and charged isolation with respect to the worst vertex choice. The validity of the shower shape corrections is checked using tag

and probe procedures for  $Z \rightarrow e^+e^-$  events where the probe electron mimics a photon and  $Z \rightarrow \mu^+\mu^-$  events with an FSR photon. A comparison of the agreement between uncorrected and corrected simulation with data is shown in Figures [FIGS]. Simulation comparisons with data for individual shower shape features before and after correction can be found in Appendix [APPENDIX], Figures [FIGS].

After deriving the corrected photon MVA discriminator, 90% signal efficiency working point cuts for barrel and endcap photons are determined for the  $H \rightarrow Z\gamma$  analysis. The working points are defined based on real photons from SM  $Z\gamma$  simulation, and correspond to discriminator scores greater than -0.4 (-0.59) for barrel (endcap) photons. A comparison of these working points with the standard EGamma POG 90% efficiency working points, plotted on the receiver operator characteristic (ROC) curve for SM  $Z\gamma$  and  $Z$  plus jets simulation, is shown in Figure [FIG]. The efficiency of the photon identification is measured with  $Z \rightarrow e^+e^-$  data using a tag and probe technique. The tag electron must pass the single electron trigger with  $p_T$  threshold 27 (32) GeV in 2016 (2017/2018), tight cut-based identification, have  $p_T > 30$  (35) GeV for 2016 (2017/2018), and have  $|\eta| < 2.5$ . The probe electron must pass the photon MVA identification with shower shape corrections described above. The scale factors, measured and applied in bins of  $p_T$  and supercluster  $\eta$ , are shown in Figure [FIG]. The efficiencies and scale factors for the conversion-safe electron veto are measured by the EGamma POG using  $Z\mu^+\mu^-$  events with an FSR photon and applied in the  $H \rightarrow Z\gamma$  analysis. The scale factors are catalogued in Table [TAB].

#### 6.4. Electron Selection

Electrons are identified using a boosted decision tree multivariate (MVA) discriminator trained on Drell-Yan plus jets simulation with prompt electrons matched to generator-level objects as signal and unmatched and non-prompt electrons as background. The features used in the training include  $p_T$ , supercluster  $\eta$ , shower shape variables, ratio of hadronic to electromagnetic energy, track and pixel hit variables, and isolation variables [REF]. These features are sensitive to bremsstrahlung along the electron trajectory, momentum-energy matching between electron trajectory and ECAL cluster, shower shape, and electrons from photon conversions. Since isolation features are included in the training of the discriminator, there is no need for separate isolation cuts. Electrons pass the identification requirement of the  $H \rightarrow Z\gamma$  analysis if their discriminator score is higher than a loose working point value, corresponding to 98% signal efficiency. In addition to the MVA cut, electrons must have  $|d_{xy}| < 0.5$  cm and  $|d_z| < 1$  cm with respect to the primary vertex. Electrons with  $p_T \leq 7$  GeV are rejected.

Electron identification efficiencies and scale factors are measured using a tag and probe method on dielectron events near the Z boson peak. These electrons must pass the single electron trigger with a  $p_T$  threshold of 27 (32) GeV for 2016 (2017/2018) data. The dielectron mass must be between 60 and 120 GeV. The tag electron must pass a tight cut-based electron identification and have  $p_T > 30$  (35) GeV for 2016 (2017) and  $|\eta| < 2.5$ . The probe electron must pass the loose MVA identification cut and associated impact parameter and  $p_T$  cuts described above. Identification efficiencies and scale factors are measured and applied in bins of  $p_T$  and supercluster  $\eta$ . The scale factors for each data-taking year are shown in Figures [FIGS].

### 6.5. Muon Selection

A loose cut-based muon identification is used in the analysis. This identification was originally developed by the  $H \rightarrow ZZ$  analysis [REF] and is well-suited for  $H \rightarrow Z\gamma$  due to the similarity of the multiple, potentially soft, muons in the final state. All muons are first required to pass a set of common cuts, followed by a separate set of cuts for high  $p_T$  (greater than 200 GeV) and low  $p_T$  muons. All muons must satisfy  $p_T > 5$  GeV,  $|\eta| < 2.4$ ,  $|d_{xy}| < 0.5$  cm, and  $|d_z| < 1$  cm, where  $d_{xy}$  and  $d_z$  are impact parameters defined with respect to the primary vertex of the interaction using the best muon track. Additionally, the three dimensional impact parameter (analogously defined) must have a magnitude less than four times its uncertainty. Muons must either be reconstructed as global muons or tracker muons. Muons with standalone tracks (tracks only in the muon system) are rejected. Muons must pass a particle flow-based isolation requirement, where the relative particle flow isolation within a  $\Delta R = 0.3$  cone is defined as

$$(6.1) \quad \mathcal{I} \equiv \left( \sum p_T^{\text{charged}} + \max [0, \sum p_T^{\text{neutral}} + \sum p_T^\gamma - p_T^{\text{PU}}(\ell)] \right) / p_T^\ell.$$

Muons must satisfy  $\mathcal{I} < 0.35$ .

For muons with  $p_T < 200$  GeV, muons satisfying the common requirements and identified by the particle flow identification algorithm are selected. For muons with  $p_T > 200$  GeV, muons are selected if they pass the particle flow identification or if they pass a set of high- $p_T$  requirements. These requirements are the following: the muon must be matched to segments in at least two muon stations; satisfy  $\frac{p_T}{\sigma_{p_T}} < 0.3$ ,  $|d_{xy}| < 0.2$  cm,  $|d_z| < 0.5$  cm, have at least one pixel hit, and have tracker hits in at least six tracker layers.

In 2016, data-taking was affected by a problem in which the level one trigger sent only one candidate per  $60^\circ$  sector instead of up to three [REF?]. As a result, when two muons in the same endcap had a low  $\Delta\phi$  separation, only one would fire the trigger. To account for this, 2016 data events containing identified muons with  $\Delta\phi < 70^\circ$  in the same endcap region are rejected.

Identification efficiencies and scale factors corresponding to the muon identification above are measured and provided by the  $H \rightarrow ZZ$  analysis working group, and are shown in Figure [FIG].

## 6.6. Jet Selection

Jets are selected in order to categorize events coming from potential VBF Higgs production, but no jet multiplicity requirement is present for the  $H \rightarrow Z\gamma$  selection. In fact, the majority of simulation and data events selected in the analysis have no jets. However, identifying and selecting jets to categorize VBF events can still significantly improve the sensitivity of the search. Jets are required to pass a loose cut-based identification in 2016 and a tight cut-based identification in 2017 and 2018. These sets of identification cuts are determined and provided by the JetMET POG. Additionally, jets must satisfy  $p_T > 30$  GeV,  $|\eta| < 4.7$ , and  $\Delta R > 0.4$  with respect to each lepton and the photon selected in the analysis. An issue with noise in the ECAL endcap in 2017 caused an artificial increase in jet multiplicity in data within a specific kinematic phase space [REF]. To mitigate this, jets are rejected if they have raw  $p_T < 50$  GeV and  $2.65 < |\eta| < 3.139$ . This cut reduces the efficiency to reconstruct dijet pairs by 12% in the specified region. To tag VBF events,

we are interested in dijet pairs. To this end, if there are more than two jets satisfying the above criteria, only the two jets with highest  $p_T$  are selected.

## 6.7. Object Corrections

Several standard corrections are applied to the physics objects selected in the analysis. Rochester muon momentum scale and resolution corrections are applied to both data and simulation [REF]. Energy and momentum scale and resolution corrections for electrons and photons are provided by the EGamma POG and applied in the analysis. Jet momentum scale and resolution correctors are provided by the JetMET POG and applied in the analysis.

An additional muon momentum correction is obtained using an FSR photon recovery procedure, and is based on the procedure used by the  $H \rightarrow ZZ$  analysis [REF]. FSR photons are not considered during the standard CMS muon reconstruction. As the  $p_T$  of FSR photons from radiating muons is generally very low, particle flow photons are considered, in contrast with the fully reconstructed photons used in the main analysis selection. A particle flow photon must pass a set of cuts in order to be identified as an FSR photon associated with one of the muons selected by the analysis. The photon must have  $p_T > 2$  GeV,  $|\eta| < 2.4$ , and relative particle flow isolation less than 1.8. It must also satisfy  $\Delta R(\gamma, \mu)/p_{T,\gamma}^2 < 0.012$  and  $\Delta R(\gamma, \mu) < 0.4$ . If multiple particle flow photons pass these requirements, the photon with the smallest  $\Delta R(\gamma, \mu)/p_{T,\gamma}^2 < 0.012$  is chosen. Then, the four momentum of the FSR photon is added back to the four momentum of the muon, and the muon kinematics reevaluated. In simulation, we find that the selected FSR photon matches the generator level FSR photon with 93% efficiency. Figure [FIG]

shows the dilepton and three body invariant mass among FSR photon-containing signal events with and without applying the FSR recovery correction. The procedure yields a 1% improvement on the three body mass resolution in the muon channel.

## 6.8. Event Selection

Events are required to have at least one good primary vertex with a reconstructed longitudinal position within 24 cm of the geometric center of the detector and a transverse position within 2 cm of the nominal beam collision point. The vertex with the largest value of summed physics-object  $p_T^2$  is taken to be the primary interaction vertex. The physics objects used in this calculation are derived from information obtained from the charged-particle tracking detectors only. These objects include jets reconstructed by clustering charged particle tracks; the associated missing transverse momentum, defined as the negative vector sum of the  $p_T$  of those jets.

Events with two same-flavor opposite sign leptons ( $e$  or  $\mu$ ) and a photon are selected. The leading muon (electron) is required to have  $p_T$  greater than 25 (20) GeV, and the trailing lepton must have  $p_T$  greater than 15 (10) GeV. Electrons (muons) must have  $|\eta|$  less than 2.5 (2.4). The photon must have  $p_T$  greater than 15 GeV and must satisfy  $0 < |\eta| < 1.4442$  or  $1.566 < |\eta| < 2.5$ . This avoids the calorimeter transition region, in which photon reconstruction is more difficult. The invariant mass of the dilepton system is required to be greater than 50 GeV. In events with multiple dilepton pairs, the pair with mass closest to the nominal Z boson mass [REF] is selected.

Events are required to have a photon satisfying  $p_T^\gamma/m_{\ell\ell\gamma} > 0.14$ , which suppresses the Z plus jets background without significantly reducing signal efficiency and without

significantly shaping the three body mass spectrum. Each lepton must have  $\Delta R > 0.4$  with respect to the photon. To reject events with final-state radiation from Drell-Yan processes, we require  $m_{\ell\ell\gamma} + m_{\ell\ell} > 185$  GeV. Finally, the three body mass is required to satisfy  $105 < m_{\ell\ell\gamma} < 170$  GeV.

### 6.9. Kinematic Fit

A kinematic fit technique is used to improve the dilepton mass resolution. This constrains the dilepton mass based on the true Z boson lineshape while accounting for the known detector resolution. The position information of the leptons has a negligible impact on the fit, so only the lepton transverse momenta and dilepton mass are included in the fit. The procedure is based on previous studies by the  $H \rightarrow ZZ$  analysis [REF]. We do not carry out an analogous kinematic fit for the  $m_{\ell\ell\gamma}$  invariant mass in order to avoid any potential bias due to reshaping the mass distribution. The kinematic fit is a maximum likelihood fit defined by the likelihood function below:

$$(6.2) \quad \begin{aligned} & \mathcal{L}(P_T^1, P_T^2 | P_T^{reco1}, \sigma_{P_T}^{-1}, P_T^{reco2}, \sigma_{P_T}^{-2}) \\ & = \text{Gauss}(P_T^{reco1} | P_T^1, \sigma_{P_T}^{-1}) \cdot \text{Gauss}(P_T^{reco2} | P_T^2, \sigma_{P_T}^{-2}) \cdot \mathcal{L}(m_{12} | m_Z) \end{aligned}$$

Here,  $p_T^{reco1}$  and  $p_T^{reco2}$  are the reconstructed transverse momenta of the two leptons,  $\sigma_{p_T}^1$  and  $\sigma_{p_T}^2$  are the per-lepton transverse momentum resolutions,  $p_T^1$  and  $p_T^2$  are the parameters to optimize, and  $m_{12}$  is the invariant mass calculated from  $p_T^1$  and  $p_T^2$ .  $\mathcal{L}(m_{12} | m_Z)$  is the likelihood given the true Z mass lineshape. The outputs of the fit are  $p_T^1$  and  $p_T^2$ . Their uncertainties are saved and used to calculate a per-event uncertainty for the

mass measurement. To optimize this procedure, we determine the true generator level Z lineshape from a gluon-gluon fusion  $H \rightarrow Z\gamma$  sample. Generator level Z lineshapes for the dielectron and dimuon final states are fit with a single-sided Crystal Ball function plus three gaussian functions. For each event, the likelihood is maximized and the  $p_T$  information of the refit leptons is updated. Then, the dilepton mass and three body mass are recalculated based on the refit leptons. The mass distributions for before and after the kinematic fit are shown in Figure [FIG]. The level of improvement in the dilepton mass resolution, as measured by  $\sigma_{eff}$ , is shown in Table 6.1.

Table 6.1. Percent improvement in dilepton and three-body mass resolution (measured by  $\sigma_{eff}$ ) after the kinematic fit.

	<b>electron channel</b>		<b>muon channel</b>	
	$m_{ee}$	$m_{e\gamma}$	$m_{\mu\mu}$	$m_{\mu\mu\gamma}$
<b>2016</b>	20%	20%	17%	12%
<b>2017</b>	28%	27%	21%	11%
<b>2018</b>	24%	24%	20%	10%

We also perform a closure test on the background as shown in Figure [FIG]. The dilepton mass should be constrained, while the three body mass should not be distorted by the kinematic constraint. Indeed, this is what is observed.

## CHAPTER 7

### Event Categorization

The sensitivity of the  $H \rightarrow Z\gamma$  search can be significantly improved by leveraging the differences between signal events arising from different Higgs boson production modes with different final state topologies and kinematics. This is achieved by defining mutually exclusive categories targeting these different types of signal events. In previously published CMS searches for  $H \rightarrow Z\gamma$ [REFS], cut-based approaches were used to define the categories. Broadly, the presence of at least one additional lepton was used to tag VH and ttH production, a dijet system was used to tag VBF production, a boosted jet was used to tag gluon-gluon fusion events recoiling off of a jet, and properties of the photon were used to divide the remaining events into kinematically distinct untagged categories. The specific definitions of the categories used in the 2016 search at 13 TeV are shown in Figure [FIG].

The categorization of the present search is inspired by the previous CMS searches, but is significantly more sophisticated and better optimized than those searches. The lepton tag category remains, as do the concepts of dijet and untagged categories. However, the dijet and untagged categories are determined via the training of two BDTs, with category boundaries optimized using the resulting BDT scores. Moreover, the boosted category is studied and found to yield no improvement over the categories determined by the BDTs, so the boosted category is dropped in the present analysis. We proceed with a description of the BDT procedures for the untagged and dijet cases. Then, with our BDTs in hand,

we describe the procedure for defining an optimal set of categories to be used in the analysis. We show that this categorization procedure is more optimal than what has been done in past CMS searches.

### 7.1. Kinematic Boosted Decision Tree

The kinematic BDT is used to separate  $H \rightarrow Z\gamma$  signal events from background events. Its primary purpose is to define the untagged categories, which are distinguished primarily by kinematics and the quality of the final state photon. Its secondary use is as an input to the dijet BDT (described below) in order to increase signal to background discrimination in the dijet BDT training. The kinematic BDT is trained on  $H \rightarrow Z\gamma$  signal events from all Higgs production modes and background events from SM  $Z\gamma$ ,  $Z$  plus jets, and  $t\bar{t}$ . The training is restricted to use half of the simulated events, with the other half reserved for category optimization and signal modeling. All training events are required to pass the basic object and event selection described in Chapter 6. Events from 2016, 2017, and 2018 simulation samples in both the muon and electron channels are combined for the training. These samples are weighted by their respective cross sections and weighted by the luminosity of each year.

The following features are used in the training of the kinematic BDT: photon MVA score, photon energy resolution,  $\eta$  of each lepton and the photon, smallest and largest of the two  $\Delta R$  values of the photon with respect to each lepton,  $p_T/m$  for the three body system, and three theoretically motivated decay angles. The angular quantities are defined in Figure [FIG] and their theoretical, generator-level, and reconstructed distributions are shown in Figure [FIG]. The signal and background distributions used in the

training for all kinematic BDT features are shown in Figure [FIG]. Table [TAB] ranks the kinematic BDT training features in order of importance for the final discriminator. We observe that the most discriminating feature is the photon MVA identification score which mainly serves distinguish real photons from misreconstructed jets. A comparison of the BDT distributions in the training and test samples allows us to check for any possible overtraining of the model. This is shown in Figure [FIG], where no significant overtraining is observed.

The output of the kinematic BDT in data and simulation is shown in Figure [FIG]. We observe good agreement between data and simulation. For additional validation of the method, two control regions are defined. A background-enriched control region is defined by  $m_{\ell\ell} + m_{\ell\ell\gamma} < 185$  GeV or  $p_T^\gamma < 15/110$ , and an irreducible background-enriched (SM  $Z\gamma$ ) control region is defince by  $m_{\ell\ell} + m_{\ell\ell\gamma} < 185$  GeV and  $80 < m_{\ell\ell} < 100$  GeV. The data and simulation distributions for the kinematic BDT for these control regions are shown in Figures [FIGS]. Again, we observe good agreement between data and simulation.

## 7.2. Vector Boson Fusion Boosted Decision Tree

The dijet BDT is used to discriminate signal from VBF Higgs production from other dijet signal and background events. The training and evaluation of the dijet BDT is only carried out on events which have two jets selected according to the requirements in section [SEC]. The training is performed using VBF  $H \rightarrow Z\gamma$ events as signal and gluon-gluon fusion  $H \rightarrow Z\gamma$ , SM  $Z\gamma$ ,  $Z$  plus jets, and  $t\bar{t}$  events as background.  $VH$  and  $ttH$   $H \rightarrow Z\gamma$ events are neglected, as their contribution is negligibly small. Only about 65% of the signal jets correspond to the true VBF jets in which we are interested. Because of this,

we perform an additional matching procedure for these jets. In order for an event to be used in the training, both reconstructed jets must be matched to generator-level partons, which must also be matched to the generator-level  $p_T$  values of the true VBF partons. As with the kinematic BDT, only half of the simulated events are used for the training procedure, with the remainder used for category optimization and signal modeling. Again, events from all data-taking years in both the muon and electron channels are combined for the training. The samples are weighted by their respective cross sections and weighted by the luminosity of each year.

The input features to the dijet BDT training include  $\Delta\eta(j, j)$ ,  $\Delta\phi(j, j)$ ,  $\Delta\phi(Z\gamma, jj)$ ,  $\Delta R(j, \gamma)$ ,  $p_T^{j1,2}$ ,  $p_T^t$  (defined as  $\frac{2|p_{Tx}^Z p_{Ty}^\gamma - p_{Ty}^Z p_{Tx}^\gamma|}{p_T^H}$ ), system  $p_T$  balance (defined as  $|\frac{\sum_{Z,\gamma,j_1,j_2} \vec{p}_{transverse}^t}{\sum_{Z,\gamma,j_1,j_2} p_T}|$ ), photon Zeppenfeld variable (defined as  $|\eta_\gamma - \frac{\eta_{j1} + \eta_{j2}}{2}|$ ), and kinematic MVA score. The angular quantities provide discrimination between VBF and non-VBF events, while the kinematic MVA score and  $p_T^t$  provide discrimination between signal and background events. The  $p_T^t$  [REF] is the component of the transverse momentum of the  $Z\gamma$  system that is perpendicular to the difference of the 3-momenta of the  $Z$  boson and the photon candidate.

### 7.3. Categorization Procedure

The procedure used for event categorization is described below.

- (1) Events with at least one additional electron or muon with  $p_T > 7$  (5) GeV are assigned to the lepton-tagged category.
- (2) Events not assigned to the lepton-tagged category, but which contain two jets satisfying the selection requirements described in Section 6, are classified as dijet events, indicative of possible VBF production. If multiple dijet pairs exist within

an event, the two jets with highest  $p_T$  are considered. The subdivision of dijet events into a set of three dijet categories is described later in this section. The VBF BDT classifier is trained to separate VBF signal events from ggH+jets and background events. The distribution of  $\mathcal{D}_{\text{VBF}}$  is shown in Fig. 7.1 (left) for both simulated event samples and data.

- (3) Events not assigned to the lepton-tagged or dijet categories are classified as untagged events. The subdivision of untagged events into a set of four untagged categories is described later in this section. The kinematic BDT classifier is trained to distinguish signal events from background events based on the kinematics of the leptons and photon in the  $Z\gamma$  candidate system, as well as on the measured properties of these physics objects. The distribution of  $\mathcal{D}_{\text{kin}}$  is shown in Fig. 7.1 (right) for both simulated samples and data.

