

Higgs Boson Properties and Search for Additional Resonances

by

Can You

A dissertation submitted to The Johns Hopkins University in conformity with the
requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

January, 2016

© Can You 2016

All rights reserved

Abstract

Abstract goes here.

Primary Reader: Andrei Gritsan

Secondary Reader: Morris Swartz

Acknowledgments

I would like to first thank my advisor, Andrei Gritsan, without whom none of these is possible. You helped me to see all the things that I could improve, not only about research, but most importantly, about styles and attitudes toward work and life. It is your patient and careful guidance and constant pursue of excellence that lead me to become a better researcher and a better person, and gave me more confidence in facing future challenges. I also like to thank Meng Xiao, who has kindly give me a lot of guidance and encouragements on research, and shared with me many of her brilliant ideas. Your courage and determination in facing all challenges has inspired me a lot. I would like to thank Andrew, Ian and Chris, who has helped me to start my first research projects, patiently explained to me all the details, and encouraged me. I'd like to thank Xiaozhou, Jonathan, Ulascan and Heshy, it is very fortunate to work with great researchers like you. I learned a lot from each of you. I also wish to thank Carol, Wenzer and Robert, It was a very fulfilling experience to work with you. Your motivation, hard-working, and passion for research have been a great encouragement for me. I would also like to give thanks to many people outside JHU , in the HZZ

group and alignment group, and many researchers I met in CERN.

I'd like to thank my undergraduate advisors Hoi Fung Chau, Jason Pun and John Leung, for guiding me through my years in the University of Hong Kong, teaching me to deal with various problems in my student life, helping me to find my interest in particle physics and lead me to become a researcher. I must also thank my high school physics teachers Yunhui Li and Liang Zhang, who cultivate in me the passion for science, teach me to tackle problems outside the textbooks and showed me the exciting experiences in solving problems.

I cannot thank my family enough for supporting me to pursue this PhD. I want to thank my parents, who always encouraged me to seek challenge rather than comfort, and always believed in me when I face difficulties. I wish to thank my husband and colleague, Lei Feng, for all he has done for me in the past ten years, with unconditional love and support.

For all the friends I met in these years, thank you, you made my PhD life a warm and delightful memory to me.

Dedication

This thesis is dedicated to ...

Contents

Abstract	ii
Acknowledgments	iii
List of Tables	xii
List of Figures	xiii
1 Introduction	1
1.1 Theoretical Motivation	1
1.1.1 the Elementary Particles	1
1.1.2 Electroweak Symmetry Breaking	1
1.1.3 Discovery of the Higgs Boson	1
1.1.4 Beyond the SM Higgs	1
1.2 Summary	1
2 Experimental Setup	2

2.1	The Large Hadron Collider	2
2.2	The Compact Muon Solenoid	2
2.2.1	The Magnet	2
2.2.2	Inner Tracker	2
2.2.3	Electromagnetic Calorimeter	3
2.2.4	Hadronic Calorimeter	3
2.2.5	Muon System	3
2.3	Alignment of CMS tracker	3
2.3.1	Track Based Alignment Techniques	4
2.3.1.1	Parameterisation	5
2.3.1.2	Global and Local Approach	5
2.3.1.3	Hierarchical Alignment	5
2.3.1.4	Differential Alignment	5
2.3.2	Weak Modes	6
2.3.2.1	Weighted Tracks	6
2.3.2.2	Zuu Constraints	6
2.3.2.3	Survey Constraints	6
2.3.3	Validation Methods	6
2.3.3.1	Track Splitting	6
2.3.3.2	Distribution of the Median of Residuals	7
2.3.3.3	Primary Vertex	8

2.3.4	Conclusion	10
2.4	Summary	11
3	Higgs Boson Phenomenology at the LHC	12
3.1	Higgs Boson Production	12
3.2	Higgs Boson Decay	12
3.3	Kinematics for Properties Measurement	12
4	Higgs Boson Exotic Spin	15
4.1	Datasets	16
4.1.1	Experimental data	16
4.1.2	Simulated samples	16
4.2	Event selection and backgrounds	18
4.3	Yields	19
4.4	Kinematic Distributions	21
4.5	Hypothesis Testing	26
4.5.1	Spin-1	27
4.5.2	Spin-2	29
4.6	Systematic Uncertainties	34
4.6.1	Theoretical uncertainties	34
4.6.2	Experimental uncertainties	34
4.7	Results	37

4.7.1	Testing of mixed spin-1 hypotheses	37
4.7.2	Testing of spin-2 hypotheses	41
4.7.3	Constrains on mixed non-interfering states	46
4.8	Summary	54
4.9	$H \rightarrow ZZ \rightarrow 4l$	55
4.9.1	Datasets	55
4.9.1.1	Experimental data	55
4.9.1.2	Simulated samples	56
4.9.2	Event Selection	59
4.9.3	Event Categorization	61
4.9.4	Kinematic discriminants	62
4.9.5	Signal Modeling	65
4.9.5.1	Signal model: reconstruction efficiency	66
4.9.5.2	Signal mass model: resolution	67
4.9.5.3	Signal mass model: lineshape prior to detector effects	71
4.9.5.4	Signal mass model: closure tests	74
4.9.5.5	Signal model: kinematics	76
4.9.6	Background Estimation	76
4.9.7	Systematic uncertainties	82
4.9.8	Results	84
4.10	$H \rightarrow ZZ \rightarrow 2l2q$	87

4.10.1	Data and Simulation	87
4.10.1.1	Experimental Data	87
4.10.1.2	Simulated Samples	88
4.10.2	Event Selection	89
4.10.2.1	Leptonic Z selection	89
4.10.2.2	Hadronic Z selection	90
4.10.3	Kinematic fitting and Kinematic discriminants	93
4.10.3.1	Kinematic fitting	93
4.10.3.2	Kinematic discriminants	94
4.10.4	Event categorisation and final selection	101
4.10.5	Signal Modeling	103
4.10.5.1	Signal selection efficiency	103
4.10.5.2	Signal mass model: resolution	108
4.10.5.3	Signal model with kinematics	111
4.10.6	Background Estimation	111
4.10.6.1	Estimation of ttbar from data	111
4.10.6.2	Estimation of Z+jets	112
4.10.7	Statistical Analysis and Systematics	115
4.10.8	Results	118
4.10.8.1	Observed Data	118
4.10.8.2	Spin-zero Limits	120

4.10.8.3	Spin-two Limits	124
4.11	Summary	124
5	Conclusion	126
	Bibliography	127

List of Tables

4.1	Spin 0 samples at 8TeV and 7TeV.	17
4.2	Spin 1 samples at 8TeV and 7TeV.	17
4.3	Spin 2 samples at 8TeV and 7TeV.	17
4.4	The number of estimated background and signal events, and number of observed candidates, after final inclusive selection.	20
4.5	List of kinematic discriminants used in this analyses.	25
4.6	Summary of the systematics.	36
4.7	A summary of models used in the analysis of spin-one hypotheses. Events are generated with SM expectation for the signal cross section ($\mu=1$).	37
4.8	List of models used in the analysis of the spin-two hypotheses corresponding to the pure states of the type noted.	42
4.9	Signal Monte Carlo samples and cross sections.	57
4.10	Background Monte Carlo samples and cross sections.	58
4.11	Summary of selection requirements and categorisation. Signal and side-band regions are defined by ranges in the hadronic Z boson candidate mass Z_{had} , after all other selection criteria are applied.	102
4.12	Summary of the experimental systematic uncertainties.	119
4.13	Expected yields in the categories passing resolved-jet selection.	119
4.14	Expected yields in the categories passing merged-jet selection.	120

