

# Extraction of CP Properties of the H(125) Boson Discovered in Proton-Proton Collisions at $\sqrt{s} = 7$ and 8 TeV with the CMS Detector at the LHC

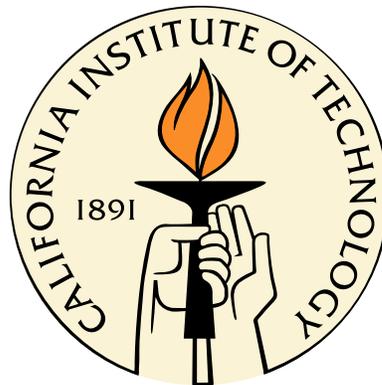
Thesis by

Yi Chen

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2015

(Submitted December 1, 2014)

© 2015

Yi Chen

All Rights Reserved

# Acknowledgements

To Maria Spiropulu - For being my advisor and support which makes this research possible.

To Harvey Newman - For his advice and help during the past six years.

To my parents Ming-huei Chen and Mei-yu Shih - For their support for me to study abroad.

To my friends at Caltech and in CMS - For your friendship and help which helps make the PhD program an enjoyable experience.

# Abstract

In this thesis we build a novel analysis framework to perform the direct extraction of all possible effective Higgs boson couplings to the neutral electroweak gauge bosons in the  $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$  channel also referred to as the golden channel. We use analytic expressions of the full decay differential cross sections for the  $H \rightarrow VV' \rightarrow 4\ell$  process, and the dominant irreducible standard model  $q\bar{q} \rightarrow 4\ell$  background where  $4\ell = 2e2\mu, 4e, 4\mu$ . Detector effects are included through an explicit convolution of these analytic expressions with transfer functions that model the detector responses as well as acceptance and efficiency effects. Using the full set of decay observables, we construct an unbinned 8-dimensional detector level likelihood function which is continuous in the effective couplings, and includes systematics. All potential anomalous couplings of  $HVV'$  where  $V = Z, \gamma$  are considered, allowing for general  $CP$  even/odd admixtures and any possible phases. We measure the  $CP$ -odd mixing between the tree-level  $HZZ$  coupling and higher order  $CP$ -odd couplings to be compatible with zero, and in the range  $[-0.40, 0.43]$ , and the mixing between  $HZZ$  tree-level coupling and higher order  $CP$ -even coupling to be in the ranges  $[-0.66, -0.57] \cup [-0.15, 1.00]$ ; namely compatible with a standard model Higgs. We discuss the expected precision in determining the various  $HVV'$  couplings in future LHC runs. A powerful and at first glance surprising prediction of the analysis is that with  $100\text{-}400 \text{ fb}^{-1}$ , the golden channel will be able to start probing the couplings of the Higgs boson to diphotons in the  $4\ell$  channel. We discuss the implications and further optimization of the methods for the next LHC runs.

# Contents

Acknowledgements	iii
Abstract	iv
1 Preface	1
<b>I Higgs and the Golden Channel</b>	<b>5</b>
2 The Standard Model And The Higgs Mechanism	6
2.1 Particle And Force Content . . . . .	6
2.2 Electroweak Symmetry Breaking and Gauge Boson Masses . . . . .	7
2.3 Fermion Masses . . . . .	10
2.4 Parity And Charge Conjugation . . . . .	12
2.5 Higgs Production and Decays at the LHC . . . . .	13
2.6 The “Golden Channel” . . . . .	14
<b>3 Effective Theory Approach and the Properties Measurements of the Higgs in the Four Lepton Channel</b>	<b>16</b>
3.1 parametrization Of The $HVV'$ Vertex . . . . .	16
3.2 Effective Theory: Range Of Validity . . . . .	19
<b>4 Helicity Amplitude Formulation for the <math>ZZ</math> On-Shell Case</b>	<b>20</b>
4.1 $CP$ Eigenstates and Lorentz Structures . . . . .	20
<b>5 <math>CP</math> Violation in the <math>HVV'</math> Sector</b>	<b>24</b>

5.1	<i>CP</i> Violation Within The Same $VV'$ Sector . . . . .	24
5.2	<i>CP</i> Violation Across Different $VV'$ Sectors - Strong Vs. Weak Phase . . . . .	26
<b>6</b>	<b>The Four Lepton Standard Model Backgrounds</b>	<b>29</b>
6.1	Background Processes And Components . . . . .	29
<b>II</b>	<b>The Higgs Discovery and the Properties Measurements</b>	
	<b>32</b>	
<b>7</b>	<b>Description of the LHC and the CMS Detector</b>	<b>33</b>
7.1	The Large Hadron Collider . . . . .	33
7.2	The Compact Muon Solenoid Detector . . . . .	35
<b>8</b>	<b>The Higgs Discovery</b>	<b>39</b>
8.1	Higgs Decay Channels . . . . .	39
8.2	The Higgs to Four Lepton Analysis . . . . .	40
8.2.1	Lepton Selection . . . . .	42
8.2.2	Final-state radiation recovery . . . . .	45
8.2.3	Kinematic selections . . . . .	45
8.2.4	The Standard Model Background Estimation . . . . .	46
8.3	The Four Lepton Analysis Results . . . . .	49
<b>9</b>	<b>The Higgs Boson Properties</b>	<b>53</b>
9.1	Introduction . . . . .	53
9.2	The Higgs Mass Measurement . . . . .	54
9.3	The Higgs Couplings Measurements . . . . .	55
9.4	The Higgs Width Measurement . . . . .	59
9.5	The Higgs Spin Inference . . . . .	59
9.6	The Higgs <i>CP</i> Properties . . . . .	63
<b>10</b>	<b>The Golden Channel Challenges</b>	<b>65</b>
10.1	The Challenge of the Twelve Observables . . . . .	65

10.2	The Four Lepton Parameter Space . . . . .	66
<b>11</b>	<b>Super Fast Detector Parametrization</b>	<b>72</b>
11.1	parametrization Of The Lepton Response . . . . .	73
11.2	Transfer Functions Validation . . . . .	74
<b>12</b>	<b>8D Likelihood Effective Higgs Couplings Extraction Framework in the Golden Channel</b>	<b>79</b>
12.1	The Multi-Dimensional Method Overview . . . . .	79
12.2	The Higgs and Other Standard Model Four Lepton Differential Cross Sections . . . . .	80
12.3	Integration Towards The 8D Likelihood . . . . .	85
12.4	Signal Likelihood for an Arbitrary Model . . . . .	87
12.5	Templated Background Components . . . . .	89
12.6	Systematics Uncertainties . . . . .	93
12.7	Total Dataset Likelihood . . . . .	94
12.8	Measurement and Inference . . . . .	95
<b>13</b>	<b>The Convolution Integral: Technical Description</b>	<b>98</b>
13.1	General Strategy . . . . .	98
13.2	Jacobian Factors . . . . .	101
13.2.1	From the Observable Basis to the Lepton Momentum Basis . .	101
13.2.2	The Background and Signal Factors . . . . .	104
13.3	Integration Strategy . . . . .	105
13.3.1	Mass Dimensions parametrization . . . . .	106
13.3.2	Mass Dimensions Grid Spacing . . . . .	107
13.3.3	Modified 2nd order Newton-Cotes Formula . . . . .	108
13.3.4	The Central Grid Point Optimization . . . . .	109
13.3.5	Differential cross section expansion: Background case . . . . .	110
13.4	Recursive Integration over Scaling Factors . . . . .	114
13.5	The Signal Case . . . . .	115

13.6 Convolution Integral Validation . . . . .	117
13.6.1 Validation Samples Generation . . . . .	119
13.6.2 Projection and Likelihood Validation . . . . .	120
13.6.3 Validation Fit Results . . . . .	124
<b>14 Four-Lepton 8D Analysis Result</b>	<b>130</b>
<b>III The Power of the Four Lepton Analysis</b>	<b>138</b>
<b>15 Four Lepton Analysis Sensitivity to Higgs Couplings beyond the ZZ Sector</b>	<b>139</b>
15.1 Sensitivity to $Z\gamma$ and $\gamma\gamma$ couplings . . . . .	139
15.1.1 High Luminosity Projections . . . . .	144
<b>16 Summary and outlook</b>	<b>151</b>
<b>A Standard Model <math>Z</math> Boson Couplings and Quark Charge Measurement</b>	<b>154</b>
A.1 $q\bar{q} \rightarrow 4\ell$ process . . . . .	154
A.2 Sensitivity Projection . . . . .	156
<b>B HCAL Noise</b>	<b>158</b>
B.1 The CMS HCAL Detector . . . . .	158
B.2 Characterization of Anomalous Noise in HCAL . . . . .	159
B.3 Fit-based Filters . . . . .	160
B.4 The 2012 Baseline Filter package . . . . .	165
B.5 Development for the 25ns data . . . . .	169
<b>C HCAL Test Beam Studies</b>	<b>172</b>
C.1 Experiment Setup . . . . .	172
C.2 Analysis . . . . .	173
C.3 Result . . . . .	176

<b>D Measurement of the Production Rate of Vector Boson in Association with Jets in CMS</b>	<b>183</b>
D.1 Description . . . . .	183
D.2 Methods . . . . .	185
D.3 Results . . . . .	187
<b>E Leptoquark Search</b>	<b>189</b>
E.1 Final State . . . . .	189
E.2 Razor Variables . . . . .	190
E.3 Methods . . . . .	193
E.4 Results . . . . .	195

# List of Figures

- 2.1 Shape of the “Mexican hat” Higgs potential as a function of the  $\phi$  field represented in the  $Re(\phi)$  and  $Im(\phi)$  axes. The height denotes the potential energy. The equilibrium at zero is not stable, as the energy is higher than its neighborhood. This triggers spontaneous symmetry breaking: the vacuum takes on a specific value along the minimum and loses its rotational symmetry. . . . . 10
- 2.2 Branching fractions of the Higgs boson to different final states. For a Higgs boson at a mass of 125 GeV the situation is especially interesting experimentally, because both fermionic and bosonic decays of the Higgs boson are accessible. . . . . 14
- 2.3 The relevant Feynman diagrams for the signal and background in the  $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$  analysis. The upper left diagram shows the dominant production mode at the LHC for the Higgs boson from gluon fusion. The Higgs then decays into four leptons through intermediate vector bosons. The upper right diagram shows the dominant standard model background  $4\ell$  process through a  $t/u$ -channel diagram from a  $q\bar{q}$  initial state. The lower left diagram is the  $s$ -channel  $Z \rightarrow 4\ell$  resonant process, which contributes roughly 10% of the background at  $m_{4\ell} = 125$  GeV. The lower right diagram is the sub-dominant background coming from gluon fusion, which at  $m_{4\ell} \sim 125$  GeV comprises about 3% of all the background. . . . . 15

- 3.1 Shape of the different terms in the effective Lagrangian, as a function of the dilepton masses. The upper left figure is the standard model  $HZZ$  tree-level coupling, and other three panels are higher order  $HVV'$  couplings. The upper right figure is  $A^{ZZ}$ , and the lower row is  $A^{Z\gamma}$  and  $A^{\gamma\gamma}$  respectively. The pink line indicates the lepton mass-pair ranges covered by typical analyses in CMS. . . . . 18
- 4.1 Three configurations of the spin lineup for the vector bosons decaying from the Higgs. Since the Higgs boson is spinless, the total projection of the spin along the decay axis must add up to zero. The projection of spins can be both parallel to the motion, both anti-parallel to the motion, or zero for both vector bosons. The polarization vectors for the vector bosons are indicated in Equation 4.2. . . . . 21
- 4.2 The distribution of the opening angle between the two  $Z \rightarrow \ell\bar{\ell}$  planes for the  $CP$  even and  $CP$  odd terms for a 200 GeV heavy resonance. The size of the coupling is chosen such that the black and the purple curves have the same normalization. The definition of the opening angle  $\phi$  is shown on the right. . . . . 23
- 5.1 Demonstration of the phase shift in the  $\phi$  distribution. Here we compare  $A_1^{ZZ}$  and  $A_3^{ZZ}$ , as well as interference between the two terms. When both terms are present, the interference term contributes and causes an overall phase shift in the  $\phi$  distribution. . . . . 25

- 5.2 Shape of the  $CP$  violating term in the case of  $H \rightarrow \ell\bar{\ell}\gamma$  as a function of the di-lepton mass and the decay angle of the leptons. In the left plot, the projection of the absolute value of the  $CP$  violating term is shown. As the contribution from forward decay and the backward decay of the lepton with respect to the  $Z$  boson direction cancel, the  $CP$  violating term does not contribute to the total cross section. This is better illustrated in the 2D distribution, where red is positive and blue is negative; green represents zero. The contribution is largest at the  $Z$  mass, and where the direction of the leptons is close to the direction of the  $Z$  boson. . . . . 28
- 6.1 Contributions from different sub-process in the standard model  $d\bar{d} \rightarrow 2e2\mu$  process, including interferences between the different terms. The red dots are events generated with MADGRAPH [36], to be compared with the total differential cross section shown in black. The different subprocesses shown include a t/u-channel  $d\bar{d} \rightarrow ZZ \rightarrow 2e2\mu$ ,  $d\bar{d} \rightarrow Z\gamma \rightarrow 2e2\mu$ ,  $d\bar{d} \rightarrow \gamma\gamma \rightarrow 2e2\mu$  and the s-channel resonant process  $d\bar{d} \rightarrow Z \rightarrow 2e2\mu$ . The dashed lines for the interference terms indicate a negative contribution. This comparison is done at leading order with the CTEQ6l1[34] PDF set. . . . . 31
- 7.1 The amount of  $pp$  collision data collected by the CMS detector as a function of time during Run I. The total integrated luminosity for  $pp$  collision data is  $23.3 \text{ fb}^{-1}$  at 8 TeV and  $6.1 \text{ fb}^{-1}$  at 7 TeV. . . . . 34
- 7.2 The peak instantaneous luminosity is shown in the top panel during Run I data taking in CMS, in 2010-2012. As shown, the peak intensity is progressively higher towards the end of data taking in 2012. This resulted in a higher number of concurrent collisions, as shown in the bottom panel, where up to an average of 34 interactions in each event were reached during one fill of the LHC. . . . . 35

7.3	Schematic view of the CMS detector. The collision point is on the left at the center of the inner circle. The produced particles travel through many subdetectors. From the inside: silicon tracking system, electromagnetic and hadron calorimeters, superconducting magnet, and finally the muon chambers and iron return yoke. . . . .	38
8.1	Summary of the measured Higgs signal strength ( $\mu \equiv \sigma/\sigma_{\text{SM}}$ ) from CMS, in different channels [35, 55, 58–61]. The results from all channels are compatible with the standard model prediction, as indicated by the dashed line. . . . .	41
8.2	Illustration of the lepton 3D impact parameter significance. The impact parameter is defined as the distance of the closest approach between the primary vertex and the lepton track. Depending on the track quality, we can define an uncertainty on the impact parameter. The analysis selects leptons based on the ratio between the impact parameter and its uncertainty. In other words, it is a test of the compatibility between the lepton and the reconstructed primary vertex. . . . .	44
8.3	The validation of the lepton calibration, showing the comparison of the simulation and the data to the known masses of the $Z$ , $\Upsilon$ , and $J/\Psi$ resonances. The dilepton resonances are reconstructed and the extracted masses serve as a measure of the accuracy of the calibration. In the kinematic ranges we are interested in, we observe a good agreement between simulation and data at the 0.3% level compared to the best known masses for the resonances, indicating that the calibration obtained by using the simulation successfully reproduces the calibration obtained with data. . . . .	44
8.4	The efficiency of the analysis selection for different final states as a function of the Higgs boson mass. For a standard model Higgs at 125 GeV, the overall selection efficiency including acceptance effects is 30%, 43% and 62% for the $4e$ , $2e2\mu$ and $4\mu$ final states respectively . . . . .	47

8.5	The validation of the reducible background estimation using the opposite-sign lepton method. The prediction of the reducible background is in agreement with the data, within uncertainties. . . . .	48
8.6	The $H \rightarrow 4\ell$ mass spectrum from the full CMS Run I dataset. The expectation is compared directly to data without a fit to extract the normalization from the data. A clear excess is observed at 125 GeV over the standard model background. The background modeling is good over the full $m_{4\ell}$ mass range, including the $Z$ resonance peak at $m_Z$ , as well as the $q\bar{q} \rightarrow Z\gamma \rightarrow 4\ell$ and $q\bar{q} \rightarrow ZZ \rightarrow 4\ell$ $t/u$ -channel contributions. . . . .	49
8.7	The limit on the $H \rightarrow 4\ell$ production cross section is shown. With the analysis with CMS Run I data, we are able to exclude a standard model Higgs-like resonance between 114.5 – 119.0 GeV and 129.5 – 832.0 GeV at 95% C.L. The limit on the production cross section at 125 GeV is 1.8 times the standard model prediction. . . . .	50
8.8	The one-sided local $p$ -value is shown as a function of different $m_H$ hypotheses. The $p$ -value for the data is defined as the probability of the background-only hypothesis fluctuating to mimic what we observe in the data. The points on the dashed line represent the expectation of the $p$ -value for a standard-model like signal hypothesis, with signal strength as predicted by the standard model. . . . .	51
8.9	The display of a $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ candidate event from the CMS Run I data taken in 2012. . . . .	52
9.1	Summary of the CMS Higgs mass measurement in the $4\ell$ final state. Curves of different colors represent the measurement done with only one final state lepton combination. Measurements from different final states agree with each other at the one sigma level. The combination of all channels is shown as the black line. . . . .	55

- 9.2 Summary of the CMS Higgs mass measurement in the  $\gamma\gamma$  final state. The green and yellow regions represent the  $1\sigma$  and  $2\sigma$  intervals respectively. There is a slight tension of  $1.6\sigma$  between the results from the  $4\ell$  channel and the  $\gamma\gamma$  channel. . . . . 56
- 9.3 Summary of the CMS Higgs mass measurement from the high resolution channels. The red line is the result from the  $4\ell$  channel, while the green line is the result from the  $\gamma\gamma$  channel. Assuming the two channels arise from the same resonance, the combination between two channels is shown in black. The combined mass measurement is  $125 \pm 0.3$  GeV. 57
- 9.4 Selected CMS results on measuring the Higgs couplings. On the left the combined result using all channels is shown, assuming a universal  $\kappa_f$  and a universal  $\kappa_V$ . The standard model point  $(1, 1)$  is indicated as the yellow diamond, which is in agreement with the combined  $1\sigma$  contour in gray. In the right plot the result from a combined fit floating  $\kappa_W, \kappa_Z, \kappa_t, \kappa_b$  and  $\kappa_\tau$  simultaneously is shown. The result is again compatible with the standard model expectation. . . . . 58
- 9.5 The ratio of the  $gg \rightarrow ZZ$  to the  $q\bar{q} \rightarrow ZZ$  background, as a function of  $m_{4\ell}$ . In the range relevant to the measurement of Higgs boson properties, the  $gg$  initial state continuum background is order of 4% compared to the  $q\bar{q}$  initial state. Compared to the analysis leading to the Higgs discovery, it is more relevant in analyses where the  $m_{4\ell} > 2m_Z$  region is important, for example the high-mass second Higgs search, and the measurement of the Higgs boson width. . . . . 60
- 9.6 The result on Higgs width measurement from interference between signal and background in the high  $m_{4\ell}$  region. The purple and the red lines indicate result from individual final state:  $4\ell$  and  $\ell\bar{\ell}\nu\bar{\nu}$  respectively. The combined result from both channels is shown in blue. The upper limit on Higgs width at 95% C.L. is  $5.4 \times \Gamma_H$ . . . . . 60

- 9.7 An example of signal and background template used in CMS analyses involving hypothesis testing on spin- $CP$  properties. This background discriminant is designed to separate signal and background. Standard model background components, shown in blue and green, favor a smaller value of the discriminant, while standard model signal in white favors a larger value. The expectation from Monte Carlo is compared with collected data. . . . . 62
- 10.1 Definition of the 12 variables. In the rest frame of each  $V(V')$  we define a decay angle  $\theta_1(\theta_2)$  of the lepton with respect to the  $V(V')$  direction, as shown in the upper-left figure. In the upper-right figure we show the definition of the  $\phi$  angle in the  $4\ell$  rest frame, which is the opening angle between two  $\ell\bar{\ell}$  planes. The dilepton masses are also indicated. Adding in the incoming partons, in the lower-left figure, we can define three additional angles:  $\Phi_1$ , the opening angle between one of the  $V \rightarrow \ell\bar{\ell}$  planes and the plane formed by the incoming parton and  $VV'$ ;  $\Theta$ , the decay angle of the  $VV'$  with respect to the incoming parton; and  $\phi_{\text{offset}}$ , the rotation along the incoming parton axis. Finally in the lab frame there is the 4-momentum of the  $4\ell$  system, which we parametrize as  $m_{4\ell}$ ,  $y_{4\ell}$  and  $\vec{p}_{T,4\ell}$ , as can be seen in the lower-right figure. . . . . 67
- 10.2 Ambiguity in assigning extra activities in the event with respect to the  $4\ell$  system. In the top plots we have two choices of assigning the extra activity (denoted collectively as “Jet”) on either side of the  $4\ell$  system. The choice changes the direction of parton lines attached to the  $4\ell$  system, and thus changes the values of some of the observables. One popular choice of assigning the extra activity is to imagine the vectorial sum of the extra activities are evenly split on two “sides”. . . . . 68

- 10.3 Toy study in the  $Z\gamma$  sector, to show that incomplete modeling of potential fluctuations can lead to artificial features in the result. Here the true model is the tree-level  $A_1^{ZZ}$  coupling, and the fluctuations are modeled by floating the  $A_2^{Z\gamma}$  and  $A_3^{Z\gamma}$  operators. In the case where the fluctuations are better described by  $A_3^{Z\gamma}$ , artificial features are seen when only  $A_2^{Z\gamma}$  is floated, since the model space probed cannot adequately sample the fluctuations. . . . . 71
- 11.1 Examples of the extracted smearing transfer functions plotted as  $c - 1$ , for muons and electrons with different momenta. On the top row muons with  $p_T = 15, 30, 45$  GeV and  $|\eta| = 0.6$  are shown, while in the bottom row the same is shown for electrons. The width of the transfer function for electrons with higher energy is smaller, relative to those with lower energy. . . . . 75
- 11.2 Width of the transfer function as a function of the lepton transverse momentum and pseudorapidity. On the left are the plots for electrons, and on the right are the muon results. The top row shows the 7 TeV transfer functions and the bottom row shows the transfer functions for 8 TeV. The binning used in the plots shown here is the same as the binning of the transfer functions used in the analysis. The color scale indicates the percentage relative to the lepton momenta. . . . . 76
- 11.3 The efficiency for the leptons as a function of the lepton transverse momenta and pseudo-rapidities. The results for 7 TeV are shown in the top row, and for 8 TeV are shown in the bottom row. The effect of the boundary between the barrel and endcap calorimeters is visible in the electron efficiency function on the left. For the muons, on the right, the efficiency is good over the full range of kinematics relevant to this analysis. 77
- 11.4 Comparison of the  $m_{4\ell}$  and  $m_1$  projections, showing the comparison between the fully simulated sample, and the smeared sample. . . . . 78

- 12.1 Examples of projections of differential cross sections for different pure  $CP$ -even models. The models shown include the standard model tree-level  $HZZ$  coupling, and the leading order loop-induced  $HVV'$  couplings. For each model  $m_1, m_2, \phi$  and  $\cos \theta_2$  are plotted, showing the distinctly different shapes predicted by the different models. . . . . 81
- 12.2 Extracted spectrum for the signal and the background for the  $2e2\mu$  final state. In the first two panels the distribution with  $m_{4\ell} = 125$  is plotted, comparing among the different components of the signal and background. In the rapidity distribution, the initial state is directional. The difference in the initial state causes a shift in the final state distributions. In the bottom left panel the  $m_{4\ell}$  distribution is shown for the background components. For a standard model Higgs signal, it is almost a delta function and is not shown here. In the last panel, the 2D distribution of  $p_T$  and  $y_{4\ell}$  for the standard model Higgs is shown. . . . 83
- 12.3 Examples of diagrams not accounted for in the calculation of SM backgrounds, at higher order in QCD. Among these, the upper left diagram (and diagrams related by crossing symmetry) does not contain a loop, and is therefore expected to contribute more than the loop diagrams. . . 84
- 12.4 Decomposition of the different initial state contributions in  $q\bar{q} \rightarrow$  processes, as a function of the  $4\ell$  system rapidity. In the left plot, an immediate  $q$  is coming from the negative side, and a  $\bar{q}$  is coming from the right. It is symmetrized in the right plot, which shows a more realistic view of the contribution, as we do not know which proton gives rise to which parton in the experiment. . . . . 84

12.5	Estimation of the precision of the calculation of the convolution integral normalization (see Equation 12.6). Since a Monte Carlo integration method is employed, we can estimate the precision by grouping individual samples into sub-samples of different sample size, and use the RMS as a measure of the precision at any given sample size. This is repeated many times and a $1/\sqrt{N}$ curve is fitted to the result. It allows us to estimate the precision for a large sample size, where there are insufficient statistics to evaluate the RMS spread with confidence. The lack of statistics is more pronounced at larger sample sizes, where we see a larger spread in the RMS values. . . . .	88
12.6	Fake rate for jets faking leptons as a function of $p_T$ and $\eta$ . The top row shows the fake rates for 7 TeV, and the bottom row shows the rates for 8 TeV data. The left-hand plots are the rates of jets faking an electron, and the right-hand plots are the corresponding rates for muons. The binning in the plots reflects the actual binning used in the analysis. . .	91
12.7	Geometrical visualization of linear interpolation in multiple dimensions. Suppose there is a value attached to each of the 8 vertices, and we would like to linearly interpret the values for a point $P$ in the center. The interpolated value is a weighted sum of all the vertices. The weight of each vertex can be calculated as the ratio between the volume of the smaller cube away from vertex, and the total volume. . . . .	92
12.8	Example projection of different check points . . . . .	94
13.1	The Jacobian factor can be thought as the ratio of an infinitesimal volume around a given point of interest between two different bases. Lines on the left correspond to lines on the right in a different basis. The volume $V$ is translated into $V'$ on the right. The Jacobian factor at that point is therefore $J = V'/V$ . . . . .	102

- 13.2 Validation of the 12D Jacobian calculation. Events are generated uniformly in both bases, and compared to each other by weighting those from the lepton basis by their respective 12x12 Jacobian factors. The small differences are due to fluctuations. We see excellent agreement for all the variables of interest. . . . . 103
- 13.3 Contour of  $m_+$  (left) and  $m_-$  (right) for one example event as a function of  $m_1$  and  $m_2$ . The contour for  $m_+$  changes from event to event, but is always roughly diagonal.  $R_m$  in this case is 3.38 (see Equation 13.13). . . . . 107
- 13.4 Demonstration of the mass grid attractor. Different sets of grid points are plotted in each line with varying mass grid attractor strength. With the attractor strength set to zero, we recover uniform spacing, while large values of the attractor strength cause points to be concentrated in the center of the line. . . . . 108
- 13.5 Demonstration of the mass grid attractor and the scan direction in the integration. Each bin is one grid point, and the color indicates the value of the inner integral. In the top two plots the attractor strength is set to zero. In the middle, a moderate attractor strength is turned on, and in the bottom plots the center grid point optimization is turned on. The left column is without diagonal scanning direction, and the right column is with the modified  $m_+$  and  $m_-$ . By using all optimizations, there are more grid points concentrated in the core, increasing the quality of integration, as shown in the bottom right plot. . . . . 111
- 13.6 Illustration of the recursive integration. In each plot, the line shows the true function, and the dots are points evaluated from the integration. In the left panel we perform the integration on a second-order polynomial  $f(x) = \frac{1}{75}x^2$ , where the method is exact; therefore the grid spacings are always the same. In the right panel the integrand is modified to be  $f(x) = \frac{1}{75}x^2 + e^{-\frac{1}{2}x^2}$ . We see that in places where the functions are rapidly varying, more points are used. . . . . 115

- 13.7 Demonstration of effect on the contribution as a function of dilepton mass due to the narrow-width approximation. In both plots the inner integral as a function of  $m_1$  and  $m_2$  is plotted for the exact same event. On the left we do not enforce the narrow-width approximation, while on the right we require that  $m_{4\ell} = 125$ . The shape is strongly elongated for the case with the narrow-width approximation. . . . . 118
- 13.8 Examples of 6  $c_1 - c_3$  ellipses for a few mass grid points on an example event are shown in the left panel, where each contour represents one mass grid point. Dots represent possible solutions to all mass constraints. For each grid point, all possible solutions form an ellipsoidal shape, as expected. In the right we show the slicing of the ellipse while doing the integral. The four extreme points (left, right, up, down) are calculated, and the curve is divided into four regions by the red lines going vertically between the extreme points. In order to avoid artificial infinities during the integration, we integrate along the horizontal direction for the two segments in the middle, and vertically up along the  $c_3$  direction for the two segments on the side, as shown by the blue arrows. . . . . 118
- 13.9 Comparison of the projections on the 12 observables for the standard model signal case, between toys generated from the reconstructed level *pdf* and the smeared sample using the same input transfer function. This serves as a cross check of the integration method. . . . . 121
- 13.10 Comparison of the projections on the 12 observables for the standard model  $q\bar{q} \rightarrow 4\ell$  background, between the reconstructed *pdf* and the smeared sample. The goal of this comparison is to cross check the background integration method. . . . . 122
- 13.11 Cross check on the projections on 12 observables, between toys generated from reconstructed level *pdf* and smeared sample on a “tofu” scenario. The sharp edges present in the generator level *pdf* in this case allows us to investigate and validate the effect of smearing due to detector effects in the integration process. . . . . 123

- 13.12 Comparison of the projections on the 8 decay observables, between toys generated from reconstructed level *pdf* and a CMS full simulation sample for the SM Higgs ( $A_1^{ZZ} = 2$ ). In addition to validating the integration process, this also validates the input transfer function. . . . . 124
- 13.13 Comparison of the projections on the 8 decay observables, between toys generated from reconstructed level *pdf*, and a CMS full simulation sample for an exotic signal model ( $A_1^{ZZ} = 2, A_2^{ZZ} = 6.552$ ), with mixing between the tree-level  $A_1^{ZZ}$  term and the higher-order *CP*-even coupling  $A_2^{ZZ}$  term. . . . . 125
- 13.14 Comparison of projections on the 8 decay observables, between toys generated from reconstructed level *pdf*, and a CMS full simulation sample for an exotic signal model ( $A_1^{ZZ} = 2, A_3^{ZZ} = 10.084$ ), with mixing between the tree-level  $A_1^{ZZ}$  term and the higher-order *CP*-odd coupling  $A_3^{ZZ}$  term. . . . . 126
- 13.15 Fit results for toys generated from the reconstructed level *pdf* with the standard model. In each plot the distributions of the fit result for different dataset sizes are shown. With larger dataset sizes, the distribution of the fit result converges to the true value, indicated by the black line. On the top row the results for *CP*-even operators for  $ZZ$ ,  $Z\gamma$  and  $\gamma\gamma$ , respectively are shown. In the second row the results for *CP*-odd operators are shown. Finally we show the fitted background fraction in the bottom plot. The color in each bin represents the number of pseudoexperiments. . . . . 127

- 13.16 Fit results for toys generated from the reconstructed level *pdf*, for an exotic model with  $A_1^{ZZ} = 2$ ,  $A_2^{ZZ} = 7.7$  and  $A_3^{ZZ} = 10.2$ . In each plot the distributions of the fit result for different dataset sizes are shown. With larger dataset sizes, the distribution of the fit result converges to the true value, indicated by the black line. On the top row the results for *CP*-even operators for  $ZZ$ ,  $Z\gamma$  and  $\gamma\gamma$ , respectively are shown. In the second row the results for *CP*-odd operators are shown. We show the fitted background fraction in the bottom plot. . . . . 128
- 13.17 Comparison of the fit results between different frameworks using a set of pseudodatasets formed from CMS fully simulated samples, with statistics similar to that of the CMS Run I data. On the left 1D distribution for the fitted  $f_{a3}$  for this ensemble of datasets is shown. The red histogram is the result of the template fitting method, and the blue is from the method described in this thesis. On the right a scatter plot for the fitted result is shown. Each point represents one pseudoexperiments, with the result from method in this thesis in plotted as the  $x$ -axis and fit result from the template fitting method as the  $y$ -axis. We observe a statistically reasonable spread of the fitted result, as these are two different approaches with different event selections. The population in the anti-diagonal line is from the pseudodatasets with near-degeneracy between the positive and negative values of the coupling parameter  $f_{a3}$ , or  $A_3^{ZZ}$ . The only term to break the degeneracy is the interference term between  $A_1^{ZZ}$  and  $A_3^{ZZ}$ , and with the CMS Run I statistics we are not yet sensitive to the interference terms in some cases. . . . . 129

- 14.1 The projections of the 8 decay observables of the data on top of the predictions from a Monte Carlo sample for the irreducible  $q\bar{q} \rightarrow 4\ell$  background, and the prediction from the data for the reducible fake background. The standard model Higgs prediction at 125 GeV is also shown as the white histogram. The observed data is in agreement with the prediction for the total of the signal and the background. . . . . 133
- 14.2 The expected (dashed) and observed (solid) likelihood scan on  $A_2^{ZZ}$  and  $A_3^{ZZ}$  for two different parametrizations. On the top the ratios of  $A_2^{ZZ}$  and  $A_3^{ZZ}$  to  $A_1^{ZZ}$  are plotted, and on the bottom the  $f_{a2}$  and  $f_{a3}$  are shown. They contain the exact same information, only the axes are different. There are some interesting features explained in text. . . . . 134
- 14.3 The comparison of the projection on the  $m_2$  and  $\phi$  observables, between the standard model Higgs signal and exotic signals with non-zero  $A_2^{ZZ}$ . The shape of standard model signal (black) is more similar in shape to that of an exotic model with large  $A_2^{ZZ}$  (green), while an intermediate model (red, purple) looks very different from the other two models due to interference effects. This is the main reason why there is a region with a large likelihood differences in the likelihood scan along the  $A_2^{ZZ}/A_1^{ZZ}$  axis. . . . . 135
- 14.4 The overlay of the likelihood scans for 1000 toy pseudoexperiments are shown in the top plots for  $f_{a2}$  and  $f_{a3}$  respectively. The color scale indicates the density of likelihood scans from the toy pseudoexperiments. The observed data is shown as a solid curve, and it is in a region where the pseudoexperiment likelihood scans are relatively dense, indicating that the height of the observed likelihood scan is common. In the bottom plots we also show the percentage of the toys that lie below the observed likelihood curve for each model point. For all the exotic signal model points (where  $f_{a2}$  and  $f_{a3}$  are not zero) we do not see any extreme  $p$ -values. 136

- 14.5 Evaluation of how probable it is for the observed data from pseudoexperiments. A set of 1000 pseudoexperiments with statistics similar to data is generated and fitted, and we compare the result from the data to the distribution of fit results in the top row for  $A_2^{ZZ}/A_1^{ZZ}$  and  $A_3^{ZZ}/A_1^{ZZ}$ . The one-sided  $p$ -values for the two cases are 45% and 50% for  $A_2^{ZZ}$  and  $A_3^{ZZ}$  respectively. Since the fit result is very close to zero, we also investigate the probability of closeness to zero by looking at the absolute value of the log of these parameters. A semi-log scale is plotted in order to better show the details close to zero. The  $p$ -values in this case are 11% and 2% for  $\log|A_2^{ZZ}/A_1^{ZZ}|$  and  $\log|A_3^{ZZ}/A_1^{ZZ}|$  respectively. . . . . 137
- 15.1 The summary of the partial widths from different terms. The partial width is defined as in Equation 15.2. The value in each bin is the size of the cross term between the two coupling parameters labeled on the  $x$  and the  $y$  axes. All coupling parameters are set to the same value, and are normalized to the standard model tree-level term. The terms involving photons are much larger primarily due to the larger phase space. 141
- 15.2 The summary of the absolute partial widths from different terms, as defined in Equation 15.2. The value in each bin is the size of the cross term between the two coupling parameters labeled on the  $x$  and the  $y$  axes. All coupling parameters are set to the same value and are normalized to the standard model tree-level term. Similar to the case of partial width in Figure 15.1, terms involving photon couplings are much larger because of phase space effect. The diagonal terms are the square terms for each coupling parameter, and since it is always positive, the values are identical to those in the case of partial width in Figure 15.1. . . . . 142

- 15.3 The summary of the partial widths, assuming standard model-like couplings. The value for  $A_1^{ZZ}$  is set to 2, and  $A_2^{ZZ}$ ,  $A_2^{Z\gamma}$  and  $A_2^{\gamma\gamma}$  are set to 0.005, 0.014 and  $-0.008$  respectively. The overall normalization is again chosen such that the  $A_1^{ZZ} \times A_1^{ZZ}$  term is one. In terms of the total  $H \rightarrow 4\ell$  cross section, the most important terms other than the tree-level  $A_1^{ZZ} \times A_1^{ZZ}$  are  $A_2^{Z\gamma} \times A_1^{ZZ}$ ,  $A_2^{Z\gamma} \times A_2^{Z\gamma}$  and  $A_2^{ZZ} \times A_1^{ZZ}$ . Again, when the sensitivity to different couplings is concerned, this only gives partial information, since the shape information is lost in these partial width numbers. . . . . 143
- 15.4 Projection of the precision for the coupling parameters, as a function of the dataset size. On the top axis the equivalent luminosity at 14 TeV is shown. All couplings are assumed real in this projection, and all are floated at the same time. A hierarchy of sensitivity  $\gamma\gamma > Z\gamma > ZZ$  is observed. . . . . 147
- 15.5 The distribution of the extracted  $A_2^{\gamma\gamma}$  parameter with all parameters floating, as a function of the dataset size. The true model used to generate datasets is  $A_1^{ZZ} = 2$ ,  $A_2^{\gamma\gamma} = -0.008$  and all other couplings set to zero. The distribution of the fitted  $A_2^{\gamma\gamma}$  parameter converges to the true value as the dataset size increases. . . . . 148
- 15.6 Constraints on the  $\gamma\gamma$  couplings from different measurements, with  $3000 \text{ fb}^{-1}$  of LHC data at 14 TeV. In blue, the 68% and 95% C.L. levels from the  $4\ell$  measurement is shown, while in magenta the projection of the sensitivity from the direct  $H \rightarrow \gamma\gamma$  measurement is plotted. The constraints from the electron dipole moment is shown in green. Note that the constraint from the electron dipole moment measurement is strongly dependent on the assumptions on first-generation Yukawa couplings and other potential new physics, on which we won't have measurements in the foreseeable future. The standard-model value is indicated by the star marker. . . . . 149

15.7	Constraints on the $\gamma\gamma$ couplings from different measurements, with $300 \text{ fb}^{-1}$ of LHC data at 14 TeV. In blue, the 68% and 95% C.L. levels from the $4\ell$ measurement is shown, while in magenta the projection of the sensitivity from the direct $H \rightarrow \gamma\gamma$ measurement is plotted. The constraints from the electron dipole moment is shown in green. Note that the constraint from the electron dipole moment measurement is strongly dependent on the assumptions on first-generation Yukawa couplings and other potential new physics, on which we won't have measurements in the foreseeable future. The standard-model value is indicated by the star marker. . . . .	150
A.1	Relevant diagrams in the measurement of SM parameter from $q\bar{q} \rightarrow 4\ell$ events. . . . .	155
A.2	Extraction of the standard model parameters using only decay information from $q\bar{q} \rightarrow 4\ell$ events. On the left we show the distribution of extracted Z coupling strength modifier as a function of dataset size. The color in the figures indicate number of pseudoexperiments. When the dataset size is small, there is near-degeneracy between positive values and negative values of the coupling strength modifier. The only way to break the degeneracy is through the interferences between $s$ -channel and $t/u$ -channel diagrams, and with small dataset sizes we are not sensitive to the interference terms. On the right where we show the extracted up quark change, there is a similar feature at small dataset sizes. . . . .	157
B.1	Example of flat noise: (left) the pulse shape in time; (right) an event display as a function of $\eta$ and $\phi$ showing that several channels are affected. Each bin corresponds to a single HCAL channel. The height of the histogram indicates deposited energy in the channel. . . . .	160

B.2	Example of triangle noise: (left) pulse shape in time; (right) event display of the same event showing the collective behavior of this type of noise. Each bin represents the reconstructed energy in a single channel in GeV. . . . .	161
B.3	Example of spike-like noise: (left) pulse shape; (right) the event display of an example event containing a spike-like noise. This type of noise usually affects only small number of channels. . . . .	162
B.4	Two examples of events containing noise. The left plots are the energy deposits as a function of $\eta$ and $\phi$ of the HCAL tower. Each bin represents one single HCAL tower. On the right is the energy measured in each of the pixels in the corresponding problematic HPD. As seen in these examples, noise-like energy deposits which are randomized in the physical $\eta - \phi$ space tend to be neighbors in the HPD pixels. . . . .	163
B.5	Ideal HCAL pulse shape as measured and parametrized during the test beam measurements. . . . .	165
B.6	Fit-based filter discriminants for a typical noisy dataset on the right, and a dataset with less noise on the left. . . . .	166
B.7	The performance of the fit-based filter, comparing the missing transverse energy (MET) after the filter to before the filter, for several types of events. . . . .	167
B.8	Distributions of R45 variable used in the filter are shown on the top for a dataset containing noise (right), and a dataset with less noise (left). In the bottom plots the HPD count distribution between a noise-enriched dataset (red) and a clean dataset (black) is shown. . . . .	168
B.9	An example of time sample ratio functions. . . . .	170
B.10	Corrected energy distribution after the correction (black), compared to that before the correction (red). The mean of the energy from pileup after correction shifts back to zero. . . . .	170

B.11	Demonstration of the negative energy filter by looking at reconstructed energy in different time samples. In the SingleMu dataset, which is clean of noise, there is no significant negative energy tail. On the other hand, in the MET dataset enriched in noise, due to the presence of spike-like noises, a large negative energy tail is seen. A large negatively reconstructed energy indicate that the pulse shape differs greatly compared to the expectation, which is a sign of noise. . . . .	171
C.1	Test beam layout and detector setup. The beam comes from the left, and passes through the wire chambers and triggering scintillators before hitting the HCAL tower. There is a lead absorber and a muon chamber behind the HCAL tower. . . . .	173
C.2	Wire chamber alignment. Left: before alignment, and right: after alignment. The coordinates from the wire chambers are measured in units of mm. By this alignment procedure, we are able to align the wire chambers to an accuracy of $O(1)$ mm. . . . .	176
C.3	The concentration of energy deposits in the middle tower, for a run with a muon beam. The $y$ -axis shows the energy concentration in the middle tower, defined as the energy measured in the middel towel divided by the total measured energy in all three towers with 16-layer readout. It is plotted with respect to the $x$ -coordinate of the extrapolated hit position. For muons with hit position above 8 mm, the majority of energy is deposited in the neighbor tower due to an misalignment of the HCAL modules with beam direction. . . . .	177

- C.4 The linear correlation coefficients of pedestal values in different channels in the three towers with 16-layer readout. The  $x$  and  $y$  axes are the channel index for each individual channel (48 in total). Before the pedestal subtraction procedure, a correlation is seen in the pedestal values for the different channels (left). The correlation is reduced a lot after the pedestal subtraction, showing a maximum of correlation coefficient of 0.1 (right) . . . . . 177
- C.5 The fit to muon calibration runs to extract the minimum ionizing potential (MIP). We use a Landau function convoluted with a Gaussian as a model to fit the spectrum. The median value of the function is then used as the measurement of the MIP for the channel. . . . . 178
- C.6 The unwinding procedure. In the top row, the ADC value distributions of two pion runs with the same beam energy are shown. On the left is the distribution with pre-amplifier on, and on the right is the corresponding spectrum with the pre-amplifier off. We observe a clear saturation effect in the spectrum with pre-amplifier on. By comparing the two distributions and matching the corresponding ADC values with the same *cdf* value, we were able to obtain a mapping between pre-amplifier on ADC value and the corresponding pre-amplifier off ADC value. This mapping is then used to “unwind” the spectrum with pre-amplifier on for the energy measurement. One such mapping function is shown in the bottom row. There is a long linear regime when the ADC value is small, and the saturation effect is seen with larger ADC values. . . . . 179

- C.7 Lepton contamination studies. On the left the energy concentration in the middle tower is shown for a run with pion beam. Since the hadron shower is not contained in a single HCAL tower, the energy concentration is in general not 100%. There is, however, a population with 100% energy concentration in the pion data. It is compatible with expectations from electrons. The longitudinal profile of the events with 100% energy concentration in the middle tower is shown on the right. We see a shallower profile compared to an average hadron shower. . . . 180
- C.8 The HCAL longitudinal profile. On the left the longitudinal profiles for pions with different energies are shown. There is a dead channel (layer 12) in the readout. With the increasing energy, position of maximum shower is deeper. The average depth of the shower is shown on the right as a function of energy. With pion beams of energy above 100 GeV, the average depth deviates from the expected logarithmic dependence, which could be due to the worse shower containment at large beam energies. . . . . 181
- C.9 Measurement of the HCAL interaction length. The position of the first energy deposit is plotted, showing an exponential dependence above the initial rise. We can fit the exponential constant and obtain a measurement of the interaction length in HCAL. It is compatible with the expectation of brass material interleaved with plastic scintillators. . . . 181
- C.10 The comparison of HCAL precision with full layer readout with an earlier test beam (2006) studies with CMS Run I setup. We observe an improvement in resolution with full layer readout. . . . . 182
- D.1 Some of the lower order Feynman diagrams for  $Z$ +jets. With each additional jet, there is an extra QCD vertex, and we expect the rate to decrease by roughly a factor of  $\alpha_S$ . . . . . 185
- D.2 The distribution of jet multiplicities in  $Z$ +jet events; comparing the data to a Monte Carlo simulation. . . . . 186

D.3	The results on the $Z (\rightarrow \mu^- \mu^+) + \text{jet}$ differential cross section measurement. On the left the differential cross section is shown, comparing to Monte Carlo simulations. The ratio $R_N^Z$ is also shown in the bottom. On the right we report the measurement on the $\alpha$ and $\beta$ parameters (Equation D.2). The $1\text{-}\sigma$ ellipse of data (green) is in agreement with the MADGRAPH simulation (red). . . . .	188
E.1	Feynman diagram for LQ and sbottom . . . . .	190
E.2	Scaling behavior of the razor variables. The exponential tail in $M_R$ drops faster with tighter cut on $R^2$ variable (top-left), and the extracted slope is consistent with a linear function on $R^2$ (top-right). The converse is also true: the tail in $R^2$ drops faster with tighter cut on $M_R$ (bottom-left), and the exponential constant is again modeled by a simple linear function of $M_R$ (bottom-right). The slopes on the linear dependence of exponential constants in $M_R$ and $R^2$ agree with each other within numerical precision. . . . .	192
E.3	Demonstration of the shape extraction procedure for different processes. On the top we fit $V + \text{jet}$ Monte Carlo samples to extract the $M_R$ shape as a function of different $R^2$ cuts. This extracted shape is then used in the fits to extract the shape of the $t\bar{t}$ process, as shown in the middle row. In the selection with a tight muon, the data contains almost exclusively events from $t\bar{t}$ and $V + \text{jet}$ processes. By fitting to the data with a tight muon and fixing the shape of the $V + \text{jet}$ sample, we are able to extract the shape of the $t\bar{t}$ contribution. The extracted shapes for the $t\bar{t}$ and $V + \text{jet}$ are then used to fit in the data containing a loose muon, where the multijet contribution is greatly enriched, as shown in the bottom row. On the left the fits on $M_R$ shape as a function of $R^2$ cut is shown, and on the right only the inclusive $R^2$ bin is plotted, but broken down to different components. . . . .	196

- E.4 Validation of the background prediction process using **ELE** sample. The extracted  $M_R$  shape from the muon samples are used to predict the  $M_R$  shape in a signal-like region in the **ELE** sample. Since we do not expect contribution from LQ signal in the **ELE** sample, it can be served as a good cross check of the background prediction process. We see good agreement for both the shape as a function of  $M_R$  and the overall yield. 197
- E.5 Comparison of the predictions and the data in the signal region, as a function of the razor variables  $M_R$  and  $R^2$ . The total prediction is in good agreement with data in both variables, and we did not see any significant excess. The distribution for two excluded models are shown as the dotted and dashed lines. . . . . 198
- E.6 LQ exclusion results from CMS, compared to the previous results from the D0 experiment at the Tevatron. The limit on production cross section (left) is shown as a function of the LQ mass assuming 100% branching fractions to  $b\nu_\tau$ . The limit on mass is reinterpreted as a function of different branching fractions (right). The result greatly extends the previous limit from the Tevatron, shown as the grey shaded area. . . . . 198
- E.7 Exclusion of a simplified model where sbottoms are pair-produced and each decays to a  $b$  quark and a neutralino ( $\tilde{\chi}^0$ , LSP) from CMS compared to the previous results from the CDF and D0 experiments at the Tevatron. 199

# List of Tables

2.1	Fundamental particles in the Standard Model. The fermions which make up normal matter span over 10 orders of magnitude in mass, from the current limit on neutrino mass of less than an electron volt, up to the heaviest known (top) quark of 173 GeV. The force carrier bosons with spin 1 include gluons for the strong force, photon for the electromagnetic force, and the $Z$ and $W$ bosons for the weak force. The final spin-0 Higgs particle, discovered in 2012 at the LHC, is responsible for the so-called Higgs mechanism that generates mass. . . . .	7
8.1	Summary of the $H$ observation status from CMS [15, 35, 53–55, 58–65] and ATLAS [16, 56, 57, 66], expressed in terms of the significance obtained in each channel. The numbers shown in the last two lines of the table, where the statistics are insufficient to measure the production rate, are the 95% CL upper limits. . . . .	40
12.1	Expected yields and the corresponding percentage uncertainties, for each contribution to the Higgs boson signal and to the backgrounds. . . . .	95
C.1	Summary of the data taken in the test beam. The first 6 runs are $\mu$ calibration runs for towers with the same setup as CMS Run I, taken by other groups. . . . .	174
E.1	Summary of samples used in the analysis . . . . .	195

E.2	Summary of the predicted and observed yields in the signal region. The uncertainty on the expected yield includes both statistical and systematic uncertainties. . . . .	197
-----	--	-----

# Chapter 1

## Preface

In the early 1960's a number of breakthrough developments and intellectual leaps occurred which allowed the application of condensed matter ideas about spontaneous symmetry breaking [1–3] to particle physics.

At first it was thought that the most suitable application of spontaneous symmetry breaking to particle physics was in the area of strong interactions [4, 5].

Only in 1967 did Weinberg [6] and, independently, Salam [7] realize that the Higgs mechanism [8–11] offered an elegant explanation of the unified structure of the electromagnetic and weak interactions. In their model, which is now the electroweak portion of the standard model, four Higgs fields are related by a gauge symmetry of the type introduced by Yang and Mills [12]. Three Goldstone bosons are “eaten” to give large masses to the  $W^+$ ,  $W^-$ , and  $Z$  bosons that mediate the weak interactions. An added bonus, not foreseen, is that the Higgs field also gives mass to quarks and leptons, the elementary fermions that make up matter. The mass of the Higgs boson is not predicted, but the interactions of the Higgs with other elementary particles can be precisely computed as a function of its mass and the masses of the other particles.

A very extended experimental program targeting the search for the Higgs particle was carried out in particle colliders over the past 30 years. The large electron-positron collider (LEP) hosted in the tunnel where the LHC is currently operating, ran at center of mass energies between 189 and 209 GeV, and produced a lower bound on the Higgs boson mass of 114.4 GeV/ $c^2$  [13].

The Tevatron collected close to 10 fb<sup>-1</sup> of  $p\bar{p}$  collisions at 1.98 TeV, and the exper-

iments CDF and D0 were able to exclude a mass range between 147 and 179  $\text{GeV}/c^2$  by early 2012 [14].

The LHC experiments, ATLAS and CMS with  $pp$  collision data collected at 7 and 8 TeV, began to observe an excess of events at around 120-130  $\text{GeV}/c^2$  in two of the main Higgs decay channels to bosons. On July 4, 2012, the CMS and the ATLAS experiments announced jointly the discovery of a Higgs-like boson at 125  $\text{GeV}/c^2$  [15, 16]. The discovery marks an important milestone in the 40-year long search for this scalar particle.

While concluding my publication of the search analysis looking for 3rd generation leptoquarks [17] around the time of the Higgs discovery, I decided that I wanted to focus the rest of my thesis research on the characterization of the newly discovered particle. At the time of discovery, the methods used to characterize the properties of the discovered particle were based mainly on pair-wise testing between models. Following work from De Rujula et al. [18], I started by implementing a multi-dimensional template method, whereby we attempt to capture the potential correlations among all observables by making a large binned template from simulated samples, and we study if this approach enhances the analyzing power relative to the existing kinematic discriminant method where many observables are combined into one discriminant.

A kinematic discriminant is usually specialized to a model one wants to probe. At an early stage of a discovery this approach is useful to exclude a number of theoretical models without being able to accurately determine the model to which the data is pointing. Extracting simultaneously multiple parameters with this method is suboptimal, and can sometimes produce misleading results. This motivates the multi-dimensional template approach, which is built based on all kinematic observables and which thereby presents the optimal precision measurement framework.

In the four lepton analysis there are 12 available observables with non-trivial correlations, and a high-dimensional template is therefore constructed as the basis of an optimized analysis. By using a high dimensional template based on all observables, constructed with simulated events, all information including correlations between the observables is retained in the analysis. However, in the course of my analysis, it soon

became obvious that the coarse binning used in previous studies could not satisfactorily carry all the information available in the  $4\ell$  channel, and that simply increasing the number of bins to deal with this issue becomes computationally intractable.

Upon examining the multi-dimensional method carefully I realized that in order to perform fits to the observed data, we do not need to keep all the information in the full multi-dimensional template. The likelihood used in the fit is evaluated from the multi-dimensional template but only for the observed data. Therefore if there is a way to calculate the likelihood directly, without storing the full template as an intermediate step, we can overcome the computational obstacle.

I then showed that it is indeed possible if we do not use the simulated sample as a means to fill the template and then extract likelihood from the template. We can instead perform a convolution integral from the differential cross section itself, combined with a functional modeling of the detector response (a so-called transfer function).

There are two key ingredients in the calculation. One is the analytic expression of the differential cross section, and the other is the 12 dimensional convolution integral to incorporate detector effects. The calculation of the fully analytic differential cross section was the subject of the thesis of theory graduate student and collaborator Dr. Vega-Morales [19]. Independently I addressed the convolution integral. During the summer of 2013 we worked out the last details of the numerical integration, and we obtained the continuous likelihood.

This new approach has since changed the overall picture of Higgs properties measurements. Instead of measuring a few parameters at a time, we are for the first time able to measure all potential coupling parameters simultaneously, without loss of information due to either binning effects or the choice of kinematic discriminants.

The main focus of this thesis is to develop and document this new approach to Higgs property measurements. We first give a short summary of the standard model and the Higgs mechanism in Chapter 2, and in Chapter 3 we focus on the coupling parameters that we can probe in the  $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$  channel. Chapter 4 gives a brief overview of the helicity amplitude formulation in the case where both

intermediate vector bosons are on-shell. The  $CP$  violating observables in this channel are summarized in Chapter 5. The background processes relevant to the properties measurements are discussed in Chapter 6.

A description of the Large Hadron Collider (LHC) and the Compact Muon Solenoid (CMS) experiment is provided in Chapter 7, leading up to the description of the  $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$  analysis in Chapter 8. All current measurements of the Higgs properties are summarized in Chapter 9.

In Chapter 10 we outline the challenges and opportunities in fully utilizing all the information available in the  $4\ell$  channel. A custom fast detector parametrization is described in Chapter 11. The convolution integral is described in Chapter 12, and the details in the implementation are described in Chapter 13. The result of my analysis using the data collected by CMS is shown in Chapter 14. Finally in Chapter 15 we discuss the future potential in the  $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$  channel, with this approach and more data.

## Part I

# Higgs and the Golden Channel

## Chapter 2

# The Standard Model And The Higgs Mechanism

### 2.1 Particle And Force Content

The standard model (SM) includes three families of quarks and leptons, and four gauge bosons as the force carriers. The quarks in each family are categorized into up-type and down-type. From lightest to heaviest the up-type quarks are the up quark (u), the charm quark (c) and the top quark (t), each of which carries  $+2/3e$  charge. The down-type quarks consist of the down quark (d), strange quark (s) and bottom quark (b), each carrying charge  $-1/3e$ . The leptons include the charged leptons: electrons (e), muons ( $\mu$ ), and taus ( $\tau$ ), and their neutral counterparts called the neutrinos ( $\nu$ ): electron neutrinos ( $\nu_e$ ), muon neutrinos ( $\nu_\mu$ ), and tau neutrinos ( $\nu_\tau$ ). Stable normal matter is made up of the up and down quarks, and electrons.

In addition to matter, there are 4 fundamental forces. The strong force which holds quarks together inside nuclei is the strongest of these, and is mediated by gluons (g). The strong force charges are carried by quarks and gluons, while leptons are neutral to the strong force. The electromagnetic force, mediated by photons ( $\gamma$ ), is responsible for all phenomena involving electric charges and electromagnetic waves. Yet weaker is the weak force, mediated by the  $W^+$ ,  $W^-$  and  $Z^0$  bosons. The weak force can explain for example the radioactive decays of unstable atoms. The weakest of all forces is gravity, hypothesized to be mediated by gravitons ( $G$ ). Gravity is weak

Type	Name	Mass	Spin	Electric charge	Weak isospin	Weak hypercharge
Quark	up (u)	2.3 MeV/c <sup>2</sup>	1/2	2/3e	+1/2 (0)	+1/3 (+4/3)
	down (d)	4.8 MeV/c <sup>2</sup>	1/2	-1/3e	-1/2 (0)	+1/3 (-2/3)
	charm (c)	1.275 GeV/c <sup>2</sup>	1/2	2/3e	+1/2 (0)	+1/3 (+4/3)
	strange (s)	95 MeV/c <sup>2</sup>	1/2	-1/3e	-1/2 (0)	+1/3 (-2/3)
	top (t)	173.07 GeV/c <sup>2</sup>	1/2	2/3e	+1/2 (0)	+1/3 (+4/3)
	bottom (b)	4.18 GeV/c <sup>2</sup>	1/2	-1/3e	-1/2 (0)	+1/3 (-2/3)
Lepton	electron (e)	511 keV/c <sup>2</sup>	1/2	-1e	-1/2 (0)	-1 (-2)
	electron neutrino ( $\nu_e$ )	< 2.2 eV/c <sup>2</sup>	1/2	0	+1/2 (-)	-1 (-)
	muon ( $\mu$ )	105.7 MeV/c <sup>2</sup>	1/2	-1e	-1/2 (0)	-1 (-2)
	muon neutrino ( $\nu_\mu$ )	< 0.17 MeV/c <sup>2</sup>	1/2	0	+1/2 (-)	-1 (-)
	tau ( $\tau$ )	1.777 GeV/c <sup>2</sup>	1/2	-1e	-1/2 (0)	-1 (-2)
	tau neutrino ( $\nu_\tau$ )	< 15.5 MeV/c <sup>2</sup>	1/2	0	+1/2 (-)	-1 (-)
Boson	gluon (g)	0	1	0		
	photon ( $\gamma$ )	0	1	0		
	Z (Z)	91.19 GeV/c <sup>2</sup>	1	0		
	W (W)	80.4 GeV/c <sup>2</sup>	1	$\pm 1e$		
	Higgs (H)	$\sim 125$ GeV/c <sup>2</sup>	0	0		

Table 2.1: Fundamental particles in the Standard Model. The fermions which make up normal matter span over 10 orders of magnitude in mass, from the current limit on neutrino mass of less than an electron volt, up to the heaviest known (top) quark of 173 GeV. The force carrier bosons with spin 1 include gluons for the strong force, photon for the electromagnetic force, and the Z and W bosons for the weak force. The final spin-0 Higgs particle, discovered in 2012 at the LHC, is responsible for the so-called Higgs mechanism that generates mass.

enough that we haven't been able to detect the existence of its force mediator, and so far only a classical theory exists for gravity. A summary of the currently discovered particle content of the standard model is shown in Table 2.1.

## 2.2 Electroweak Symmetry Breaking and Gauge Boson Masses

The standard model is formulated in the language of gauge field theories. In order to preserve gauge symmetry, the fields are required to be massless; i.e., mass terms such as  $m\Psi\bar{\Psi}$  for fermions are not gauge invariant, and we cannot simply include this term by itself. This poses a problem as particles are observed to have mass.

The electroweak theory, now part of the standard model, proposed by Glashow [20], Weinberg [6] and Salam [7] introduces an additional complex scalar field  $\phi$  which solves this problem through the Higgs mechanism [8–11, 21, 22]. The scalar field is in a spinor representation of the  $SU(2)$  symmetry, with an additional  $U(1)$  gauge symmetry. Since we know that photons are massless, without the additional  $U(1)$  gauge symmetry there would be no gauge transformation that leaves the field invariant, and as a consequence all gauge bosons would be massive. The gauge transformation can be written as

$$\phi \rightarrow e^{i\alpha^a \sigma^a / 2} e^{i\beta / 2} \phi, \quad (2.1)$$

where  $\sigma^a$  denotes the usual  $2 \times 2$   $\sigma$  matrices of  $SU(2)$ . The most general gauge-invariant explicit renormalizable Lagrangian of this new scalar field sector can be written as

$$\mathcal{L} = D_\mu \phi^\dagger D^\mu \phi + \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2. \quad (2.2)$$

with  $\mu^2 > 0, \lambda > 0$ , the potential term has an interesting ‘‘Mexican hat’’ shape, as illustrated in Figure 2.1. Due to this shape, the vacuum expectation value (the so-called *vev*) of the field, which corresponds to the minimum of the potential, is nonzero and since the ring of lowest potential energy is degenerate, the location of the vacuum on the ring is arbitrary:

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} U(x) \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (2.3)$$

The vacuum expectation value  $v$  occurs at  $\sqrt{\frac{\mu^2}{\lambda}}$ , and the rotation matrix  $U(x)$  is arbitrary and does not affect physics observables. The symmetry of this field is *spontaneously broken*. The covariant derivative in Equation 2.2 can be written as

$$D_\mu \phi = \left( \partial_\mu - \frac{1}{2} i g A_\mu^a \sigma^a - \frac{1}{2} i g' B_\mu \right) \phi. \quad (2.4)$$

The gauge bosons of the  $SU(2)$  and  $U(1)$  fields are denoted by  $A_\mu^a$  and  $B_\mu$  respectively.

The mass terms for the Gauge bosons can be obtained from the kinematic term of the covariant derivative:

$$\begin{aligned} \mathcal{L} \supset \frac{1}{8} \begin{pmatrix} 0 & v \end{pmatrix} (gA_\mu^a \sigma^a + g' B_\mu) (gA^{b\mu} \sigma^b + g' B^\mu) \begin{pmatrix} 0 \\ v \end{pmatrix} \\ = \frac{v^2}{8} [g^2 A_\mu^1 A^{1\mu} + g^2 A_\mu^2 A^{2\mu} + (-gA_\mu^3 + g' B_\mu)(-gA^{3\mu} + g' B^\mu)]. \end{aligned} \quad (2.5)$$

We can see that three out of four combinations of the gauge boson fields obtain mass.

The different combinations are listed as follows:

1. Two combinations of  $A_\mu^1$  and  $A_\mu^2$  have mass  $\frac{1}{2}gv$ . We can identify  $\frac{1}{2}(A_\mu^1 \mp iA_\mu^2)$  as the  $W^\pm$  bosons.
2. The  $Z$  boson can be identified as  $\frac{1}{\sqrt{g^2 + g'^2}}(-gA_\mu^3 + g' B_\mu)$ . It has a mass of  $\frac{1}{2}\sqrt{g^2 + g'^2}v$ .
3. The last combination,  $\frac{1}{\sqrt{g^2 + g'^2}}(g'A_\mu^3 + gB_\mu)$  remains massless, and is identified with the photon.

Finally, we can expand the scalar field around its potential minimum:

$$\phi(x) = \frac{1}{\sqrt{2}} U(x) \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \quad (2.6)$$

and characterize perturbations of the  $\phi(x)$  field induced by  $H(x)$ . We call  $H(x)$  the *Higgs field*. The rotation matrix  $U(x)$  can be eliminated through gauge transformations. Expanding the potential energy part of Lagrangian (Equation 2.2), we obtain

$$\mathcal{L} \supset -\lambda v^2 H^2 - \lambda v H^3 - \frac{1}{4} \lambda H^4. \quad (2.7)$$

We can then identify the Higgs boson mass and self-interaction terms, with a mass given from the  $H^2$  term as

$$m_H = \sqrt{2\lambda}v. \quad (2.8)$$

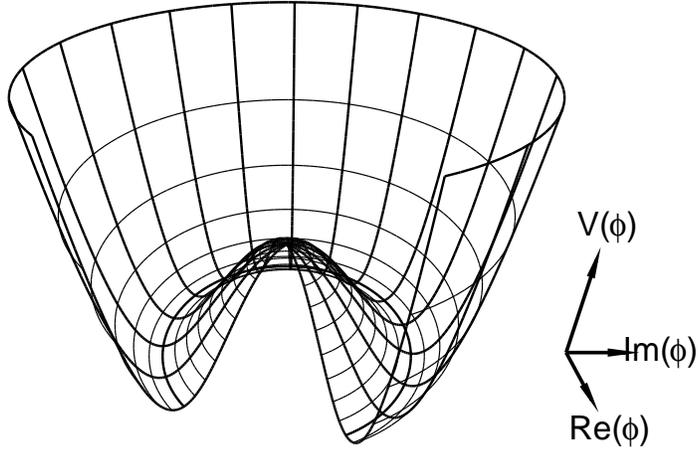


Figure 2.1: Shape of the “Mexican hat” Higgs potential as a function of the  $\phi$  field represented in the  $Re(\phi)$  and  $Im(\phi)$  axes. The height denotes the potential energy. The equilibrium at zero is not stable, as the energy is higher than its neighborhood. This triggers spontaneous symmetry breaking: the vacuum takes on a specific value along the minimum and loses its rotational symmetry.

All boson masses are specified by the four parameters  $g, g', v$  and  $\lambda$ . The determination of the exact numerical values of these parameters still awaits further experimental measurements of the self-coupling of the Higgs field, as we only have three constraints from the gauge boson masses, and this does not yet fully specify all four parameters.

## 2.3 Fermion Masses

As originally proposed by Weinberg [6], Fermions obtain mass through coupling with the  $\phi$  field which has a vacuum expectation value. The up and down type of left-handed fermions are grouped and assigned to a doublet representation of  $SU(2)$ . Right-handed fermions on the other hand are assigned as singlets in  $SU(2)$ . The  $SU(2)$  charge is conventionally named as the weak isospin: for left-handed fermions it is  $\pm 1/2$ , while it is zero for right-handed fermions. The charge under  $U(1)$  (the

weak hypercharge) is also different between left- and right-handed fermions. These assignments for fermions are summarized in Table 2.1.