The subdivision of dijet and untagged events into categories is based on the VBF BDT and kinematic BDT discriminants. Category boundaries are defined as mutually exclusive regions of  $\mathcal{D}_{\text{VBF}}$  and  $\mathcal{D}_{\text{kin}}$ . The locations of the boundaries defining the categories are optimized by iterating over all possible combinations of boundaries using  $\sum_i^n S_i^2/B_i$  as a figure-of-merit. The variables  $S_i$  and  $B_i$  represent the number of expected signal and background events in the  $i^{th}$  category, and  $n$  is the total number of categories. We consider categories with boundaries corresponding to signal efficiencies between 0–100% in 10% increments. The optimization procedure results in three dijet categories for the VBF BDT and four untagged categories for the kinematic BDT. The lowest  $\mathcal{D}_{\text{kin}}$  boundary corresponds to the 10% point in integrated signal efficiency, and events below the 10% point are excluded from the analysis to preserve the stability of the background model.

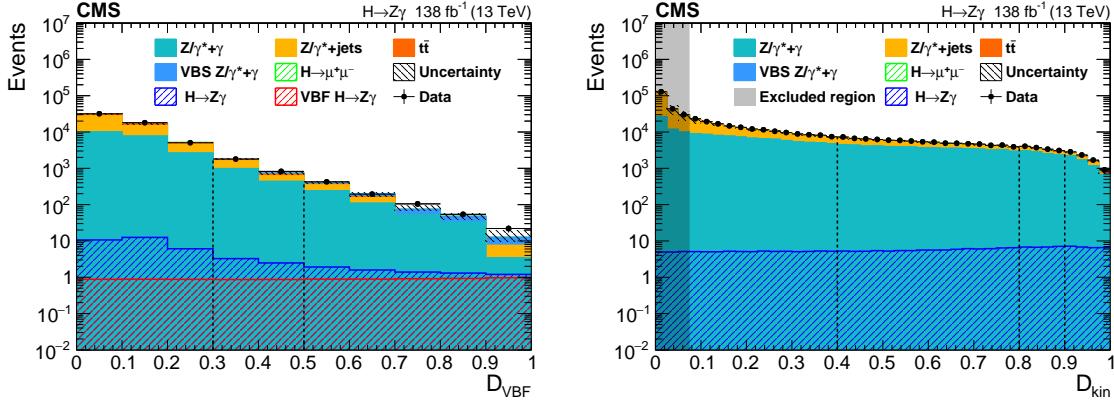


Figure 7.1. The  $\mathcal{D}_{VBF}$  (left) and  $\mathcal{D}_{kin}$  (right) distributions for signal, simulated background, and data. The  $\mathcal{D}_{VBF}$  distribution includes only dijet-tagged events, and the  $\mathcal{D}_{kin}$  distribution includes only untagged events. The sum of contributions from all signal production mechanisms is shown by the blue line, while the contribution from only the VBF mechanism is shown by the red line. The uncertainty band incorporates all statistical and systematic uncertainties in the expected background. The dashed lines indicate the boundaries for the dijet and untagged categories. The gray shaded region in the  $\mathcal{D}_{kin}$  distribution is excluded from the analysis.

Table 7.1. Summary of the category definitions. The lepton-tagged category requires at least one additional electron or muon. Dijet categories are defined by regions of  $\mathcal{D}_{VBF}$  and untagged categories are defined by regions of  $\mathcal{D}_{kin}$ .

Lepton	Dijet 1	Dijet 2	Dijet 3	Untagged 1	Untagged 2	Untagged 3	Untagged 4
	$\mathcal{D}_{VBF}$ selection			$\mathcal{D}_{kin}$ selection			
$\geq 1 e, \mu$	0.5–1.0	0.3–0.5	0.0–0.3	0.9–1.0	0.8–0.9	0.4–0.8	0.1–0.4

The full categorization and optimization procedure results in the following eight mutually exclusive categories: one lepton-tagged category, three dijet categories, and four untagged categories. The category definitions are summarized in Table 7.1.

Table 7.2 lists the event categories used in the analysis, along with the expected event yields for an  $m_H = 125.38$  GeV signal arising from ggH, VBF, VH, and  $t\bar{t}H$  production, as well as a 6% resonant background contribution from FSR from  $H \rightarrow \mu^+\mu^-$ . The

Table 7.2. Yields and approximate significance ( $S/\sqrt{B}$ ) for each category, where  $S$  and  $B$  are the expected number of signal and background events in the narrowest  $m_{\ell^+\ell^-\gamma}$  interval containing 95% of the expected signal distribution. Also shown is the  $m_{\ell^+\ell^-\gamma}$  resolution, computed using the narrowest interval containing 68% of the expected signal distribution.

$138 \text{ fb}^{-1}$	Lepton		Dijet 1	Dijet 2	Dijet 3	Untagged 1	Untagged 2	Untagged 3	Untagged 4
<b>SM signal yield</b>									
ggH	0.51	$e^+e^-$ $\mu^+\mu^-$	1.10 1.41	1.62 2.05	9.44 12.1	6.89 8.52	7.35 9.17	29.8 38.0	22.5 29.0
VBF	0.09	$e^+e^-$ $\mu^+\mu^-$	1.94 2.40	0.76 0.97	1.13 1.43	0.71 0.89	0.35 0.43	0.92 1.18	0.51 0.65
VH + $t\bar{t}H$	1.84	$e^+e^-$ $\mu^+\mu^-$	0.04 0.05	0.13 0.16	1.89 2.36	0.31 0.39	0.17 0.21	0.45 0.57	0.27 0.33
<b>SM resonant background</b>									
$H \rightarrow \mu^+\mu^-$	0.14	$\mu^+\mu^-$	0.27	0.27	0.43	0.62	0.49	2.02	1.78
<b>Mass resolution (GeV)</b>		2.12	$e^+e^-$ $\mu^+\mu^-$	1.91 1.52	2.06 1.61	2.15 1.72	1.80 1.37	1.97 1.42	2.12 1.62
<b>Data yield</b>	1485		168	589	11596	1485	1541	2559	17608
$S/\sqrt{B}$	0.06		0.54	0.24	0.26	0.45	0.35	0.53	0.30

expected event yield from  $H \rightarrow \tau^+\tau^-$  is estimated to be below the 1% level relative to the  $H \rightarrow Z\gamma$  yield and is neglected. The dominant contribution to the signal yield is generally from ggH production, except in the lepton-tagged category, in which VH and  $t\bar{t}H$  events dominate, and in the dijet 1 category, in which VBF events dominate. The categorization procedure increases the sensitivity of the analysis by 24% with respect to an inclusive event selection. The product of signal acceptance and efficiency for  $pp \rightarrow H \rightarrow Z\gamma \rightarrow \ell^+\ell^-\gamma$  for  $m_H = 125.38 \text{ GeV}$  is 23 (29)% in the electron (muon) channel.

#### 7.4. Dropping of Boosted Category

In the previous 2016 analysis, a boosted Higgs category was used. This category is defined by a simple cut on the  $p_T$  of the  $\ell\ell\gamma$  system. However, for the current MVA-based analysis, the Higgs kinematics are already included in the BDT training. In this case, it stands to reason that the boosted category may no longer be useful for improving the analysis sensitivity. We therefore test the possibility of constructing a boosted category before constructing the untagged categories to see whether the combined significance is increased or decreased. The left plot of Figure 7.2 shows the  $p_T$  spectrum of the  $\ell\ell\gamma$  system for various signal MC samples. We see that the  $p_T$  spectrum for the gluon-gluon fusion process is similar to the main backgrounds, while the other signal processes are more boosted. The right plot shows the upper limit on  $\sigma/\sigma_{SM}$  as a function of a  $p_T$  cut on the  $\ell\ell\gamma$  system, where the background estimation is taken from data. The upper limit for the boosted category (shown in blue) worsens as the  $p_T$  cut increases. This is expected since we lose a substantial amount of gluon-gluon fusion signal. However, the combined limit of the untagged categories improves. The overall combined limit (shown in black) is rather flat as a function of the  $p_T$  of the  $\ell\ell\gamma$  system. For the overall combined limit, we observe fluctuations beyond a  $p_T$  cut of 60 GeV, so we choose 60 GeV as the cut defining a potential boosted category. Comparing the combined significance of the boosted category plus untagged categories to the untagged categories alone, we see that the combined significance is substantially higher for the untagged categories alone. An explanation for this behavior is as follows. In the current analysis flow, dijet events are selected before boosted events. We loosen the dijet selection requirements, which causes the boosted

events to migrate to the dijet categories and reduces the significance of boosted category.

Therefore, we choose not to use a boosted category in this version of the analysis.

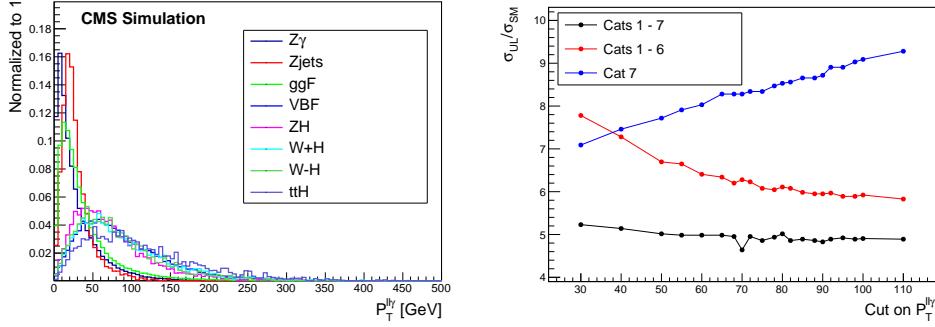


Figure 7.2. Left:  $\ell\ell\gamma p_T$  for different signal MC processes. Right: optimization of the boosted category by scanning the expected limit as function of  $\ell\ell\gamma p_T$ .

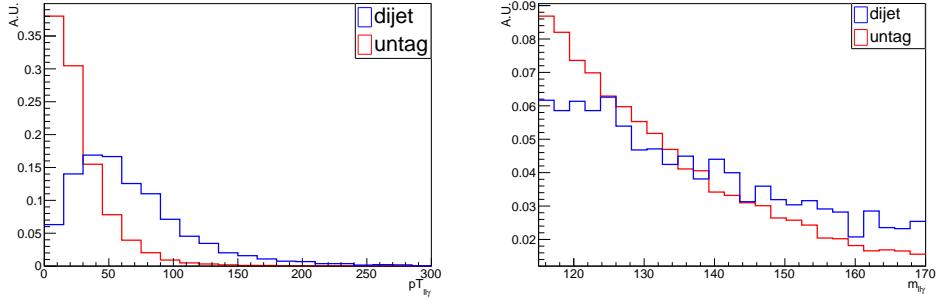


Figure 7.3. Left:  $\ell\ell\gamma p_T$  for untagged events and dijet events. Right: The  $m_{\ell\ell\gamma}$  spectrum of untagged events and dijet events.

## CHAPTER 8

### Signal and Background Modeling

In each category, we carry out a shape analysis to search for a signal peak in the  $m_{\ell\ell\gamma}$  spectrum. The signal and background mass shapes are modeled using parametric functions.

#### 8.1. Signal Modeling

The signal model is defined as the sum of Crystal Ball [78] and Gaussian functions. The signal shape parameters are determined by fitting this model to simulated signal events in each category. To account for differences in mass resolution, these fits are performed separately for the event samples used to model each data-taking year, as well as for muon and electron channel events. This results in six signal models that are summed to give the total signal expectation in a given category. Separate sets of parameter values are found by fitting simulated events with  $m_H$  of 120, 125, and 130 GeV. Using linear interpolation, parameter values are also determined at 1 GeV intervals in  $m_H$  from 120–130 GeV, as well as at 125.38 GeV. In the fit to data, the mean and resolution parameters are allowed to vary subject to constraints from several systematic uncertainties, described in Section 9, while the remaining parameters are held fixed.

Figure [REF] (REF) shows the signal fits for the  $m_H = 125$  GeV simulated samples in the electron (muon) channel for the 2016 data-taking period. Figure [REF] shows the signal fits for the lepton-tagged category for the 2016 data-taking period, in which electron

and muon channel events are combined. In the lepton-tagged category, the signal shape is modeled separately for ZH and WH production in order to account for differences in potential lepton mispairing. Signal fits for the 2017 and 2018 data-taking periods are shown in Figs. [REF] in Appendix [REF].

## 8.2. Resonant Background Modeling

The process  $H \rightarrow \mu^+ \mu^-$ , where at least one muon radiates an FSR photon, is expected to contribute at the 6% level relative to the  $H \rightarrow Z\gamma$  signal yield. Because of this, we treat it as a resonant background. To model the  $H \rightarrow \mu^+ \mu^-$  shape, we perform fits to the  $m_{\ell^+ \ell^- \gamma}$  distributions in simulated event samples, following the same procedure used for the signal fits described above. As for the signal, the analytic model is taken to be the sum of Crystal Ball and Gaussian functions. The resulting fits are shown in Figs. [REF].

## 8.3. Nonresonant Background Modeling

As in previous versions of the analysis, we use a parametric fit to the  $m_{\ell^+ \ell^- \gamma}$  spectrum in data to estimate the background. The background model in each category is obtained from the data using the discrete profiling method [79]. This technique accounts for the systematic uncertainty associated with choosing an analytic functional form to fit the background. The background function is chosen from a set of candidate functions via a discrete nuisance parameter in the fit. These functions are derived from the data in each category, with muon and electron events from all data-taking years combined. The discrete profiling method is described in further detail later in this section.

In Run 1 [80], the background fitting started from 100 GeV. At the lower end of this fit range, below about 115 GeV, there is a notable turn-on in the mass distribution that

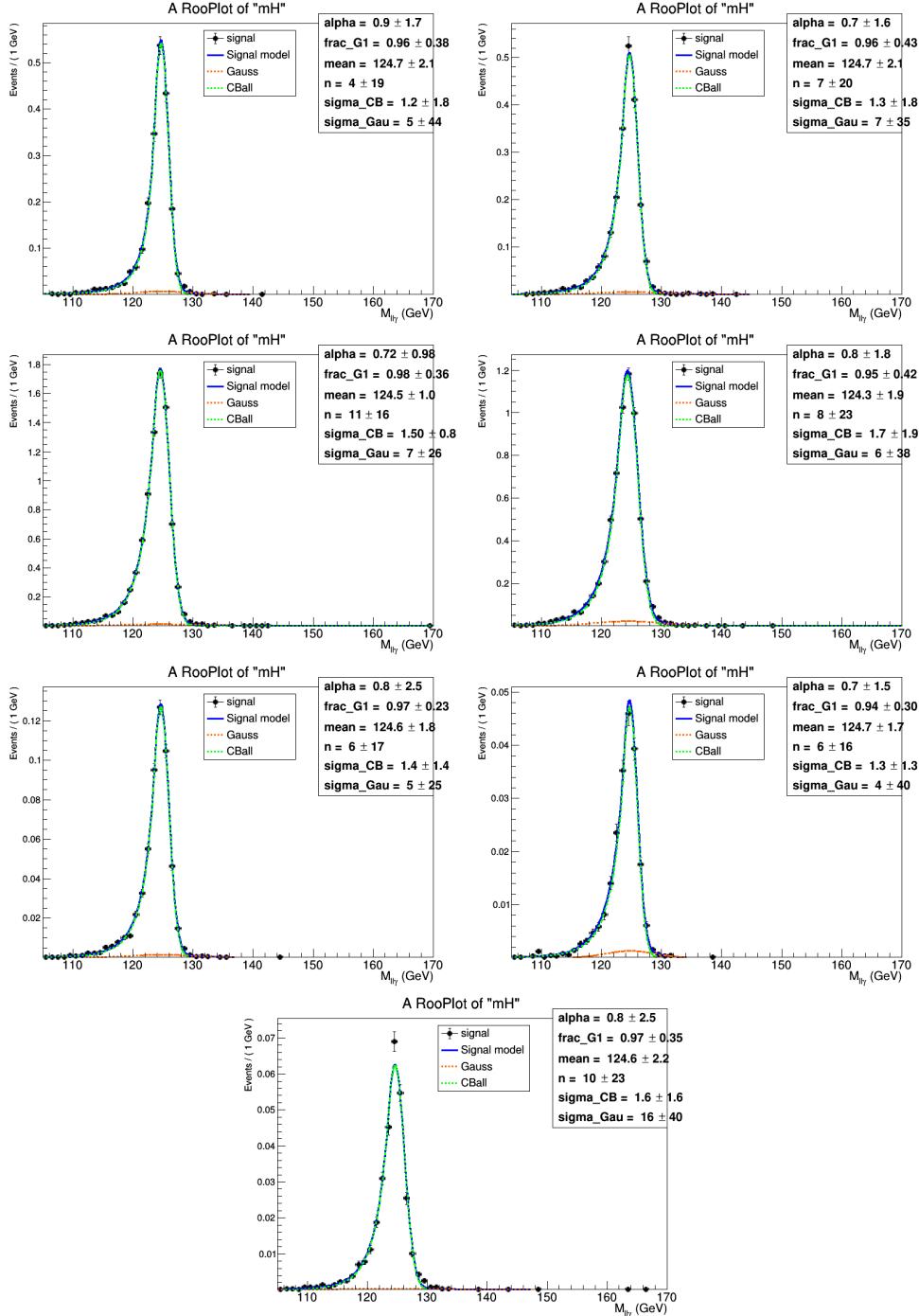


Figure 8.1. Fits to simulated  $m_{\ell^+ \ell^- \gamma}$  signal distributions in the electron channel for  $m_H = 125$  GeV for the 2016 data-taking period. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component. The top four plots correspond to the untagged categories, and the bottom three plots correspond to the dijet categories.

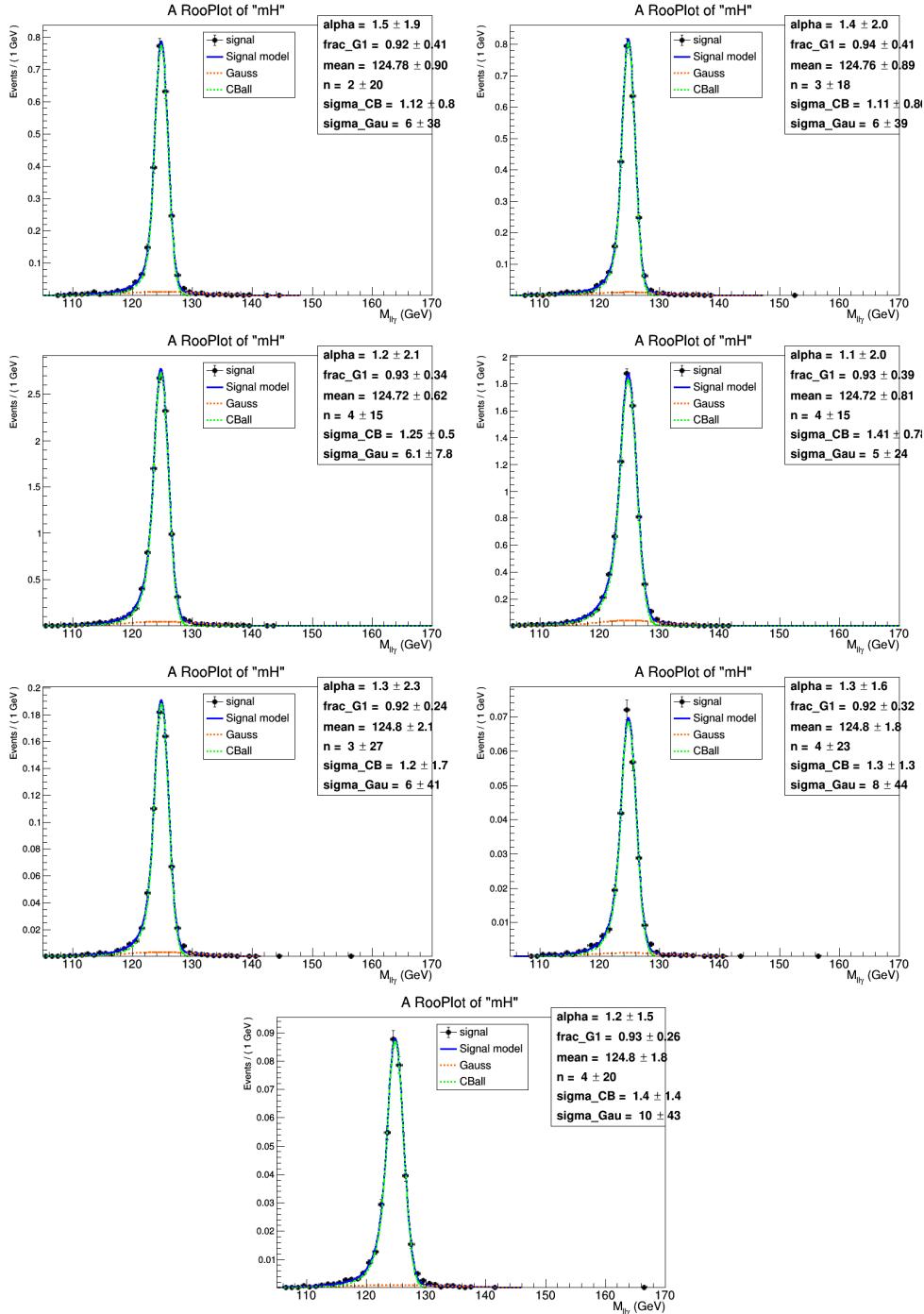


Figure 8.2. Fits to simulated  $m_{\ell^+\ell^-\gamma}$  signal distributions in the muon channel for  $m_H = 125$  GeV for the 2016 data-taking period. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component. The top four plots correspond to the untagged categories, and the bottom three plots correspond to the dijet categories.