List of Figures

2.1	The normalized differences between two halves of a cosmic track, split at the point of closest approach to the interaction region, in the track's azimuthal angle and polar angle	7
2.2	The distribution of median residuals is plotted for the local x- and y-directions in the barrel pixel detector, using 1 million collision events collected with the magnetic field at 3.8T. The blue line shows the geometry obtained at the end of 2015, and the alignment shown in magenta was produced with the Millepede-II and HipPy algorithms using 3.8T cosmic ray and collision data.	9
4.1	Distribution of $\mathcal{D}_{\text{bkg}}^{\text{dec}}$ vs. $\mathcal{D}_{1-}^{\text{dec}}$ for the qqZZ background (a) and for a signal resonance consistent with the SM Higgs boson with $m_{0+} = 125.6$ GeV. (b)	23
4.2	Observed distributions of the discriminants in data and MC expectations for the background and for a signal resonance at $m_H = 125.6$ GeV, either the SM Higgs boson or an alternative spin-1 hypothesis indicated. Top row from left to right: $\mathcal{D}_{\text{bkg}}(J^P = 1^-)$, $\mathcal{D}_{\text{bkg}}(J^P = 1^+)$, \mathcal{D}_{1-} , \mathcal{D}_{1+} ; bottom row from left to right: $\mathcal{D}_{\text{bkg}}^{\text{dec}}(J^P = 1^-)$, $\mathcal{D}_{\text{bkg}}^{\text{dec}}(J^P = 1^+)$, $\mathcal{D}_{1-}^{\text{dec}}$, $\mathcal{D}_{1+}^{\text{dec}}$. All distributions, except \mathcal{D}_{bkg} and $\mathcal{D}_{\text{bkg}}^{\text{dec}}$, are shown with the requirement $\mathcal{D}_{\text{bkg}} > 0.5$ ($\mathcal{D}_{\text{bkg}}^{\text{dec}} > 0.5$) to enhance signal purity.	24
4.3	Distribution of $\mathcal{D}_{\text{bkg}}^{\text{dec}}$ vs. $\mathcal{D}_{1-}^{\text{dec}}$ for the qqZZ background (a) and for a signal resonance consistent with the SM Higgs boson with $m_{0+} = 125.6$ GeV. (b)	28
4.4	Distribution of $D_{\text{bkg}}^{\text{dec}}$ vs. $D_{J^P}^{\text{dec}}$ for the $gg \rightarrow ZZ$ background (left) and for the signal resonance consistent with the SM Higgs boson $m_H = 125.6$ GeV model.	30
4.5	Distribution of D_{bkg} in data and MC expectations for the background and for a signal resonance consistent with SM Higgs boson at $m_H = 125.6$ GeV (left). The $D_{\text{bkg}}^{\text{dec}}$ distribution for the production independent scenario (right).	31

4.6	Distributions of D_{JP} with a requirement $D_{bkg} > 0.5$. Distributions for background and signal are shown.	32
4.7	Distributions of D_{JP} with a requirement $D_{bkg} > 0.5$. Distributions for background and signal are shown.	33
4.8	Distribution of the test-statistic $q = -2 \ln(\mathcal{L}_{JP}/\mathcal{L}_{0+})$ of the spin-1 hypothesis tested against the SM Higgs boson hypothesis, for the production independent scenario. J^P = pure 1^- (top left), mixture with $f_{b2}=0.2$ (top middle), $f_{b2}=0.4$ (top right), $f_{b2}=0.6$ (bottom left), $f_{b2}=0.8$ (bottom middle), and pure 1^+ (bottom right). Distributions for the SM Higgs boson are represented by the yellow histogram and for the alternative J^P hypotheses by the blue histogram.	38
4.9	Distribution of the test-statistic $q = -2 \ln(\mathcal{L}_{JP}/\mathcal{L}_{0+})$ of the spin-1 hypothesis tested against the SM Higgs boson hypothesis, for the production dependent scenario. J^P = pure 1^- (top left), mixture with $f_{b2}=0.2$ (top middle), $f_{b2}=0.4$ (top right), $f_{b2}=0.6$ (bottom left), $f_{b2}=0.8$ (bottom middle), and pure 1^+ (bottom right). Distributions for the SM Higgs boson are represented by the yellow histogram and for the alternative J^P hypotheses by the blue histogram.	39
4.10	Distribution of median test-statistic q , as a function of f_{b2} . The green(red) and blue(black) bands represent the 1σ and 2σ around the median expected value for the SM(spin-1) Higgs boson hypothesis.	40
4.11	Distribution of test-statistics $q = -2 \ln(\mathcal{L}_{JP}/\mathcal{L}_{0+})$ for two signal types (0^+ represented by the yellow histogram and alternative hypothesis by the blue histogram) for $m_H = 125.6$ GeV shown with a large number of generated experiments. Red arrow represents the observed value of the test-statistics.	43
4.12	Distribution of $q = -2 \ln(\mathcal{L}_{JP}/\mathcal{L}_{0+})$ for two signal types (0^+ represented by the yellow histogram and alternative hypothesis by the blue histogram) for $m_H = 125.6$ GeV shown with a large number of generated experiments. Ten alternative hypotheses are tested in gg , $q\bar{q}$ and production model independent way.	44
4.13	Summary of the expected and observed values for the test-statistic q distributions for the twelve alternative hypotheses tested with respect to the SM Higgs boson. The orange (blue) bands represent the 1σ , 2σ , and 3σ around the median expected value for the SM Higgs boson hypothesis (alternative hypothesis). The black point represents the observed value.	45

4.14	Expected and observed distributions of $-2\Delta \ln \mathcal{L}$ as a function of $f(J^P)$, in the production independent (any $\rightarrow J^P$) case. $J^P =$ pure 1^- (top left), mixture with $f_{b2}=0.2$ (top middle), $f_{b2}=0.4$ (top right), $f_{b2}=0.6$ (bottom left), $f_{b2}=0.8$ (bottom middle), and pure 1^+ (bottom right). The horizontal lines at $-2\Delta \ln \mathcal{L} = 1$ and 3.84 represent the 68% and 95% C.L., respectively.	47
4.15	Expected and observed distributions of $-2\Delta \ln \mathcal{L}$ as a function of $f(J^P)$, in the $\bar{q}q$ case. $J^P =$ pure 1^- (top left), mixture with $f_{b2}=0.2$ (top middle), $f_{b2}=0.4$ (top right), $f_{b2}=0.6$ (bottom left), $f_{b2}=0.8$ (bottom middle), and pure 1^+ (bottom right). The horizontal lines at $-2\Delta \ln \mathcal{L} = 1$ and 3.84 represent the 68% and 95% C.L., respectively.	48
4.16	Summary of expected and observed constraints on the non-interfering fraction measurements for spin-one.	49
4.17	Non-interfering fraction expectations for various spin-2 models. The above results are production-independent.	50
4.18	Non-interfering fraction expectations for various spin-2 models ($gg \rightarrow J^P$).	51
4.19	Non-interfering fraction expectations for various spin-2 models ($\bar{q}q \rightarrow J^P$).	52
4.20	Summary of expected and observed constraints on the non-interfering fraction measurements for spin-two.	53
4.21	Signal relative purity of the six event categories in terms of the 5 main production mechanisms of the H(125) boson in a $118 < m_{4\ell} < 130$ GeV window. The WH , ZH and $t\bar{t}H$ processes are split according to the decay of associated objects, whereby X denotes anything else than a lepton.	63
4.22	Parameterization of reconstruction efficiency $\mathcal{E}(m_{4\ell})$ (including acceptance) for a SM-like signal with $J^P = 0^+$ in the gluon fusion production (solid circles); for a SM-like signal with $J^P = 0^+$ in the VBF production (solid triangles); and for a spin-2 signal with $J^P = 2_b^+$ in the model of KK graviton propagating in the bulk in the gluon fusion production (open circles). Dependence as a function of mass $m_{4\ell}$ is shown in three channels: $X \rightarrow 4\mu$ (red, top); $2e2\mu$ (blue, middle); $X \rightarrow 4e$ (green, bottom). Points show results from full simulation, curves show analytical fits.	67
4.23	Representative $m_{4\ell}$ distributions of a resolution function $\mathcal{R}(m_{4\ell} m_{4\ell}^{\text{ideal}})$ in $2e2\mu$ (left), $4e$ (middle), and 4μ (right) channels. Several examples are shown from top to bottom: $m_{4\ell} = 120, 200, 740, 1000$ GeV.	69
4.24	Dependence of six parameters of the resolution function in the $2e2\mu$ (left), $4e$ (middle), and 4μ (right) channels. Point show results of individual fits, smooth curves show parameterization from a simultaneous fit of all distributions.	70