The Yukawa coupling terms in the Lagrangian for leptons can be written as

$$\mathcal{L} \supset -\lambda_f \bar{F}_L \phi f_R + \text{h.c.}, \quad (2.9)$$

where  $\lambda_f$  is the coupling strength, which is different for each particle. The values are taken as input to the theory; we do not know yet if there is a deeper relation between these values (which would imply that there is a yet-to-be discovered underlying symmetry). The fields  $F_L$  and  $f_R$  denote a left-handed doublet field and a right-handed singlet field respectively. Substituting the field  $\phi$  by its vacuum expectation value plus a perturbing field  $H$ , two terms arise:

$$\mathcal{L} \supset -\frac{1}{2}\lambda_f v \bar{f}_L f_R - \frac{1}{2}\lambda_f H \bar{f}_L f_R + \text{h.c.}, \quad (2.10)$$

where we can identify  $\frac{1}{2}\lambda_f v \equiv m_f$  as the fermion mass, and rewrite the Yukawa coupling strength between the fermion and  $H$  field as  $m_f/v$ .

For quarks, one extra non-negligible term arises since down-type quarks also have mass:

$$\mathcal{L} \supset -\lambda_d \bar{Q}_L \phi d_R - \lambda_u \epsilon^{ab} \bar{Q}_{La} \phi_b^\dagger u_R + \text{h.c.} \quad (2.11)$$

Similar to the lepton case, this leads to masses  $m_d = \frac{1}{2}\lambda_d v$  and  $m_u = \frac{1}{2}\lambda_u v$ . Note that since neutrinos are found to have non-zero mass, their masses would be described by terms similar to the latter term in the above Lagrangian, if the group representation properties of the lepton sector are the same as the quark sector.

With the inclusion of three generations in the quark sector however, the mass eigenstates are not diagonal in the electroweak couplings. There is mixing between different families, related by unitary transformations between two bases. If we collectively write the left-handed up-type quarks as a triplet  $u_L$  and the left-handed

down-type quarks as a triplet  $d_L$ , the transformation between the two bases can be characterized by unitary matrices  $U_u$  and  $U_d$ :  $u'_L = U_u u_L$ ,  $d'_L = U_d d_L$ , where the primed fields are the ones that diagonalize the Higgs couplings, and the non-primed ones are the mass eigenstates. These mixings lead to flavor changing charged currents, parametrized by the Cabibbo-Kobayashi-Maskawa (CKM) matrix  $V \equiv U_u^\dagger U_d$ .

## 2.4 Parity And Charge Conjugation

In addition to the continuous gauge groups, there are a few discrete symmetries of interest: parity ( $P$ ), charge conjugation ( $C$ ) and time reversal ( $T$ ). The parity operation transforms a system into its mirror image; a particle with momentum  $\vec{p}$  is transformed into  $-\vec{p}$  under parity. The helicity state however is not altered. This symmetry, which is preserved under the electromagnetic interaction, the strong interaction and gravity, was hypothesized in the early 20<sup>th</sup> century to be conserved everywhere. However in 1957 Wu, Ambler, Hayward, Hoppes and Hudson [23] discovered that parity is violated in  $^{60}\text{Co}$  beta decays, and thus invalidated the hypothesis that parity is a fundamental symmetry of nature.

Charge conjugation symmetry is the symmetry that relates a particle with its anti-particle. It was proposed, after the observation of parity symmetry violation, that the combination of charge conjugation and parity,  $CP$ , might be a fundamental symmetry of nature. In 1964, Christenson, Cronin, Fitch and Turlay [24] discovered evidence that  $CP$  might be broken in the Kaon system.  $CP$  violation is further confirmed by many collider experiments [25, 26].

$CP$  violation is connected to some of the biggest physics puzzles today. For example, it is one of the necessary conditions for the observed matter-antimatter asymmetry in the universe, as pointed out by Sakharov [27]. However, the amount of  $CP$  violation discovered in the standard model cannot explain the large baryon asymmetry in the observable universe. Additional sources of  $CP$  violation are needed, which makes the measurement of  $CP$  properties especially interesting.

The last discrete symmetry is time reversal symmetry. It relates processes by

flipping the arrow of time. It is postulated that the  $CPT$  symmetry combination which is a basic characteristic of the formulation of quantum field theories, might be a fundamental symmetry of nature, and to this date there is no experimental evidence of  $CPT$  violation. It is interesting to note that the existence of  $CP$  violation implies  $T$ -symmetry violation if  $CPT$  is a symmetry of nature, and  $T$ -symmetry violation has been recently observed experimentally [28].

## 2.5 Higgs Production and Decays at the LHC

Since the Higgs field plays a fundamental role in the standard model, it has been very important to experimentally verify the existence of the Higgs boson with particle colliders.

There are three primary production mechanisms of the Higgs boson in the standard model: gluon fusion through a heavy quark loop, vector boson fusion, and associated production with a vector boson. In linear lepton colliders, as there is no gluon in the initial state, the Higgs bosons are mainly produced from the associated production. At the large collision energy of the LHC, the gluon contribution to the proton parton distribution functions (PDF) is much more important than quarks (and antiquarks) in producing a resonance at 125 GeV. The production of Higgs bosons is thus almost exclusively from gluon fusion at the LHC.

There are many ways a standard model  $H$  can decay. Since the coupling involves masses, a rich landscape is seen as the boson mass  $m_H$  varies. A summary of the decay branching fractions as a function of  $m_H$  is shown in Figure 2.2. The relative size of the couplings to fermions can be easily understood: since the coupling is proportional to the fermion mass, the cross section is expected to be proportional to the square of the mass, in regions where on-shell decays are kinematically favorable.

The decays to leptons and quarks through vector bosons are fixed through the Higgs mechanism, and the branching fraction to these modes keeps growing as  $m_H$  increases through the kinematic thresholds.

A Higgs mass at around 125 GeV is experimentally especially interesting, since it

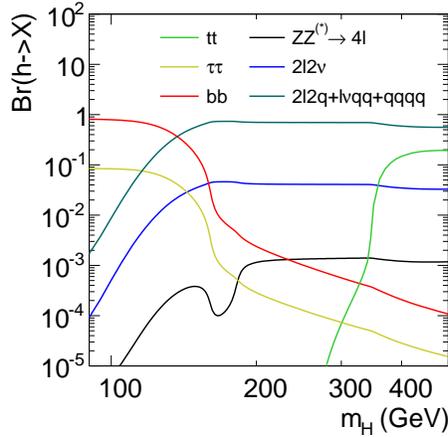


Figure 2.2: Branching fractions of the Higgs boson to different final states. For a Higgs boson at a mass of 125 GeV the situation is especially interesting experimentally, because both fermionic and bosonic decays of the Higgs boson are accessible.

is right in the transition region where decays through vector bosons are comparable to decays to fermions. This allows us to probe the properties of the Higgs boson using many different channels.

## 2.6 The “Golden Channel”

Among the numerous decay channels, the decay to four leptons, predominantly from  $H \rightarrow ZZ^{(*)} \rightarrow \ell^+ \ell^- \ell^+ \ell^-$  ( $\ell = e, \mu$ ), is of special interest. At  $m_H \sim 125$  GeV and 8 TeV collision energy, the branching ratio for  $H \rightarrow ZZ \rightarrow 4\ell$  ( $\ell = e, \mu$ ) is only  $1.26 \times 10^{-4}$  and the cross section is 2.43 fb. While the decay branching fraction and cross section is low, the signature is striking and the signal to background ratio is good. With  $20 \text{ fb}^{-1}$  we expect a few dozen signal events, and similar number from background processes.

The Feynman diagrams for the processes relevant to this channel are summarized in Figure 2.3. The main irreducible background comes from the  $t/u$ -channel  $q\bar{q}$  initial state, with a smaller contribution from  $s$ -channel process, and an even smaller contribution from the  $gg$  initial state. The composition of the background processes

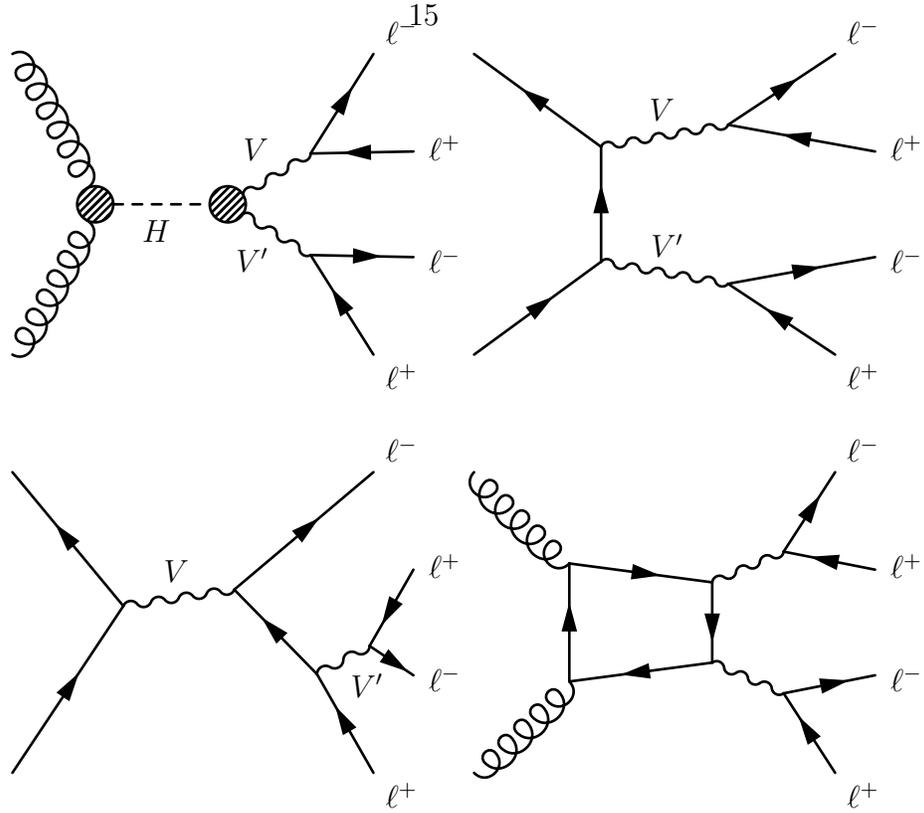


Figure 2.3: The relevant Feynman diagrams for the signal and background in the  $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$  analysis. The upper left diagram shows the dominant production mode at the LHC for the Higgs boson from gluon fusion. The Higgs then decays into four leptons through intermediate vector bosons. The upper right diagram shows the dominant standard model background  $4\ell$  process through a  $t/u$ -channel diagram from a  $q\bar{q}$  initial state. The lower left diagram is the  $s$ -channel  $Z \rightarrow 4\ell$  resonant process, which contributes roughly 10% of the background at  $m_{4\ell} = 125$  GeV. The lower right diagram is the sub-dominant background coming from gluon fusion, which at  $m_{4\ell} \sim 125$  GeV comprises about 3% of all the background.

is discussed in more detail in Section 6.1.

Another advantage of this channel is the richness of the decay information. With four leptons in the final state, it is a closed (fully reconstructible) final state that contains 12 degrees of freedom. This allows us to study the Higgs resonance in great detail, and gives us the ability to unambiguously establish the discovery and measure the Higgs boson's properties.

## Chapter 3

# Effective Theory Approach and the Properties Measurements of the Higgs in the Four Lepton Channel

### 3.1 parametrization Of The $HVV'$ Vertex

A framework useful for the measurement of the Higgs boson properties starts by parametrizing the effective  $HVV'$  vertex [29]. There are in principle three possible Lorentz structures in each of the  $VV'$  sectors, where  $V$  and  $V'$  are neutral gauge bosons ( $Z, \gamma$ ):

$$i\Gamma_{VV'}^{\mu\nu} = v^{-1} \left( A_1^{VV'} m_Z^2 g^{\mu\nu} + A_2^{VV'} (k_1 \cdot k_2 g^{\mu\nu} - k_1^\mu k_2^\nu) + A_3^{VV'} \epsilon^{\mu\nu\rho\sigma} k_{1\rho} k_{2\sigma} \right), \quad (3.1)$$

where  $k_1$  and  $k_2$  are the two 4-momenta of  $V$  and  $V'$ . The couplings  $A_1^{VV'}$ ,  $A_2^{VV'}$  and  $A_3^{VV'}$  are dimensionless complex form factors. The first coupling,  $A_1^{VV'}$ , is related to the mass term. In the standard model only the  $A_1^{ZZ}$  coupling is present, while  $A_1^{Z\gamma}$  and  $A_1^{\gamma\gamma}$  are protected by gauge symmetry and vanish to all orders.

Among the three couplings,  $A_3^{VV'}$  is odd under a  $CP$  operation, while the other two are even under  $CP$ . Since these are the only three allowed Lorentz structures, we can view these as one basis for the effective theory. Any physics of a heavier scale (typically some new particles running in loops between  $H$  and  $VV'$ ) will generate some combination of these couplings. Any deviations in these couplings would thus

serve as a smoking gun for new physics.

Each of these terms generates a distinctive kinematic distribution. The mass distributions of the two intermediate vector bosons are shown in Figure 3.1.

The couplings  $A_{1,2,3}^{VV'}$  are in general contain  $q^2$ -dependent form factors. In most high mass scenarios, since further terms are suppressed by scale factors of the form  $\frac{1}{\Lambda^2}$  where  $\Lambda$  is an effective mass scale or higher powers of this scale factor, we can consider the form factors to be constant.

The form factors can also be complex, and could be generated for example if a new unseen light particle is running in a loop. This generates non-trivial momentum dependence, and one would need to consider both the complex phase and momentum dependence at the same time for a better picture.

To the lowest order in the momentum expansion, the effective vertex parametrization can be mapped one-to-one from the effective Lagrangian:

$$\begin{aligned} \mathcal{L} \sim \frac{H}{4v} & (A_1^{ZZ} Z_\mu Z^\mu + A_2^{ZZ} Z_{\mu\nu} Z^{\mu\nu} + A_3^{ZZ} Z^{\mu\nu} \tilde{Z}_{\mu\nu} \\ & + A_2^{Z\gamma} Z_{\mu\nu} F^{\mu\nu} + A_3^{Z\gamma} Z^{\mu\nu} \tilde{F}_{\mu\nu} \\ & + A_2^{\gamma\gamma} F_{\mu\nu} F^{\mu\nu} + A_3^{\gamma\gamma} F^{\mu\nu} \tilde{F}_{\mu\nu}). \end{aligned} \quad (3.2)$$

In the expression above, the vector boson fields with two indices are the anti-symmetric field strength tensor

$$V_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu, \quad (3.3)$$

and the tilde indicates the dual of the field strength tensor:

$$\tilde{V}_{\mu\nu} = \varepsilon_{\mu\nu\rho\sigma} V^{\rho\sigma}. \quad (3.4)$$

While the vertex parametrization is unique and complete, this choice of operators in the Lagrangian is not. There are many equivalent operators which map back to the same vertex factors. Momentum dependent terms can be mapped from additional operators as well.

A similar formulation [30, 31] used in CMS results parametrizes the decay ampli-

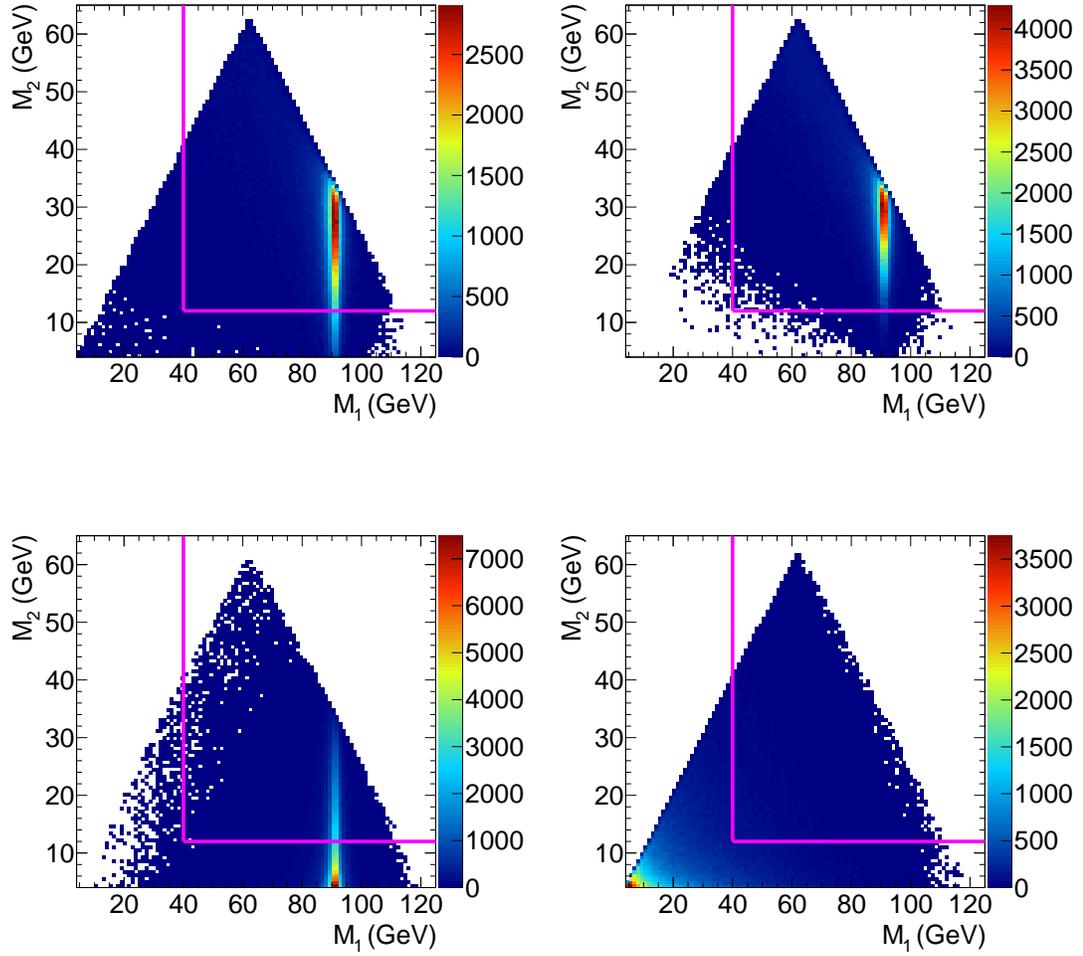


Figure 3.1: Shape of the different terms in the effective Lagrangian, as a function of the dilepton masses. The upper left figure is the standard model  $HZZ$  tree-level coupling, and other three panels are higher order  $HVV'$  couplings. The upper right figure is  $A^{ZZ}$ , and the lower row is  $A^{Z\gamma}$  and  $A^{\gamma\gamma}$  respectively. The pink line indicates the lepton mass-pair ranges covered by typical analyses in CMS.

tude as follows:

$$\begin{aligned}
A \sim v^{-1} & (a_1 m_Z^2 \epsilon_1^* \epsilon_2^* + a_2 f_{(Z)\mu\nu}^* f_{(Z)}^{*\mu\nu} + a_3 f_{(Z)\mu\nu}^* \tilde{f}_{(Z)}^{*\mu\nu} \\
& + a_2^{Z\gamma} f_{(Z)\mu\nu}^* f_{(\gamma)}^{*\mu\nu} + a_3^{Z\gamma} f_{(Z)\mu\nu}^* \tilde{f}_{(\gamma)}^{*\mu\nu} \\
& + a_2^{\gamma\gamma} f_{(\gamma)\mu\nu}^* f_{(\gamma)}^{*\mu\nu} + a_3^{\gamma\gamma} f_{(\gamma)\mu\nu}^* \tilde{f}_{(\gamma)}^{*\mu\nu}). \tag{3.5}
\end{aligned}$$

Here the field strength tensor of the Lagrangian formulation is replaced by  $f_{(V)}^{\mu\nu}$ , defined as  $f_{(V)}^{\mu\nu} = \epsilon^\mu k^\nu - \epsilon^\nu k^\mu$ , where  $k$  is the vector boson momentum.

## 3.2 Effective Theory: Range Of Validity

The range of validity of this effective theory approach depends on how we would like to interpret the result. We can measure the parameter set that best fits the data within a designated model space, where all model points in the space can be valid alternatives. The measurement is then simply the range of parameters within the model space that best describes the data.

The interpretation of the result, however, is much harder. There are many possible considerations that can affect the results of the measurements. For example, one can consider the contribution of the momentum-dependent terms to the couplings. Although it is expected that these terms are suppressed by some heavy scale of new physics, there is no formal study to assess how small momentum-dependent contributions might affect our measurement, where we only take the constant term in the momentum expansion. In order to interpret the result in terms of *the* Lagrangian parameters, one needs to carefully consider all possibilities, including the  $b$ -quark loop which generates some small complex phase, and momentum-dependent form factors.

It is therefore important not to mix the interpretation with the measurement. Since for the standard model we don't have a full analytic calculation of these momentum-dependent form factors (past leading order), it is prudent to report the best fit values we see obtain from the data in the defined model space, *as is*.

## Chapter 4

# Helicity Amplitude Formulation for the $ZZ$ On-Shell Case

### 4.1 $CP$ Eigenstates and Lorentz Structures

In  $H \rightarrow ZZ$  when  $H$  is  $> 2m_Z$ , the two  $Z$  bosons are both on-shell, and the helicity states are well-defined. One can then identify the three  $CP$  eigenstates, and therefore the three terms in the effective theory formulation, from the configuration of the helicity states. Since  $H$  is a scalar and  $Z$  is a vector, there are three ways to line up the helicity of the  $Z$ s so that the total projection of the spin is zero. In the rest frame of the  $H$ , we can have the projection of the spin of each  $Z$  both parallel to the direction of each  $Z$ , or both anti-parallel to the motion, as shown in Figure 4.1. The third possibility is to have zero in the spin projection along the momentum direction. A  $CP$ -even eigenstate and a  $CP$ -odd eigenstate can then be formed by combining the first two configurations. The third is even under  $CP$  operation.

The transverse and longitudinal polarization vectors of the two  $Z$ s in their respective rest frames can be written as

$$\begin{aligned}
 \epsilon^{(1)} &= (0, 0, 0, \pm 1) \\
 \epsilon^{(2)} &= \frac{1}{\sqrt{2}}(0, 1, i, 0) \\
 \epsilon^{(3)} &= \frac{1}{\sqrt{2}}(0, 1, -i, 0).
 \end{aligned} \tag{4.1}$$

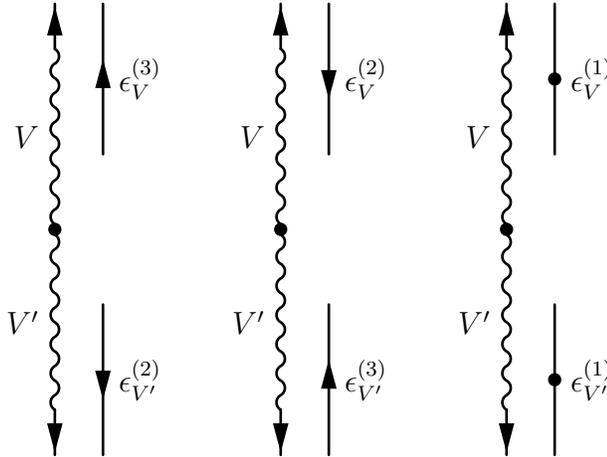


Figure 4.1: Three configurations of the spin lineup for the vector bosons decaying from the Higgs. Since the Higgs boson is spinless, the total projection of the spin along the decay axis must add up to zero. The projection of spins can be both parallel to the motion, both anti-parallel to the motion, or zero for both vector bosons. The polarization vectors for the vector bosons are indicated in Equation 4.2.

We choose the boost direction of one of the  $Z$ s to be  $z$  axis, with the boost into the rest frame of the heavy resonance ( $H$ ):

$$\begin{aligned}
 \epsilon_V^{(1)} &= \epsilon_{V'}^{(1)} = (\gamma\beta, 0, 0, \pm\gamma) \\
 \epsilon_V^{(2)} &= \epsilon_{V'}^{(3)} = \frac{1}{\sqrt{2}}(0, 1, i, 0) \\
 \epsilon_V^{(3)} &= \epsilon_{V'}^{(2)} = \frac{1}{\sqrt{2}}(0, 1, -i, 0),
 \end{aligned} \tag{4.2}$$

where the  $\epsilon^{(1)}$ 's for the two  $Z$ 's have opposite signs in the  $z$ -projection. Since the resonance that gives rise to the  $Z$ 's is a scalar, the transverse polarization vectors follow the following relation:

$$\epsilon_V = \kappa \epsilon_{V'}^*, \tag{4.3}$$

where  $\kappa \in \mathbb{R}$ . The Lorentz structure (Equation 3.1) is real when all the couplings are

real. For a symmetric  $VV'$  tensor structure the helicity amplitude is real:

$$\begin{aligned}
\epsilon_{V\mu}\epsilon_{V'\nu}M^{\mu\nu} &= \kappa\epsilon_{V\mu}\epsilon_{V'\nu}^*M^{\mu\nu} \\
&= \frac{1}{2}\kappa(\epsilon_{\mu}\epsilon_{\nu}^* + \epsilon_{\nu}\epsilon_{\mu}^*)M^{\mu\nu} \\
&= \kappa \operatorname{Re}(\epsilon_{\mu}\epsilon_{\nu}^*)M^{\mu\nu}.
\end{aligned} \tag{4.4}$$

Note that from here on, the subscript  $V$  is dropped for brevity.

Similarly we can show that if the tensor structure is antisymmetric, the helicity amplitude is purely imaginary. Hence we can write the helicity amplitude as follows:

$$A_T = f_R(A_1^{ZZ}, A_2^{ZZ}) + if_I(A_3^{ZZ}) = |A_T|e^{i\delta_T}, \tag{4.5}$$

where the functions  $f_R$  and  $f_I$  are real functions of the coupling parameters, and vanish when all arguments are zero. The total helicity amplitude contains a non-zero phase only when the  $CP$ -odd coupling is non-zero. The longitudinal polarization vector is real, and because of the similar sign in the energy component and the opposite sign in the  $z$  component, both the symmetric and antisymmetric parts of the tensor structure contribute. The helicity amplitude  $A_L = f_L(A_1^{ZZ}, A_2^{ZZ}, A_3^{ZZ})$  for such a state is a real number, which depends on all the coupling parameters.

The phase in the helicity amplitude enters as a phase shift in the  $\phi$  distribution, which is the opening angle between the two  $Z \rightarrow \ell\bar{\ell}$  planes. Distributions of the different terms are illustrated in Figure 4.2, for a 200 GeV resonance decaying into two on-shell  $Z$ 's.

This formulation is less well-defined when one of the  $Z$ 's is off-shell. The phenomenological behavior however remains similar for lower mass resonances. A helicity amplitude formulation in such cases remains as a future project.

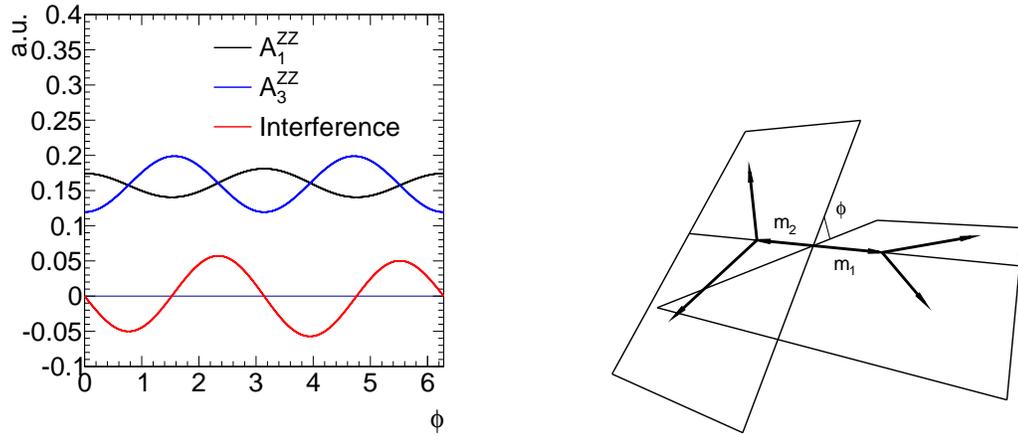


Figure 4.2: The distribution of the opening angle between the two  $Z \rightarrow \ell\bar{\ell}$  planes for the  $CP$  even and  $CP$  odd terms for a 200 GeV heavy resonance. The size of the coupling is chosen such that the black and the purple curves have the same normalization. The definition of the opening angle  $\phi$  is shown on the right.

# Chapter 5

## *CP* Violation in the $HVV'$ Sector

One topic of special interest for the measurement in this channel is the *CP* property of the  $HVV'$  vertex. There are two types of potential *CP* violation [32]: one occurs when there is interference between the *CP*-even and *CP*-odd operators. The interference term is *CP*-odd in nature. It could happen both in the same  $VV'$  sector as well as across different  $VV'$  sectors. Another type of potential *CP* violation occurs when both strong and weak phases are present. Other than the possibility of having complex couplings, one natural source of a strong phase is the phase from the  $Z$  Breit-Wigner line shape. These two types of potential *CP* violation are discussed in more detail in this chapter.

### 5.1 *CP* Violation Within The Same $VV'$ Sector

As described in Chapter 4, the helicity amplitudes for the *CP*-even operators are real, and imaginary for the *CP*-odd operators. This imaginary phase enters as a phase shift in the distribution of the opening angle between the two  $Z \rightarrow \ell\bar{\ell}$  decay planes. This angle  $\phi$  can be written as a triple product of the momentum vectors of the four leptons:

$$\cos \phi = \frac{(\vec{p}_1 \times \vec{p}_2) \cdot (\vec{p}_3 \times \vec{p}_4)}{|(\vec{p}_3 \times \vec{p}_4)| |(\vec{p}_3 \times \vec{p}_4)|}. \quad (5.1)$$

When both *CP*-even and *CP*-odd terms are present, interference effects cause the phase of the  $\cos 2\phi$  component to shift (as shown in the figure below): the square terms from the *CP* even parts have a phase of 0, while the square terms from the

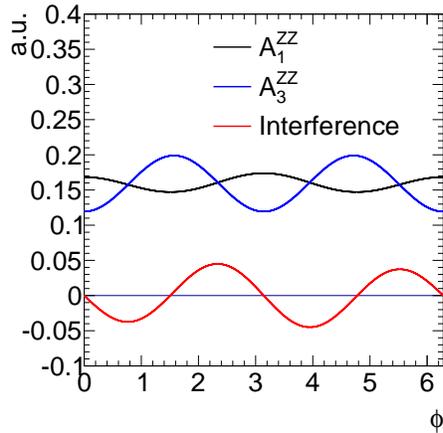


Figure 5.1: Demonstration of the phase shift in the  $\phi$  distribution. Here we compare  $A_1^{ZZ}$  and  $A_3^{ZZ}$ , as well as interference between the two terms. When both terms are present, the interference term contributes and causes an overall phase shift in the  $\phi$  distribution.

$CP$  odd part have a phase of  $\pi$ , and the terms coming from the interference between parts of different  $CP$  natures have a phase  $\pi/2$ . One example of breaking down the different terms in the  $\phi$  distribution is shown in Figure 5.1.

The size of the interference term in general is correlated with the overlap between the two square terms. If the total amplitude is  $(A_1 + A_2)$  for the two terms, the size of interference term is  $2|A_1^* A_2| \sim O(\sqrt{|A_1|^2 |A_2|^2})$ . Therefore this type of  $CP$  violation is most significant within the same  $VV'$  sector, where the mass distributions are similar between the  $CP$ -even and  $CP$ -odd terms.

Additionally, since the observable is the opening angle between the two vector boson decay planes, both planes are required for the observation of this type of  $CP$  violation. By contrast, in the case of on-shell final state photons, one could in principle measure the polarization of the photon or the plane of conversion. However, this is difficult to perform experimentally, and there is no experimental analysis yet that attempts to do this measurement by using the final state photons.

## 5.2 $CP$ Violation Across Different $VV'$ Sectors - Strong Vs. Weak Phase

Another type of  $CP$  violation can arise when both a strong phase and a weak phase are present.  $CP$  violation is provided in this case by interference between the strong phase  $\delta_S$ , which does not change sign under  $CP$  operation, with the weak phase  $\delta_W$ , which changes sign. One example of the weak phase is the  $CP$ -odd term which has an imaginary amplitude. It changes sign under a  $CP$  operation.

To see that the presence of both strong and weak phases leads to  $CP$  violation, suppose there is an amplitude where we have both types of phases:

$$A = (A_S e^{i\delta_S} + A_W e^{i\delta_W}) e^{i\delta_0}, \quad (5.2)$$

where  $A_S$  and  $A_W$  real and non-zero, and where  $\delta_0$  is an overall phase which could contain both weak and strong phases. The phases  $\delta_S$  and  $\delta_W$  in this equation are to be understood as the difference in the strong and weak phases respectively between the two terms in the amplitude. Any overall common phase can be factored out and is not physical. The interference term in the square of the amplitude can then be written as

$$|A|_{\text{int.}}^2 = A_S A_W (e^{i(-\delta_W + \delta_S)} + e^{i(\delta_W - \delta_S)}) = 2A_S A_W \cos(\delta_W - \delta_S). \quad (5.3)$$

Under a  $CP$  operation, the weak phase changes sign:

$$CP [ |A|_{\text{int.}}^2 ] = 2A_S A_W \cos(\delta_W + \delta_S), \quad (5.4)$$

and we can write the  $CP$  violating term as

$$\frac{1}{2} (2A_S A_W \cos(\delta_W - \delta_S) - 2A_S A_W \cos(\delta_W + \delta_S)) \propto A_S A_W \sin(\delta_W) \sin(\delta_S). \quad (5.5)$$

We see that when both the strong and weak phases are present, there is a non-zero

$CP$  violating term.

The strong phase can come from different sources. One possibility is to have complex couplings, where the coupling phase does not change under a  $CP$  operation. However due to the reasons described in Chapter 3, a complex coupling is less well-motivated theoretically.

A strong phase can be present even with all couplings real. Since the  $Z$  boson is massive and shortlived, the propagator is imaginary and carries a nontrivial phase dependence:

$$\frac{-i}{k^2 - m_Z^2 + im_Z\Gamma_Z} \quad (5.6)$$

This imaginary phase of the propagator doesn't change under a  $CP$  operation. It can be the source of the strong phase in  $CP$  violating terms.

As seen in the existence condition for a  $CP$  violating term in this case (Equation 5.5), we require that  $\sin \delta_S$  is nonzero, indicating that the strong phase of the two terms (in Equation 5.2) in the amplitude are different. If both terms in the amplitude come from the same  $VV'$  sector, the phase of the propagator is factored out into the  $\delta_0$  part, and there is no  $CP$  violation of this type.

As an example, we can look at the interference between the  $Z\gamma$  sector and the  $\gamma\gamma$  sector. When all the couplings are real, we can rewrite the effective Lagrangian (Equation 3.2) in these two sectors as

$$\begin{aligned} \mathcal{L} \sim \frac{h}{4v} & \left( A^{Z\gamma} \cos \phi_{Z\gamma} Z^{\mu\nu} F_{\mu\nu} + A^{Z\gamma} \sin \phi_{Z\gamma} Z^{\mu\nu} \tilde{F}_{\mu\nu} \right. \\ & \left. + A^{\gamma\gamma} \cos \phi_{\gamma\gamma} F^{\mu\nu} F_{\mu\nu} + A^{\gamma\gamma} \sin \phi_{\gamma\gamma} F^{\mu\nu} \tilde{F}_{\mu\nu} \right), \end{aligned} \quad (5.7)$$

where  $\phi_{VV'}$  is the  $CP$  phase in the  $VV'$  sector. Here the weak phase can be provided by the difference between the  $CP$  phase of the two sectors:  $\phi_{Z\gamma} - \phi_{\gamma\gamma}$ .

Note that this type of  $CP$  violation does not require one to construct two planes and look at the opening angle. One can look at  $H \rightarrow \ell\bar{\ell}\gamma$  events, for example, and observe this type of  $CP$  violation. In the differential cross section of  $H \rightarrow \ell\bar{\ell}\gamma$ , there

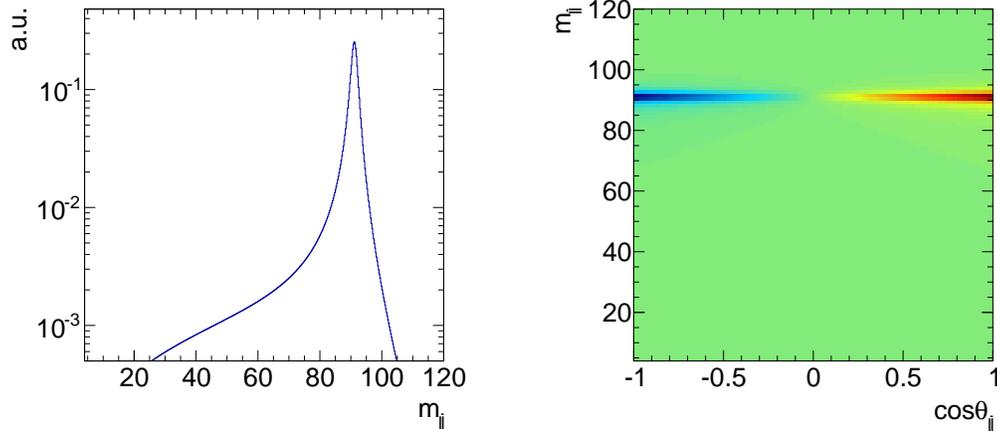


Figure 5.2: Shape of the  $CP$  violating term in the case of  $H \rightarrow \ell\bar{\ell}\gamma$  as a function of the di-lepton mass and the decay angle of the leptons. In the left plot, the projection of the absolute value of the  $CP$  violating term is shown. As the contribution from forward decay and the backward decay of the lepton with respect to the  $Z$  boson direction cancel, the  $CP$  violating term does not contribute to the total cross section. This is better illustrated in the 2D distribution, where red is positive and blue is negative; green represents zero. The contribution is largest at the  $Z$  mass, and where the direction of the leptons is close to the direction of the  $Z$  boson.

is a  $CP$  violating term as follows:

$$\frac{d\Gamma_{\text{CPV}}}{dm_{\ell\bar{\ell}}^2 d\cos\theta_{\ell\bar{\ell}}} = \cos\theta_{\ell\bar{\ell}} A^{Z\gamma} A^{\gamma\gamma} \sin(\phi_{Z\gamma} - \phi_{\gamma\gamma}) \frac{e(g_{Z,R} - g_{Z,L})m_Z\Gamma_Z (m_H^2 - m_{\ell\bar{\ell}}^2)^3}{512\pi^3 m_H^3 v^2 ((m_{\ell\bar{\ell}}^2 - m_Z^2)^2 + m_Z^2 \Gamma_Z^2)}. \quad (5.8)$$

This term is proportional to the width of  $Z$  resonance, and has a characteristic Breit-Wigner piece in the expression. The contribution is largest when  $m_{\ell\bar{\ell}}$  is close to  $m_Z$ . It is also a function of the relative weak phase of the two sectors  $\phi_{Z\gamma} - \phi_{\gamma\gamma}$ . Finally, it is proportional to  $\cos\theta_{\ell\bar{\ell}}$ , the decay angle in the di-lepton rest frame with respect to the direction of the photon, indicating that this decay angle is a  $CP$ -violation observable. The shape of this term as a function of  $\cos\theta_{\ell\bar{\ell}}$  and  $m_{\ell\bar{\ell}}$  is shown in Figure 5.2.

## Chapter 6

# The Four Lepton Standard Model Backgrounds

### 6.1 Background Processes And Components

The background processes [29, 33] to the Higgs boson signal are discussed in this section. There are a number of standard model processes that can generate a four lepton final state. Those relevant to the measurement of the properties for a resonance around 125 GeV are shown in Figure 2.3. The largest contribution to the standard model background is from the  $q\bar{q}$  initial state. A small contribution comes from the  $gg$  initial state, where the two vector bosons attach to an internal fermion loop. For the  $q\bar{q}$  initial state there exist both  $t$ - and  $u$ -channel processes, where the fermion line couples directly to two vector bosons; and the  $s$ -channel process, where the  $q\bar{q}$  annihilates into a vector boson which decays into two leptons, and one of the leptons radiates another off-shell vector boson which subsequently decays into two leptons.

At a center of mass energy  $\sim 125$  GeV, the  $s$ -channel resonance mainly goes through a  $Z$ , and the contribution from  $\gamma$  is small. The radiated vector boson is much more likely to be a  $\gamma$ , since the  $Z$  is too heavy. The differential cross section in  $m_{4\ell}$  from this sub-process has a resonance peak at  $m_{4\ell} = m_Z$  over a falling spectrum.

For the  $t/u$ -channel sub-process, the largest contribution is through  $Z\gamma$  that generates a broad shoulder above the kinematic threshold of  $m_Z$ . A residual contribution comes from the  $ZZ$  and  $\gamma\gamma$  intermediate states.

The relative contributions of different  $q\bar{q}$  initial state processes are shown [29] in Figure 6.1 at leading order, combining the analytical differential cross section in the  $4\ell$  center of mass frame with the CTEQ611 [34] PDF set to account for the global boost of the  $4\ell$  system. A CMS-like cut [35] on the di-lepton mass and lepton acceptance is applied in this comparison. Different colors indicate different contributions:  $s$ -channel process,  $t/u$ -channel  $\gamma\gamma$ ,  $t/u$ -channel  $ZZ$ ,  $t/u$ -channel  $Z\gamma$ ,  $t/u$ -channel  $ZZ/Z\gamma/\gamma\gamma$  interference, and interference between the  $s$ -channel and the  $t/u$ -channel. The dotted lines indicate a negative contribution, as the interference between different the terms might contribute negatively to the cross section.

In the region we are interested in, the standard model background is dominated by the  $t/u$ -channel diagrams with  $Z\gamma$  intermediate state. The contribution from  $ZZ$  rises continually as a function of  $m_{4\ell}$ , until it passes the kinematic threshold  $2m_Z$ , after which it dominates. The resonant  $s$ -channel process dominates at  $m_{4\ell} \simeq m_Z$ , but decreases when  $m_{4\ell}$  is far from  $m_Z$ . The contribution from the  $t/u$ -channel  $\gamma\gamma$  processes is a few percent of the total rate, throughout the plotted mass range.

The processes initiated from a  $gg$  initial state are not significant in the low-mass region. This contribution rises to about 30% when  $m_{4\ell}$  is larger.

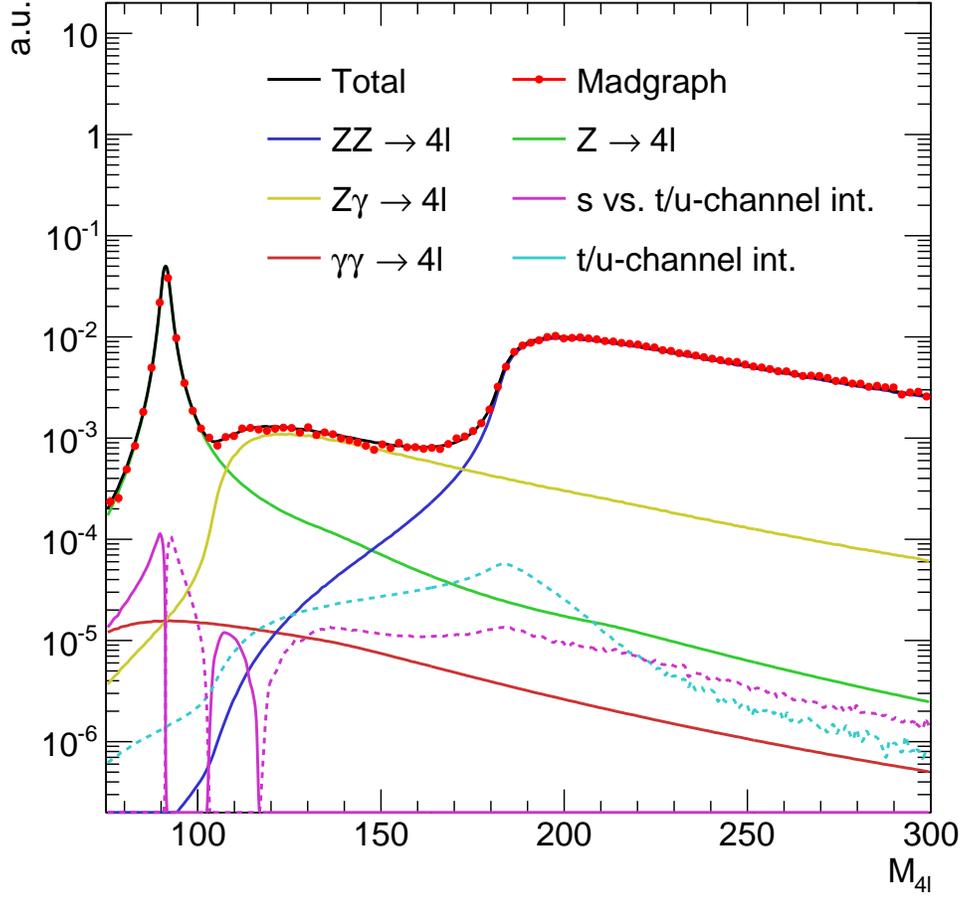


Figure 6.1: Contributions from different sub-process in the standard model  $d\bar{d} \rightarrow 2e2\mu$  process, including interferences between the different terms. The red dots are events generated with MADGRAPH [36], to be compared with the total differential cross section shown in black. The different subprocesses shown include a t/u-channel  $d\bar{d} \rightarrow ZZ \rightarrow 2e2\mu$ ,  $d\bar{d} \rightarrow Z\gamma \rightarrow 2e2\mu$ ,  $d\bar{d} \rightarrow \gamma\gamma \rightarrow 2e2\mu$  and the s-channel resonant process  $d\bar{d} \rightarrow Z \rightarrow 2e2\mu$ . The dashed lines for the interference terms indicate a negative contribution. This comparison is done at leading order with the CTEQ611[34] PDF set.

## Part II

# The Higgs Discovery and the Properties Measurements

# Chapter 7

## Description of the LHC and the CMS Detector

### 7.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [37] is a proton-proton collider located at CERN, spanning the border of Switzerland and France near Geneva. The main ring is 100 m to 150 m underground and has a circumference of 27 km. Superconducting magnets are used together with RF cavities to accelerate and guide the beams, including 1232 dipole magnets and 392 quadrupole magnets. With a large magnetic field in the dipoles of 8.2 T, the LHC is designed to accelerate protons up to 7 TeV per beam. During Run I of the LHC, 3.5 to 4.0 TeV per beam was delivered, with prospects of 13 to 14 TeV collisions foreseen for Run II in 2015-2018.

There are four main detectors along the LHC ring. The two general-purpose detectors, CMS [38] and ATLAS [39], focus on having almost  $4\pi$  detector coverage, and good tracking and calorimetry. These detectors were optimized to detect the Higgs, as well as any new physics signatures, and to study standard model processes. The LHCb detector [40] is specialized in studying  $b$ -physics, which boasts an excellent tracking detector with outstanding particle identification. Since the goal of LHCb is to study heavy flavor physics and it uses an internal gas jet target intercepted by one beam, the detector covers only the forward region. ALICE [41] is a detector that specializes in studying heavy ion collisions. It has excellent particle identification

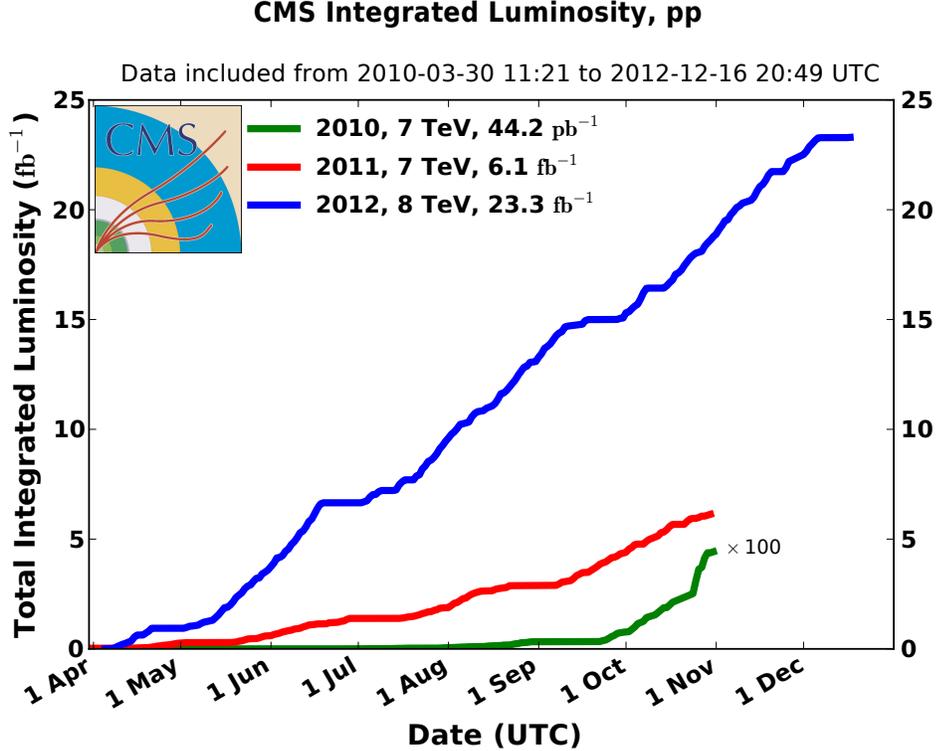


Figure 7.1: The amount of  $pp$  collision data collected by the CMS detector as a function of time during Run I. The total integrated luminosity for  $pp$  collision data is  $23.3 \text{ fb}^{-1}$  at 8 TeV and  $6.1 \text{ fb}^{-1}$  at 7 TeV.

capacity, even in busy environments like the one produced by heavy ion collisions, which allows one to study the properties of the quark gluon plasma [42] in depth. With the recent and future planned addition of electromagnetic calorimeter modules in ALICE, the study of jet physics is made possible, enlarging the physics scope.

During Run I  $6.13 \text{ fb}^{-1}$  of  $pp$  collision at 7 TeV and  $23.3 \text{ fb}^{-1}$  at 8 TeV was collected by CMS [43]. For heavy ion studies, the LHC provided  $36.14 \text{ nb}^{-1}$  of  $pPb$  collision at  $\sqrt{s_{NN}} = 5.02 \text{ TeV}$  and  $166.67 \mu\text{b}^{-1}$  of  $PbPb$  collision data to CMS at  $\sqrt{s_{NN}} = 2.76 \text{ TeV}$ . The evolution of the data collection as a function of time is shown in Figure 7.1.

With the increased instantaneous luminosity towards the end of Run I (up to  $7.7 \text{ Hz/nb}$ ), for each beam crossing there are simultaneous multiple collisions (pileup). Since the cross section for soft interactions is much higher than for the hard processes we want to study, to a very good approximation we can treat each event to be at

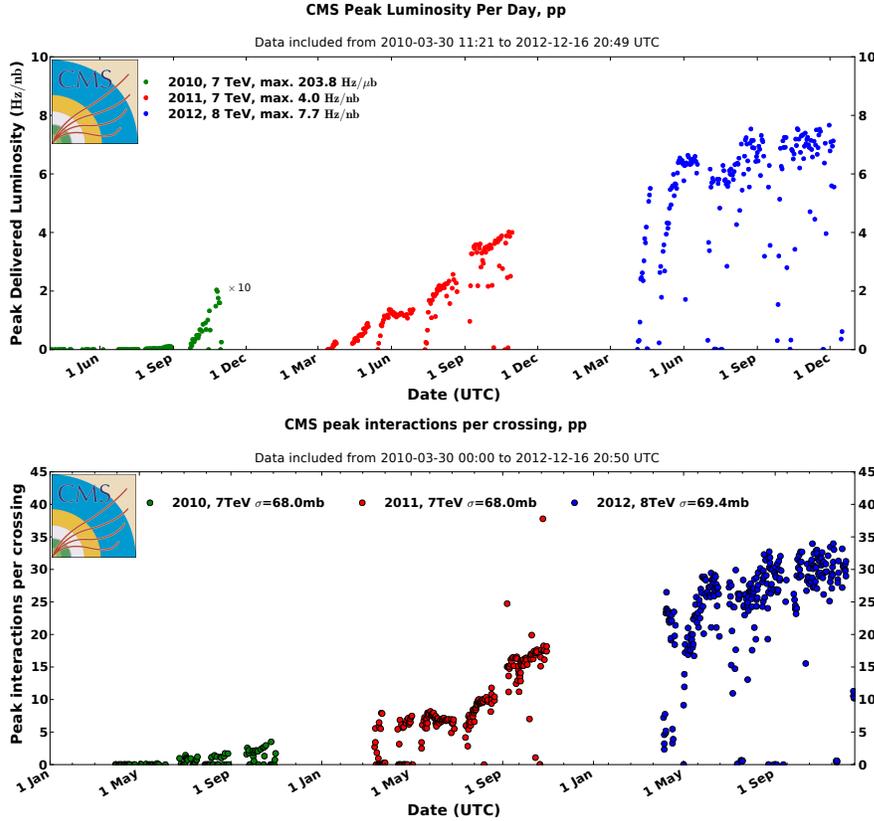


Figure 7.2: The peak instantaneous luminosity is shown in the top panel during Run I data taking in CMS, in 2010-2012. As shown, the peak intensity is progressively higher towards the end of data taking in 2012. This resulted in a higher number of concurrent collisions, as shown in the bottom panel, where up to an average of 34 interactions in each event were reached during one fill of the LHC.

most one hard interaction, with the rest of the event coming from soft interactions which are of less interest. The peak number of concurrent interactions experienced in CMS during Run I, averaged over one fill of the LHC, was 34.

## 7.2 The Compact Muon Solenoid Detector

CMS is a general-purpose detector with excellent all-round physics performance and enhanced muon detection [44], located at the interaction region in the LHC ring opposite to the CERN Meyrin site, close to Cessy, France. The detector is  $\sim 5$ -stories high weighing 14 kton. From the inside to the outside, as shown in Figure 7.3, there is the silicon pixel and strip tracking system, electromagnetic and hadron calorime-

ters, superconducting solenoidal magnet, and a set of muon chambers interleaved with the layers of the iron flux return yoke enclosing the whole detector. There are several forward detectors along the beamline: the forward hadron calorimeter, the CASTOR [45] experiment, the zero-degree calorimeter, as well as the TOTEM [46] experiment.

Closest to the collision point, the silicon pixel tracker has a granularity of  $100 \times 150 \mu\text{m}^2$  with a cylindrical design for a hermetic  $2\pi$  coverage. It is complemented by a set of strip tracking systems just outside of the pixel system, with granularity of  $10 \text{ cm} \times 80 \mu\text{m}$  within 55 cm of the beam line, and  $25 \text{ cm} \times 180 \mu\text{m}$  in the outer region of the tracker. The high granularity, together with a magnetic field of 3.8 T, allows us to have excellent tracking resolution, up to 2% momentum resolution over the whole tracker coverage of the pseudorapidity region up to  $|\eta| = 2.4$ , for a track with  $p_T \sim 10 \text{ GeV}/c$ .

The electromagnetic calorimeter [44] (ECAL) is a hermetic and homogeneous calorimeter made of lead tungstate crystals ( $\text{PbWO}_4$ ). It covers the pseudorapidity region up to  $|\eta| = 3$  with 61,200 crystals in the barrel region  $|\eta| < 1.479$ , and 7,324 crystals in each of the endcap regions with  $1.479 < |\eta| < 3$ . During test beam studies it was measured that the energy resolution of electrons from a  $3 \times 3$  crystal matrix can be parametrized as  $2.83\%/\sqrt{E} \oplus 12.4\%/E \oplus 0.26\%$  (with  $E$  in units of GeV). The crystals are about 24 radiation lengths [47] long, for shower containment. The excellent energy resolution of the ECAL allows us to study electrons and photons with excellent precision, making the discovery of the Higgs boson possible.

The hadron calorimeter [48, 49] (HCAL) is located outside of the ECAL. It is made of brass absorbers interleaved with plastic scintillators equipped with wavelength-shifting fibers for detector readout. The detector has a 10 interaction length [47] depth [44], and an outer hadron calorimeter (HO) in the back to catch the tail of hadron showers. By using the ECAL and HCAL, we are able to achieve a hadron resolution of 10% for a single 100 GeV charged pion.

The forward hadron calorimeter covers  $|\eta| = 3.0 - 5.0$ . It is made of steel absorber with quartz fibers for light detection. Fibers ending at different depth allow one to

measure the electromagnetic component and the hadronic component separately. A total of 1800 channels are present in both sides of the forward HCAL. The resolution in the forward HCAL for QCD dijet events can be approximated by  $\frac{\sigma}{E} \sim 1.0/\sqrt{\sum E_T}$ , where the energies are expressed in GeV.

The muon chambers [50] consist of three types of detectors: drift tube (DT) chambers in the barrel and cathode strip chambers (CSC) in the endcap region, as well as resistive plate chambers (RPCs) [51] in the endcap for triggering. Situated outside of the magnet, the barrel muon chambers are interleaved with the iron yoke that returns the magnetic flux. The drift tubes consist of several layers parallel to the beam line, each measuring either the  $x$  or  $y$  coordinate in the plane. A position resolution of  $200 \mu\text{m}$  is achieved in the drift tube chambers. The cathode strip chamber consists of 7 planes of cathode strips, with wires in between the planes of strips. By reading out the ionization signal in each of the strips and in the wires, we are able to reconstruct the location where the muon passes through each CSC layer. The location resolution is typically about  $200 \mu\text{m}$  (and  $100 \mu\text{m}$  for the plane closest to the interaction point) in the  $x - y$  plane. The RPCs consist of double-gap Bakelite chambers [51] operating in avalanche mode. The chambers are large and cannot be used for precise location measurements, but the fast response allows for a more efficient triggering.

We employ a multi-layered trigger system [52] for data acquisition. The first layer, the level-1 trigger, operates on reduced-granularity data for a fast decision time. The latency is designed to be less than  $1 \mu\text{s}$ , with a rate of up to 100 kHz. During the decision-making of the level-1 trigger, the data is held in a pipeline, and sent to the high-level trigger (HLT) system if accepted. The HLT operates on the output of the level-1 trigger, and is designed to reduce the total trigger rate from 100 kHz down to about 400 Hz for storage. Since the rate is controlled already by the level-1 trigger, we are able to run higher precision HLT algorithms on the events for a better trigger decision. The algorithms used in the HLT are similar to those used in the offline physics analyses, reducing the potential inefficiencies due to an algorithm mismatch between the triggering system and the event selections used in

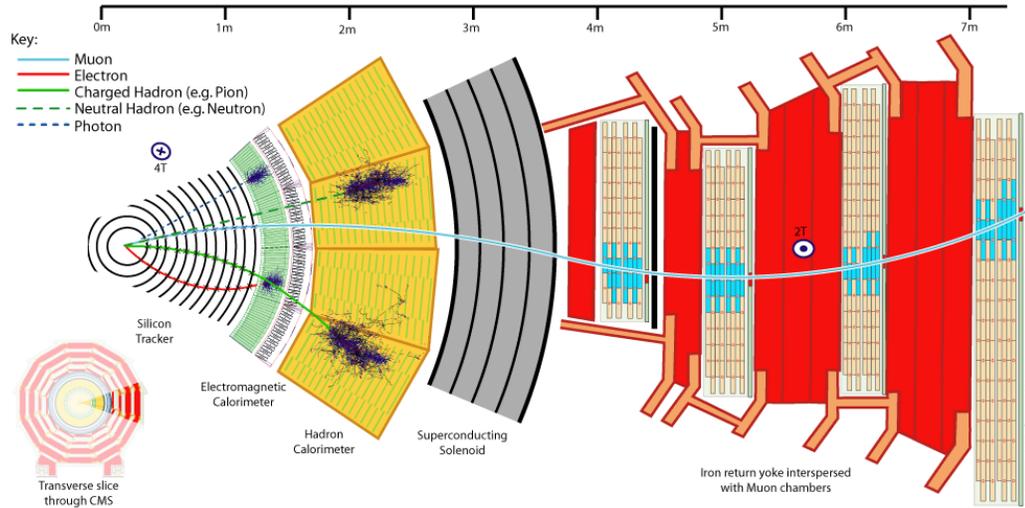


Figure 7.3: Schematic view of the CMS detector<sup>1</sup>. The collision point is on the left at the center of inner circle. The produced particles travel through many subdetectors. From inside: silicon tracking system, electromagnetic and hadron calorimeters, superconducting magnet, and finally the muon chambers and iron return yoke.

various physics analyses.

CMS had a very good overall data taking efficiency [43] in Run I. During Run I  $pp$  data taking, the overall efficiency was 93.5%, corresponding to a total of  $21.79 \text{ fb}^{-1}$  of  $pp$  collision data at 8 TeV and  $5.55 \text{ fb}^{-1}$  at 7 TeV.

<sup>1</sup>For an interactive version of the figure, see <https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/ShowDocument?docid=4172>

# Chapter 8

## The Higgs Discovery

### 8.1 Higgs Decay Channels

The CMS and ATLAS experiments at the LHC announced the discovery [15, 16, 53] of a standard model Higgs-like boson at around 125 GeV on July 4, 2012. Since then the experiments have made observations and studied the properties of the new boson in a number of channels, as summarized in Table 8.1.

The channels with the most discovery power come from closed final states with high precision objects such as leptons and photons [35, 54–57]. The cross section is low for these channels, but due to the distinctive final state signature, the background level is manageable, allowing CMS to unambiguously discover it early on.

The decay to  $W^+W^-$  is harder to search for, mainly due to the two unseen neutrinos in the dilepton channel [56, 58]. Compared to the  $ZZ$  decay channel for example, it is harder to distinguish the Higgs signal from the background, mostly due to the irreducible standard model  $W^+W^-$  background.

There are also channels with direct decays to fermions which have been explored [59–61]. The decays of the Higgs boson to the  $b\bar{b}$  final state has a large branching fraction, but it is difficult to measure as the QCD  $b\bar{b}$  production rate is large. The analysis in this channel is only feasible when the Higgs is produced in association with a vector boson  $pp \rightarrow VH$ ,  $H \rightarrow b\bar{b}$ , in order to reduce the background level.

Searches for  $H \rightarrow \tau^-\tau^+$  [60] and  $\mu^-\mu^+$  [61] are very challenging. In both exper-

Channel	CMS		ATLAS	
	Expected	Observed	Expected	Observed
$H \rightarrow ZZ$	$6.3\sigma$	$6.5\sigma$	$4.4\sigma$ (124.3)	$6.6\sigma$ (124.3)
$H \rightarrow \gamma\gamma$	$5.3\sigma$	$5.6\sigma$	$7.9\sigma$ (126.5)	$4.1\sigma$ (126.5)
$H \rightarrow W^+W^-$	$5.4\sigma$	$4.7\sigma$	$3.7\sigma$ (125.0)	$3.8\sigma$ (125.0)
$H \rightarrow \tau^-\tau^+$	$3.9\sigma$	$3.8\sigma$	$3.2\sigma$ (125.0)	$4.1\sigma$ (125.0)
$H \rightarrow b\bar{b}$	$2.3\sigma$	$2.0\sigma$	$1.4 \times \text{SM}$ (125.0)	$1.3 \times \text{SM}$ (125.0)
$H \rightarrow Z\gamma$	$10 \times \text{SM}$	$9 \times \text{SM}$	$13.5 \times \text{SM}$ (125.0)	$18.2 \times \text{SM}$ (125.0)
$H \rightarrow \mu^-\mu^+$	$5.1 \times \text{SM}$	$5.9 \times \text{SM}$	$7.0 \times \text{SM}$ (125.5)	$7.2 \times \text{SM}$ (125.5)

Table 8.1: Summary of the  $H$  observation status from CMS [15, 35, 53–55, 58–65] and ATLAS [16, 56, 57, 66], expressed in terms of the significance obtained in each channel. The numbers shown in the last two lines of the table, where the statistics are insufficient to measure the production rate, are the 95% CL upper limits.

iments multivariate techniques are used for better separation of the signal and the background. Nevertheless experiments have reported evidence of the decay in the  $\tau^-\tau^+$  channel, and we are approaching sensitivity to the decay to  $\mu^-\mu^+$  as well.

The Higgs signal strength ( $\mu \equiv \sigma/\sigma_{\text{SM}}$ ) measurements in CMS are summarized in Figure 8.1.

## 8.2 The Higgs to Four Lepton Analysis

The  $4\ell$  channel is among the highest precision channels and the main focus of this thesis.

Events used in this analysis are collected by triggers requiring double lepton triggers, where all of the  $\mu\mu$ ,  $e\mu$  and  $ee$  combinations are triggered upon. The trigger thresholds are  $p_T = 17$  GeV and  $p_T = 8$  GeV for the two leptons. For the  $4e$  final state an additional three-electron trigger is used, which requires leptons with  $p_T$  of at least 15 GeV, 8 GeV and 5 GeV to recover some of the loss in efficiency because of low energy electrons. The trigger efficiency for a Higgs-like signal with  $m_H \sim 125$  GeV is greater than 98% [35].

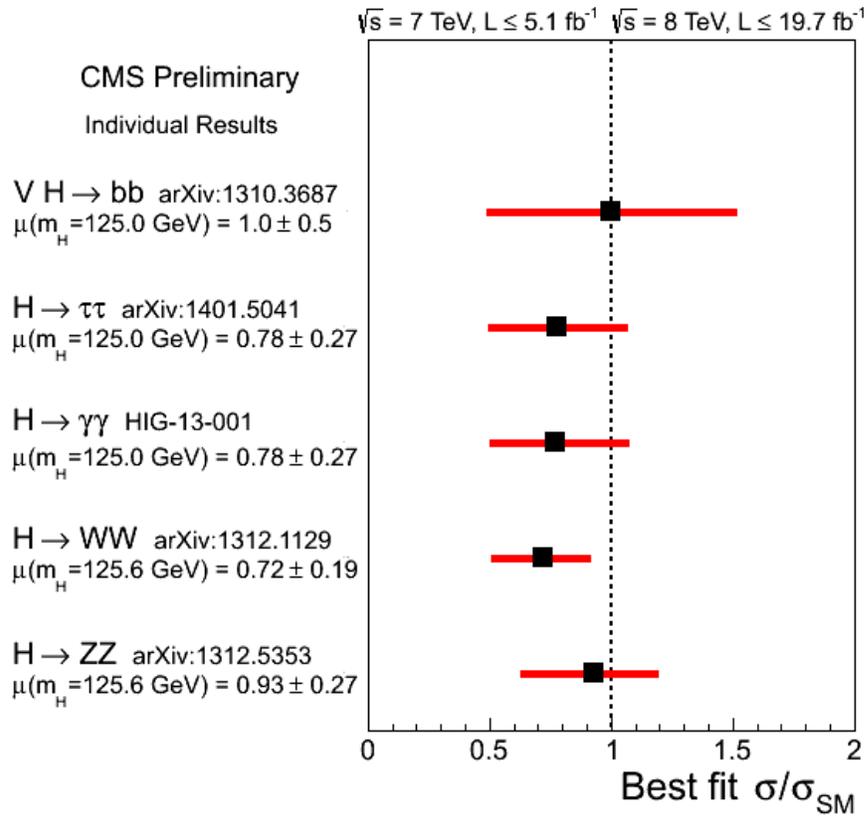


Figure 8.1: Summary of the measured Higgs signal strength ( $\mu \equiv \sigma/\sigma_{SM}$ ) from CMS, in different channels [35, 55, 58–61]. The results from all channels are compatible with the standard model prediction, as indicated by the dashed line.

### 8.2.1 Lepton Selection

We use a “particle flow” set of algorithms [67] in the analysis, that exploits information from all the subdetectors in CMS to determine the energies of jets and leptons, as well as the missing transverse energy. Electron candidates are required to have a transverse momentum of at least 7 GeV and to be within the acceptance of  $|\eta| < 2.5$ . The reconstruction of electrons combines clusters in the ECAL with the tracks matched to it. Nearly collinear final state photon radiation of the electron is recovered in the energy of the candidate electron (or positron). A “loose” selection criterion on the quality of the track-fits is applied, to preserve high efficiency for the electron candidates, at the expense of allowing part of the reducible background to remain.