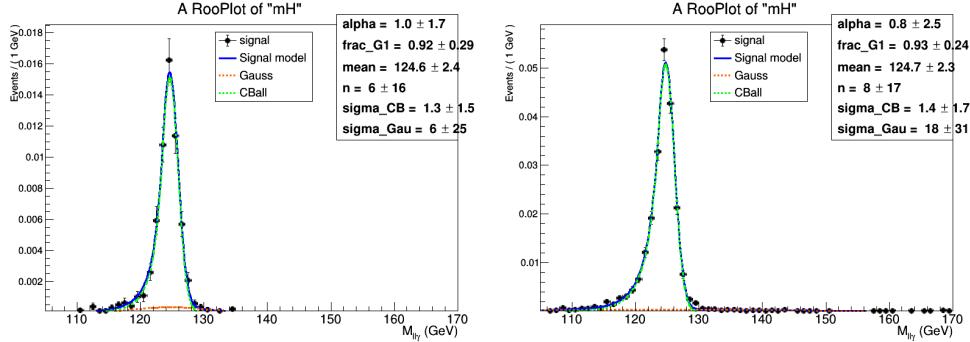


Figure 8.3. Fits to simulated  $m_{\ell^+\ell^-\gamma}$  signal distributions in the electron and muon channels combined in the lepton-tagged category for  $m_H = 125$  GeV for the 2016 data-taking period. The left plot shows the fit to simulated ZH production events, and the right plot shows the fit to simulated WH production events. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component.

arises due to the presence of the on-shell Z boson. As such, the Run 1 analysis used PDFs composed of a Gaussian turn-on component convoluted with a falling spectrum component. In the first Run 2 analysis using only the 2016 data, this procedure was significantly changed. The lower bound of the fit range was updated to 115 GeV in order to avoid the turn-on. This meant that the functional forms used to model the background could be simplified, as they no longer required the turn-on component. Now, in the full Run 2 analysis with 2016, 2017, and 2018 data, we have returned to the Run 1 approach of fitting the turn-on. The reason for this is that the turn-on can have an impact on the mass distribution, even above 115 GeV. As a consequence, avoiding the turn-on component of the fit leads to some mismodeling in the low mass region. Moreover, this mismodeling can be enough in some cases to lead to significant bias in the signal strength measurement. The solution has been to set the lower bound at 105 GeV such that the turn-on is roughly Gaussian and can be modeled and fit properly.

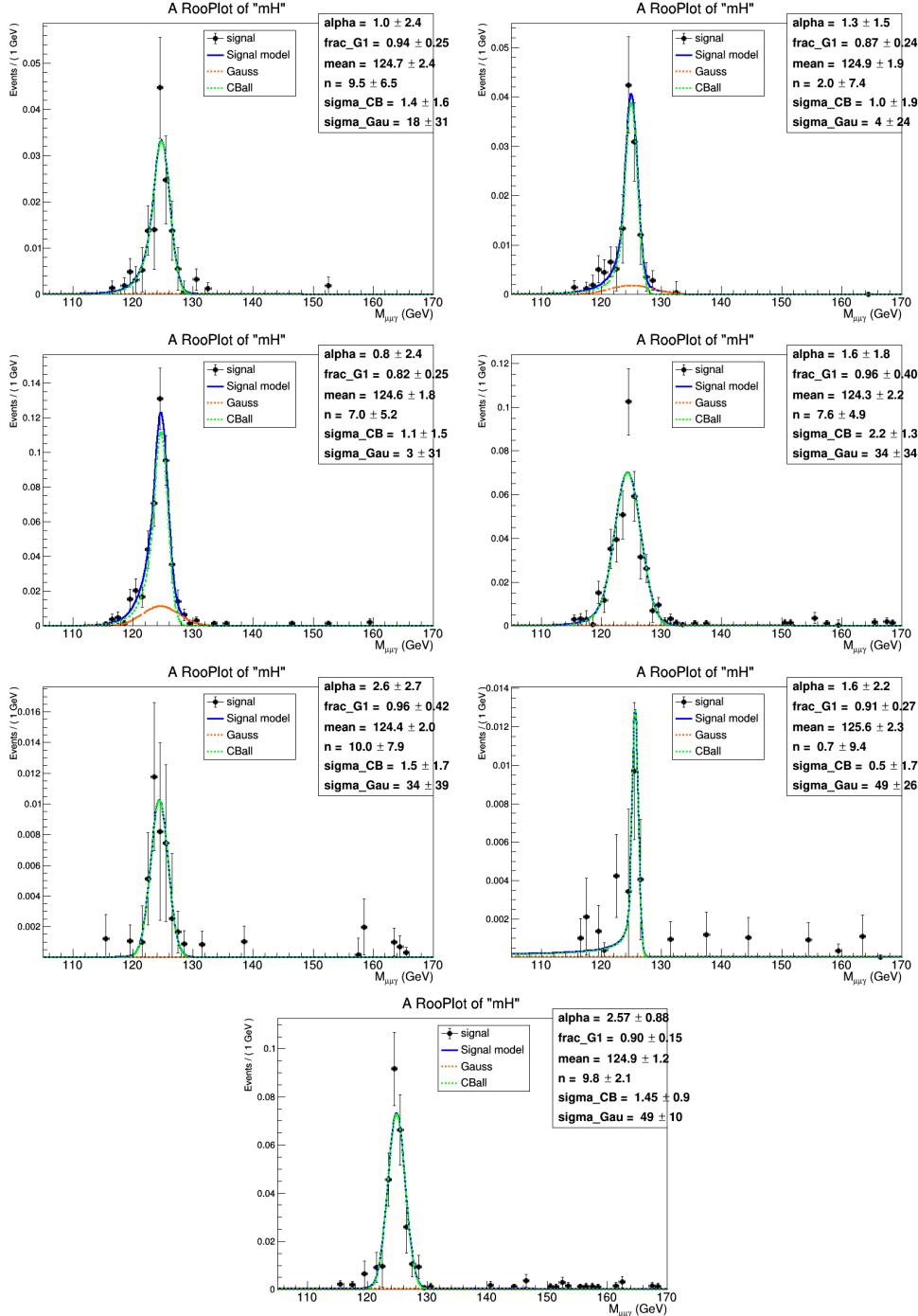


Figure 8.4. Fits to simulated  $m_{\mu^+\mu^-\gamma}$  resonant background distributions from  $H \rightarrow \mu^+\mu^-$  for  $m_H = 125$  GeV for the 2016 data-taking period. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component. The top four plots correspond to the untagged categories, and the bottom three plots correspond to the dijet categories.

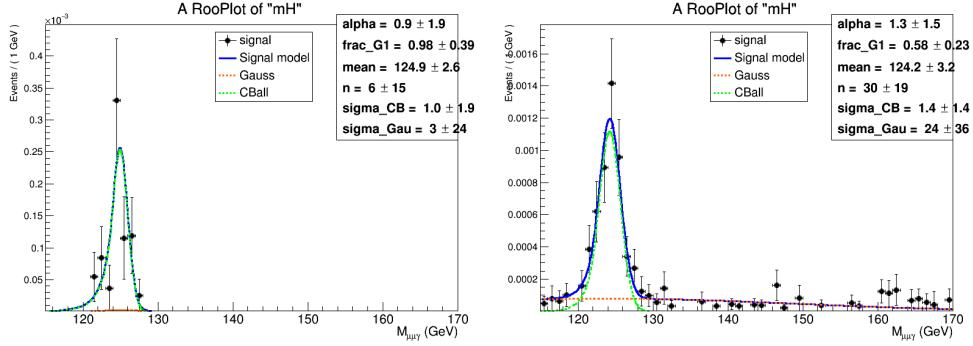


Figure 8.5. Fits to simulated  $m_{\ell^+\ell^-\gamma}$  resonant background distributions from  $H \rightarrow \mu^+\mu^-$  in the electron and muon channels combined in the lepton-tagged category for  $m_H = 125$  GeV for the 2016 data-taking period. The left plot shows the fit to simulated ZH production events, and the right plot shows the fit to simulated WH production events. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component.

### 8.3.1. General Form of PDFs

The  $m_{\ell^+\ell^-\gamma}$  spectrum consists of a turn-on peak around 110–115 GeV and a falling spectrum in the high-mass tail, where the turn-on peak is driven by the photon  $p_T$  selection. These features are modeled by the convolution of a Gaussian function with a step function multiplied by one of several falling spectrum functions. The complete function has the general form:

$$(8.1) \quad \mathcal{F}(m_{\ell^+\ell^-\gamma}; \mu_G, \sigma_G, s, \vec{\alpha}) = \int_{105}^{170} \mathcal{N}(m_{\ell^+\ell^-\gamma} - t; \mu_G, \sigma_G) \Theta(t; s) f(t; \vec{\alpha}) dt,$$

where  $t$  is the integration variable for the convolution,  $\mathcal{N}(m_{\ell^+\ell^-\gamma} - t; \mu_G, \sigma_G)$  is the Gaussian function with mean  $\mu_G$  and standard deviation  $\sigma_G$ ,  $\Theta(t; s)$  is the Heaviside step function with step location  $s$ , and  $f(t; \vec{\alpha})$  is the falling spectrum function with shape parameters  $\vec{\alpha}$ . The total number of parameters in the fit is equal to three plus the number of parameters in the falling spectrum component.

### 8.3.2. Falling Spectrum Component Families

The choice of falling spectrum component to model the background in a given category is not a priori known. For this reason, we consider several families of PDFs for the falling spectrum. The families tested include Bernstein polynomials, exponential series, power law series, and Laurent series. These families are described in more detail as follows:

- exponential series of order  $N$ :

$$(8.2) \quad \text{Exp}_N(m_{\ell\ell\gamma}) = \sum_{i=1}^N f_i e^{p_i m_{\ell\ell\gamma}}$$

with  $2N$  free parameters:  $p_i < 0$  and  $f_i$ . The lowest order considered has  $N = 1$ , i.e. one term. The next order has 2 exponential terms, but 3 parameters.

- power law series of order  $N$ :

$$(8.3) \quad \text{Pow}_N(m_{\ell\ell\gamma}) = \sum_{i=1}^N f_i m_{\ell\ell\gamma}^{p_i}$$

with  $2N$  free parameters:  $p_i < 0$  and  $f_i$ . The lowest order considered has  $N = 1$ , i.e. one term.

- Bernstein polynomial of order N:

$$(8.4) \quad \text{Bern}_N(m_{\ell\ell\gamma}) = \sum_{i=0}^N f_i b_{i,N}(m_{\ell\ell\gamma})$$

$$(8.5) \quad b_{i,N}(m_{\ell\ell\gamma}) = \binom{N}{i} m_{\ell\ell\gamma}^i (1 - m_{\ell\ell\gamma}^{N-i})$$

with  $N$  free parameters  $f_i$ . The lowest order considered has  $N = 1$ .

- Laurent series with 2, 3, 4, 5, or 6 terms, where  $N + 1$  equals the number of free parameters:

$$(8.6) \quad \text{Lau}_1(m_{\ell\ell\gamma}) = f_2 m_{\ell\ell\gamma}^{-4} + f_3 m_{\ell\ell\gamma}^{-5}$$

$$(8.7) \quad \text{Lau}_2(m_{\ell\ell\gamma}) = f_1 m_{\ell\ell\gamma}^{-3} + f_2 m_{\ell\ell\gamma}^{-4} + f_3 m_{\ell\ell\gamma}^{-5}$$

$$(8.8) \quad \text{Lau}_3(m_{\ell\ell\gamma}) = f_1 m_{\ell\ell\gamma}^{-3} + f_2 m_{\ell\ell\gamma}^{-4} + f_3 m_{\ell\ell\gamma}^{-5} + f_4 m_{\ell\ell\gamma}^{-6}$$

$$(8.9) \quad \text{Lau}_4(m_{\ell\ell\gamma}) = f_1 m_{\ell\ell\gamma}^{-2} + f_2 m_{\ell\ell\gamma}^{-3} + f_3 m_{\ell\ell\gamma}^{-4} + f_4 m_{\ell\ell\gamma}^{-5} + f_5 m_{\ell\ell\gamma}^{-6}$$

$$(8.10) \quad \text{Lau}_5(m_{\ell\ell\gamma}) = f_1 m_{\ell\ell\gamma}^{-2} + f_2 m_{\ell\ell\gamma}^{-3} + f_3 m_{\ell\ell\gamma}^{-4} + f_4 m_{\ell\ell\gamma}^{-5} + f_5 m_{\ell\ell\gamma}^{-6} + f_6 m_{\ell\ell\gamma}^{-7}$$

### 8.3.3. Discrete Profiling Method

In each category, there is ambiguity about which PDF to choose to model the background. In previous versions of this analysis, a detailed bias study was performed for each falling

spectrum component family, and the family with the least bias was chosen. However, no uncertainty was assigned due to the choice of background PDF. In principle, the choice of PDF is a source of systematic uncertainty in the measurement. Therefore, the discrete profiling method was since developed as a way to assess this uncertainty in a formal way. In this method, the choice of PDF for the background fit is included as a discrete nuisance parameter in the full likelihood function. All reasonable families of functions should be considered, and additionally, it is found that certain orders within the same family should be considered. The set of functions profiled in a given category is referred to as an *envelope*. The exact composition of the envelope in each category is based on a set of selection requirements that account for goodness of fit, an F-test procedure, and an assessment of bias. The details of this selection will be described in the following section. When fitting the background, all functions in the envelope are tried, with a penalty term added to the likelihood to account for the number of free parameters in the fit. When a measurement is made of a parameter of interest, for example the signal strength, the PDF with the smallest negative log likelihood, profiled as a function of the parameter of interest, is used. The profiling allows the best fit function to change for different values of the parameter of interest, so we can say that we carry out the fit with the full envelope of functions. The envelope yields a profile likelihood curve that is broader than the profile likelihood curve obtained from any individual function. This increase in breadth reflects the uncertainty associated with the choice of background PDF.

### 8.3.4. Envelope Selection

Individual functions are added to the envelope in a given category based on a series of selection requirements. In general, the requirements are related to three main factors: goodness of fit, F-test, and bias. The simplest requirement is a basic goodness of fit cut. For each function considered, the chi-squared per number of degrees of freedom is calculated. This is then converted into a p-value, or chi-squared probability. If the chi-squared probability is greater than 0.01, we consider the quality of the fit good enough for the envelope. Any function failing this requirement is thrown away, as it cannot be considered a reasonable description of the data.

The second consideration is the result of an F-test. The F-test is a way to compare functions within a specific family of falling spectrum component PDFs. It compares the fit of a lower order function in the family to that of a higher order function. While it is expected that the higher order function will always provide a better fit, the question is whether the higher order fit is better in a statistically significant way. If it is not significantly better, there is a motivation to stick to the lower order function and exclude the higher order function from the envelope. The details of the F-test procedure are as follows. First, the lowest order PDF in a given family is fit to the data in a given category. Then, the next highest order PDF is fit to the same data. The difference between twice the negative log likelihood between the two fits,  $2\Delta NLL_{N+1} = 2(NLL_N - NLL_{N+1})$ , can be used to determine whether the data are better supported by the higher order PDF. More precisely,  $2\Delta NLL_{N+1}$  is distributed as a  $\chi^2$  PDF with  $M$  degrees of freedom, where  $M$  is the difference in the number of free parameters between the order  $N+1$  PDF and the order  $N$  PDF. For example, in the case of the first two orders of the exponential family,

$M = 4 - 2 = 2$ , and for the Bernstein polynomial family,  $M = 3 - 2 = 1$ . After carrying out the fits and computing the negative log likelihoods, a p-value is then calculated as

$$(8.11) \quad p = P(2\Delta NLL > 2\Delta NLL_{N+1} | \chi^2(M)).$$

If the p-value is less than 0.05, we state that the higher order PDF is supported by the data, and the procedure is repeated for the next highest order PDF. Alternatively, if the p-value is greater than 0.05, the higher order PDF is assumed to be too flexible given the data, and the F-test ends, having found the highest order PDF supported by the data. This procedure is done to find the highest-order PDF supported by the data for each PDF family. In this analysis, we consider fits up to order 5 for each falling spectrum family. In each family, we identify the highest order function with an F-test probability below 0.05 and include it in the envelope. We then also include all lower order functions within the family that pass the chi-squared goodness of fit criterion.

The final consideration in selecting functions for the envelope is bias and coverage. After constructing an initial envelope in a given category based on the chi-squared and F-test results, an envelope bias study is performed. The full details of the envelope bias studies for each category can be reviewed in Section 8.3.6. In a nutshell, toys are generated from each PDF contained in the envelope, and these toys are subsequently fit with the full envelope. The collection of toy fits allows us to compute the bias and coverage of the signal strength. Should the bias and coverage be at an acceptable level, the envelope can be used in the analysis. However, in the presence of large bias or low coverage, we can consider modifying the composition of the envelope in order to reduce the bias. In practice, this is often achievable by adding higher order functions that may have been excluded by the

F-test, but nevertheless fit the data well. These more flexible functions are often able to reduce the bias and improve the coverage of the envelope. This is considered category by category. The fits of the chosen envelope functions in each category are shown in Figure 8.6.

### 8.3.5. Comparison with Previous Approach

Previously, in the 2016 analysis and in earlier versions of the full Run 2 analysis, the background was modeled on the three body mass range 115 to 170 GeV. After extending the range down to 105 GeV and studying the fits in each category, it is worth checking how the new modeling compares to the old. One simple way to compare is by overlaying the fit plots using the best fit function in each range. This is shown in Figure 8.7. From the plots, it is clear that modeling the turn-on properly yields a better fit around 115 GeV, the previous lower bound of the fit.

### 8.3.6. Envelope Bias Studies

Upon the selection of an initial envelope of PDFs to model the background in a given category of the analysis, a bias study is performed for that category. The principal goal of the bias study is to measure the bias and coverage of the full model, the envelope of functions, for different choices of the true underlying PDF. Of course, we do not know the true underlying PDF *a priori*, so we rely on the chi-squared and F-test probabilities used to construct the envelope to identify likely truth functions. In practice, this means that each individual PDF in the envelope is considered separately as a possible truth function in our bias studies. There are several possible outcomes of the bias studies in

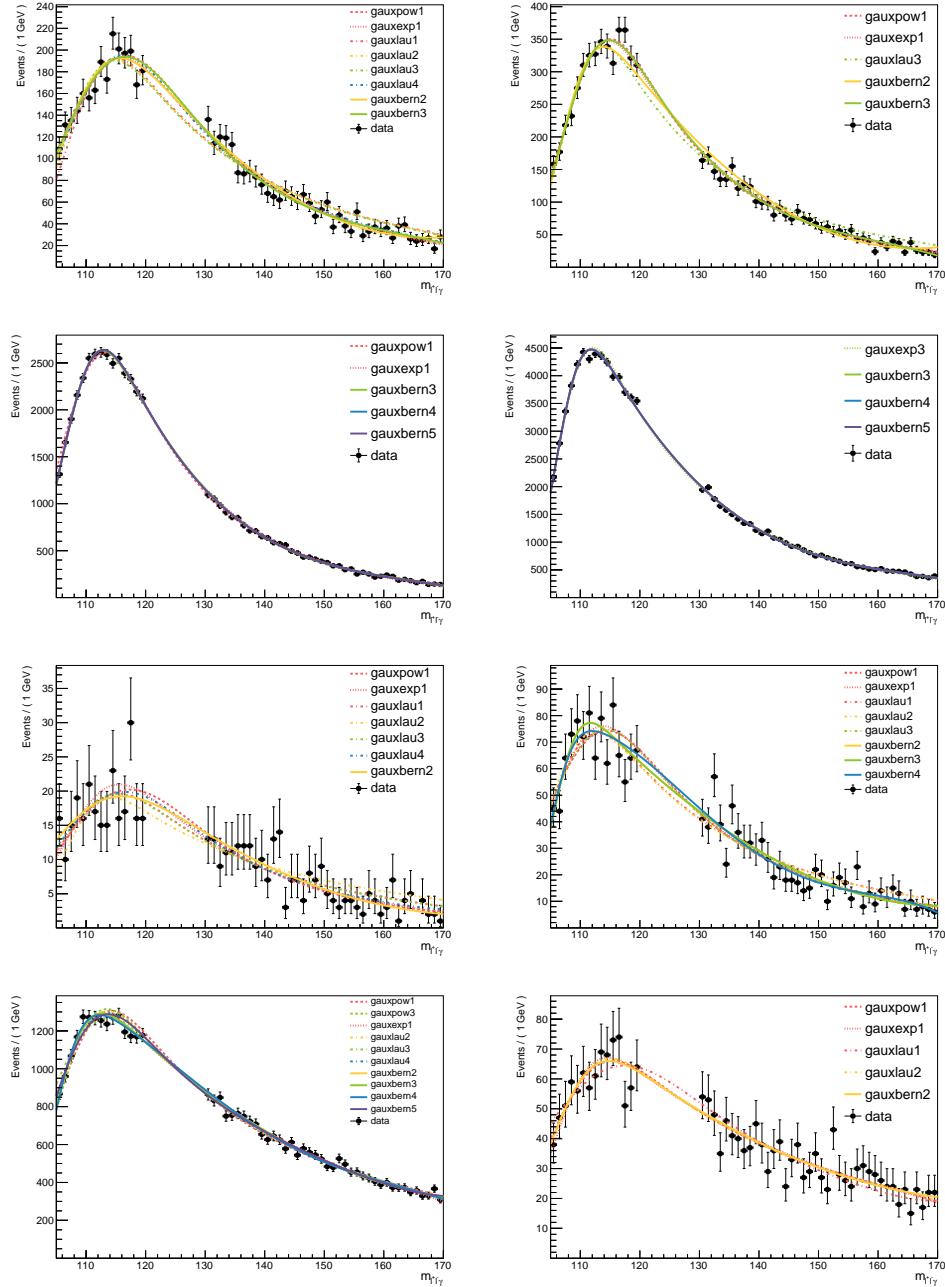


Figure 8.6. Background functions determined by the discrete profiling method. The top four plots correspond to the untagged categories, and the bottom four plots correspond to the dijet categories and lepton tag category.