4.25 Distribution of events (arbitrary units) in the process $gg \rightarrow 4\ell$ (left) and $qq \rightarrow 4\ell$ (right) as a function of invariant mass $m_{4\ell}$ generated with the MCFM+JHUGen and Phantom framework. Various components in the $X(450) + H(125)^* + \text{bkg}$ interference are shown and explained in the legend.	73
4.26 Left: comparison of POWHEG, Phantom simulation, analytical model in VBF and gluon fusion model. Right: closure test on VBF full simulation.	74
4.27 Closure test if $H(125)$ width parameterization with off-shell tail. Left: SM width, right: 10 times SM width. Statistics box shows fit results for the mass and width, in agreement with the input.	75
4.28 Top: Observed and expected upper limits at the 95% CL on the $X \rightarrow ZZ \rightarrow 4\ell$ cross section σ_X (including four-lepton branching fraction) as a function of m_X at several Γ_X values with f_{VBF} floated, with 12.9 fb^{-1} of data at 13 TeV. Bottom: Same as above but with $f_{\text{VBF}} = 1$ fixed.	86
4.29 Distribution of \mathcal{D}_{Zjj} with events with two resolved jets (left) and with a merged jet (right). Spin-0 at top, spin-2 at bottom.	96
4.30 Distribution of \mathcal{D}_{Zjj} vs $m_{2\ell 2q}$ with conditional normalisation for spin-0 signal (top left), DY + jets (top right), $t\bar{t} + WW$ (bottom left), ZW, ZZ diboson backgrounds (bottom left), for the merged selection. When conditionally normalised, each slice of \mathcal{D}_{Zjj} is normalised to unit area for a given value of $m_{2\ell 2q}$	97
4.31 Distribution of \mathcal{D}_{Zjj} vs $m_{2\ell 2q}$ with conditional normalisation for spin-0 signal (top left), DY + jets (top right), $t\bar{t} + WW$ (bottom left), ZW, ZZ diboson backgrounds (bottom right), for the resolved selection. When conditionally normalised, each slice of \mathcal{D}_{Zjj} is normalised to unit area for a given value of $m_{2\ell 2q}$	98
4.32 Distribution of \mathcal{D}_{Zjj} vs $m_{2\ell 2q}$ with conditional normalisation for spin-2 signal (top left), DY + jets (top right), $t\bar{t} + WW$ (bottom left), ZW, ZZ diboson backgrounds (bottom right), for the merged selection. When conditionally normalised, each slice of \mathcal{D}_{Zjj} is normalised to unit area for a given value of $m_{2\ell 2q}$	99
4.33 Distribution of \mathcal{D}_{Zjj} vs $m_{2\ell 2q}$ with conditional normalisation for spin-2 signal (top left), DY + jets (top right), $t\bar{t} + WW$ (bottom left), ZW, ZZ diboson backgrounds (bottom right), for the resolved selection. When conditionally normalised, each slice of \mathcal{D}_{Zjj} is normalised to unit area for a given value of $m_{2\ell 2q}$	100
4.34 Parameterization of reconstruction efficiency $\mathcal{E}(m_{2\ell 2q})$ (including acceptance) (left: $Z \rightarrow ee$, right: $Z \rightarrow \mu\mu$) for a SM-like signal with $J^P = 0^+$ in the gluon fusion production. Lines refer to the 6 jet categories.	105

4.35 Parameterization of reconstruction efficiency $\mathcal{E}(m_{2\ell 2q})$ (including acceptance) (left: $Z \rightarrow ee$, right: $Z \rightarrow \mu\mu$) for a SM-like signal with $J^P = 0^+$ in the VBF production. Lines refer to the 6 jet categories.	106
4.36 Parameterization of reconstruction efficiency $\mathcal{E}(m_{2\ell 2q})$ (including acceptance) (left: $Z \rightarrow ee$, right: $Z \rightarrow \mu\mu$) for a spin-2 signal with $J^P = 2_b^+$ in the model of KK graviton . Lines refer to the 6 jet categories.	107
4.37 Representative $m_{2\ell 2q}$ distributions of a resolution function $\mathcal{R}(m_{2\ell 2q} m_{2\ell 2q}^{\text{ideal}})$ two-jet (left) and merged-jet (right) categories. Several examples are shown from top to bottom: $m_{2\ell 2q} = 750, 1000, 2000$ GeV.	109
4.38 Dependence of six parameters of the resolution function in the two-jet (left) and merged-jet (right) categories. Point show results of individual fits, smooth curves show parameterization from a simultaneous fit of all distributions.	110
4.39 Expected and observed upper limits on the $pp \rightarrow X \rightarrow ZZ$ cross section under narrow width resonance assumption for 12.9 fb^{-1} of CMS data at 13 TeV where the fraction of VBF production is allowed to float. Left plot is done using m_{ZZ} as the observable and right plot is done with statistical model including both m_{ZZ} and D_{jet} information.	121
4.40 Expected upper limits on the $pp \rightarrow X \rightarrow ZZ$ cross section under narrow width resonance assumption for 12.9 fb^{-1} of CMS data at 13 TeV assuming 100% gluon fusion production. Left plot is done using m_{ZZ} as the observable and right plot is done with statistical model including both m_{ZZ} and D_{jet} information.	122
4.41 Expected upper limits on the $pp \rightarrow X \rightarrow ZZ$ cross section under narrow width resonance assumption for 12.9 fb^{-1} of CMS data at 13 TeV assuming 100% VBF production. Left plot is done using m_{ZZ} as the observable and right plot is done with statistical model including both m_{ZZ} and D_{jet} information.	123
4.42 Expected and observed upper limits on the $pp \rightarrow X \rightarrow ZZ$ cross section under narrow width resonance assumption for 12.9 fb^{-1} of CMS data at 13 TeV with a spin-2 graviton assumption. Left plot is done using m_{ZZ} as the observable and right plot is done with statistical model including both m_{ZZ} and D_{jet} information.	124

Chapter 1

Introduction

1.1 Theoretical Motivation

1.1.1 the Elementary Particles

1.1.2 Electroweak Symmetry Breaking

1.1.3 Discovery of the Higgs Boson

1.1.4 Beyond the SM Higgs

1.2 Summary

Chapter 2

Experimental Setup

2.1 The Large Hadron Collider

2.2 The Compact Muon Solenoid

2.2.1 The Magnet

2.2.2 Inner Tracker

Figure .. shows an r-z view of the CMS Tracker.

2.2.3 Electromagnetic Calorimeter

2.2.4 Hadronic Calorimeter

2.2.5 Muon System

2.3 Alignment of CMS tracker

The task of CMS tracker is to measure the trajectories of charged particle with high resolution in track position, angle and momentum. Excellent tracking performance is essential for analysis of physics processes. For example, searches for high mass resonances in leptonic final states require good momentum resolution for transverse momenta above 1TeV. In addition, excellent impact parameter resolution of reconstructed tracks is important for reconstruction of beam spot, primary vertices and b-jets tagging.

Tracker geometry refers to a complete set of parameters describing the geometrical properties (both location and angles) for the modules composing the tracker, and it is one of the most important input for track reconstruction. The uncertainty of tracker geometry should remain below the intrinsic silicon hit resolution of around $10\mu\text{m}$ for pixels and $30\mu\text{m}$ for strips.

There are many challenges to determine to tracker geometry, such as the limited accessibility of the tracker, the large number of modules to align, the high precision

required, and the constant changing in geometry due to changes in environment conditions. Track-based internal alignment is effective in solving these problems, and it is responsible for adjusting the positions and angles of the tracker modules relative to each other. Survey measurements of TOB is used to determine its global position relative to the beam axis and the other sub-components are aligned relative to TOB by track-based alignment algorithms.

2.3.1 Track Based Alignment Techniques

Track-hit residuals is sensitive to the tracker geometry used in track reconstruction relative to the true positions and angles of tracker modules. Track based alignment determine the module positions using the tracks reconstructed by the tracker geometry in situ. It can be treated as a least square minimization problem, where sum of the squares of track-hit residuals from a collection of tracks is minimized.

The objective function to minimize is: .. where ..

An initial geometry is used to determine the approximate track parameters. Since the corrections from the initial geometry can be assumed to be small, f_{ij} can be linearised around these initial values. The linearisation leads to normal equation of least squares.

Due to the large number of parameters (16588 modules with 6 parameters each), the alignment problem is hard to solve exactly, and several approximation methods are used.

2.3.1.1 Parameterisation

2.3.1.2 Global and Local Approach

global local combine

2.3.1.3 Hierarchical Alignment

The CMS tracker is build in a hierarchical way. HL - ML

2.3.1.4 Differential Alignment

IOV

2.3.2 Weak Modes

2.3.2.1 Weighted Tracks

2.3.2.2 Zuu Constraints

2.3.2.3 Survey Constraints

2.3.3 Validation Methods

2.3.3.1 Track Splitting

Cosmic tracks are split in half at the point of closest approach to beam line, then both halves are reconstructed independently and their parameters are compared at the splitting point. The normalized differences between two halves are histogrammed, and the width of the distribution measures the achieved alignment precision, while deviations from zero indicate possible biases.

Figure. 2.1 shows an example of track splitting validation results.

2.3.3.2 Distribution of the Median of Residuals

Each track is refitted using the alignment constants under consideration, and the hit prediction for each module is obtained from all of the other track hits. The median of the distribution of unbiased hit residuals (DMR) is then taken for each module and is histogrammed.