The electron identification is done with a multivariate technique including observables sensitive to electron bremsstrahlung, matching of tracks and ECAL clusters, and shower shape. It is trained on simulated Drell-Yan [68] events as signal, and on a  $W$ +jet data sample where the jets are mis-identified as electrons, as background. It is then validated using the  $Z \rightarrow \ell\ell$  sample. This procedure is done in 6 bins of  $p_T$  and  $|\eta|$  of the electron, to capture the different behavior in different detector regions.

The muons are required to have at least 5 GeV transverse momentum with  $|\eta| < 2.4$ . The reconstruction uses both information from the tracking system and the muon chambers. Higher quality muons, which have track segments in both systems that are matched together, as well as the lower quality muons where a well-reconstructed track is matched to hits in the muon chambers, are used in this analysis. The quality selection is done on the track segments (and on the muon chamber segments for the “global” muons), while requiring small energy deposits in the calorimeters.

The leptons are required to be isolated. The total isolation variable is defined as the sum in a cone of size  $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$  around the lepton candidate, of additional charged candidate and neutral candidates, while subtracting out the

expected pileup contribution:

$$I = \frac{1}{p_T^\ell} \left( \sum p_T^{\text{charged}} + \max \left[ 0, \sum p_T^{\text{neutral}} + \sum p_T^\gamma - p_T^{\text{PU}} \right] \right), \quad (8.1)$$

where the sum runs over charged hadron candidates (charged), neutral hadron candidates (neutral) and photon candidates ( $\gamma$ ) within the cone. The summations are the scalar sum of the transverse momenta of the candidates.

For electrons, the pileup contribution is estimated using the median energy density in an event inside of clustered jets, multiplied by the “effective area” of the lepton [69–71], tuned as a function of pseudorapidity. For muons, the pileup contribution is estimated by averaging all the transverse momenta of the charged candidates not originating from the primary vertex. The isolation requirement for both types of leptons is that the total isolation is less than 40% of the lepton transverse momentum.

Additionally, a lepton 3D impact parameter significance criterion is applied, in order to reject lepton candidates coming from pileup collisions. The impact parameter significance is defined as the impact parameter divided by its uncertainty, with respect to the primary vertex, as illustrated in Figure 8.2. If the significance is greater than 4, the lepton candidate is discarded.

The CMS detector simulation offers an outstanding ability to accurately represent leptons in the apparatus, which results in a precise lepton calibration. In order to validate the calibration, the masses of various dilepton resonances ( $Z, \Upsilon, J/\Psi$ ) are extracted and compared to the best known values [47]. The result is shown in Figure 8.3. For both electrons and muons in the kinematic ranges we are interested in, the reconstructed dilepton resonance mass agrees with the known values to the level of 0.3%, indicating that the simulation successfully reproduces the calibration in data to a good precision. We also verified that the resolution for each type of lepton is well reproduced in the simulation, at the 5% level.

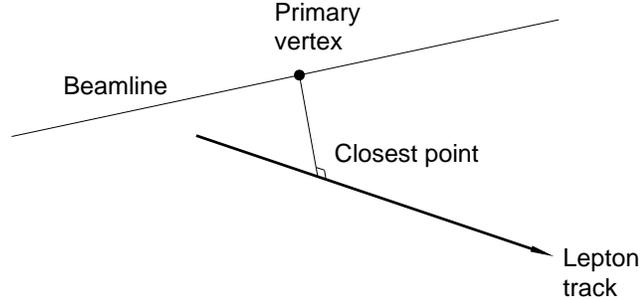


Figure 8.2: Illustration of the lepton 3D impact parameter significance. The impact parameter is defined as the distance of the closest approach between the primary vertex and the lepton track. Depending on the track quality, we can define an uncertainty on the impact parameter. The analysis selects leptons based on the ratio between the impact parameter and its uncertainty. In other words, it is a test of the compatibility between the lepton and the reconstructed primary vertex.

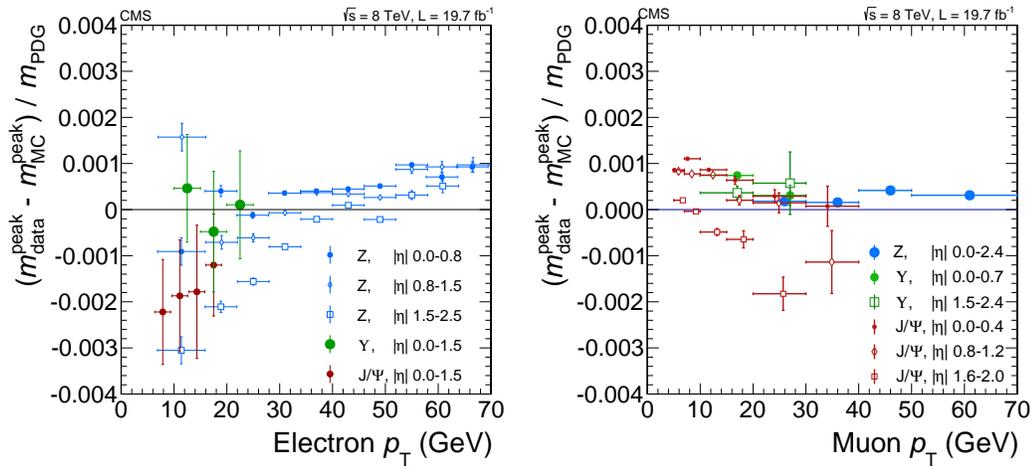


Figure 8.3: The validation of the lepton calibration, showing the comparison of the simulation and the data to the known masses of the  $Z$ ,  $\Upsilon$ , and  $J/\Psi$  resonances. The dilepton resonances are reconstructed and the extracted masses serve as a measure of the accuracy of the calibration. In the kinematic ranges we are interested in, we observe a good agreement between simulation and data at the 0.3% level compared to the best known masses for the resonances, indicating that the calibration obtained by using the simulation successfully reproduces the calibration obtained with data.

### 8.2.2 Final-state radiation recovery

A lepton can have final state radiation (FSR) where  $Z/\gamma^* \rightarrow \ell^+\ell^-\gamma$ . In this case it is useful to assign the energy of the radiated photons back to the lepton. Photons are selected in close proximity to the selected lepton candidate in  $\eta-\phi$ : either  $p_T > 2$  GeV within a cone of  $\Delta R < 0.07$ , or  $p_T > 4$  GeV within a cone of  $\Delta R < 0.5$  from the lepton, where  $\Delta R$  is defined as

$$\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2}. \quad (8.2)$$

Photons are required to have a relative isolation smaller than 1. In the case of photons the isolation variable is calculated by summing up all the transverse momenta of charged hadrons and neutral candidates within a cone size of  $\Delta R < 0.3$ :

$$I = \frac{1}{p_T^\gamma} \left( \sum p_T^{\text{charged}} + \sum p_T^{\text{neutral}} + \sum p_T^\gamma \right), \quad (8.3)$$

where the summation of  $p_T$  over photons excludes the candidate photon. All photons satisfying the above conditions are potential candidate final state radiation (FSR) photons.

We select at most one photon for the FSR recovery: if there is more than one candidate FSR photon with  $p_T > 4$  GeV, the highest  $p_T$  photon is selected; otherwise the one closest to the lepton candidate is used.

### 8.2.3 Kinematic selections

Once we have selected the set of candidate leptons per event, all possible pairings of the same flavor and opposite charge are examined. The pairing closest to  $m_Z$  is chosen as the first lepton pair that comes from a  $Z$  candidate, and the second pair is formed using the remaining leptons.

We consider the addition of FSR photons only if the addition of the photon brings the invariant mass of the three particles closer to  $m_Z$ . If the three-body invariant mass is above 100 GeV, the FSR candidate photon is discarded. Since the muon momentum does not include any radiated photons, final state recovery procedure is

expected to recover resolution.

The first lepton pair is required to have an invariant mass within  $40 < m_{\ell\bar{\ell}} < 120$  GeV, while the second pair is required to satisfy  $12 < m_{\ell\bar{\ell}} < 120$  GeV. Since the analysis is targeted at discovering a Higgs boson above 110 GeV, one of the  $Z$ 's is preferably on-shell. The first lepton pair invariant mass selection does not reduce the Higgs signal efficiency, while controlling reducible backgrounds from misidentified leptons. On the other hand, the invariant mass of the second  $Z$  is expected to be highly off-shell, and so pushing the mass cut down helps to increase the sensitivity. The lower bound is chosen such that the  $\Upsilon$  resonances do not contribute to our signal sample. If there are multiple combinations of leptons satisfying the requirement, the ambiguity is resolved by selecting the combination with the highest sum of transverse momentum.

Among all the selected leptons, we require one lepton to be above 20 GeV and one to be above 10 GeV. The cuts are driven by the trigger efficiency. In order to reject decays from low-mass resonances we also apply  $m_{\ell+\ell^-} > 4$  GeV, irrespective of the lepton flavor.

The overall signal efficiency for this selection within the geometrical acceptance is 30%, 43% and 62% for the  $4e$ ,  $2e2\mu$  and  $4\mu$  final states respectively, for a standard model Higgs boson signal with  $m_H = 125$  GeV. The signal efficiency as a function of the Higgs mass is shown in Figure 8.4.

## 8.2.4 The Standard Model Background Estimation

The normalization and shape of the irreducible backgrounds are estimated from simulations. The cross sections at next to leading order (NLO) for the  $q\bar{q}$  initial state, and at leading order (LO) for the  $gg$  initial state, are evaluated with MCFM [72]. The uncertainty from the simulation prediction is completely dominated by Monte Carlo statistics.

The irreducible background from double-parton interactions is estimated using the PYTHIA generator [73], with an overall cross section calculated as the product of  $Z$

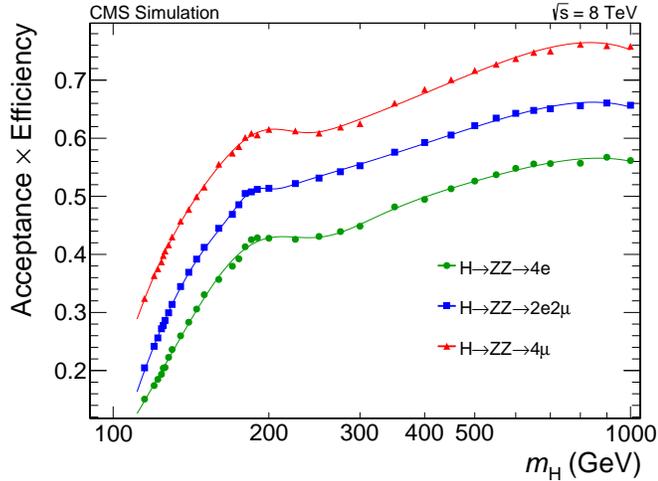


Figure 8.4: The efficiency of the analysis selection for different final states as a function of the Higgs boson mass. For a standard model Higgs at 125 GeV, the overall selection efficiency including acceptance effects is 30%, 43% and 62% for the  $4e$ ,  $2e2\mu$  and  $4\mu$  final states respectively

cross section with the probability of a second Drell-Yan (DY) interaction. The probability of the second DY interaction is estimated by the ratio between the Drell-Yan (DY) cross section and the phenomenological minimum bias effective cross section, which is measured at the LHC at 7 TeV to be 15 mb. This assumes that the second partonic interaction is independent of the first DY interaction, and that the probability of having a DY interaction can be estimated by the ratio of the DY cross section to the total  $pp$  cross section.

For the reducible background two methods are used. We extrapolate the yield from a control region (defined below) to the signal region using a parametrization of the lepton mis-identification probability as a function of the lepton transverse momentum.

The first method uses oppositely signed leptons. The control region in this case is defined by using a first lepton pair candidate with two same flavor oppositely charged leptons, that pass a relaxed set of selection criteria. Here we define two subsamples: 2-pass-2-fail (2P2F), where two leptons pass all the selections, and the other two pass the loose selection but fail further requirements; the other analogous sample is the 3-pass-1-fail (3P1F) sample, where only one lepton fails the lepton identification

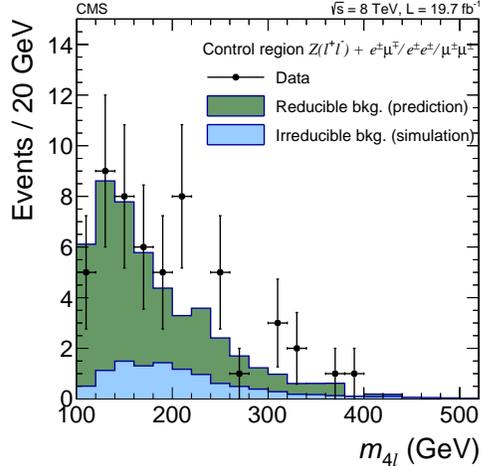


Figure 8.5: The validation of the reducible background estimation using the opposite-sign lepton method. The prediction of the reducible background is in agreement with the data, within uncertainties.

criteria.

The 2P2F sample is used predominantly to estimate events with two prompt leptons, while the other control sample is used to estimate the background contributions with three prompt leptons, such as leptonic decays of  $WZ$  production. Yields in each of the control samples are weighted by the lepton fake rate.

The second method uses same-sign leptons. The control region in this case consists of a good first lepton pair candidate, with two additional leptons with the same flavor but same charge. Since there is no signal contribution in this control sample, we expect the leptons to be fake leptons. The expected contribution from the reducible background is then obtained from the same-sign control sample weighted by the lepton fake rate, and the expected ratio of same sign to opposite sign pairs, taken from simulations.

A validation of the method is shown in Figure 8.5, where we compare the prediction of background with data, in a control region where the selection is the same as the main analysis selection, except that the second dilepton pair is formed from a wrong combination of flavor or charges ( $\mu^\pm\mu^\pm, e^\pm e^\pm, e^\pm\mu^\mp$ ). We see a reasonable agreement between the prediction and the data, within the uncertainties.

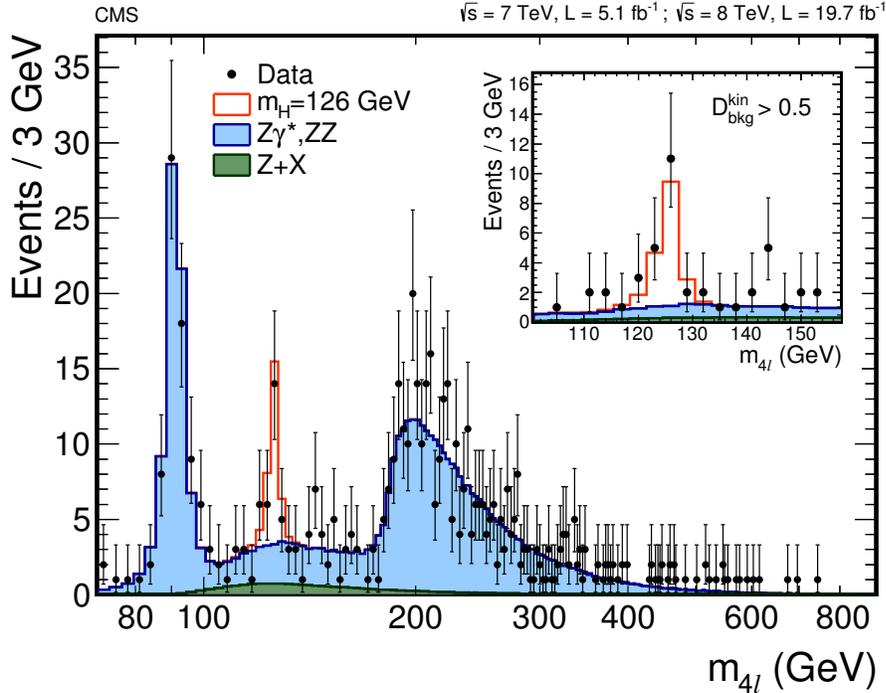


Figure 8.6: The  $H \rightarrow 4\ell$  mass spectrum from the full CMS Run I dataset. The expectation is compared directly to data without a fit to extract the normalization from the data. A clear excess is observed at 125 GeV over the standard model background. The background modeling is good over the full  $m_{4\ell}$  mass range, including the  $Z$  resonance peak at  $m_Z$ , as well as the  $q\bar{q} \rightarrow Z\gamma \rightarrow 4\ell$  and  $q\bar{q} \rightarrow ZZ \rightarrow 4\ell$   $t/u$ -channel contributions.

### 8.3 The Four Lepton Analysis Results

An excess is observed relative to the background in the  $4\ell$  channel around 125 GeV, as shown in Figure 8.6. We show the cross section limit in Figure 8.7 and the local significance of the excess in Figure 8.8. A standard model-like Higgs boson of mass between 114.5 – 119.0 GeV and 129.5 – 832.0 GeV is excluded at the 95% confidence level using the  $CL_s$  [74, 75] statistical approach. The local significance of the excess is expected to be  $6.6\sigma$ , while the observed significance is  $6.9\sigma$ .

We show an event display of one of the selected candidates for a  $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$  event in Figure 8.9.

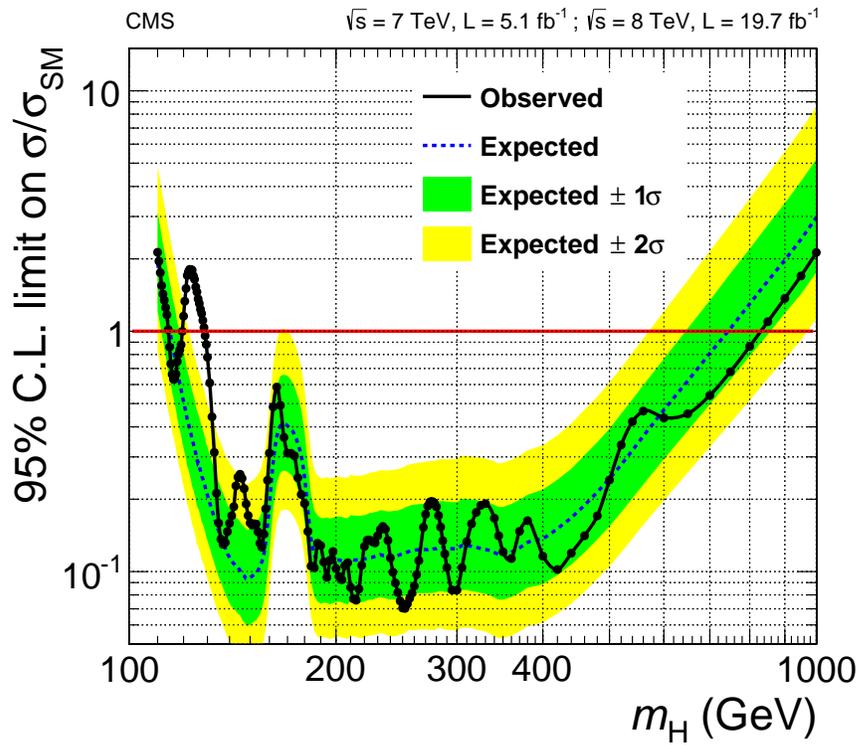


Figure 8.7: The limit on the  $H \rightarrow 4\ell$  production cross section is shown. With the analysis with CMS Run I data, we are able to exclude a standard model Higgs-like resonance between 114.5–119.0 GeV and 129.5–832.0 GeV at 95% C.L. The limit on the production cross section at 125 GeV is 1.8 times the standard model prediction.

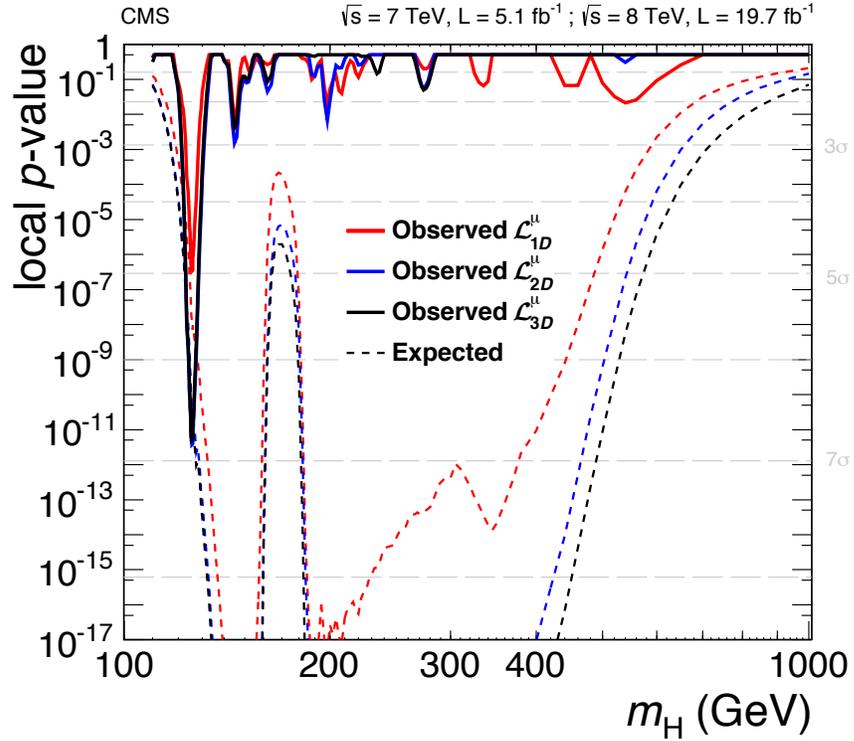


Figure 8.8: The one-sided local  $p$ -value is shown as a function of different  $m_H$  hypotheses. The  $p$ -value for the data is defined as the probability of the background-only hypothesis fluctuating to mimic what we observe in the data. The points on the dashed line represent the expectation of the  $p$ -value for a standard-model like signal hypothesis, with signal strength as predicted by the standard model.

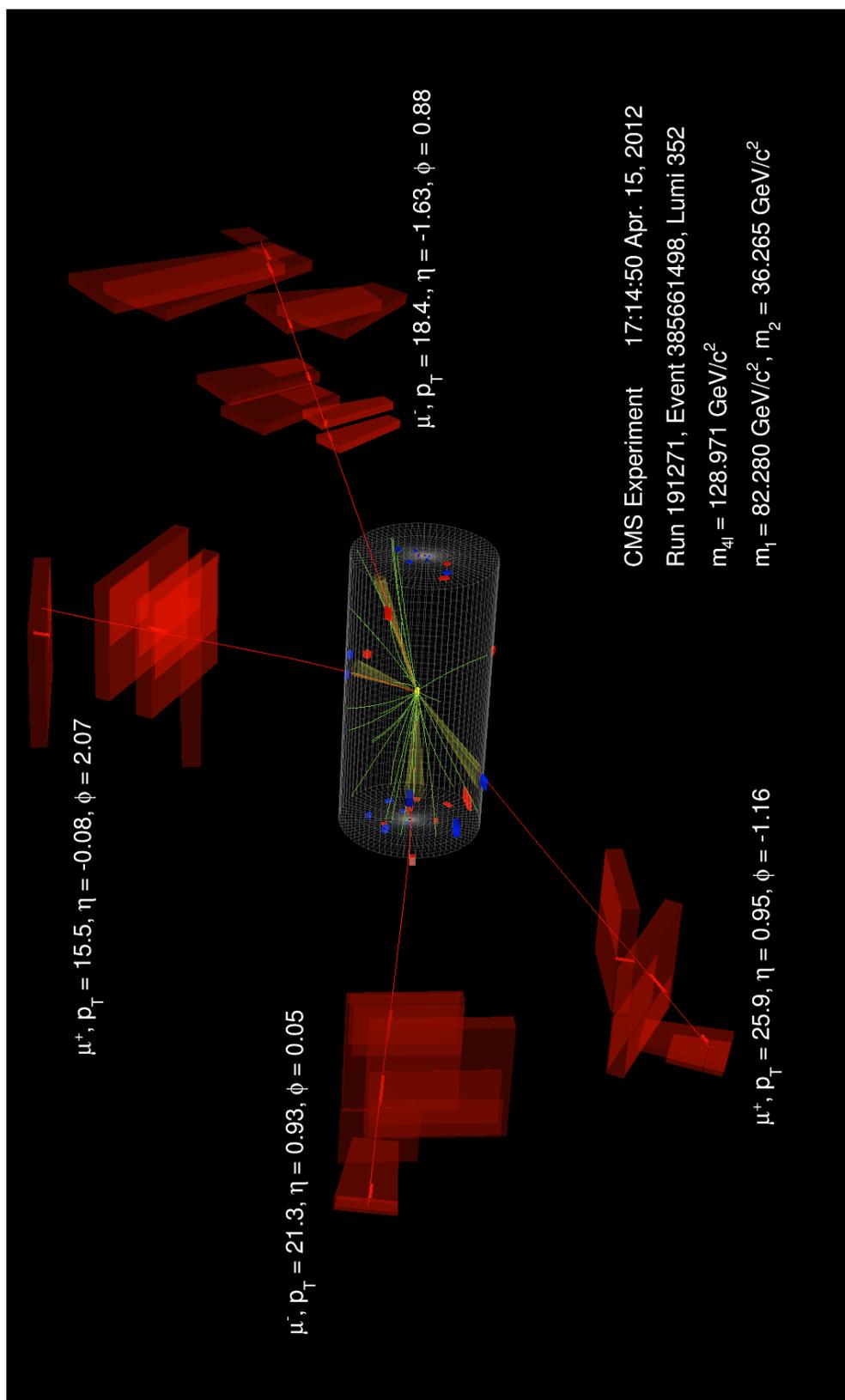


Figure 8.9: The display of a  $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$  candidate event from the CMS Run I data taken in 2012.

# Chapter 9

## The Higgs Boson Properties

### 9.1 Introduction

With the observation of a Higgs-like boson established in multiple channels, the focus of the work has moved to characterizing the properties of the resonance. Of particular interest are the mass, spin/ $CP$  properties, cross section and couplings to other standard model particles. Efforts have been pursued on all fronts, which we summarize in this chapter.

Before we describe the ongoing work to characterize this resonance, there are a few inferences on the properties of this particle that can be made from the decay products.

The particle is a boson and is electrically neutral. This can be seen from the  $4\ell$  and  $\gamma\gamma$  final states. Since the final state in these channels is closed (fully reconstructed), and a clear peak is observed in these final states, the total spin can only be an integer.

The spin of this particle can be deduced from the observation of it decaying into two photons. Since photons are massless there is no longitudinal polarization. It is impossible to line up the helicity of photons to match any spin state of a spin-1 particle, also known as the Landau-Yang theorem [76, 77].

## 9.2 The Higgs Mass Measurement

To measure the mass of the resonance we use the channels with the best measured final state particles, namely leptons and photons. Therefore the  $4\ell$  and  $\gamma\gamma$  channels are the best candidates for the mass measurement. Analyses [78, 79] have been carried out in CMS in these channels.

The  $4\ell$  channel analysis [79] is done primarily by using three observables. In order to distinguish the  $H$  signal from other standard model background processes, a discriminant is constructed as the ratio between the signal and background likelihoods. The shape of this distribution gives good separation power between the signal and background. In addition to the discriminant, the two other observables include the  $4\ell$  mass and an estimator of the event-by-event mass resolution  $\sigma_M$ , estimated by the kinematics and quality of the identified leptons. The introduction of this observable allows us to more heavily weight events with good mass resolution, and to give less weight to events with leptons in the kinematic regime with lower expected resolution. The analysis proceeds by constructing templates from simulated Monte Carlo events based on these three observables, and performing a fit to the data. The result for the mass measurement in the  $4\ell$  channel is shown in Figure 9.1.

The mass measurement in the di-photon channel [55] is done by carefully modeling the line shape, and by performing a fit to the signal using this shape, on top of the large background in the  $m_{\gamma\gamma}$  spectrum. The signal line shape is modeled by a Gaussian distribution, while the non-peaking background shape is modeled by a 5-th order polynomial. The fit is performed simultaneously in different sub-samples classified by the number of jets and the quality of photons. The uncertainty in the mass measurement is performed by a scan of the hypothetical mass and find the region with 68% C.L. We show the result in Figure 9.2.

A combination analysis has been carried out using the  $\gamma\gamma$  and  $ZZ$  channels, shown in Figure 9.3. We observe some tension between the mass measured in these two channels, at the  $1.6\sigma$  level. The best fit combined mass is  $125.0^{+0.26}_{-0.27}$  (stat.)  $^{+0.13}_{-0.15}$  (syst.) GeV.

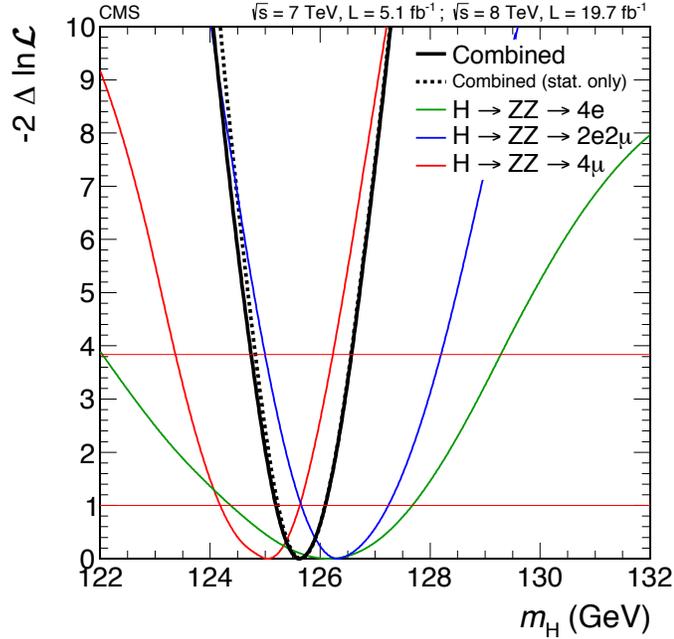


Figure 9.1: Summary of the CMS Higgs mass measurement in the  $4\ell$  final state. Curves of different colors represent the measurement done with only one final state lepton combination. Measurements from different final states agree with each other at the one sigma level. The combination of all channels is shown as the black line.

### 9.3 The Higgs Couplings Measurements

A common framework [80] to parametrize the potential deviation of the Higgs couplings to various particles from standard model predictions has been formulated by the LHC Higgs cross section working group (LHCHSWG). This framework assumes one single resonance in the narrow width approximation, and a standard model like tensor structure in the Lagrangian.

The ratio of the observed rate of a Higgs boson decaying to different final states, relative to their respective standard model expectations, is denoted by  $\kappa^2$  with different subscripts. There is a  $\kappa^2$  parameter for each of the observable tree-level Higgs couplings to vector boson and fermion pairs:  $\kappa_W^2, \kappa_Z^2, \kappa_t^2, \kappa_b^2, \kappa_\tau^2$ , as well as  $\kappa$  parameters for the effective couplings involving loops of other particles, for example  $\kappa_{gg}^2$  and  $\kappa_{\gamma\gamma}^2$ . One can also assume some additional constraint to reduce the dimensionality, for example  $\kappa_f^2$  which assumes a uniform deviation of Higgs couplings to all fermions, as well as  $\kappa_V^2$  to probe universal deviations of couplings to vector bosons ( $Z$  and  $W$ )

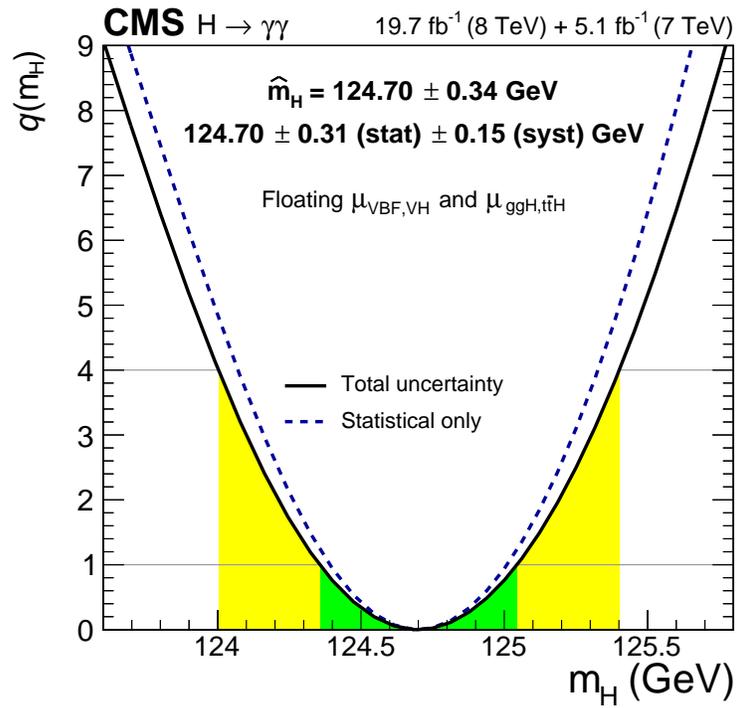


Figure 9.2: Summary of the CMS Higgs mass measurement in the  $\gamma\gamma$  final state. The green and yellow regions represent the  $1\sigma$  and  $2\sigma$  intervals respectively. There is a slight tension of  $1.6\sigma$  between the results from the  $4\ell$  channel and the  $\gamma\gamma$  channel.

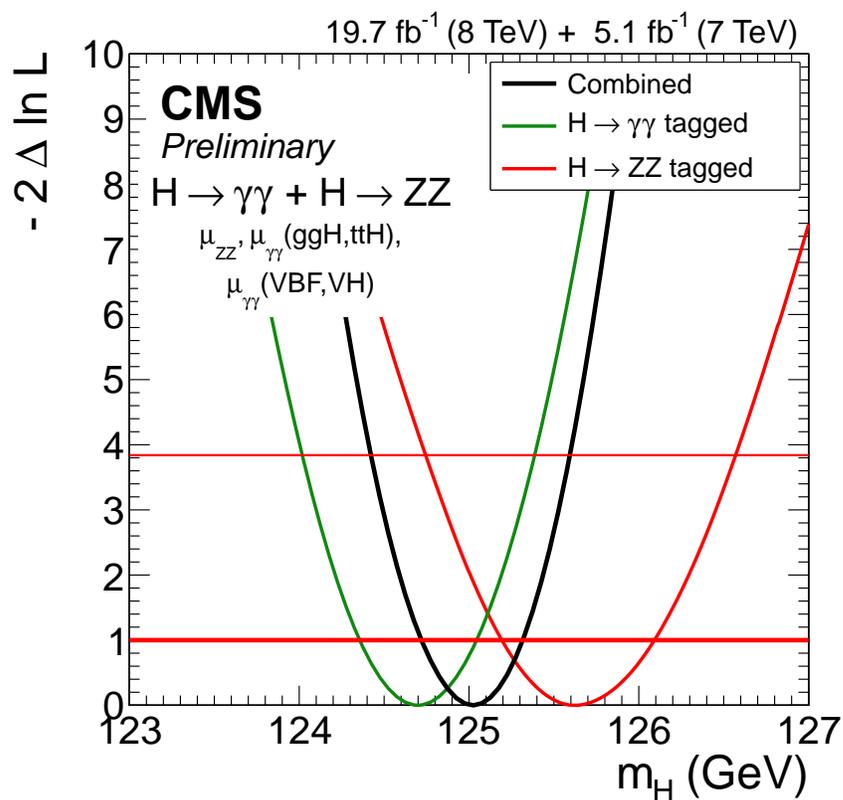


Figure 9.3: Summary of the CMS Higgs mass measurement from the high resolution channels. The red line is the result from the  $4\ell$  channel, while the green line is the result from the  $\gamma\gamma$  channel. Assuming the two channels arise from the same resonance, the combination between two channels is shown in black. The combined mass measurement is  $125 \pm 0.3$  GeV.

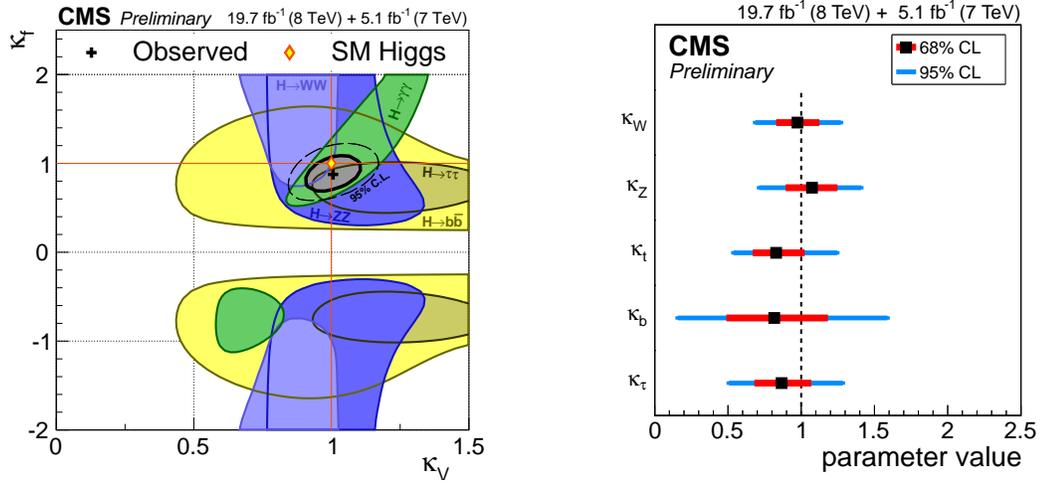


Figure 9.4: Selected CMS results on measuring the Higgs couplings. On the left the combined result using all channels is shown, assuming a universal  $\kappa_f$  and a universal  $\kappa_V$ . The standard model point (1, 1) is indicated as the yellow diamond, which is in agreement with the combined  $1\sigma$  contour in gray. In the right plot the result from a combined fit floating  $\kappa_W$ ,  $\kappa_Z$ ,  $\kappa_t$ ,  $\kappa_b$  and  $\kappa_\tau$  simultaneously is shown. The result is again compatible with the standard model expectation.

assuming custodial symmetry [81, 82].

This framework is useful to parametrize deviations from the standard model. However, because of the many assumptions behind the definition of these parameters, if a significant deviation is seen, we cannot interpret it directly as arising from modifications to the couplings themselves. We should therefore formally examine all potential alternative models, in order to have a better picture of the true nature of any coupling deviations.

The results from all channels in CMS [83, 84] are combined and fitted globally to obtain results for these coupling deviation parameters  $\kappa^2$ . Selected results are shown in Figure 9.4. In the left panel a global fit allowing only  $\kappa_f^2$  and  $\kappa_V^2$  to float is shown. As shown in the figure, the symmetry between positive and negative values of  $\kappa_f$  is broken only by the measurement of the  $\gamma\gamma$  final state, as a result of the destructive interference between the  $W$  loop and the  $t$  loop. In the right panel the fit to the five tree-level  $\kappa^2$  is shown. We do not see a significant deviation from the standard model expectations, in any of these parameters.

## 9.4 The Higgs Width Measurement

A sophisticated analysis [85] has been designed to probe the tiny Higgs width (4.2 MeV) using the interference between the signal and the background. The standard model Higgs at 125 GeV is mostly produced from the  $gg$  initial state, which interferes with the background process initiated by the same initial state. In the Higgs resonance region the contribution from the  $gg$ -initiated background process is small, but as shown in Figure 9.5, it increases as the invariant mass becomes larger. The key observation that makes this measurement possible is that while the resonance cross section is a function of the Higgs width  $\Gamma_H$ , the interference effect does not depend on the Higgs width in the highly off-shell region where  $m_{4\ell} > 2m_Z$ . The ratio of the resonant cross section to the size of the interference allows us to measure the width of the Higgs.

CMS has carried out this analysis with a cut-based approach [86] that uses the mass spectrum in both the  $4\ell$  and  $\ell\bar{\ell}\nu\bar{\nu}$  channels. Since we do not have sensitivity to measure the Higgs width yet, we currently set a limit on the size of Higgs width, assuming that everything is standard model, and with the mass set to the best-measured value 125.6 GeV in the  $4\ell$  channel. We obtain a limit of to be  $5.4 \times \Gamma_H$  at 95% C.L.. The likelihood scan is shown in Figure 9.6.

## 9.5 The Higgs Spin Inference

We can explicitly test the spin of the Higgs-like resonance by use of pair-wise model testing [87]. The technique compares two model hypotheses, and use a measure of “distance to each model” to quantify whether the data is closer to one model or the other.

This distance of measure is often formed with kinematic discriminants, as done in CMS’ and ATLAS’ analyses [35, 79, 88]. The kinematic discriminant employed in this case is the ratio of the truth-level differential cross section, recast into the range of 0 and 1:

$$D_1 \equiv \frac{P_0}{P_0 + cP_1}. \quad (9.1)$$

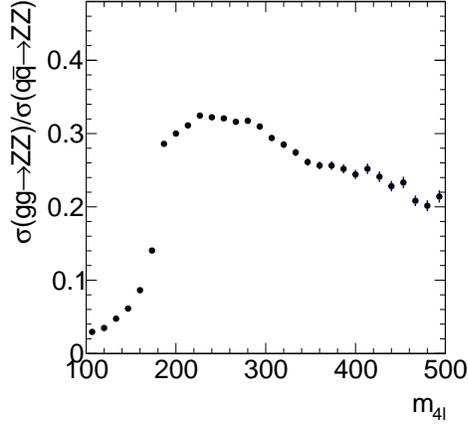


Figure 9.5: The ratio of the  $gg \rightarrow ZZ$  to the  $q\bar{q} \rightarrow ZZ$  background, as a function of  $m_{4\ell}$ . In the range relevant to the measurement of Higgs boson properties, the  $gg$  initial state continuum background is order of 4% compared to the  $q\bar{q}$  initial state. Compared to the analysis leading to the Higgs discovery, it is more relevant in analyses where the  $m_{4\ell} > 2m_Z$  region is important, for example the high-mass second Higgs search, and the measurement of the Higgs boson width.

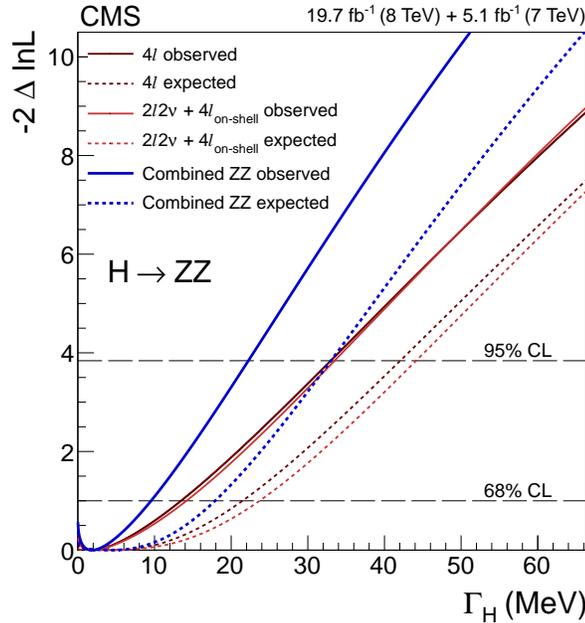


Figure 9.6: The result on Higgs width measurement from interference between signal and background in the high  $m_{4\ell}$  region. The purple and the red lines indicate result from individual final state:  $4\ell$  and  $\ell\bar{\ell}\nu\bar{\nu}$  respectively. The combined result from both channels is shown in blue. The upper limit on Higgs width at 95% C.L. is  $5.4 \times \Gamma_H$ .

The subscript 0 indicates null hypothesis, or the standard model signal hypothesis, while the subscript 1 indicates alternative hypothesis. Since we don't have access to the quantities before the detector acceptance and smearing, the reconstructed quantities are used for the evaluation of the truth-level probabilities. An additional discriminant is constructed to distinguish between the standard-model and the background hypotheses:

$$D_{Bkg} \equiv \frac{P_0}{P_0 + c'P_B}, \quad (9.2)$$

where  $P_B$  is the background likelihood. Having constructed these quantities, the analysis proceeded by constructing templates using these quantities as axes, and filled with Monte Carlo generated events for different processes. One such example is shown in Figure 9.7. The height of this template is then used as the measure of distance to each model:

$$\mathcal{L}_i = \mathcal{L}_i(D_1, D_B), \quad (9.3)$$

where  $i$  represents templates from different hypotheses. The total “distance” of the dataset is the product of all single-event distances. Finally, the ratio of the dataset distance is used to distinguish one hypothesis from another.

There are many possible Lorentz structures for each of the spin-1 and spin-2 possibilities. All combinations of different terms with varying strength, even though not necessarily theoretically well-motivated, represent viable alternative models. The CMS analysis has tested through 12 spin-1 models and 30 spin-2 models as of date (Nov. 7, 2014). However in order to fully exclude the possibility that the underlying resonance has a certain spin (in the framework of hypothesis testing), one must test against all possible such alternative models.

We can also perform analyses that extract parameters for a more-complete picture on the property of this Higgs-like resonance, as described in later sections.

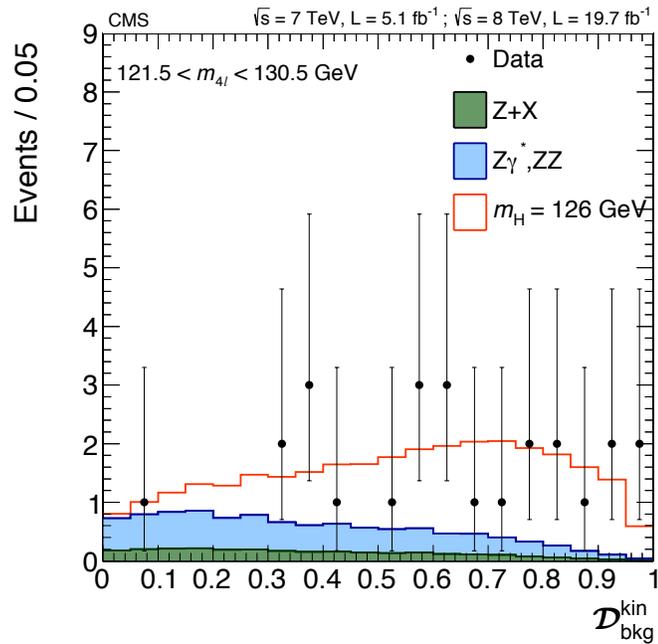


Figure 9.7: An example of signal and background template used in CMS analyses involving hypothesis testing on spin- $CP$  properties. This background discriminant is designed to separate signal and background. Standard model background components, shown in blue and green, favor a smaller value of the discriminant, while standard model signal in white favors a larger value. The expectation from Monte Carlo is compared with collected data.

## 9.6 The Higgs $CP$ Properties

Even in the realm of spin-0 models, there are the three Lorentz structures we can probe, as was described in Section 3.1. The LHC experiments have proceeded to measure mixtures between different terms, specifically in the  $ZZ$  sector where there are three terms:  $A_1^{ZZ}$ ,  $A_2^{ZZ}$  and  $A_3^{ZZ}$ .

An experimentally inspired way of defining the mixture, employed by CMS and ATLAS experiments, recasts the coupling parameters in terms of ratios of cross sections [79, 88]:

$$f_{a_i} \equiv \frac{|a_i|^2 \sigma_i}{\sum_j |a_j|^2 \sigma_j}$$

$$\phi_{a_i} \equiv \text{Arg} \left( \frac{a_i}{a_1} \right), \quad (9.4)$$

where  $a_i$  is the coupling parameter for the term of interest, representing the set  $A_{1,2,3}^{VV'}$  (discussed in more detail in Chapter 3) in the case where all coupling parameters are constant, and  $\sigma_i$  is the corresponding cross section where  $a_i = 1$  and the other couplings are set to zero. The term  $A_1^{ZZ}$  corresponds to the standard model tree-level  $HZZ$  coupling. This has the advantage of being reparametrization invariant, i.e., one can reparametrize the Lagrangian in many different ways, but this quantity would stay the same. The disadvantage is that while the coupling parameters are theoretically directly interpretable, these fractions are not. One would need to know the cross sections, as well as the selections used in evaluating these cross sections, making the interpretation of the result less straightforward. When there is a significant contribution to the cross section from the interference term, as is the case between  $A_1^{ZZ}$  and  $A_2^{ZZ}$ , this fraction loses the simple interpretation as the ratio of the contribution from different terms. It is nevertheless a viable parametrization to present the result and is used by the experiments.

While this method on fitting using templates works well to measure up to three parameters simultaneously, it becomes difficult to perform fits to more parameters

simultaneously, as the number of dimensions of the template increases quadratically to the number of parameters to fit. In order to perform such an analysis a novel method is needed, which is the main work of this thesis. Before describing the analysis framework, in the next chapter we discuss some of the challenges in the Higgs property measurement in the golden channel.

# Chapter 10

## The Golden Channel Challenges

### 10.1 The Challenge of the Twelve Observables

While the  $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$  channel offers a good signal to background ratio, and a large number of handles that have helped in the discovery of the Higgs boson resonance, a number of challenges need to be tackled. There are four leptons in the final state, each with three degrees of freedom (three momentum). In total there are 12 degrees of freedom. It is a challenge to accurately model all 12 highly-correlated observables; a challenge that was the focus of this thesis.

One common parametrization of the 12 degrees of freedom starts by categorizing observables into “decay” observables and “production” observables. Production observables refer to observables related to  $H$  production, and decay observables refer to the decay of  $H$ . There are 8 observables in the decay side and all of them are global-boost invariant. It is best visualized in the rest frame of 4 leptons, where we have incoming protons  $pp$  producing a back to back gauge boson pair  $VV'$  where each boson subsequently decays into 2 leptons:

1.  $m_{4\ell}$ : invariant mass of the 4 leptons.
2.  $m_1, m_2$ : invariant mass of the two lepton pairs.
3.  $\Phi_1$ : azimuthal angle between the  $pp \rightarrow VV'$  plane and the first  $V \rightarrow \ell\bar{\ell}$  plane.
4.  $\phi$ : azimuthal angle between the two  $V \rightarrow \ell\bar{\ell}$  planes.

5.  $\cos \Theta$ : angle between the  $pp$  and the  $VV'$  directions.
6.  $\cos \theta_1, \cos \theta_2$ : the decay angle of each of the negatively-charged leptons in their respective dilepton rest frame with respect to the boost direction.

The definition of these variables are summarized in Figure 10.1. The production observables consist of 4 observables:

1.  $\vec{p}_{T,4\ell}$ : the transverse momentum vector of the 4-lepton system in the lab frame. The two components are usually written as  $p_{T,4\ell}$  and  $\phi_{4\ell}$ ; namely the size and direction of the vector  $\vec{p}_{T,4\ell}$ .
2.  $y_{4\ell}$ : rapidity of the 4-lepton system in the lab frame
3.  $\phi_{\text{offset}}$ : rotation of the  $4\ell$  system along the  $z$ -axis in the  $4\ell$  rest frame before boosting to the lab frame. It is defined as the azimuthal angle of the lepton of the first lepton pair in the  $4\ell$  rest frame.

The linear angle combination  $\phi_{4\ell} + \phi_{\text{offset}}$  corresponds to an overall rotation of the system. We expect it to be flat for all models we are considering. The other combination,  $\phi_{4\ell} - \phi_{\text{offset}}$ , however, could in principle be non-trivial.

When there is extra activity in the event, for example extra jets, there is an ambiguity in how we define the incoming partons that attach to the  $4\ell$  system, as can be seen in Figure 10.2, the choice of diagrams affect some of the reconstructed observables. One common definition [89] is to “distribute” the vectorial sum extra activity evenly in both “sides” of the  $4\ell$  system. If the jets sum up to a total momentum  $\vec{p}$ , we would attach  $\vec{p}/2$  at the two sides, and calculate the parton momentum accordingly. This is illustrated in the last diagram in Figure 10.2.

## 10.2 The Four Lepton Parameter Space

Another challenge in this channel is the large number of parameters of interest. Even if we take the lowest order in the momentum expansion of the form factor, there are 7 parameters across different  $VV'$  sectors, and each can be complex in principle.

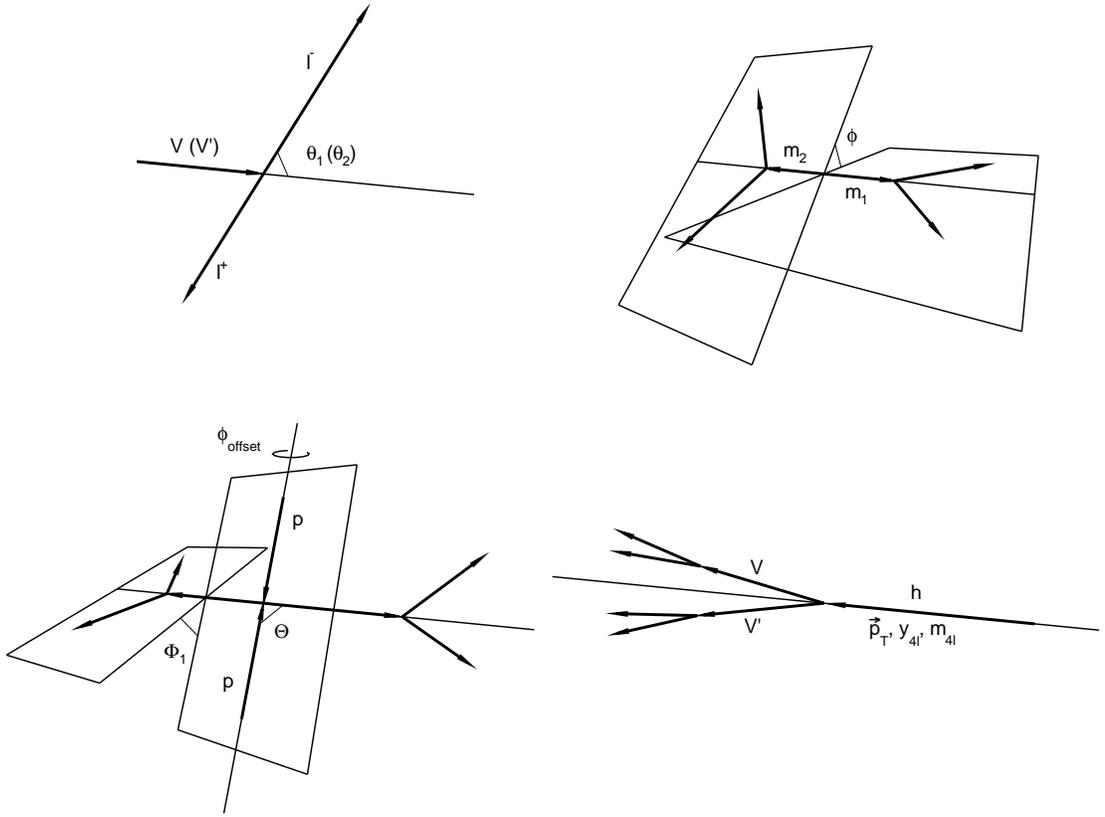


Figure 10.1: Definition of the 12 variables. In the rest frame of each  $V(V')$  we define a decay angle  $\theta_1(\theta_2)$  of the lepton with respect to the  $V(V')$  direction, as shown in the upper-left figure. In the upper-right figure we show the definition of the  $\phi$  angle in the  $4\ell$  rest frame, which is the opening angle between two  $\ell\bar{\ell}$  planes. The dilepton masses are also indicated. Adding in the incoming partons, in the lower-left figure, we can define three additional angles:  $\Phi_1$ , the opening angle between one of the  $V \rightarrow \ell\bar{\ell}$  planes and the plane formed by the incoming parton and  $VV'$ ;  $\Theta$ , the decay angle of the  $VV'$  with respect to the incoming parton; and  $\phi_{\text{offset}}$ , the rotation along the incoming parton axis. Finally in the lab frame there is the 4-momentum of the  $4\ell$  system, which we parametrize as  $m_{4\ell}$ ,  $y_{4\ell}$  and  $\vec{p}_{T,4\ell}$ , as can be seen in the lower-right figure.

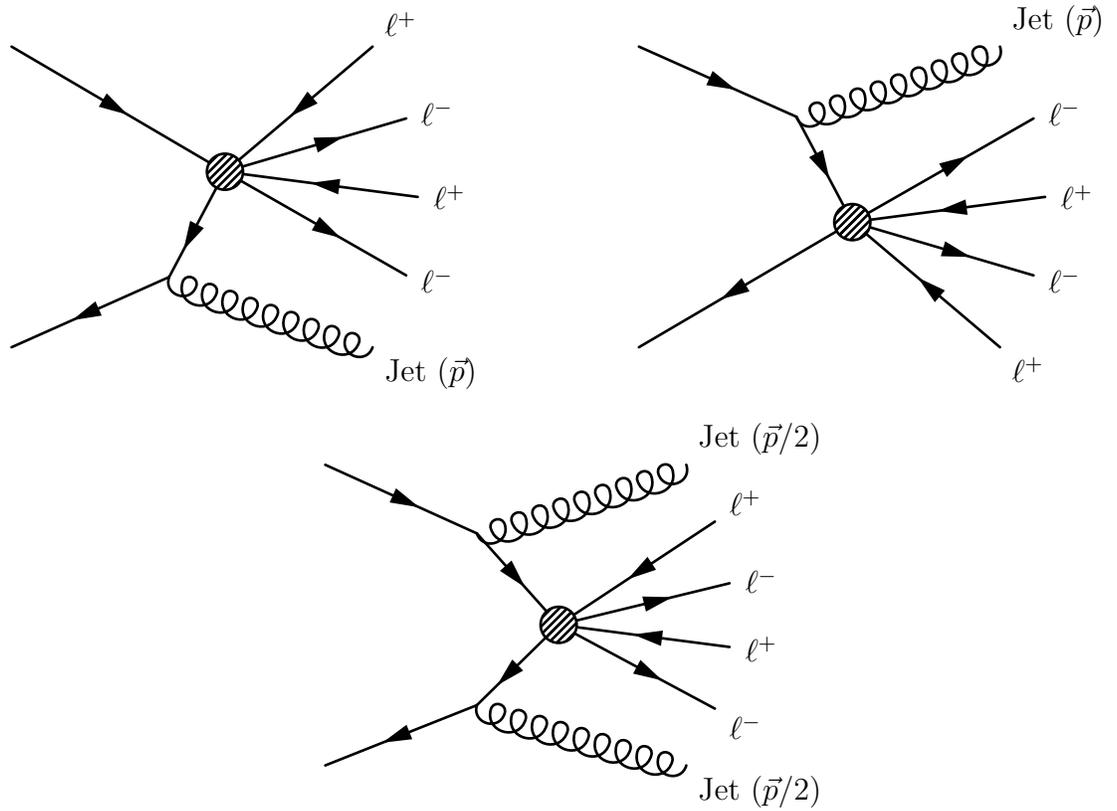


Figure 10.2: Ambiguity in assigning extra activities in the event with respect to the  $4\ell$  system. In the top plots we have two choices of assigning the extra activity (denoted collectively as “Jet”) on either side of the  $4\ell$  system. The choice changes the direction of parton lines attached to the  $4\ell$  system, and thus changes the values of some of the observables. One popular choice of assigning the extra activity is to imagine the vectorial sum of the extra activities are evenly split on two “sides”.

The problem can be simplified by observing that the overall phase is not a physical parameter. It can be fixed to any arbitrary value we choose. Furthermore the overall scale of all the parameters is fixed by the observed rate, and is not our primary focus of the property measurements in this analysis. This leaves us with 12 effective parameters (or 6 in the case where all parameters are real). However, with the momentum dependent terms the number of parameters grows rapidly, and measuring all the parameters then becomes challenging.

With the large number of parameters of interest, a natural question arises: how many of the parameters should we measure at the same time? One possible strategy is to measure one parameter at a time, while fixing the other parameters to the standard model value. While this is the simplest measurement one can do, it can lead to potentially misleading results if not accompanied by complementary measurements. It is rare for a model to predict non standard model values for just one (or two) of the effective couplings. It is usually the case that we expect deviations in all the  $VV'$  sectors. In any given measurement, by setting the coupling parameters which we do not measure to standard model values, we neglect potential contributions that may lead to a bias.

In the case where the true underlying model is the standard model, another view of the parameter extraction procedure is that the statistical fluctuations of the true model could be modeled by the anomalous coupling parameters. Artificial results could thus be seen, if the model space we probe with the fit cannot properly describe the fluctuations. For example, if the underlying distribution of an angle observable is in fact flat, one cannot expect a good fit if we try to model the statistical fluctuations using only cosine functions and no sine functions.

To better illustrate this point we consider the following toy example, which is a 1D projected version of the  $Z\gamma$  sector, together with a tree-level term  $A_1^{ZZ}$ . Suppose there is only one observable  $\phi$ , and the true model is

$$f_0(\phi) = 350 - 7 \cos \phi + 42 \cos 2\phi, \quad (10.1)$$

and the function used to model fluctuation is

$$\begin{aligned}
f_1(\phi) &= (35680 - 140 \cos \phi + 980 \cos 2\phi)A_2^2 + (28490 - 7660 \cos 2\phi)A_3^2 \\
&+ (300.5 - 821.4 \cos \phi + 52.6 \cos 2\phi)A_2 + (-569 \sin \phi + 39.9 \sin 2\phi)A_3 \\
&+ (17180 \sin 2\phi)A_2A_3,
\end{aligned} \tag{10.2}$$

where the terms on the first line are the square terms of  $CP$ -even and  $CP$ -odd operators, and the terms on the second and third lines are the interferences between various terms. We can generate the data assuming a shape of  $f_0(\phi)$ , and have the option to fit for both  $A_2$  and  $A_3$  together, or just  $A_2$  alone, using  $f_0(\phi) + f_1(\phi)$ .

The result is shown in Figure 10.3. Since the true model is with  $A_2 = A_3 = 0$ , we see that the extracted values across many pseudoexperiments for  $A_2$  and  $A_3$  when both are floated, form a simple 2D Gaussian-like shape, indicating that this is an adequate basis to model the fluctuations. However if we fix  $A_3$  to zero and float only  $A_2$ , an interesting pattern emerges. For pseudoexperiments with  $A_3$  fitted to be close to zero, when both are floated, the distribution of  $A_2$  remains normal - even if we float only one parameter. However when the fit for  $A_3$  yields larger values, a bias starts to appear in  $A_2$  if it's the only one floated. This kind of behavior happens even with a simple toy 1D analysis. In the full analysis which contains multiple observables, the biases that result from the use of an inadequate basis can be complex to disentangle.

While we can never exhaust all possible terms in all effective theories, we should also perform a complementary analysis where more terms are fitted at the same time. The precision at which we can pin down values of effective couplings might be worse, but by following this procedure we are better protected against potential biases caused by an overly restrictive choice of the model. For understanding the underlying physics, providing both types of measurements, with different model space assumptions, is useful.

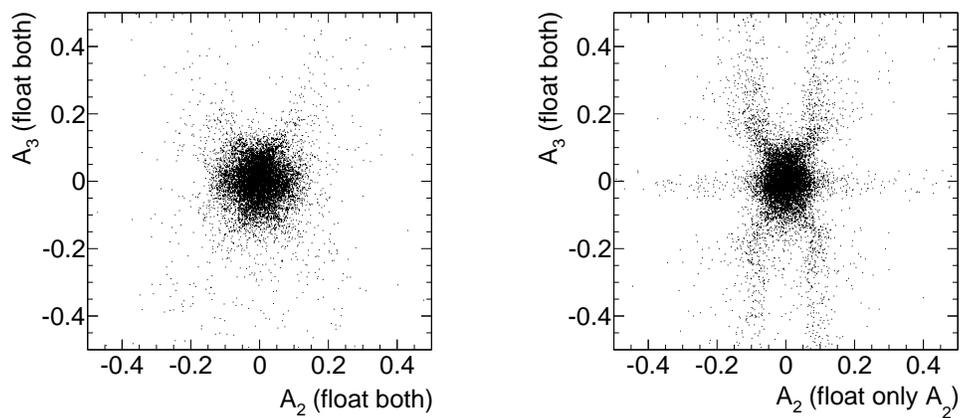


Figure 10.3: Toy study in the  $Z\gamma$  sector, to show that incomplete modeling of potential fluctuations can lead to artificial features in the result. Here the true model is the tree-level  $A_1^{ZZ}$  coupling, and the fluctuations are modeled by floating the  $A_2^{Z\gamma}$  and  $A_3^{Z\gamma}$  operators. In the case where the fluctuations are better described by  $A_3^{Z\gamma}$ , artificial features are seen when only  $A_2^{Z\gamma}$  is floated, since the model space probed cannot adequately sample the fluctuations.

# Chapter 11

## Super Fast Detector Parametrization

In CMS there is an official event simulation and reconstruction software package [90] that utilizes GEANT4 [91, 92] for simulation. It propagates the generated particles through a full description of the CMS detector, simulates showers in the calorimeters, and digitizes the responses in a manner closely approximating the digitization of the real data. It has been proven to be very successful in describing the digitized responses of the detector. The simulated and digitized detector responses are then passed to a chain of event reconstruction algorithms which interpret the digitized detector responses into energy deposits, and physical objects such as leptons and jets.

The full simulation takes a considerable amount of CPU time. It takes order of few minutes to simulate and reconstruct a single event in 2012 conditions. For some searches where a large number of model points need to be scanned through, it becomes too time consuming to prepare all the samples.

For studies where detailed detector responses are not the primary focus, a fast simulation [93] has been developed. The fast simulation parametrizes the detector responses as a function of the generated particles, and thus bypasses the time-consuming part of the event simulation through GEANT4. Although this yields an approximation to the detector responses, it is sufficient for many studies. Through the use of this technique, CPU usage is reduced by a factor of 100-1000.

For the purpose of studying  $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$  events, the fast simulation is still not fast enough for the algorithms described in the next few sections. However, since only the leptons are of interest in the event, we have developed a custom super fast simulation which parametrizes only the lepton response, and which goes from the generator level to the reconstructed level directly. The CPU usage for each event in this case is reduced to the order of tens of nanoseconds.

## 11.1 parametrization Of The Lepton Response

One important ingredient in the coupling measurement using the  $4\ell$  channel is the modeling of the lepton response in the detector. The modeling includes both the smearing of the momenta, and the efficiency/acceptance effects on the leptons. The model is extracted from CMS simulated samples for  $H \rightarrow VV' \rightarrow 4\ell$  and from simulated  $ZZ$  di-boson sample. Since leptons are well-measured objects in the CMS detector, the assumption that each lepton is independent of the others is reasonable.

The excellent CMS tracker directional resolution allows us to assume that the directions of the leptons are perfectly measured, so that we need only model changes in the magnitude of the lepton momenta. We define smearing factors  $c$  which are the ratios of detector level lepton momentum magnitude divided by the corresponding generator level momentum. The transfer function  $\tilde{T}$  is then defined as the probability density function that describes the chance that a lepton of a certain generator level momentum gets smeared by a factor  $c$ :

$$\frac{dP}{dc_i} = \tilde{T}(c_i|\vec{p}_i^G), \quad (11.1)$$

where the index  $i$  labels the lepton we examine. The momentum smearing transfer function  $\tilde{T}$  is normalized, so that integrating along all the possible smearing factors yields unity. In addition, the efficiency and acceptance effects are encapsulated in an

$\epsilon(\vec{p}_i^G)$  function. The total transfer function  $T$  is thus expressed as follows:

$$T(c_i|\vec{p}_i^G) \equiv \tilde{T}(c_i|\vec{p}_i^G)\epsilon(\vec{p}_i^G). \quad (11.2)$$

We construct each transfer function in bins of lepton pseudorapidity and transverse momentum. We use 15 bins in  $p_T$  and 16 bins in  $|\eta|$ , to better capture the dependence of the transfer function on the lepton momentum. For each bin, a histogram of the smearing factor  $c_i$  is filled. A double-sided crystal ball function provides an adequate description of the shape of these distributions. Some examples of the extracted transfer function are shown in Figure 11.1, for muons and electrons with various momenta.