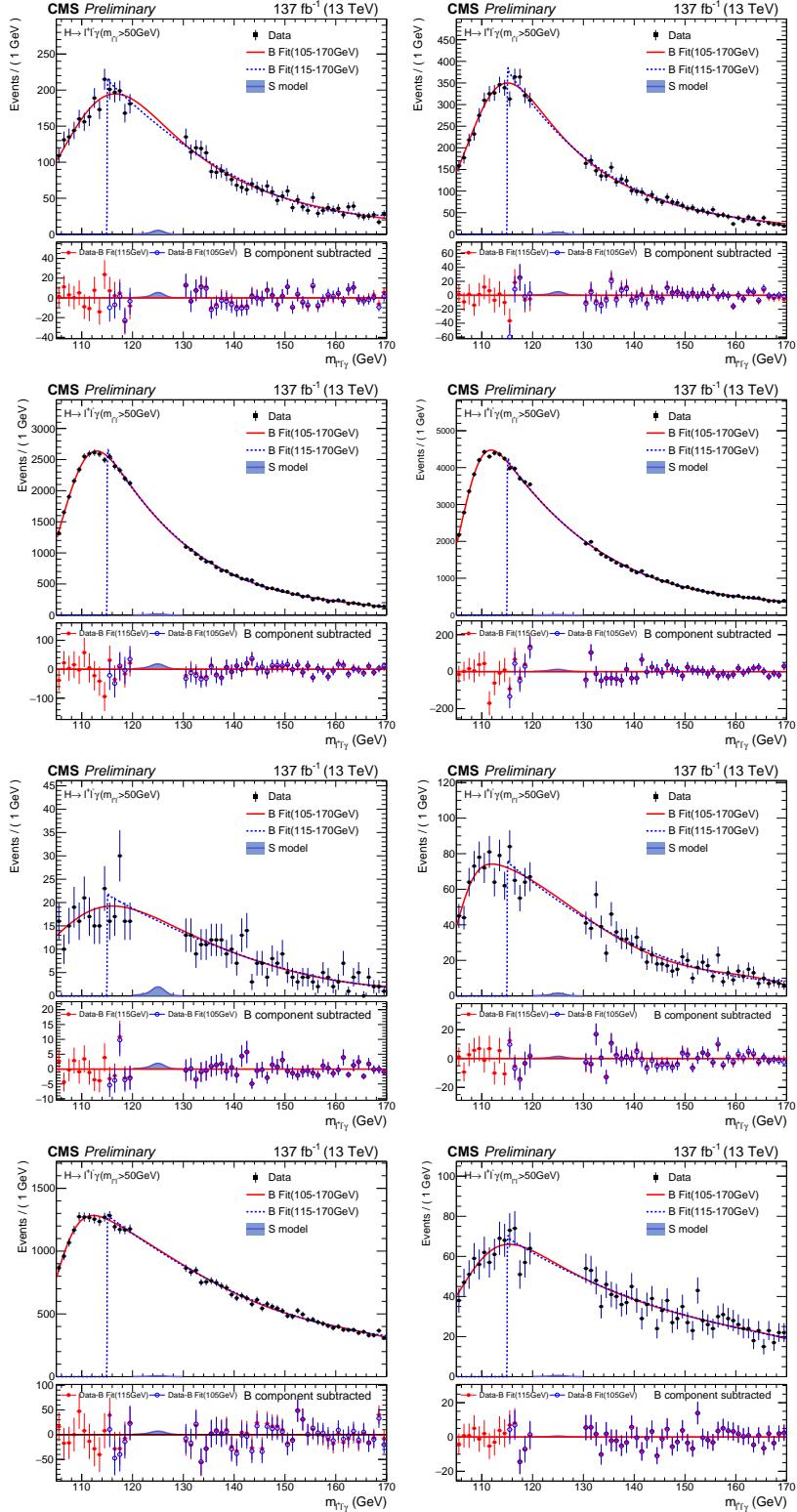


Figure 8.7. Comparison of background fits with and without modeling the turn-on. The top four plots correspond to the untagged categories, and the bottom four plots correspond to the dijet categories and lepton tag category.

each category. One outcome is that the bias is minimal and the coverage good for all truth functions. In this case, we accept the composition of the envelope as good for the analysis. Another possible outcome is that one or more truth functions is associated with a large (significantly greater than 14%) bias or low (significantly less than 68%) coverage. In this case, we can consider adding more functions to the envelope to increase the flexibility of the model, and then retry the bias study using the extended envelope.

A more precise description of the bias study procedure follows. First, we generate a set of 1,000 toy datasets for a given truth function. The parameters of the truth function are fixed by fitting the data. Then, this PDF is used to generate the toys, with the same level of statistical uncertainty as the data and assuming a true signal strength of zero (background only). Next, each toy is fit using the full envelope, where the index of the best fit function is profiled with the signal strength. In general, this means that different toys may be fit by different constituent PDFs within the envelope, and these PDFs may be different from the generating PDF. For each toy fit, we obtain a best fit signal strength  $\hat{r}$  and a 68% confidence interval  $(\hat{r}_{down}, \hat{r}_{up})$ . We define the pull for each toy as  $pull = (\hat{r} - r_{true})/\sigma_{\hat{r}}$ . In the background only case,  $r_{true} = 0$ . The uncertainty in the denominator accounts for the assymetric error bars coming from the confidence interval. The pull distribution is expected to be a unit Gaussian. We fit the pull distribution with a Gaussian, and the best fit mean value is taken to be the bias. The coverage is defined as the frequency with which zero falls within the confidence interval  $(\hat{r}_{down}, \hat{r}_{up})$ . It is expected to be near 68%. If overcoverage is observed, this implies the uncertainty in the signal strength may be overly conservative. Within a few percent of 68%, this is acceptable. If undercoverage of a significant amount is observed, this implies a problem with the model,

and may mean that the individual functions are biased or that the envelope should be made more flexible.

The bias and coverage results for the final envelope in each category are shown in Tables 8.1a through 8.1h. Note that in all categories except untagged 3, untagged 4, and dijet 3, the envelope functions are exactly those identified by the chi-squared and F-test probabilities previously described. However, in untagged 3, untagged 4, and dijet 3, it was observed that the initial functions chosen by the chi-squared and F-test procedure showed large biases (greater than 20%) and low coverages (less than 60%) for certain truth functions. Therefore, the envelope was extended to include higher order Bernstein functions up to order 5 for each of these categories. The bias tables for these categories reflect the results for the final extended envelope.

### 8.3.7. Interpretation of Bias Study Results

We now turn to the interpretation of the initial envelope bias studies. First, it can be seen that most categories show no signs of significant bias or undercoverage for any potential truth functions. Specifically, we see that the envelopes for categories untagged 1, untagged 4, and dijet 1 can immediately be accepted as good for the analysis. In untagged 2, untagged 3, dijet 2, dijet 3, and the lepton tag category, specific truth functions deserve a bit more scrutiny.

In the untagged 2, dijet 2, dijet 3, and lepton tag categories, we observe slightly larger bias for the Laurent polynomial truth functions. From the heat map of the chosen functions, we can see the toys generated with Laurent functions are most often fit with Bernstein polynomials. Additionally, the Laurent functions give lower probabilities than

Table 8.1. Bias and coverage results for each category.

Generating PDF	Exp1	Pow1	Lau1	Lau2	Lau3	Lau4	Bern2	Bern3
<b>Bias</b>	-0.03	0.02	0.02	0.06	0.04	-0.09	-0.07	0.00
<b>Coverage</b>	$\pm 0.03$	$\pm 0.04$	$\pm 0.03$	$\pm 0.04$				
	0.68	0.66	0.71	0.68	0.71	0.68	0.65	0.70

(a) Untagged 1 bias and coverage results from the final envelope bias study.

Generating PDF	Exp1	Pow1	Lau3	Bern2	Bern3
<b>Bias</b>	-0.10	-0.09	-0.27	-0.07	0.07
<b>Coverage</b>	$\pm 0.04$	$\pm 0.05$	$\pm 0.04$	$\pm 0.03$	$\pm 0.04$
	0.67	0.66	0.67	0.67	0.65

(b) Untagged 2 bias and coverage results from the final envelope bias study.

Generating PDF	Exp1	Pow1	Bern3	Bern4	Bern5
<b>Bias</b>	-0.05	-0.11	0.03	-0.02	0.27
<b>Coverage</b>	$\pm 0.04$	$\pm 0.03$	$\pm 0.04$	$\pm 0.03$	$\pm 0.04$
	0.67	0.67	0.70	0.67	0.66

(c) Untagged 3 bias and coverage results from the final envelope bias study.

Generating PDF	Exp3	Bern3	Bern4	Bern5
<b>Bias</b>	-0.08	0.07	0.16	0.10
<b>Coverage</b>	$\pm 0.03$	$\pm 0.03$	$\pm 0.03$	$\pm 0.03$
	0.68	0.67	0.67	0.68

(d) Untagged 4 bias and coverage results from the final envelope bias study.

Generating PDF	Exp1	Pow1	Lau1	Lau2	Lau3	Lau4	Bern2
<b>Bias</b>	0.03	0.09	-0.02	-0.06	-0.02	0.02	-0.04
<b>Coverage</b>	$\pm 0.03$	$\pm 0.03$	$\pm 0.04$	$\pm 0.04$	$\pm 0.03$	$\pm 0.03$	$\pm 0.03$
	0.64	0.66	0.66	0.68	0.67	0.68	0.66

(e) Dijet 1 bias and coverage results from the final envelope bias study.

Generating PDF	Exp1	Pow1	Pow3	Lau2	Lau3	Lau4	Bern2	Bern3	Bern4	Bern5
<b>Bias</b>	-0.12	-0.29	-0.14	-0.22	-0.16	-0.18	-0.06	-0.05	0.05	-0.13
<b>Coverage</b>	$\pm 0.04$	$\pm 0.03$	$\pm 0.03$	$\pm 0.03$						
	0.67	0.65	0.68	0.66	0.66	0.67	0.68	0.64	0.67	0.65

(f) Dijet 2 bias and coverage results from the final envelope bias study.

Generating PDF	Exp1	Pow1	Pow3	Lau2	Lau3	Lau4	Bern2	Bern3	Bern4	Bern5
<b>Bias</b>	-0.12	-0.29	-0.14	-0.22	-0.16	-0.18	-0.06	-0.03	0.05	-0.13
<b>Coverage</b>	$\pm 0.03$	$\pm 0.04$	$\pm 0.04$	$\pm 0.04$	$\pm 0.03$	$\pm 0.03$	$\pm 0.03$	$\pm 0.05$	$\pm 0.03$	$\pm 0.03$
	0.67	0.65	0.68	0.67	0.66	0.67	0.68	0.64	0.67	0.65

(g) Dijet 3 bias and coverage results from the final envelope bias study.

Generating PDF	Exp1	Pow1	Lau1	Lau2	Bern2
<b>Bias</b>	-0.04	0.05	0.27	0.03	-0.03
<b>Coverage</b>	$\pm 0.03$	$\pm 0.03$	$\pm 0.04$	$\pm 0.03$	$\pm 0.03$
	0.71	0.69	0.66	0.70	0.70

(h) Lepton tag bias and coverage results from the final envelope bias study.

the other functions in the envelope. This means that the Laurent family is irrevelant in the final background model. Given this consideration, along with the fairly good coverage for these Laurent functions, we expect the impact of these biases on the analysis to be negligible. A similar line of logic also applies to the case of the power law 1 truth function in categories dijet 2 and dijet 3. Again, we expect the impact of these biases to be negligible.

In untagged 3, we observe a small, but somewhat significant bias of 27% for the truth function with falling spectrum component Bernstein 5. In this case, we see that the coverage of 66% is still quite good, only slightly below 68%. Looking at the fit frequency heat map for this category, we see that this bias arises when the fit function is a lower order Bernstein polynomial. This type of behavior is expected. Generating with a higher order and fitting with a lower order within the same function family will usually yield some level of bias. This is less of a concern than a case in which one truth family shows a large bias when fit with a different family. Given the small size of the bias for a single truth function, and the fact that the behavior is somewhat expected for lower order fits within the Bernstein family, we expect the overall impact on the analysis to be negligible. Therefore, the untagged 3 envelope is also deemed fine for the analysis.

## CHAPTER 9

**Systematic Uncertainties**

The uncertainties associated with the choice of background shape are incorporated into the fit to the data through the use of the discrete profiling method. They are, therefore, reflected in the statistical uncertainties obtained from the fit. The systematic uncertainties, affecting either the normalization or the shape of the signal expectation, are listed below, and the numerical values are summarized in Table 9.1, which also indicates whether the effect is correlated between the data-taking periods.

- Theoretical cross section calculations: These include the effects of the choice of PDFs, the value of the strong coupling constant ( $\alpha_S$ ), and the effect of missing higher orders in the perturbative cross section calculations, evaluated from variations of the renormalization and factorization scales ( $\mu_R$ ,  $\mu_F$ ) [81, 82, 83]. The uncertainties are treated as independent for each Higgs boson production mechanism. The uncertainty in  $\mathcal{B}(H \rightarrow Z\gamma)$  is also considered [12].
- Underlying event and parton shower modeling: The uncertainty associated with the choice and tuning of the generator is estimated with dedicated samples which are generated by varying the parameters of the tune used to generate the original signal samples. The uncertainties are treated as correlated for the 2017 and 2018 samples, which use the CP5 tune [72], while being uncorrelated with the 2016 sample, which uses the CUETP8M1 tune [71].

- Integrated luminosity: The integrated luminosities for the 2016, 2017, and 2018 data-taking years have uncertainties of 1.2%, 2.3% and 2.5% [59, 60, 61], respectively, with an overall uncertainty for the 2016–2018 period of 1.6%, the improvement in precision reflecting the (uncorrelated) time evolution of some systematic effects.
- L1 trigger: During the 2016 and 2017 data-taking periods, a gradual shift in the timing of the inputs of the ECAL L1 trigger in the  $|\eta| > 2.4$  region led to a specific inefficiency. A correction of approximately 1% is applied to the simulation along with the corresponding uncertainty in the inefficiency measurement.
- Trigger: There are uncertainties in the corrections applied to the simulation to match the trigger efficiencies measured in data with  $Z \rightarrow e^+e^-$  and  $Z \rightarrow \mu^+\mu^-$  events.
- Photon identification and isolation: There are uncertainties in the corrections applied to the simulation to match the selection efficiencies in data measured with  $Z \rightarrow e^+e^-$  events.
- Lepton identification and isolation: There are uncertainties in the corrections applied to the simulation to match electron and muon selection efficiencies in data measured with  $Z \rightarrow e^+e^-$  and  $Z \rightarrow \mu^+\mu^-$  events.
- Pileup modeling: The uncertainty in the description of the pileup in the signal simulation is estimated by varying the total inelastic cross section by  $\pm 4.6\%$  [84].
- Kinematic BDT: The uncertainties in the photon and lepton energy and the correction of the photon MVA discriminant are propagated to  $\mathcal{D}_{\text{kin}}$ . Changes in  $\mathcal{D}_{\text{kin}}$  cause the migration of signal events across category boundaries.

- VBF BDT: The uncertainties in the jet energy and the uncertainty in  $\mathcal{D}_{\text{kin}}$  are propagated to  $\mathcal{D}_{\text{VBF}}$ . Changes in  $\mathcal{D}_{\text{VBF}}$  cause the migration of signal events across category boundaries.
- Photon energy scale and resolution: The photon energy in the simulation is varied due to the ECAL energy scale and resolution uncertainties, and the effects on the signal mean and resolution parameters are propagated to the fits.
- Lepton momentum scale and resolution: The lepton momentum in the simulation is varied due to the lepton momentum scale and resolution uncertainties, and the effects on signal mean and resolution parameters are propagated to the fits.

In the  $\mathcal{B}(\text{H} \rightarrow \text{Z}\gamma)/\mathcal{B}(\text{H} \rightarrow \gamma\gamma)$  measurement, the common sources of theoretical and systematic uncertainty in the two analyses are treated as correlated in the fit. These are the theoretical uncertainties in the Higgs production cross section calculations, and the systematic uncertainties in the underlying event and parton shower modeling, the integrated luminosity, and the L1 trigger inefficiency. The remaining uncertainties are treated as uncorrelated.

Table 9.1. Sources of systematic uncertainty affecting the simulated signal. The normalization effect on the expected yield, or the effect on the signal shape parameters, is given as indicated, with the values averaged over all event categories. The third column shows the uncertainties that have a correlated effect across the three data-taking periods.

Sources	Uncertainty (%)	Year-to-year correlation
<b>Normalization</b>		
Theoretical	—	✓
– $\mathcal{B}(H \rightarrow Z\gamma)$	5.7	
– ggH cross section ( $\mu_F, \mu_R$ )	3.9	
– ggH cross section ( $\alpha_S$ )	2.6	
– ggH cross section (PDF)	1.9	
– VBF cross section ( $\mu_F, \mu_R$ )	0.4	
– VBF cross section ( $\alpha_S$ )	0.5	
– VBF cross section (PDF)	2.1	
– WH cross section ( $\mu_F, \mu_R$ )	$+0.6$ $-0.7$	
– WH cross section (PDF)	1.7	
– ZH cross section ( $\mu_F, \mu_R$ )	$+3.8$ $-3.1$	
– ZH cross section (PDF)	1.3	
– WH/ZH cross section ( $\alpha_s$ )	0.9	
– $t\bar{t}H$ cross section ( $\mu_F, \mu_R$ )	$+5.8$ $-9.2$	
– $t\bar{t}H$ cross section ( $\alpha_s$ )	2.0	
– $t\bar{t}H$ cross section (PDF)	3.0	
Underlying event and parton shower	3.7–4.4	Partial
Integrated luminosity	1.2–2.5	Partial
L1 trigger	0.1–0.4	—
Trigger	—	—
– Electron channel	0.9–1.9	
– Muon channel	0.1–0.4	
Photon identification and isolation	0.2–5.0	✓
Lepton identification and isolation	—	✓
– Electron channel	0.5–0.7	
– Muon channel	0.3–0.4	
Pileup	0.4–1.0	✓
Kinematic BDT	2.5–3.7	✓
VBF BDT	5.9–14	✓
<b>Shape</b>		
Photon energy and momentum		✓
– Signal mean	0.1–0.4	
– Signal resolution	3.1–5.9	
Lepton energy and momentum		✓
– Signal mean	0.007	
– Signal resolution	0.007–0.01	

## CHAPTER 10

### Statistical Analysis

The signal search is performed using a simultaneous fit to the  $m_{\ell^+\ell^-\gamma}$  distribution in the eight event categories. The expected SM  $H \rightarrow Z\gamma$  distributions, scaled by a factor of 10, are also shown. The fit uses a binned maximum likelihood method in the range  $105 < m_{\ell^+\ell^-\gamma} < 170$  GeV. In each category, a likelihood function is defined using analytic models of signal and background events, along with nuisance parameters for systematic uncertainties. The combined likelihood function is the product of the likelihood functions in each category. The parameter of interest in the maximum likelihood fit is the signal strength  $\mu$ , defined as the product of the cross section and the branching fraction [ $\sigma(pp \rightarrow H)\mathcal{B}(H \rightarrow Z\gamma)$ ], relative to the SM expectation.