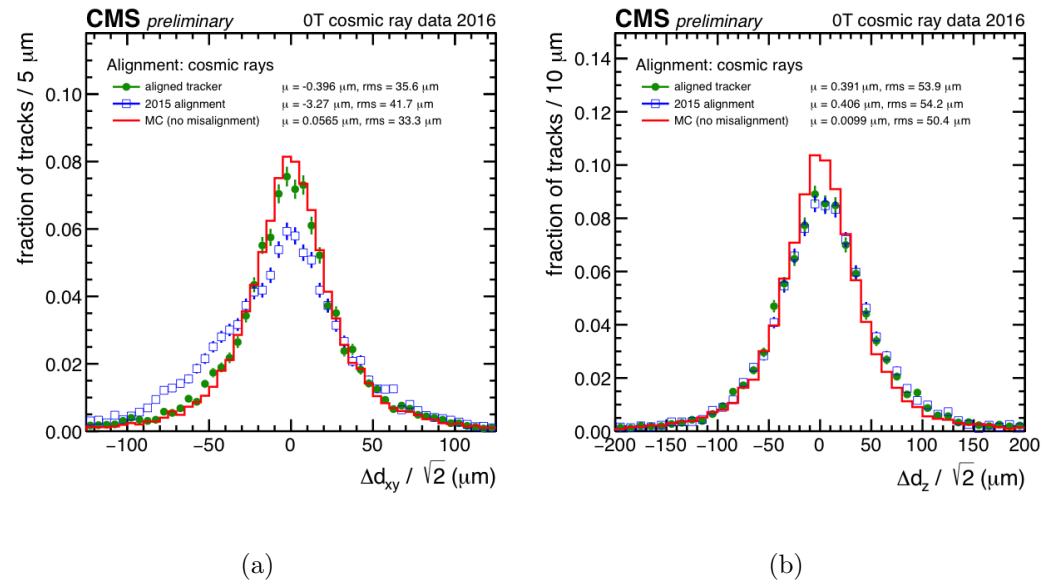


Figure 2.1: The normalized differences between two halves of a cosmic track, split at the point of closest approach to the interaction region, in dxy , the xy distance, and dz , the z distance, between the track and the origin. The observed precision using the aligned geometry (green circles), produced with the Millepede-II and HipPy algorithms using cosmic ray data at 0T, is a major improvement over the 2015 EOY geometry (blue empty squares) which is no longer valid for 2016 data, primarily because of temperature and magnetic field changes. The mean and rms values, calculated in the plotting range, show that the bias is reduced compared with 2015 geometry, and the precision comes close to that of the ideal Monte Carlo, illustrating that the tracker has almost reached its design spatial and angular resolution for 0T cosmic ray data.

The width of this distribution of the medians of residuals is a measure of the statistical precision of alignment results; deviations from zero indicate possible biases. The width also has an intrinsic component due to the limited number of tracks, meaning that distributions can only be compared if they are produced with the same number of tracks, as is the case within each set of plots here.

Figure. 2.2 shows an example the DMR validation results.

2.3.3.3 Primary Vertex

The resolution of the reconstructed vertex position is driven by the pixel detector since it is the closest detector to the interaction point and has the best hit resolution. The primary vertex residual method is based on the study the distance between the track and the vertex, the latter reconstructed without the track under scrutiny (unbiased track-vertex residual).

Events used in this analysis are selected online with minimum bias triggers. The fit of the vertex must have at least 4 degrees of freedom. For each of the vertices, the impact parameters are measured for tracks with more than 6 hits in the tracker, of which at least two are in the pixel detector, and at least one hit in the first layer of the Barrel Pixel or the first disk of the Forward Pixel, with Chi2/ndof of the track fit ≤ 5 . The vertex position is recalculated excluding the track under scrutiny. A deterministic annealing clustering algorithm is used in order to make the method robust against pileup, as in the default reconstruction sequence.

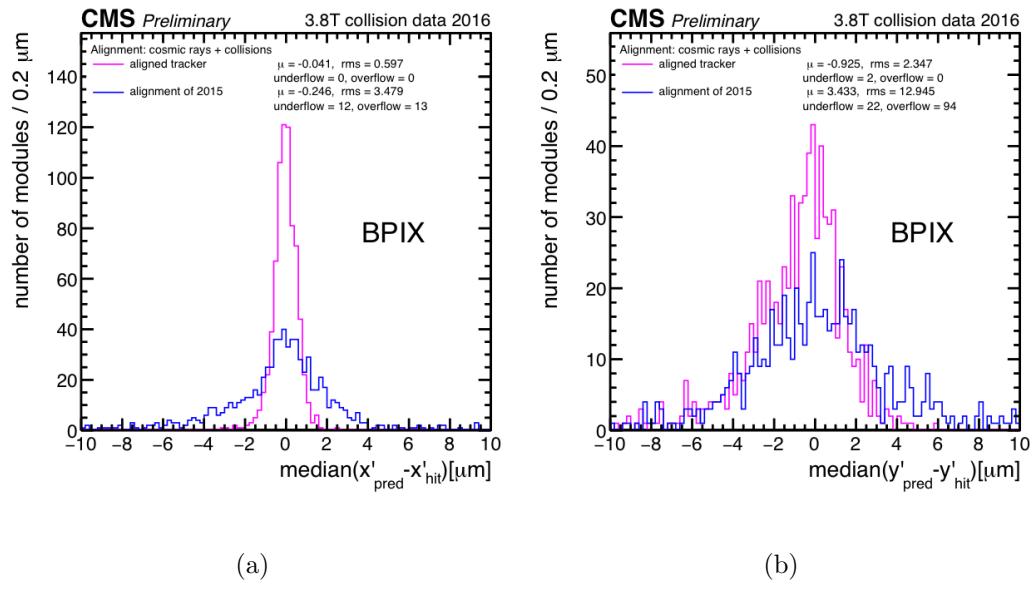


Figure 2.2: The distribution of median residuals is plotted for the local x- and y-directions in the barrel pixel detector, using 1 million collision events collected with the magnetic field at 3.8T. The blue line shows the geometry obtained at the end of 2015, which is no longer valid for 2016 data, primarily because of temperature and magnetic field changes. The alignment shown in magenta was produced with the Millepede-II and HipPy algorithms using 3.8T cosmic ray and collision data. The position of the pixel detector is known to be very sensitive to the change of conditions. The rms values, calculated using modules both inside and outside the plot range, show improvement over the 2015 geometry by a factor of 5.

The distributions of the unbiased track-vertex residuals in the transverse plane, and in the longitudinal direction, are studied in bins of track azimuth and pseudo-rapidity . Random misalignments of the modules affect only the resolution of the unbiased track-vertex residual, increasing the width of the distributions, but without biasing their mean. Systematic movements of the modules will bias the distributions in a way that depends on the nature and size of the misalignment and the and of the selected tracks.

Figure. 2.1 shows an example the primary vertex validation results.

2.3.4 Conclusion

The track-based alignment has been performed. The local and global methods deliver similar results and show improvements and the combined approach give the best results.

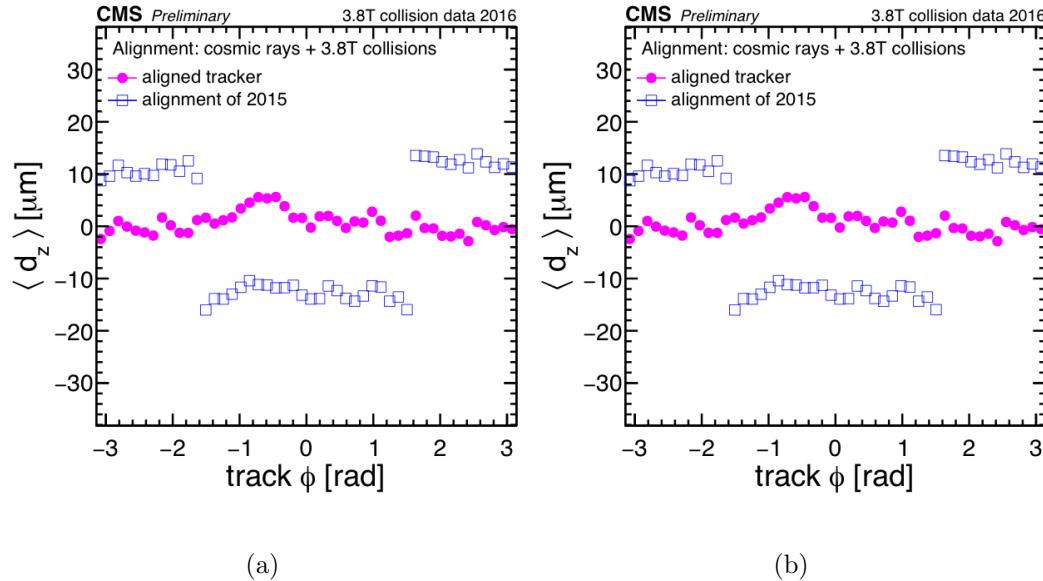


Figure 2.3: The distance in the transverse (d_{xy}) and longitudinal (d_z) plane of the track at its closest approach to a refit unbiased primary vertex is studied in bins of track azimuth ϕ using a sample of 1 million collision events collected by the CMS detector at full magnetic field (3.8T) selected online through minimum bias triggers. The performance of a dedicated alignment achieved with the Millepede-II and HipPy algorithms using cosmic ray and collision data at 3.8T is compared to the one of the alignment used to reprocess the collision data collected by CMS during 2015.