We summarize the extracted width of the transfer function  $\tilde{T}$  for electrons and muons as a function of the lepton  $p_T$  and  $|\eta|$  in Figure 11.2. The width is expressed as percentages of the generator level lepton momentum. For electrons, the momentum resolution is good in the central region, and we see a worsening of resolution near  $|\eta| = 1.5$  which corresponds to the boundary between the barrel and endcap calorimeters. For muons this feature is not present since we do not rely on the calorimeters to measure muon energy. The resolution for muons is good throughout the whole region of interest, with some worsening close to the detector boundary at  $|\eta| = 2.4$ .

The efficiency function  $\epsilon(\vec{p}_i^G)$  is shown in Figure 11.3. Again, for electrons the boundary between barrel and endcap calorimeters is clearly seen, while for muons the efficiency is good overall. The efficiency is slightly worse at 8 TeV compared to 7 TeV, mainly due to a busier environment in the 8 TeV data, as a result of higher pileup.

## 11.2 Transfer Functions Validation

In order to validate the extracted transfer functions, we apply them to the standard model Higgs samples and compare distributions to CMS simulated samples. The distributions of  $m_{4\ell}$  and  $m_1$  are shown in Figure 11.4. These two observables are the most sensitive to lepton smearing effects, because of the sharp peak in the generator

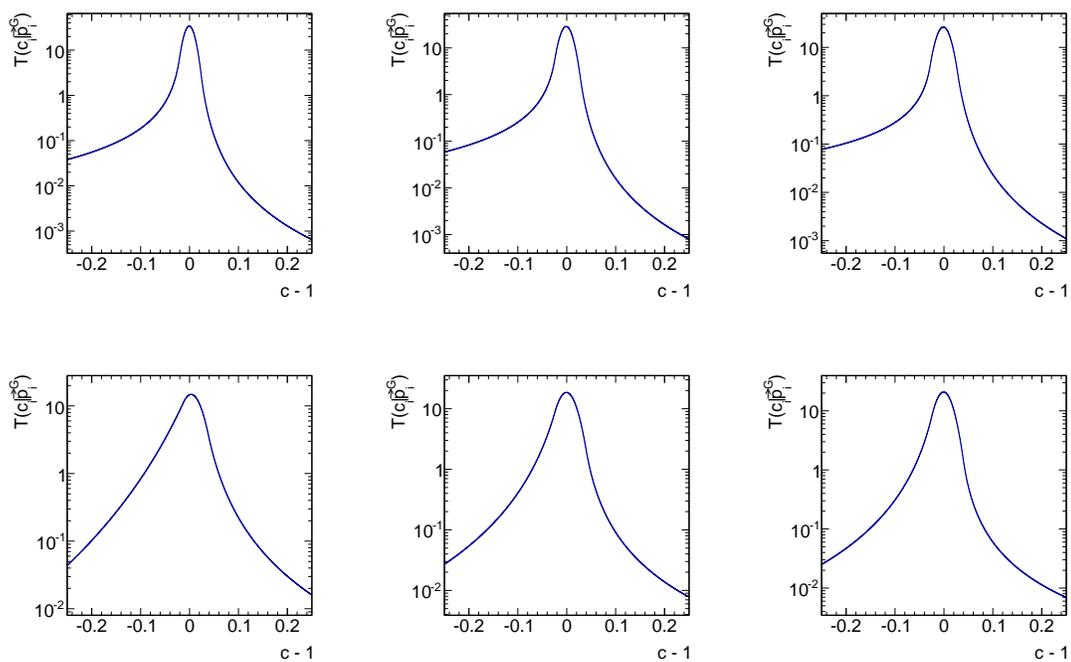


Figure 11.1: Examples of the extracted smearing transfer functions plotted as  $c - 1$ , for muons and electrons with different momenta. On the top row muons with  $p_T = 15, 30, 45$  GeV and  $|\eta| = 0.6$  are shown, while in the bottom row the same is shown for electrons. The width of the transfer function for electrons with higher energy is smaller, relative to those with lower energy.

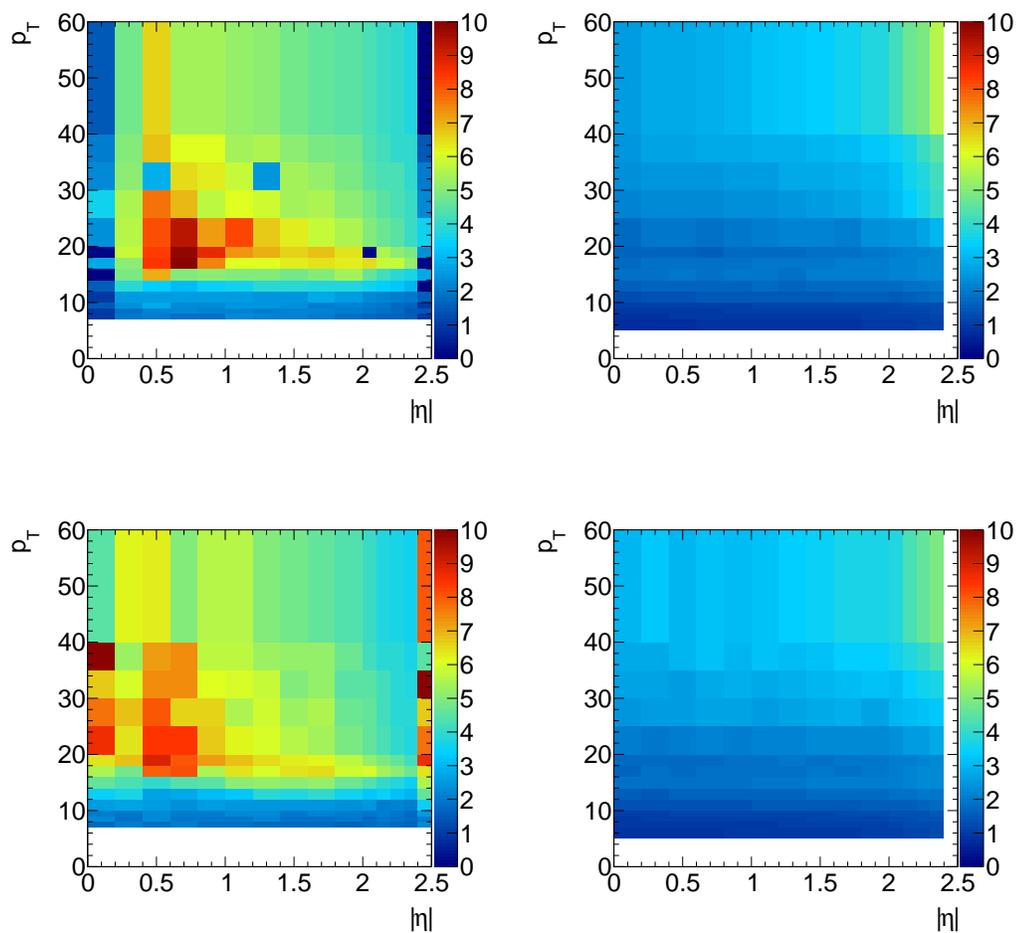


Figure 11.2: Width of the transfer function as a function of the lepton transverse momentum and pseudorapidity. On the left are the plots for electrons, and on the right are the muon results. The top row shows the 7 TeV transfer functions and the bottom row shows the transfer functions for 8 TeV. The binning used in the plots shown here is the same as the binning of the transfer functions used in the analysis. The color scale indicates the percentage relative to the lepton momenta.

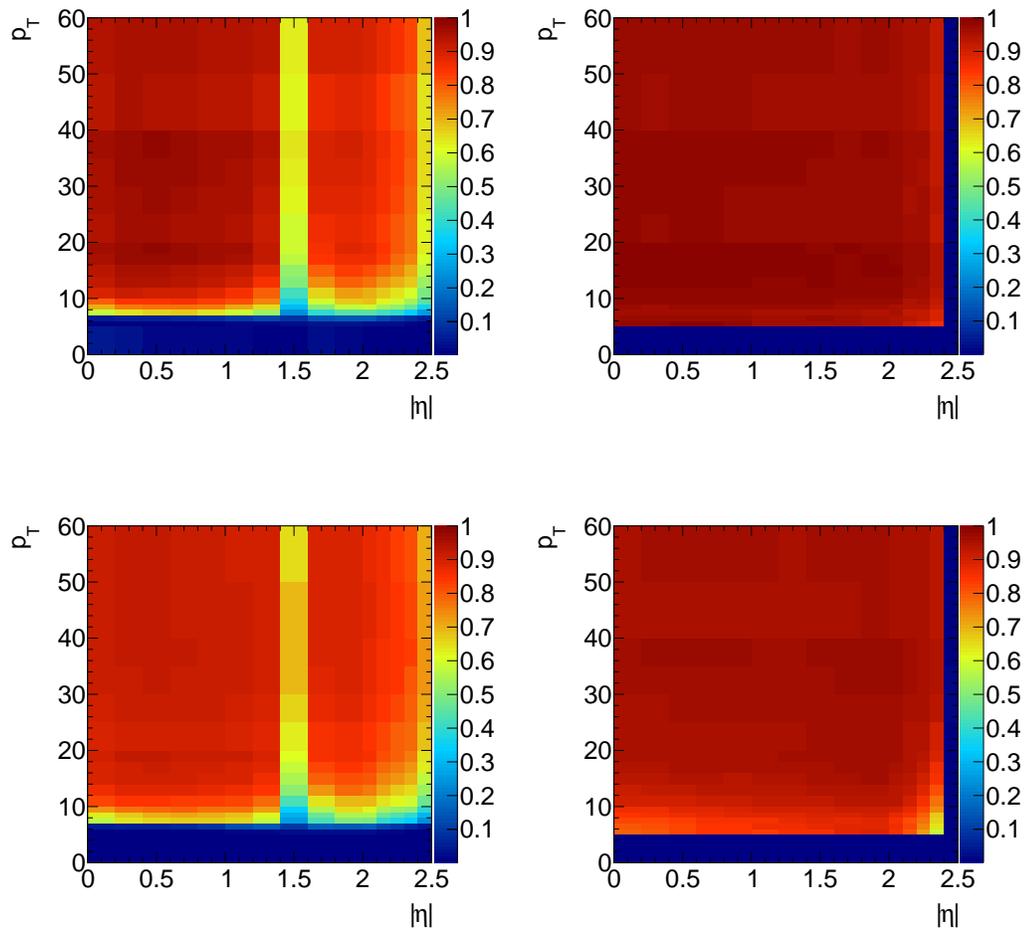


Figure 11.3: The efficiency for the leptons as a function of the lepton transverse momenta and pseudo-rapidities. The results for 7 TeV are shown in the top row, and for 8 TeV are shown in the bottom row. The effect of the boundary between the barrel and endcap calorimeters is visible in the electron efficiency function on the left. For the muons, on the right, the efficiency is good over the full range of kinematics relevant to this analysis.

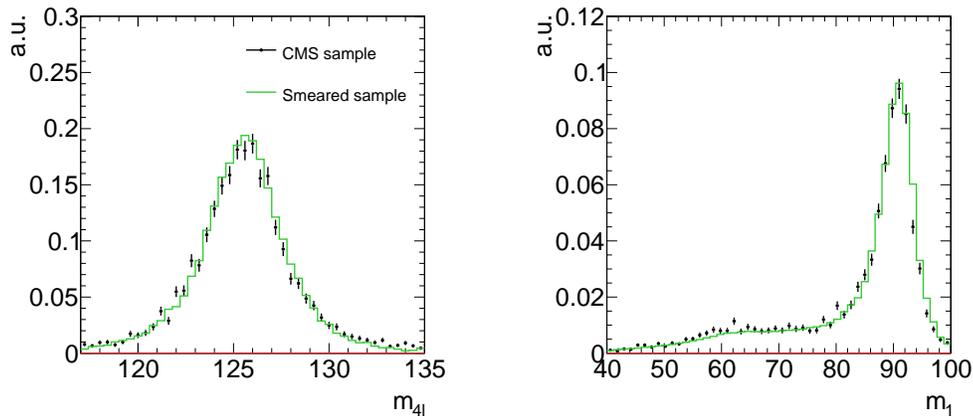


Figure 11.4: Comparison of the  $m_{4\ell}$  and  $m_1$  projections, showing the comparison between the fully simulated sample, and the smeared sample.

level distribution. The  $m_{4\ell}$  is modeled as a delta function at the generator level, and the shape comes entirely from the smearing of leptons. We see excellent agreement between the two samples, which are both distinctively different compared to the generator-level distributions. A stronger check will be described in Section 13.6.3, where we extract parameters using the reconstructed level *pdf* which incorporates the transfer functions as part of the detector modeling.

## Chapter 12

# 8D Likelihood Effective Higgs Couplings Extraction Framework in the Golden Channel

### 12.1 The Multi-Dimensional Method Overview

The goal of this analysis is to exploit the full potential of this channel. In order to do that we can first ask ourselves what we wanted to achieve if everything was possible. It is clear that the current existing methods have some limitations. Since a discriminant is constructed by taking the ratio of the likelihoods for two distinct hypotheses, it is only useful in measuring any mixture of the two hypotheses. In the case of multiple parameters of interest, for the existing approach one would need to construct many such discriminants in order to span the full parameter space. It is further complicated by the interference between different coupling parameters, which, in the current approach, uses additional discriminants to help in describing the shape. The number of discriminant dimensions needed to span the full space of parameters of interest increases rapidly, and the use of Monte Carlo-based templates is thus not computationally feasible.

There would be three types of potential improvements if everything were possible: (1) we would have a continuous 12 dimensional *pdf* as a function of the reconstructed quantities. Since there are only 12 degrees of freedom in the  $4\ell$  system, this would exhaust all possible correlations among the observables, and allow us to measure all

potential parameters of interest at once, without the need to pick discriminants, and (2) The *pdf* would be a continuous function of the parameters of interest, and (3) Evaluation of the *pdf* would be extremely fast, as there are a great many evaluations involved in the minimization process.

Once all of these are realized, one would simply perform a fit to the existing data and extract all the parameters of interest at the same time. In the following sections I will document the methods used in this thesis in detail, which make most of these improvements possible.

## 12.2 The Higgs and Other Standard Model Four Lepton Differential Cross Sections

Differential cross sections for the decay to four leptons have been calculated by Vega-Morales et al [29, 33] to leading order. Instead of choosing a different language to calculate different terms in the Lagrangian including interferences, a novel approach was developed to calculate all the terms in a covariant way, including contributions from different intermediate vector bosons. The expression obtained in this way contains all possible distinct interference signal terms. Some example distributions are shown in Figure 12.1.

All contributions with the  $q\bar{q}$  initial state from the SM are calculated (again at LO). This includes t/u-channel di-boson processes (including  $ZZ$ ,  $Z\gamma$  and  $\gamma\gamma$ ) as well as the  $s$ -channel process  $q\bar{q} \rightarrow Z/\gamma \rightarrow 4\ell$ , and all the interferences among these terms.

For the measurement of a Higgs-like boson at 125 GeV, the background is dominated by  $q\bar{q}$  initial state, while the SM signal is dominated by the  $gg$  initial state. As a result, we can ignore the interference between the signal and the background for this analysis. This signal-background interference, however, will need to be considered in the era of HL-LHC.

It is not possible to write down the differential cross section for the production observables, the four-vector of the  $4\ell$  system, since it involves parton distribution

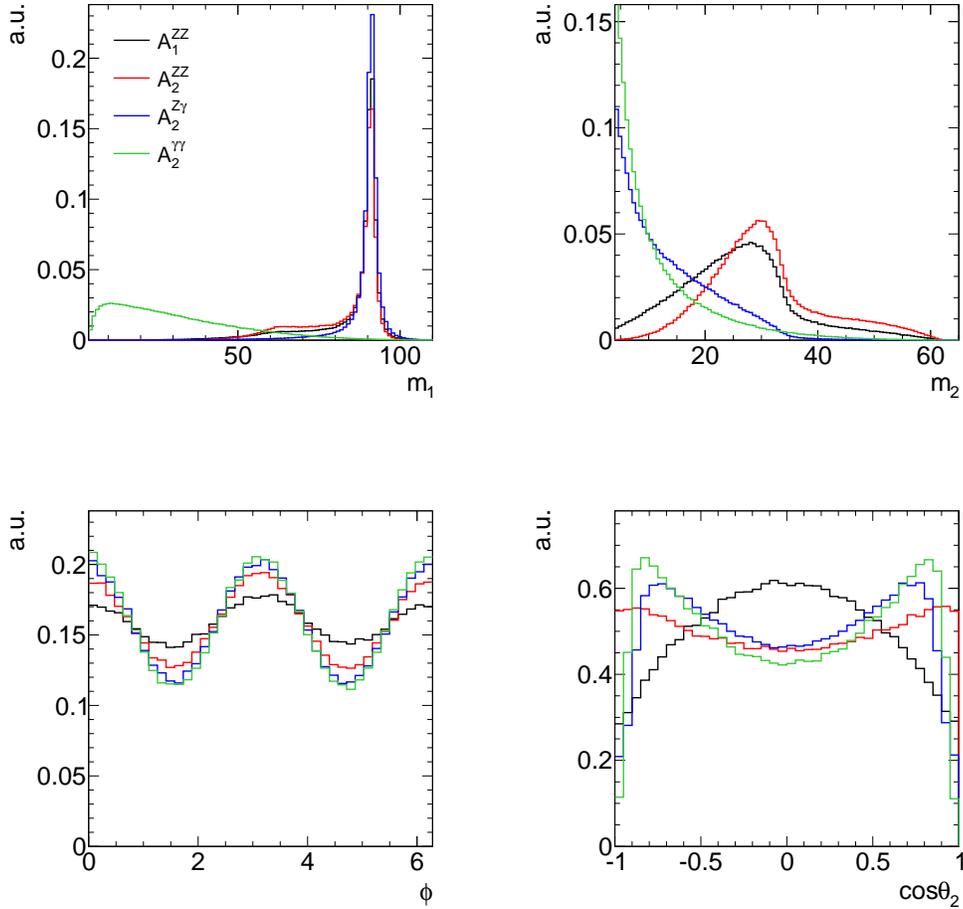


Figure 12.1: Examples of projections of differential cross sections for different pure  $CP$ -even models. The models shown include the standard model tree-level  $HZZ$  coupling, and the leading order loop-induced  $HVV'$  couplings. For each model  $m_1, m_2, \phi$  and  $\cos\theta_2$  are plotted, showing the distinctly different shapes predicted by the different models.

functions (PDF). In order to model these observables, we performed a fit to events generated using Monte Carlo generators.

In the signal case, we can safely assume factorization between the production side and the decay side, which is calculated analytically:

$$\begin{aligned} & \frac{d^{12}\sigma_S}{dm_{4\ell}^2 dp_{T,4\ell} dy_{4\ell} d\phi_{4\ell} d\phi_{\text{offset}} dm_1^2 dm_2^2 d\Phi_1 d\cos\Theta d\phi d\cos\theta_1 d\cos\theta_2} \\ = & \frac{d^5\sigma_S}{dm_{4\ell}^2 dp_{T,4\ell} dy_{4\ell} d\phi_{4\ell} d\phi_{\text{offset}}} \times \frac{d^7\sigma_S}{dm_1^2 dm_2^2 d\Phi_1 d\cos\Theta d\phi d\cos\theta_1 d\cos\theta_2}, \end{aligned} \quad (12.1)$$

where the production piece includes contributions from all possible initial states, including the small non- $gg$  contributions. We perform extraction from the NLO generators to account for some of the acceptance effects. The extracted spectrum for the signal is shown in Figure 12.2

Since the decay part of the background process depends on the exact initial state, we have to extract the production spectra for different initial states, and combine each of them with the appropriate decay differential cross section:

$$\frac{d^{12}\sigma_B}{d\vec{X}} = \sum_i \left( \frac{d^{12}\sigma_B^{q_i\bar{q}_i}}{d\vec{X}} + \frac{d^{12}\sigma_B^{\bar{q}_i q_i}}{d\vec{X}} + \frac{d^{12}\sigma_B^{gq_i}}{d\vec{X}} + \frac{d^{12}\sigma_B^{g\bar{q}_i}}{d\vec{X}} + \frac{d^{12}\sigma_B^{q_i g}}{d\vec{X}} + \frac{d^{12}\sigma_B^{\bar{q}_i g}}{d\vec{X}} \right), \quad (12.2)$$

while looping over the different quark types. Note that since the decay kinematics is not symmetric under a parity operation, the labels on the differential cross section are directional and the two letters are not interchangeable. Here the NLO contribution in the initial states are mostly accommodated, with an exception of a couple of diagrams which we expect to give a small contribution. Some examples of such diagrams are shown in Figure 12.3. Decomposition of the contributions from different initial states are shown in Figure 12.4.

In order to fully include NLO effects, a different, more elaborate analysis strategy needs to be developed, and we leave this to future studies.

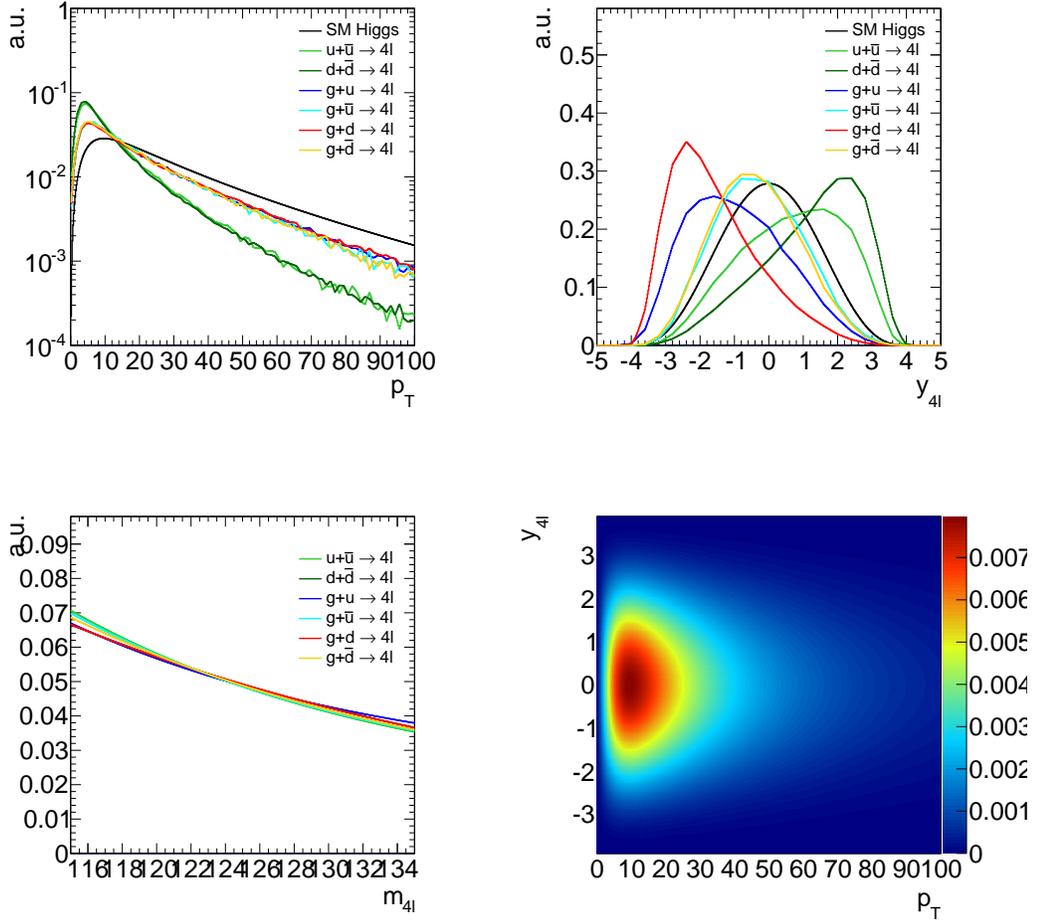


Figure 12.2: Extracted spectrum for the signal and the background for the  $2e2\mu$  final state. In the first two panels the distribution with  $m_{4\ell} = 125$  is plotted, comparing among the different components of the signal and background. In the rapidity distribution, the initial state is directional. The difference in the initial state causes a shift in the final state distributions. In the bottom left panel the  $m_{4\ell}$  distribution is shown for the background components. For a standard model Higgs signal, it is almost a delta function and is not shown here. In the last panel, the 2D distribution of  $p_T$  and  $y_{4l}$  for the standard model Higgs is shown.

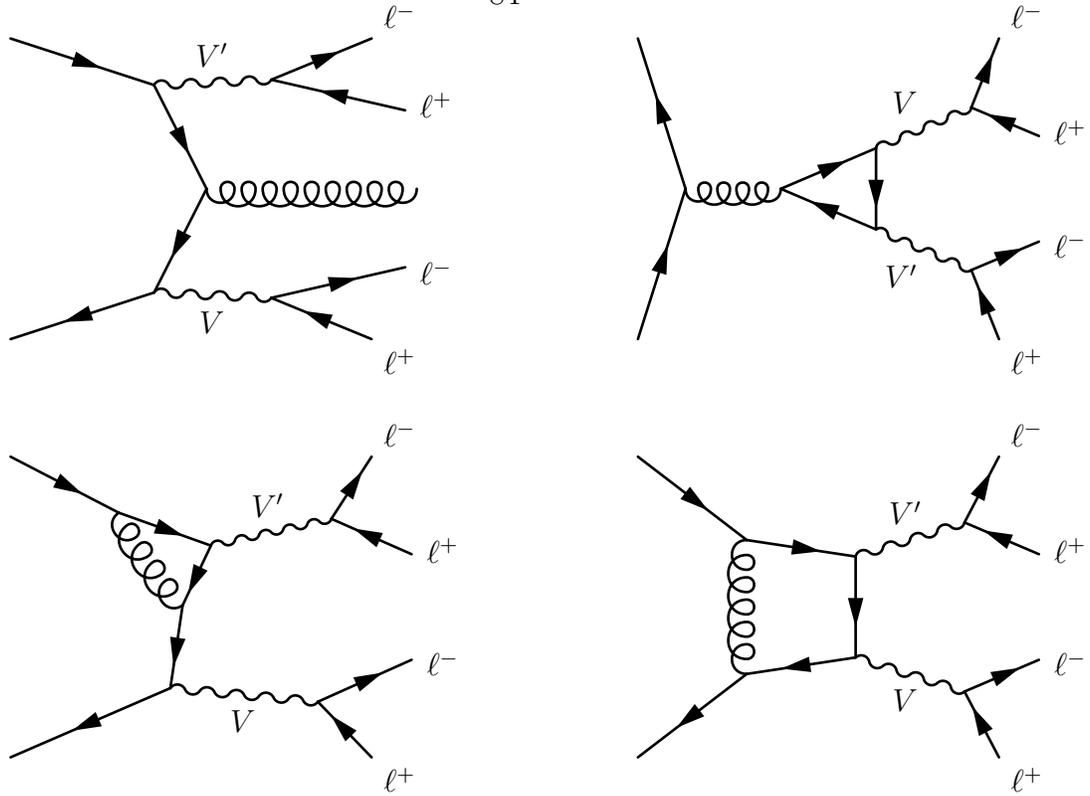


Figure 12.3: Examples of diagrams not accounted for in the calculation of SM backgrounds, at higher order in QCD. Among these, the upper left diagram (and diagrams related by crossing symmetry) does not contain a loop, and is therefore expected to contribute more than the loop diagrams.

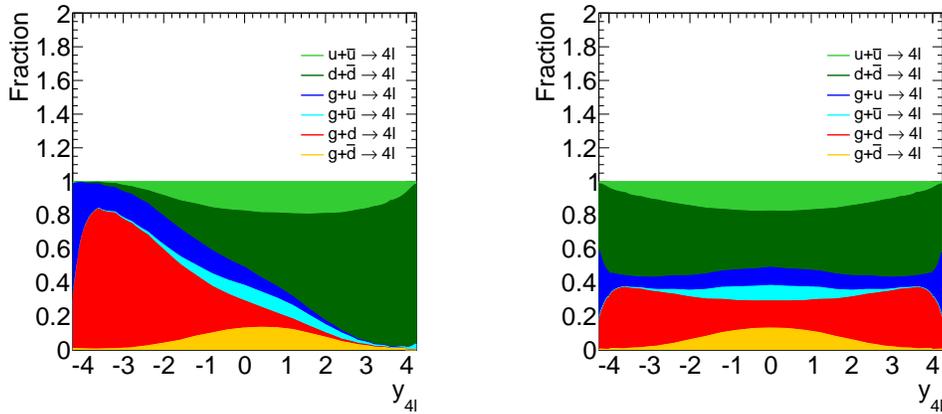


Figure 12.4: Decomposition of the different initial state contributions in  $q\bar{q} \rightarrow$  processes, as a function of the  $4\ell$  system rapidity. In the left plot, an immediate  $q$  is coming from the negative side, and a  $\bar{q}$  is coming from the right. It is symmetrized in the right plot, which shows a more realistic view of the contribution, as we do not know which proton gives rise to which parton in the experiment.

## 12.3 Integration Towards The 8D Likelihood

Having both the differential cross section and the transfer functions to describe the detector response to leptons, we can proceed with calculating the detector-level likelihood via a convolution integral [94], as a function of all parameters of interest  $\vec{A}$ <sup>1</sup>:

$$P(\vec{X}^R|\vec{A}) \propto F(\vec{X}^R|\vec{A}) = \int P(\vec{X}^G|\vec{A})T(\vec{X}^R|\vec{X}^G)d\vec{X}^G, \quad (12.3)$$

where  $\vec{X}^R$  denotes detector level quantities, and  $\vec{X}^G$  denotes generator-level quantities. Equation 12.3 states that the detector-level probability is proportional to the integral of all generator-level configurations each multiplied by its respective smearing function, to obtain the corresponding detector-level configuration. Since the transfer function  $T(\vec{X}^R|\vec{X}^G)$  (see Equation 11.2) in general also includes detector efficiencies, the integral  $F$  shown above is not properly normalized. The normalization procedure is described in the next section.

There are three levels for each event:

Generator level  $\rightarrow$  Detector level  $\rightarrow$  Interpreted level

The generator level stage refers to the ‘‘truth level’’ quantities before any detector effects. These quantities is then passed through the detector to get the detector level quantities. In this intermediate stage we assume that we know the correct lepton pairing, from the truth information. However since we only observe the four leptons in the detector, ‘‘assign’’ the leptons to pairs, and consider all possible pairs and the corresponding likelihoods.

The final quantities we reconstruct are at the interpreted level, where we choose a specific lepton pairing to be interpreted as the ‘first pair’ and ‘second pair’. It is a many-to-one mapping from detector level to interpreted level, where we consider all

---

<sup>1</sup>In this case,  $\vec{A}$  denotes  $A_{1,2,3}^{VV'}$ . Also note that this compact notation will allow us to add further parameters in future analyses

possible pairings when calculating the final likelihood for a certain event:

$$P(\vec{X}^I|\vec{A}) = \sum_i P(\vec{X}_i^R|\vec{A}), \quad (12.4)$$

where  $\vec{X}^I$  represents the interpreted quantities, obtained by summing over all possible detector-level configurations that would be interpreted into the same quantities. There are 2 possible pairings for each event with a  $2e2\mu$  final state, and 4 possible pairings for each  $4e$  and each  $4\mu$  final state events. The implementation of the convolution integral will be discussed in detail in Chapter 13.

It is necessary to normalize the *pdf* properly in order to treat it as a likelihood. The normalization can be calculated by integrating the  $F(\vec{X}^R|\vec{A})$  over the reconstructed level configurations  $\vec{X}^R$ :

$$I(\vec{A}) \equiv \int F(\vec{X}^R|\vec{A})d\vec{X}^R. \quad (12.5)$$

The evaluation of the convolution integral  $F(\vec{X}^R|\vec{A})$  is computationally intensive, and it is not possible to repeat it many different times to get a numerical average via Monte Carlo methods. It's also non-trivial to calculate it via numerical methods as there are 12 dimensions, and there is no simple way to cut down the dimensions, as will be shown in the next chapter for the convolution integral. We can, however, evaluate the equation differently and avoid the numerically intensive part:

$$\begin{aligned} I(\vec{A}) &= \int F(\vec{X}^R|\vec{A})d\vec{X}^R \\ &= \int \left( \int P(\vec{X}^G|\vec{A})T(\vec{X}^R|\vec{X}^G)d\vec{X}^G \right) d\vec{X}^R \\ &= \int P(\vec{X}^G|\vec{A}) \left( \int T(\vec{X}^R|\vec{X}^G)d\vec{X}^R \right) d\vec{X}^G \\ &\equiv \int P(\vec{X}^G|\vec{A})\bar{\epsilon}(\vec{X}^G)d\vec{X}^G, \end{aligned} \quad (12.6)$$

where in the last step we define the ‘‘average efficiency’’ given a certain generator level configuration as  $\bar{\epsilon}(\vec{X}^G)$ . The average efficiency gives the probability that a certain generator-level event will survive all the analysis cuts.

Equation 12.6 follows from the observation that we are integrating over both the generator level quantities and the reconstructed level quantities. In the first line of the definition of the integral, the integration is first done over generator level quantities inside  $F(\vec{X}^R|\vec{A})$ , and then done over reconstructed level quantities  $\vec{X}^R$ . For the correct evaluation of  $F(\vec{X}^R|\vec{A})$ , a careful calculation of the integration trajectory in  $\vec{X}^G$  is needed, which increases the computing time considerably. However, since the integration is done over the whole phase space, it doesn't matter which integration we do first. If we do the integration over  $\vec{X}^R$  first, there is no need to calculate any integration trajectory. It is then straightforward to do the integral, and we avoid the computationally difficult parts. A simple Monte Carlo algorithm works well in this case.

In Figure 12.5 we show an example of the convergence of the normalization calculation as a function of sample count and CPU time. A large Monte Carlo sample is generated with a uniform distribution in  $\vec{X}^G$ , and the events are then grouped into pseudodatasets of a given sample size. By calculating the normalization from each pseudodataset, and evaluating the RMS of the results relative to the mean value, we can get an estimate of the precision for a given sample size. This is then repeated many times for different sample sizes, and the result is plotted. The precision of the normalization follows a simple  $N^{-1/2}$  dependence. When the sample size is large there is some spread due to insufficient statistics, but  $N^{-1/2}$  fits well. We find that we can reach precision of 0.1% in a few CPU-hours of run time for all components.

## 12.4 Signal Likelihood for an Arbitrary Model

Due to the simple structure of the differential equations, it is possible to pre-calculate different pieces of the detector level likelihood separately, and later combine the different pieces to get the final detector-level likelihood. If the generator level differential cross section can be written in this form:

$$P(\vec{X}^G|\vec{A}) = \sum_i f_i^X(\vec{X}^G) f_i^A(\vec{A}), \quad (12.7)$$

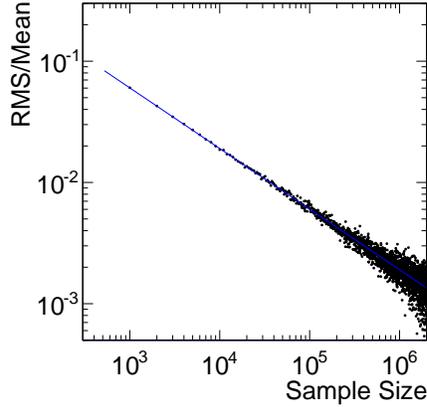


Figure 12.5: Estimation of the precision of the calculation of the convolution integral normalization (see Equation 12.6). Since a Monte Carlo integration method is employed, we can estimate the precision by grouping individual samples into sub-samples of different sample size, and use the RMS as a measure of the precision at any given sample size. This is repeated many times and a  $1/\sqrt{N}$  curve is fitted to the result. It allows us to estimate the precision for a large sample size, where there are insufficient statistics to evaluate the RMS spread with confidence. The lack of statistics is more pronounced at larger sample sizes, where we see a larger spread in the RMS values.

where the index runs over different terms in the expression, the convolution integral can be done on pieces that do not involve  $\vec{A}$ , and are later combined together:

$$\begin{aligned}
 F(\vec{X}^R|\vec{A}) &= \int P(\vec{X}^G|\vec{A})T(\vec{X}^R|\vec{X}^G)d\vec{X}^G \\
 &= \int \left( \sum_i f_i^X(\vec{X}^G)f_i^A(\vec{A}) \right) T(\vec{X}^R|\vec{X}^G)d\vec{X}^G \\
 &= \sum_i \left( \int f_i^X(\vec{X}^G)T(\vec{X}^R|\vec{X}^G)d\vec{X}^G \right) f_i^A(\vec{A}). \quad (12.8)
 \end{aligned}$$

Since  $f_i^A(\vec{A})$  does not depend on any of the observables  $\vec{X}^G$ , it can be taken out of the integral. What's left inside the integral does not depend on the coupling parameters  $\vec{A}$ , which allows us to calculate it only once per event. The integration in Equation 12.8 thus does not need to be repeated for the evaluation of likelihood at different points of  $\vec{A}$ .

In the case of the current analysis, where we have calculated up to leading order

in these couplings, the differential cross section is a second-order polynomial. The same structure applies to the convolution integral used to obtain the normalization, which is also a polynomial of the parameters we want to measure:

$$\begin{aligned}
 I(\vec{A}) &= \int P(\vec{X}^G|\vec{A})\bar{\epsilon}(\vec{X}^G)d\vec{X}^G \\
 &= \int \left( \sum_i f_i^X(\vec{X}^G)f_i^A(\vec{A}) \right) \bar{\epsilon}(\vec{X}^G)d\vec{X}^G \\
 &= \sum_i \left( \int f_i^X(\vec{X}^G)\bar{\epsilon}(\vec{X}^G)d\vec{X}^G \right) f_i^A(\vec{A}). \tag{12.9}
 \end{aligned}$$

Once we have the coefficients (shown inside the parentheses in the last line of Equation 12.8 and Equation 12.9) pre-calculated, it's straightforward and fast to find the final likelihood, for arbitrary values of the parameters of interest.

## 12.5 Templated Background Components

The convolution integral construction includes the components where we have an analytic form for the generator level differential cross section, and a good description of the transfer function, as is the case for the signal ( $P_S(\vec{X}|\vec{A})$ ) and the  $q\bar{q} \rightarrow 4\ell$  background ( $P_{q\bar{q}}(\vec{X})$ ). These conditions do not apply to all the components. One such example is the reducible background from jets faking leptons in Drell-Yan events ( $P_Z(\vec{X})$ ). Since the statistics is small for LHC Run I, it is useful to choose cuts that maximize the signal efficiency, at the expense of letting more background into our final event selection. In the future, when statistics will not be a problem, the event selection strategy could be changed so that the signal purity is emphasized, where we tighten the selection cuts to reduce the reducible background as much as possible. With the current selection, the reducible background contributes 15% of all events in the window of 115 GeV to 135 GeV.

In order to model this background properly, we construct a large 8D template by weighting the generator-level Drell-Yan events by their respective lepton fake rates. The rate of jets faking leptons is extracted from Drell-Yan events simulated in CMS,

as a function of the lepton momentum and direction. The extracted fake rate is shown in Figure 12.6. It is validated by comparing the distributions from CMS simulated samples to those from reweighted generator-level events.

A large sample of PYTHIA [73] (version 6.426) Drell-Yan events are generated to construct the template. To account for possible differences of the kinematics between PYTHIA events and that from other NLO generators, a simple reweighting is done on the PYTHIA sample to match the dilepton mass distributions with POWHEG [95–99] samples. The dimensions of the template are chosen to maximize the number of bins while maintaining reasonable statistics. We choose to have 12 bins in the  $m_{4\ell}$  dimension, and 10 bins in each of the other decay observables.

The binning effect is reduced by performing a linear interpolation of the bins in all dimensions. Each bin in the template is represented by a point at the bin center, with the bin content as the value. In order to do the interpolation we use all neighboring bins: for an 8D template there are  $2^8$  neighboring bins for any given point. The interpreted value is a weighted sum of the values for all the neighboring points. The weight assigned to each neighboring point has a simple geometrical interpretation, as illustrated in Figure 12.7. The weight is the ratio between the volume of the away-side hypercube and the volume of the total hypercube.

We performed a closure test in which we assess the precision of this procedure by following the same procedure on the  $q\bar{q} \rightarrow 4\ell$  background components, and use the calculated likelihood from the convolution integral as a benchmark. The precision on the convolution integral is much better than this procedure, so the spread on the ratio of likelihood values obtained from the different procedures can be assigned to the big template procedure. We see O(20%) precision on this component. A systematic uncertainty is needed to cover possible imperfect modeling from this procedure.

The  $gg \rightarrow 4\ell$  irreducible background component,  $P_{gg}(\vec{X})$ , is treated in the same way. As it consists of only  $< 1\%$  of all events, the effect from imprecise modeling of this background component is negligible.

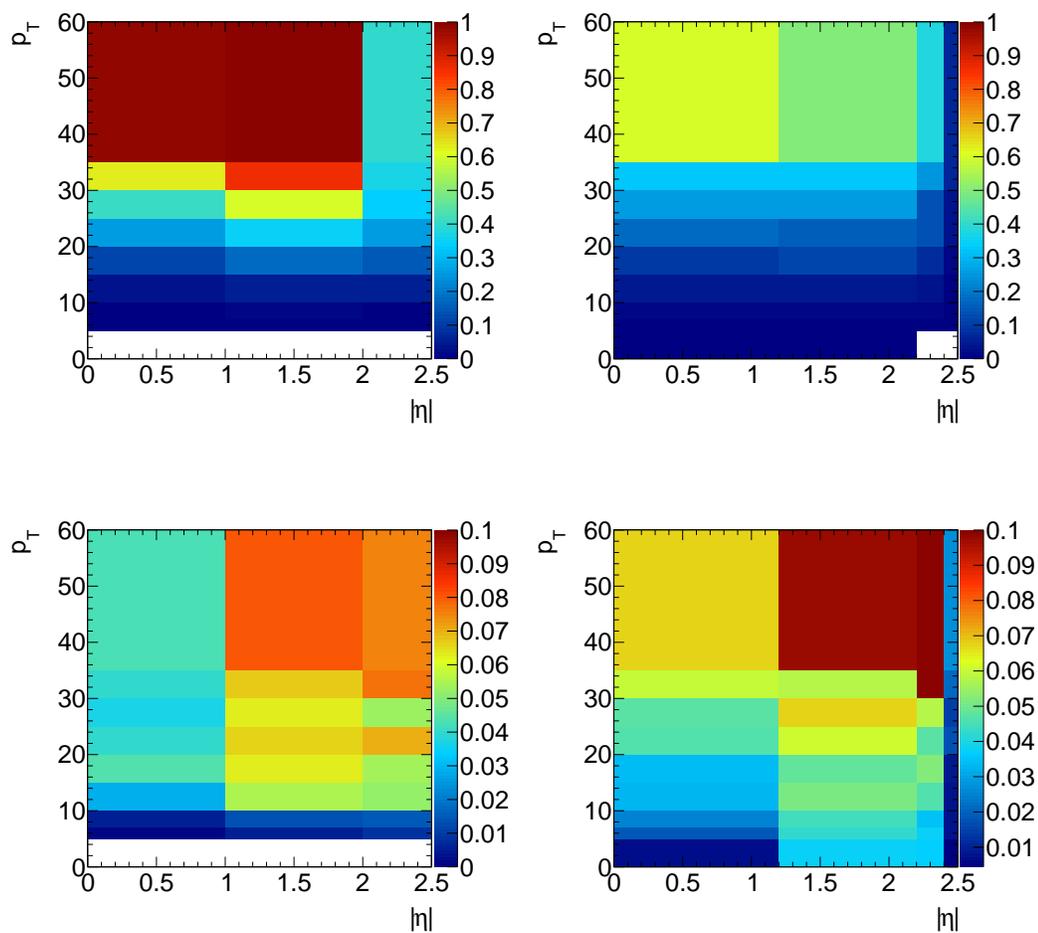


Figure 12.6: Fake rate for jets faking leptons as a function of  $p_T$  and  $\eta$ . The top row shows the fake rates for 7 TeV, and the bottom row shows the rates for 8 TeV data. The left-hand plots are the rates of jets faking an electron, and the right-hand plots are the corresponding rates for muons. The binning in the plots reflects the actual binning used in the analysis.

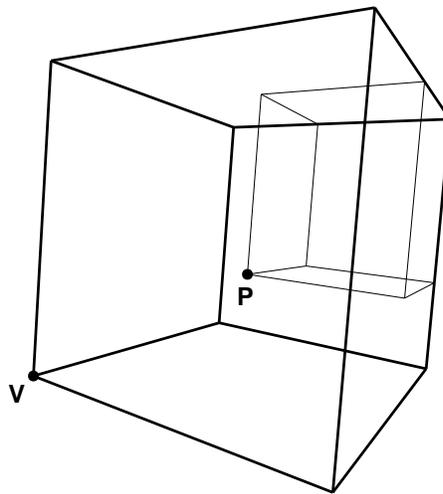


Figure 12.7: Geometrical visualization of linear interpolation in multiple dimensions. Suppose there is a value attached to each of the 8 vertices, and we would like to linearly interpolate the values for a point  $P$  in the center. The interpolated value is a weighted sum of all the vertices. The weight of each vertex can be calculated as the ratio between the volume of the smaller cube away from vertex, and the total volume.

## 12.6 Systematics Uncertainties

We consider a list of systematic uncertainties in the analysis framework: (1) lepton scale and resolution uncertainties, (2) scale and centrality of the production spectrum, (3) systematic uncertainties on the templated background components, and (4) assumptions on the Higgs mass  $m_H$ .

The systematic uncertainties are treated with different versions of the *pdfs* and with nuisance parameters. For each systematic we consider, different versions of the *pdfs* are constructed to represent potential variations. For example for the lepton scale uncertainty, in addition to the central *pdf*  $P_0(\vec{X}^R|\vec{A})$ , two additional *pdfs*  $P_{\pm}(\vec{X}^R|\vec{A})$  are constructed by using the transfer functions with different scales:

$$\tilde{T}_{\pm}(c_i|\vec{p}_i^G) \equiv \tilde{T}(c_i \pm \sigma(\vec{p}_i^G)|\vec{p}_i^G), \quad (12.10)$$

where the  $\sigma$  is the uncertainty on the lepton momentum scale. A nuisance parameter  $n$  is introduced to represent the systematic variation, and the final *pdf* is linearly interpolated between the different versions of the *pdfs*, namely:

$$P(\vec{X}^R|\vec{A}) = \begin{cases} nP_+(\vec{X}^R|\vec{A}) + (1-n)P_0(\vec{X}^R|\vec{A}), & \text{if } n \geq 0 \\ -nP_-(\vec{X}^R|\vec{A}) + (1+n)P_0(\vec{X}^R|\vec{A}) & \text{otherwise.} \end{cases} \quad (12.11)$$

When  $n = 0$  this reproduces  $P_0(\vec{X}^R|\vec{A})$  and when  $n = \pm 1$  the final *pdf* is the same as  $P_{\pm}(\vec{X}^R|\vec{A})$ . The nuisance parameter  $n$  is then floated in the fit, enlarging the uncertainties on the fitted parameters.

For a systematic uncertainty with large variations, we put many “checkpoints” along the axis of nuisance parameter, and piece-wise interpolate linearly between the different checkpoints, as illustrated in Figure 12.8 taking as an example the lepton scale uncertainty. One other example of this is the assumption of Higgs mass. While we use the CMS measured value as the central value in the *pdf*, potential variations must be accounted for by including different “checkpoints”, with different Higgs masses.

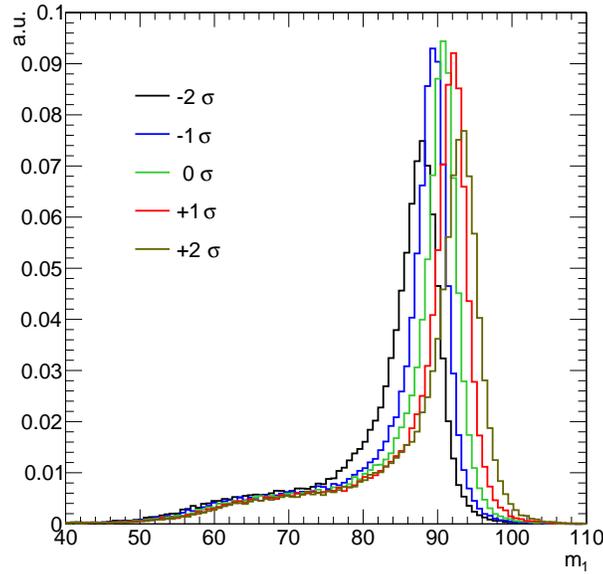


Figure 12.8: Example projection of different check points

## 12.7 Total Dataset Likelihood

Now that we have all the ingredients needed to obtain the final likelihood value for every event, we can combine the likelihoods as follows:

$$\begin{aligned}
 P_i(\vec{A}) &= (1 - f_{q\bar{q}} - f_Z - f_{gg})P_S(\vec{X}_i^R|\vec{A}) + f_{q\bar{q}}P_{q\bar{q}}(\vec{X}_i^R) + f_ZP_Z(\vec{X}_i^R) + f_{gg}P_{gg}(\vec{X}_i^R) \\
 P(\vec{A}) &= \prod_j \text{Poisson}(N_{obs}^j | \mu N_S^j + N_{q\bar{q}}^j + N_Z^j + N_{gg}^j) \prod_i P_i(\vec{A}) \\
 &\quad \prod_{j'} L(N_S^{j'} | N_{S,exp}^{j'}, \sigma_{S,exp}^{j'}) L(N_{q\bar{q}}^{j'} | N_{q\bar{q},exp}^{j'}, \sigma_{q\bar{q},exp}^{j'}) \\
 &\quad L(N_Z^{j'} | N_{Z,exp}^{j'}, \sigma_{Z,exp}^{j'}) L(N_{gg}^{j'} | N_{gg,exp}^{j'}, \sigma_{gg,exp}^{j'}), \tag{12.12}
 \end{aligned}$$

where  $i$  loops over the observed events, and  $j$  ( $j'$ ) loops over the final states and the CM energy. The fractions of different background components are represented by  $f_{q\bar{q}}^j$ ,  $f_Z^j$  and  $f_{gg}^j$ . Prior knowledge of the expected yield is represented by the  $L(N_{fit}^j | N_{exp}^j, \sigma_{exp}^j)$  terms, where for each such term we use a log-normal distribution to implement this constraint.

The expected yield in the  $4\ell$  final state between  $m_H = 115$  and  $135$  GeV for CMS

		7 TeV		8 TeV	
		$N$	$\sigma$	$N$	$\sigma$
$2e2\mu$	$gg \rightarrow H \rightarrow 4\ell$	0.95	4.7%	4.36	4.8%
	$q\bar{q} \rightarrow 4\ell$	0.87	25%	3.52	25%
	$gg \rightarrow 4\ell$ continuum	0.027	40%	0.19	40%
	Reducible $Z + X$ background	0.11	11%	0.72	11%
$2\mu 2e$	$gg \rightarrow H \rightarrow 4\ell$	0.74	4.7%	3.38	4.8%
	$q\bar{q} \rightarrow 4\ell$	0.55	25%	2.02	25%
	$gg \rightarrow 4\ell$ continuum	0.020	40%	0.15	40%
	Reducible $Z + X$ background	0.58	11%	2.08	11%
$4e$	$gg \rightarrow H \rightarrow 4\ell$	0.70	4.7%	3.07	4.7%
	$q\bar{q} \rightarrow 4\ell$	0.52	25%	1.79	25%
	$gg \rightarrow 4\ell$ continuum	0.021	40%	0.12	40%
	Reducible $Z + X$ background	0.41	11%	1.69	11%
$4\mu$	$gg \rightarrow H \rightarrow 4\ell$	1.27	4.7%	5.98	4.7%
	$q\bar{q} \rightarrow 4\ell$	1.07	25%	4.64	25%
	$gg \rightarrow 4\ell$ continuum	0.037	40%	0.25	40%
	Reducible $Z + X$ background	0.17	11%	0.82	11%

Table 12.1: Expected yields and the corresponding percentage uncertainties, for each contribution to the Higgs boson signal and to the backgrounds.

Run I running is summarized in Table 12.1.

## 12.8 Measurement and Inference

We proceed to perform a fit and extract the parameters and their errors with the constructed likelihood function.

When interpreting the results there are two complementary fundamental questions: (1) if the data follows standard model, how likely is it that it fluctuates to different fit central values? (2) once we have a set of data observed, what is the possibility that each exotic model is true? These are distinct questions and should not be mixed.

To answer the first question, we generate an ensemble of pseudo-experiments, each of similar statistics to the (expected) data. By extracting the fit results repeatedly, one can obtain a distribution of the central fit values. Statistical statements can then be made based on the distribution of fit results.

For the second question, one widely-adopted way of representing this probability of a certain exotic model fluctuating into a standard model-like dataset is to use the likelihood differences. We scan over different model spaces, and for each model we calculate this likelihood difference. This likelihood scan is done on the observed dataset, or on a special dataset that corresponds to the expected value for each model.

One can then construct an “average” dataset (or conventionally called an “Asimov dataset” [100]). This dataset is defined to represent the “average” expectation value. It is constructed by generating 1000 times the nominal statistics, and by reweighting the events so that the total yield matches what we expect. In this way the possible fluctuations are averaged out.

While this is a useful representation of the expectation, the result from this special dataset could be misleading, since it represents the central value of the expectation and doesn’t say much about potential fluctuations. In order to evaluate the effect of potential fluctuations, an ensemble of pseudoexperiments is again used. We repeat the likelihood scan on each of the pseudoexperiments, and overlay all the curves on top of each other. The observed curve is then compared with this ensemble of curves, and we make statistically meaningful statements based on this comparison.

With the statistical interpretation in place, we have described the outline of the essential parts of this novel analysis method, and a new multidimensional, computationally tractable framework for Higgs property measurements. Through the use of the new framework we are able to fully utilize the power of the golden channel  $H \rightarrow 4\ell$  by constructing continuous likelihood functions in all observables as a function of the parameters of interest, for the Higgs signal and for the major background components.

This analysis method is distinctively different compared to other established methods. Methods which utilize templates of discriminants offer many advantages such as simplicity in carrying out the analysis, but they are more dependent on the availability of Monte Carlo samples and the choice of discriminant. This method on the other hand requires (arguably) more computing resources to do the analysis, but offers great flexibility and a more complete picture of the extracted result. We describe

more detail on the implementation of this new technique in the next chapter.

## Chapter 13

# The Convolution Integral: Technical Description

### 13.1 General Strategy

Before describing the details of the implementation of the convolution integral, we first outline the three relevant bases. The generator level differential cross section is defined in the 8 decay observables and the 4 production observables as presented in Section 10.1. These 12 observables are collectively denoted as  $\vec{X}$ . While this basis is convenient for the theoretical calculation, detector effects are not modeled as a function of  $\vec{X}$ . A more natural basis for the integration of the lepton momentum resolution effect is the basis of lepton momenta, denoted as  $\vec{P}$  collectively. The 12 degrees of freedom can then be parametrized by the parallel components  $\vec{P}^{\parallel}$  and the perpendicular components  $\vec{P}^{\perp}$  of the lepton momenta. Thanks to the excellent directional resolution of leptons in CMS and ATLAS, considering only the detector resolution on the parallel component is a good approximation. The  $\vec{P}$  basis is useful to model detector effects, but it is not easy to use to perform the integration. A final change of variable is thus performed, from the parallel components to smearing factors and dilepton masses (together with  $m_{4\ell}$  in the case of the signal).

As described above, the first basis is the basis in which our differential cross section is defined, named as the “observable basis”. It consists of the 8 decay observables (3 masses and 5 angles) where an analytic expression is available, together with

other production observables which we extract from generators, denoted as  $\vec{X}$ . The convolution integral can then be written symbolically as:

$$P(\vec{X}^R|\vec{A}) \propto F(\vec{X}^R|\vec{A}) = \int P(\vec{X}^G|\vec{A})T(\vec{X}^R|\vec{X}^G)d\vec{X}^G, \quad (13.1)$$

where the superscripts  $G$  and  $R$  represent generator level and reconstructed level quantities respectively. The generator level differential cross section  $P(\vec{X}^G|\vec{A})$  is a function of the parameters of interest ( $\vec{A}$ ). It is convoluted with the transfer function  $T(\vec{X}^R|\vec{X}^G)$  which encodes detector effects including momentum resolution, efficiency and acceptance. As mentioned in the previous chapter, the result of the convolution integral  $F(\vec{X}^R|\vec{A})$  is in general not normalized, since the transfer function includes efficiency effects. An additional normalization is thus needed to arrive at the *pdf*,  $P(\vec{X}^R|\vec{A})$ .

The second basis, the ‘‘lepton momentum basis’’, is defined based on lepton momenta. For each lepton there are three axes defined: one parallel  $p_i^\parallel$  to the lepton momentum and two perpendicular  $\vec{p}_i^\perp$  to it. A total of 12 degrees of freedom are covered with the three axes for each of the four leptons. The four axes parallel to each lepton can also be defined in terms of scaling factors  $c_i = p_i^R/p_i^G$  for the sake of convenience in integration, while the directions perpendicular to the lepton directions are in terms of momentum components (in units of GeV/c). The leptons are labelled by index 1 through 4, where lepton 1 and lepton 2 are paired together as the first pair, and the other two as the second pair. Leptons with odd indices are negatively charged and even indices are positively-charged. After the change of variables from the observable basis to the lepton momentum basis, the expression becomes:

$$\begin{aligned} F(\vec{X}^R|\vec{A}) &= \int P(\vec{X}^G|\vec{A})T(\vec{c}|\vec{P}^G)\frac{|J^G|}{|J^R|}d\vec{P}^\perp d\vec{c} \\ &= \int P(\vec{X}^G|\vec{A})T(\vec{c}|\vec{P}^G)\frac{|J^G|}{|J^R|}d\vec{c}, \end{aligned} \quad (13.2)$$

where two Jacobian factors emerge, one ( $J^G$ ) from the change of variable in the volume element, and one ( $J^R$ ) from the change of variable in the transfer function. Since

directional resolution is assumed to be perfect, integrations along the perpendicular momenta  $\vec{P}^\perp$  can be trivially done with the delta functions in the transfer function in the second line of Equation 13.2. The four smearing factors of the leptons are collectively written as  $\vec{c}$ .

The third and final basis (“integration basis”) is defined by doing a change of parameters from the 4 scaling factors  $c_i$  into the dilepton masses  $m_1$  and  $m_2$ , and in case of the signal, the 4-lepton mass, while leaving the components perpendicular to the lepton directions unchanged. Specifically for the background we change into the basis of  $(m_1^2, m_2^2, c_1, c_3)$ :

$$\begin{aligned}
F_B(\vec{X}^R) &= \int P_B(\vec{X}^G) T(\vec{c} | \vec{P}^G) \frac{|J^G|}{|J_R|} d\vec{c} \\
&= \iiint P_B(\vec{X}^G) T(\vec{c} | \vec{P}^G) \frac{|J^G|}{|J_R|} |J_B| dc_1 dc_3 dm_1^2 dm_2^2 \\
&= \frac{1}{|J_R|} \iiint P_B(\vec{X}^G) T(\vec{c} | \vec{P}^G) |J^G| |J_B| dc_1 dc_3 dm_1^2 dm_2^2, \tag{13.3}
\end{aligned}$$

while for the signal the basis is changed to  $(m_1^2, m_2^2, c_1, m_{4\ell}^2)$  in order to integrate out the delta function on  $m_{4\ell}^2$  given the narrow width approximation:

$$\begin{aligned}
F_S(\vec{X}^R | \vec{A}) &= \int P_S(\vec{X}^G | \vec{A}) T(\vec{c} | \vec{P}^G) \frac{|J^G|}{|J_R|} d\vec{c} \\
&= \iiint P_S(\vec{X}^G | \vec{A}) T(\vec{c} | \vec{P}^G) \frac{|J^G|}{|J_R|} |J_S| dc_1 dm_1^2 dm_2^2 dm_{4\ell}^2 \\
&= \iiint P_S(\vec{X}^G | \vec{A}) |_{m_{4\ell}=m_H} T(\vec{c} | \vec{P}^G) \frac{|J^G|}{|J_R|} |J_S| dc_1 dm_1^2 dm_2^2 \\
&= \frac{1}{|J_R|} \iiint P_S(\vec{X}^G | \vec{A}) |_{m_{4\ell}=m_H} T(\vec{c} | \vec{P}^G) |J^G| |J_S| dc_1 dm_1^2 dm_2^2. \tag{13.4}
\end{aligned}$$

It is now computationally feasible to perform the integral in this form.

## 13.2 Jacobian Factors

### 13.2.1 From the Observable Basis to the Lepton Momentum Basis

Many Jacobian factors arise during the process of changing bases in order to reconcile the observable basis where generator level differential cross section is defined, to the lepton momentum basis where the detector smearing is defined, and to the integration basis where the integration is carried out. There is a 12x12 Jacobian factor in the change of variables from the observable basis to the lepton momentum basis. The transformation is non-trivial to write down analytically, and we take a numerical approach to calculate the Jacobian factor.

Jacobian factors have a simple geometrical interpretation, as illustrated in Figure 13.1. It is the ratio of some small infinitesimal volume elements before and after change of variable. In this case we construct a mini 12D cube, transform it to the new basis, and calculate the volumes of the parallelepiped in 12 dimensions. There are many readily available algorithms to calculate volume in hyperspace. In the current implementation, we take a simple approach and sequentially subtract the perpendicular components of previously examined edges from the current edge, to construct an equivalent hypercube with the same volume.

The calculation of this Jacobian factor can be validated with pseudoexperiment distributions. For any given test distribution  $f(\vec{X})$ , we can generate events in two different bases, and compare the distribution with one weighted by the Jacobian factor:

$$\int f(\vec{X})d\vec{X} = \int f(\vec{X})|J|d\vec{P}^\perp d\vec{P}^\parallel, \quad (13.5)$$

where  $\vec{P}^\perp$  and  $\vec{P}^\parallel$  denote parallel and perpendicular components of the lepton vectors. Notice that the lepton momentum basis is dependent on  $\vec{X}$ . Different observable points  $\vec{X}$  and  $\vec{X}'$  have different lepton orientations, and the lepton momentum axes  $\vec{P}$  and  $\vec{P}'$  are different. Therefore we can only check the validity of the Jacobian on a slightly different basis where the parallel components are also expressed in units

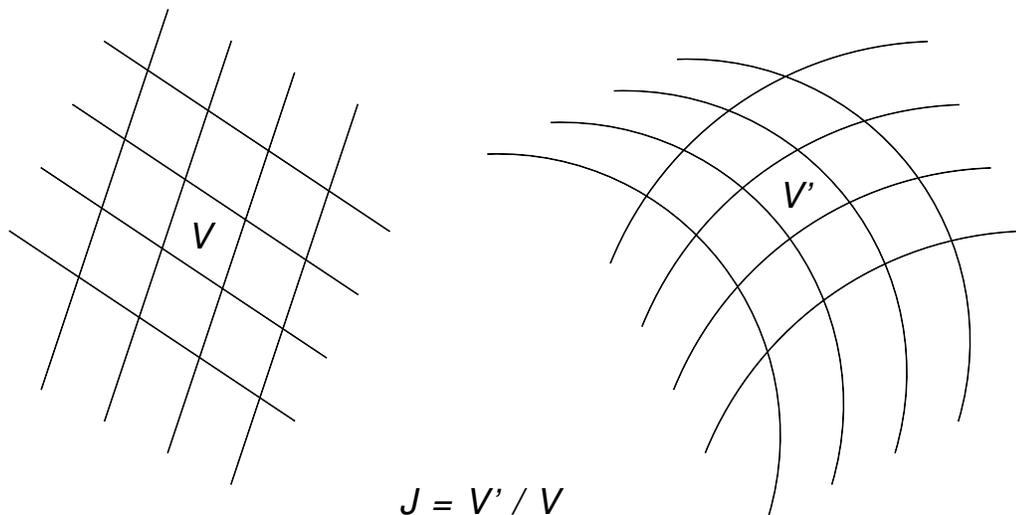


Figure 13.1: The Jacobian factor can be thought as the ratio of an infinitesimal volume around a given point of interest between two different bases. Lines on the left correspond to lines on the right in a different basis. The volume  $V$  is translated into  $V'$  on the right. The Jacobian factor at that point is therefore  $J = V'/V$ .

of GeV, and verify the change of variable to scaling factors later. In this slightly modified basis, even though the axes  $\vec{P}$  and  $\vec{P}'$  are different, they relate to each other by a rotation in space. The volume element all have the same size under a rotation, and we are able to perform this validation by choosing a particular lepton momentum axes as the representative basis.

The result of the 12D Jacobian validation is shown in Figure 13.2. It is done on a flat toy matrix element with  $100^2 < m_{4\ell}^2 < 140^2$ ,  $4^2 < m_1^2, m_2^2 < 100^2$ ,  $|\vec{p}_{T,4\ell}| < 100$ ,  $-4 < y_{4\ell} < 4$  and  $|p_\ell| < 100$ . We see excellent agreement between the events generated in the two different bases, indicating that the calculation is performed correctly.

The change of variable from parallel components  $\vec{P}^\parallel$  to scaling factors  $\vec{c}$  is diagonal in the lepton momentum basis. It can therefore be calculated trivially to be either  $|c_i/P_i^\parallel|$ , or  $|1/P_i^\parallel|$ , depending on whether we are changing the variable in the volume element or the transfer function.

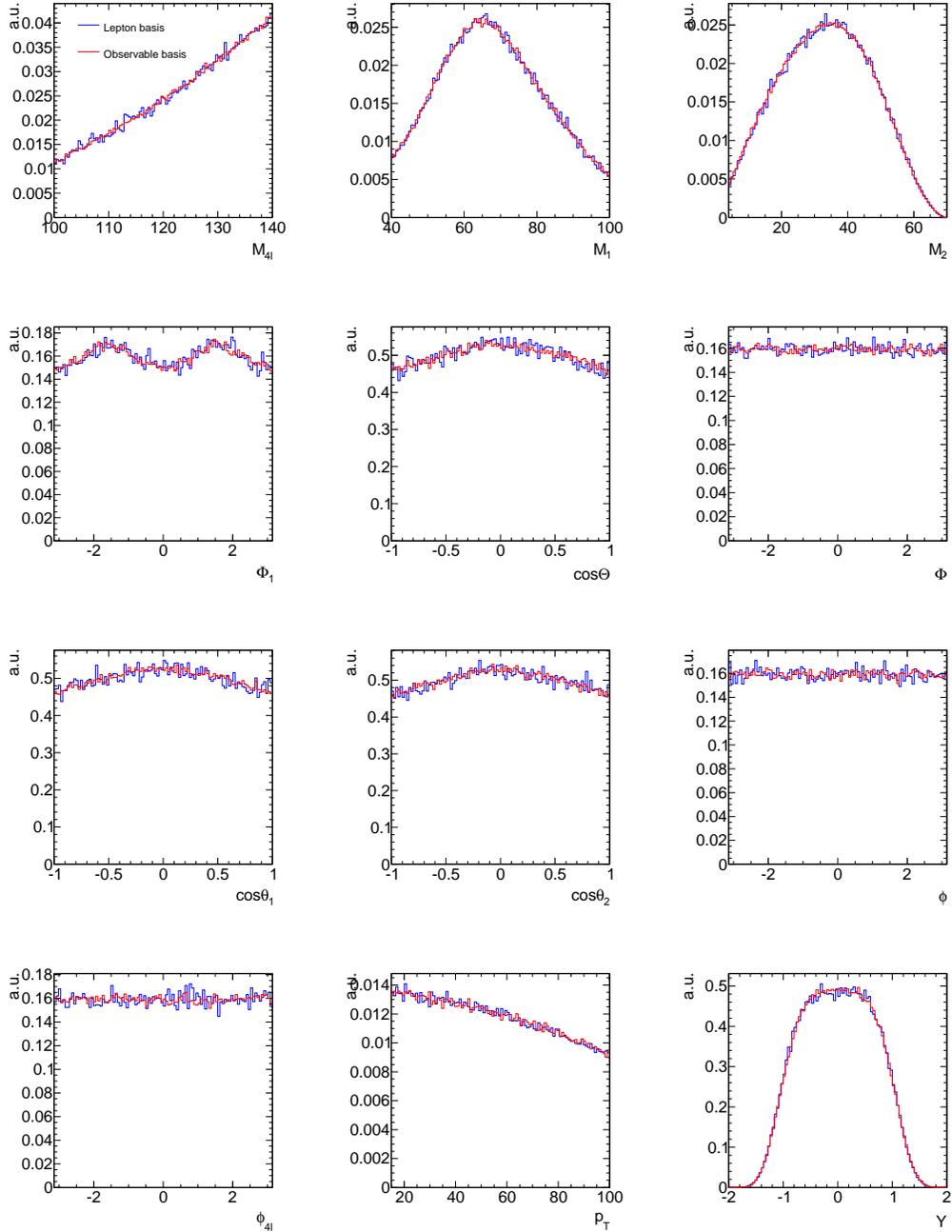


Figure 13.2: Validation of the 12D Jacobian calculation. Events are generated uniformly in both bases, and compared to each other by weighting those from the lepton basis by their respective 12x12 Jacobian factors. The small differences are due to fluctuations. We see excellent agreement for all the variables of interest.