The signal model is defined as the sum of Crystal Ball [78] and Gaussian functions. The signal shape parameters are determined by fitting this model to simulated signal events in each category. To account for differences in mass resolution, these fits are performed separately for the event samples used to model each data-taking year, as well as for muon and electron channel events. This results in six signal models that are summed to give the total signal expectation in a given category. Table 7.2 gives these mass resolutions for  $H \rightarrow Z\gamma$ , summed over the three years, as obtained from simulation. The mass resolutions range from 1.4–2.3 GeV, depending on the category. Separate sets of parameter values are found by fitting simulated events with  $m_H$  of 120, 125, and 130 GeV. Using linear interpolation, parameter values are also determined at 1 GeV intervals in  $m_H$  from

120–130 GeV, as well as at 125.38 GeV. In the fit to data, the mean and resolution parameters are allowed to vary subject to constraints from several systematic uncertainties, described in Section 9, while the remaining parameters are held fixed.

The background model in each category is obtained from the data using the discrete profiling method [79]. This technique accounts for the systematic uncertainty associated with choosing an analytic functional form to fit the background. The background function is chosen from a set of candidate functions via a discrete nuisance parameter in the fit. These functions are derived from the data in each category, with muon and electron events from all data-taking years combined. The  $m_{\ell^+\ell^-\gamma}$  spectrum consists of a turn-on peak around 110–115 GeV and a falling spectrum in the high-mass tail, where the turn-on peak is driven by the photon  $p_T$  selection. These features are modeled by the convolution of a Gaussian function with a step function multiplied by one of several falling spectrum functions. The complete function has the general form:

$$(10.1) \quad \mathcal{F}(m_{\ell^+\ell^-\gamma}; \mu_G, \sigma_G, s, \vec{\alpha}) = \int_{105}^{170} \mathcal{N}(m_{\ell^+\ell^-\gamma} - t; \mu_G, \sigma_G) \Theta(t; s) f(t; \vec{\alpha}) dt,$$

where  $t$  is the integration variable for the convolution,  $\mathcal{N}(m_{\ell^+\ell^-\gamma} - t; \mu_G, \sigma_G)$  is the Gaussian function with mean  $\mu_G$  and standard deviation  $\sigma_G$ ,  $\Theta(t; s)$  is the Heaviside step function with step location  $s$ , and  $f(t; \vec{\alpha})$  is the falling spectrum function with shape parameters  $\vec{\alpha}$ . The falling spectrum function families considered include exponential functions, power law functions, Laurent series, and Bernstein polynomials. Functions from each family are selected based on a chi-squared goodness-of-fit criterion ( $p$ -value > 0.01) as well as an  $\mathcal{F}$ -test [85], which determines the highest order function to be used. A penalty term is added to the final likelihood to take into account the number of parameters in

each function, ensuring that higher-order functions will not be preferred a priori. The set of profiled background functions in each category is checked to ensure that any bias introduced into the fit results is small and that the associated CL intervals have the appropriate frequentist coverage. For each function, pseudo-data sets are generated under a fixed signal strength hypothesis. A signal-plus-background fit is performed on each pseudo-data set, with the choice of background function profiled, and the distribution of best fit signal strengths is used to assess the bias and frequentist coverage.

The best fit value of the signal strength,  $\hat{\mu}$ , is determined by maximizing the likelihood, accounting for all nuisance parameters. The uncertainty in  $\hat{\mu}$  and the observed significance are derived from the profile likelihood test statistic [86],

$$(10.2) \quad q(\mu) = -2\ln\left(\frac{\mathcal{L}(\mu, \hat{\vec{\theta}}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\vec{\theta}})}\right),$$

where  $\vec{\theta}$  is the set of nuisance parameters,  $\hat{\mu}$  and  $\hat{\vec{\theta}}$  are unconditional best fit values, and  $\hat{\vec{\theta}}_\mu$  is the set of conditional best fit values of the nuisance parameters for a given value of  $\mu$ . An upper limit on  $\mu$  is determined using the profile likelihood statistic with the CL<sub>s</sub> criterion. The asymptotic approximation for the sampling distribution of  $q(\mu)$  is assumed in the derivation of these results [87, 88, 86, 89]. The expected significance under the SM hypothesis and the expected upper limits under the background-only hypothesis are also reported. These are obtained by fitting to the corresponding Asimov data sets [89].

In addition, a combined maximum likelihood fit with the CMS measurement [13] of  $H \rightarrow \gamma\gamma$  using the same data sample is performed to determine the ratio  $\mathcal{B}(H \rightarrow Z\gamma)/\mathcal{B}(H \rightarrow \gamma\gamma)$ . The  $H \rightarrow \gamma\gamma$  analysis obtained a signal strength for  $\sigma(pp \rightarrow H)\mathcal{B}(H \rightarrow$

$\gamma\gamma$ ) of  $1.12 \pm 0.09$ . In this combined fit, the branching fraction  $\mathcal{B}(H \rightarrow \gamma\gamma)$  is an additional free parameter. The uncertainty in the measured ratio of the two branching fractions is dominated by the statistical uncertainties in the branching fractions. Common sources of theoretical and experimental uncertainty in the two measurements, described in the next section, are treated as correlated in the fit. The combination is performed at  $m_H = 125.38\text{ GeV}$ , and the discrete profiling method is used for the background modeling in both cases.

## CHAPTER 11

### Results and Interpretation

Figures 11.1 and 11.2 show the  $m_{\ell^+\ell^-\gamma}$  distributions of the data events in each category. The expected SM  $H \rightarrow Z\gamma$  distributions, scaled by a factor of 10, are also shown. Figure 11.3 shows the signal-plus-background fit to the data and the corresponding distribution after background subtraction for the sum of all categories. Each category is weighted by the factor  $S/(S + B)$ , where  $S$  is the signal yield and  $B$  is the background yield in the narrowest mass interval containing 95% of the signal distribution.

The best fit value of the signal strength is  $2.4^{+0.8}_{-0.9}$  (stat)  $^{+0.3}_{-0.2}$  (syst) at  $m_H = 125.38$  GeV. The corresponding measured value of  $\sigma(pp \rightarrow H)\mathcal{B}(H \rightarrow Z\gamma)$  is  $0.21^{+0.07}_{-0.08}$  (stat)  $^{+0.03}_{-0.02}$  (syst) pb. This measurement is consistent with the SM prediction of  $0.09 \pm 0.01$  pb at the 1.6 standard deviation level. Figure 11.4 shows the signal strengths obtained for each category separately, corresponding to the fit results shown in Figs. 11.1 and 11.2, as well as from simultaneous fits to the dijet categories, the untagged categories, and all categories combined. Among the eight categories, dijet 1 is the most sensitive. A category compatibility  $p$ -value, under the hypothesis of a common signal strength in all categories, is calculated from the likelihood ratio between the nominal combined fit, in which all categories have the same signal strength parameter, and a separate fit, in which each category has its own signal strength parameter. This  $p$ -value is found to be 0.02, corresponding to 2.3 standard deviations, and is driven by the dijet 3 category, which has a signal strength of  $\hat{\mu} = 12.3^{+3.7}_{-3.5}$ . The observed (expected) local significance is 2.7 (1.2) standard

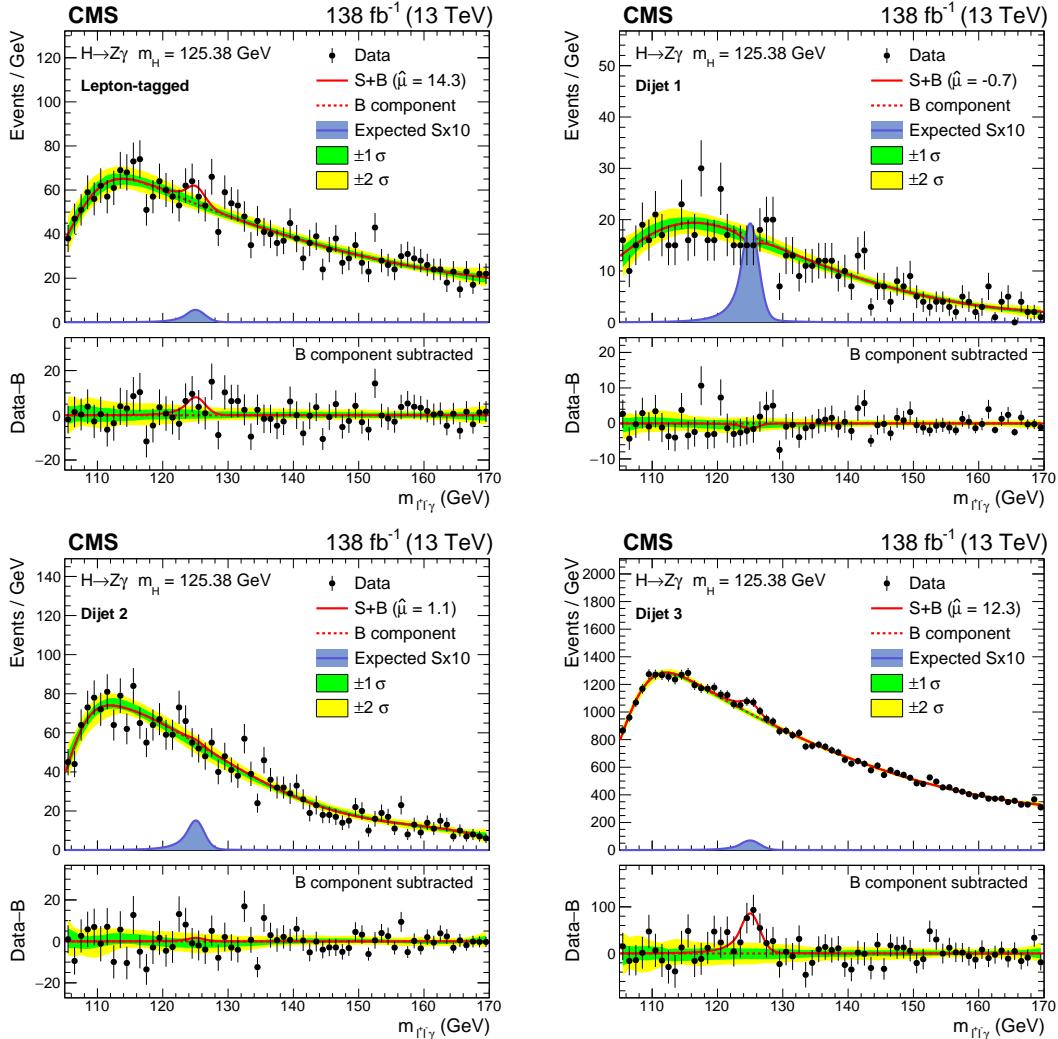


Figure 11.1. Fits to the  $m_{\ell^+\ell^-\gamma}$  data distribution in the lepton-tagged (upper left), dijet 1 (upper right), dijet 2 (lower left), and dijet 3 (lower right) categories. In the upper panel, the red solid line shows the result of a signal-plus-background fit to the given category. The red dashed line shows the background component of the fit. The green and yellow bands represent the 68 and 95% CL uncertainties in the fit. Also plotted is the expected SM signal, scaled by a factor of 10. In the lower panel, the data minus the background component of the fit is shown.

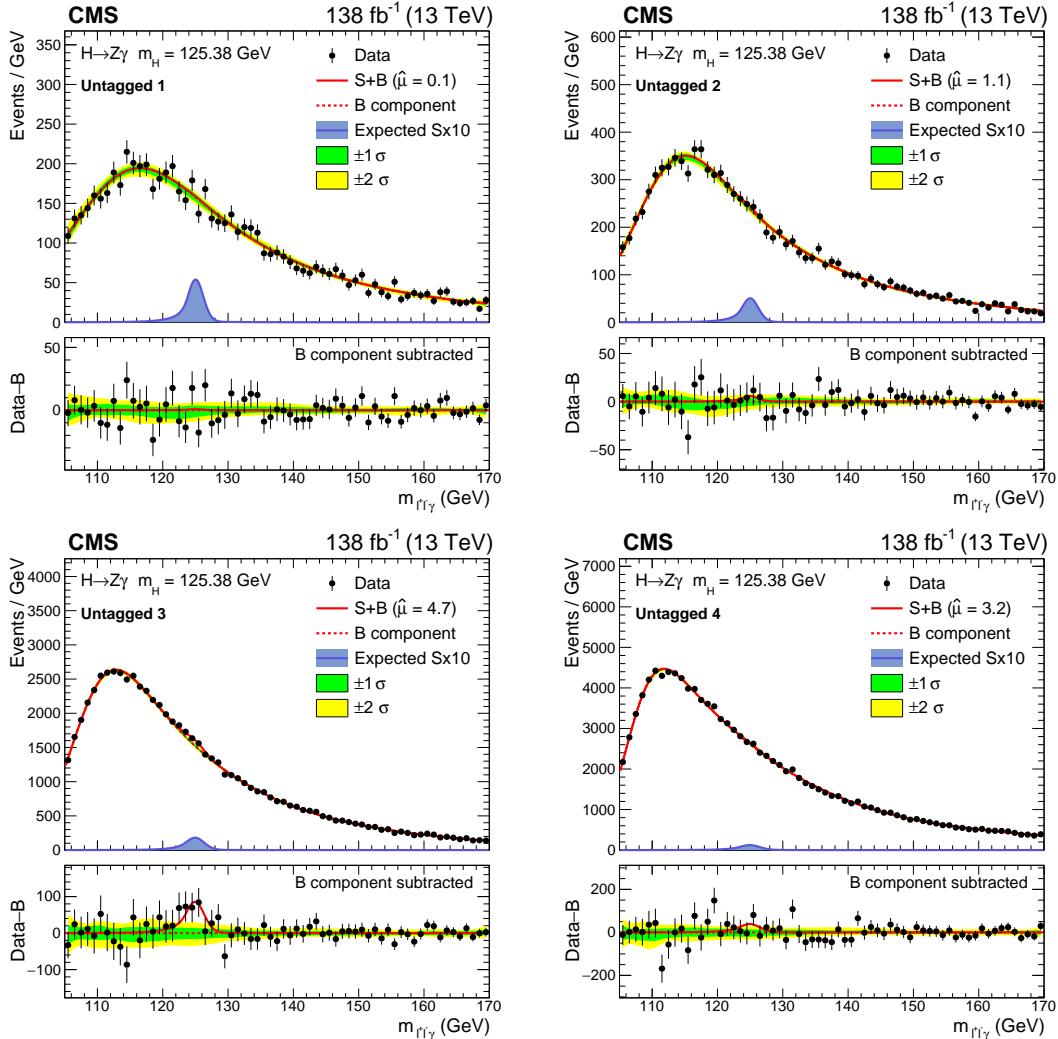


Figure 11.2. Fits to the  $m_{\ell^+\ell^-\gamma}$  data distribution in the untagged 1 (upper left), untagged 2 (upper right), untagged 3 (lower left), and untagged 4 (lower right) categories. In the upper panel, the red solid line shows the result of a signal-plus-background fit to the given category. The red dashed line shows the background component of the fit. The green and yellow bands represent the 68 and 95% CL uncertainties in the fit. Also plotted is the expected SM signal, scaled by a factor of 10. In the lower panel, the data minus the background component of the fit is shown.

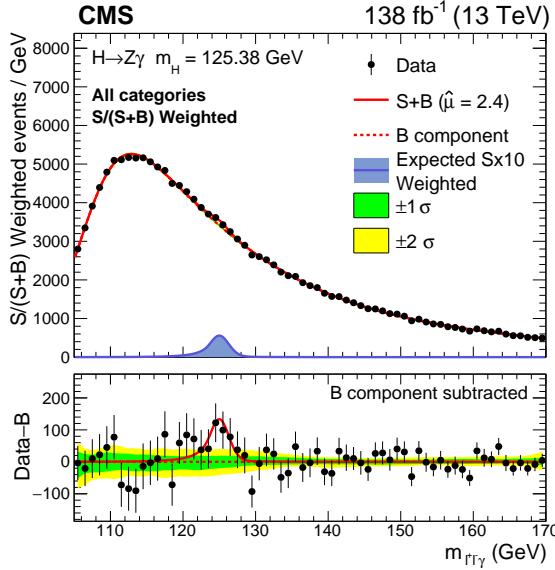


Figure 11.3. Sum over all categories of the data points and signal-plus-background model after the simultaneous fit to each  $m_{\ell^+\ell^-\gamma}$  distribution. The contribution from each category is weighted by  $S/(S + B)$ , as defined in the text. In the upper panel, the red solid line shows the signal-plus-background fit. The red dashed line shows the background component of the fit. The green and yellow bands represent the 68 and 95% CL uncertainties in the fit. Also plotted is the expected SM signal weighted by  $S/(S + B)$  and scaled by a factor of 10. In the lower panel, the data minus the background component of the fit is shown.

deviations. Upper limits on  $\mu$  are calculated at 1 GeV intervals in the mass range of  $120 < m_{\ell^+\ell^-\gamma} < 130$  GeV and at  $m_H = 125.38$  GeV, as shown in Fig. 11.5. The observed (expected) limit at 95% CL relative to the SM expectation for  $m_H = 125.38$  GeV is 4.1 (1.8).

### 11.1. Ratio Measurement

The ratio  $\mathcal{B}(H \rightarrow Z\gamma)/\mathcal{B}(H \rightarrow \gamma\gamma)$  is theoretically interesting, as it is potentially sensitive to beyond BSM physics, such as supersymmetry and extended Higgs sectors [15, 17, 16]. BSM effects from models such as these typically shift the  $H \rightarrow Z\gamma$  and  $H \rightarrow \gamma\gamma$

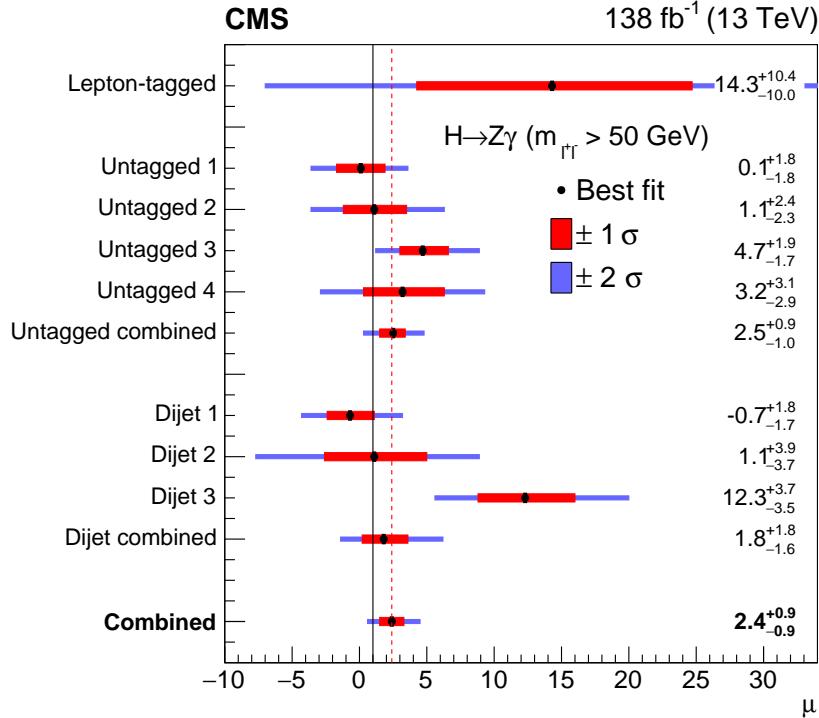


Figure 11.4. Observed signal strength ( $\mu$ ) for a SM Higgs boson with  $m_H = 125.38$  GeV. The labels “untagged combined,” “dijet combined,” and “combined” represent the results obtained from simultaneous fits of the untagged categories, dijet categories, and full set of categories, respectively. The black solid line shows  $\mu = 1$ , and the red dashed line shows the best fit value  $\hat{\mu} = 2.4 \pm 0.9$  of all categories combined.

branching fractions by different amounts, making the ratio the most sensitive observable. By combining the  $H \rightarrow Z\gamma$  analysis with the existing CMS  $H \rightarrow \gamma\gamma$  analysis, described in Ref. [13], we can obtain a measurement of this ratio and compare it to the SM prediction of  $0.69 \pm 0.04$ . An experimental advantage of this measurement is the cancellation of common theoretical uncertainties related to Higgs boson production and common experimental uncertainty sources like the L1 prefiring issue.