2.4 Summary

Chapter 3

Higgs Boson Phenomenology at the LHC

3.1 Higgs Boson Production

3.2 Higgs Boson Decay

3.3 Kinematics for Properties Measure- ment

Kinematics of the Higgs decay to 4ℓ final state and its application to the studies of the Higgs properties has been extensively studied in the literature.^{11, 14, 15, 20, 21, 23, 31, 32, 32, 37, 40–43, 48, 52, 67}

The four-momenta of $H \rightarrow 4\ell$ decay products carry eight independent degrees of freedom which fully describe the kinematic configuration of a four-lepton system in its center-of-mass frame, up to an arbitrary rotation around the beam axis. These can be expressed in terms of five angles $\vec{\Omega} \equiv (\theta^*, \Phi_1, \theta_1, \theta_2, \Phi)$ defined in Fig. ??,^{24,42} the invariant masses of the lepton pairs, m_{Z_1} and m_{Z_2} , and the invariant masses of the 4ℓ system. These are defined in the center-of-mass frame as follows:

$m_{4\ell}$ is an invariant mass of the four-lepton system.

m_1 is the invariant mass of the Z_1 .

m_2 is the invariant mass of the Z_2 .

θ_1 is an angle between the Z_1 boost direction and its daughter particle's momentum in the Z_1 rest frame.

θ_2 is an angle between the Z_2 boost direction and its daughter particle's momentum in the Z_2 rest frame.

Φ is an angle between the Z_1 and Z_2 decay planes in the CM frame.

Φ_1 is a rotation angle of Z_1 lepton (e^- or μ^-) along the $-z$ -beam axis when Z_1 is aligned along the beamline, or a rotation around $-\hat{z}_{\text{beam}}$. This is a production-dependent variable.

θ^* is an angle between the beam axis and a Z_1 – Z_2 axis in the CM frame. This is a production-dependent variable.

These angular observables, information on their correlations and the dilepton pairs invariant masses can be used to experimentally establish the consistency of the spin and parity quantum numbers with respect to the SM, and to further discriminate signal from background and thus increase the signal sensitivity and reduce the statistical uncertainty in measurements.

Chapter 4

Higgs Boson Exotic Spin

The observation of a new boson consistent with the standard model (SM) Higgs boson^{38,47,49–51,56} has been reported by ATLAS and CMS Collaborations in 2012 .^{6,27} It has been followed by a comprehensive set of measurements of the properties of the new boson with focus on answering if the new boson is the SM Higgs boson and if there are any indications for the physics beyond SM.

Four-lepton final state provides a particularly sensitive channel for the measurement of the coupling structures which determine the spin and CP properties of the boson. Hypothesis testing can be used to evaluate if the data is compatible with the SM Higgs boson or some exotic spin states, like a graviton. In this chapter, the tests of mixed spin-one resonances and spin-two resonances will be presented.

4.1 Datasets

4.1.1 Experimental data

The datasets used was recorded by the CMS experiment corresponding to $\mathcal{L} = 5.1 \text{ fb}^{-1}$ in 2011 at 7 TeV and $\mathcal{L} = 19.7 \text{ fb}^{-1}$ in 2012 at 8 TeV. Events are selected online requiring the presence of a pair of electrons or muons, or a triplet of electrons. Triggers requiring an electron and a muon are also used. The minimal pT of the first and second lepton are 17 and 8 GeV, respectively, for the double lepton triggers, while they are 15, 8 and 5 GeV for the triple electron trigger.

4.1.2 Simulated samples

Spin-zero signal samples are simulated using at next-to-leading order (NLO) using the **POWHEG**^{17,39,63} generator and decayed through **JHUGen**^{14,21,42} to correctly model all spin correlations and interferences of all contributing amplitudes. These samples were then interfaced with **PYTHIA**⁶⁶ and processed through a dedicated simulation of the CMS detector based on **GEANT4**.⁸ This production contains various samples to adequately cover the spin-zero terms considered in this analysis. For alternate spin-one and spin-two signals the simulation was done at LO with **JHUGen** where production and decay spin correlations are important.

The samples are presented in Tables 4.1, 4.2, and 4.3.

Table 4.1: Spin 0 samples at 8TeV and 7TeV.

Sample	scenario
Higgs0PMToZZTo4L_M-125p6_7/8TeV-powheg15-JHUGenV3	0_m^+ SM

Table 4.2: Spin 1 samples at 8TeV and 7TeV.

Sample	scenario
Vector1PToZZTo4L_M-125p6_7/8TeV-JHUGenV3	1^+
Vector1MToZZTo4L_M-125p6_7/8TeV-JHUGenV3	1^-
Vector1Mf05ph01Pf05ph0ToZZTo4L_M-125p6_7/8TeV-JHUGenV3	$f_{b_1} = 0.5$, phase=0
Vector1Mf05ph01Pf05ph90ToZZTo4L_M-125p6_7/8TeV-JHUGenV3	$f_{b_1} = 0.5$, phase= $\pi/2$

Table 4.3: Spin 2 samples at 8TeV and 7TeV.

Sample	scenario
Graviton2PH2ToZZTo4L_M-125p6_7/8TeV-JHUGenV3	$2_{h_2}^+$
Graviton2PH3ToZZTo4L_M-125p6_7/8TeV-JHUGenV3	$2_{h_3}^+$
Graviton2PH6ToZZTo4L_M-125p6_7/8TeV-JHUGenV3	$2_{h_6}^+$
Graviton2PH7ToZZTo4L_M-125p6_7/8TeV-JHUGenV3	$2_{h_7}^+$
Graviton2MH9ToZZTo4L_M-125p6_7/8TeV-JHUGenV3	$2_{h_9}^-$
Graviton2MH10ToZZTo4L_M-125p6_7/8TeV-JHUGenV3	$2_{h_{10}}^-$
q \bar{q}	
Graviton2BPqqbarToZZTo4L_M-125p6_7/8TeV-JHUGenV3	2_b^+
Graviton2HPqqbarToZZTo4L_M-125p6_7/8TeV-JHUGenV3	2_h^+
Graviton2MPqqbarToZZTo4L_M-125p6_7/8TeV-JHUGenV3	2_h^-
Graviton2PH2qqbarToZZTo4L_M-125p6_7/8TeV-JHUGenV3	$2_{h_2}^+$
Graviton2PH3qqbarToZZTo4L_M-125p6_7/8TeV-JHUGenV3	$2_{h_3}^+$
Graviton2PH6qqbarToZZTo4L_M-125p6_7/8TeV-JHUGenV3	$2_{h_6}^+$
Graviton2PH7qqbarToZZTo4L_M-125p6_7/8TeV-JHUGenV3	$2_{h_7}^+$
Graviton2MH9qqbarToZZTo4L_M-125p6_7/8TeV-JHUGenV3	$2_{h_9}^-$
Graviton2MH10qqbarToZZTo4L_M-125p6_7/8TeV-JHUGenV3	$2_{h_{10}}^-$

4.2 Event selection and backgrounds

Events with at least four identified and isolated electrons or muons, compatible with being produced at the primary vertex, are then selected offline. We require a Z candidate originating from a pair of leptons of the same flavor and opposite charge. The pair with an invariant mass closest to the nominal Z mass is retained and is denoted Z_1 if it satisfies $40 \leq m_{Z_1} \leq 120\text{GeV}$. A second $\ell^+\ell^-$ pair, denoted Z_2 , is required to satisfy $12 \leq m_{Z_2} \leq 120\text{GeV}$. If more than one Z_2 candidate satisfies all criteria, the pair of leptons with the highest scalar p_T sum is chosen. At least one lepton should have $p_T \geq 20\text{GeV}$, another one $p_T \geq 10\text{GeV}$ and any opposite-charge pair of leptons among the four selected must satisfy $m_{\ell\ell} \geq 4\text{GeV}$. Finally events are restricted to a narrow window around the observed 125.6 GeV resonance ($105.6 < m_{4\ell} < 140.6\text{GeV}$).