### 13.2.2 The Background and Signal Factors

In this section we describe the calculation of the  $|J_B|$  Jacobian factor in Equation 13.3. It arises from the change of variables from  $(c_1, c_2, c_3, c_4)$  to  $(c_1, m_1^{G^2}, c_3, m_2^{G^2})$ . Assuming that the leptons are massless, and observing that

$$\begin{aligned} m_1^{R^2} &= (p_1^R + p_2^R)^2 = 2p_1^R p_2^R = 2c_1 c_2 p_1^G p_2^G = c_1 c_2 m_1^{G^2} \\ m_2^{R^2} &= (p_3^R + p_4^R)^2 = 2p_3^R p_4^R = 2c_3 c_4 p_3^G p_4^G = c_3 c_4 m_2^{G^2}, \end{aligned} \quad (13.6)$$

we can write down the value of the smearing factors  $c_2$  and  $c_4$  as

$$\begin{aligned} c_2 &= \frac{1}{c_1} \frac{m_1^{R^2}}{m_1^{G^2}} \\ c_4 &= \frac{1}{c_3} \frac{m_2^{R^2}}{m_2^{G^2}}, \end{aligned} \quad (13.7)$$

and therefore

$$\begin{aligned} dc_2 &= -\frac{1}{c_1} \frac{m_1^{R^2}}{m_1^{G^4}} dm_1^{G^2} \\ dc_4 &= -\frac{1}{c_3} \frac{m_2^{R^2}}{m_2^{G^4}} dm_2^{G^2}. \end{aligned} \quad (13.8)$$

Finally, the Jacobian can be written as

$$|J_B| = \frac{1}{c_1} \frac{m_1^{R^2}}{m_1^{G^4}} \frac{1}{c_3} \frac{m_2^{R^2}}{m_2^{G^4}}. \quad (13.9)$$

In the Higgs signal case, due to the narrow width of the Higgs boson (6 MeV) a delta function  $\delta(m_{4\ell}^2 - m_H^2)$  is used as an approximation. This introduces an additional constraint on the generator level  $m_{4\ell}$ , which induces a correlation between  $c_1$  and  $c_3$ . In order to integrate out the delta function on  $m_{4\ell}^{G^2}$ , we choose the final basis to be  $(c_1, m_1^{G^2}, m_{4\ell}^{G^2}, m_2^{G^2})$ . Similar to the dilepton mass case, assuming the leptons are massless, for  $m_{4\ell}$  we can write down an additional simple relation between the

generator level and reconstructed level quantities:

$$\begin{aligned}
m_{4\ell}^{G^2} &= (p_1^G + p_2^G + p_3^G + p_4^G)^2 = \sum_{i>j} 2p_i^G p_j^G \\
&= \sum_{i>j} 2c_i^{-1} c_j^{-1} p_i^R p_j^R = \sum_{i>j} c_i^{-1} c_j^{-1} m_{ij}^{R^2},
\end{aligned} \tag{13.10}$$

where  $m_{ij}^R$  is the dilepton mass formed by the reconstructed lepton  $i$  and lepton  $j$  momenta. This allows us to write down an analytic form of the Jacobian matrix:

$$\hat{M} = \begin{bmatrix} \frac{\partial m_1^2}{\partial c_2} & \frac{\partial m_{4\ell}^2}{\partial c_2} & 0 \\ 0 & \frac{\partial m_{4\ell}^2}{\partial c_3} & \frac{\partial m_2^2}{\partial c_3} \\ 0 & \frac{\partial m_{4\ell}^2}{\partial c_4} & \frac{\partial m_2^2}{\partial c_4} \end{bmatrix}, \tag{13.11}$$

from which the Jacobian factor can be obtained by the determinant of the matrix  $\hat{M}$ . With this in place we can start discussing the integration strategy, as described in the next section.

### 13.3 Integration Strategy

The integral is performed by numerical methods based on Gaussian quadrature. We proceed by first scanning over a grid in the two dilepton mass dimensions ( $m_1^2$  and  $m_2^2$ ), and within each grid point we integrate over the remaining degrees of freedom ( $c_1$  and  $c_3$  in the background case, and  $c_1$  in the signal case). Since the transfer function is largest when the size of generator level momentum is the same as the reconstructed level momentum, it is natural to expect the largest contribution during the integration of mass dimensions to come from the configuration where  $m_1^G = m_1^R$  and  $m_2^G = m_2^R$ . The configuration with the largest contribution is slightly changed in the signal case with the constraint on  $m_{4\ell}^G$ . In both cases, the contribution drops rapidly with masses away from the maximum point, because of the drop in the transfer

function. A strategy has been developed to maximize the precision of convolution integration.

### 13.3.1 Mass Dimensions parametrization

We observe that in the background case the contribution from different dilepton masses is not correlated between  $m_1$  and  $m_2$ . The variation in one dilepton mass dimension is not dependent on the other dilepton mass dimension. Therefore integration along the two dilepton mass dimensions is a good choice. In the signal case, however, the correlation becomes much stronger. This is mainly due to the narrow width of the resonance. Since  $m_{4\ell}^2$  is the sum of all pairs of dilepton masses, then by having a constraint on  $m_{4\ell}^2$ , there is a negative correlation introduced between the variations in  $m_1^2$  and the  $m_2^2$ . It is advantageous in this case to pick a set of ‘‘diagonal masses’’ as the axes of the mass dimensions:

$$\begin{aligned} m_+ &= \sqrt{m_1^2 + m_2^2 + R_m m_1 m_2} \\ m_- &= m_1 - m_2, \end{aligned} \tag{13.12}$$

where  $R_m$  is formed by using the reconstructed level dilepton masses as

$$R_m \equiv \frac{m_{13}^{R^2} + m_{14}^{R^2} + m_{23}^{R^2} + m_{24}^{R^2}}{m_{12}^R m_{34}^R}, \tag{13.13}$$

and where  $m_{ij}^R$  is the invariant mass formed by reconstructed lepton  $i$  and lepton  $j$ . The reconstructed masses are fixed and taken as input to the integration process. By doing the scan in these directions, we obtain an additional Jacobian factor:

$$dc_1 dm_{4\ell}^2 dm_1^2 dm_2^2 = |J_m| dc_1 dm_{4\ell}^2 dm_+ dm_-. \tag{13.14}$$

The Jacobian factor  $J_m$  is calculated to be

$$J_m = (R_m + 2)(m_1^2 + m_2^2). \tag{13.15}$$

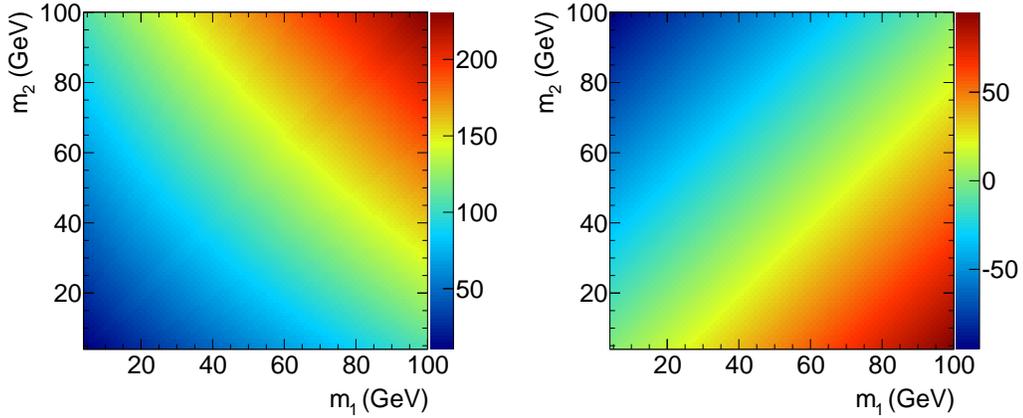


Figure 13.3: Contour of  $m_+$  (left) and  $m_-$  (right) for one example event as a function of  $m_1$  and  $m_2$ . The contour for  $m_+$  changes from event to event, but is always roughly diagonal.  $R_m$  in this case is 3.38 (see Equation 13.13).

One example of the grid lines along these modified directions are shown in Figure 13.3.

### 13.3.2 Mass Dimensions Grid Spacing

In addition to optimizing the mass axes, we also use a non-uniform grid for further optimization. Due to the nature of contributions concentrated around a certain parameter point, we choose the grid to be more dense in the center. With a uniform grid, we can write the location  $x_i^{\text{uniform}}$  of each grid point as

$$\begin{aligned}
 x_i^{\text{uniform}} &= x_{\text{center}} + \frac{x_{\text{window}}}{N_{\text{grid}}} \left( i - \frac{1}{2} N_{\text{grid}} \right) \\
 &\equiv x_{\text{center}} + \Delta_i,
 \end{aligned}
 \tag{13.16}$$

where  $x_{\text{center}}$  is the center point of the scan,  $x_{\text{window}}$  is the distance between the leftmost point and rightmost point of the grid, and  $N_{\text{grid}}$  is the number of grid points. In the last line we rewrite the displacement from the center of the scanning window as  $\Delta_i$  for convenience.

With the spacing modification, the grid spacing is characterized by an “attractor”

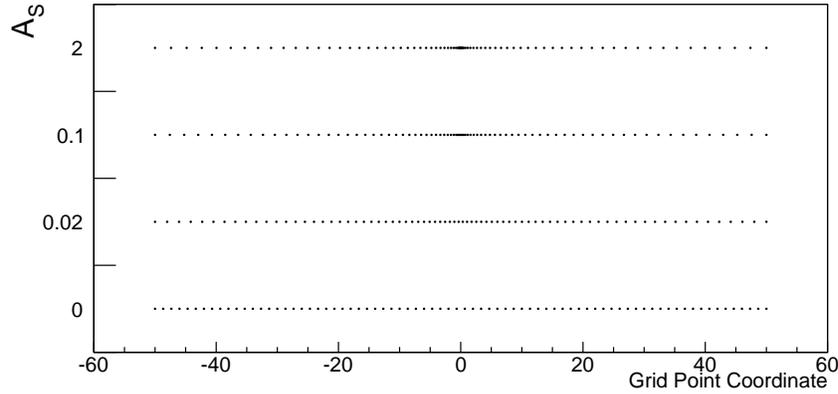


Figure 13.4: Demonstration of the mass grid attractor. Different sets of grid points are plotted in each line with varying mass grid attractor strength. With the attractor strength set to zero, we recover uniform spacing, while large values of the attractor strength cause points to be concentrated in the center of the line.

parameter  $A_S$ . Grid point locations  $x_i^{\text{modified}}$  are now defined as

$$x_i^{\text{modified}} = x_{\text{center}} + \frac{\Delta_i(|A_S \Delta_i| + 1)}{\frac{1}{2}|A_S x_{\text{window}}| + 1}. \quad (13.17)$$

The uniform linear spacing is now modified to be quadratic, with the center point and end points the same as before. A larger  $|A_S|$  value results in denser grid spacing in the center. Uniform linear spacing is recovered with  $A_S = 0$ . This is illustrated in Figure 13.4.

### 13.3.3 Modified 2nd order Newton-Cotes Formula

With the modification of the grid spacing used in the integration, it is necessary to derive the equivalent of the Newton-Cotes formula [101] for a non-uniform grid. We work with 2nd order closed integration: to each interval of three points a second-order polynomial is fitted, and an integral of the polynomial is used. Furthermore, the area can be written as a weighted sum of the points used in the fit. Suppose we have three points located at  $-\delta, 0$  and  $+\delta$ , with height  $f(-\delta), f(0)$  and  $f(\delta)$ . The integral estimator can then be written as

$$I(-\delta, +\delta) = 2\delta \frac{f(-\delta) + 4f(0) + f(\delta)}{6}. \quad (13.18)$$

In the case of non-uniform spacing where the three points are located at  $-\delta_-, 0$  and  $\delta_+$ , the integral estimator can be written as

$$I(-\delta_-, +\delta_+) = \frac{1}{3}A(\delta_+^3 + \delta_-^3) + \frac{1}{2}B(\delta_+^2 - \delta_-^2) + C(\delta_+ + \delta_-), \quad (13.19)$$

where the coefficients  $A, B$  and  $C$  can be written as

$$\begin{bmatrix} A \\ B \\ C \end{bmatrix} = \frac{1}{\delta_+ \delta_- (\delta_+ + \delta_-)} \begin{bmatrix} \delta_- & -\delta_+ - \delta_- & \delta_+ \\ \delta_-^2 & \delta_+^2 - \delta_-^2 & -\delta_+^2 \\ 0 & \delta_+ \delta_- (\delta_+ + \delta_-) & 0 \end{bmatrix} \begin{bmatrix} f(\delta_+) \\ f(0) \\ f(\delta_-) \end{bmatrix}. \quad (13.20)$$

Equation 13.19 reduces to the linear formula (Equation 13.18) when  $\delta_-$  and  $\delta_+$  are equal. The corresponding two-dimensional formula can be trivially obtained by using Equation 13.19 and Equation 13.20 multiplicatively on the two dimensions.

### 13.3.4 The Central Grid Point Optimization

We do not know *a priori* where the location of the maximum contribution will be in the integration. Therefore a trial-and-error approach is adopted. The integral is started with a best guess of the central point for the grid. In the background case it is the reconstructed dilepton masses, and in the signal case it is the reconstructed dilepton masses scaled by ratio of 4-lepton mass and  $m_H$ .

The integral is carried out repeatedly with the central point updated each time to be the point with largest contribution. This process terminates when the point with the best contribution lies in the 5% most central grid points.

We then adjust the grid window. During the integration we also keep track of the RMS of the integral contribution, in units of the number of grid points. If the RMS is

found to be less than 4 grid points in either direction, the integration is repeated with a reduced window. This ensures that we capture the core correctly, since the grid spacing at the center is small enough that there are enough grid points to capture the shape of the core, which could be narrow.

A demonstration of the whole process, on one example signal event is shown in Figure 13.5. Each bin in the plot is one grid point that we scan over, and the color indicates the value of the inner integral from that given bin. On the left we have the result of using the usual scanning directions, and on the right is the result when we scan along the modified diagonal direction, as was explained in Section 13.3.1. In the top two panels the scanning is done without the mass grid attractor, and without central the grid point optimization. In the middle two panels the mass grid attractor is turned on, but without the central grid point optimization. In the bottom panels both the mass grid attractor and the central grid point optimizations are turned on. By using all optimizations, the number of grid points with a significant contribution is greatly improved, and the correlation between the two scanning directions is decreased, thereby increasing the quality and the stability of integration.

### 13.3.5 Differential cross section expansion: Background case

In the background case, at each grid point of masses, there is a double-integral over  $c_1$  and  $c_3$  to do. Recursive integration in two dimensions would take too much computing resources, so we start by making some observations.

If the integrand of the  $c_1$  and  $c_3$  double integral can be factorized into a product of terms concerning only  $c_1$ , and terms concerning only  $c_3$ , the integral can be written as the product of two single integrals:

$$\iint f(c_1)f'(c_3)dc_1dc_3 = \int f(c_1)dc_1 \int f'(c_3)dc_3. \quad (13.21)$$

In the final integration expression for the background (Equation 13.3), the transfer functions can then be factorized into terms on the first lepton pair and terms on the

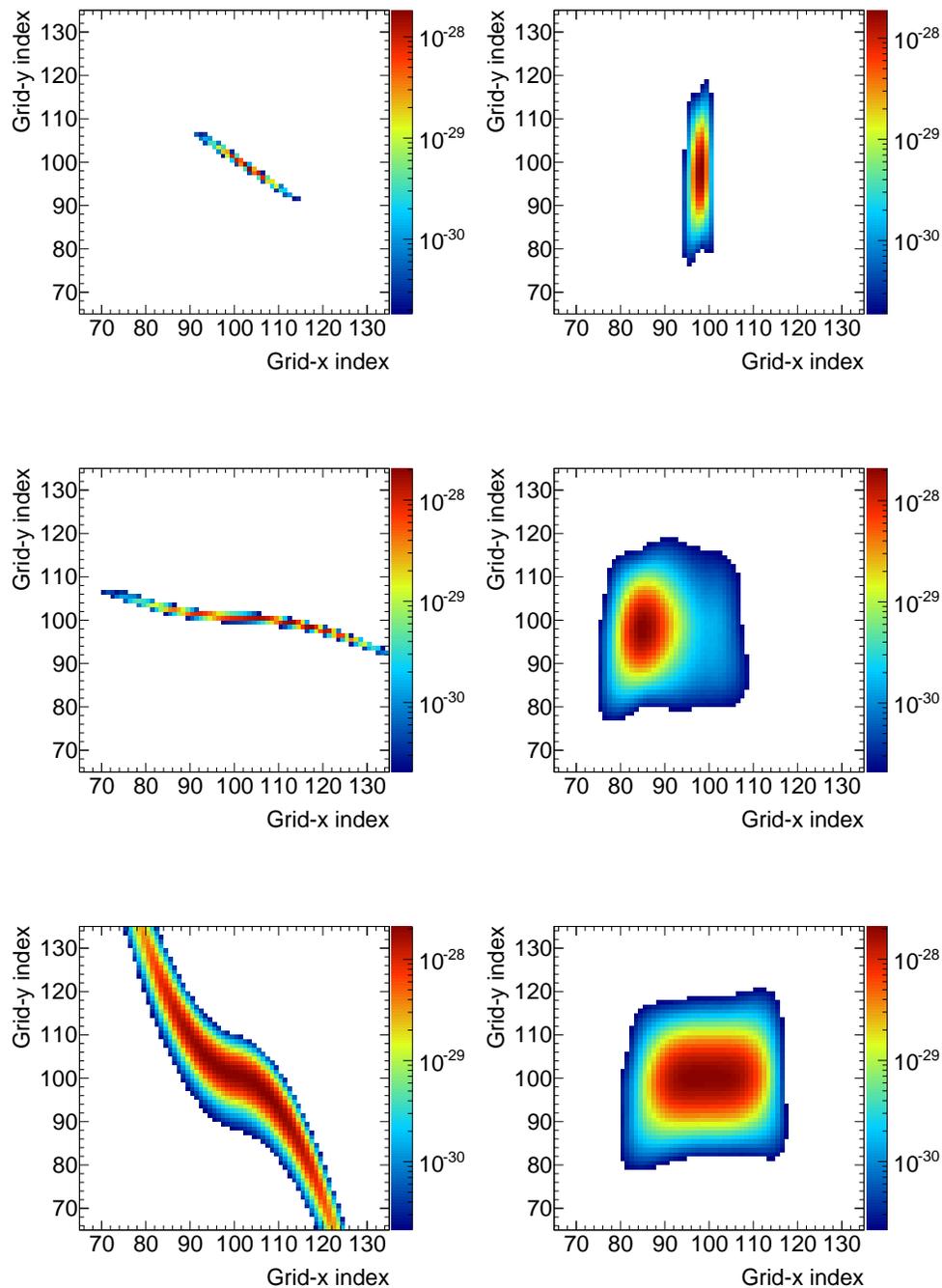


Figure 13.5: Demonstration of the mass grid attractor and the scan direction in the integration. Each bin is one grid point, and the color indicates the value of the inner integral. In the top two plots the attractor strength is set to zero. In the middle, a moderate attractor strength is turned on, and in the bottom plots the center grid point optimization is turned on. The left column is without diagonal scanning direction, and the right column is with the modified  $m_+$  and  $m_-$ . By using all optimizations, there are more grid points concentrated in the core, increasing the quality of integration, as shown in the bottom right plot.

second lepton pair, since the transfer functions are products of functions on each lepton individually.

The differential cross section and Jacobian factors however cannot be factorized. Therefore we perform an expansion around  $c_1 = c_3 = 1$ :

$$P_B(\vec{X}^G)|J^G||J^B| \equiv B_0 + B_1(c_1 - 1) + B_3(c_3 - 1) \\ + B_{11}(c_1 - 1)^2 + B_{13}(c_1 - 1)(c_3 - 1) + B_{33}(c_3 - 1)^2 + \dots \quad (13.22)$$

In the limit of expanding to infinite order, the coefficients  $B_0, B_1, B_3, B_{11}, B_{13}, B_{33}$  are the usual Taylor expansion coefficients:

$$B_0 = P_B(\vec{X}^G)|J^G||J^B| \Big|_{c_1=c_3=1} \\ B_1 = \frac{\partial P_B(\vec{X}^G)|J^G||J^B|}{\partial c_1} \Big|_{c_1=c_3=1} \\ B_3 = \frac{\partial P_B(\vec{X}^G)|J^G||J^B|}{\partial c_3} \Big|_{c_1=c_3=1}, \quad (13.23)$$

and similarly for the second order terms with double subscript. Note that these expansion coefficients do not depend on either  $c_1$  or  $c_3$ . In the case of expanding only to a certain order, we are using the polynomial on the right-hand side of Equation 13.22 to approximate the integrand  $P_B(\vec{X}^G)|J^G||J^B|$ . In the end, if precision is shown to be a problem, one can increase the order of this polynomial to obtain a better representation of the integrand.

Now let's examine the inner double integral over  $c_1$  and  $c_3$ :

$$F_B(\vec{X}) = \frac{1}{|J^R|} \iiint P_B(\vec{X}^G)T(\vec{c}|\vec{P}^G)|J^G||J_B|dc_1dc_3dm_1^2dm_2^2 \\ = \frac{1}{|J^R|} \iint \left( \iint P_B(\vec{X}^G)T(\vec{c}|\vec{P}^G)|J^G||J_B|dc_1dc_3 \right) dm_1^2dm_2^2.$$

It can be rewritten as a summation of products of single integrals:

$$\begin{aligned}
& \iint P_B(\vec{X}^G) T(\vec{c}|\vec{P}^G) |J^G| |J^B| dc_1 dc_3 \\
&= \iint T(\vec{c}|\vec{P}^G) (B_0 + B_1(c_1 - 1) + B_3(c_3 - 3) + \dots) dc_1 dc_3 \\
&= B_0 \iint T(\vec{c}|\vec{P}^G) dc_1 dc_3 \\
&+ B_1 \iint T(\vec{c}|\vec{P}^G)(c_1 - 1) dc_1 dc_3 \\
&+ B_3 \iint T(\vec{c}|\vec{P}^G)(c_3 - 1) dc_1 dc_3 + \dots \\
&= B_0 \int T_{12}(\vec{c}|\vec{P}^G)(c_1 - 1)^0 dc_1 \int T_{34}(\vec{c}|\vec{P}^G)(c_3 - 1)^0 dc_3 \\
&+ B_1 \int T_{12}(\vec{c}|\vec{P}^G)(c_1 - 1)^1 dc_1 \int T_{34}(\vec{c}|\vec{P}^G)(c_3 - 1)^0 dc_3 \\
&+ B_3 \int T_{12}(\vec{c}|\vec{P}^G)(c_1 - 1)^0 dc_1 \int T_{34}(\vec{c}|\vec{P}^G)(c_3 - 1)^1 dc_3 + \dots \\
&\equiv B_0 F_{12}^{(0)} F_{34}^{(0)} + B_1 F_{12}^{(1)} F_{34}^{(0)} + B_3 F_{12}^{(0)} F_{34}^{(1)} + \dots \tag{13.24}
\end{aligned}$$

In the last step we renamed the single integrals to  $F_{ij}^{(n)}$  for convenience, defined as follows:

$$F_{ij}^{(n)} \equiv \int T_{ij}(\vec{c}|\vec{P}^G)(c_i - 1)^n dc_i. \tag{13.25}$$

Transfer functions for the first and second lepton pairs are denoted as  $T_{12}(\vec{c}|\vec{P}^G)$  and  $T_{34}(\vec{c}|\vec{P}^G)$  respectively. The double integral over  $c_1$  and  $c_3$  is reduced to a series of single integrals over  $c_1$  and over  $c_3$  separately, which is computationally easier to control.

The coefficients in front of each term can be obtained by finding an approximate 2D polynomial to the target function over the whole integration range. In the current implementation this is done by picking a few points as representative, and finding the polynomial that goes through all the points.

## 13.4 Recursive Integration over Scaling Factors

To have a better handle on the precision, the integrations over scaling factors are done by a recursive algorithm. The algorithm starts by splitting the integration range into multiple segments, and by applying a recursive algorithm [101] to each segment with some tolerance level  $\varepsilon$ . The recursive algorithm proceeds as follows:

1. Apply second order Newton-Cotes quadrature to the whole range, to obtain an estimate  $I_0$  of integral.
2. Split the segment in half, and apply the quadrature on each of the half-segments, to obtain a second estimate  $I_1 + I_2$ .
3. Estimate the integration error  $E$  by comparing  $I_0$  to  $I_1 + I_2$ .
4. If  $|E| < \varepsilon$ , terminate the algorithm and use  $I_1 + I_2$  as the integral.
5. Otherwise, repeat the procedure on each of the two half-segments, and require that each satisfy a tolerance level of  $\varepsilon/2$ .

This ensures that the overall error is at most  $\varepsilon$ , assuming that the error estimation is reasonable. Suppose the step size is small enough that we don't have a drastic change of landscape inside each integration segment. Then since the estimator we use is a fourth order method<sup>1</sup> [101], the difference between the actual integral and the estimator can be written as

$$\int_{segment} f(x)dx = I_0 + kh^4 + O(h^5), \quad (13.26)$$

where  $h$  is the size of segment, and  $k$  is a constant characteristic of the method and independent of  $h$ . Applying the same formula to the total of the two half-segments, we arrive at the following:

$$\int_{segment} f(x)dx = I_1 + I_2 + 2k \left(\frac{h}{2}\right)^4 + O(h^5). \quad (13.27)$$

---

<sup>1</sup>the leading order of error scales as the fourth power of the step size during integration

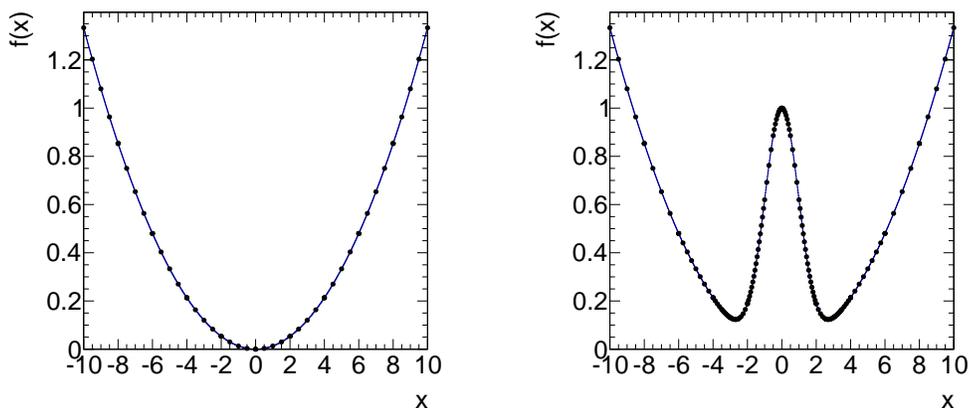


Figure 13.6: Illustration of the recursive integration. In each plot, the line shows the true function, and the dots are points evaluated from the integration. In the left panel we perform the integration on a second-order polynomial  $f(x) = \frac{1}{75}x^2$ , where the method is exact; therefore the grid spacings are always the same. In the right panel the integrand is modified to be  $f(x) = \frac{1}{75}x^2 + e^{-\frac{1}{2}x^2}$ . We see that in places where the functions are rapidly varying, more points are used.

Assuming that the term inside  $O(h^5)$  is negligible, by comparing the two formulae we obtain an estimate of integration error to be

$$E \equiv 2k \left(\frac{h}{2}\right)^4 = \frac{1}{7} (I_1 + I_2 - I_0). \quad (13.28)$$

Figure 13.6 shows an example of the integration using the adaptive method. In regions where the function varies rapidly, more evaluations are used.

## 13.5 The Signal Case

As explained in the previous sections, in the signal case, due to the presence of narrow-width resonance, anti-correlation between the two lepton pairs is induced. Furthermore by using a delta function to model the width of resonance, the double integral over  $c_1$  and  $c_3$  that is present in the background case is reduced to a single integral over  $c_1$  only.

While this makes it easier computationally, an additional complication arises since

we have to integrate along a trajectory in which  $m_{4\ell}$  is kept constant. Specifically, for any given value of  $c_1$  we have to calculate the corresponding value of other smearing factors that will keep all of  $m_1$ ,  $m_2$  and  $m_{4\ell}$  constant. It is easy to keep the dilepton masses constant, since

$$c_1 c_2 m_1^2 = m_{12}^{R^2}, \quad (13.29)$$

and we can choose  $c_2 = m_{21}^{R^2}/m_1^2/c_1$  to satisfy the condition (and similarly for the second lepton pair). For the last condition where  $m_{4\ell}$  is kept constant, more calculation is involved. From the equation

$$\sum_{i>j} c_i^{-1} c_j^{-1} m_{ij}^{R^2} = m_{4\ell}^2 \quad (13.30)$$

we can write down the solution for  $c_3$ :

$$\left( R_{34} m_{14}^{R^2} + c_1^2 R_{34} R_{12} m_{24}^{R^2} \right) c_3^2 - (m_H^2 - m_1^2 - m_2^2) c_1 c_3 - \left( m_{13}^{R^2} + c_1^2 R_{12} m_{23}^{R^2} \right) = 0, \quad (13.31)$$

where  $R_{12} = m_1^2/m_{12}^{R^2}$  and  $R_{34} = m_2^2/m_{34}^{R^2}$ . The condition where solutions exist for  $c_3$  as a function of  $c_1$  is

$$\begin{aligned} & \left[ (m_H^2 - m_1^2 - m_2^2)^2 - 4R_{34}R_{12}m_{14}^{R^2}m_{23}^{R^2} - 4R_{12}R_{34}m_{13}^{R^2}m_{23}^{R^2} \right] c_1^2 \\ & - 4R_{34}m_{14}^{R^2}m_{13}^{R^2} - 4R_{12}^2R_{34}m_{24}^{R^2}m_{23}^{R^2}c_1^4 \geq 0, \end{aligned} \quad (13.32)$$

which is a quadratic function of  $c_1^2$ . We can make a few observations based on Equation 13.32. First of all, the quadratic coefficient is negative definite, indicating that the parabola concaves downward. The constant term is also negative definite. If there exists a positive solution for  $c_1^2$ , both solutions are positive. We also require that the linear coefficient is positive, which ensures that if there is a solution, there exists a positive solution. To summarize, the two conditions that need to be met in order to

have some solutions of the scaling factor that will meet all mass requirements are

$$\begin{aligned} & (m_H^2 - m_1^2 - m_2^2)^2 - 4R_{34}R_{12}m_{14}^{R^2}m_{23}^{R^2} - 4R_{12}R_{34}m_{13}^{R^2}m_{23}^{R^2} \\ & > 8R_{12}R_{34}m_{13}^Rm_{14}^Rm_{23}^Rm_{24}^R > 0. \end{aligned} \quad (13.33)$$

The first inequality is the necessary condition for Equation 13.32 to be true, combined with the condition that linear coefficient is to be positive. If these conditions are met, all allowed solutions form an ellipsoidal contour in the positive quadrant in the  $c_1 - c_3$  plane. Some examples of such ellipses are shown in Figure 13.8.

Due to the shape of the ellipse, we integrate different parts of the ellipse separately. We “slice” this ellipse into four regions. First, two points are picked by calculating the average value of  $c_1$  between the leftmost point of the ellipse and the smaller of two points where  $dc_3/dc_1 = 0$ . Similarly we can also define another point using the rightmost point and the larger of two points where  $dc_3/dc_1 = 0$ . The four regions are thus defined by the two horizontal segments between the two aforementioned points and two vertical segments on the two sides. An example is shown in Figure 13.8. We integrate the horizontal segments normally over  $c_1$ . For the vertical segments we flip the role of  $c_1$  and  $c_3$  and integrate along the  $c_3$  direction, adapting  $c_1$  so that all mass requirements are met. This is necessary since the signal Jacobian factor  $|J_S|$  diverges at the edge of the ellipse where the direction of the curve approaches the vertical.

## 13.6 Convolution Integral Validation

As a first step we validate the convolution integral by comparing the “smeared” sample with the  $pdf$  obtained from the convolution integral. In the smeared sample we take generator-level events and smear them according to the same transfer function, including efficiency and acceptance effects. It is difficult to integrate out other dimensions in the reconstructed level  $pdf$  with other numerical methods, so Monte Carlo techniques are used where we generate samples from the  $pdf$ , and compare projections of the decay and production observables.

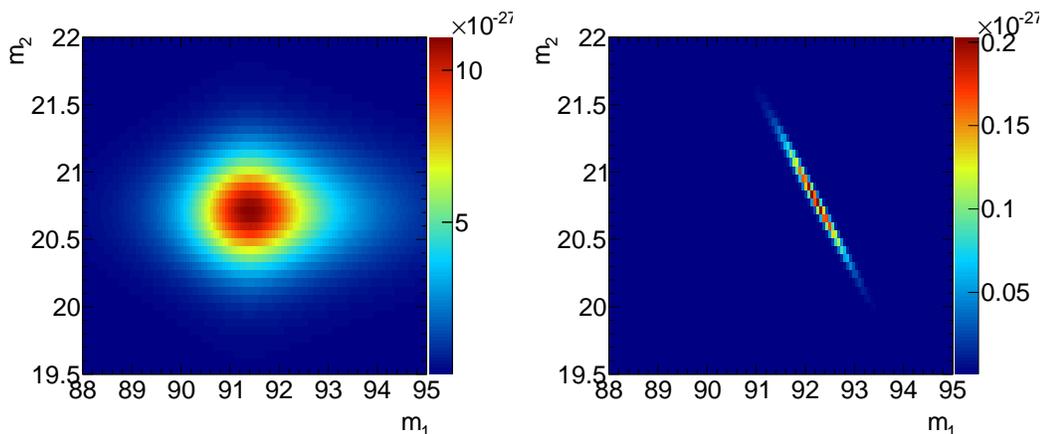


Figure 13.7: Demonstration of effect on the contribution as a function of dilepton mass due to the narrow-width approximation. In both plots the inner integral as a function of  $m_1$  and  $m_2$  is plotted for the exact same event. On the left we do not enforce the narrow-width approximation, while on the right we require that  $m_{4\ell} = 125$ . The shape is strongly elongated for the case with the narrow-width approximation.

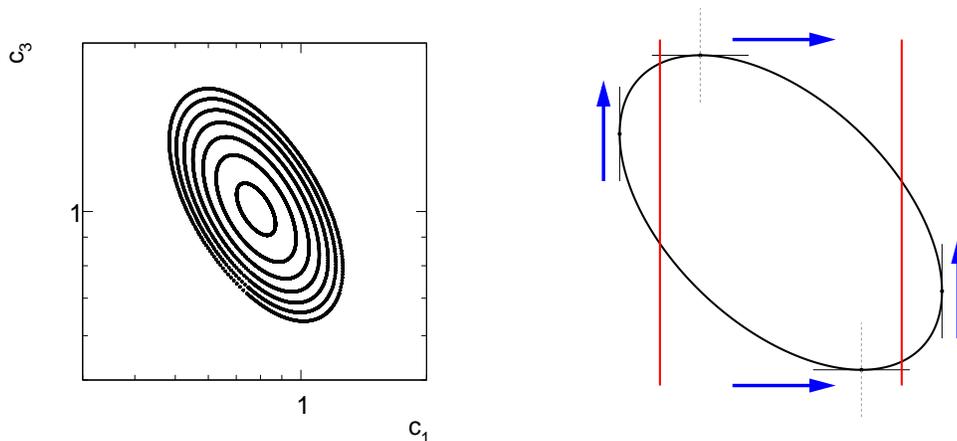


Figure 13.8: Examples of 6  $c_1 - c_3$  ellipses for a few mass grid points on an example event are shown in the left panel, where each contour represents one mass grid point. Dots represent possible solutions to all mass constraints. For each grid point, all possible solutions form an ellipsoidal shape, as expected. In the right we show the slicing of the ellipse while doing the integral. The four extreme points (left, right, up, down) are calculated, and the curve is divided into four regions by the red lines going vertically between the extreme points. In order to avoid artificial infinities during the integration, we integrate along the horizontal direction for the two segments in the middle, and vertically up along the  $c_3$  direction for the two segments on the side, as shown by the blue arrows.

Once the convolution integral is validated, we proceed with validating the transfer function, as is discussed in the earlier sections on detector modeling. Combining the two steps gives us confidence that the analysis method is robust. Finally we validate everything simultaneously by comparing the final reconstructed level *pdf* with the CMS simulated samples.

In the following sections we document the different aspects of the validation procedure.

### 13.6.1 Validation Samples Generation

Sample generation is an essential part of the validation process. A *weighted* sample can be easily generated by first generating random points uniformly in the available phase space, and weighting each sample by its respective *pdf* value. This can be carried out at either the reconstructed stage or the interpreted stage<sup>2</sup> of the event, and the algorithm is slightly different in each case. One way to do the sampling is to scan over the *detector stage* events. For these events we know the lepton pairing, and in order to calculate weight there is no need to sum over different possible pairings. Another, equivalent, way is to scan over the *interpreted stage* events. Here we throw away events landing in the parts of phase space where the configurations get reinterpreted. However for the weight we need to sum over all pairing possibilities, with the correct volume element adjustment:

$$w(\vec{X}^I|\vec{A}) = \sum_i P(\vec{X}_i^R|\vec{A}) \frac{|J_{\vec{X}^I}|}{|J_{\vec{X}^R}|}, \quad (13.34)$$

where  $J_{\vec{X}^I}$  and  $J_{\vec{X}^R}$  are the 12x12 Jacobian factors (described in Section 13.2.1), evaluated at  $\vec{X}^I$  and  $\vec{X}^R$ .

To increase the efficiency it is also possible to use importance sampling, where instead of a uniformly distributed set of random points, the random number is drawn

---

<sup>2</sup>There are three stages of each event: generated stage, reconstructed stage, and interpreted stage. The reconstructed stage is with detector effects, but before the interpretation in terms of lepton pairings. For a more detailed discussion, refer to Section 12.3

from a predefined sampling function. In this case an unbiased sample is obtained by multiplying the weight extracted from the *pdf* by the inverse of the sampling function.

### 13.6.2 Projection and Likelihood Validation

The first validation compares the “smeared sample” with the reconstructed level *pdf* for both the signal and the background, with an additional toy scenario. Specifically, we consider the following scenarios:

1. Standard model Higgs signal model point
2. Standard model  $q\bar{q} \rightarrow 4\ell$  model
3. Toy “tofu” scenario

The differential cross section for the “tofu” scenario is defined to be unity if  $120^2 < m_{4\ell}^2 < 125^2$ ,  $50^2 < m_1^2 < 80^2$ ,  $30^2 < m_2^2 < 50^2$ ,  $0.8 < \Phi_1 < 5.2$ ,  $-0.7 < \cos \Theta < 0.7$ ,  $0.5 < \phi < 5.5$ ,  $-0.6 < \cos \theta_1 < 0.8$ ,  $-0.6 < \cos \theta_2 < 0.8$ ,  $-2 < Y_{4\ell} < 1$ ,  $0 < pT_{4\ell} < 100$ , and zero otherwise. This toy scenario contains many sharp edges in the differential cross section, and it is useful to validate the modeling of smearing effects from the convolution integral.

The results are shown in Figure 13.9, Figure 13.10, and Figure 13.11. Good agreement is seen in all cases.

A second validation involves the full procedure on CMS simulated samples to verify that both the convolution integral calculation and the detector modeling are adequate.

Here we compare a number of different models, including the standard model  $H \rightarrow 4\ell$  distribution and two exotic points. The comparisons are shown in Figure 13.12- Figure 13.14. In all cases we observe good agreement between samples from the reconstructed level *pdf* and the CMS simulated samples.

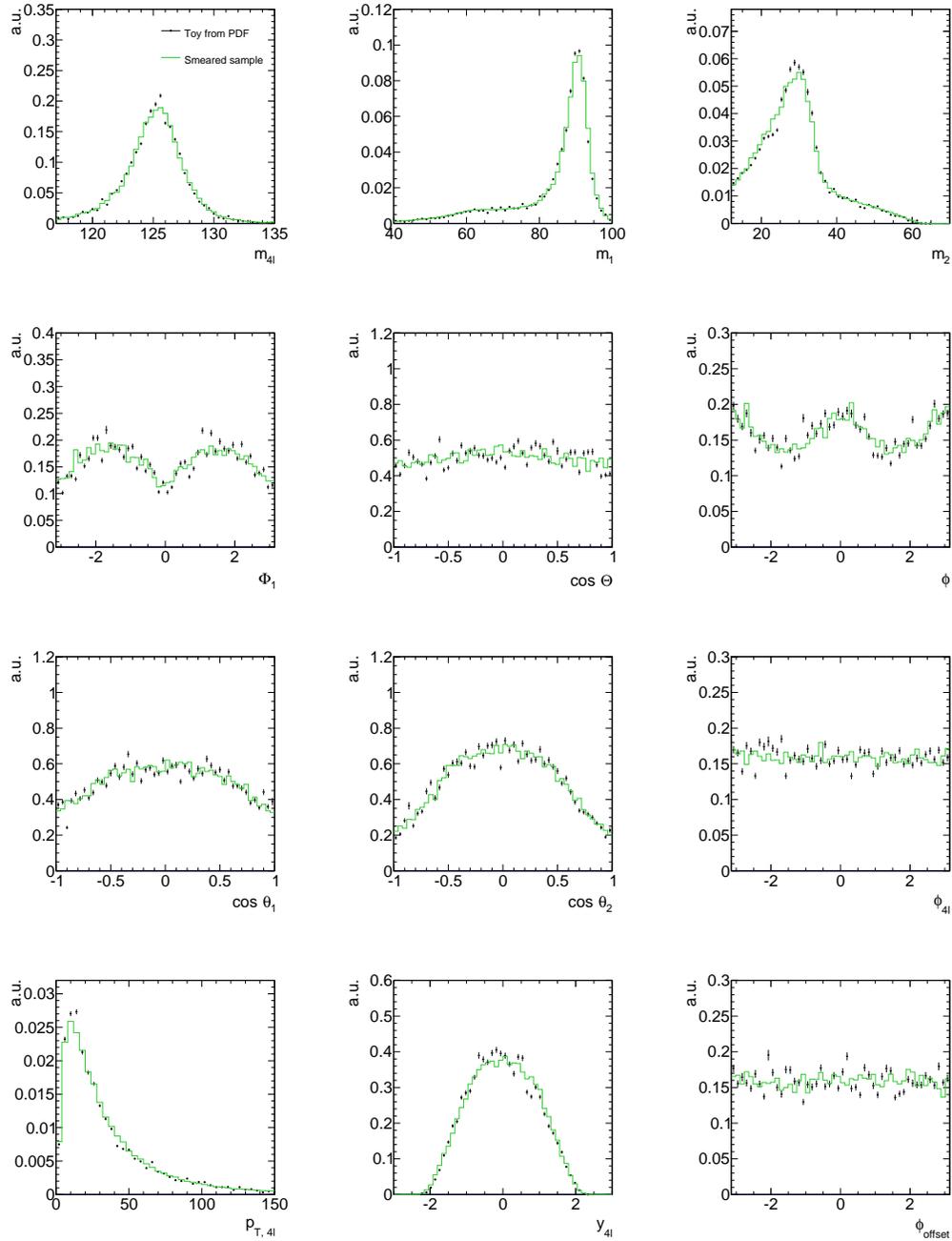


Figure 13.9: Comparison of the projections on the 12 observables for the standard model signal case, between toys generated from the reconstructed level *pdf* and the smeared sample using the same input transfer function. This serves as a cross check of the integration method.

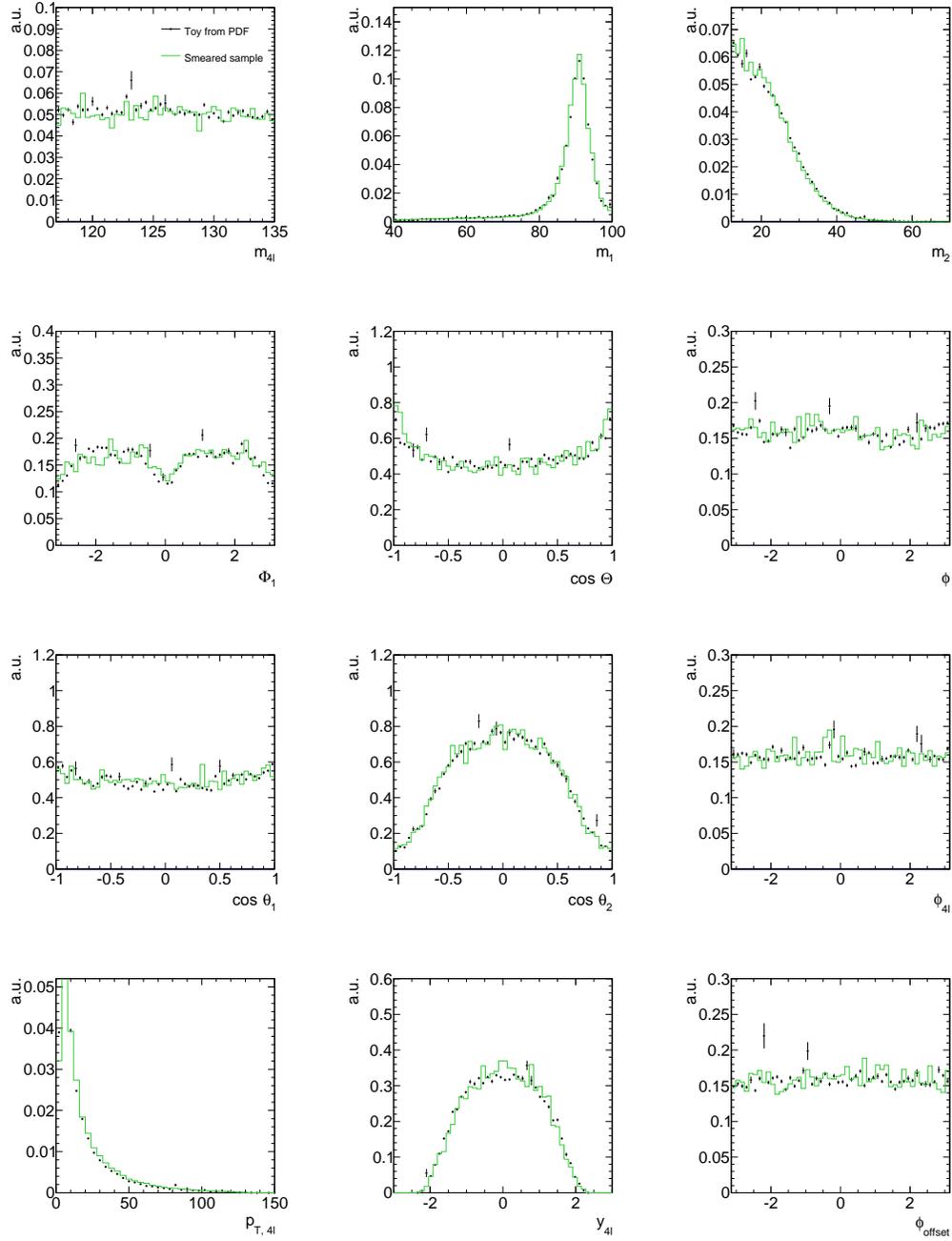


Figure 13.10: Comparison of the projections on the 12 observables for the standard model  $q\bar{q} \rightarrow 4\ell$  background, between the reconstructed  $pdf$  and the smeared sample. The goal of this comparison is to cross check the background integration method.

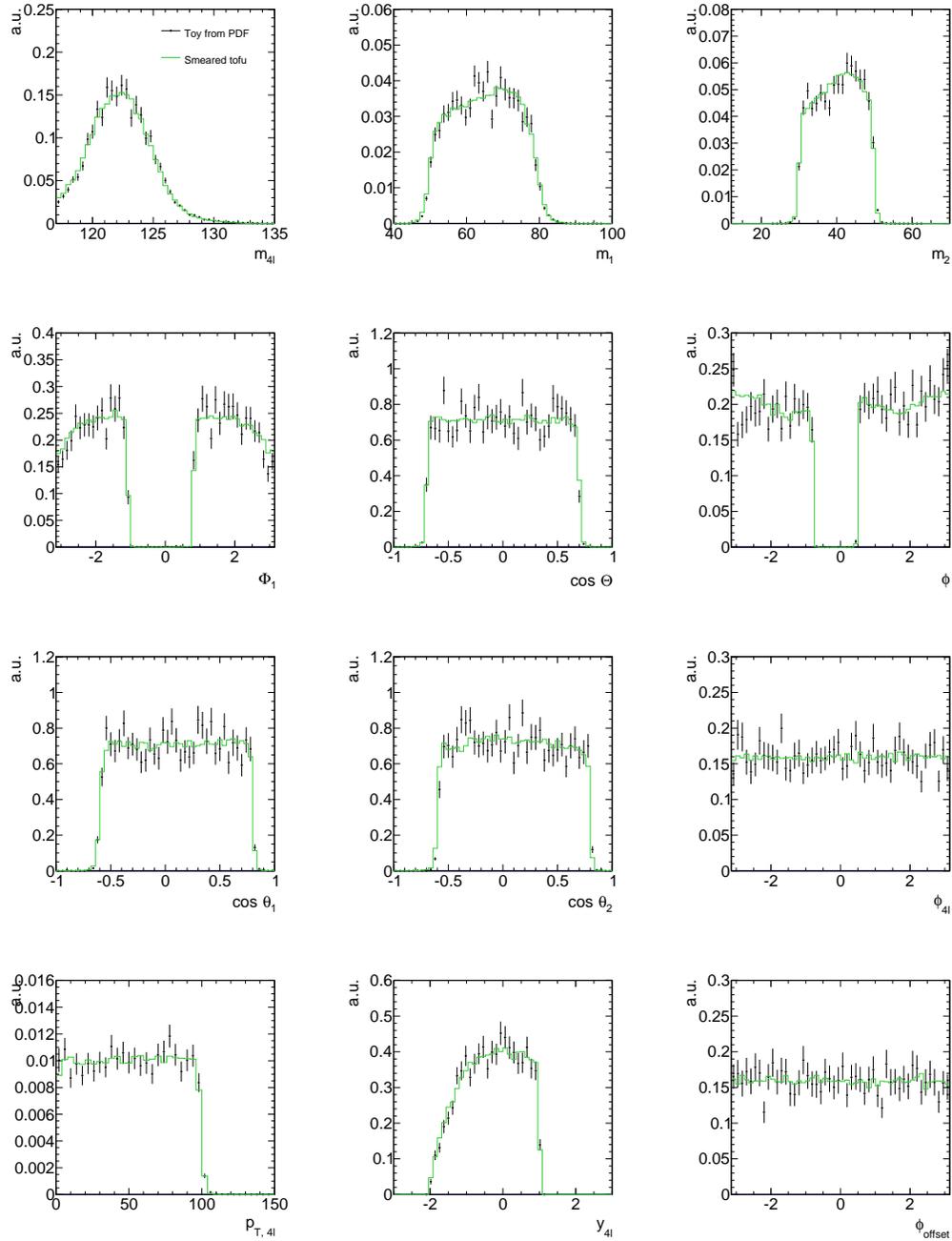


Figure 13.11: Cross check on the projections on 12 observables, between toys generated from reconstructed level *pdf* and smeared sample on a “tofu” scenario. The sharp edges present in the generator level *pdf* in this case allows us to investigate and validate the effect of smearing due to detector effects in the integration process.

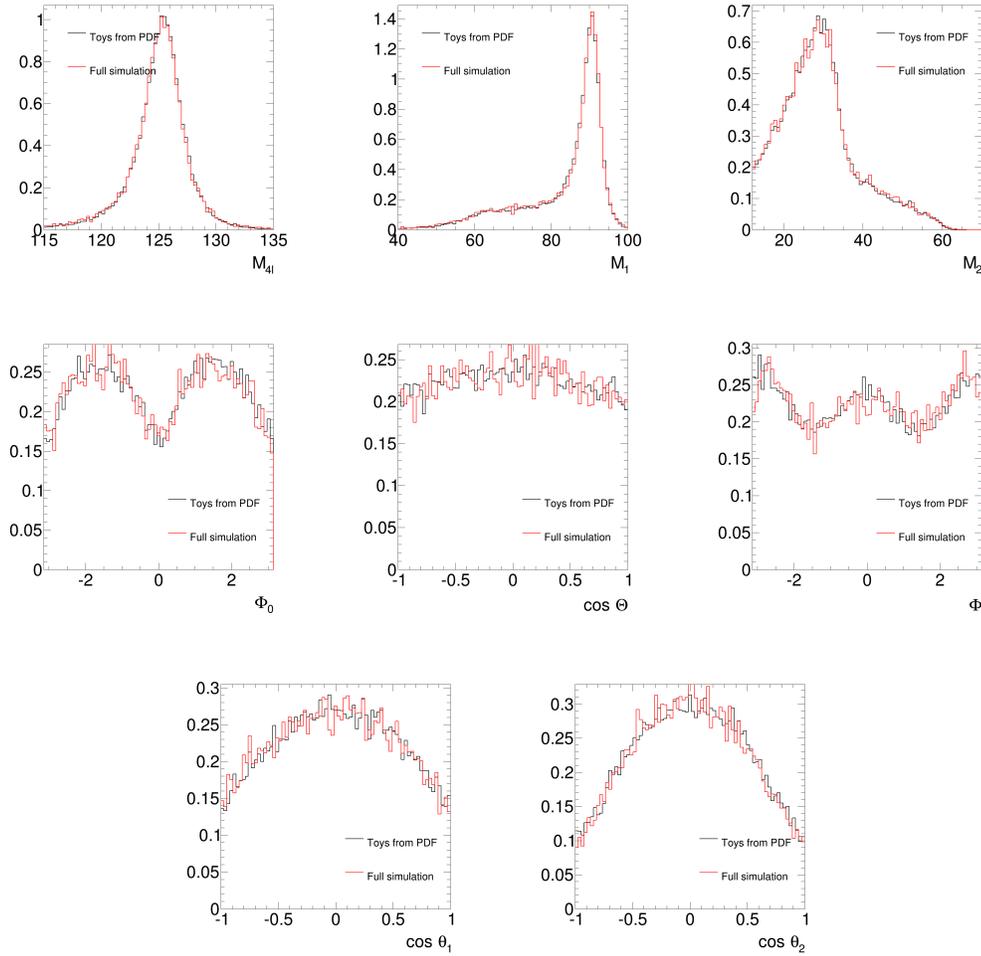


Figure 13.12: Comparison of the projections on the 8 decay observables, between toys generated from reconstructed level  $pdf$  and a CMS full simulation sample for the SM Higgs ( $A_1^{ZZ} = 2$ ). In addition to validating the integration process, this also validates the input transfer function.

### 13.6.3 Validation Fit Results

In addition to a comparison of the observable distributions, a set of fits is performed to ensure that there is no bias in the fit results. To show that there is no intrinsic bias in the fit procedure, pseudodatasets are formed by generating unweighted events from the reconstructed-level likelihood. Two model points are tested: a model with only  $A_1^{ZZ}$  present, and a model with  $A_1^{ZZ} = 2$ ,  $A_2^{ZZ} = 7.7$ ,  $A_3^{ZZ} = 10.5$ . The contributions from the square terms for  $A_1^{ZZ}$ ,  $A_2^{ZZ}$  and  $A_3^{ZZ}$  are similar for the latter model. For each model, a fit is performed on a set of statistically independent datasets of various sizes,

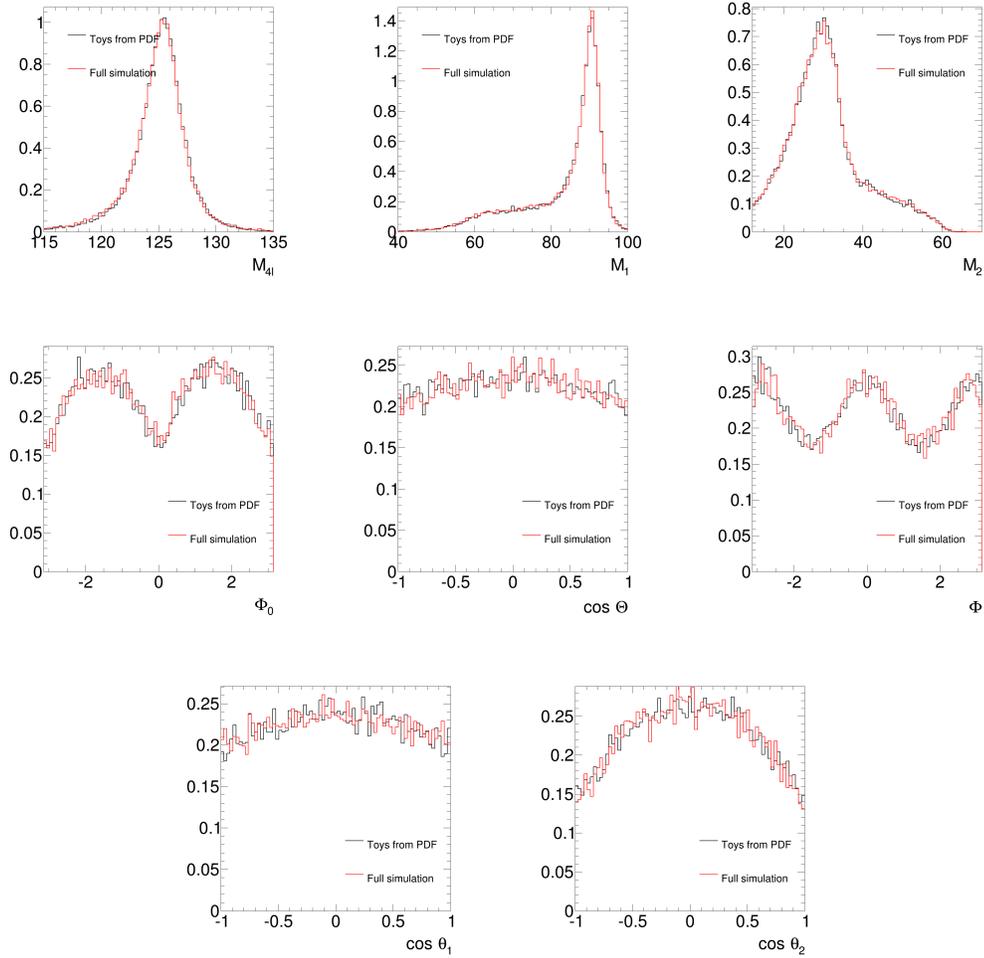


Figure 13.13: Comparison of the projections on the 8 decay observables, between toys generated from reconstructed level  $pdf$ , and a CMS full simulation sample for an exotic signal model ( $A_1^{ZZ} = 2, A_2^{ZZ} = 6.552$ ), with mixing between the tree-level  $A_1^{ZZ}$  term and the higher-order  $CP$ -even coupling  $A_2^{ZZ}$  term.

and the result is shown in Figure 13.15 and Figure 13.16. The fit results converge to the correct input value as indicated by the black line, indicating that there is no intrinsic bias in the fit procedure.

Next, fits are performed on CMS simulated samples. A set of 1000 pseudoexperiments are generated with statistics comparable to CMS Run I. For each pseudo-dataset, fits are performed by this method and compared with the fit results obtained from the template-fitting method as explained in Section 9.6. The comparison is shown in Figure 13.17. A statistically reasonable correlation is seen between the

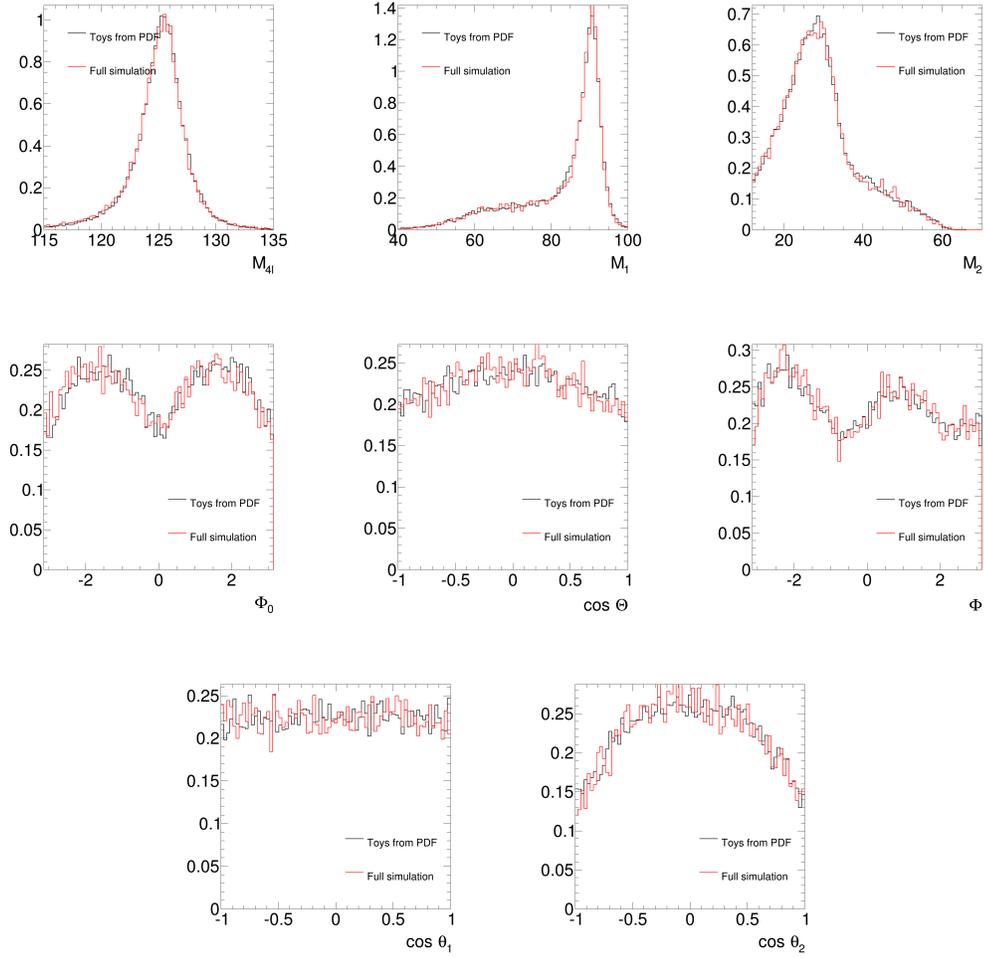


Figure 13.14: Comparison of projections on the 8 decay observables, between toys generated from reconstructed level  $pdf$ , and a CMS full simulation sample for an exotic signal model ( $A_1^{ZZ} = 2, A_3^{ZZ} = 10.084$ ), with mixing between the tree-level  $A_1^{ZZ}$  term and the higher-order  $CP$ -odd coupling  $A_3^{ZZ}$  term.

different methods.

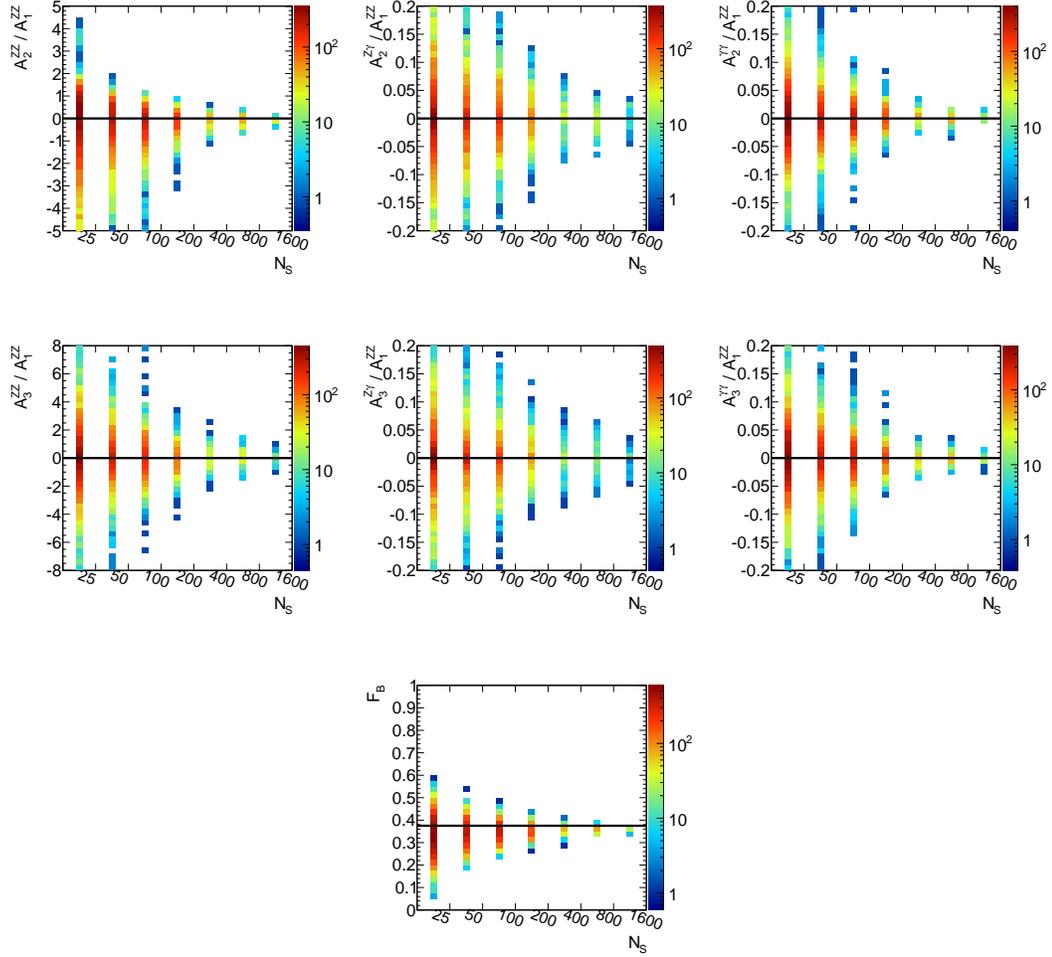


Figure 13.15: Fit results for toys generated from the reconstructed level *pdf* with the standard model. In each plot the distributions of the fit result for different dataset sizes are shown. With larger dataset sizes, the distribution of the fit result converges to the true value, indicated by the black line. On the top row the results for *CP*-even operators for  $ZZ$ ,  $Z\gamma$  and  $\gamma\gamma$ , respectively are shown. In the second row the results for *CP*-odd operators are shown. Finally we show the fitted background fraction in the bottom plot. The color in each bin represents the number of pseudoexperiments.