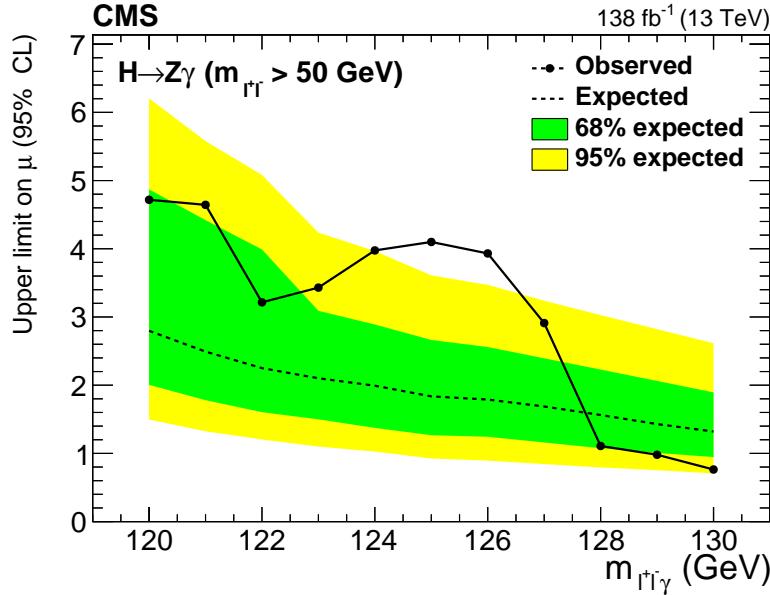


Figure 11.5. Upper limit (95% CL) on the signal strength ( $\mu$ ) relative to the SM prediction, as a function of the assumed value of the Higgs boson mass used in the fit.

The ratio measurement is obtained via a combined fit of the two analyses with two parameters of interest: the  $H \rightarrow \gamma\gamma$  signal strength with respect to the SM expectation ( $\mu_{\gamma\gamma}$ ) and the ratio of the signal strengths for the two channels ( $\mu_{Z\gamma}/\mu_{\gamma\gamma}$ ). After obtaining the best fit values for these parameters of interest, the ratio of signal strengths can be scaled to obtain the ratio of branching fractions. Common systematic uncertainties in the two analyses are correlated in the combined fit, and are listed in Table 11.1. The combination is performed at a Higgs boson mass of 125.38 GeV, and discrete profiling of the background models is enabled for both analyses.

The profile likelihood scans for the parameters of interest are shown in Figure 11.6. The best fit values of the parameters of interest are  $\mu_{\gamma\gamma} = 1.121^{+0.095}_{-0.090}$  and  $\mu_{Z\gamma}/\mu_{\gamma\gamma} = 2.225^{+0.926}_{-0.825}$ . The value of  $\mu_{\gamma\gamma}$  agrees well with the standalone  $H \rightarrow \gamma\gamma$  fit. Multiplying

Source	
Theoretical uncertainties	
Scale uncertainties	QCD scale ZH QCD scale WH QCD scale ggH QCD scale qqH QCD scale ttH
PDF uncertainties	pdf_Higgs_ggH pdf_Higgs_VH pdf_Higgs_qqH pdf_Higgs_ttH
Systematic uncertainties	
Luminosity	lumi_13TeV_Uncorrelated_2016,2017,2018 lumi_13TeV_Beam_Current_Calibration lumi_13TeV_Beam_Beam_Deflection lumi_13TeV_Dynamic_Beta lumi_13TeV_Ghosts_And_Satellite lumi_13TeV_Length_Scale lumi_13TeV_X_Y_Factorization
Parton showering	PartonShower_norm
Underlying events	UnderlyingEvent_norm
L1-prefire	2016,2017 L1 prefire uncertainty

Table 11.1. Correlated uncertainties between  $H \rightarrow Z\gamma$  and  $H \rightarrow \gamma\gamma$  for the  $\mathcal{B}(H \rightarrow Z\gamma)/\mathcal{B}(H \rightarrow \gamma\gamma)$  measurement.

through, the value of  $\mu_{Z\gamma}$  is 2.49, which agrees well with the result of the standalone  $H \rightarrow Z\gamma$  fit. The measured value of  $\mathcal{B}(H \rightarrow Z\gamma)/\mathcal{B}(H \rightarrow \gamma\gamma)$  from the combined fit with the  $H \rightarrow \gamma\gamma$  analysis is  $1.5^{+0.7}_{-0.6}$ . This measurement is consistent with the SM prediction for the ratio at the 1.5 standard deviation level. The nuisance parameter impacts on the signal strength ratio are shown in Figure 11.6.

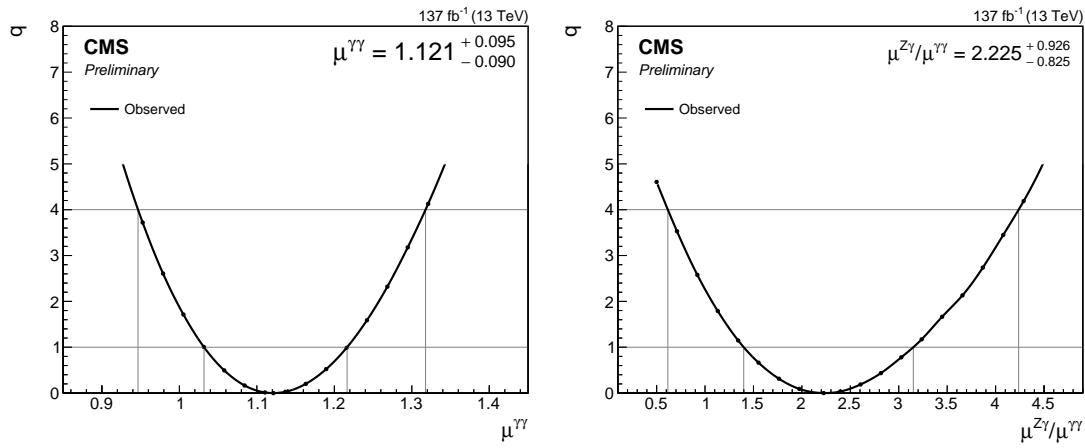
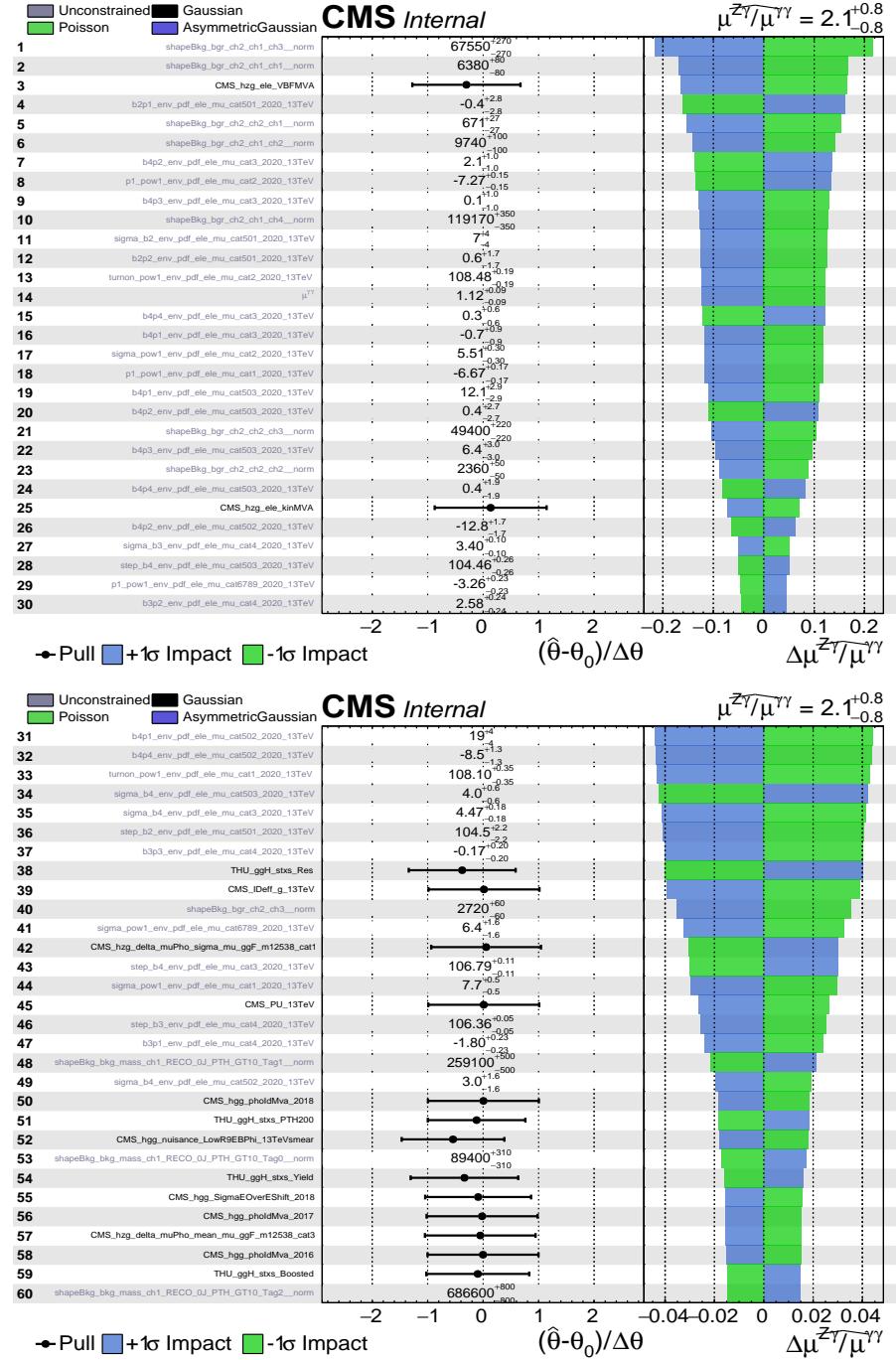


Figure 11.6. Profile likelihood scans of  $\mu_{\gamma\gamma}$  and  $\mu_{Z\gamma}/\mu_{\gamma\gamma}$  with fixed  $m_H = 125.38 \text{ GeV}$ .

Figure 11.6. Nuisance parameter impacts on  $\mu_{Z\gamma}/\mu_{\gamma\gamma}$ .

## CHAPTER 12

### Conclusion

A search is performed for a standard model (SM) Higgs boson decaying into a lepton pair ( $e^+e^-$  or  $\mu^+\mu^-$ ) and a photon with  $m_{\ell^+\ell^-} > 50$  GeV. The analysis is performed using a sample of CMS proton-proton (pp) collision data at  $\sqrt{s} = 13$  TeV, corresponding to an integrated luminosity of  $138 \text{ fb}^{-1}$ . The main contribution to this final state is from Higgs boson decays to a Z boson and a photon ( $H \rightarrow Z\gamma \rightarrow \ell^+\ell^-\gamma$ ). The contributions from final-state radiation of  $H \rightarrow \mu^+\mu^-$  or  $H \rightarrow \tau^+\tau^-$  are at the 6 and  $< 1\%$  levels, respectively, and are not included in the measured  $H \rightarrow Z\gamma$  signal. The best fit value of the signal strength  $\hat{\mu}$  for  $m_H = 125.38$  GeV is  $\hat{\mu} = 2.4^{+0.8}_{-0.9}$  (stat)  $^{+0.3}_{-0.2}$  (syst)  $= 2.4 \pm 0.9$ . This measurement corresponds to  $\sigma(\text{pp} \rightarrow H)\mathcal{B}(H \rightarrow Z\gamma) = 0.21 \pm 0.08$  pb. The measured value is 1.6 standard deviations higher than the SM prediction. The observed (expected) local significance is 2.7 (1.2) standard deviations, where the expected significance is determined for the SM hypothesis. The observed (expected) upper limit at 95% confidence level on  $\mu$  is 4.1 (1.8). In addition, a combined fit with the  $H \rightarrow \gamma\gamma$  analysis of the same data set [13] is performed to measure the ratio  $\mathcal{B}(H \rightarrow Z\gamma)/\mathcal{B}(H \rightarrow \gamma\gamma) = 1.5^{+0.7}_{-0.6}$ , which is consistent with the ratio of  $0.69 \pm 0.04$  predicted by SM at the 1.5 standard deviation level for a Higgs boson mass of 125.38 GeV.

## References

- [1] G. Aad *et al.* “Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC.” *Phys. Lett. B* (2012) **716**:1. doi:10.1016/j.physletb.2012.08.020. 1207.7214.
- [2] S. Chatrchyan *et al.* “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC.” *Phys. Lett. B* (2012) **716**:30. doi:10.1016/j.physletb.2012.08.021. 1207.7235.
- [3] S. Chatrchyan *et al.* “Observation of a new boson with mass near 125 GeV in pp collisions at  $\sqrt{s} = 7$  and 8 TeV.” *JHEP* (2013) **06**:081. doi:10.1007/JHEP06(2013)081. 1303.4571.
- [4] Particle Data Group, P. A. Zyla, *et al.* “Review of particle physics.” *Prog. Theor. Exp. Phys.* (2020) **2020**:083C01. doi:10.1093/ptep/ptaa104.
- [5] A. Abbasabadi, D. Bowser-Chao, D. A. Dicus, and W. W. Repko. “Radiative Higgs boson decays  $H \rightarrow f\bar{f}\gamma$ .” *Phys. Rev. D* (1997) **55**:5647. doi:10.1103/PhysRevD.55.5647. hep-ph/9611209.
- [6] L. B. Chen, C. F. Qiao, and R. L. Zhu. “Reconstructing the 125 GeV SM Higgs boson through  $\ell\bar{\ell}\gamma$ .” *Phys. Lett. B* (2013) **726**:306. doi:10.1016/j.physletb.2013.08.050. 1211.6058.
- [7] Y. Sun, H. Chang, and D. Gao. “Higgs decays to  $\gamma\ell^+\ell^-$  in the standard model.” *JHEP* (2013) **05**:061. doi:10.1007/JHEP05(2013)061. 1303.2230.
- [8] G. Passarino. “Higgs boson production and decay: Dalitz sector.” *Phys. Lett. B* (2013) **727**:424. doi:10.1016/j.physletb.2013.10.052. 1308.0422.
- [9] J. M. Campbell, R. K. Ellis, W. T. Giele, and C. Williams. “Finding the Higgs boson in decays to  $Z\gamma$  using the matrix element method at next-to-leading order.” *Phys. Rev. D* (2013) **87**:073005. doi:10.1103/PhysRevD.87.073005. 1301.7086.

- [10] G. Degrassi and M. Vitti. “The effect of an anomalous Higgs trilinear self-coupling on the  $h \rightarrow \gamma Z$  decay.” *Eur. Phys. J. C* (2020) **80**:307. doi:10.1140/epjc/s10052-020-7860-7. 1912.06429.
- [11] I. Low, J. Lykken, and G. Shaughnessy. “Singlet scalars as Higgs imposters at the Large Hadron Collider.” *Phys. Rev. D* (2011) **84**:035027. doi:10.1103/PhysRevD.84.035027. 1105.4587.
- [12] D. de Florian *et al.* “Handbook of LHC Higgs cross sections: 4. deciphering the nature of the Higgs sector.” (2016) **2/2017**. doi:10.23731/CYRM-2017-002. 1610.07922.
- [13] A. M. Sirunyan *et al.* “Measurements of Higgs boson production cross sections and couplings in the diphoton decay channel at  $\sqrt{s} = 13$  TeV.” *JHEP* (2021) **07**:027. doi:10.1007/JHEP07(2021)027. 2103.06956.
- [14] A. M. Sirunyan *et al.* “A measurement of the Higgs boson mass in the diphoton decay channel.” *Phys. Lett. B* (2020) **805**:135425. doi:10.1016/j.physletb.2020.135425. 2002.06398.
- [15] A. Djouadi, V. Driesen, W. Hollik, and A. Kraft. “The Higgs-photon-Z boson coupling revisited.” *Eur. Phys. J. C* (1998) **1**:163. doi:10.1007/BF01245806. hep-ph/9701342.
- [16] M. Carena, I. Low, and C. E. M. Wagner. “Implications of a modified Higgs to diphoton decay width.” *JHEP* (2012) **08**:060. doi:10.1007/JHEP08(2012)060. 1206.1082.
- [17] C.-W. Chiang and K. Yagyu. “Higgs boson decays to  $\gamma\gamma$  and  $Z\gamma$  in models with Higgs extensions.” *Phys. Rev. D* (2013) **87**:033003. doi:10.1103/PhysRevD.87.033003. 1207.1065.
- [18] C.-S. Chen, C.-Q. Geng, D. Huang, and L.-H. Tsai. “New scalar contributions to  $h \rightarrow Z\gamma$ .” *Phys. Rev. D* (2013) **87**:075019. doi:10.1103/PhysRevD.87.075019. 1301.4694.
- [19] G. Aad *et al.* “Search for Higgs boson decays to a photon and a Z boson in pp collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS detector.” *Phys. Lett. B* (2014) **732**:8. doi:10.1016/j.physletb.2014.03.015. 1402.3051.
- [20] S. Chatrchyan *et al.* “Search for a Higgs boson decaying into a Z and a photon in pp collisions at  $\sqrt{s} = 7$  and 8 TeV.” *Phys. Lett. B* (2013) **726**:587. doi:10.1016/j.physletb.2013.09.057. 1307.5515.

- [21] A. M. Sirunyan *et al.* “Search for the decay of a Higgs boson in the  $\ell\ell\gamma$  channel in proton-proton collisions at  $\sqrt{s} = 13$  TeV.” *JHEP* (2018) **11**:152. doi:10.1007/JHEP11(2018)152. 1806.05996.
- [22] G. Aad *et al.* “A search for the  $Z\gamma$  decay mode of the Higgs boson in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector.” *Phys. Lett. B* (2020) **809**:135754. doi: 10.1016/j.physletb.2020.135754. 2005.05382.
- [23] G. Aad *et al.* “Evidence for Higgs boson decays to a low-mass dilepton system and a photon in pp collisions at  $\sqrt{s}=13$  TeV with the ATLAS detector.” *Phys. Lett. B* (2021) **819**:136412. doi:10.1016/j.physletb.2021.136412. 2103.10322.
- [24] V. Khachatryan *et al.* “Search for a Higgs boson decaying into  $\gamma^*\gamma \rightarrow \ell\ell\gamma$  with low dilepton mass in pp collisions at  $\sqrt{s} = 8$  TeV.” *Phys. Lett. B* (2016) **753**:341. doi:10.1016/j.physletb.2015.12.039. 1507.03031.
- [25] Wikipedia Commons. “File:standard model of elementary particles.svg.” [https://en.wikipedia.org/wiki/File:Standard\\_Model\\_of\\_Elementary\\_Particles.svg#file](https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg#file) (2019).
- [26] F. Hasert, H. Faissner, W. Krenz, J. Von Krogh, D. Lanske, *et al.* “Search for elastic muon-neutrino electron scattering.” *Physics Letters B* (1973) **46** (1):121–124. ISSN 0370-2693. doi:10.1016/0370-2693(73)90494-2.
- [27] F. Hasert, S. Kabe, W. Krenz, J. Von Krogh, D. Lanske, *et al.* “Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment.” *Physics Letters B* (1973) **46** (1):138–140. ISSN 0370-2693. doi:[https://doi.org/10.1016/0370-2693\(73\)90499-1](https://doi.org/10.1016/0370-2693(73)90499-1).
- [28] G. Arnison *et al.* “Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at  $\sqrt{s} = 540$  GeV.” *Phys. Lett. B* (1983) **122**:103–116. doi:10.1016/0370-2693(83)91177-2.
- [29] M. Banner *et al.* “Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider.” *Phys. Lett. B* (1983) **122**:476–485. doi:10.1016/0370-2693(83)91605-2.
- [30] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons.” *Phys. Rev. Lett.* (1964) **13**:321–323. doi:10.1103/PhysRevLett.13.321.
- [31] P. W. Higgs. “Broken symmetries, massless particles and gauge fields.” *Phys. Lett.* (1964) **12**:132–133. doi:10.1016/0031-9163(64)91136-9.

- [32] P. W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons.” *Phys. Rev. Lett.* (1964) **13**:508–509. doi:10.1103/PhysRevLett.13.508.
- [33] A. Cho. “Spontaneous symmetry breaking decanted.” *Science* (2012) **337** (6100):1289–1289. doi:10.1126/science.337.6100.1289. <https://www.science.org/doi/pdf/10.1126/science.337.6100.1289>  
URL <https://www.science.org/doi/abs/10.1126/science.337.6100.1289>
- [34] L. Evans and P. Bryant. “LHC machine.” *Journal of Instrumentation* (2008) **3** (08):S08001–S08001. doi:10.1088/1748-0221/3/08/s08001.  
URL <https://doi.org/10.1088/1748-0221/3/08/s08001>
- [35] S. Myers. *The LEP Collider, from design to approval and commissioning*. John Adams’ memorial lecture. CERN, Geneva (1991). doi:10.5170/CERN-1991-008. Delivered at CERN, 26 Nov 1990.  
URL <https://cds.cern.ch/record/226776>
- [36] “LHC Design Report. 3. The LHC injector chain.” (2004) doi:10.5170/CERN-2004-003-V-3.
- [37] <https://home.cern/science/accelerators/accelerator-complex>.
- [38] S. Chatrchyan *et al.* “The CMS experiment at the CERN LHC.” *JINST* (2008) **3**:S08004. doi:10.1088/1748-0221/3/08/S08004.
- [39] A. M. Sirunyan *et al.* “Performance of the CMS Level-1 trigger in proton-proton collisions at  $\sqrt{s} = 13$  TeV.” *JINST* (2020) **15**:P10017. doi:10.1088/1748-0221/15/10/P10017. 2006.10165.
- [40] V. Khachatryan *et al.* “The CMS trigger system.” *JINST* (2017) **12**:P01020. doi:10.1088/1748-0221/12/01/P01020. 1609.02366.
- [41] V. Khachatryan *et al.* “Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at  $\text{sqrt}(s) = 8$  TeV.” *JINST* (2015) **10**:P08010. doi:10.1088/1748-0221/10/08/P08010. 1502.02702.
- [42] V. Khachatryan *et al.* “Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV.” *JINST* (2015) **10**:P06005. doi:10.1088/1748-0221/10/06/P06005. 1502.02701.
- [43] A. M. Sirunyan *et al.* “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at  $\sqrt{s} = 13$  TeV.” *JINST* (2018) **13**:P06015. doi:10.1088/1748-0221/13/06/P06015. 1804.04528.