After the selection, the irreducible backgrounds originates from the $\bar{q}q \rightarrow ZZ$ and $gg \rightarrow ZZ$ processes. While the reducible background, comes from the production of Z and WZ in association with jets, as well as $t\bar{t}$, with one or two jets misidentified as an electron or a muon. The dominant $\bar{q}q \rightarrow ZZ$ and $gg \rightarrow ZZ$ backgrounds are evaluated from simulation. The reducible background is evaluated using a tight-to-loose “fake rate” method. Two independent implementations are combined, which differ by the definition of the control regions in the data and by the difference in the composition of the samples where the fake rates are measured. The shape of the $m_{4\ell}$ distribution for the reducible background is obtained by fitting the $m_{4\ell}$ distributions in two different

control regions with empirical functional forms.

4.3 Yields

The number of estimated background and signal events and number of observed candidates after final inclusive selection in data in the narrow mass region around 125.6 GeV, the counting in data compared with signal expectation in a small range around 125.6 GeV($105.6 < m_{4\ell} < 140.6 \text{ GeV}$), is given in Table 4.4, separately for 2011 (7 TeV) and 2012 (8TeV) and all combined.

Table 4.4: The number of estimated background and signal events, and number of observed candidates, after final inclusive selection.

Channel	4e	4 μ	2e2 μ
5.1 fb ⁻¹ @ 7 TeV			
$\bar{q}q \rightarrow ZZ$	0.84 \pm 0.10	1.80 \pm 0.11	2.24 \pm 0.28
$Z + X$	0.62 \pm 0.14	0.22 \pm 0.09	1.06 \pm 0.29
$gg \rightarrow ZZ$	0.03 \pm 0.01	0.06 \pm 0.02	0.07 \pm 0.02
All background expected	1.49 \pm 0.17	2.08 \pm 0.14	3.37 \pm 0.40
$m_H = 125.6$ GeV	0.70 \pm 0.11	1.24 \pm 0.14	1.67 \pm 0.26
Observed	1	3	6
19.7 fb ⁻¹ @ 8 TeV			
$\bar{q}q \rightarrow ZZ$	2.94 \pm 0.33	7.65 \pm 0.49	8.86 \pm 0.68
$Z + X$	2.77 \pm 0.62	1.19 \pm 0.48	4.29 \pm 1.10
$gg \rightarrow ZZ$	0.20 \pm 0.05	0.41 \pm 0.10	0.50 \pm 0.13
All background expected	5.91 \pm 0.71	9.25 \pm 0.69	13.65 \pm 1.30
$m_H = 125.6$ GeV	3.09 \pm 0.47	5.95 \pm 0.71	7.68 \pm 0.98
Observed	9	15	16
5.1 fb ⁻¹ @ 7 TeV and 19.7 fb ⁻¹ @ 8 TeV			
$\bar{q}q \rightarrow ZZ$	3.78 \pm 0.34	9.45 \pm 0.50	11.10 \pm 0.73
$Z + X$	3.39 \pm 0.64	1.41 \pm 0.49	5.36 \pm 1.14
$gg \rightarrow ZZ$	0.23 \pm 0.05	0.47 \pm 0.11	0.57 \pm 0.13
All background expected	7.40 \pm 0.73	11.33 \pm 0.71	17.03 \pm 1.36
$m_H = 125.6$ GeV	3.79 \pm 0.48	7.19 \pm 0.73	9.35 \pm 1.01
Observed	10	18	22

4.4 Kinematic Distributions

Separation between the two types of four-lepton processes can be achieved by constructing kinematic discriminants following the matrix element method approach.

Kinematic discriminants used in this study are computed using JHUGen^{14,21,42} matrix elements for signal and MCFM²⁵ matrix elements for background processes, both implemented with the MELA package.^{14,21,42}

Given several signal (SM and J^P) and the main background ($q\bar{q} \rightarrow ZZ$) hypotheses, there are several effective probabilities that one can calculate for each event:

$$\mathcal{P}_{\text{SM}} = \mathcal{P}_{\text{SM}}^{\text{kin}}(\vec{\Omega}, m_1, m_2 | m_{4\ell}) \times \mathcal{P}_{\text{sig}}^{\text{mass}}(m_{4\ell} | m_H)$$

$$\mathcal{P}_{J^P} = \mathcal{P}_{J^P}^{\text{kin}}(\vec{\Omega}, m_1, m_2 | m_{4\ell}) \times \mathcal{P}_{\text{sig}}^{\text{mass}}(m_{4\ell} | m_H)$$

$$\mathcal{P}_{q\bar{q}ZZ} = \mathcal{P}_{q\bar{q}ZZ}^{\text{kin}}(\vec{\Omega}, m_1, m_2 | m_{4\ell}) \times \mathcal{P}_{q\bar{q}ZZ}^{\text{mass}}(m_{4\ell})$$

Observables (kinematic discriminants) are constructed from the probability listed above:

$$\mathcal{D}_{\text{bkg}} = \frac{\mathcal{P}_{\text{SM}}}{\mathcal{P}_{\text{SM}} + c \times \mathcal{P}_{\text{bkg}}} = \left[1 + c(m_{4\ell}) \times \frac{\mathcal{P}_{\text{bkg}}^{\text{kin}}(m_1, m_2, \vec{\Omega} | m_{4\ell}) \times \mathcal{P}_{\text{bkg}}^{\text{mass}}(m_{4\ell})}{\mathcal{P}_{\text{SM}}^{\text{kin}}(m_1, m_2, \vec{\Omega} | m_{4\ell}) \times \mathcal{P}_{\text{sig}}^{\text{mass}}(m_{4\ell} | m_H)} \right]^{-1} \quad (4.1)$$

$$\mathcal{D}_{J^P} = \frac{\mathcal{P}_{\text{SM}}^{\text{kin}}}{\mathcal{P}_{\text{SM}}^{\text{kin}} + c_{J^P} \times \mathcal{P}_{J^P}^{\text{kin}}} = \left[1 + c_{J^P} \times \frac{\mathcal{P}_{J^P}^{\text{kin}}(m_1, m_2, \vec{\Omega} | m_{4\ell})}{\mathcal{P}_{\text{SM}}^{\text{kin}}(m_1, m_2, \vec{\Omega} | m_{4\ell})} \right]^{-1} \quad (4.2)$$

Here, the constants c_x are tuned to adjust relative normalization of probabilities, for the optimal appearance of the event distributions after the detector acceptance effects. The \mathcal{D}_{bkg} discriminant is used to separate signal from background. The \mathcal{D}_{J^P} discriminant is used to separate the SM signal form an alternative pure J^P state.

It is desirable to extend the hypothesis testing approach in a way that does not depend on the production model. The spin-zero kinematics is already independent of the production mechanism due to lack of spin correlations for any spin-zero particle; as the result $\cos \theta^*$ and Φ_1 distributions are isotropic for any production model. For spin-one and spin-2 discriminants, this feature can be achieved by considering the unpolarized X -boson production by either averaging over the spin degrees of freedom of the produced X -boson or, equivalently, integrating over the two production angles $\cos \theta^*$ and Φ_1 , defined in Fig. ??, in the J_x^P probability expectation $\mathcal{P}_{J_x^P}$.⁴² This leads to the spin-averaged matrix element squared for the X -decay as the probability \mathcal{P} in the kinematic discriminant. This method is also applied to calculate the $\mathcal{D}_{\text{bkg}}^{\text{dec}}$ discriminant for signal-to-background separation, where we employ numerical integration over the two production angles.

$$\mathcal{D}_{\text{bkg}}^{\text{dec}} = \left[1 + c(m_{4\ell}) \times \frac{\frac{1}{4\pi} \int d\Phi_1 d \cos \theta^* \mathcal{P}_{\text{bkg}}^{\text{kin}}(m_1, m_2, \vec{\Omega} | m_{4\ell}) \times \mathcal{P}_{\text{bkg}}^{\text{mass}}(m_{4\ell})}{\mathcal{P}_{\text{SM}}^{\text{kin}}(m_1, m_2, \vec{\Omega} | m_{4\ell}) \times \mathcal{P}_{\text{sig}}^{\text{mass}}(m_{4\ell} | m_H)} \right]^{-1} \quad (4.3)$$

$$\mathcal{D}_{J_x^P}^{\text{dec}} = \left[1 + c_{J_x^P} \times \frac{\frac{1}{4\pi} \int d\Phi_1 d \cos \theta^* \mathcal{P}_{J_x^P}^{\text{kin}}(m_1, m_2, \vec{\Omega} | m_{4\ell})}{\mathcal{P}_{\text{SM}}^{\text{kin}}(m_1, m_2, \vec{\Omega} | m_{4\ell})} \right]^{-1} \quad (4.4)$$

Some examples of the distributions as expected from simulation and as observed in data can be seen in Figure 4.2 for the discriminants used in the spin-1 study. A complete list of all discriminants used in the analysis is presented in Table 4.5.