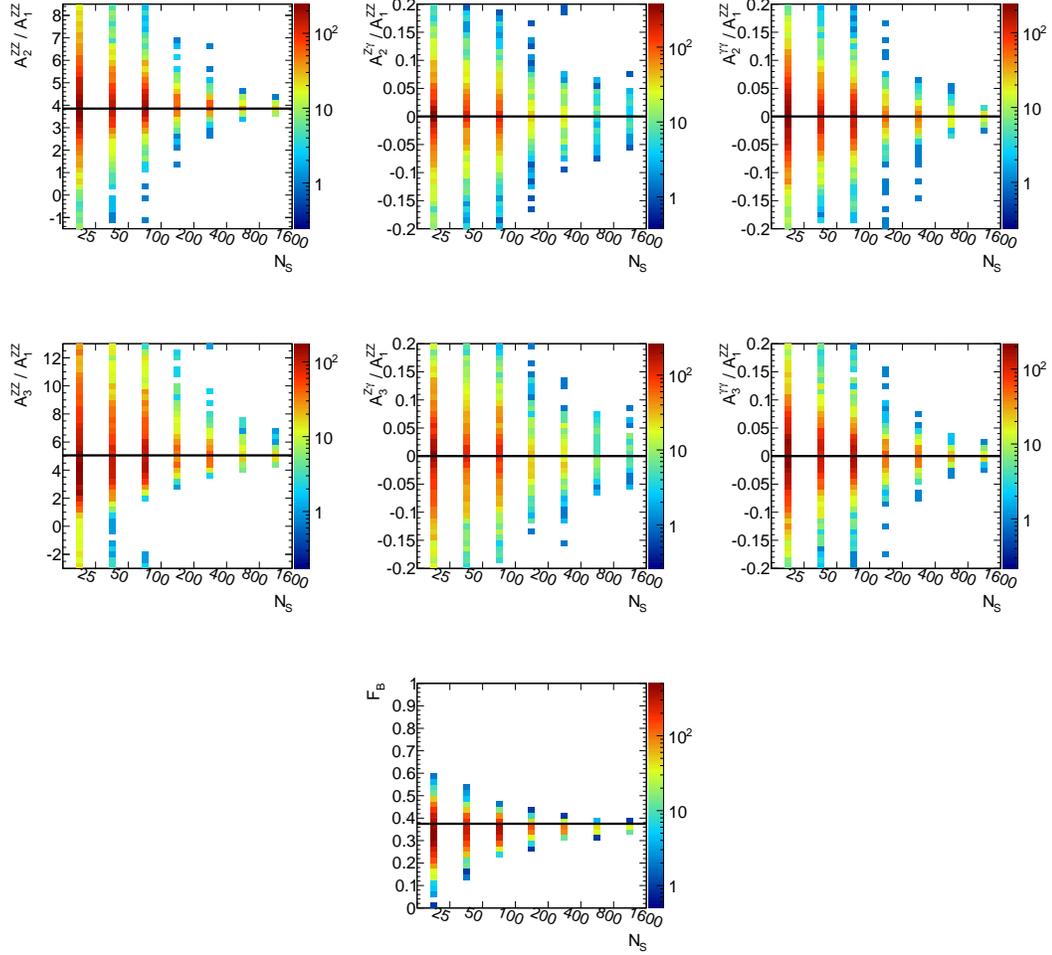


Figure 13.16: Fit results for toys generated from the reconstructed level  $pdf$ , for an exotic model with  $A_1^{ZZ} = 2$ ,  $A_2^{ZZ} = 7.7$  and  $A_3^{ZZ} = 10.2$ . In each plot the distributions of the fit result for different dataset sizes are shown. With larger dataset sizes, the distribution of the fit result converges to the true value, indicated by the black line. On the top row the results for  $CP$ -even operators for  $ZZ$ ,  $Z\gamma$  and  $\gamma\gamma$ , respectively are shown. In the second row the results for  $CP$ -odd operators are shown. We show the fitted background fraction in the bottom plot.

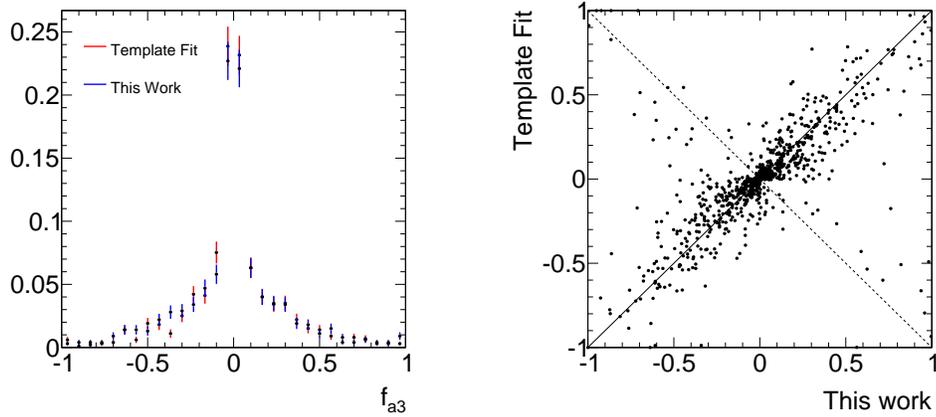


Figure 13.17: Comparison of the fit results between different frameworks using a set of pseudodatasets formed from CMS fully simulated samples, with statistics similar to that of the CMS Run I data. On the left 1D distribution for the fitted  $f_{a3}$  for this ensemble of datasets is shown. The red histogram is the result of the template fitting method, and the blue is from the method described in this thesis. On the right a scatter plot for the fitted result is shown. Each point represents one pseudoexperiments, with the result from method in this thesis in plotted as the  $x$ -axis and fit result from the template fitting method as the  $y$ -axis. We observe a statistically reasonable spread of the fitted result, as these are two different approaches with different event selections. The population in the anti-diagonal line is from the pseudodatasets with near-degeneracy between the positive and negative values of the coupling parameter  $f_{a3}$ , or  $A_3^{ZZ}$ . The only term to break the degeneracy is the interference term between  $A_1^{ZZ}$  and  $A_3^{ZZ}$ , and with the CMS Run I statistics we are not yet sensitive to the interference terms in some cases.

## Chapter 14

# Four-Lepton 8D Analysis Result

Having built and validated the method for the Higgs property measurements, we proceed to perform the analysis on the CMS data and measure the anomalous couplings  $A_2^{ZZ}$  and  $A_3^{ZZ}$  in the  $ZZ$  sector.

During LHC Run I, CMS has collected  $19.7 \text{ fb}^{-1}$  of data at 8 TeV and  $5.1 \text{ fb}^{-1}$  at 7 TeV. The full dataset is used in our analysis for the Higgs property measurements. The projections of the observed data as a function of the decay observables are shown in Figure 14.1, where good agreement is seen between the expectations and observations.

A likelihood scan on the coupling parameters is performed, and the result is shown in Figure 14.2. Two different parametrizations of  $A_2^{ZZ}$  and  $A_3^{ZZ}$  are scanned: the ratio of the anomalous coupling to the standard model, and the fractions as defined in Section 9.6 and used in recent CMS and ATLAS analyses. The results are compatible with a standard model Higgs, which predicts on the order of  $10^{-2}$  for  $A_2^{ZZ}$  and less than  $10^{-7}$  for  $A_3^{ZZ}$ .

There are a few interesting features in the likelihood scans. For  $A_3^{ZZ}$  the likelihood difference is larger away from minimum, indicating that exotic models with large  $A_3^{ZZ}$  are less similar to a standard model Higgs boson. In  $A_2^{ZZ}$ , things are more interesting: there is a special point at  $A_2^{ZZ}/A_1^{ZZ} = 3$  where the likelihood difference is the largest. The projections of the differential likelihood on  $\phi$  and  $m_2$  around this region of parameter space are examined to explain this behavior. As shown in Figure 14.3, the characteristic shoulder at  $m_H - m_Z$  in the  $m_2$  distribution disappears

due to the negative interference between the  $A_1^{ZZ}$  and  $A_2^{ZZ}$  terms. The coupling parameter value where this happens is dependent on the selection cuts, and there is no known simple theoretical prediction where this happens. Another phenomenon in this region of parameter space is the flattening of the  $\phi$  distribution, which is the opening angle between the two  $V \rightarrow \ell\bar{\ell}$  planes. The reason of the flattening is also due to the destructive interference between the two CP-even terms. Similar to the dilepton mass case, the exact location of the flattening is sensitive to the analysis selection requirement. The combination of these two effects, happening close to each other, leads to the large observed likelihood difference shown in Figure 14.2. For even larger  $A_2^{ZZ}/A_1^{ZZ}$ , the interference effect is smaller, and the distributions in both  $m_2$  and  $\phi$  are closer to the standard model case.

We note that in the likelihood scans, the observed limit appears to be much better than the “expected” limits, shown as dashed lines in Figure 14.2. The likelihood difference is always larger in the data compared to the result from the “average” Asimov dataset. In order to investigate if the effect arises from potential statistical fluctuations in the likelihood scan curves, an ensemble of pseudodatasets with comparable statistics to the CMS Run I data is used. For each pseudodataset the likelihood scan is performed, and the results are all plotted together. The result is shown in Figure 14.4. In the upper row, the color indicates the density of the pseudodataset scan results. By looking at the ensemble of results a clearer picture emerges. The observed curve from the CMS Run I data, shown as the black line, is not an extreme case in the ensemble of results. In order to better quantify the result, for each point in the parameter space, we calculate the percentage of pseudoexperiments having a likelihood value smaller than the observed data. There is no extreme values for the calculated percentage, indicating that the observed likelihood scan curve is not unusual.

In addition to the likelihood scan, which concerns the region of the model space compatible with the data, we also examine whether the observed data is compatible with a fluctuation, assuming the true model is the standard model. The fit results from an ensemble of pseudodatasets are compared with the data, as shown in the top

row of Figure 14.5. The fitted result from the data is in the center of the distribution of the toy results. The one-sided  $p$ -values are 45% and 50% respectively for  $A_2^{ZZ}/A_1^{ZZ}$  and  $A_3^{ZZ}/A_1^{ZZ}$ . The two-sided  $p$ -values can be small since both fitted results are very close to 0. In order to better evaluate if it is normal for the fit result to be so close to 0, we plot the log of the absolute value of  $A_2^{ZZ}/A_1^{ZZ}$  and  $A_3^{ZZ}/A_1^{ZZ}$ , as shown in the bottom row of the figure. The  $p$ -values in this case are 11% and 2% for  $A_2^{ZZ}/A_1^{ZZ}$  and  $A_3^{ZZ}/A_1^{ZZ}$  respectively. While the  $p$ -values are relatively close to zero, they could still be attributed to fluctuations.

If with more data the  $p$ -value decreases and the result is still too close to zero, one possibility is that the model space we probe with the fit does not cover the true underlying model. In such a case the result would be a smoking gun of physics beyond the standard model, and a fit with a larger parameter space would be needed, and would be launched, to better understand the result. At this stage, what we can say is that more data is needed to understand whether there is a potential disagreement with the standard model.

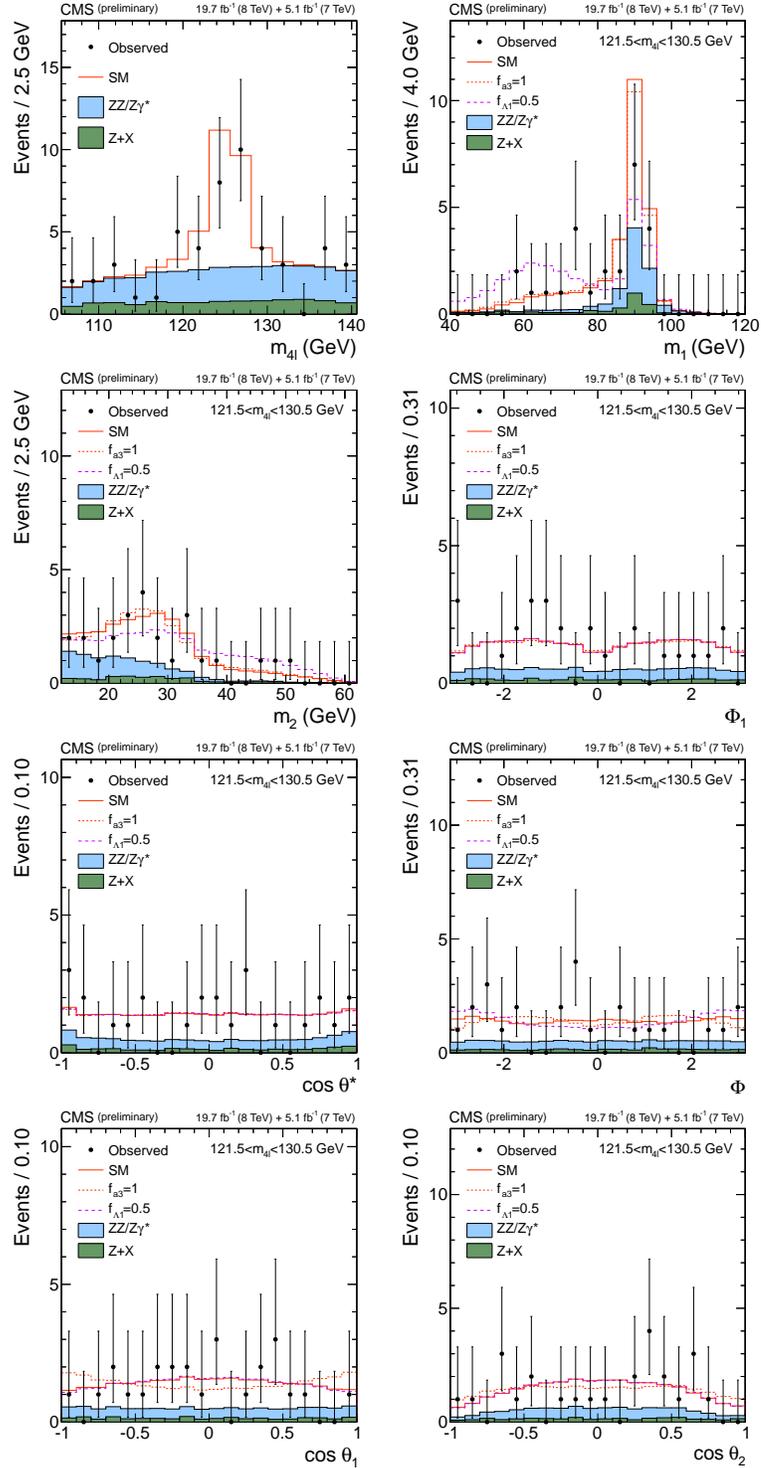


Figure 14.1: The projections of the 8 decay observables of the data on top of the predictions from a Monte Carlo sample for the irreducible  $q\bar{q} \rightarrow 4\ell$  background, and the prediction from the data for the reducible fake background. The standard model Higgs prediction at 125 GeV is also shown as the white histogram. The observed data is in agreement with the prediction for the total of the signal and the background.

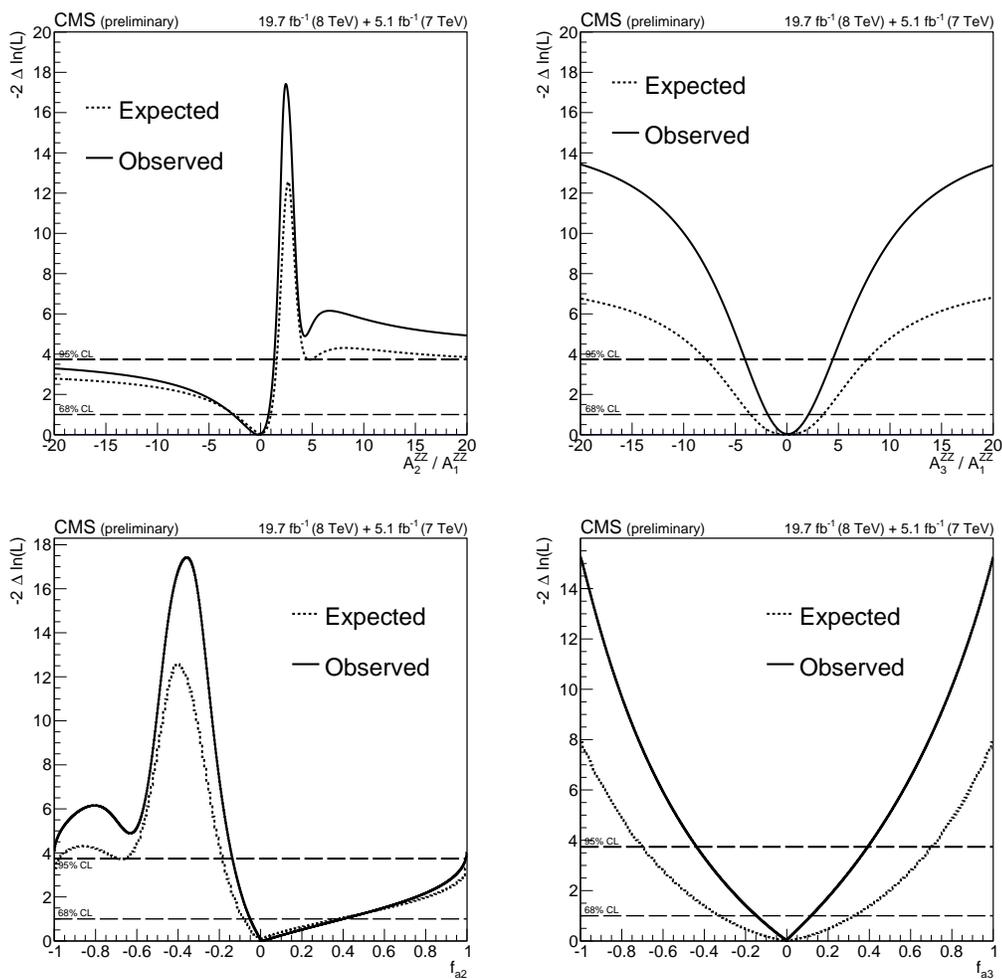


Figure 14.2: The expected (dashed) and observed (solid) likelihood scan on  $A_2^{ZZ}$  and  $A_3^{ZZ}$  for two different parametrizations. On the top the ratios of  $A_2^{ZZ}$  and  $A_3^{ZZ}$  to  $A_1^{ZZ}$  are plotted, and on the bottom the  $f_{a2}$  and  $f_{a3}$  are shown. They contain the exact same information, only the axes are different. There are some interesting features explained in text.

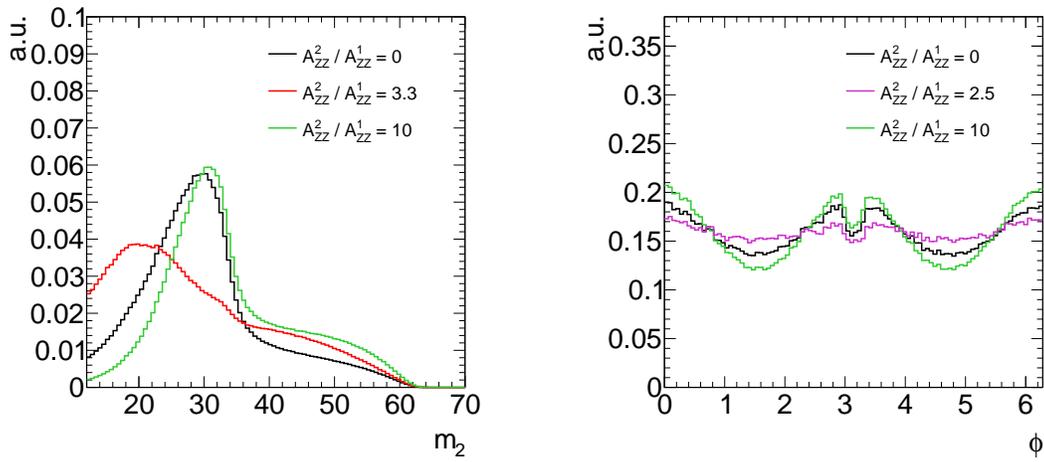


Figure 14.3: The comparison of the projection on the  $m_2$  and  $\phi$  observables, between the standard model Higgs signal and exotic signals with non-zero  $A_2^{ZZ}$ . The shape of standard model signal (black) is more similar in shape to that of an exotic model with large  $A_2^{ZZ}$  (green), while an intermediate model (red, purple) looks very different from the other two models due to interference effects. This is the main reason why there is a region with a large likelihood differences in the likelihood scan along the  $A_2^{ZZ}/A_1^{ZZ}$  axis.

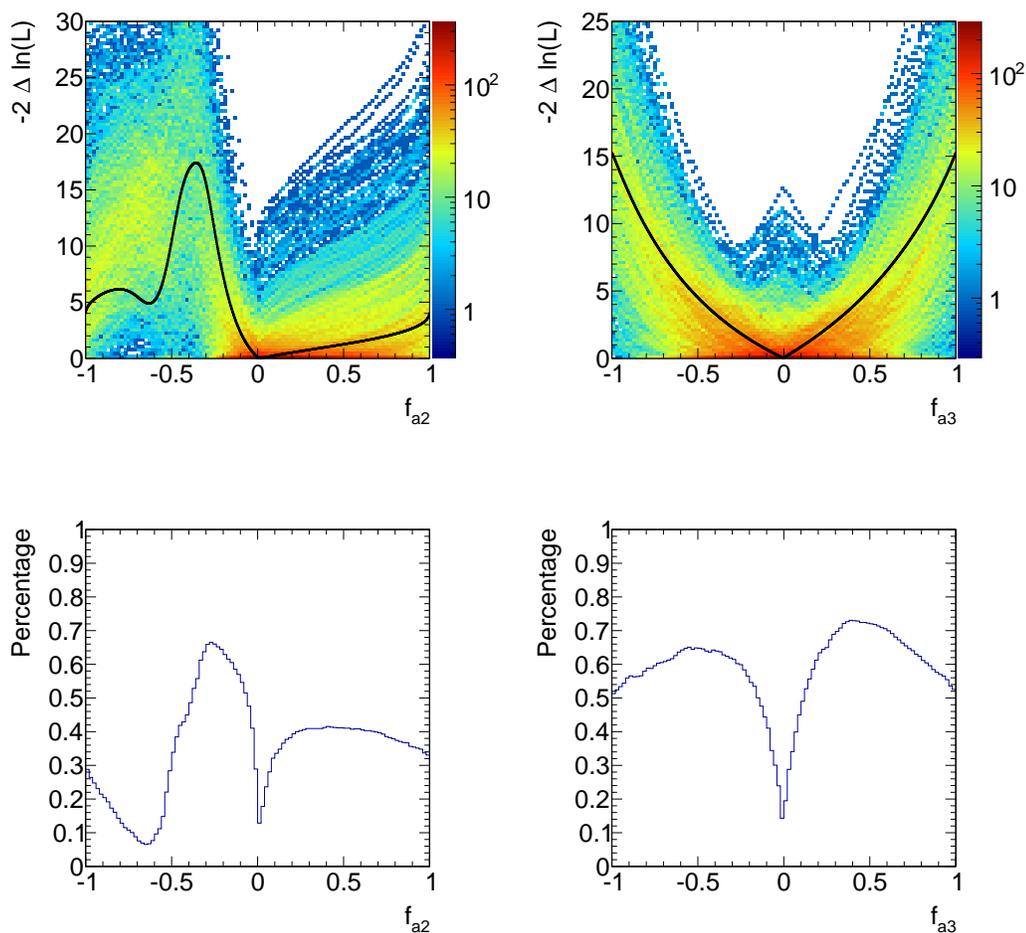


Figure 14.4: The overlay of the likelihood scans for 1000 toy pseudoexperiments are shown in the top plots for  $f_{a2}$  and  $f_{a3}$  respectively. The color scale indicates the density of likelihood scans from the toy pseudoexperiments. The observed data is shown as a solid curve, and it is in a region where the pseudoexperiment likelihood scans are relatively dense, indicating that the height of the observed likelihood scan is common. In the bottom plots we also show the percentage of the toys that lie below the observed likelihood curve for each model point. For all the exotic signal model points (where  $f_{a2}$  and  $f_{a3}$  are not zero) we do not see any extreme  $p$ -values.

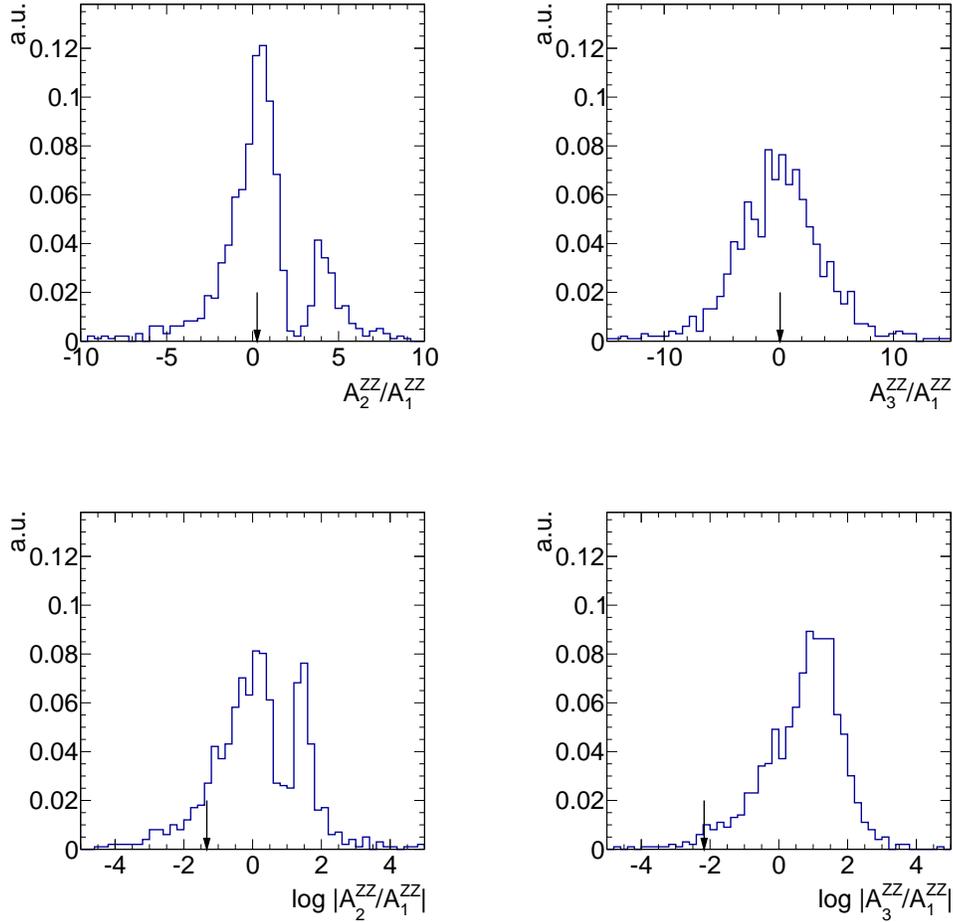


Figure 14.5: Evaluation of how probable it is for the observed data from pseudoexperiments. A set of 1000 pseudoexperiments with statistics similar to data is generated and fitted, and we compare the result from the data to the distribution of fit results in the top row for  $A_2^{ZZ}/A_1^{ZZ}$  and  $A_3^{ZZ}/A_1^{ZZ}$ . The one-sided  $p$ -values for the two cases are 45% and 50% for  $A_2^{ZZ}$  and  $A_3^{ZZ}$  respectively. Since the fit result is very close to zero, we also investigate the probability of closeness to zero by looking at the absolute value of the log of these parameters. A semi-log scale is plotted in order to better show the details close to zero. The  $p$ -values in this case are 11% and 2% for  $\log |A_2^{ZZ}/A_1^{ZZ}|$  and  $\log |A_3^{ZZ}/A_1^{ZZ}|$  respectively.

## Part III

# The Power of the Four Lepton Analysis

## Chapter 15

# Four Lepton Analysis Sensitivity to Higgs Couplings beyond the ZZ Sector

### 15.1 Sensitivity to $Z\gamma$ and $\gamma\gamma$ couplings

Even though the focus of the analysis presented so far has been mainly the  $ZZ$  sector, the  $4\ell$  channel has the potential to also probe the  $Z\gamma$  and  $\gamma\gamma$  sectors [102]. Since the couplings for  $A_2^{VV'}$  and  $A_3^{VV'}$  are dimensionless, and there is no arbitrary energy scale in the definition of the terms, and the values of these couplings are directly comparable between different sectors.

We can study our sensitivity to these couplings by generating Monte-Carlo samples assuming the standard model, and perform fits to datasets of different sizes to estimate the expected sensitivity. It is useful to discuss the size of the different terms. At leading order, the fully differential decay width for  $H \rightarrow 4\ell$  is a sum of terms quadratic in the couplings which we can write schematically as

$$\frac{d\Gamma_{H \rightarrow 4\ell}}{d\vec{X}} \sim \sum A_n^i A_m^{j*} \times \frac{d\hat{\Gamma}_{nm}^{ij}}{d\vec{X}}, \quad (15.1)$$

where the summation is over  $n, m = 1, 2, 3$  and  $i, j = ZZ, Z\gamma, \gamma\gamma$ . The couplings  $A_{1,2,3}^{VV'}$  are the ones appearing in the vertex factor as explained in Section 3.1, where  $A_1^{VV'}$  and  $A_2^{VV'}$  are  $CP$ -even couplings and  $A_3^{VV'}$  is  $CP$ -odd. The 12 observables in

the  $4\ell$  channel (Section 10.1) are collectively written as  $\vec{X}$ . We also define the “partial width” as the integrated differential cross section:

$$\Gamma_{nm}^{ij} = A_n^i A_m^{j*} \times \int \frac{d\hat{\Gamma}_{nm}^{ij}}{d\vec{X}} d\vec{X}. \quad (15.2)$$

The partial widths of all the terms assuming all higher order couplings are the same as  $A_1^{ZZ}$  are summarized in Figure 15.1. The numbers are normalized to the standard model partial width. The standard model values for these couplings are  $A_1^{ZZ} = 2$ ,  $A_2^{Z\gamma}$  and  $A_2^{\gamma\gamma} \sim O(10^{-2})$ . The partial widths including photons are much larger than the  $ZZ$ -only terms, mainly due to the larger phase space of the photons. Interference terms between different  $CP$  properties integrate to zero, as expected.

The partial widths only give part of the picture, because while the overall sizes are small, their effect on the shape of the distributions of the observables can still be large. For example, one can imagine an interference term distributed as  $\sin \phi$ , where the integral is zero but nevertheless detectable. In order to better understand the phenomenology, it is helpful to look at the “absolute partial width”, which is the integral of the absolute value of the differential cross section:

$$\Pi_{nm}^{ij} = A_n^i A_m^{j*} \times \int \left| \frac{d\hat{\Gamma}_{nm}^{ij}}{d\vec{X}} \right| d\vec{X}, \quad (15.3)$$

which is summarized in Figure 15.2 for the case where all couplings are of the same size. The normalization in this figure is chosen such that the standard model value is one, as we are only concerned with the relative size of the terms. For the square terms on the diagonal, where the differential cross section is never negative, partial widths and absolute partial widths are the same. For the interference terms a clearer picture emerged: both the interference between terms of the same  $CP$  nature and terms of different  $CP$  nature contributes. In most cases, the contributions from  $CP$ -even and  $CP$ -odd terms are of the same order.

In addition to understanding the relative size of the different partial widths, it is useful to look at the relative size in the standard model case. The partial widths

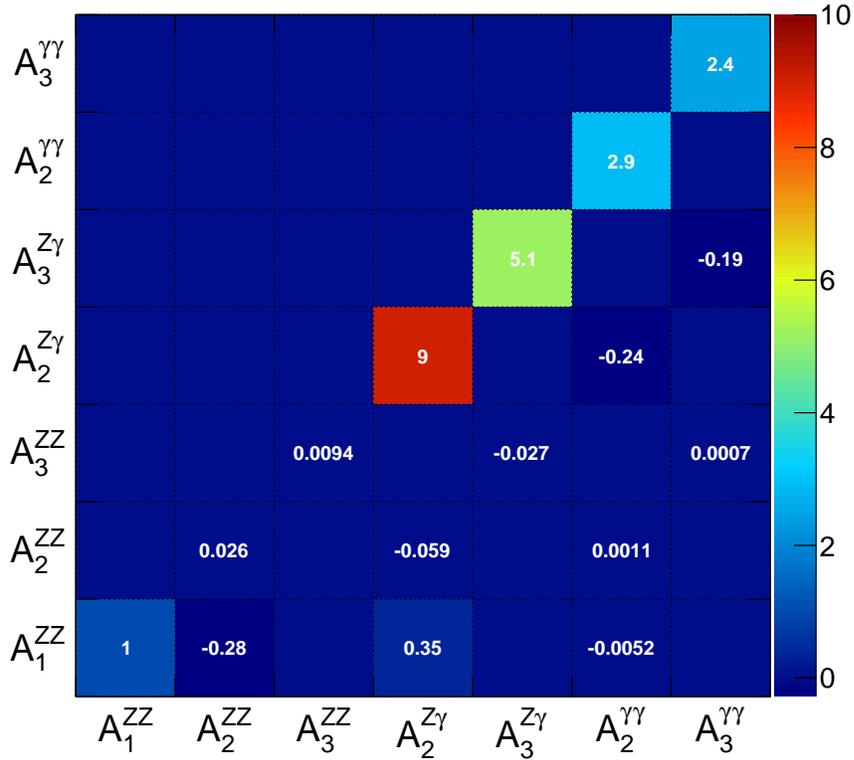


Figure 15.1: The summary of the partial widths from different terms. The partial width is defined as in Equation 15.2. The value in each bin is the size of the cross term between the two coupling parameters labeled on the  $x$  and the  $y$  axes. All coupling parameters are set to the same value, and are normalized to the standard model tree-level term. The terms involving photons are much larger primarily due to the larger phase space.

where we insert the standard model-like value of  $A_1^{ZZ} = 2$ , and CP-even couplings  $A_2^{ZZ}$ ,  $A_2^{Z\gamma}$  and  $A_2^{\gamma\gamma}$  of 0.005, 0.014 and -0.008 respectively, are summarized in Figure 15.3. The most important terms, other than the tree-level  $A_1^{ZZ}$  coupling, are the interference terms between  $A_1^{ZZ}$  and the other couplings. Note that the other terms not involving  $A_1^{ZZ}$  are non-negligible because of the large photon phase space, and are also due to the drastically different shape (for example in the  $m_1$  and  $m_2$  distributions). While it's still not a complete representation of the contribution, it is helpful for the understanding of the sensitivity projection.

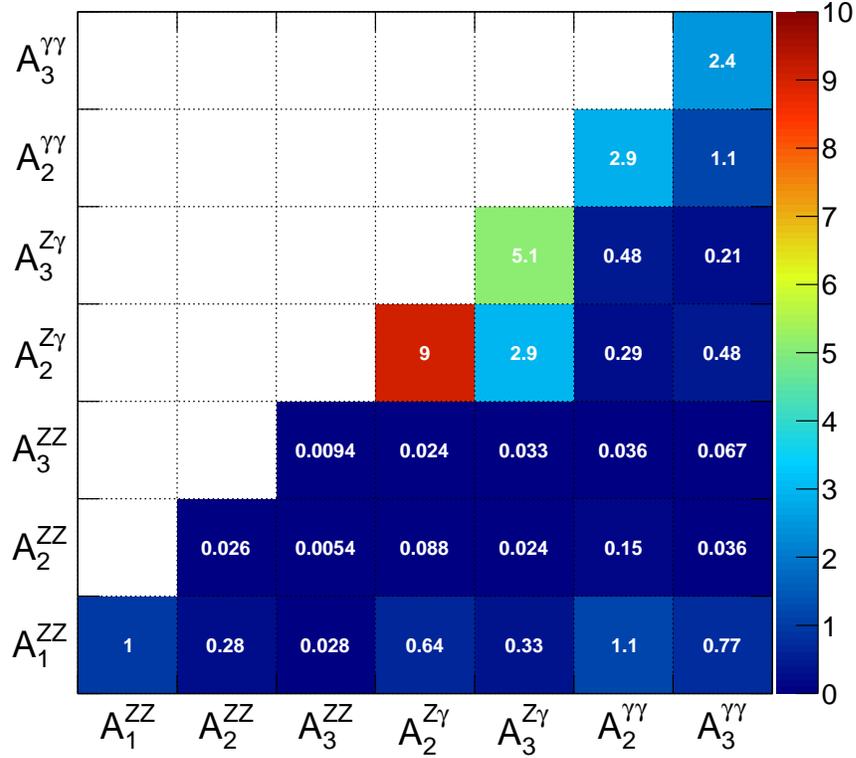


Figure 15.2: The summary of the absolute partial widths from different terms, as defined in Equation 15.2. The value in each bin is the size of the cross term between the two coupling parameters labeled on the  $x$  and the  $y$  axes. All coupling parameters are set to the same value and are normalized to the standard model tree-level term. Similar to the case of partial width in Figure 15.1, terms involving photon couplings are much larger because of phase space effect. The diagonal terms are the square terms for each coupling parameter, and since it is always positive, the values are identical to those in the case of partial width in Figure 15.1.

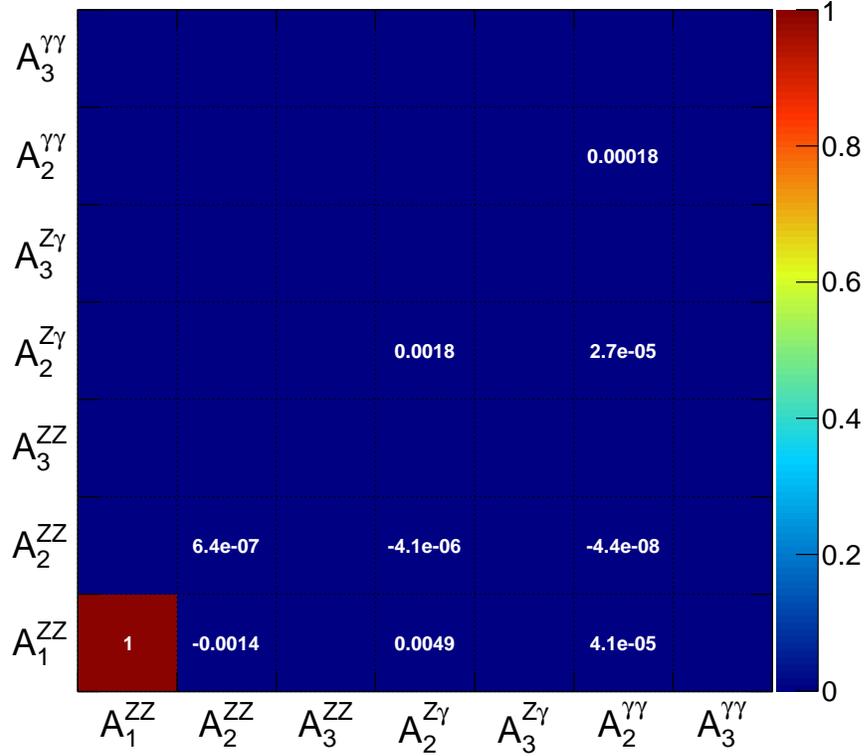


Figure 15.3: The summary of the partial widths, assuming standard model-like couplings. The value for  $A_1^{ZZ}$  is set to 2, and  $A_2^{ZZ}$ ,  $A_2^{Z\gamma}$  and  $A_2^{\gamma\gamma}$  are set to 0.005, 0.014 and  $-0.008$  respectively. The overall normalization is again chosen such that the  $A_1^{ZZ} \times A_1^{ZZ}$  term is one. In terms of the total  $H \rightarrow 4\ell$  cross section, the most important terms other than the tree-level  $A_1^{ZZ} \times A_1^{ZZ}$  are  $A_2^{Z\gamma} \times A_1^{ZZ}$ ,  $A_2^{Z\gamma} \times A_2^{Z\gamma}$  and  $A_2^{ZZ} \times A_1^{ZZ}$ . Again, when the sensitivity to different couplings is concerned, this only gives partial information, since the shape information is lost in these partial width numbers.

### 15.1.1 High Luminosity Projections

The sensitivity to each coupling, assuming all couplings are real, is summarized in Figure 15.4. The sensitivity is extracted for different dataset sizes, corresponding to different equivalent LHC luminosities at 14 TeV  $pp$  collision energy. A green line is shown, indicating the size of the standard model predicted value for the constant term in  $A_2^{\gamma\gamma}$ . All coupling constants are floated at the same time. The sensitivity is characterized by an “effective width”, which is defined as

$$\sigma_{\text{eff}} = \sqrt{\frac{\pi}{2}} \langle | \hat{A} - A_0 | \rangle, \quad (15.4)$$

where  $A_0$  is the value used to generate the datasets, and  $\hat{A}$  is the value of best fit parameter point. In the case of normally-distributed fit results, this quantity will reproduce the usual  $\sigma$  parameter of the Gaussian distribution.

There are a number of interesting features to note in the sensitivity curves. We observe a hierarchy of sensitivity between different coupling parameters. The parameters with best sensitivity are  $A_{2,3}^{\gamma\gamma}$ , followed by  $A_{2,3}^{Z\gamma}$  and then  $A_{2,3}^{ZZ}$ . This can be understood as a result of the fact that the phase space factor is much larger for the couplings involving photons ( $A_{2,3}^{\gamma\gamma}$ ,  $A_{2,3}^{Z\gamma}$ ), as shown in Figure 15.1 and Figure 15.2, which leads to a tighter constraint on the coupling parameters. Furthermore, there is a large difference in the shape of the dilepton mass distribution between  $A_{2,3}^{Z\gamma}$  and  $A_{2,3}^{ZZ}$ , and even more so between  $A_{2,3}^{\gamma\gamma}$  and  $A_{2,3}^{ZZ}$ , as can be seen in Figure 12.1.

In fact we see that a sensitivity of  $\sim O(10^{-2})$  can be reached for the  $A_2^{\gamma\gamma}$  coupling with a data sample of 800 events or more, which corresponds to 100 fb $^{-1}$  assuming 100% efficiency. We estimate that this number of events can be reached with  $\sim 300 - 400$  fb $^{-1}$  after accounting for detector efficiencies.

The next interesting feature is that there are many different slope regimes in the sensitivity curves. In the region when the number of events is small, all sensitivity curves have steep slopes, compared to those with larger numbers of events. This is mainly due to the highly asymmetrical nature of  $A_2^{ZZ}$ . As described in Chapter 14, there is a region of parameter space (around  $A_2^{ZZ}/A_1^{ZZ} \simeq 3$ ) where the distributions of

$m_2$  and  $\phi$  are much different, compared to the rest of the parameter space. Therefore when  $\sigma(A_2^{ZZ})$  is of order  $O(1)$  or larger, the distribution of  $A_2^{ZZ}$  across different pseudoexperiments becomes asymmetric. This causes other couplings to have larger effective widths, since all parameters are floated in the fit. Indeed if we instead fix  $A_2^{ZZ}$  to zero, and float all other parameters, the sensitivity curves do not show this feature.

With a larger number of events we do not see the non-symmetric behavior. Looking at the  $Z\gamma$  curves we see a bending of the sensitivity slopes. The bending is caused by the interplay between effects that dominates in each regime. The squared terms (the diagonal terms in Figure 15.2) drive the sensitivity when the dataset size  $N$  is small, and fluctuations can lead to larger extracted values. Since the fluctuation of the true model (the standard model coupling with  $A_1^{ZZ} = 2$  in this case) scales as  $N^{-1/2}$ , this implies that when the squared terms dominate we expect  $|A_{2,3}^{Z\gamma}|^2 \sim N^{-1/2}$ . Therefore the sensitivity scales as  $\sigma(A_{2,3}^{Z\gamma}) \sim N^{-1/4}$ .

On the other hand when the dataset size is large, fluctuations are small enough that the interference terms dominate:  $A_1^{ZZ} \times A_{2,3}^{Z\gamma} = 2 \times A_{2,3}^{Z\gamma} \gg A_{2,3}^{ZZ} \times A_{2,3}^{Z\gamma}$ . Since the fluctuation still scales as  $N^{-1/2}$ , we expect the sensitivity to scale as  $\sigma(2 \times A_{2,3}^{Z\gamma}) \sim N^{-1/2}$ .

As a result when the dataset size is smaller, the square terms dominate, and the slope is shallower, compared to the regime with larger dataset size, where the interference terms dominate. This transition between the two regimes happens for all coupling parameter values, as the dataset size increases. For  $A_{2,3}^{ZZ}$  and  $A_{2,3}^{\gamma\gamma}$  this effect is hidden in the non-Gaussian regime at very low dataset sizes. In order to see it one needs to examine the fit without floating  $A_2^{ZZ}$ .

Note, however, that these considerations only describe the dominant behavior, and the precise shape of the sensitivity curve depends on the details of the distributions of the observables.

The results shown in Figure 15.4 indicate that the golden channel is able to establish the  $CP$  nature and the overall sign of the  $H\gamma\gamma$  coupling within lifetime of the LHC. In order to explore this further we consider a second model given by

$A_1^{ZZ} = 2, A_2^{\gamma\gamma} = -0.008$ , which are the standard model predicted values for these couplings for a 125 GeV Higgs and on-shell external photons [103]. Even though the photons are off-shell in the golden channel, this is a good enough approximation for the purpose of this study.

We perform a fit floating all parameters, and examine the  $A_2^{\gamma\gamma}$  and  $A_3^{\gamma\gamma}$  couplings. The resulting distribution for  $A_2^{\gamma\gamma}$  as a function of the dataset size is shown in Figure 15.5, which demonstrates that we can extract the true value without bias. We show a further demonstration of the power of the golden channel in Figure 15.6 and Figure 15.7. The results from a large set of pseudoexperiments each containing 12,800 (1,600) events is shown, corresponding to an integrated luminosity of 3000 fb<sup>-1</sup> (300 fb<sup>-1</sup>) assuming a 60% (75%) detector efficiency for all  $4\ell$  channels. The expected 68% and 95% interval on the  $A_{2,3}^{\gamma\gamma}$  plane from the golden channel fits is shown as cyan ellipses. The magenta ring indicates the projection from the cross section measurement of the  $H \rightarrow \gamma\gamma$  channel. It is clear that with the diphoton channel alone, one would not be able to probe the  $CP$  nature of this coupling. Another constraint, shown as the thin green band, comes from the electron EDM measurements [104, 105]. The constraint is strongly model dependent, since it depends on the coupling between Higgs and the electron, which cannot be probed experimentally in the foreseeable future. However, even with the strong (model-dependent) constraint, the overall sign of the  $A_2^{\gamma\gamma}$  coupling cannot be determined without the input from the golden channel  $H \rightarrow VV' \rightarrow 4\ell$ .

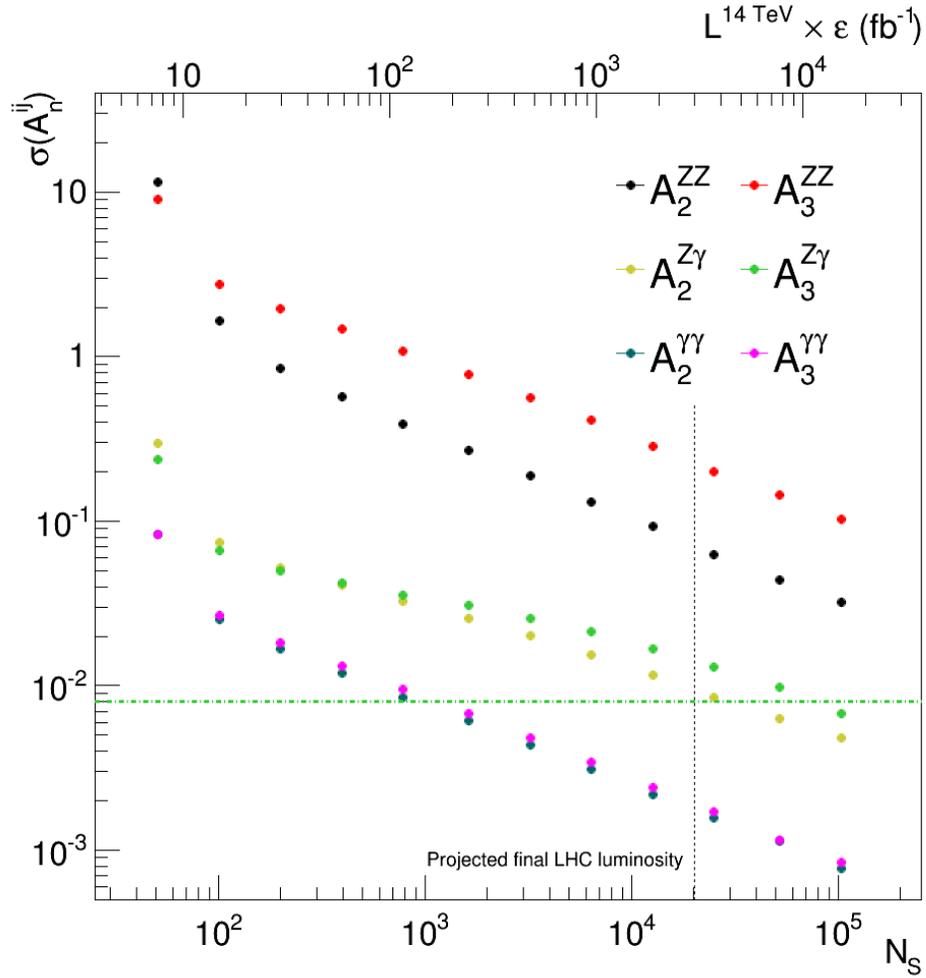


Figure 15.4: Projection of the precision for the coupling parameters, as a function of the dataset size. On the top axis the equivalent luminosity at 14 TeV is shown. All couplings are assumed real in this projection, and all are floated at the same time. A hierarchy of sensitivity  $\gamma\gamma > Z\gamma > ZZ$  is observed.

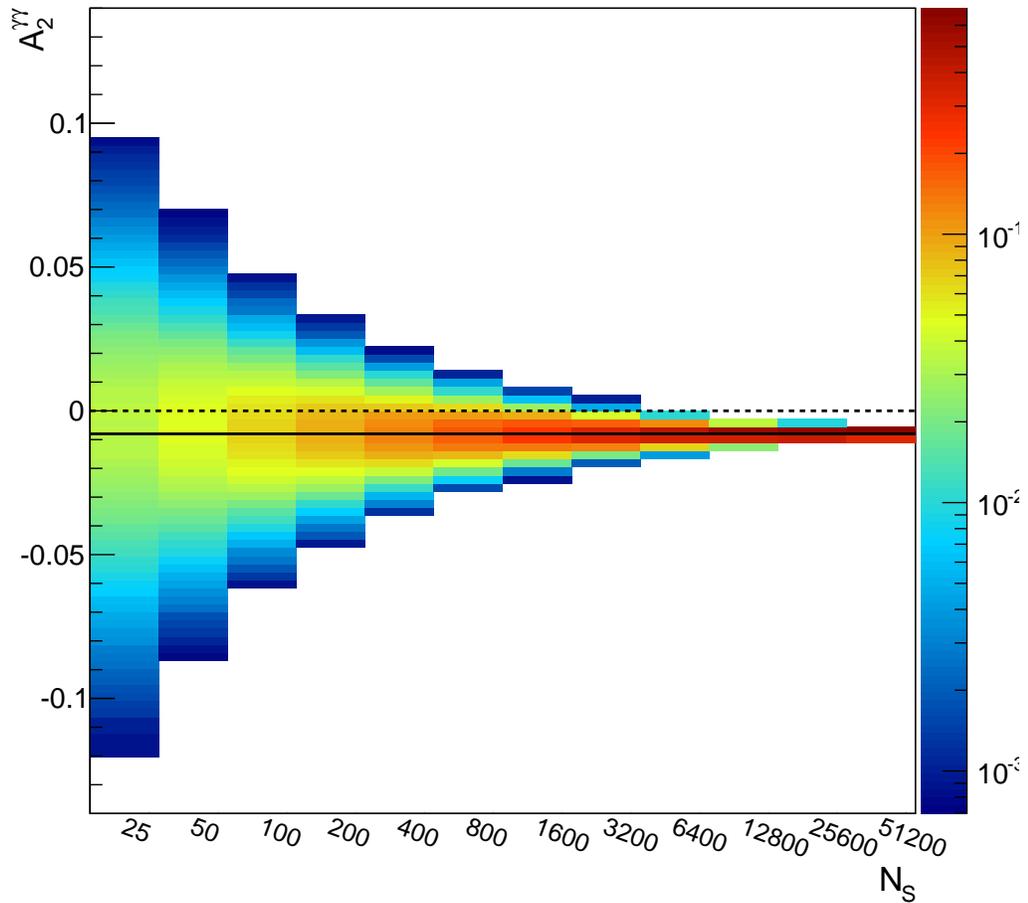


Figure 15.5: The distribution of the extracted  $A_2^{\gamma\gamma}$  parameter with all parameters floating, as a function of the dataset size. The true model used to generate datasets is  $A_1^{ZZ} = 2$ ,  $A_2^{\gamma\gamma} = -0.008$  and all other couplings set to zero. The distribution of the fitted  $A_2^{\gamma\gamma}$  parameter converges to the true value as the dataset size increases.

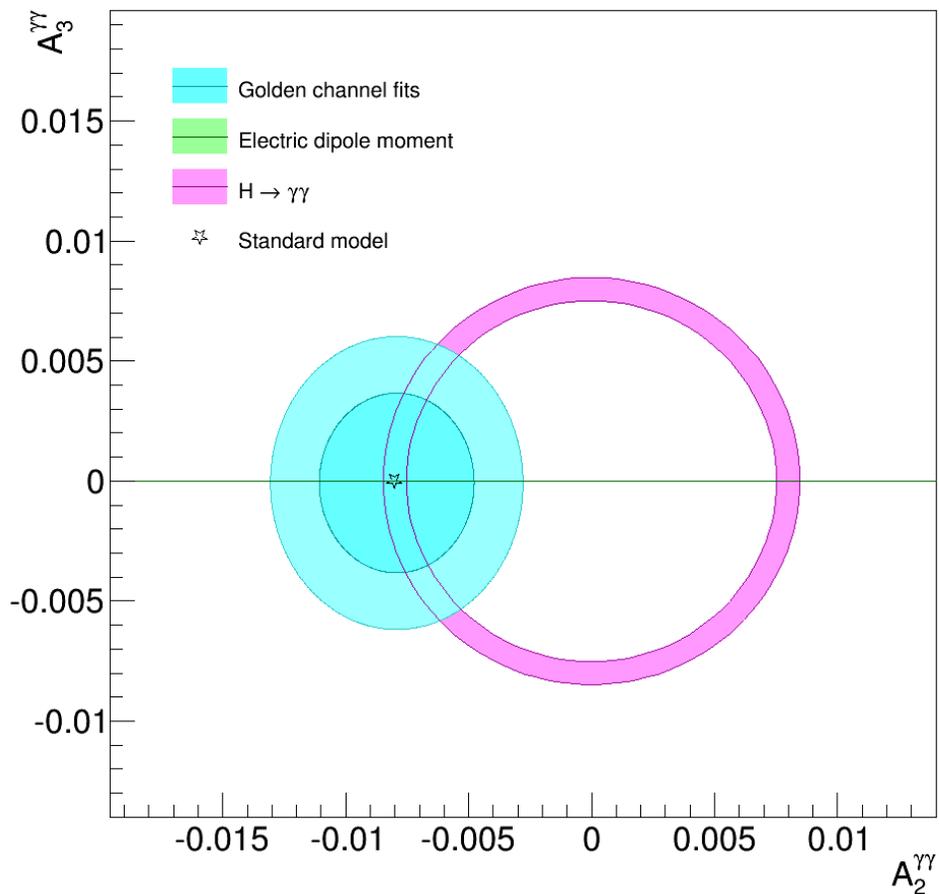


Figure 15.6: Constraints on the  $\gamma\gamma$  couplings from different measurements, with  $3000 \text{ fb}^{-1}$  of LHC data at 14 TeV. In blue, the 68% and 95% C.L. levels from the  $4\ell$  measurement is shown, while in magenta the projection of the sensitivity from the direct  $H \rightarrow \gamma\gamma$  measurement is plotted. The constraints from the electron dipole moment is shown in green. Note that the constraint from the electron dipole moment measurement is strongly dependent on the assumptions on first-generation Yukawa couplings and other potential new physics, on which we won't have measurements in the foreseeable future. The standard-model value is indicated by the star marker.

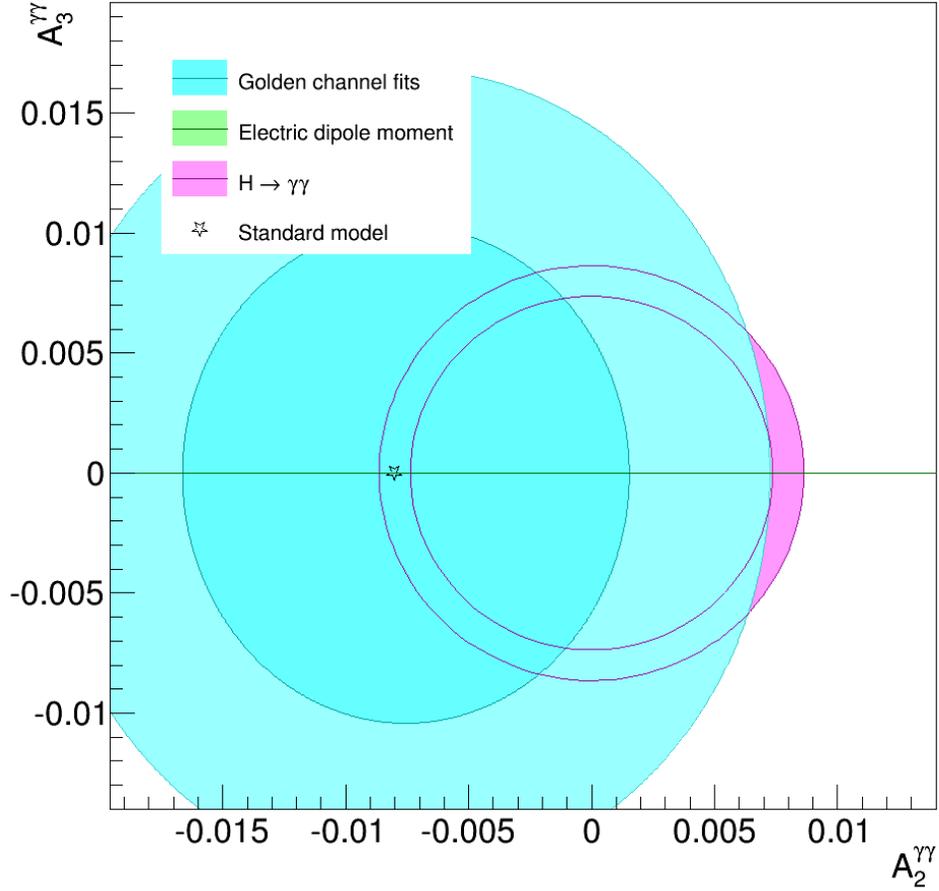


Figure 15.7: Constraints on the  $\gamma\gamma$  couplings from different measurements, with  $300 \text{ fb}^{-1}$  of LHC data at 14 TeV. In blue, the 68% and 95% C.L. levels from the  $4\ell$  measurement is shown, while in magenta the projection of the sensitivity from the direct  $H \rightarrow \gamma\gamma$  measurement is plotted. The constraints from the electron dipole moment measurement is shown in green. Note that the constraint from the electron dipole moment measurement is strongly dependent on the assumptions on first-generation Yukawa couplings and other potential new physics, on which we won't have measurements in the foreseeable future. The standard-model value is indicated by the star marker.

# Chapter 16

## Summary and outlook

In this thesis I first reviewed the standard model and the Higgs mechanism. I then proceeded to describe the effective theory approach and the allowed Lorentz structure of the  $HVV'$  effective vertex which depends on the seven coupling parameters to the lowest order: the standard model tree-level  $HZZ$  coupling, and higher order  $CP$ -even ( $A_2^{VV'}$ ) and  $CP$ -odd couplings ( $A_3^{VV'}$ ).

The helicity amplitude formulation in the case of both  $Z$ s being on-shell was described, to gain more insight into the decay kinematics of the  $H \rightarrow 4\ell$  channel. I discussed the distribution of the opening angle between the two  $Z$  decay planes when there are both  $CP$ -even and  $CP$ -odd couplings present, and the corresponding phase shift.

The potential sources of  $CP$  violation in  $H \rightarrow VV' \rightarrow 4\ell$  were also reviewed. There are two types of  $CP$  violating observables: one that arises typically in the same  $VV'$  sector, in a triple product formed from the lepton momenta  $\cos\phi = \frac{(\vec{p}_1 \times \vec{p}_2) \cdot (\vec{p}_3 \times \vec{p}_4)}{|\vec{p}_1 \times \vec{p}_2| |\vec{p}_3 \times \vec{p}_4|}$ , where  $\hat{p}_i$  is the unit vector along lepton  $i$  direction, and the other that arises from the presence of both the weak and the strong phases, for example the imaginary phase of the coupling parameters, or the Breit-Wigner line shape of the  $Z$  propagator.

I described the dominant irreducible background  $q\bar{q} \rightarrow 4\ell$  in the four lepton channel. Using the analytic differential cross section calculation for this process, I decomposed the full  $m_{4\ell}$  spectrum into pieces arising from different diagrams:  $t/u$ -channel  $VV'$  and  $s$ -channel  $q\bar{q} \rightarrow Z \rightarrow 4\ell$ .

I reviewed the current status of the LHC experiments' Higgs properties measurements. A description of the LHC and the CMS detector was followed by a summary of the Higgs discovery in different channels, and a detailed description of the  $4\ell$  analysis. I reviewed all the current measurements of the different properties of the Higgs boson, namely the mass, the couplings, the width, the spin, and the  $CP$  properties.

I discussed the main challenges in the  $4\ell$  channel: the twelve observables with non-trivial correlations, and the vast parameter space we can probe. In order to address these challenges, I developed a super fast custom detector parametrization on the leptons, reducing the computing time for each event by a factor of  $10^8$ , from a few seconds down to tens of nanoseconds. Together with the full analytic differential cross sections for the general  $H \rightarrow VV' \rightarrow 4\ell$  signal and the main  $q\bar{q} \rightarrow 4\ell$  irreducible background, the necessary ingredients to address the challenge were completed.

I then described a novel method to measure the  $CP$  properties of the Higgs. I obtained an unbinned detector-level likelihood by performing a 12D convolution integration, between the generator level differential cross section and the lepton transfer functions that model the detector effects, including acceptance and efficiencies. I described how to efficiently build the full signal likelihood for an arbitrary model, without the need to repeat the convolution integral. With the detector-level likelihood calculation in place, I performed a fit and extracted the parameters of interest using all available information.

The technical description on the convolution integration is described in terms of numerous Jacobian factors and the optimization of the integration. I validated the convolution integral by comparing the *pdf* projections with both the custom super fast simulated samples, and the CMS full simulation samples, and by the performing parameter extractions on a large ensemble of pseudodatasets, to verify that we do not have any biases.

Finally the method was applied to the  $pp$  collision data collected by the CMS detector during Run I. I discussed the various features in the result, including the apparently larger likelihood values of the observed data compared to the average (Asimov) expected likelihood value, and also how likely it is that the fitted value is

so close to zero. The worst  $p$ -value in this case is 2%, which could still be ascribed to statistical fluctuations.

I described the additional power and potential I found in the four lepton analysis. Specifically, I found that we will be able to probe the  $Z\gamma$  and  $\gamma\gamma$  effective couplings during the lifetime of the LHC. With 100-400  $\text{fb}^{-1}$  of  $pp$  collision data at 14 TeV we can begin observing the  $H\gamma\gamma$  couplings in the  $4\ell$  channel. This will serve as an important measurement to complement the direct diphoton measurement. We will also have the unique ability, in the  $4\ell$  channel, to measure the Higgs  $CP$  properties in the  $H\gamma\gamma$  couplings.

There are further measurements we can perform using the analysis framework I developed in this thesis. One example is described in the Appendix A, which shows that we can measure the standard model  $Z\ell\bar{\ell}$  coupling using  $q\bar{q} \rightarrow 4\ell$  events.

In the other appendices I describe the research work I have been involved in, during my time as a student at Caltech. This work includes the characterization and rejection of the noise in the hadron calorimeter in CMS, a measurement of the silicon photomultipliers (SiPM) with a full layer readout in the hadron calorimeter, a measurement of the differential cross section of vector boson production in association with jets, and a search analysis looking for hypothetical leptoquarks which would be an important step in uniting the quark and lepton sectors.

With the discovery of the Higgs boson, we have entered a new and exciting era in particle physics. Much work will be carried out to fully characterize the new particle, and LHC Run II is the beginning of a long experimental program. I'm looking forward to more surprises and discoveries in the coming new data.

# Appendix A

## Standard Model $Z$ Boson Couplings and Quark Charge Measurement

### A.1 $q\bar{q} \rightarrow 4\ell$ process

The exact same framework I developed in this thesis can be used to measure the standard model parameters in various channels. One good example is measurement of couplings from  $q\bar{q} \rightarrow 4\ell$  events, where relevant diagrams are summarized in Figure A.1. In the  $t/u$ -channel diagrams of  $q\bar{q} \rightarrow 4\ell$ , there are two vertices with quarks coupling to photon or  $Z$ , and two vertices with leptons coupling to a vector boson. For the  $s$ -channel diagrams there is one quark-vector boson coupling, and three lepton-vector boson couplings. The mixture between the  $s$ -channel and  $t/u$ -channels allows us to extract the relative couplings, factoring out overall yield from the list of measurements. This mixture is most useful around the  $Z \rightarrow 4\ell$  resonance peak, with the  $Z\gamma$  shoulder above it, excluding the regions where the Higgs “background” contributes. The contribution from the  $gg$  initial state increases with larger  $m_{4\ell}$ , and since the analytic expression for such a process has not been calculated yet, the high mass region in this study is excluded from consideration.

For all relevant diagrams, there is symmetry of the  $Z$  couplings to left- and right-handed fermions in the analytic expression. In both the  $t/u$ -channel  $Z\gamma$  process and the  $s$ -channel  $Z$  process where the subsequent radiation is a photon, we have only

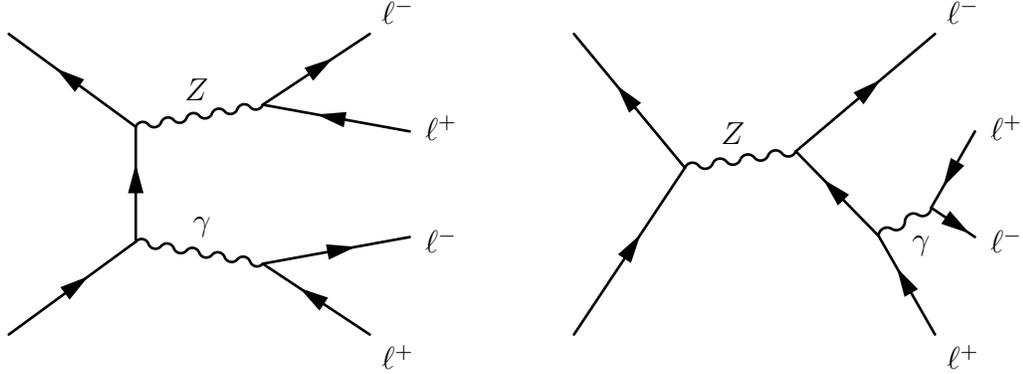


Figure A.1: Relevant diagrams in the measurement of SM parameter from  $q\bar{q} \rightarrow 4\ell$  events.

one  $Z$  coupling to quarks and one to leptons. Unless we exploit the relative size of the  $Z\gamma$  contribution to the  $ZZ$  or  $\gamma\gamma$  contribution, it is not possible to resolve this degeneracy. In the larger  $m_{4\ell}$  regime, the contribution from the  $gg$  initial state is non-negligible, while in the smaller  $m_{4\ell}$  regime there are the challenges of higher reducible background as well as lepton efficiencies at low masses.