- [44] S. Chatrchyan *et al.* “Description and performance of track and primary-vertex reconstruction with the CMS tracker.” *JINST* (2014) **9**:P10009. doi:10.1088/1748-0221/9/10/P10009. 1405.6569.
- [45] <https://cms.cern/detector>.
- [46] F. Cavallari and C. I. Rovelli. “Calibration and Performance of the CMS Electromagnetic Calorimeter in LHC Run2.” (2021) .  
URL <http://cds.cern.ch/record/2798128>
- [47] A. M. Sirunyan *et al.* “Particle-flow reconstruction and global event description with the CMS detector.” *JINST* (2017) **12**:P10003. doi:10.1088/1748-0221/12/10/P10003. 1706.04965.
- [48] W. Adam, R. Frühwirth, A. Strandlie, and T. Todor. “Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC.” (2005) .  
URL <https://cds.cern.ch/record/815410>
- [49] A. M. Sirunyan *et al.* “Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC.” *JINST* (2021) **16**:P05014. doi:10.1088/1748-0221/16/05/P05014. 2012.06888.
- [50] 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* (1987) **262** (2):444–450. ISSN 0168-9002. doi:[https://doi.org/10.1016/0168-9002\(87\)90887-4](https://doi.org/10.1016/0168-9002(87)90887-4).  
URL <https://www.sciencedirect.com/science/article/pii/0168900287908874>
- [51] CMS Collaboration. “Muon reconstruction performance during run 2.” (2019) .  
URL [https://cds.cern.ch/record/2682902/files/DP2019\\_022.pdf](https://cds.cern.ch/record/2682902/files/DP2019_022.pdf)
- [52] M. Cacciari, G. P. Salam, and G. Soyez. “The anti- $k_T$  jet clustering algorithm.” *JHEP* (2008) **04**:063. doi:10.1088/1126-6708/2008/04/063. 0802.1189.
- [53] M. Cacciari, G. P. Salam, and G. Soyez. “FastJet user manual.” *Eur. Phys. J. C* (2012) **72**:1896. doi:10.1140/epjc/s10052-012-1896-2. 1111.6097.
- [54] A. M. Sirunyan *et al.* “Pileup mitigation at CMS in 13 TeV data.” *JINST* (2020) **15**:P09018. doi:10.1088/1748-0221/15/09/P09018. 2003.00503.

- [55] V. Khachatryan *et al.* “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV.” *JINST* (2017) **12**:P02014. doi:10.1088/1748-0221/12/02/P02014. 1607.03663.
- [56] K. Ackerstaff *et al.* “Search for anomalous production of dilepton events with missing transverse momentum in  $e^+e^-$  collisions at  $\sqrt{s} = 161$  GeV and 172 GeV.” *Eur. Phys. J. C* (1998) **4**:47. doi:10.1007/PL00021655. [hep-ex/9710010](#).
- [57] M. Vesterinen and T. R. Wyatt. “A novel technique for studying the Z boson transverse momentum distribution at hadron colliders.” *Nucl. Instrum. Meth. A* (2009) **602**:432. doi:10.1016/j.nima.2009.01.203. 0807.4956.
- [58] A. M. Sirunyan *et al.* “Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at  $\sqrt{s} = 13$  TeV.” *JHEP* (2017) **11**:047. doi:10.1007/JHEP11(2017)047. 1706.09936.
- [59] A. M. Sirunyan *et al.* “Precision luminosity measurement in proton-proton collisions at  $\sqrt{s} = 13$  TeV in 2015 and 2016 at CMS.” *Eur. Phys. J. C* (2021) **81**:800. doi:10.1140/epjc/s10052-021-09538-2. 2104.01927.
- [60] “CMS luminosity measurement for the 2017 data-taking period at  $\sqrt{s} = 13$  TeV.” *CMS Physics Analysis Summary CMS-PAS-LUM-17-004* (2018). URL <https://cds.cern.ch/record/2621960>
- [61] “CMS luminosity measurement for the 2018 data-taking period at  $\sqrt{s} = 13$  TeV.” *CMS Physics Analysis Summary CMS-PAS-LUM-18-002* (2019). URL <https://cds.cern.ch/record/2676164>
- [62] S. Alioli, P. Nason, C. Oleari, and E. Re. “NLO Higgs boson production via gluon fusion matched with shower in POWHEG.” *JHEP* (2009) **04**:002. doi:10.1088/1126-6708/2009/04/002. 0812.0578.
- [63] P. Nason and C. Oleari. “NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG.” *JHEP* (2010) **02**:037. doi:10.1007/JHEP02(2010)037. 0911.5299.
- [64] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, *et al.* “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations.” *JHEP* (2014) **07**:079. doi:10.1007/JHEP07(2014)079. 1405.0301.

- [65] S. Frixione, P. Nason, and G. Ridolfi. “A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction.” *JHEP* (2007) **09**:126. doi:10.1088/1126-6708/2007/09/126. 0707.3088.
- [66] C. Anastasiou, C. Duhr, F. Dulat, E. Furlan, T. Gehrmann, *et al.* “High precision determination of the gluon fusion Higgs boson cross-section at the LHC.” *JHEP* (2016) **05**:058. doi:10.1007/JHEP05(2016)058. 1602.00695.
- [67] M. Cacciari, F. A. Dreyer, A. Karlberg, G. P. Salam, and G. Zanderighi. “Fully differential vector-boson-fusion higgs production at next-to-next-to-leading order.” *Phys. Rev. Lett.* (2015) **115**:082002. doi:10.1103/PhysRevLett.115.082002. [Erratum: doi:10.1103/PhysRevLett.120.139901].  
URL <https://link.aps.org/doi/10.1103/PhysRevLett.115.082002>
- [68] O. Brein, A. Djouadi, and R. Harlander. “NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders.” *Phys. Lett. B* (2004) **579**:149. doi:10.1016/j.physletb.2003.10.112. hep-ph/0307206.
- [69] S. Dawson, C. Jackson, L. H. Orr, L. Reina, and D. Wackerlo. “Associated Higgs production with top quarks at the large hadron collider: NLO QCD corrections.” *Phys. Rev. D* (2003) **68**:034022. doi:10.1103/PhysRevD.68.034022. hep-ph/0305087.
- [70] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, *et al.* “An introduction to PYTHIA 8.2.” *Comput. Phys. Commun.* (2015) **191**:159. doi:10.1016/j.cpc.2015.01.024. 1410.3012.
- [71] V. Khachatryan *et al.* “Event generator tunes obtained from underlying event and multiparton scattering measurements.” *Eur. Phys. J. C* (2016) **76**:155. doi:10.1140/epjc/s10052-016-3988-x. 1512.00815.
- [72] A. M. Sirunyan *et al.* “Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements.” *Eur. Phys. J. C* (2020) **80**:4. doi:10.1140/epjc/s10052-019-7499-4. 1903.12179.
- [73] R. D. Ball *et al.* “Parton distributions for the LHC run II.” *JHEP* (2015) **04**:040. doi:10.1007/JHEP04(2015)040. 1410.8849.
- [74] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, *et al.* “Parton distributions from high-precision collider data.” *Eur. Phys. J. C* (2017) **77**:663. doi:10.1140/epjc/s10052-017-5199-5. 1706.00428.
- [75] “Geant4—a simulation toolkit.” *Nucl. Instrum. Meth. A* (2003) **506**:250. ISSN 0168-9002. doi:10.1016/S0168-9002(03)01368-8.

- [76] W. Erdmann. “Offline primary vertex reconstruction with deterministic annealing clustering.” *CMS Internal note* (2011) **2011/014**.
- [77] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/L1PrefiringWeightRecipe>.
- [78] M. J. Oreglia. “A study of the reactions  $\psi' \rightarrow \gamma\gamma\psi$ .” Ph.D. thesis, Stanford University (1980). SLAC Report SLAC-R-236, see Appendix D.  
URL <http://www.slac.stanford.edu/pubs/slacreports/slac-r-236.html>
- [79] P. D. Dauncey, M. Kenzie, N. Wardle, and G. J. Davies. “Handling uncertainties in background shapes.” *JINST* (2015) **10**:P04015. doi:10.1088/1748-0221/10/04/P04015. 1408.6865.
- [80] CMS Collaboration. “Search for a higgs boson decaying into a z boson and a photon in pp collisions at  $\sqrt{s} = 7$  and 8 tev.” *Phys. Lett. B* (2013) **726**:587––609.
- [81] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. “Parton distributions for the LHC.” *Eur. Phys. J. C* (2009) **63**:189. doi:10.1140/epjc/s10052-009-1072-5. 0901.0002.
- [82] H. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, *et al.* “New parton distributions for collider physics.” *Phys. Rev. D* (2010) **82**:74024. doi:10.1103/PhysRevD.82.074024. 1007.2241.
- [83] J. Butterworth *et al.* “PDF4LHC recommendations for LHC run II.” *J. Phys. G* (2016) **43**:023001. doi:10.1088/0954-3899/43/2/023001. 1510.03865.
- [84] A. M. Sirunyan *et al.* “Measurement of the inelastic proton-proton cross section at  $\sqrt{s} = 13$  TeV.” *JHEP* (2018) **07**:161. doi:10.1007/JHEP07(2018)161. 1802.02613.
- [85] R. A. Fisher. “On the mathematical foundations of theoretical statistics.” *Phil. Trans. Roy. Soc. Lond.* (1922) **222**:309. doi:10.1098/rsta.1922.0009.
- [86] ATLAS and CMS Collaborations, The LHC Higgs combination group. “Procedure for the LHC Higgs boson search combination in summer 2011.” *Technical Report CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11* (2011).  
URL <https://cds.cern.ch/record/1379837>
- [87] A. L. Read. “Presentation of search results: The  $CL_s$  technique.” *J. Phys. G* (2002) **28**:2693. doi:10.1088/0954-3899/28/10/313.

- [88] T. Junk. “Confidence level computation for combining searches with small statistics.” *Nucl. Instrum. Meth. A* (1999) **434**:435. doi:10.1016/S0168-9002(99)00498-2. [hep-ex/9902006](#).
- [89] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. “Asymptotic formulae for likelihood-based tests of new physics.” *Eur. Phys. J. C* (2011) **71**:1554. doi: 10.1140/epjc/s10052-011-1554-0. [Erratum: doi:10.1140/epjc/s10052-013-2501-z], [1007.1727](#).

## APPENDIX A

### Signal and Resonant Background Fits

#### A.1. 2017 Signal Fits

#### A.2. 2018 Signal Fits

#### A.3. 2017 Resonant Background Fits

#### A.4. 2018 Resonant Background Fits

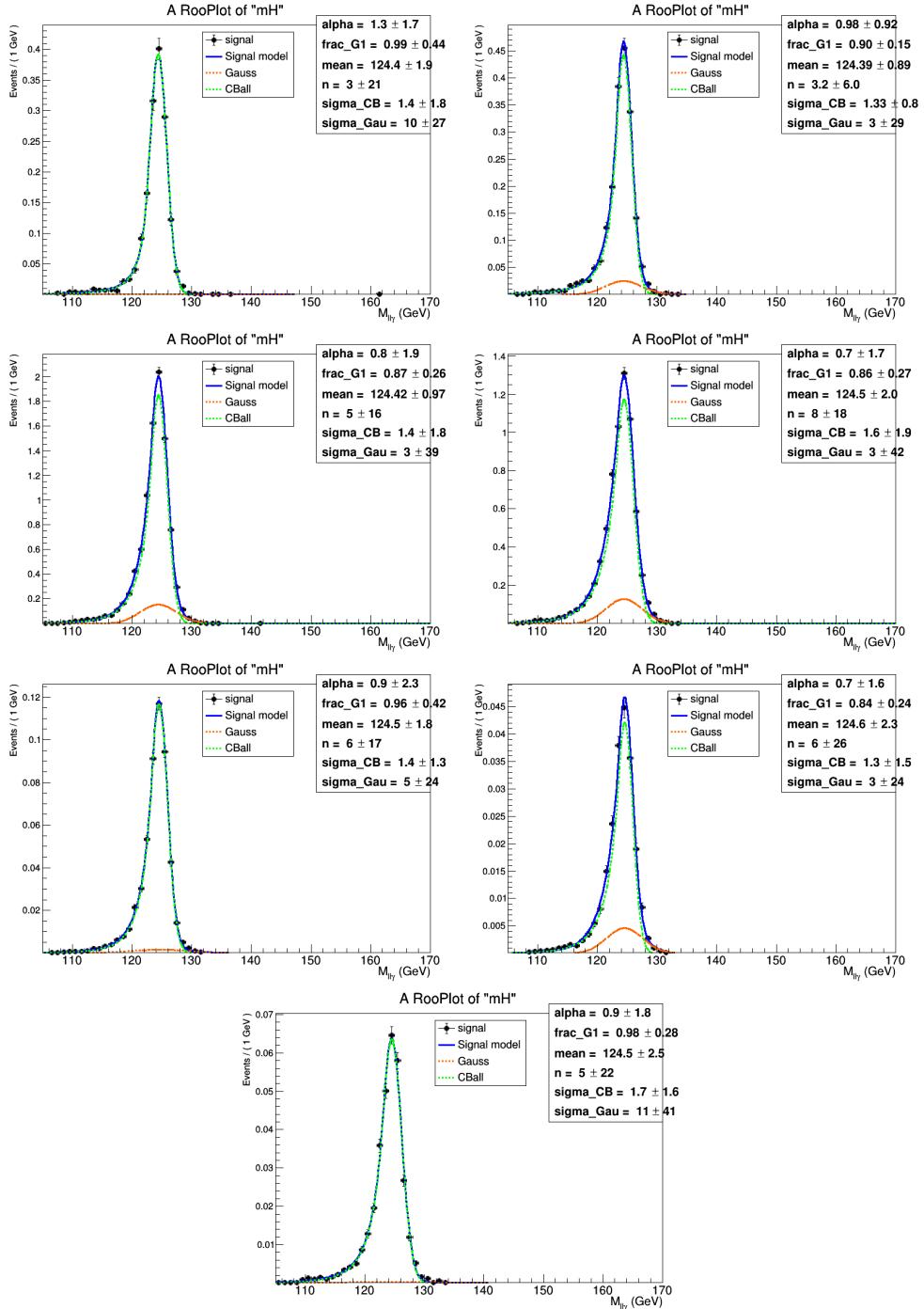


Figure A.1. Fits to simulated  $m_{\ell^+\ell^-\gamma}$  signal distributions in the electron channel for  $m_H = 125$  GeV for the 2017 data-taking period. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component. The top four plots correspond to the untagged categories, and the bottom three plots correspond to the dijet categories.

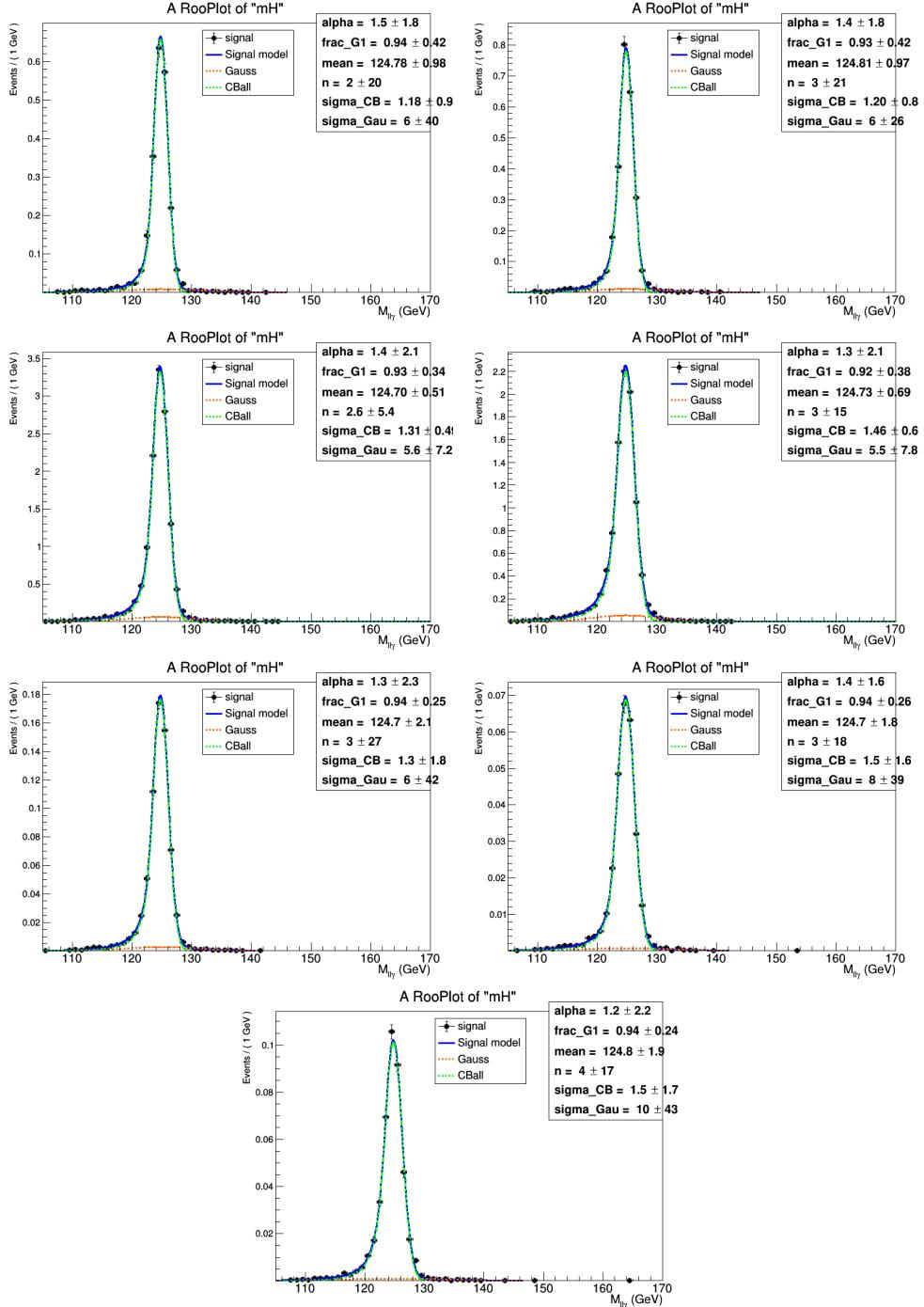


Figure A.2. Fits to simulated  $m_{\ell^+\ell^-\gamma}$  signal distributions in the muon channel for  $m_H = 125$  GeV for the 2017 data-taking period. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component. The top four plots correspond to the untagged categories, and the bottom three plots correspond to the dijet categories.

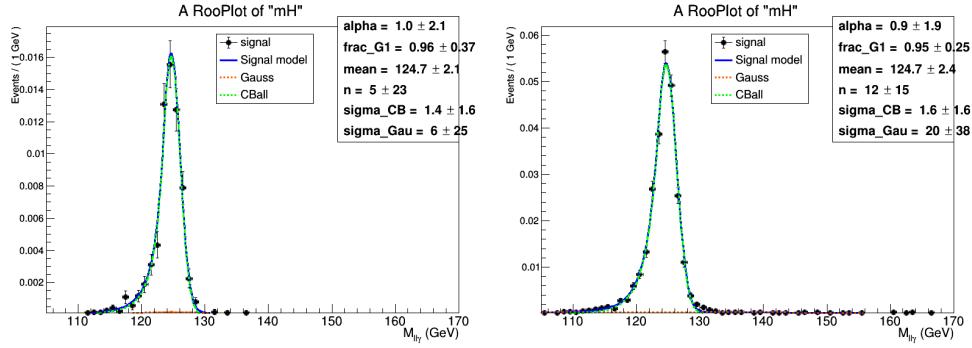


Figure A.3. Fits to simulated  $m_{\ell^+\ell^-\gamma}$  signal distributions in the electron and muon channels combined in the lepton-tagged category for  $m_H = 125$  GeV for the 2017 data-taking period. The left plot shows the fit to simulated ZH production events, and the right plot shows the fit to simulated WH production events. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component.