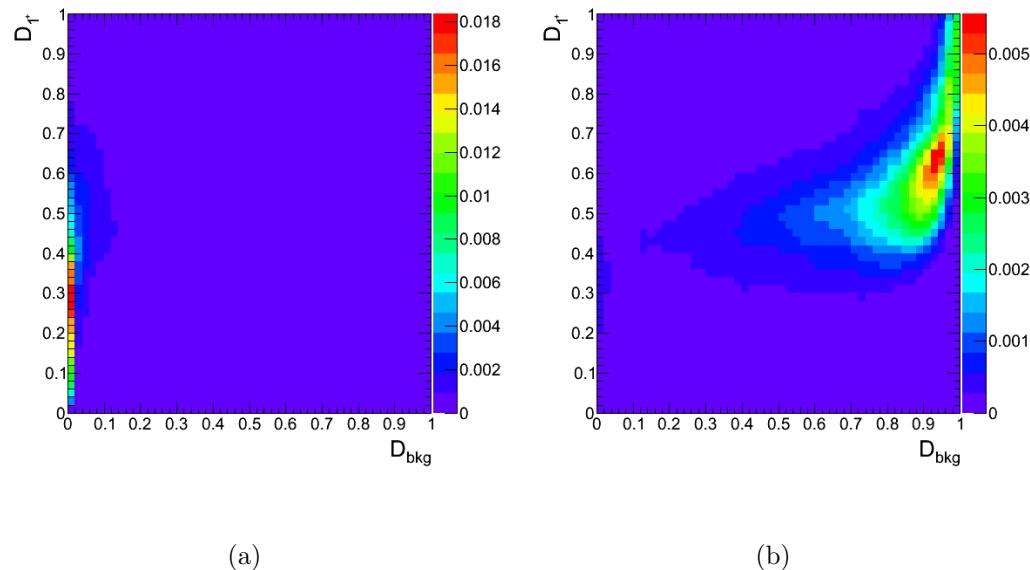


Figure 4.1: Distribution of $\mathcal{D}_{\text{bkg}}^{\text{dec}}$ vs. $\mathcal{D}_{1^-}^{\text{dec}}$ for the qqZZ background (a) and for a signal resonance consistent with the SM Higgs boson with $m_{0^+} = 125.6$ GeV. (b)

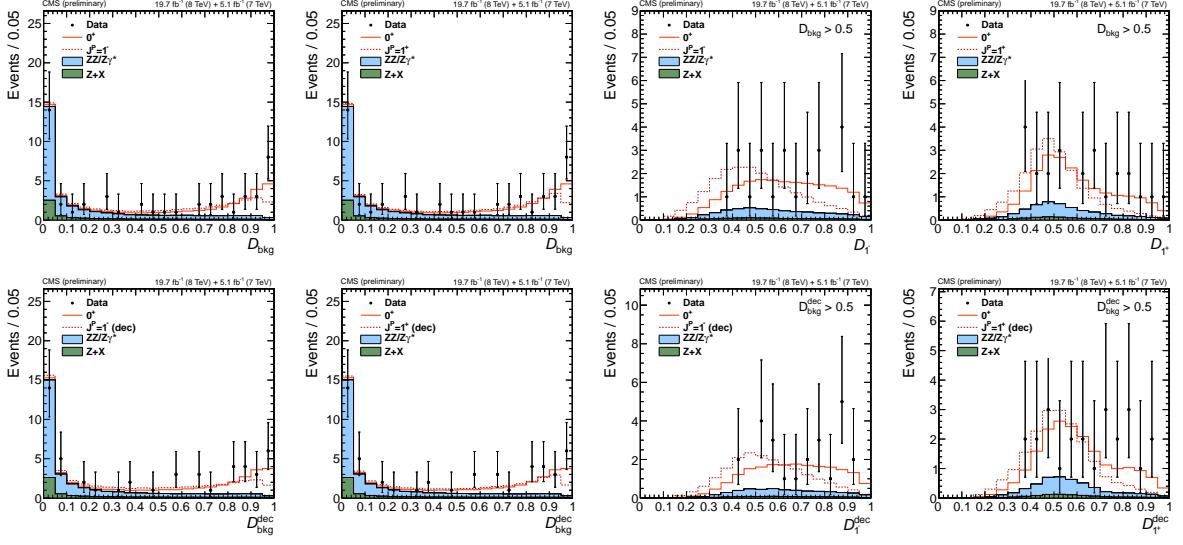


Figure 4.2: Observed distributions of the discriminants in data and MC expectations

for the background and for a signal resonance at $m_H = 125.6$ GeV, either the SM Higgs boson or an alternative spin-1 hypothesis indicated. Top row from left to right: $\mathcal{D}_{\text{bkg}}(J^P = 1^-)$, $\mathcal{D}_{\text{bkg}}(J^P = 1^+)$, \mathcal{D}_{1-} , \mathcal{D}_{1+} ; bottom row from left to right: $\mathcal{D}_{\text{bkg}}^{\text{dec}}(J^P = 1^-)$, $\mathcal{D}_{\text{bkg}}^{\text{dec}}(J^P = 1^+)$, $\mathcal{D}_{1-}^{\text{dec}}$, $\mathcal{D}_{1+}^{\text{dec}}$. All distributions, except \mathcal{D}_{bkg} and $\mathcal{D}_{\text{bkg}}^{\text{dec}}$, are shown with the requirement $\mathcal{D}_{\text{bkg}} > 0.5$ ($\mathcal{D}_{\text{bkg}}^{\text{dec}} > 0.5$) to enhance signal purity.

Table 4.5: List of kinematic discriminants used in this analyses.

Observables used for the study of the exotic models	
$\mathcal{D}_{\text{bkg}}^{\text{dec}}$	discriminate against ZZ background, include $m_{4\ell}$, exclude $\cos \theta^*$, Φ_1
$\mathcal{D}_1^{\text{dec}}$	Exotic vector (1^-), $q\bar{q} \rightarrow X$, decay only
$\mathcal{D}_{1+}^{\text{dec}}$	Exotic pseudovector (1^+), $q\bar{q} \rightarrow X$, decay only
$\mathcal{D}_{2_b^+}^{\text{dec}}$	KK Graviton-like with SM in the bulk (2_b^+), $q\bar{q} \rightarrow X$, decay only
$\mathcal{D}_{2_h^+}^{\text{dec}}$	BSM tensor with higher dim operators (2_h^+), $q\bar{q} \rightarrow X$, decay only
$\mathcal{D}_{2_h^-}^{\text{dec}}$	BSM pseudotensor with higher dim operators (2_h^-), $q\bar{q} \rightarrow X$, decay only
$\mathcal{D}_{2_{h_2}^+}^{\text{dec}}$	BSM tensor with higher dim operators ($2_{h_2}^+$), $gg \rightarrow X, q\bar{q} \rightarrow X$, decay only
$\mathcal{D}_{2_{h_3}^+}^{\text{dec}}$	BSM tensor with higher dim operators ($2_{h_3}^+$), $gg \rightarrow X, q\bar{q} \rightarrow X$, decay only
$\mathcal{D}_{2_{h_6}^+}^{\text{dec}}$	BSM tensor with higher dim operators ($2_{h_6}^+$), $gg \rightarrow X, q\bar{q} \rightarrow X$, decay only
$\mathcal{D}_{2_{h_7}^+}^{\text{dec}}$	BSM tensor with higher dim operators ($2_{h_7}^+$), $gg \rightarrow X, q\bar{q} \rightarrow X$, decay only
$\mathcal{D}_{2_{h_9}^-}^{\text{dec}}$	BSM pseudotensor with higher dim operators ($2_{h_9}^-$), $gg \rightarrow X, q\bar{q} \rightarrow X$, decay only
$\mathcal{D}_{2_{h_{10}}^-}^{\text{dec}}$	BSM pseudotensor with higher dim operators ($2_{h_{10}}^-$), $gg \rightarrow X, q\bar{q} \rightarrow X$, decay only

4.5 Hypothesis Testing

To distinguish an alternative signal hypothesis from the SM Higgs boson, hypothesis testing based on kinematic discriminants is used. It is based on probability density functions packed into the 2D or 3D templates, built out of the kinematic discriminant distributions D as described before. The technical implementation of the hypothesis testing is done in the framework of the RooStats based CMS Higgs combination tools.