There is also another degeneracy between the  $Z$  to quark coupling strength and the  $Z$  to lepton coupling strength, since in all relevant diagrams in this mass region we have the same order of these couplings.

When the analysis selection is limited to the mass region between the  $Z$  and  $\sim 135$  GeV, it is only feasible to measure the ratio of quark charge to the electron charge. There are 4 vertices in all relevant leading-order diagrams, and thus the expressions are 8-th order in the couplings of interest, which are the vector boson to fermion couplings:

$$P(\vec{X}^G | \vec{A}_{\text{VFF}}) = \sum_i g_q^{L n_1^i} g_q^{R n_2^i} g_e^{L n_3^i} g_e^{R n_4^i} q_q^{n_5^i} q_e^{n_6^i} P_i(\vec{X}^G), \quad (\text{A.1})$$

where sum of powers of the couplings  $n_1^i + n_2^i + n_3^i + n_4^i + n_5^i + n_6^i = 8$  for all terms, and where  $P_i(\vec{X}^G)$  is the coefficient in differential cross section before the couplings. The couplings of the  $Z$  to quarks are denoted by  $g_q^L$  and  $g_q^R$ , and the couplings to leptons are denoted as  $g_e^L$  and  $g_e^R$ . The quark charge and lepton charge are represented by  $q_q$  and  $q_e$  respectively. All couplings are collectively written as  $\vec{A}_{\text{VFF}}$ .

## A.2 Sensitivity Projection

We can proceed with the measurement by assuming the ratio between the left and right couplings, and the ratio between the quark and lepton couplings to the  $Z$  boson are given by their standard model values. This is equivalent to using an overall “ $Z$  coupling strength modifier” which multiplicatively changes all the couplings strengths of the  $Z$  to fermions. We therefore have two parameters, other than the overall yield which is proportional to the overall scale of the coupling strength. The two parameters are chosen to be as follows:

1. Ratio of the quark charge to the lepton charge
2. Ratio of the  $Z$  coupling strength modifier to the lepton charge

The most discriminating variable in this case is the  $m_{4\ell}$  spectrum, in particular the relative size of the  $Z$  resonance peak and the  $Z\gamma$  continuum. The angular information also helps in extracting the parameters of interest. To demonstrate the potential of extracting the parameters from the angular information, MADGRAPH events are generated with a  $u\bar{u}$  initial state at fixed  $\sqrt{s} = 125$  GeV, and grouped into pseudodatasets. For each dataset the parameters are extracted, and we examine the sensitivity as a function of the dataset size. The result is shown in Figure A.2.

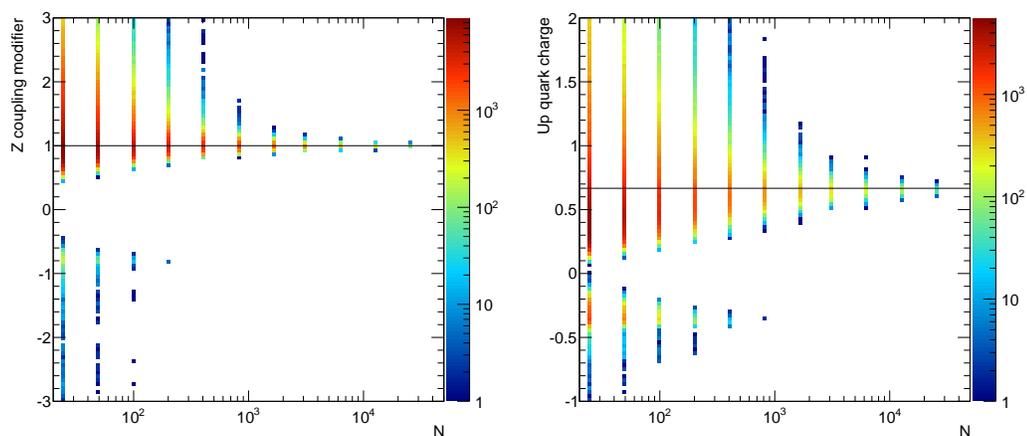


Figure A.2: Extraction of the standard model parameters using only decay information from  $q\bar{q} \rightarrow 4\ell$  events. On the left we show the distribution of extracted Z coupling strength modifier as a function of dataset size. The color in the figures indicate number of pseudoexperiments. When the dataset size is small, there is near-degeneracy between positive values and negative values of the coupling strength modifier. The only way to break the degeneracy is through the interferences between  $s$ -channel and  $t/u$ -channel diagrams, and with small dataset sizes we are not sensitive to the interference terms. On the right where we show the extracted up quark charge, there is a similar feature at small dataset sizes.

# Appendix B

## HCAL Noise

### B.1 The CMS HCAL Detector

In the barrel and endcap regions, every 18 channels are grouped and read out by a hybrid photomultiplier module (HPD), and 4 such HPDs are located inside the same readout box (RBX). The connections among the channels in  $\eta - \phi$  space are randomized when connecting to the pixels in HPDs, ie., neighboring channels in the physical detector may not be neighbors in the pixels in an HPD. This helps facilitate noise rejection, even though so far we haven't been using the full potential noise rejection capacity of all the handles we could use.

In the Run I setup there are 256 available ADC values per channel, which is converted to an equivalent deposited charge (in fC) in the reconstruction. A pedestal (of 3-5 fC) is then subtracted, based on events taken in the LHC abort gap <sup>1</sup>. The conversion from the ADC value to a deposited charge is non-linear, while the conversion from the deposited charge to energy is linear. Therefore one can realistically use deposited charge as an alias for deposited energy for many purposes, for example in looking at variables inspired by the pulse shape.

A time sample is formed every 25ns by integrating the light collected in the detector within this time window. During Run I data taking, 10 time samples were read out per channel in the barrel and endcap region, allowing us to use the pulse shape

---

<sup>1</sup>The beam structure of the LHC has a cycle of roughly  $100\mu s$ , and at the end of each cycle there is a quiet time when no collisions are happening

as a handle for further study.

## B.2 Characterization of Anomalous Noise in HCAL

Anomalous signals in the hadron calorimeter in CMS have posed challenges to physics studies related to jets and missing energy. An unphysical signal from the detector can give a large reconstructed missing energy and be mistaken as a signature of new physics. The identification and characterization of anomalous signals is thus very important, making most new physics searches possible.

Traditionally, noise in HCAL was classified into three broad categories: pedestal shifting type noise, HPD discharge, and ion feedback noise. The names are only indicative of guesses of their origin, while the actual mechanism generating the noise is never well understood.

The pedestal shifting type noise has a flat pulse shape in time, and usually happens as a collective phenomenon affecting many channels. We observe multiple channels exhibiting similar behavior at the same time. An example of the pulse shape is shown in Figure B.1. It is also called the “flat noise” due to the pulse shape.

HPD discharge type noise is also a collective behavior. There is a lot of overlap with the pedestal shifting type noise, and there is a smaller subset where the pulse shape looks different. For example, some pulses look like a real discharge: we see a flash of energy at some point in time, and the module becomes silent afterwards.

At the other extreme lies the ion feedback type noise. It is characterized by isolated signals extending up to several TeV of apparent energy. Temporally it looks like a short pulse shape, with all the energy concentrated in 25ns, in contrast to a normal physical energy deposit where a longer tail is observed, with 60% of the energy contained in 25ns and 90% of the energy contained within 50ns. One hypothesis for the source of this type of noise is that a stray ion from an hadronic shower hits the readout module, creating a flash in the detector readout. This type of “spike like noise” consists of most of the higher energy anomalous signals. An example is shown in Figure B.3.

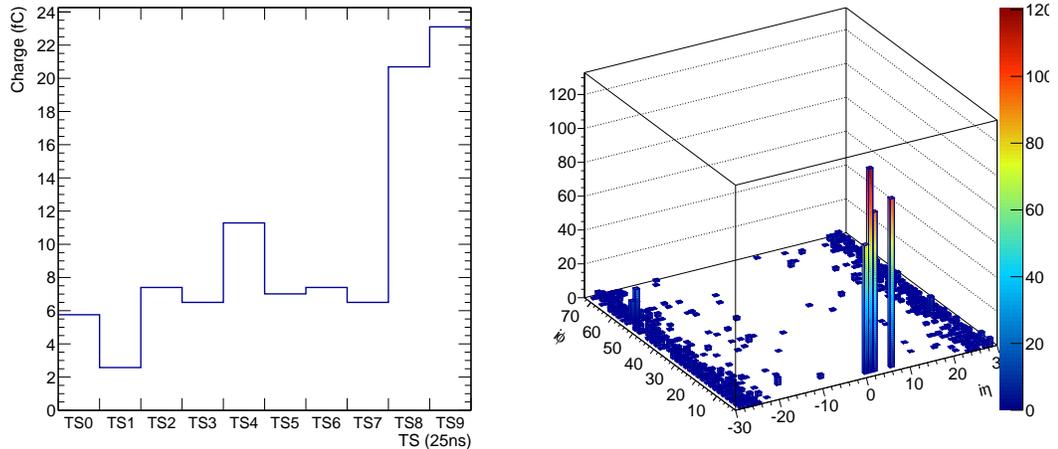


Figure B.1: Example of flat noise: (left) the pulse shape in time; (right) an event display as a function of  $\eta$  and  $\phi$  showing that several channels are affected. Each bin corresponds to a single HCAL channel. The height of the histogram indicates deposited energy in the channel.

Since the mapping of the pixels in an HPD to the location of the HCAL towers in the physical space is randomized, it is useful to look at the distribution of energy deposits mapped to the HPD pixel space, to better understand the noise behavior. In this space any physical energy deposit will be randomized, and any synchronous behavior between neighboring pixels is an indication of noise. An example of such a noise behavior is shown in Figure B.4.

### B.3 Fit-based Filters

The HCAL ideal pulse shape has been measured a long time ago in test beams, and further refined a few years ago. While not a satisfactory description of the actual pulse shape in the HCAL, it can be used as a template to distinguish noise from genuine physics signals. The “ideal pulse shape” of the HCAL is shown in Figure B.5.

I have investigated the possibility of using pulse shape fitting to describe the compatibility of a given pulse with different hypotheses, with the log of the ratio of a measure of goodness of fit acting as a discriminant, on which we place a cut to distinguish noise from real energy deposits. Since the pulse shape is by nature binned,

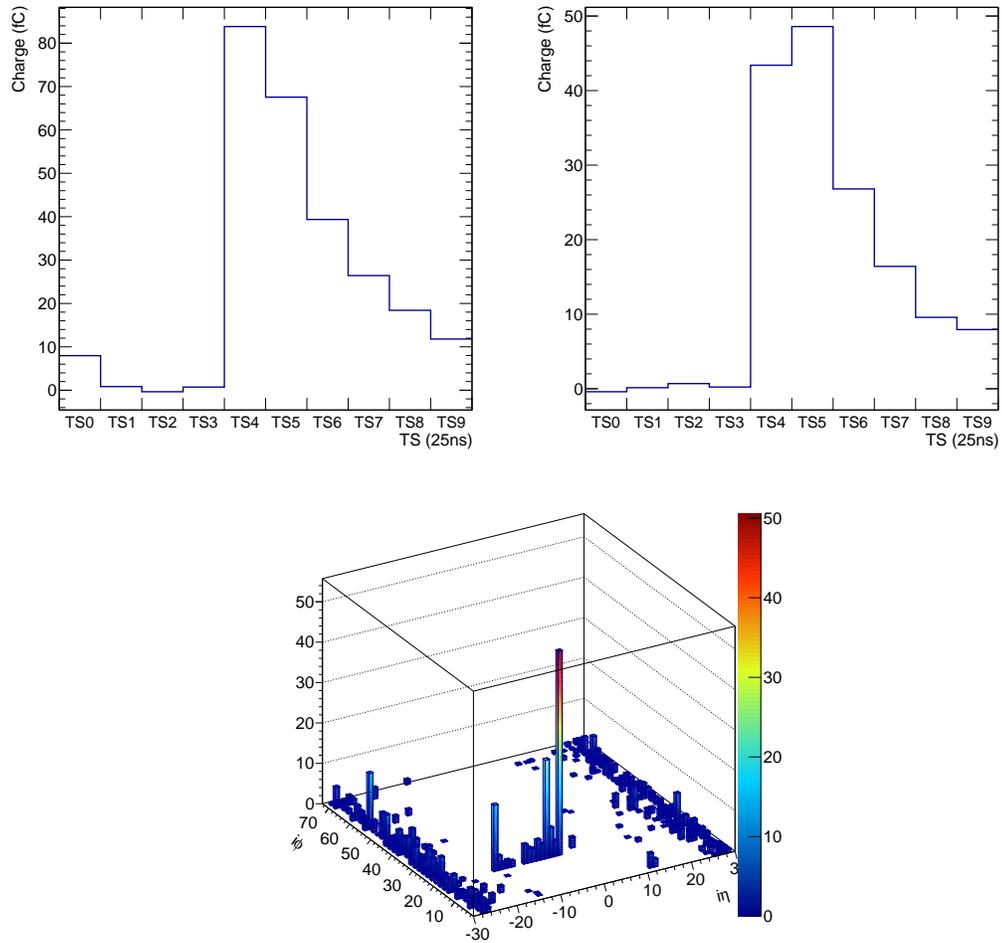


Figure B.2: Example of triangle noise: (left) pulse shape in time; (right) event display of the same event showing the collective behavior of this type of noise. Each bin represents the reconstructed energy in a single channel in GeV.

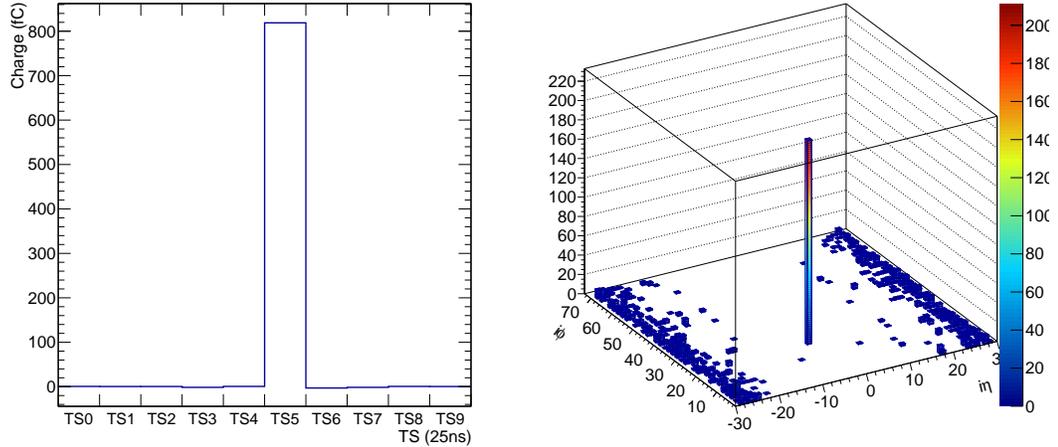


Figure B.3: Example of spike-like noise: (left) pulse shape; (right) the event display of an example event containing a spike-like noise. This type of noise usually affects only small number of channels.

we can choose  $\chi^2$  as the measure of goodness of fit:

$$\chi^2 \equiv \sum_i \left( \frac{T_i - f_i(\vec{\phi})}{\sigma_i} \right)^2, \quad (\text{B.1})$$

where  $\vec{\phi}$  is a set of parameters describing each candidate pulse shape hypothesis, where  $T_i$  is the size of the observed pulse shape at time sample  $i$ , and where  $\sigma_i$  is the error associated with the pulse height for that time sample. Due to the Poissonian nature of light collection in the detector, we expect the error  $\sigma_i$  to be as follows:

$$\sigma_i = k\sqrt{T_i}, \quad (\text{B.2})$$

where  $k$  is some constant characteristic of the detector. This constant can in principle affect the fitting and subsequently the  $\chi^2$  value, however for our purpose it doesn't matter. Observing that  $k$  factors out from the  $\chi^2$  expression, different reasonable choices of  $k$  won't make a difference in the fit result. Furthermore, since we are interested in forming the log of ratio of  $\chi^2$  values from fits to different pulse hypotheses, this factor cancels in the process.

We can form different pulse shape hypotheses:

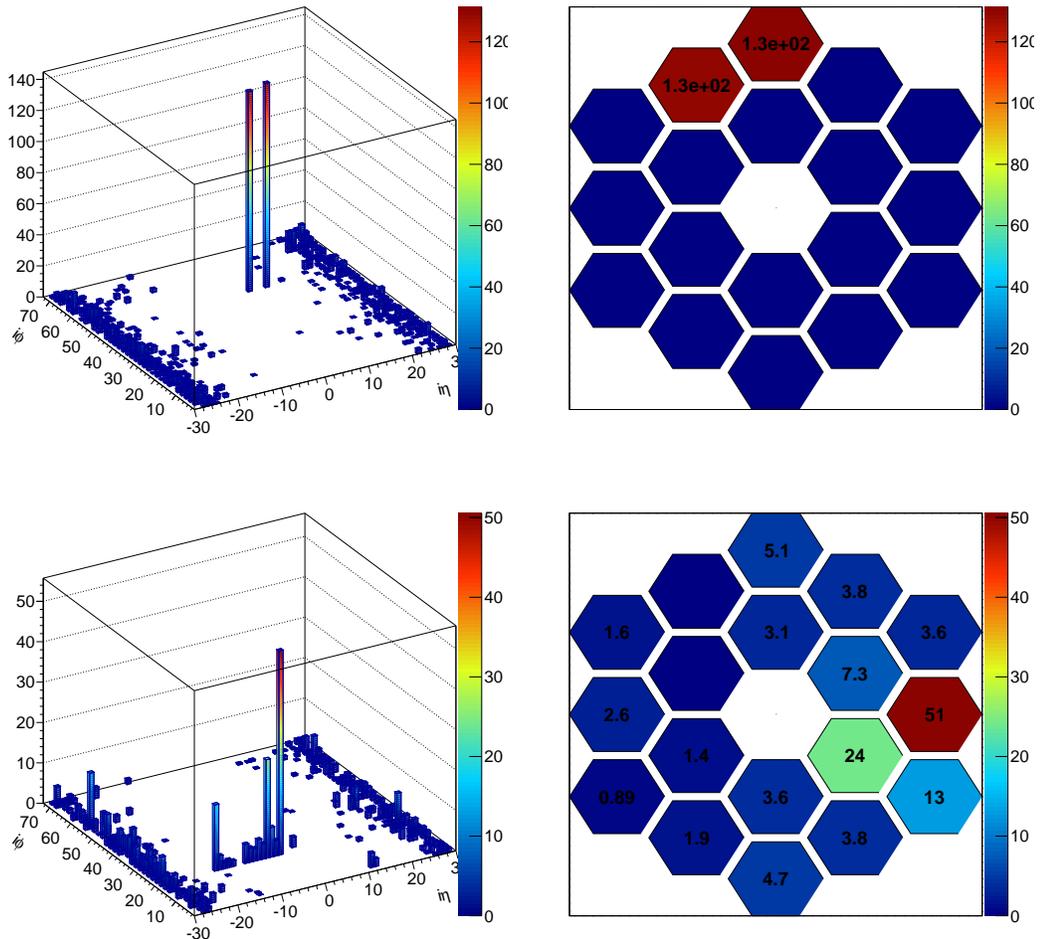


Figure B.4: Two examples of events containing noise. The left plots are the energy deposits as a function of  $\eta$  and  $\phi$  of the HCAL tower. Each bin represents one single HCAL tower. On the right is the energy measured in each of the pixels in the corresponding problematic HPD. As seen in these examples, noise-like energy deposits which are randomized in the physical  $\eta - \phi$  space tend to be neighbors in the HPD pixels.

1. Ideal pulse shape. We allow the pulse to move in time, and the pulse height to vary for different pulses. To extract the time from the fit, we build in the time slewing effect in the pulse shape hypothesis, where lower energy pulses appear to be arriving late.
2. Flat pulse shape. The pulse shape is fitted with a straight line connecting the first time sample and the last time sample. This is useful in distinguishing physical energy deposits from a flat-like pulse.
3. Spike-like pulse shape. The discriminant in this case is calculated by the RMS value of the lowest 8 time samples, divided by the pulse height summed over all time samples in the pulse.
4. Triangle-shaped pulse. This is done by fitting the pulse to a tent-like shape, where the pivot is located at the trigger time sample, with rising and falling lines on the left and the right respectively.

The goodness of fit for each noise-like hypothesis is divided by the  $\chi^2$  value for a fit to the ideal pulse shape to form discriminants. For the flat pulse hypothesis we have

$$\Lambda_{\text{flat}} \equiv \log \frac{\chi_{\text{flat}}^2}{\chi_{\text{ideal}}^2}, \quad (\text{B.3})$$

where  $\chi_{\text{flat}}^2$  is obtained from the fit to a flat pulse shape, and  $\chi_{\text{ideal}}^2$  is from the fit to the ideal pulse shape. Similarly we can form discriminants for spike-like pulse and triangle-shaped pulse hypotheses as  $\Lambda_{\text{spike}}$  and  $\Lambda_{\text{triangle}}$  respectively. The distributions of these discriminants is shown in Figure B.6 for a noise dataset and an orthogonal dataset where we don't have much noise contamination.

During the 50ns data taking in 2012, it was necessary to modify the fit to the ideal pulse, to include the contributions from out-of-time energy deposits. This is done by fitting to up to two pulses in the time window (HCAL occupancy is at the percent level with  $E_T > 1$  GeV during Run I); it's very unlikely we have three different energy deposits overlapping in the same time window in the same channel. The out of time pileup however also ruins the discriminant  $\Lambda_{\text{triangle}}$ . Therefore during the latter part of

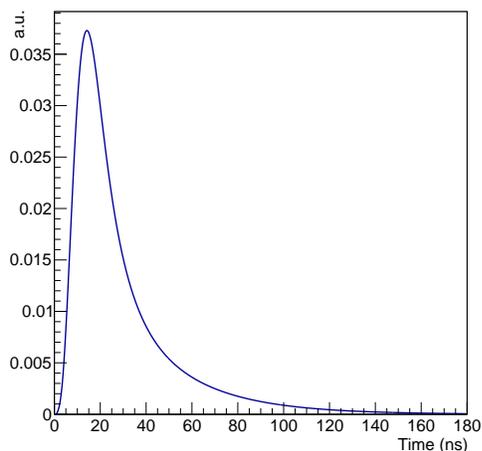


Figure B.5: Ideal HCAL pulse shape as measured and parametrized during the test beam measurements.

the data taking where there was a high level of pileup, this discriminant was tagged, but removed from consideration in the reconstruction.

The performance of the fit-based filters can be seen in Figure B.7, where we compare the missing energy distribution before and after pulse cleaning by the filter. The missing energy is reduced in all events after cleaning, indicating that the filter is very effectively cleaning out the noise. If there has been any mistagging of good pulses, we would see an increase of missing energy from time to time, and this was not observed, showing that the filter is safe, as well as efficient.

This filter was deployed in CMS data taking, and was turned on by default in the normal reconstruction during the second half of 2012.

## B.4 The 2012 Baseline Filter package

In addition to the fit-based filter, we also implemented a filter package for baseline filtering, which consists of two main filters:

1. HPD hit count filter. This filter counts how many channels are above an energy threshold (set to 1.5 GeV during the Run I data taking). It is sensitive to HPD discharge type noise, where we see the collective behavior of many channels in

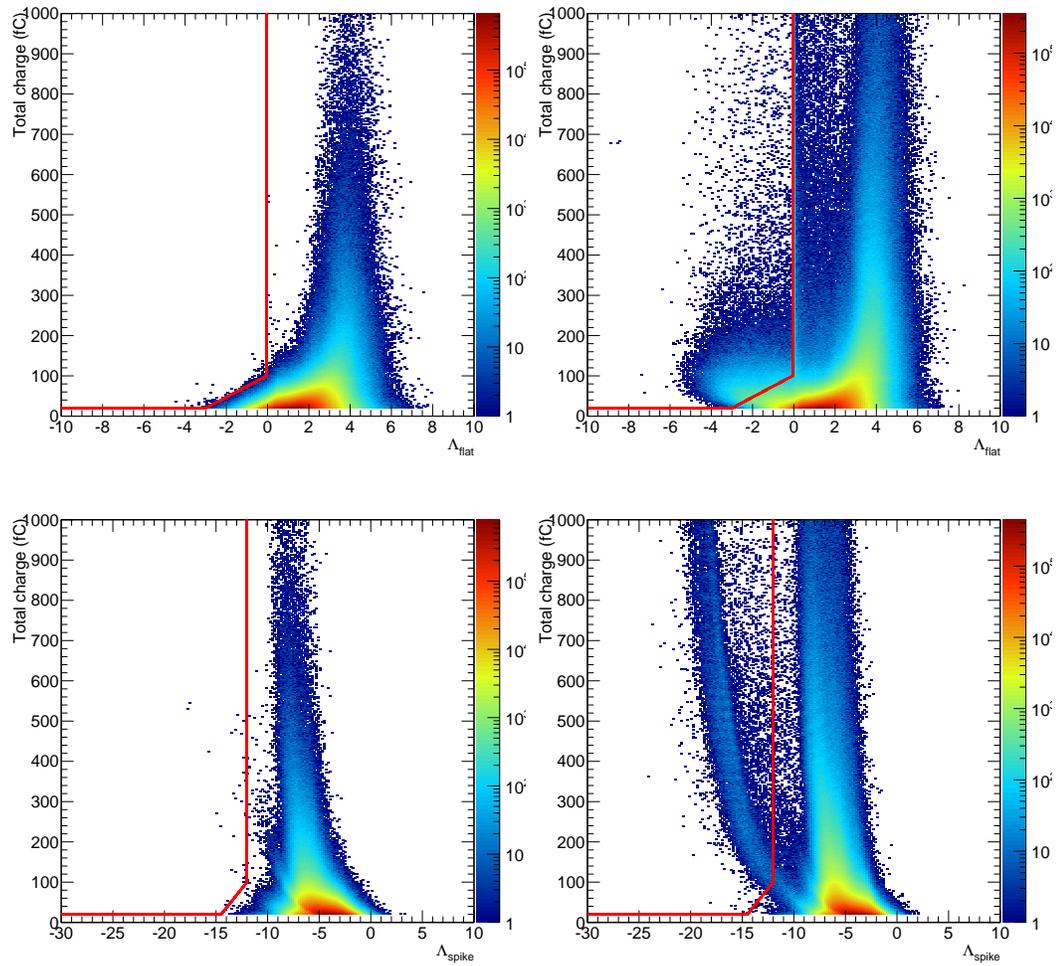


Figure B.6: Fit-based filter discriminants for a typical noisy dataset on the right, and a dataset with less noise on the left.

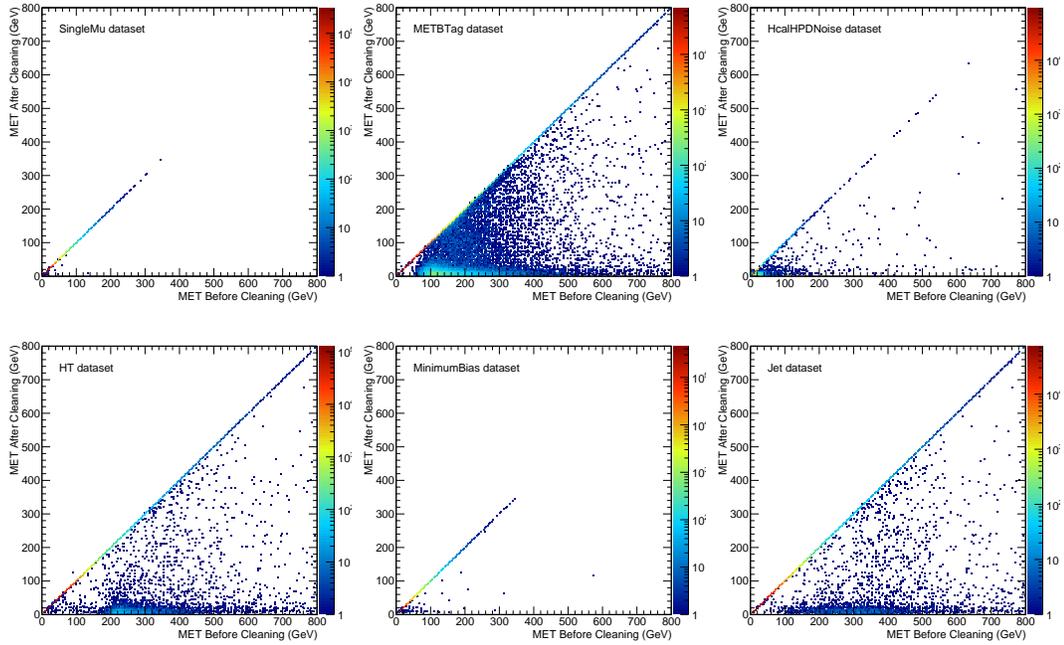


Figure B.7: The performance of the fit-based filter, comparing the missing transverse energy (MET) after the filter to before the filter, for several types of events.

an HPD. If the number of channels exceed a threshold (17 channels out of 18), the event is categorized as noise-like.

2. R45 filter. The R45 variable is defined as

$$R45 \equiv \frac{T_4 - T_5}{T_4 + T_5}, \quad (\text{B.4})$$

where  $T_4$  is the energy deposit during the time sample that contains the trigger, and  $T_5$  is the energy deposit for the subsequent time sample. This variable characterizes the energy sharing between the first 25ns and the next 25ns of the pulse. The collective pulse shape of all 72 channels in an RBX is calculated, and is used as a discriminant to filter out noise. If a whole RBX is noise-like, it creates a big-enough hole in the detector response that it is better to discard the whole event as noise-like, rather than just throwing out noise-like channels.

Distribution of variables used in the filter is shown in Figure B.8.

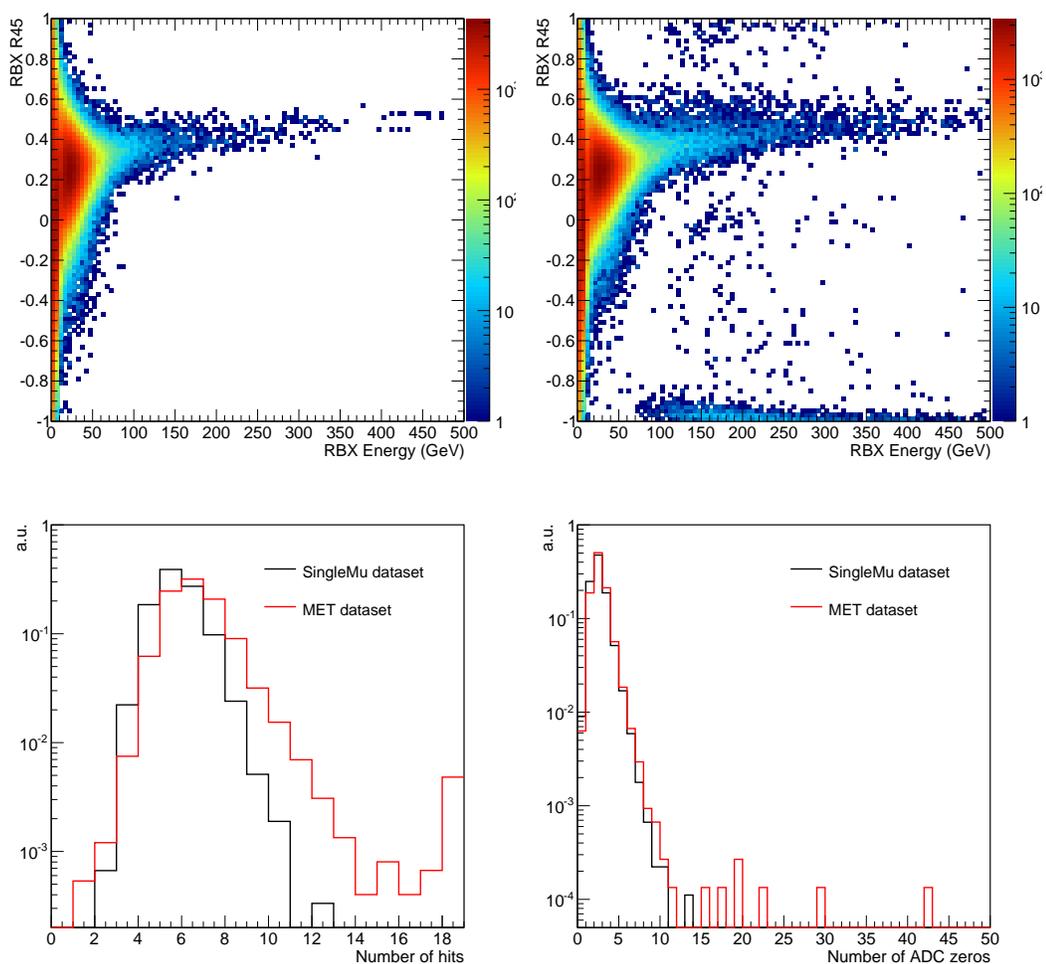


Figure B.8: Distributions of R45 variable used in the filter are shown on the top for a dataset containing noise (right), and a dataset with less noise (left). In the bottom plots the HPD count distribution between a noise-enriched dataset (red) and a clean dataset (black) is shown.

A relaxed version of this filter was implemented in the high level trigger (HLT) by gating the trigger paths where noise was a problem; for example the triggers on missing energy. This filter helped lower the energy threshold on the missing energy, enabling many physics searches.

## B.5 Development for the 25ns data

A number of developments have been made in preparation for the data taking for LHC Run II running starting 2015, where the time-spacing between bunches will be reduced to the design of 25ns. In the barrel and the endcap HCAL, where the pulses have a tail beyond 25ns, there will be a significant overlap of the pulses from different bunch crossings. The first task therefore was to study how we could mitigate the contribution from out-of-time pileup, using the fact that the HCAL pulse shape is stable for most channels.

We derived the average pulse shape using the data collected in 2012, since the ideal pulse shape did not describe the pulse shape observed in the data satisfactorily. The derivation of the shape was done by fitting an approximate parametrization to the ratio of the time samples in the pulse shape. An example of the extracted functions is shown in Figure B.9. We observe that by using the accurate pulse shapes derived from the data, the corrected energy is shifted back much closer to original value, as shown in Figure B.10.

This energy reconstruction was then used in noise rejection. Normal physical energy deposits have a characteristic long tail, while the spike-like pulse lacks the tail. Therefore if we calculate the corrected energy for the time samples in the tail, for a normal pulse without extra pileup, we expect the corrected values to be centered at zero with some spread due to fluctuation. For the spike-like pulses however, we will reconstruct large negative energies. Since any additional pileup will only increase or smear the measured energy, large negative values are indicative of noise. Figure B.11 shows the corrected energy distribution, in a number of datasets with different noise levels. The missing energy dataset contains enhanced noise due to trigger bias, while the single muon dataset has little noise, where we do not observe a tail in the distribution.

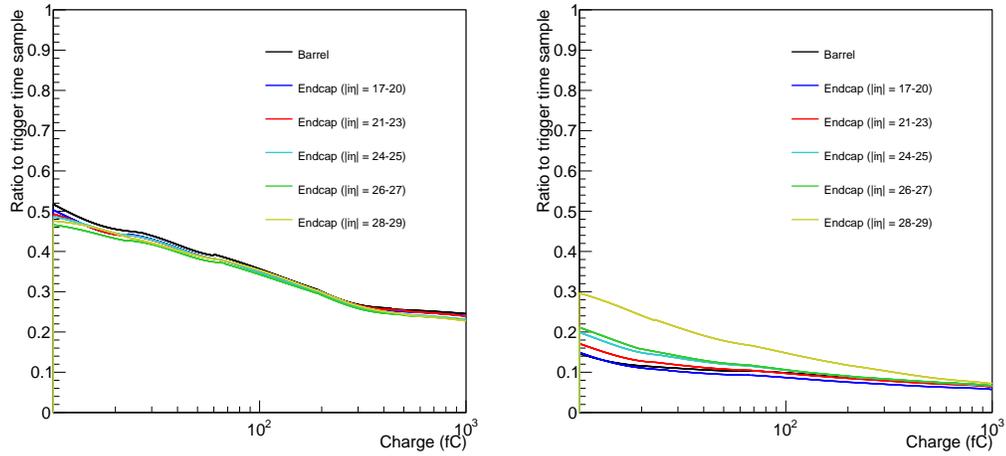


Figure B.9: An example of time sample ratio functions.

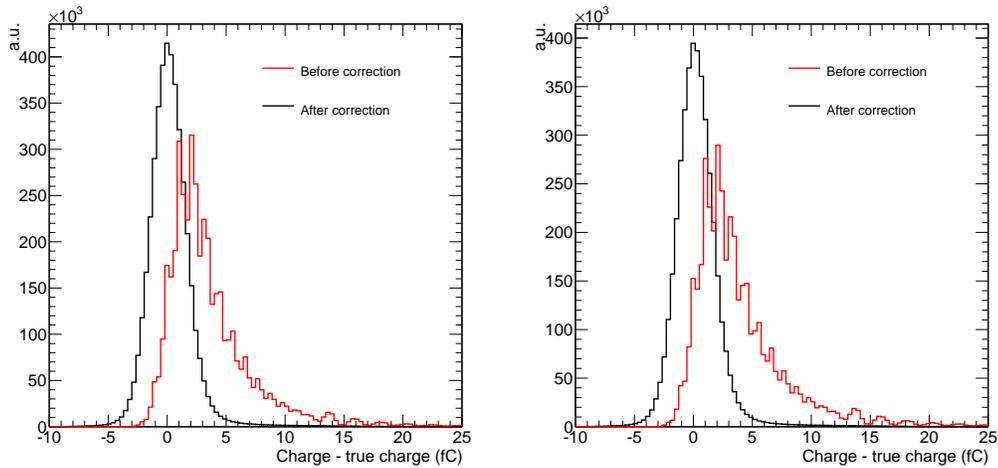


Figure B.10: Corrected energy distribution after the correction (black), compared to that before the correction (red). The mean of the energy from pileup after correction shifts back to zero.

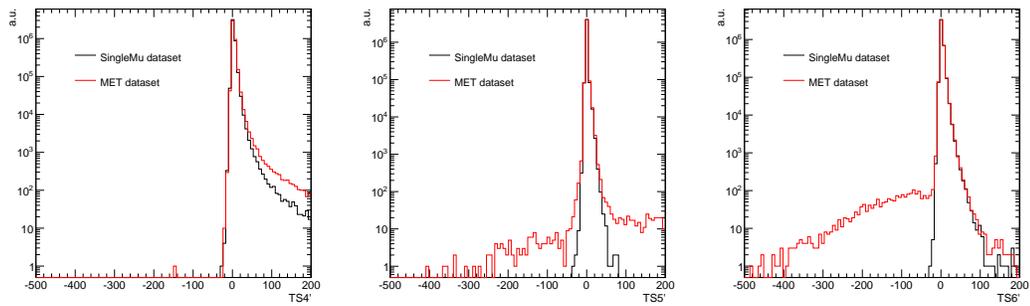


Figure B.11: Demonstration of the negative energy filter by looking at reconstructed energy in different time samples. In the SingleMu dataset, which is clean of noise, there is no significant negative energy tail. On the other hand, in the MET dataset enriched in noise, due to the presence of spike-like noises, a large negative energy tail is seen. A large negatively reconstructed energy indicate that the pulse shape differs greatly compared to the expectation, which is a sign of noise.

# Appendix C

## HCAL Test Beam Studies

### C.1 Experiment Setup

The HCAL test beam in 2009 was aimed at measuring the properties of the silicon photomultipliers (SiPM), as well as the full detector readout of the hadron calorimeter. The CMS HCAL has 16 longitudinal segments, each with a layer of brass absorber from 45mm to 75mm thick, interleaved with plastic scintillators. In the front there is steel supporting material with a layer of plastic scintillator between the steel and the brass absorbers behind it. Overall there are 17 channels per HCAL tower. In LHC Run I, light coming from all the layers was summed up by bundling the light fibers matching the bundle to a single readout pixel in a hybrid photomultiplier (HPD).

For the test beam we had 9 HCAL towers used in the analysis, arranged in a 3x3 grid. The towers corresponded to the barrel HCAL region with  $i\eta$  ranging from 7 to 9, and  $i\phi$  from 2 to 4 [44]. Three towers with  $i\phi = 3$  had 16-layers of individual readouts, while the rest had 1-channel summed readouts, as in the CMS Run I detector setup. There were three wire chambers (each with both  $x$  and  $y$  readout) upstream, with a set of triggering scintillators right before the last wire chamber. Behind the towers, there was an absorber with muon detectors behind it, acting as an event veto during the analysis. The setup is shown schematically in Figure C.1.

Data with different types of beam were taken, as summarized in Table C.1. Runs with a  $\mu^+$  beam were taken as a MIP calibration, and runs with  $\pi^+$  and  $\pi^-$  were used to measure the resolution, where we scanned through different energies. In each run

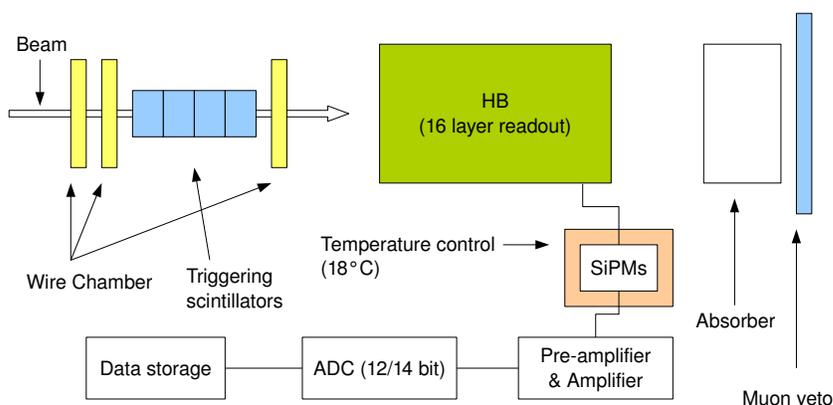


Figure C.1: Test beam layout and detector setup. The beam comes from the left, and passes through the wire chambers and triggerring scintillators before hitting the HCAL tower. There is a lead absorber and a muon chamber behind the HCAL tower.

two types of triggering were used. One was the usual trigger when there is a particle going through the scintillators, and the other was the “out-of-the-spill” trigger which fires when there is nothing happening. Events with the out-of-the-spill trigger can be used to perform pedestal calibrations. There were also other runs taken from other groups which we use to calibrate towers with the CMS Run I readout.

## C.2 Analysis

We started the analysis by doing an *in situ* alignment among the different wire chambers. This was done by looking at the measured location of the three chambers, exploiting the fact that each event should form a straight line, and lining up the different local chamber coordinates. The difference between the local coordinates of pairs of chambers are shown in Figure C.2 before and after alignment. After alignment the center of the beam spot is lined up with the local coordinate of the chambers. The alignment can be checked with  $\mu$ , data where the shower size is small, as shown

Table C.1: Summary of the data taken in the test beam. The first 6 runs are  $\mu$  calibration runs for towers with the same setup as CMS Run I, taken by other groups.

Run	Pre-amp	Type	E (GeV)	Statistics	Comments
106437	off	$\mu^-$	150	20000	Single tower only, aiming at (7, 2)
106436	off	$\mu^-$	150	20000	Single tower only, aiming at (8, 2)
106434	off	$\mu^-$	150	20000	Single tower only, aiming at (9, 2)
106427	off	$\mu^-$	150	20000	Single tower only, aiming at (7, 4)
106425	off	$\mu^-$	150	20000	Single tower only, aiming at (8, 4)
106426	off	$\mu^-$	150	20000	Single tower only, aiming at (9, 4)
108269	off	$\pi^-$	300	50000	Single tower only
108273	off	$\pi^-$	300	50000	Single tower only
108278	off	$\pi^-$	300	50000	Single tower only
108281	off	$\pi^-$	300	50000	Single tower only
108284	off	$\pi^-$	300	50000	Single tower only
108287	off	$\pi^-$	300	50000	Single tower only
108289	off	$\pi^-$	300	50000	Single tower only
108291	off	$\pi^-$	300	50000	Single tower only
108292	off	$\pi^-$	150	50000	Single tower only
108294	off	$\pi^+$	100	50000	Single tower only
108299	off	$\pi^+$	100	50000	Single tower only
108300	off	$\pi^+$	50	50000	Single tower only
108301	off	$\pi^+$	50	50000	Single tower only
108302	off	$\pi^+$	30	50000	Single tower only
108303	off	$\mu^+$	150	200000	Single tower only
108523	on	$\mu^+$	150	20000	Three tower only, aiming at (9, 3)
108525	on	$\mu^+$	150	20000	Three tower only, aiming at (7, 3)
108528	on	$\mu^+$	150	20300	Three tower only, aiming at (8, 3)
108532	on	$\pi^-$	300	50000	All 3x3 readout
108533	on	$\pi^-$	225	50000	All 3x3 readout
108535	on	$\pi^-$	150	50000	All 3x3 readout
108537	on	$\pi^-$	50	50000	All 3x3 readout
108538	on	$\pi^-$	30	50000	All 3x3 readout
108540	on	$\pi^-$	100	50000	All 3x3 readout
108541	on	$\pi^-$	100	30400	All 3x3 readout

in Figure C.3.

The pedestal fluctuates in a correlated fashion among the channels. A linear correlation coefficient between energy measured in different channels for pedestal events can be seen in the left panel of Figure C.4. We use the first two time samples as an event-by-event estimate of the pedestal, and verified that the correlation largely disappeared with the pedestal-subtracted measurement of energy, as can be seen in the right panel of Figure C.4.

Energy calibration is done with a  $\mu$  beam of 150 GeV. The measured energy spectrum is fitted by a Landau shape, convoluted with a Gaussian. The median value of the shape is taken as the minimum ionizing potential (MIP) response. The MIP of the whole tower corresponds to roughly 2.7 GeV of deposited energy [106]. Since there is no radioactive source calibration done for this setup, we do not have an absolute energy scale set. In order to set the absolute energy scale, we used 30 GeV  $\pi$  data, with a cut on the hit position so that the pion beam is centered on each tower.

An intercalibration using  $\pi^-$  data is carried out on top of the baseline  $\mu$  calibration. It is done by assigning a multiplicative constant to each layer as a correction factor, and by minimizing the sum of squares of the difference of the measured energy relative to the input energy. Note that this is good for the total energy measurement, and was not adequate to measure other quantities, such as the longitudinal shower profile.

An additional step is done for the data with the pre-amplifier on. During the data taking the bias voltage for the detector was set to a value that was too high. As a result we observed a clear saturation effect in the readout. This is shown in Figure C.6, where we compared the ADC value distribution between the data where the pre-amplifier was off, and where it was on. It was not feasible to remove events where the ADC values were observed to be saturated, as this would have removed a non-negligible fraction of the events and potentially bias our energy measurement. Instead, assuming that a larger ADC value indicates a larger energy deposit, even in the region close to saturation, we were able to “unwind” the distribution by comparing runs with the same  $\pi$  energy but with different pre-amplifier settings. This procedure

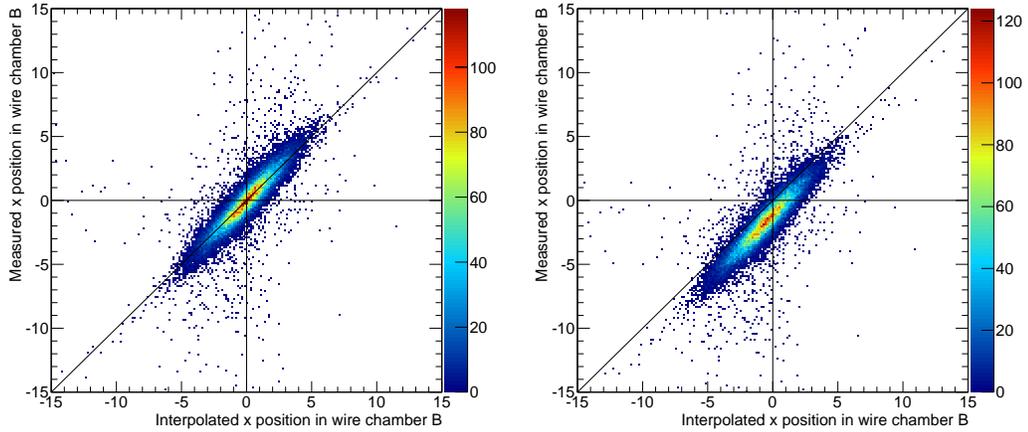


Figure C.2: Wire chamber alignment. Left: before alignment, and right: after alignment. The coordinates from the wire chambers are measured in units of mm. By this alignment procedure, we are able to align the wire chambers to an accuracy of  $O(1)$  mm.

was done on all channels, and one example is shown in Figure C.6.

By examining the transverse energy profile, it was possible to estimate the electron contamination in the pion beam. Since an electron shower is electromagnetic in nature and is very narrow (the Moliere radius  $R_M \sim 3$  cm [44]), it is well-contained in a single tower. The concentration of shower is shown in left panel of Figure C.7. The longitudinal profile of these events, shown in the right panel, where the energy is highly concentrated in the transverse direction, shows a higher energy deposit in the earlier layers, consistent with the expectations from an electron shower.

### C.3 Result

Once the calibrations and alignments were done, we examined different observables of interest. One result of interest is the  $\pi$  longitudinal profile. This is shown in Figure C.8. The average depth increases with increasing pion energy, with a characteristically long tail for the hadronic shower. When energy is large, the tail extends beyond the range of HCAL, indicating that HCAL does not fully contain all showers. The average observed depth is summarized in the right panel of Figure C.8. The de-

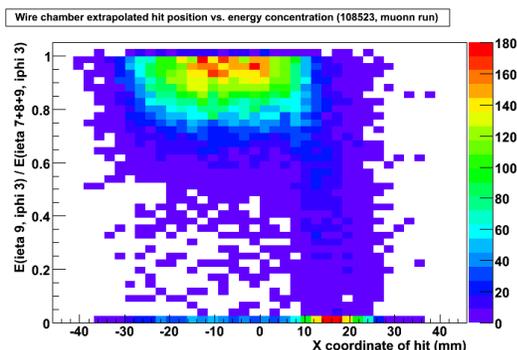


Figure C.3: The concentration of energy deposits in the middle tower, for a run with a muon beam. The  $y$ -axis shows the energy concentration in the middle tower, defined as the energy measured in the middle tower divided by the total measured energy in all three towers with 16-layer readout. It is plotted with respect to the  $x$ -coordinate of the extrapolated hit position. For muons with hit position above 8 mm, the majority of energy is deposited in the neighbor tower due to an misalignment of the HCAL modules with beam direction.

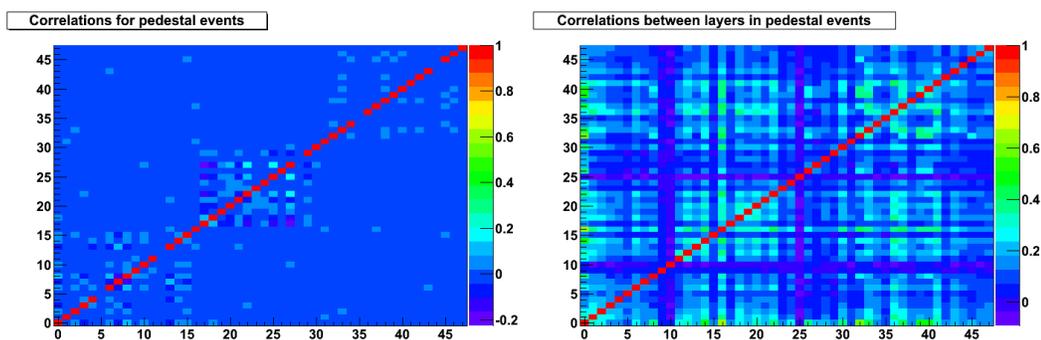


Figure C.4: The linear correlation coefficients of pedestal values in different channels in the three towers with 16-layer readout. The  $x$  and  $y$  axes are the channel index for each individual channel (48 in total). Before the pedestal subtraction procedure, a correlation is seen in the pedestal values for the different channels (left). The correlation is reduced a lot after the pedestal subtraction, showing a maximum of correlation coefficient of 0.1 (right)

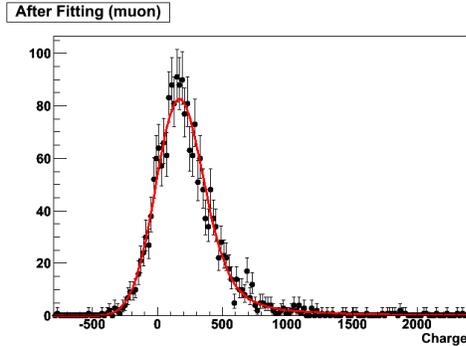


Figure C.5: The fit to muon calibration runs to extract the minimum ionizing potential (MIP). We use a Landau function convoluted with a Gaussian as a model to fit the spectrum. The median value of the function is then used as the measurement of the MIP for the channel.

viation from logarithmic dependence could be due to the worse containment of high energy showers.

The interaction length can be measured by considering the depth distribution of the first energy cluster. The distribution was shown to be a simple exponential distribution, as illustrated in Figure C.9. The exponential decay constant was determined to be 182.7 mm, consistent with what we expected for a tower composed of brass and scintillator layers of the given thickness.

Finally, the energy resolution was extracted and plotted as a function of the measured energy, as shown in Figure C.10. We were able to characterize the resolution curve as  $c_1 \oplus \frac{c_0}{\sqrt{E}}$ , with the fitted parameters:

$$c_0 = 0.896 \pm 0.026$$

$$c_1 = 0.046 \pm 0.005.$$

We carried out a similar analysis on the data from an earlier (2006) test beam run, where the towers were instrumented with standard CMS Run I setup. The fitted parameter we obtained from that analysis were:

$$c_0 = 0.936 \pm 0.029$$

$$c_1 = 0.105 \pm 0.002.$$

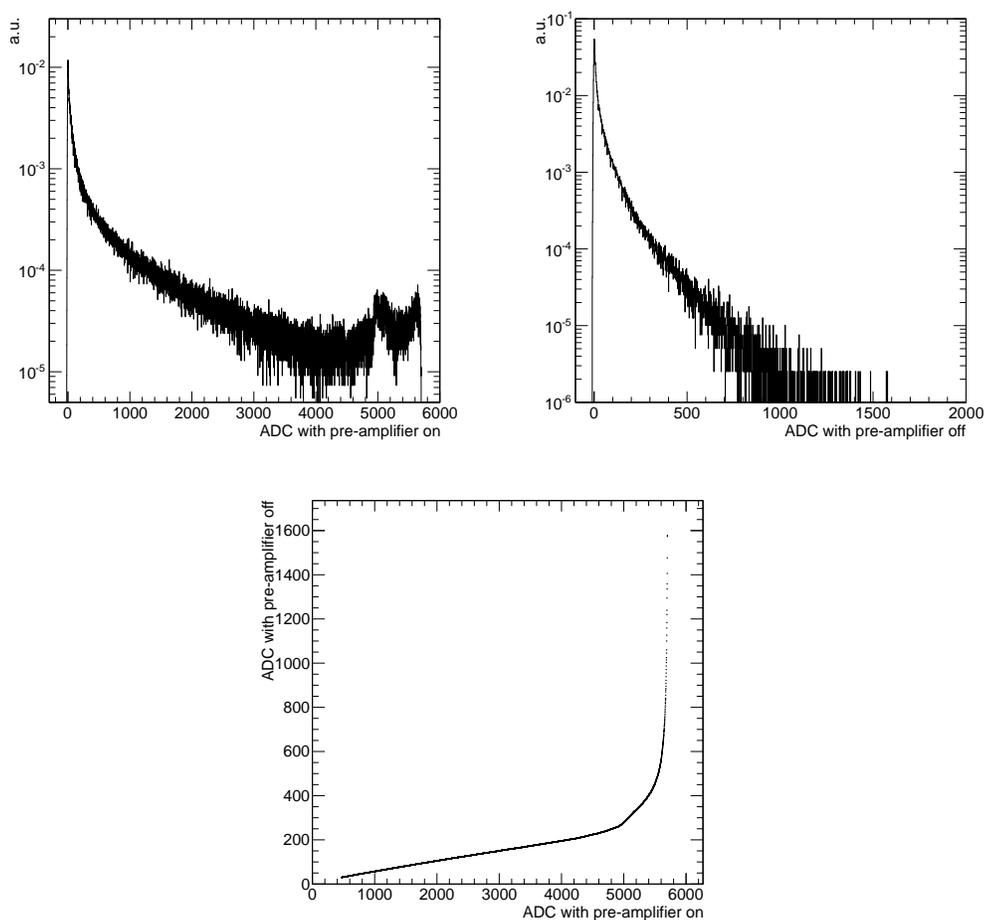


Figure C.6: The unwinding procedure. In the top row, the ADC value distributions of two pion runs with the same beam energy are shown. On the left is the distribution with pre-amplifier on, and on the right is the corresponding spectrum with the pre-amplifier off. We observe a clear saturation effect in the spectrum with pre-amplifier on. By comparing the two distributions and matching the corresponding ADC values with the same  $cdf$  value, we were able to obtain a mapping between pre-amplifier on ADC value and the corresponding pre-amplifier off ADC value. This mapping is then used to “unwind” the spectrum with pre-amplifier on for the energy measurement. One such mapping function is shown in the bottom row. There is a long linear regime when the ADC value is small, and the saturation effect is seen with larger ADC values.

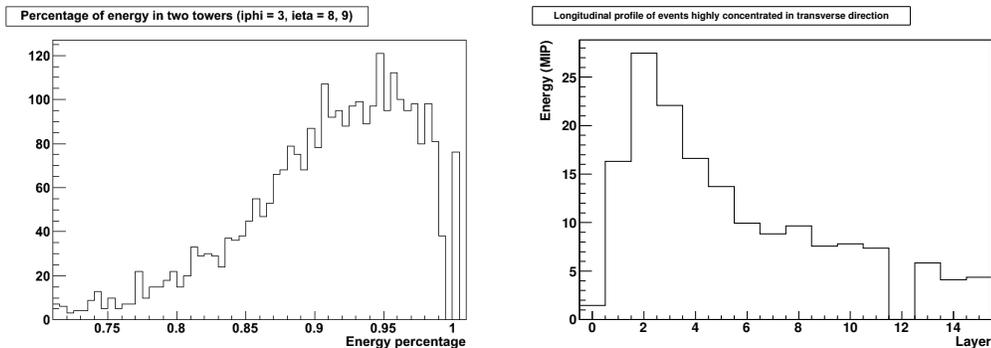


Figure C.7: Lepton contamination studies. On the left the energy concentration in the middle tower is shown for a run with pion beam. Since the hadron shower is not contained in a single HCAL tower, the energy concentration is in general not 100%. There is, however, a population with 100% energy concentration in the pion data. It is compatible with expectations from electrons. The longitudinal profile of the events with 100% energy concentration in the middle tower is shown on the right. We see a shallower profile compared to an average hadron shower.

An improvement in the HCAL resolution is thus observed with the SiPM + full layer readout.

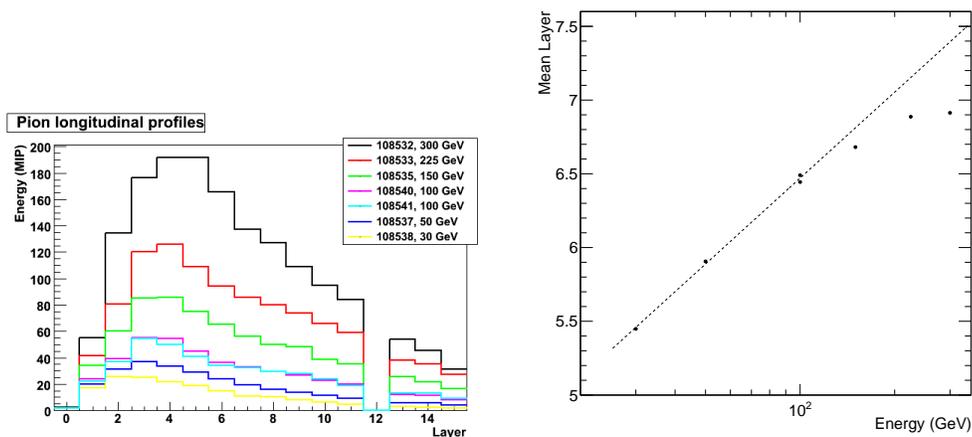


Figure C.8: The HCAL longitudinal profile. On the left the longitudinal profiles for pions with different energies are shown. There is a dead channel (layer 12) in the readout. With the increasing energy, position of maximum shower is deeper. The average depth of the shower is shown on the right as a function of energy. With pion beams of energy above 100 GeV, the average depth deviates from the expected logarithmic dependence, which could be due to the worse shower containment at large beam energies.

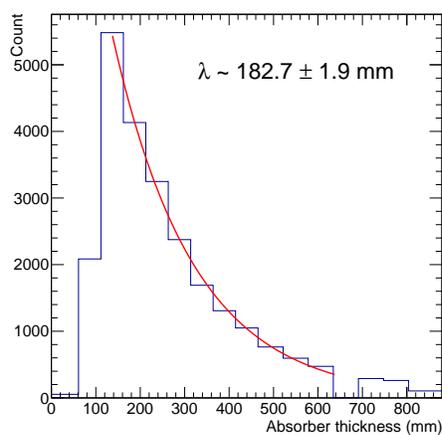


Figure C.9: Measurement of the HCAL interaction length. The position of the first energy deposit is plotted, showing an exponential dependence above the initial rise. We can fit the exponential constant and obtain a measurement of the interaction length in HCAL. It is compatible with the expectation of brass material interleaved with plastic scintillators.

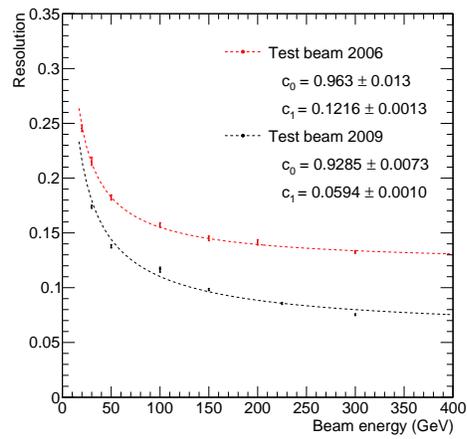


Figure C.10: The comparison of HCAL precision with full layer readout with an earlier test beam (2006) studies with CMS Run I setup. We observe an improvement in resolution with full layer readout.

# Appendix D

## Measurement of the Production Rate of Vector Boson in Association with Jets in CMS

### D.1 Description

We carried out a measurement of the  $Z$ +jet differential cross section using  $36 \text{ pb}^{-1}$  of  $pp$  collision data at 7 TeV collected by the CMS detector at the LHC [107]. I'm one of the main contributors to the  $Z \rightarrow \mu^+\mu^-$  channel. Differential cross section measurements up to  $Z$ +4 jets were carried out, as well as a test of Berends-Giele scaling.

Ever since its discovery in the 90's, the  $Z$  boson has been extensively studied at various colliders. At the LHC, it serves as a great candle for detector calibration, and also as a good probe of physics. Since there is always a jet balancing the  $Z$  boson, we can test the properties of QCD by looking at jets recoiling from the  $Z$  decay products. This study also provides important feedback to many other beyond standard model searches, since Drell-Yan events are among the most common background in such analyses.

While the  $Z$  boson has been extensively studied and is extremely well understood, its production in association with jets is more complicated. The  $Z$ +jets process involves many diagrams with different phenomenological consequences. Some of the lower order diagrams are shown in Figure D.1. A scaling behavior, known as Berend-

Giele scaling, was suggested to approximate the cross section scaling behavior, as a function of the number of associated jets. The scaling rule states that the cross section ratio between  $N$  jets and  $N+1$  jets is constant across the different numbers of jets:

$$R_N^Z = \frac{\sigma_{Z+(N+1\text{jets})}}{\sigma_{Z+(N\text{jets})}} \sim \alpha (\simeq \alpha_S), \quad (\text{D.1})$$

where  $\alpha$  is some constant smaller than 1 that could naively be expected to be similar to  $\alpha_S$ . This phenomenological model carries a simple intuitive interpretation: with each additional jet, there is in general one extra QCD vertex in the diagrams, and we therefore expect the cross section to scale down roughly by a factor of the coupling strength of the strong interaction  $\alpha_S$ .

This simple picture is however complicated by a few factors. Jets are complex objects in hadron colliders. While at parton level we can directly talk about diagrams and outgoing partons, the association between the reconstructed jets and the partons is less trivial. This association involves good modeling of both the jet fragmentation and hadronization processes, as well as the underlying soft processes accompanying the hard  $Z$ +jet process we want to study. For this reason we cannot measure jets with too small energies, lest we lose the power to interpret our result. The choice of the jet energy threshold also comes at the expense of discarding low energy jets, even if they come from our desired process.

Furthermore  $\alpha_S$  runs with the  $q^2$  scale, and it is most rapidly varying at low scales. While choosing the jet threshold fixes the scale somewhat, as a function of jet multiplicity this scale changes. Specifically for events with no associated jet (above threshold), this scale is much lower compared to the other jet multiplicities. Therefore we expect  $R_0^Z$  to be much different compared to the other ratios. A slightly modified scaling behavior has been suggested to incorporate the potential change of scale, and many other factors that might scale with number of jets:

$$R_N^Z \sim \alpha + \beta N, \quad (\text{D.2})$$

where  $\alpha$  and  $\beta$  are constants.

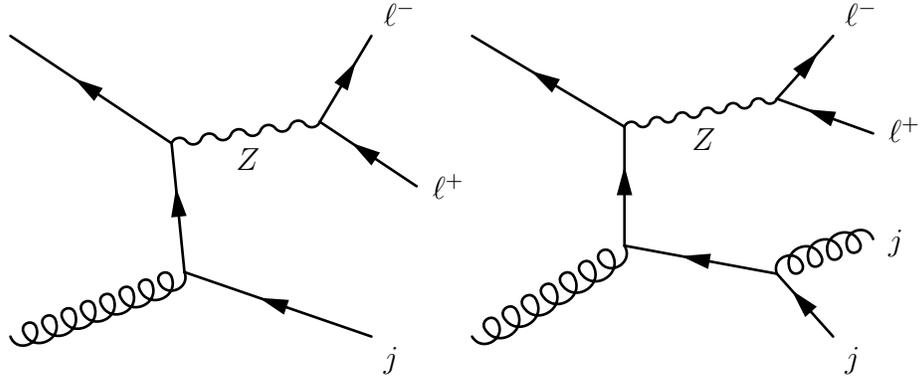


Figure D.1: Some of the lower order Feynman diagrams for  $Z$ +jets. With each additional jet, there is an extra QCD vertex, and we expect the rate to decrease by roughly a factor of  $\alpha_S$ .