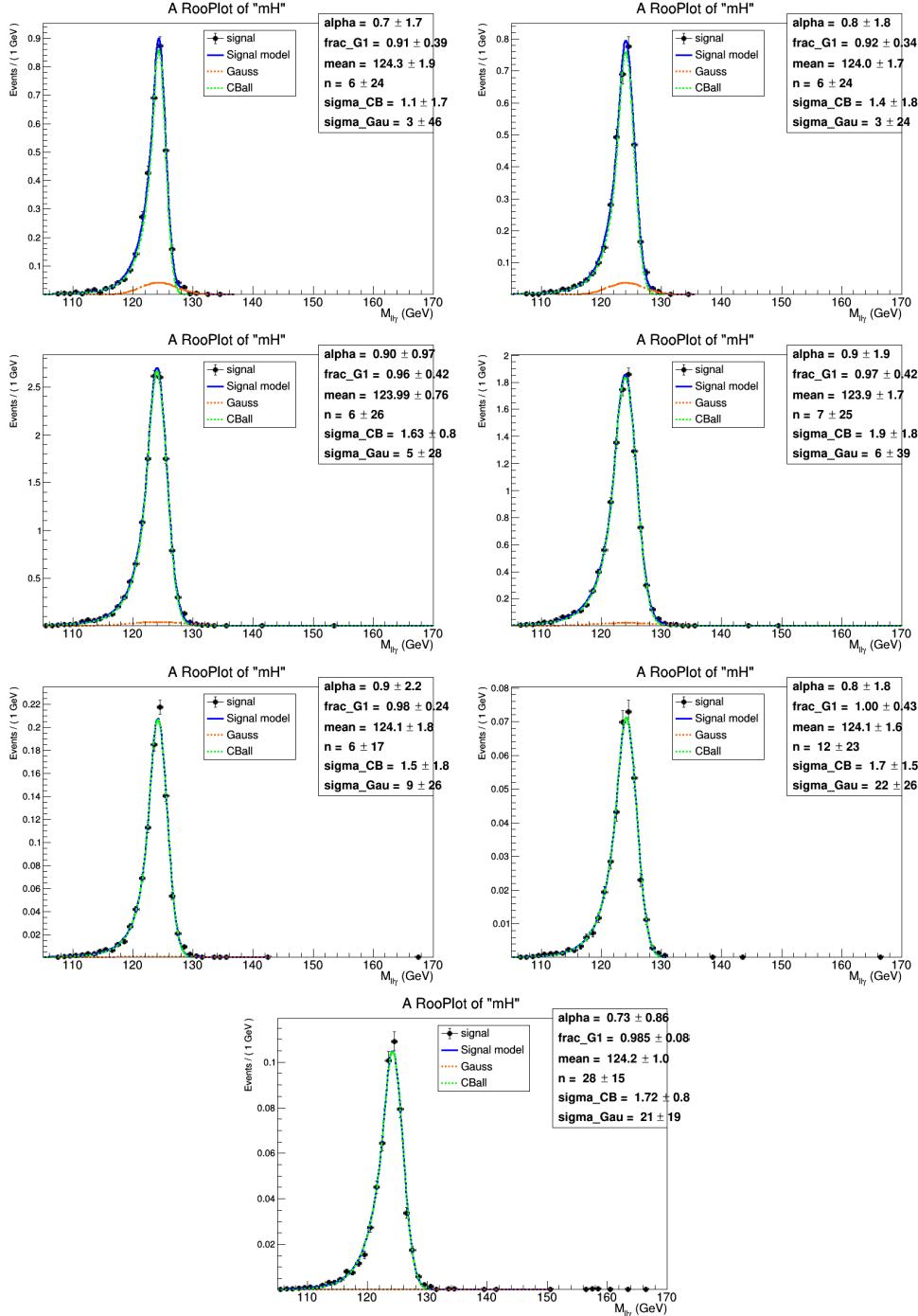


Figure A.4. Fits to simulated  $m_{\ell^+\ell^-\gamma}$  signal distributions in the electron channel for  $m_H = 125$  GeV for the 2018 data-taking period. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component. The top four plots correspond to the untagged categories, and the bottom three plots correspond to the dijet categories.

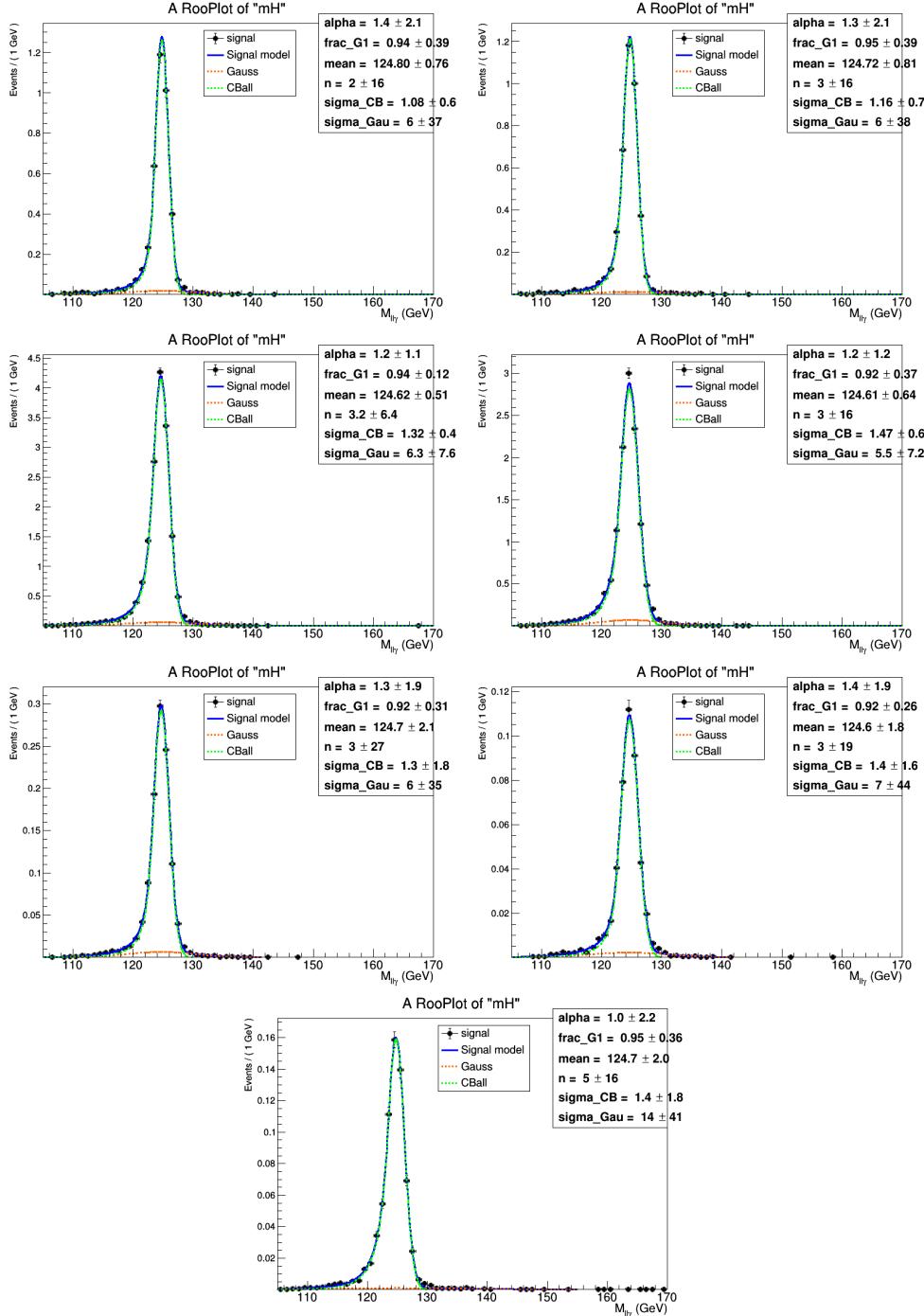


Figure A.5. Fits to simulated  $m_{\ell^+\ell^-\gamma}$  signal distributions in the muon channel for  $m_H = 125$  GeV for the 2018 data-taking period. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component. The top four plots correspond to the untagged categories, and the bottom three plots correspond to the dijet categories.

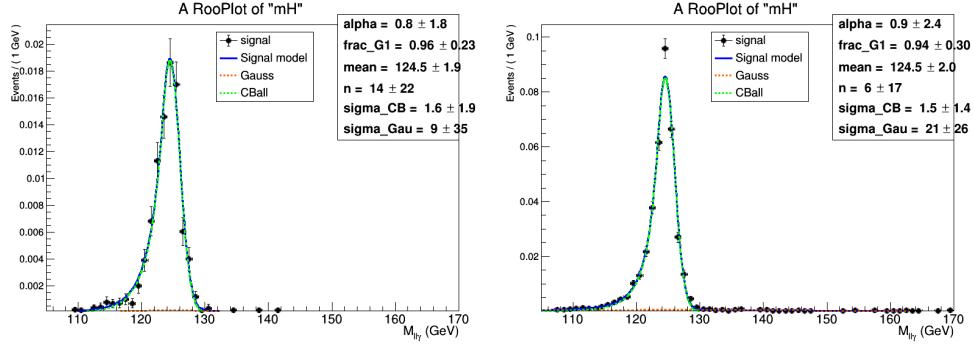


Figure A.6. Fits to simulated  $m_{\ell^+\ell^-\gamma}$  signal distributions in the electron and muon channels combined in the lepton-tagged category for  $m_H = 125$  GeV for the 2018 data-taking period. The left plot shows the fit to simulated ZH production events, and the right plot shows the fit to simulated WH production events. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component.

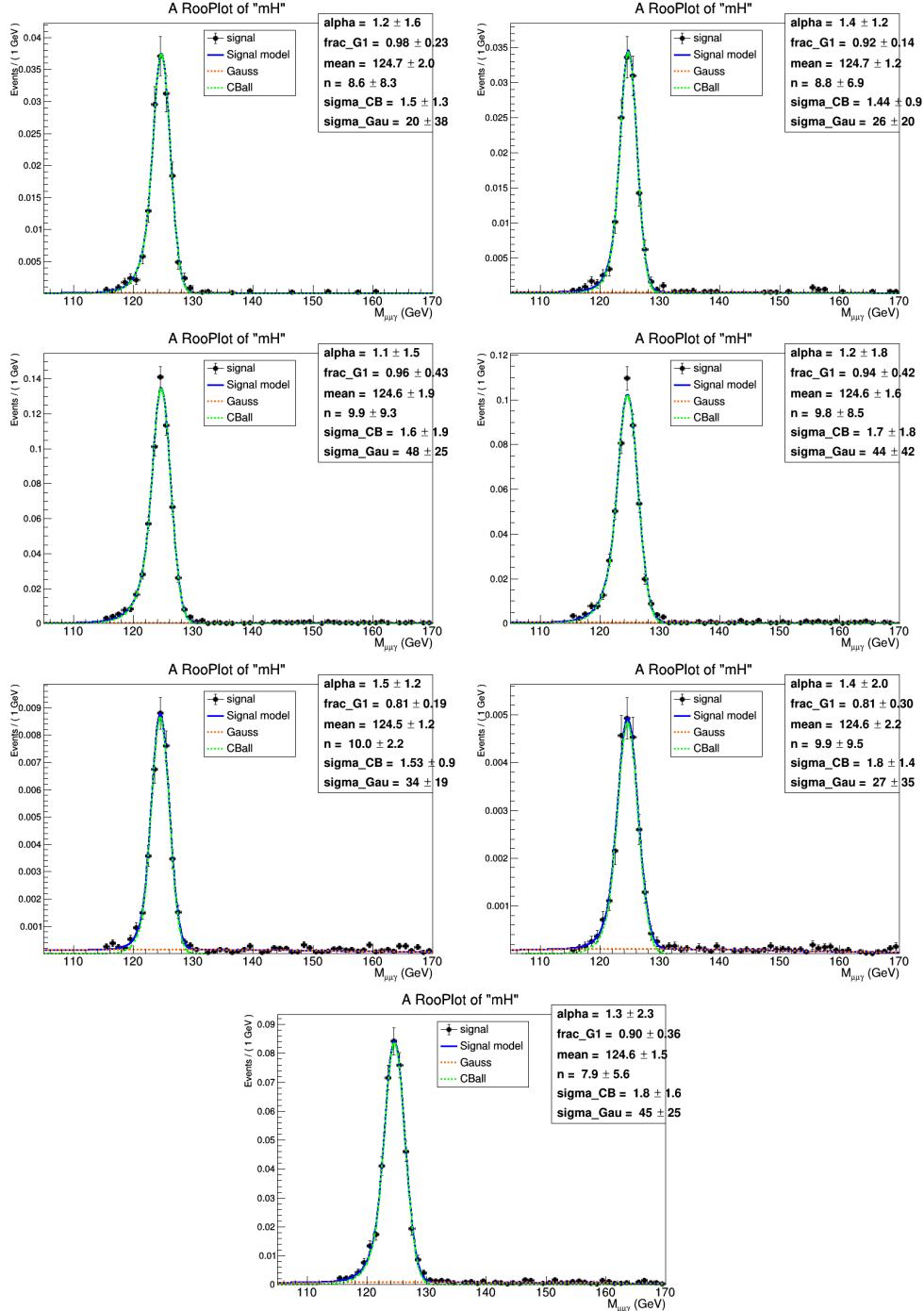


Figure A.7. Fits to simulated  $m_{\mu^+\mu^-\gamma}$  resonant background distributions from  $H \rightarrow \mu^+\mu^-$  for  $m_H = 125$  GeV for the 2017 data-taking period. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component. The top four plots correspond to the untagged categories, and the bottom three plots correspond to the dijet categories.

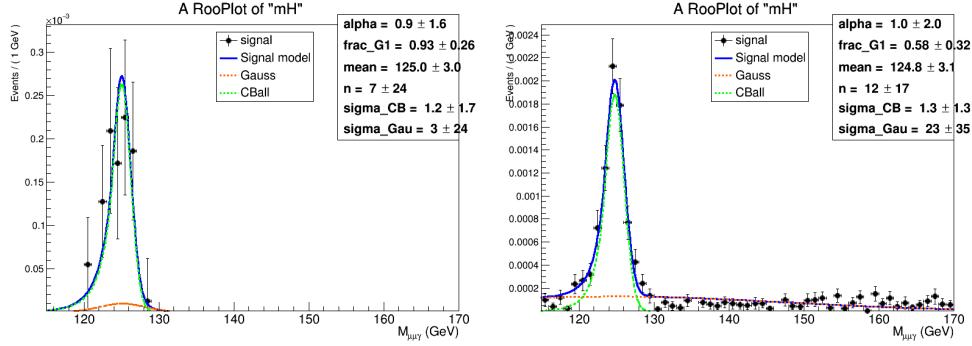


Figure A.8. Fits to simulated  $m_{\ell^+\ell^-\gamma}$  resonant background distributions from  $H \rightarrow \mu^+\mu^-$  in the electron and muon channels combined in the lepton-tagged category for  $m_H = 125$  GeV for the 2017 data-taking period. The left plot shows the fit to simulated ZH production events, and the right plot shows the fit to simulated WH production events. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component.

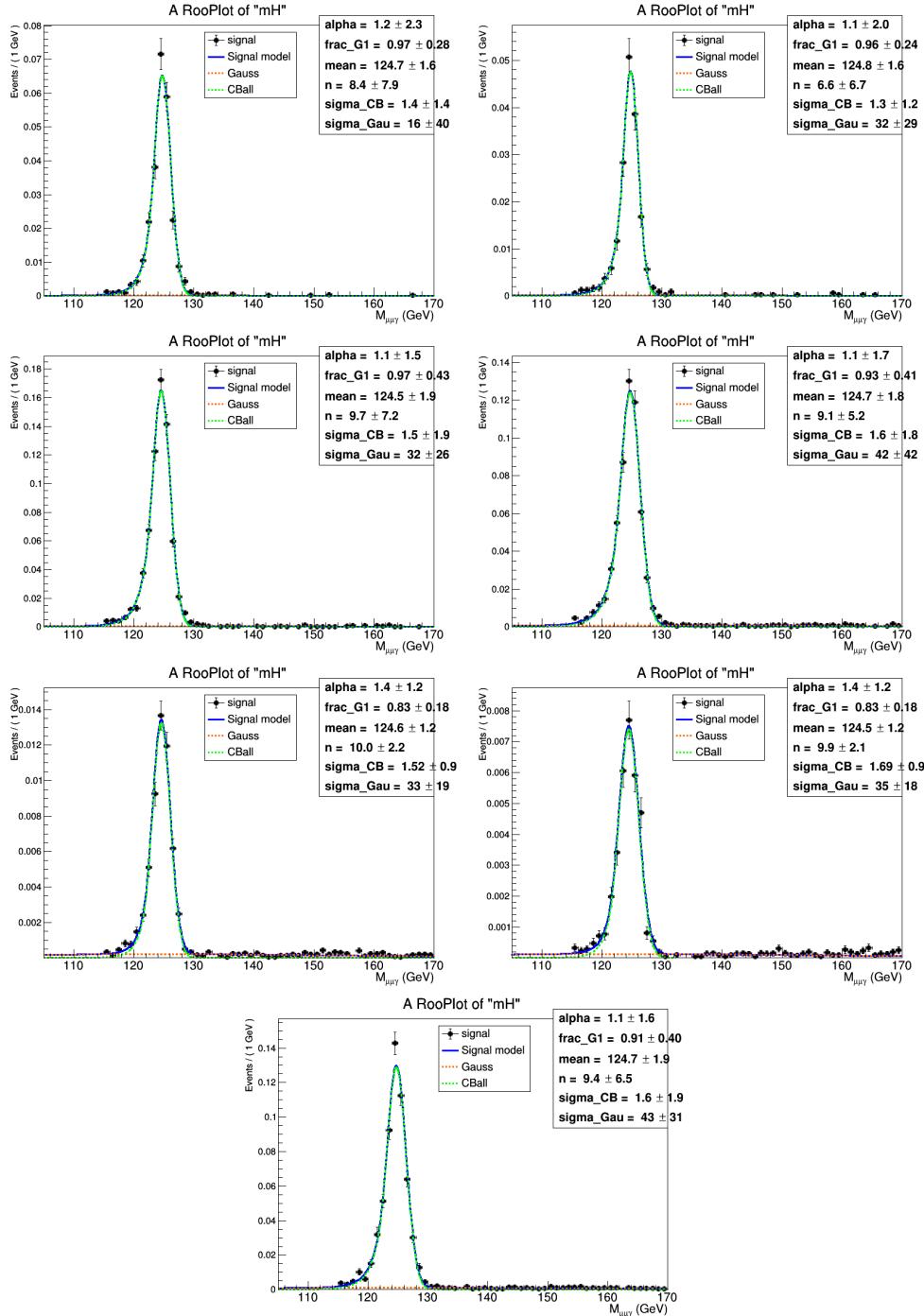


Figure A.9. Fits to simulated  $m_{\mu^+\mu^-\gamma}$  resonant background distributions from  $H \rightarrow \mu^+\mu^-$  for  $m_H = 125$  GeV for the 2018 data-taking period. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component. The top four plots correspond to the untagged categories, and the bottom three plots correspond to the dijet categories.

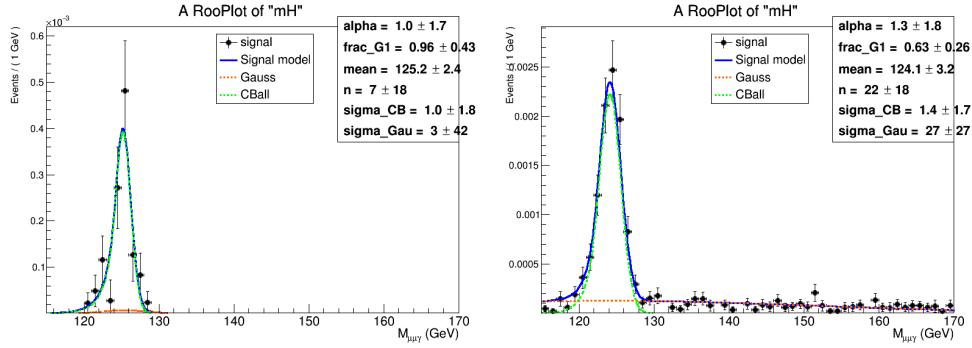


Figure A.10. Fits to simulated  $m_{\ell^+\ell^-\gamma}$  resonant background distributions from  $H \rightarrow \mu^+\mu^-$  in the electron and muon channels combined in the lepton-tagged category for  $m_H = 125$  GeV for the 2018 data-taking period. The left plot shows the fit to simulated ZH production events, and the right plot shows the fit to simulated WH production events. The blue line shows the total fit function, the green line shows the Crystal Ball function component, and the red line shows the Gaussian function component.