The non-trivial dependence of  $R_N^Z$  on  $N$  makes it a great test of the standard model at this scale. By studying it carefully we have the opportunity to learn more about QCD, and it could also serve as a potential input to fragmentation and hadronization models.

## D.2 Methods

The  $Z$ +jets signal extraction procedure is straightforward. All di-lepton events whose invariant masses are in the proximity of the  $Z$  resonance mass ( $60 \text{ GeV} < m_{\ell\bar{\ell}} < 120 \text{ GeV}$ ) are selected. We select leading muons by requiring a combined relative isolation of 0.15 in a cone of size  $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.3$  around the muon. The combined relative isolation is defined as  $(p_T^{\text{tract}} + E_T^{\text{HCAL}} + E_T^{\text{ECAL}})/p_T^\mu$ , which characterizes the amount of activity around the target muon. We require the leading muon to have  $p_T > 20 \text{ GeV}$  and  $|\eta| < 2.1$ , with a high quality requirement on the muon track, and a transverse impact parameter  $d_{xy} < 0.2 \text{ cm}$ . Once a leading muon is found, we proceed to identify the second leading muon with  $p_T > 10 \text{ GeV}$  and  $|\eta| < 2.4$ .

Jets are created from the collection of particle candidates by use of the particle-flow algorithm, clustered with the anti- $k_T$  algorithm with a size parameter of  $R = 0.5$ . All jets are required to be within the tracker acceptance of  $|\eta| < 2.4$ , and a minimum

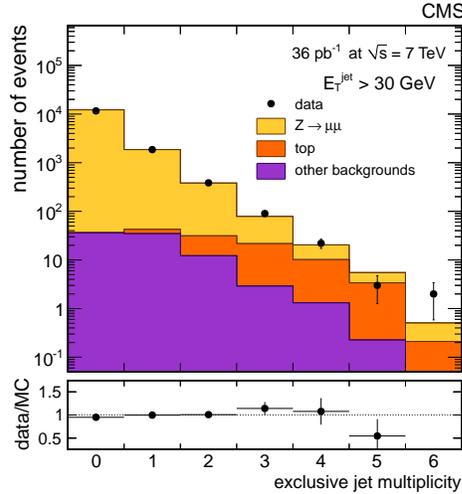


Figure D.2: The distribution of jet multiplicities in  $Z$ +jet events; comparing the data to a Monte Carlo simulation.

threshold of  $E_T > 30$  GeV is applied. We reduce the contribution from underlying pileup events by calculating the “average energy” in the non-clustered area, and subtracting it from each jet we consider.

The main background in this analysis is from accidental background, where two muons are produced in the event and happen to be in the invariant mass window we consider. At lower jet multiplicities, the background is mostly from (light) QCD processes, while for larger jet multiplicities the contributions from  $t\bar{t}$  processes dominate. The comparison of the data to the Monte Carlo simulation as a function of jet multiplicities is shown in Figure D.2.

Signal extraction is done via an unbinned maximum likelihood fit to the  $m_{\ell\bar{\ell}}$  distribution. We model the signal shape with a distribution with an asymmetric Gaussian core and a power law tail. Since the background shape is not peaking, we use an exponential distribution to model this contribution. All parameters are floated simultaneously for the lower jet multiplicity bins, while for the  $N \geq 4$ , due to insufficient statistics, the shape parameters are taken from the lower jet multiplicity bins. This procedure is used to extract the differential cross section.

Another fit is implemented which constrains yield in different jet multiplicity bins, according to the modified Berends-Giele scaling law (Equation D.2). We directly

extract the scaling parameters  $\alpha$  and  $\beta$  from the fit, correctly taking the correlations into account.

The effect on the number of jets from the jet energy resolution and the reconstruction efficiency is unfolded, while the detector acceptance is not unfolded, and we report the measurements in the fiducial phase space. The unfolding is done by using a migration matrix, which parametrizes the probability of observing the different jet multiplicities with a given number of out-going partons. This matrix is directly implemented in the fit, and we do not need any additional unfolding step.

### D.3 Results

The differential cross section as a function of the jet multiplicity is reported as a ratio relative to the inclusive  $Z$  cross section. By doing this, we cancel out many of the systematic effects that would otherwise be present in the analysis. The result is reported in the left panel of Figure D.3. It is compared with MADGRAPH [36], events showered by PYTHIA [73] with two different tunings, as well as with events generated by PYTHIA directly. We see that the MADGRAPH generator provides a good description of the data, and that  $R_N^Z$  is roughly constant, as expected from the scaling behavior.

Furthermore, we can extract the scaling parameters  $\alpha$  and  $\beta$  directly. The result is shown in the right panel of Figure D.3. The data is in good agreement with the expectation from the MADGRAPH generator.

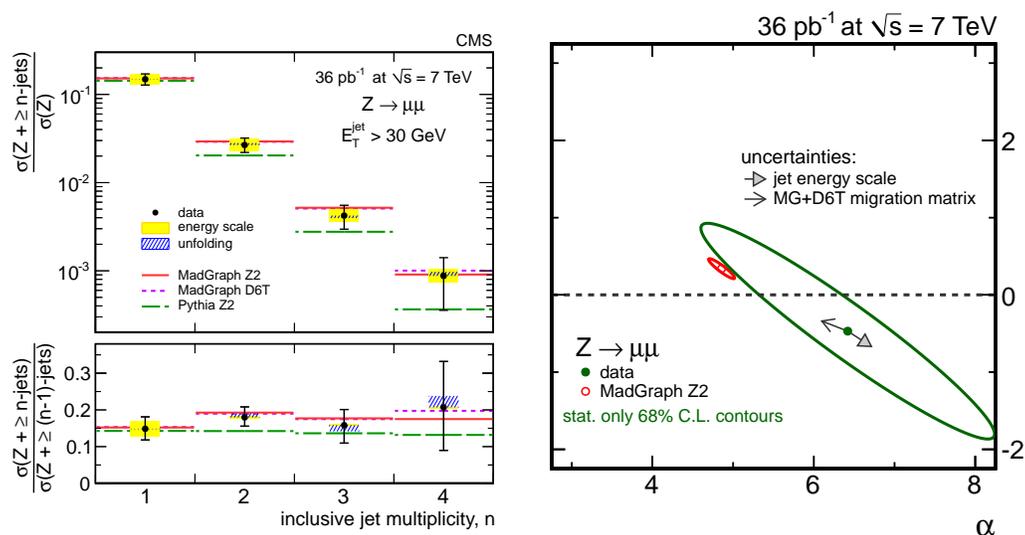


Figure D.3: The results on the  $Z (\rightarrow \mu^- \mu^+) + \text{jet}$  differential cross section measurement. On the left the differential cross section is shown, comparing to Monte Carlo simulations. The ratio  $R_N^Z$  is also shown in the bottom. On the right we report the measurement on the  $\alpha$  and  $\beta$  parameters (Equation D.2). The 1- $\sigma$  ellipse of data (green) is in agreement with the MADGRAPH simulation (red).

# Appendix E

## Leptoquark Search

A search [17] for third generation leptoquarks decaying to a  $b$  quark and a  $\tau$  neutrino was done using  $4.7 \text{ fb}^{-1}$  of  $pp$  collision data at 7 TeV, using the CMS detector at the LHC.

### E.1 Final State

Leptoquarks (LQ) are hypothetical particles that couple directly to a quark and a lepton. These particles arise in many grand unified theories [108], where some larger symmetry relating the quark sector and the lepton sector is assumed to exist. Similarly to the family structure of the fermions, LQs are traditionally classified into three families, where the couplings are much weaker across generations, partly due to the stringent experimental bounds on flavor-changing neutral currents.

There are different types of third generation LQ, each leading to different potential final states. Since they couple to a quark and a lepton, there are four possible decay products:  $b\nu_\tau$ ,  $b\tau$ ,  $t\nu_\tau$  and  $t\tau$ , corresponding to different types of LQ. Since they carry lepton number and baryon number, they are preferably produced in pairs in  $pp$  collisions, from gluon splitting. In this analysis, we perform a search for the  $b\bar{b}\nu_\tau\bar{\nu}_\tau$  final state, where each LQ decays to a  $b$  quark and a  $\tau$  neutrino. The Feynman diagram for this process is shown in Figure E.1. In the detector, we would see two high energy  $b$ -jets and a large missing energy.

Many SUSY scenarios also give rise to this same final state, for example as a result

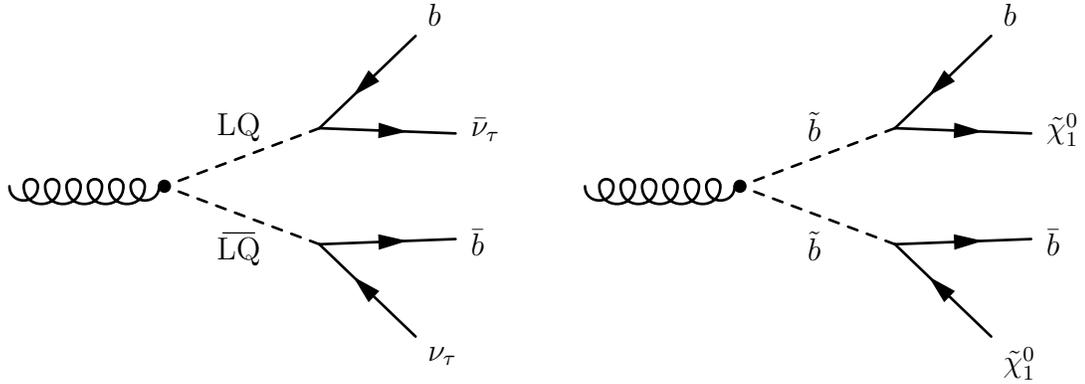


Figure E.1: Feynman diagram for LQ and sbottom

of  $\tilde{b}\bar{b}$  production. The result from this search can also be interpreted in many other contexts.

Previous bounds from the Tevatron limit this type of LQ to have a mass  $> 247$  GeV. We therefore started the search from slightly below this bound, to maintain some overlap and to check the Tevatron result.

## E.2 Razor Variables

The “razor variables” are used in this analysis to search for LQs. These are two complementary variables that characterize different aspects of a given event. The first variable,  $M_R$ , estimates the scale of the event, while the second one,  $R^2$ , describes how imbalanced the event is, in the energy transverse to the beam line. The shape of  $M_R$  peaks at its relevant scale for each component, and falls exponentially above the peak, mainly due to the PDF shape. The exponential slope in  $M_R$  changes linearly with the cut on  $R^2$ , allowing us to model the background in a simple way, as illustrated in Figure E.2.

The razor variables are derived through consideration of a set of frames referenced to the event topology we are searching for. In this search we assume that a pair of heavy particles are pair-produced, each of which decays into a jet and something invisible. In the rest frame of each heavy particle, the magnitude of the momentum of the jet has the same value. The decay products of each heavy particle are boosted

into the rest frame of the heavy system, and then boost again into the lab frame. Since the pair-produced particles are heavy, we make the assumption that they are produced roughly at rest, implying that the rest frame of the heavy particle system is related to lab frame by a simple boost along the beam direction. This can be approximated by boosting the visible system to rest, while assuming that the invisible system is similar to the visible system.

Next we approximate the boost from the rest frame of the heavy system to the respective rest frame of each of the heavy particles, by a transverse boost in the opposite direction. We can solve for the boost with which the jets will have the same magnitude of momentum as in the rest frames of the heavy particles, and use this magnitude as an estimator of the event scale. After all the algebra, we can write down the expression for  $M_R$  as

$$M_R = \sqrt{(E_1 + E_2)^2 - (p_1^z + p_2^z)^2}, \quad (\text{E.1})$$

where the subscript refers to the first and second jet.  $E$  is the energy of the jet, and  $p^z$  is the momentum component in the beam direction. Once we have all the boosts to approximate the different frames, we can write down a transverse version of the scale estimator (analogous to the transverse mass) as

$$M_R^T = \sqrt{\frac{1}{2}|\vec{M}|(|\vec{p}_1^T| + |\vec{p}_2^T|) - \frac{1}{2}\vec{M} \cdot (\vec{p}_1^T + \vec{p}_2^T)}, \quad (\text{E.2})$$

where  $\vec{M}$  is the missing transverse energy, and  $\vec{p}^T$  is the transverse momentum of the jet(s). Finally the second razor variable  $R^2$  is defined as

$$R^2 = \left( \frac{M_R^T}{M_R} \right)^2. \quad (\text{E.3})$$

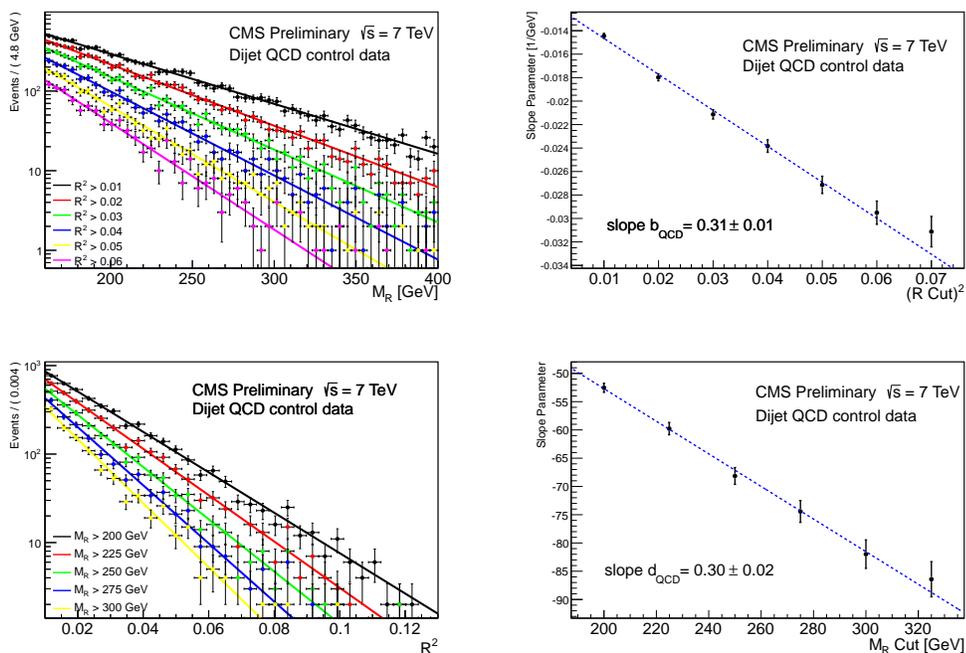


Figure E.2: Scaling behavior of the razor variables. The exponential tail in  $M_R$  drops faster with tighter cut on  $R^2$  variable (top-left), and the extracted slope is consistent with a linear function on  $R^2$  (top-right). The converse is also true: the tail in  $R^2$  drops faster with tighter cut on  $M_R$  (bottom-left), and the exponential constant is again modeled by a simple linear function of  $M_R$  (bottom-right). The slopes on the linear dependence of exponential constants in  $M_R$  and  $R^2$  agree with each other within numerical precision.

### E.3 Methods

$b$ -mesons have a characteristic long life time ( $c\tau \sim O(\text{cm})$ ). They decay in flight and produce a secondary vertex which is visible in the detector. In CMS many algorithms are developed which use different variables, based on two general ideas. One is the existence of a track with a large impact parameter (and subsequently a reconstructible secondary vertex), and the other is related to the mass of the combination of tracks within the jet cone. In most algorithms, the variables are thrown into a multivariate estimator to provide a final discriminant which increases the sensitivity over any given single variable. The efficiency of the  $b$ -tagging algorithm has a characteristic shape: the maximum efficiency is obtained for jets of around 100 GeV, and it drops mildly with increasing jet energy, due to the decrease in the ability to reconstruct a large impact parameter, as a result of the larger boost and thus the smaller spread of tracks. The efficiency drops for low energy jets below 100 GeV because of the smaller decay length  $c\gamma\tau$ , where the secondary vertex is closer to the primary collision point, and is thus harder to detect. The mistag rate on the other hand rises almost linearly as a function of the jet transverse momentum.

Most of the background relevant to this search comes from  $t\bar{t}$ +jet and  $V$ +jet processes. The decay products of a  $t$  quark provide a natural source of  $b$  jets, while the missing energy can be provided by a  $W$  decaying leptonically. For a  $V$ +jet event, there are contributions from  $V + b\bar{b}$  production, as well as from  $V$ +light jets, where the light jets are mistagged as  $b$  jets. The QCD multijet process has a small efficiency for this search, both because of the lower intrinsic event scale and the lack of large transverse momentum. However, due to the large cross section, we expect a small but non-negligible contribution from it.

Three baseline samples are defined in the analysis:

1. **MU**: events collected with the “muon razor triggers”, containing one loose muon with  $p_T > 20$  GeV,  $M_R > 400$  GeV and  $R^2 > 0.14$ .
2. **ELE**: events collected with the “electron razor triggers”, containing one loose electron with  $p_T > 20$  GeV,  $M_R > 400$  GeV and  $R^2 > 0.14$ .

3. **HAD**: events collected with the “hadronic razor triggers”, not satisfying any other requirements, and with  $M_R > 400$  GeV and  $R^2 > 0.2$ .

We also require that there are at least two jets above 60 GeV in each event, to ensure that the trigger is fully efficient for our selected events. In order to study and estimate the background contributions in the **HAD** sample, we treat muons and electrons in the **MU** and **ELE** samples as neutrinos, and recompute the  $R^2$  variable. Since there is a presence of leptons in the **MU** and **ELE** samples, they are free of the signal relevant to this analysis, which consists of final states with jets and missing energy (only). The events with leptons therefore they serve as good candidates to estimate the shape of the background contributions, without signal contamination.

The shape of  $M_R$  and  $R^2$  for different background components are derived from the data, as much as possible. The  $t\bar{t}$ +jet background shape is extracted from the **MU** selection, requiring a tight muon. This is then used in the full **MU** selection to extract the shape of the QCD multijet background. In the full **MU**, selection since we only require a very loose muon in the event, there is a significant contribution from the multijet background, which is not present in the sample where we require a tight muon. The shape of  $V$ +jets background is taken from the heavy-flavor enriched MADGRAPH samples. The extracted shape is shown in Figure E.3. We summarize the different sample selections in Table E.1.

The extraction of the shape of the background is done via an unbinned maximum likelihood fit. We split the data into different  $R^2$  bins, and model the shape for different  $R^2$  cuts as exponential functions. We constrain the slope parameter to be linear as a function of the  $R^2$  cut. All the  $R^2$  bins are fitted simultaneously.

Finally, once we have all the shapes extracted from the different control samples, the overall normalization of the background is measured in the region with  $R^2 < 0.25$  in the **HAD** sample, and is then extrapolated into the signal region with  $R^2 > 0.25$ . This procedure of background extraction is validated in the **ELE** sample, which is devoid of signal; and indeed we see good agreement between the predicted background and the data, as shown in Figure E.4. The expected contributions from each of the

Table E.1: Summary of samples used in the analysis

Sample	$R^2$ cut	Leptons	Comment
W/Z MC	$R^2 > 0.07$	tight $\mu$	shape of $V$ +jets
<b>MU</b>	$R^2 > 0.14$	tight $\mu$	shape of $t\bar{t}$ +jets
<b>MU</b>	$R^2 > 0.14$	loose $\mu$	shape of QCD multijets
<b>ELE</b>	$0.20 < R^2 < 0.25$	tight $e$	normalization measurement in <b>ELE</b> validation
<b>ELE</b>	$R^2 > 0.14$	loose $e$	signal-like region in <b>ELE</b> validation
<b>HAD</b>	$0.20 < R^2 < 0.25$	veto	normalization measurement
<b>HAD</b>	$R^2 > 0.14$	veto	signal region

processes are shown as the magenta dashed line ( $t\bar{t}$ ), green dotted line ( $W/Z$ ), and red dash-dotted line (QCD multijets). The total of the expected contribution from each process is shown as the blue solid line, in agreement in shape and yield compared to the data.

## E.4 Results

[109]

The predicted and the observed yields are summarized in Table E.2, with the distribution in the razor variables shown in Figure E.5. We did not see a significant excess in the signal region, and we proceed to set limits on the potential cross section.

The excluded cross section for a LQ pair production as a function of the mass is summarized in the left panel of Figure E.6. Since this type of LQ also has the chance of decaying into  $t\tau$ , the result is reinterpreted in terms of branching fraction to  $b\nu_\tau$ , as shown in the right panel of the plot.

We also reinterpreted the result in terms of simplified SUSY models, where sbottoms ( $\tilde{b}$ ) are pair-produced, each decaying into a  $b$ -jet and a neutralino ( $\tilde{\chi}_0$ ), which is the lowest supersymmetric particle (LSP) in this model. All other supersymmetric particles are assumed to be well above the scale we can probe. In such a simplified model, the event topology is similar to LQ, with the same visible final state. The interpreted results are shown in Figure E.7, where we give an exclusion limit in the  $\tilde{b} - \tilde{\chi}_0$  mass plane, greatly extending the previous result from the Tevatron.

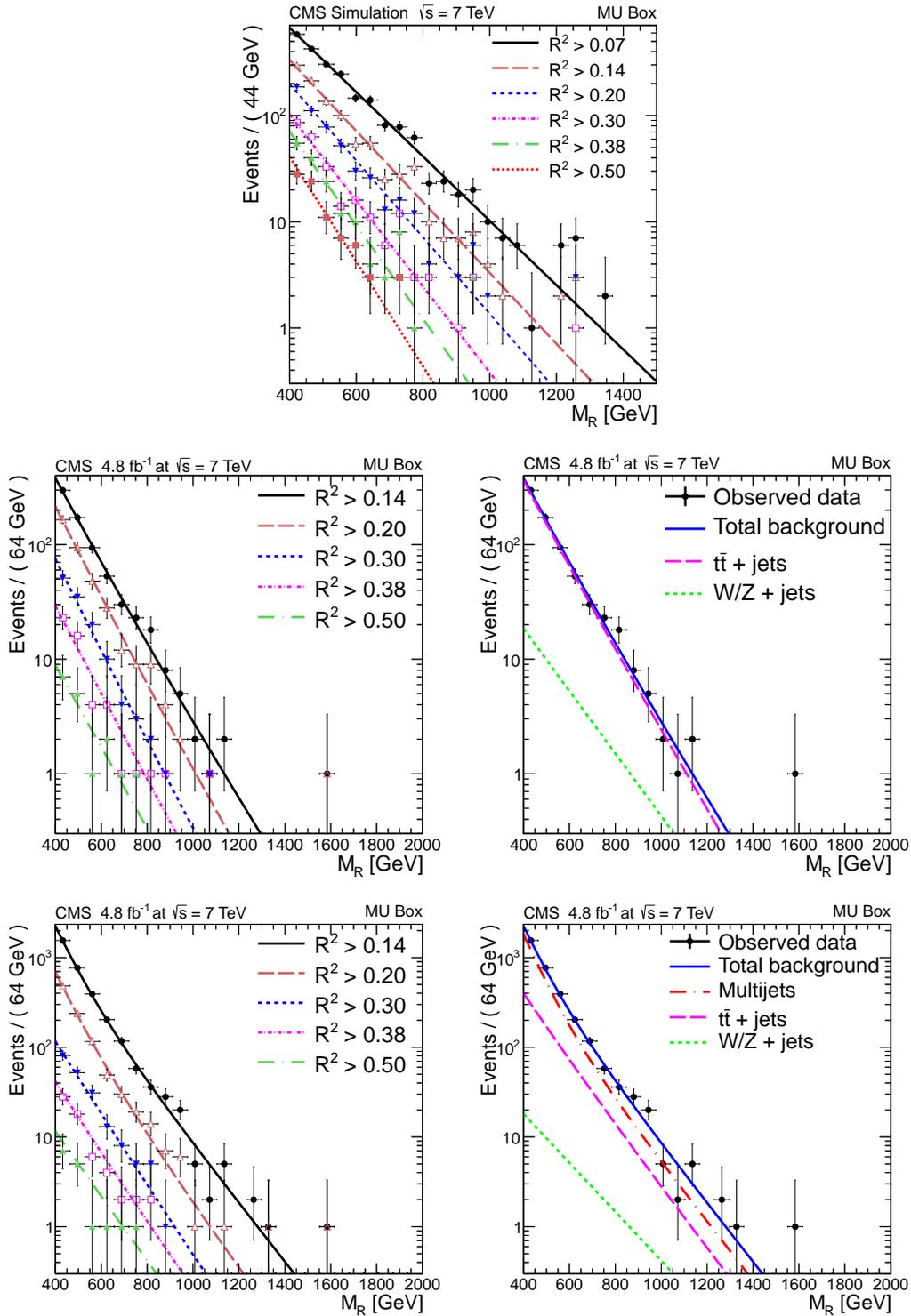


Figure E.3: Demonstration of the shape extraction procedure for different processes. On the top we fit  $V$ +jet Monte Carlo samples to extract the  $M_R$  shape as a function of different  $R^2$  cuts. This extracted shape is then used to in the fits to extract the shape of the  $t\bar{t}$  process, as shown in the middle row. In the selection with a tight muon, the data contains almost exclusively events from  $t\bar{t}$  and  $V$ +jet processes. By fitting to the data with a tight muon and fixing the shape of the  $V$ +jet sample, we are able to extract the shape of the  $t\bar{t}$  contribution. The extracted shapes for the  $t\bar{t}$  and  $V$ +jet are then used to fit in the data containing a loose muon, where the multijet contribution is greatly enriched, as shown in the bottom row. On the left the fits on  $M_R$  shape as a function of  $R^2$  cut is shown, and on the right only the inclusive  $R^2$

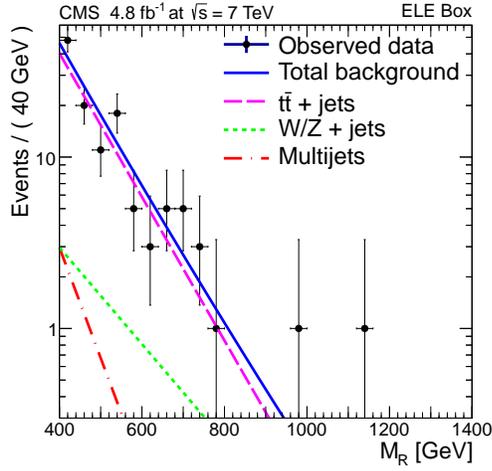


Figure E.4: Validation of the background prediction process using **ELE** sample. The extracted  $M_R$  shape from the muon samples are used to predict the  $M_R$  shape in a signal-like region in the **ELE** sample. Since we do not expect contribution from LQ signal in the **ELE** sample, it can be served as a good cross check of the background prediction process. We see good agreement for both the shape as a function of  $M_R$  and the overall yield.

Table E.2: Summary of the predicted and observed yields in the signal region. The uncertainty on the expected yield includes both statistical and systematic uncertainties.

$R^2$ cut	Expected yields	Observed yields
$> 0.25$	$119 \pm 23$	121
$0.25 - 0.30$	$51 \pm 17$	48
$0.30 - 0.35$	$30 \pm 10$	26
$0.35 - 0.38$	$9.9 \pm 5.2$	11
$0.38 - 0.42$	$11.5 \pm 5.0$	11
$> 0.42$	$16.8 \pm 4.8$	25

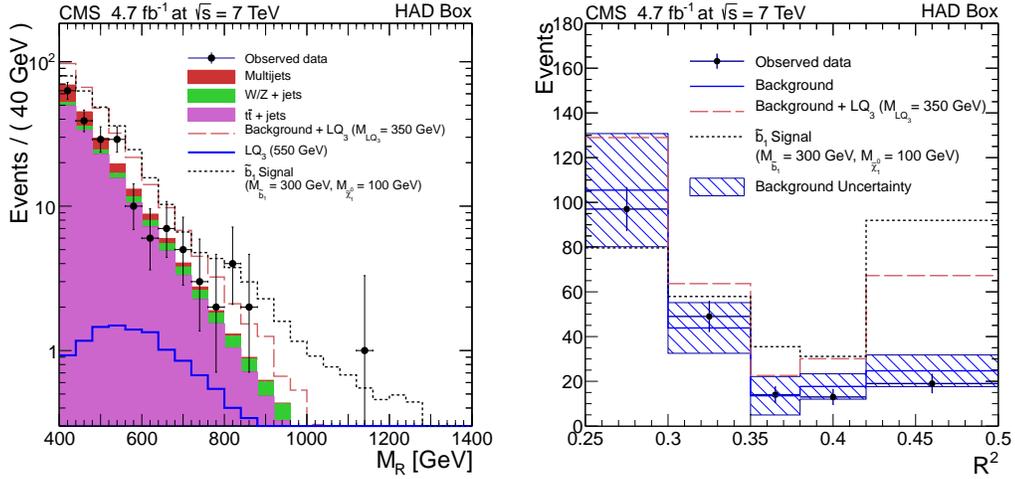


Figure E.5: Comparison of the predictions and the data in the signal region, as a function of the razor variables  $M_R$  and  $R^2$ . The total prediction is in good agreement with data in both variables, and we did not see any significant excess. The distribution for two excluded models are shown as the dotted and dashed lines.

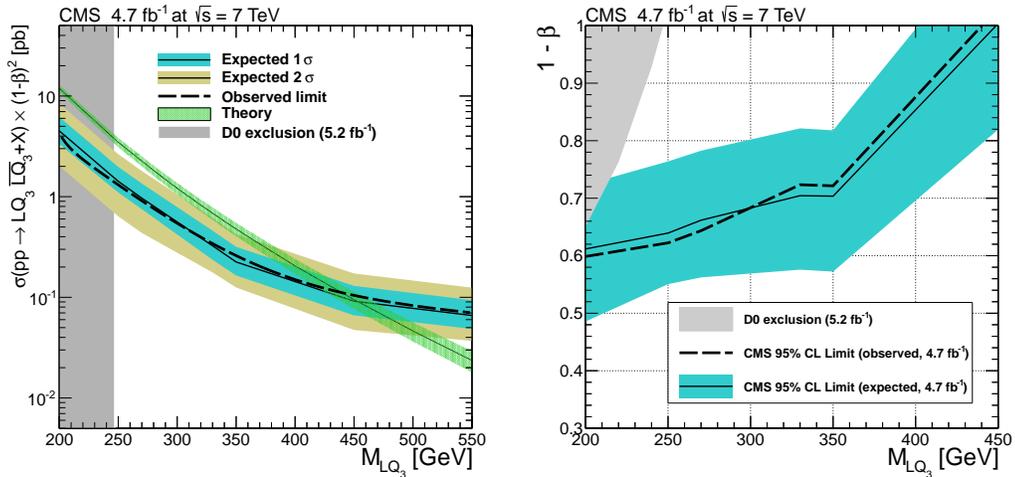


Figure E.6: LQ exclusion results from CMS, compared to the previous results from the D0 experiment at the Tevatron. The limit on production cross section (left) is shown as a function of the LQ mass assuming 100% branching fractions to  $b\nu_\tau$ . The limit on mass is reinterpreted as a function of different branching fractions (right). The result greatly extends the previous limit from the Tevatron, shown as the grey shaded area.

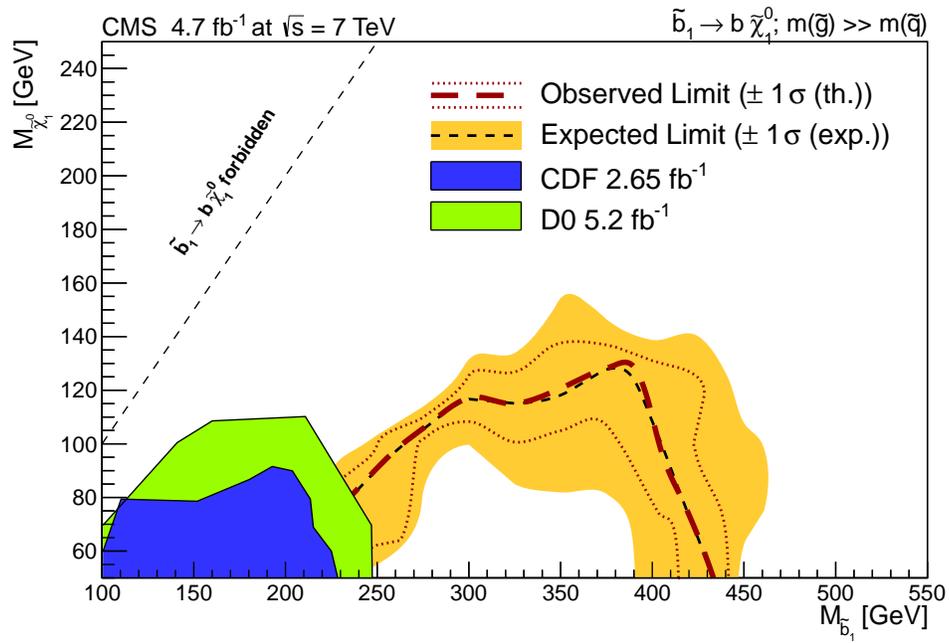


Figure E.7: Exclusion of a simplified model where sbottoms are pair-produced and each decays to a  $b$  quark and a neutralino ( $\tilde{\chi}^0$ , LSP) from CMS compared to the previous results from the CDF and D0 experiments at the Tevatron.

# Bibliography

- [1] J. Goldstone. Field theories with superconductor solutions. *Il Nuovo Cimento*, 19(1):154–164, 1961. URL: <http://dx.doi.org/10.1007/BF02812722>, doi:10.1007/BF02812722.
- [2] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. Broken symmetries. *Phys. Rev.*, 127:965–970, Aug 1962. URL: <http://link.aps.org/doi/10.1103/PhysRev.127.965>, doi:10.1103/PhysRev.127.965.
- [3] P. W. Anderson. Plasmons, gauge invariance, and mass. *Phys. Rev.*, 130:439–442, Apr 1963. URL: <http://link.aps.org/doi/10.1103/PhysRev.130.439>, doi:10.1103/PhysRev.130.439.
- [4] Y. Nambu and G. Jona-Lasinio. Dynamical model of elementary particles based on an analogy with superconductivity. i. *Phys. Rev.*, 122:345–358, Apr 1961. URL: <http://link.aps.org/doi/10.1103/PhysRev.122.345>, doi:10.1103/PhysRev.122.345.
- [5] Y. Nambu and G. Jona-Lasinio. Dynamical model of elementary particles based on an analogy with superconductivity. ii. *Phys. Rev.*, 124:246–254, Oct 1961. URL: <http://link.aps.org/doi/10.1103/PhysRev.124.246>, doi:10.1103/PhysRev.124.246.
- [6] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov

1967. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.19.1264>,  
doi:10.1103/PhysRevLett.19.1264.
- [7] Abdus Salam. Weak and Electromagnetic Interactions. *Conf.Proc.*, C680519:367–377, 1968.
- [8] P.W. Higgs. Broken symmetries, massless particles and gauge fields. *Physics Letters*, 12(2):132 – 133, 1964. URL: <http://www.sciencedirect.com/science/article/pii/0031916364911369>,  
doi:[http://dx.doi.org/10.1016/0031-9163\(64\)91136-9](http://dx.doi.org/10.1016/0031-9163(64)91136-9).
- [9] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global Conservation Laws and Massless Particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.13.585>,  
doi:10.1103/PhysRevLett.13.585.
- [10] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.13.508>,  
doi:10.1103/PhysRevLett.13.508.
- [11] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.13.321>,  
doi:10.1103/PhysRevLett.13.321.
- [12] C. N. Yang and R. L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.*, 96:191–195, Oct 1954. URL: <http://link.aps.org/doi/10.1103/PhysRev.96.191>,  
doi:10.1103/PhysRev.96.191.
- [13] Search for the standard model higgs boson at {LEP}. *Physics Letters B*, 565(0):61 – 75, 2003. URL: [http:](http://)



[//www.sciencedirect.com/science/article/pii/0029558261904692](http://www.sciencedirect.com/science/article/pii/0029558261904692),  
doi:[http://dx.doi.org/10.1016/0029-5582\(61\)90469-2](http://dx.doi.org/10.1016/0029-5582(61)90469-2).

- [21] T. W. B. Kibble. Symmetry Breaking in Non-Abelian Gauge Theories. *Phys. Rev.*, 155:1554–1561, Mar 1967. URL:  
<http://link.aps.org/doi/10.1103/PhysRev.155.1554>,  
doi:10.1103/PhysRev.155.1554.
- [22] Peter W. Higgs. Spontaneous Symmetry Breakdown without Massless Bosons. *Phys. Rev.*, 145:1156–1163, May 1966. URL:  
<http://link.aps.org/doi/10.1103/PhysRev.145.1156>,  
doi:10.1103/PhysRev.145.1156.
- [23] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental Test of Parity Conservation in Beta Decay. *Phys. Rev.*, 105:1413–1415, Feb 1957. URL:  
<http://link.aps.org/doi/10.1103/PhysRev.105.1413>,  
doi:10.1103/PhysRev.105.1413.
- [24] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the  $2\pi$  Decay of the  $K_2^0$  Meson. *Phys. Rev. Lett.*, 13:138–140, Jul 1964. URL:  
<http://link.aps.org/doi/10.1103/PhysRevLett.13.138>,  
doi:10.1103/PhysRevLett.13.138.
- [25] A. Alavi-Harati, I. F. Albuquerque, T. Alexopoulos, M. Arenton, K. Arisaka, S. Averitte, A. R. Barker, L. Bellantoni, A. Bellavance, J. Belz, R. Ben-David, D. R. Bergman, E. Blucher, G. J. Bock, C. Bown, S. Bright, E. Cheu, S. Childress, R. Coleman, M. D. Corcoran, G. Corti, B. Cox, M. B. Crisler, A. R. Erwin, R. Ford, A. Glazov, A. Golossanov, G. Graham, J. Graham, K. Hagan, E. Halkiadakis, K. Hanagaki, S. Hidaka, Y. B. Hsiung, V. Jejer, J. Jennings, D. A. Jensen, R. Kessler, H. G. E. Kobrak, J. LaDue, A. Lath, A. Ledovskoy, P. L. McBride, A. P. McManus, P. Mikelsons, E. Monnier, T. Nakaya, U. Nauenberg, K. S. Nelson, H. Nguyen, V. O'Dell, M. Pang,

R. Pordes, V. Prasad, C. Qiao, B. Quinn, E. J. Ramberg, R. E. Ray, A. Roodman, M. Sadamoto, S. Schnetzer, K. Senyo, P. Shanahan, P. S. Shawhan, W. Slater, N. Solomey, S. V. Somalwar, R. L. Stone, I. Suzuki, E. C. Swallow, R. A. Swanson, S. A. Taegar, R. J. Tesarek, G. B. Thomson, P. A. Toale, A. Tripathi, R. Tschirhart, Y. W. Wah, J. Wang, H. B. White, J. Whitmore, B. Winstein, R. Winston, J.-Y. Wu, T. Yamanaka, and E. D. Zimmerman. Observation of direct cp violation in  $k_{s,l} \rightarrow \pi\pi$  decays. *Phys. Rev. Lett.*, 83:22–27, Jul 1999. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.83.22>, doi:10.1103/PhysRevLett.83.22.

- [26] V. Fanti, A. Lai, D. Marras, L. Musa, A.J. Bevan, T.J. Gershon, B. Hay, R.W. Moore, K.N. Moore, D.J. Munday, M.D. Needham, M.A. Parker, S.F. Takach, T.O. White, S.A. Wotton, G. Barr, H. Blmer, G. Bocquet, J. Bremer, A. Ceccucci, J. Cogan, D. Cundy, N. Doble, W. Funk, L. Gatignon, A. Gianoli, A. Gonidec, G. Govi, P. Grafstrm, G. Kessler, W. Kubischta, A. Lacourt, S. Luitz, J.P. Matheys, A. Norton, S. Palestini, B. Panzer-Steindel, B. Peyaud, D. Schinzel, H. Taureg, M. Velasco, O. Vossnack, H. Wahl, G. Wirrer, A. Gaponenko, V. Kekelidze, D. Madigojine, A. Mestvirishvili, Yu. Potrebenikov, G. Tatishvili, A. Tkatchev, A. Zinchenko, L. Bertolotto, O. Boyle, I.G. Knowles, V.J. Martin, H.L.C. Parsons, K.J. Peach, C. Talamonti, M. Contalbrigo, P. Dalpiaz, J. Duclos, A. Formica, P.L. Frabetti, M. Martini, F. Petrucci, M. Savri, A. Bizzeti, M. Calvetti, G. Collazuol, G. Graziani, E. Iacopini, M. Lenti, A. Michetti, H.G. Becker, P. Buchholz, D. Coward, C. Ebersberger, H. Fox, A. Kalter, K. Kleinknecht, U. Koch, L. Kpke, B. Renk, J. Scheidt, J. Schmidt, V. Schnharting, Y. Schu, R. Wilhelm, M. Wittgen, J.C. Chollet, S. Crp, L. Fayard, L. Iconomidou-Fayard, J. Ocariz, G. Unal, D. Vattolo, I. Wingerter-Seez, G. Anzivino, F. Bordacchini, P. Cenci, P. Lubrano, A. Nappi, M. Pepe, M. Punturo, L. Bertanza, A. Bigi, P. Calafiura, R. Carosi, R. Casali, C. Cerri,

- M. Cirilli, F. Costantini, R. Fantechi, S. Giudici, B. Gorini, I. Mannelli, V. Marzulli, G. Pierazzini, F. Raffaelli, M. Sozzi, J.B. Cheze, M. De Beer, P. Debu, R. Granier de Cassagnac, P. Hristov, E. Mazzucato, S. Schanne, R. Turlay, B. Vallage, I. Augustin, M. Bender, M. Holder, M. Ziolkowski, R. Arcidiacono, C. Biino, R. Cester, F. Marchetto, E. Menichetti, N. Pastrone, J. Nassalski, E. Rondio, M. Szleper, W. Wislicki, S. Wronka, H. Dibon, G. Fischer, M. Jeitler, M. Markytan, I. Mikulec, G. Neuhofer, M. Pernicka, and A. Taurok. A new measurement of direct cp violation in two pion decays of the neutral kaon. *Physics Letters B*, 465(14):335 – 348, 1999. URL: <http://www.sciencedirect.com/science/article/pii/S0370269399010308>, doi:[http://dx.doi.org/10.1016/S0370-2693\(99\)01030-8](http://dx.doi.org/10.1016/S0370-2693(99)01030-8).
- [27] Andrei D Sakharov. Violation of CP in variance, C asymmetry, and baryon asymmetry of the universe. *Soviet Physics Uspekhi*, 34(5):392, 1991. URL: <http://stacks.iop.org/0038-5670/34/i=5/a=A08>.
- [28] J. P. et al Lees. Observation of time-reversal violation in the  $B^0$  meson system. *Phys. Rev. Lett.*, 109:211801, Nov 2012. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.109.211801>, doi:10.1103/PhysRevLett.109.211801.
- [29] Yi Chen, Nhan Tran, and Roberto Vega-Morales. Scrutinizing the Higgs signal and background in the 2e2 golden channel. *Journal of High Energy Physics*, 2013(1), 2013. URL: <http://dx.doi.org/10.1007/JHEP01%282013%29182>, doi:10.1007/JHEP01(2013)182.
- [30] Yanyan Gao, Andrei V. Gritsan, Zijin Guo, Kirill Melnikov, Markus Schulze, and Nhan V. Tran. Spin determination of single-produced resonances at hadron colliders. *Phys. Rev. D*, 81:075022, Apr 2010. URL: <http://link.aps.org/doi/10.1103/PhysRevD.81.075022>, doi:10.1103/PhysRevD.81.075022.
- [31] Sara Bolognesi, Yanyan Gao, Andrei V. Gritsan, Kirill Melnikov, Markus

- Schulze, Nhan V. Tran, and Andrew Whitbeck. Spin and parity of a single-produced resonance at the LHC. *Phys. Rev. D*, 86:095031, Nov 2012. URL: <http://link.aps.org/doi/10.1103/PhysRevD.86.095031>, doi:10.1103/PhysRevD.86.095031.
- [32] Yi Chen, Adam Falkowski, Ian Low, and Roberto Vega-Morales. New Observables for CP Violation in Higgs Decays. 2014. [arXiv:1405.6723](https://arxiv.org/abs/1405.6723).
- [33] Yi Chen and Roberto Vega-Morales. Extracting effective Higgs couplings in the golden channel. *Journal of High Energy Physics*, 2014(4), 2014. URL: <http://dx.doi.org/10.1007/JHEP04%282014%29057>, doi:10.1007/JHEP04(2014)057.
- [34] Jonathan Pumplin, Daniel Robert Stump, Joey Huston, Hung-Liang Lai, Pavel Nadolsky, and Wu-Ki Tung. New Generation of Parton Distributions with Uncertainties from Global QCD Analysis. *Journal of High Energy Physics*, 2002(07):012, 2002. URL: <http://stacks.iop.org/1126-6708/2002/i=07/a=012>.
- [35] S. Chatrchyan et al. Measurement of the properties of a Higgs boson in the four-lepton final state. *Phys. Rev. D*, 89:092007, May 2014. URL: <http://link.aps.org/doi/10.1103/PhysRevD.89.092007>, doi:10.1103/PhysRevD.89.092007.
- [36] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7), 2014. URL: <http://dx.doi.org/10.1007/JHEP07%282014%29079>, doi:10.1007/JHEP07(2014)079.
- [37] Lyndon Evans and Philip Bryant. LHC Machine. *Journal of Instrumentation*,

- 3(08):S08001, 2008. URL:  
<http://stacks.iop.org/1748-0221/3/i=08/a=S08001>.
- [38] CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004, 2008. URL:  
<http://stacks.iop.org/1748-0221/3/i=08/a=S08004>.
- [39] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08003, 2008. URL:  
<http://stacks.iop.org/1748-0221/3/i=08/a=S08003>.
- [40] LHCb Collaboration. The LHCb Detector at the LHC. *Journal of Instrumentation*, 3(08):S08005, 2008. URL:  
<http://stacks.iop.org/1748-0221/3/i=08/a=S08005>.
- [41] ALICE Collaboration. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002, 2008. URL:  
<http://stacks.iop.org/1748-0221/3/i=08/a=S08002>.
- [42] Henrik Bohr and H.B. Nielsen. Hadron production from a boiling quark soup: A thermodynamical quark model predicting particle ratios in hadronic collisions. *Nuclear Physics B*, 128(2):275 – 293, 1977. URL: <http://www.sciencedirect.com/science/article/pii/0550321377900323>,  
[doi:http://dx.doi.org/10.1016/0550-3213\(77\)90032-3](http://dx.doi.org/10.1016/0550-3213(77)90032-3).
- [43] CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update. Technical Report CMS-PAS-LUM-13-001, CERN, Geneva, 2013.
- [44] CMS Collaboration. CMS Physics Technical Design Report, Volume II: Physics Performance. *Journal of Physics G: Nuclear and Particle Physics*, 34(6):995, 2007. URL: <http://stacks.iop.org/0954-3899/34/i=6/a=S01>.
- [45] The CASTOR Collaboration. Formation of Centauro in Pb+Pb collisions at the LHC and their detection with the CASTOR detector of CMS. *Nuclear*

- Physics B - Proceedings Supplements*, 122(0):205 – 208, 2003. Proceedings of the 12th International Symposium on Very High Energy Cosmic Ray Interactions. URL: <http://www.sciencedirect.com/science/article/pii/S0920563203803791>, doi:[http://dx.doi.org/10.1016/S0920-5632\(03\)80379-1](http://dx.doi.org/10.1016/S0920-5632(03)80379-1).
- [46] TOTEM Collaboration. The TOTEM Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08007, 2008. URL: <http://stacks.iop.org/1748-0221/3/i=08/a=S08007>.
- [47] K.A. Olive and Particle Data Group. Review of Particle Physics. *Chinese Physics C*, 38(9):090001, 2014. URL: <http://stacks.iop.org/1674-1137/38/i=9/a=090001>.
- [48] S. Abdullin et al. Design, performance, and calibration of the CMS hadron-outercalorimeter. *The European Physical Journal C*, 57(3):653–663, 2008. URL: <http://dx.doi.org/10.1140/epjc/s10052-008-0756-6>, doi:10.1140/epjc/s10052-008-0756-6.
- [49] CMS Collaboration. *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.
- [50] *The CMS muon project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.
- [51] S et al Costantini. Uniformity and Stability of the CMS RPC Detector at the LHC. Technical Report arXiv:1209.1989, Sep 2012. Comments: Plenary talk presented at RPC2012. Part of JINST RPC2012 special issue.
- [52] G L et al Bayatyan. *CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems*. Technical Design Report CMS.
- [53] S. Chatrchyan et al. Observation of a new boson with mass near 125 GeV in pp collisions at  $\sqrt{s} = 7$  and 8 TeV. *Journal of High Energy Physics*, 2013(6),

2013. URL: <http://dx.doi.org/10.1007/JHEP06%282013%29081>,  
doi:10.1007/JHEP06(2013)081.
- [54] Serguei Chatrchyan and Helen F Heath. Search for a Higgs boson decaying into a  $Z$  and a photon in pp collisions at  $s = 7$  and 8 TeV. *Physics Letters B - Nuclear Physics and Particle Physics*, B726:587–609, 2013.  
doi:10.1016/j.physletb.2013.09.057.
- [55] Vardan Khachatryan et al. Observation of the diphoton decay of the Higgs boson and measurement of its properties. 2014. arXiv:1407.0558.
- [56] Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC. *Physics Letters B*, 726(13):88 – 119, 2013. URL: <http://www.sciencedirect.com/science/article/pii/S0370269313006369>,  
doi:<http://dx.doi.org/10.1016/j.physletb.2013.08.010>.
- [57] Search for Higgs boson decays to a photon and a  $Z$  boson in pp collisions at and 8 TeV with the ATLAS detector. *Physics Letters B*, 732(0):8 – 27, 2014. URL: <http://www.sciencedirect.com/science/article/pii/S0370269314001713>,  
doi:<http://dx.doi.org/10.1016/j.physletb.2014.03.015>.
- [58] S. Chatrchyan et al. Measurement of Higgs boson production and properties in the  $WW$  decay channel with leptonic final states. *Journal of High Energy Physics*, 2014(1), 2014. URL:  
<http://dx.doi.org/10.1007/JHEP01%282014%29096>,  
doi:10.1007/JHEP01(2014)096.
- [59] S. Chatrchyan et al. Search for the standard model Higgs boson produced in association with a  $W$  or a  $Z$  boson and decaying to bottom quarks. *Phys. Rev. D*, 89:012003, Jan 2014. URL:

<http://link.aps.org/doi/10.1103/PhysRevD.89.012003>,  
doi:10.1103/PhysRevD.89.012003.

- [60] S. Chatrchyan et al. Evidence for the 125 GeV Higgs boson decaying to a pair of leptons. *Journal of High Energy Physics*, 2014(5), 2014. URL: <http://dx.doi.org/10.1007/JHEP05%282014%29104>, doi:10.1007/JHEP05(2014)104.
- [61] The CMS Collaboration. Evidence for the direct decay of the 125 GeV Higgs boson to fermions. *Nat Phys*, 10:557–560, August 2014. URL: <http://dx.doi.org/10.1038/nphys3005>, doi:10.1038/nphys3005.
- [62] S. Chatrchyan et al. Search for the standard model Higgs boson produced in association with a top-quark pair in pp collisions at the LHC. *Journal of High Energy Physics*, 2013(5), 2013. URL: <http://dx.doi.org/10.1007/JHEP05%282013%29145>, doi:10.1007/JHEP05(2013)145.
- [63] Vardan Khachatryan et al. Search for the associated production of the Higgs boson with a top-quark pair. Technical Report arXiv:1408.1682. CMS-HIG-13-029. CERN-PH-EP-2014-189, CERN, Geneva, Aug 2014. Comments: Submitted to JHEP.
- [64] S. Chatrchyan et al. Search for invisible decays of Higgs bosons in the vector boson fusion and associated ZH production modes. *The European Physical Journal C*, 74(8), 2014. URL: <http://dx.doi.org/10.1140/epjc/s10052-014-2980-6>, doi:10.1140/epjc/s10052-014-2980-6.
- [65] S. Chatrchyan et al. Combined results of searches for the standard model Higgs boson in pp collisions at  $\sqrt{s} = 7$  TeV. *Physics Letters B*, 710(1):26 – 48, 2012. URL: [http:](http://)

[//www.sciencedirect.com/science/article/pii/S0370269312002055](http://www.sciencedirect.com/science/article/pii/S0370269312002055),  
doi:<http://dx.doi.org/10.1016/j.physletb.2012.02.064>.

- [66] G. Aad et al. Search for Invisible Decays of a Higgs Boson Produced in Association with a  $Z$  Boson in ATLAS. *Phys. Rev. Lett.*, 112:201802, May 2014. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.112.201802>, doi:10.1103/PhysRevLett.112.201802.
- [67] Florian Beaudette. The CMS Particle Flow Algorithm. pages 295–304, 2014. arXiv:1401.8155.
- [68] Sidney D. Drell and Tung-Mow Yan. Massive lepton-pair production in hadron-hadron collisions at high energies. *Phys. Rev. Lett.*, 25:316–320, Aug 1970. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.25.316>, doi:10.1103/PhysRevLett.25.316.
- [69] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Physics Letters B*, 659(12):119 – 126, 2008. URL: <http://www.sciencedirect.com/science/article/pii/S0370269307011094>, doi:<http://dx.doi.org/10.1016/j.physletb.2007.09.077>.
- [70] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The catchment area of jets. *Journal of High Energy Physics*, 2008(04):005, 2008. URL: <http://stacks.iop.org/1126-6708/2008/i=04/a=005>.
- [71] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet user manual. *The European Physical Journal C*, 72(3), 2012. URL: <http://dx.doi.org/10.1140/epjc/s10052-012-1896-2>, doi:10.1140/epjc/s10052-012-1896-2.
- [72] John M. Campbell and R.K. Ellis. MCFM for the tevatron and the LHC. *Nuclear Physics B - Proceedings Supplements*, 205206(0):10 – 15, 2010. Loops and Legs in Quantum Field Theory Proceedings of the 10th DESY Workshop on Elementary Particle Theory. URL: [http:](http://)

[//www.sciencedirect.com/science/article/pii/S0920563210001945](http://www.sciencedirect.com/science/article/pii/S0920563210001945),  
doi:<http://dx.doi.org/10.1016/j.nuclphysbps.2010.08.011>.

- [73] Torbjrn Sjstrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 physics and manual. *Journal of High Energy Physics*, 2006(05):026, 2006. URL: <http://stacks.iop.org/1126-6708/2006/i=05/a=026>.
- [74] Thomas Junk. Confidence level computation for combining searches with small statistics. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 434(23):435 – 443, 1999. URL: <http://www.sciencedirect.com/science/article/pii/S0168900299004982>, doi:[http://dx.doi.org/10.1016/S0168-9002\(99\)00498-2](http://dx.doi.org/10.1016/S0168-9002(99)00498-2).
- [75] A L Read. Presentation of search results: the cl s technique. *Journal of Physics G: Nuclear and Particle Physics*, 28(10):2693, 2002. URL: <http://stacks.iop.org/0954-3899/28/i=10/a=313>.
- [76] L.D. Landau. On the angular momentum of a two-photon system. *Dokl.Akad.Nauk Ser.Fiz.*, 60:207–209, 1948.
- [77] C. N. Yang. Selection Rules for the Dematerialization of a Particle into Two Photons. *Phys. Rev.*, 77:242–245, Jan 1950. URL: <http://link.aps.org/doi/10.1103/PhysRev.77.242>, doi:10.1103/PhysRev.77.242.
- [78] Georges Aad et al. Measurement of the Higgs boson mass from the  $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ^* \rightarrow 4\ell$  channels with the ATLAS detector using  $25 \text{ fb}^{-1}$  of  $pp$  collision data. 2014. arXiv:1406.3827.
- [79] S. Chatrchyan et al. Study of the Mass and Spin-Parity of the Higgs Boson Candidate via Its Decays to  $Z$  Boson Pairs. *Phys. Rev. Lett.*, 110:081803, Feb 2013. URL: <http://link.aps.org/doi/10.1103/PhysRevLett.110.081803>, doi:10.1103/PhysRevLett.110.081803.

- [80] S Heinemeyer et al. Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group. Technical Report arXiv:1307.1347. CERN-2013-004, Geneva, 2013. Comments: 404 pages, 139 figures, to be submitted to CERN Report.
- [81] M. Veltman. Limit on mass differences in the Weinberg model. *Nuclear Physics B*, 123(1):89 – 99, 1977. URL: <http://www.sciencedirect.com/science/article/pii/055032137790342X>, doi:[http://dx.doi.org/10.1016/0550-3213\(77\)90342-X](http://dx.doi.org/10.1016/0550-3213(77)90342-X).
- [82] P. Sikivie, L. Susskind, M. Voloshin, and V. Zakharov. Isospin breaking in technicolor models. *Nuclear Physics B*, 173(2):189 – 207, 1980. URL: <http://www.sciencedirect.com/science/article/pii/055032138090214X>, doi:[http://dx.doi.org/10.1016/0550-3213\(80\)90214-X](http://dx.doi.org/10.1016/0550-3213(80)90214-X).
- [83] Procedure for the LHC Higgs boson search combination in Summer 2011. Technical Report CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, Aug 2011.
- [84] Precise determination of the mass of the Higgs boson and studies of the compatibility of its couplings with the standard model. Technical Report CMS-PAS-HIG-14-009, CERN, Geneva, 2014.
- [85] Fabrizio Caola and Kirill Melnikov. Constraining the Higgs boson width with  $ZZ$  production at the LHC. *Phys. Rev. D*, 88:054024, Sep 2013. URL: <http://link.aps.org/doi/10.1103/PhysRevD.88.054024>, doi:10.1103/PhysRevD.88.054024.
- [86] Constraints on the Higgs boson width from off-shell production and decay to Z-boson pairs. *Physics Letters B*, 736(0):64 – 85, 2014. URL: <http://www.sciencedirect.com/science/article/pii/S0370269314004821>, doi:<http://dx.doi.org/10.1016/j.physletb.2014.06.077>.

- [87] J. Neyman and E. S. Pearson. On the Problem of the Most Efficient Tests of Statistical Hypotheses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 231:pp. 289–337, 1933. URL: <http://www.jstor.org/stable/91247>.
- [88] Evidence for the spin-0 nature of the Higgs boson using ATLAS data. *Physics Letters B*, 726(13):120 – 144, 2013. URL: <http://www.sciencedirect.com/science/article/pii/S0370269313006527>, doi:<http://dx.doi.org/10.1016/j.physletb.2013.08.026>.
- [89] John C. Collins and Davison E. Soper. Angular distribution of dileptons in high-energy hadron collisions. *Phys. Rev. D*, 16:2219–2225, Oct 1977. URL: <http://link.aps.org/doi/10.1103/PhysRevD.16.2219>, doi:10.1103/PhysRevD.16.2219.
- [90] David J Lange and Representing the CMS Collaboration. The cms reconstruction software. *Journal of Physics: Conference Series*, 331(3):032020, 2011. URL: <http://stacks.iop.org/1742-6596/331/i=3/a=032020>.
- [91] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, F. Behner, L. Bellagamba, J. Boudreau, L. Broglia, A. Brunengo, H. Burkhardt, S. Chauvie, J. Chuma, R. Chytracsek, G. Cooperman, G. Cosmo, P. Degtyarenko, A. Dell’Acqua, G. Depaola, D. Dietrich, R. Enami, A. Feliciello, C. Ferguson, H. Fesefeldt, G. Folger, F. Foppiano, A. Forti, S. Garelli, S. Giani, R. Giannitrapani, D. Gibin, J.J. Gmez Cadenas, I. Gonzalez, G. Gracia Abril, G. Greeniaus, W. Greiner, V. Grichine, A. Grossheim, S. Guatelli, P. Gumplinger, R. Hamatsu, K. Hashimoto, H. Hasui, A. Heikkinen, A. Howard, V. Ivanchenko, A. Johnson, F.W. Jones, J. Kallenbach, N. Kanaya, M. Kawabata, Y. Kawabata, M. Kawaguti, S. Kelner, P. Kent, A. Kimura, T. Kodama, R. Kokoulin, M. Kossov, H. Kurashige, E. Lamanna, T. Lampn,

V. Lara, V. Lefebure, F. Lei, M. Liendl, W. Lockman, F. Longo, S. Magni, M. Maire, E. Medernach, K. Minamimoto, P. Mora de Freitas, Y. Morita, K. Murakami, M. Nagamatu, R. Nartallo, P. Nieminen, T. Nishimura, K. Ohtsubo, M. Okamura, S. O’Neale, Y. Oohata, K. Paech, J. Perl, A. Pfeiffer, M.G. Pia, F. Ranjard, A. Rybin, S. Sadilov, E. Di Salvo, G. Santin, T. Sasaki, N. Savvas, Y. Sawada, S. Scherer, S. Sei, V. Sirotenko, D. Smith, N. Starkov, H. Stoecker, J. Sulkimo, M. Takahata, S. Tanaka, E. Tcherniaev, E. Safai Tehrani, M. Tropeano, P. Truscott, H. Uno, L. Urban, P. Urban, M. Verderi, A. Walkden, W. Wander, H. Weber, J.P. Wellisch, T. Wenaus, D.C. Williams, D. Wright, T. Yamada, H. Yoshida, and D. Zschesche. Geant4a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250 – 303, 2003. URL: <http://www.sciencedirect.com/science/article/pii/S0168900203013688>, doi:[http://dx.doi.org/10.1016/S0168-9002\(03\)01368-8](http://dx.doi.org/10.1016/S0168-9002(03)01368-8).

- [92] J. Allison, K. Amako, J. Apostolakis, H. Araujo, P.A. Dubois, M. Asai, G. Barrand, R. Capra, S. Chauvie, R. Chytracsek, G.A.P. Cirrone, G. Cooperman, G. Cosmo, G. Cuttone, G.G. Daquino, M. Donszelmann, M. Dressel, G. Folger, F. Foppiano, J. Generowicz, V. Grichine, S. Guatelli, P. Gumplinger, A. Heikkinen, I. Hrivnacova, A. Howard, S. Incerti, V. Ivanchenko, T. Johnson, F. Jones, T. Koi, R. Kokoulin, M. Kossov, H. Kurashige, V. Lara, S. Larsson, F. Lei, O. Link, F. Longo, M. Maire, A. Mantero, B. Mascialino, I. McLaren, P.M. Lorenzo, K. Minamimoto, K. Murakami, P. Nieminen, L. Pandola, S. Parlati, L. Peralta, J. Perl, A. Pfeiffer, M.G. Pia, A. Ribon, P. Rodrigues, G. Russo, S. Sadilov, G. Santin, T. Sasaki, D. Smith, N. Starkov, S. Tanaka, E. Tcherniaev, B. Tome, A. Trindade, P. Truscott, L. Urban, M. Verderi, A. Walkden, J.P. Wellisch, D.C. Williams, D. Wright, and H. Yoshida. Geant4 developments and applications. *Nuclear Science, IEEE Transactions on*, 53(1):270–278, Feb

2006. doi:10.1109/TNS.2006.869826.

- [93] Douglas Orbaker and the CMS collaboration. Fast simulation of the CMS detector. *Journal of Physics: Conference Series*, 219(3):032053, 2010. URL: <http://stacks.iop.org/1742-6596/219/i=3/a=032053>.
- [94] Yi Chen, Emanuele Di Marco, Joe Lykken, Maria Spiropulu, Roberto Vega-Morales, et al. 8D Likelihood Effective Higgs Couplings Extraction Framework in the Golden Channel. 2014. arXiv:1401.2077.
- [95] Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *Journal of High Energy Physics*, 2004(11):040, 2004. URL: <http://stacks.iop.org/1126-6708/2004/i=11/a=040>.
- [96] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *Journal of High Energy Physics*, 2007(11):070, 2007. URL: <http://stacks.iop.org/1126-6708/2007/i=11/a=070>.
- [97] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *Journal of High Energy Physics*, 2010(6), 2010. URL: <http://dx.doi.org/10.1007/JHEP06%282010%29043>, doi:10.1007/JHEP06(2010)043.
- [98] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. Vector boson plus one jet production in POWHEG. *Journal of High Energy Physics*, 2011(1), 2011. URL: <http://dx.doi.org/10.1007/JHEP01%282011%29095>, doi:10.1007/JHEP01(2011)095.
- [99] JohnM. Campbell, R.Keith Ellis, Paolo Nason, and Giulia Zanderighi. W and Z bosons in association with two jets using the POWHEG method. *Journal of High Energy Physics*, 2013(8), 2013. URL:

<http://dx.doi.org/10.1007/JHEP08%282013%29005>,  
doi:10.1007/JHEP08(2013)005.

- [100] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2), 2011. URL:  
<http://dx.doi.org/10.1140/epjc/s10052-011-1554-0>,  
doi:10.1140/epjc/s10052-011-1554-0.
- [101] Brian Bradie. *A friendly introduction to numerical analysis: with C and MATLAB materials on website; int. ed.* Prentice-Hall, Englewood Cliffs, NJ, 2006.
- [102] Yi Chen, Roni Harnik, and Roberto Vega-Morales. Probing the Higgs Couplings to Photons in  $h \rightarrow 4\ell$  at the LHC. 2014. arXiv:1404.1336.
- [103] Ian Low, Joseph Lykken, and Gabe Shaughnessy. Have we observed the higgs boson (imposter)? *Phys. Rev. D*, 86:093012, Nov 2012. URL:  
<http://link.aps.org/doi/10.1103/PhysRevD.86.093012>,  
doi:10.1103/PhysRevD.86.093012.
- [104] David McKeen, Maxim Pospelov, and Adam Ritz. Modified higgs branching ratios versus cp and lepton flavor violation. *Phys. Rev. D*, 86:113004, Dec 2012. URL: <http://link.aps.org/doi/10.1103/PhysRevD.86.113004>,  
doi:10.1103/PhysRevD.86.113004.
- [105] The ACME Collaboration, J. Baron, W. C. Campbell, D. DeMille, J. M. Doyle, G. Gabrielse, Y. V. Gurevich, P. W. Hess, N. R. Hutzler, E. Kirilov, I. Kozyryev, B. R. OLeary, C. D. Panda, M. F. Parsons, E. S. Petrik, B. Spaun, A. C. Vutha, and A. D. West. Order of magnitude smaller limit on the electric dipole moment of the electron. *Science*, 343(6168):269–272, 2014. URL: <http://www.sciencemag.org/content/343/6168/269.abstract>,

arXiv:<http://www.sciencemag.org/content/343/6168/269.full.pdf>,  
doi:10.1126/science.1248213.

- [106] Roberta Arcidiacono and the CMS ECAL Group. Studies of the CMS electromagnetic calorimeter performance in the electron test beam. *Journal of Physics: Conference Series*, 160(1):012048, 2009. URL: <http://stacks.iop.org/1742-6596/160/i=1/a=012048>.
- [107] S. Chatrchyan et al. Jet production rates in association with W and Z bosons in pp collisions at  $\sqrt{s} = 7$  TeV. *Journal of High Energy Physics*, 2012(1), 2012. URL: <http://dx.doi.org/10.1007/JHEP01%282012%29010>,  
doi:10.1007/JHEP01(2012)010.
- [108] Jogesh C. Pati and Abdus Salam. Lepton number as the fourth "color". *Phys. Rev. D*, 10:275–289, Jul 1974. URL: <http://link.aps.org/doi/10.1103/PhysRevD.10.275>,  
doi:10.1103/PhysRevD.10.275.
- [109] Lorenzo Moneta, Kevin Belasco, Kyle S. Cranmer, S. Kreiss, Alfio Lazzaro, et al. The RooStats Project. *PoS*, ACAT2010:057, 2010. arXiv:1009.1003.