The kinematic discriminant distributions are then used to construct the test statistic q for each point of parameter space as follows:

$$q = -2 \ln \frac{\mathcal{L}_X}{\mathcal{L}_{0^+}} = -2 \ln \prod_i \frac{\text{pdf}(D_i | X + \text{bkg})}{\text{pdf}(D_i | 0^+ + \text{bkg})}, \quad (4.5)$$

where i runs over all the events in an pseudo-experiment. The likelihoods \mathcal{L} is evaluated for the two models and assume the nominal SM Higgs boson signal strength ($\mu = 1$), while the cross sections for the alternative signal hypotheses are taken to be the same as for the SM Higgs boson (the $2e2\mu$ channel is taken as a reference). The distribution of test statistics is examined with generated samples of background and signal of two types (SM 0^+ and J^P) for $m_H = 125.6$ GeV. Here the likelihoods \mathcal{L} are calculated with the signal rates allowed to float independently for each signal type and the nuisance parameters are treated as independent. The expected distributions are generated with signal cross-section equal to that of the SM, which is consistent with observation.

To quantify the expected separation power between alternative signal hypotheses,

the “tail probability” of the TS distribution for hypothesis X at the point is computed, which is median of the TS distribution for hypothesis 0^+ , $P(q \geq q_{\text{median}}^{(X)} | 0^+)$.

Events in the mass range $105.6 < m_{4l} < 140.6$ GeV are used to perform these studies. The Higgs boson mass is assumed to be $m_H = 125.6$ GeV. Templates are obtained from simulation for signal and irreducible background and from control regions for the reducible backgrounds, using the procedure described.

4.5.1 Spin-1

In spin-1 analysis discrimination is based on 3D probability density functions $(\mathcal{D}_{1-}, \mathcal{D}_{1+}, \mathcal{D}_{\text{bkg}})$. The spin-1 hypotheses are also tested by relying only on their decay information, i.e. in a production-independent way, using kinematic discriminants $(\mathcal{D}_{1+}^{\text{dec}}, \mathcal{D}_{1-}^{\text{dec}}, \mathcal{D}_{\text{bkg}}^{\text{dec}})$.

Figure 4.3 shows expected 2D distributions of the discriminants $\mathcal{D}_{\text{bkg}}^{\text{dec}}$ vs. $\mathcal{D}_{1-}^{\text{dec}}$, for ggZZ background and SM. $\mathcal{D}_{\text{bkg}}^{\text{dec}}$ distributions differ significantly between signals and backgrounds.

The alternative signal models are defined by the tensor structure of couplings, however, the absolute values of couplings and, hence, the expected event yields are not uniquely defined. The cross sections for alternative signal hypotheses are left floating in the fit. The same approach is taken for the SM Higgs boson hypothesis, i.e. the overall SM Higgs boson signal strength μ_S is the best fit value as it comes out from data. This way, the overall signal event yield is not a part of the discrimination

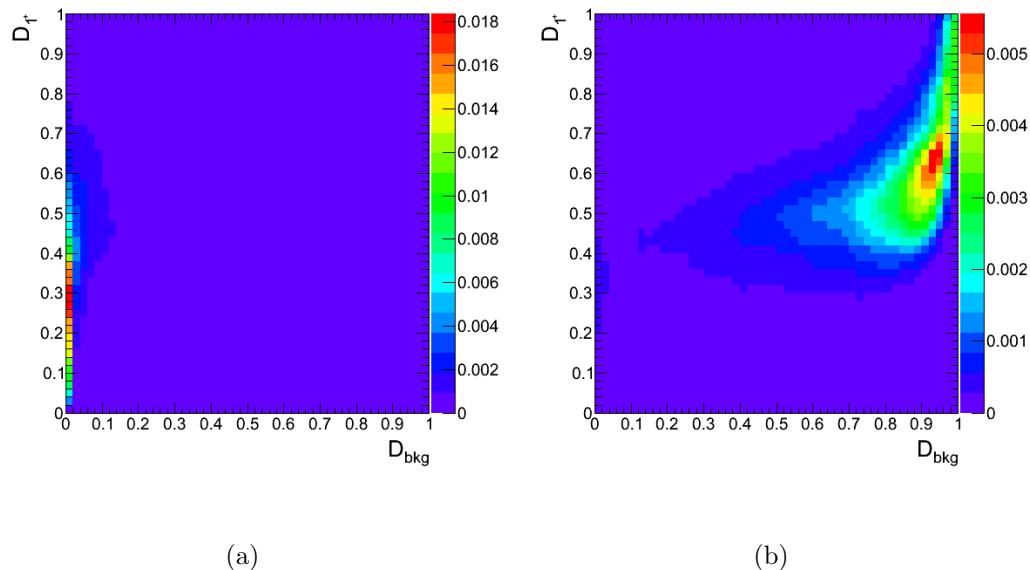


Figure 4.3: Distribution of $\mathcal{D}_{\text{bkg}}^{\text{dec}}$ vs. $\mathcal{D}_{1^-}^{\text{dec}}$ for the qqZZ background (a) and for a signal resonance consistent with the SM Higgs boson with $m_{0^+} = 125.6$ GeV. (b)

between alternative hypotheses. Consequently, for pair-wise tests of alternative signal hypotheses with respect to the SM Higgs boson, the test statistic is defined using the ratio of signal plus background likelihoods for two signal hypotheses.

In addition to pure 1^+ and 1^- , their mixtures with various fractions of 1^+ (f_{1+}) are also tested. The expected distribution of q for the spin-1 hypothesis (blue histogram) and the SM Higgs boson (orange histogram) are shown in Fig. 4.8 (production-independent) and Fig. 4.9 ($\bar{q}q$).

We assume the nominal SM Higgs boson signal strength ($\mu_S=1$), while the cross sections for the alternative signal hypotheses are taken to be the same as for the SM Higgs boson (the $2e2\mu$ channel is taken as a reference).

4.5.2 Spin-2

In spin-2 analysis, discrimination is based on 2D probability density functions (D_{bkg}, D_{JP}) packed into the 2D templates. All spin-2 models are tested in gg , $q\bar{q}$ and production model independent way. When doing production model independent tests we use decay only variables ($D_{bkg}^{dec}, D_{JP}^{dec}$).

Fig. 4.4 shows 2D distributions of the discriminants D_{bkg}^{dec} vs. D_{JP}^{dec} , for $gg \rightarrow ZZ$ background and Standard model. Distributions differ significantly between signal and background.

Fig. 4.5 shows distribution of the D_{bkg} observable, and the production independent D_{bkg}^{dec} observable, used only to test production independent hypotheses.

The test statistics q is constructed as:

The likelihood \mathcal{L} is evaluated for the two models, assuming the nominal SM Higgs boson signal strength ($\mu = 1$), while the cross sections for the alternative signal hypotheses are taken to be the same as for the SM Higgs boson (the $2e2\mu$ channel is taken as a reference). The distribution of test statistics is examined with generated samples of background and signal of two types (SM 0^+ and J^P) for $m_H = 125.6$ GeV. Here the likelihoods \mathcal{L} are calculated with the signal rates allowed to float independently for each signal type and the nuisance parameters are treated as independent. The expected distributions are generated with signal cross-section equal to that of the SM, which is consistent with observation.

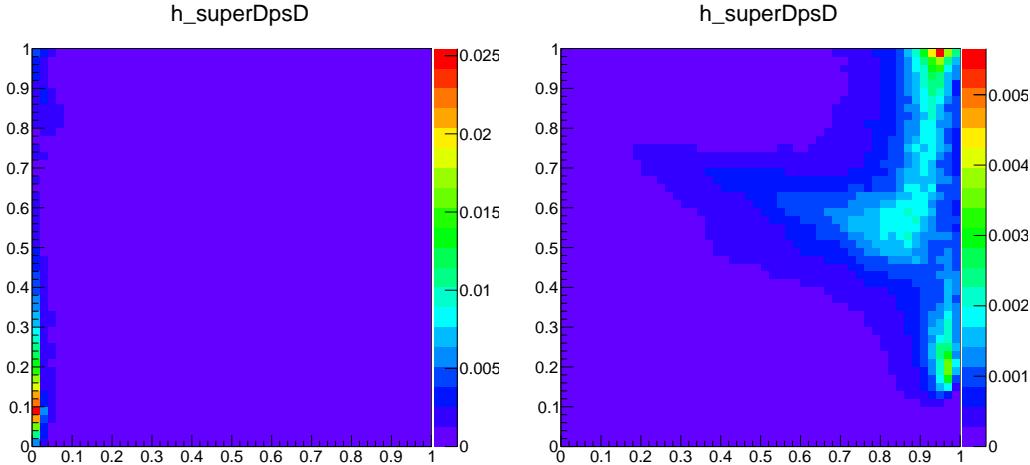


Figure 4.4: Distribution of D_{bkg}^{dec} vs. $D_{J^P}^{dec}$ for the $gg \rightarrow ZZ$ background (left) and for the signal resonance consistent with the SM Higgs boson $m_H = 125.6$ GeV model.

In Fig. 4.6 and 4.7 the D_{J^P} observables for all tested hypotheses are shown.

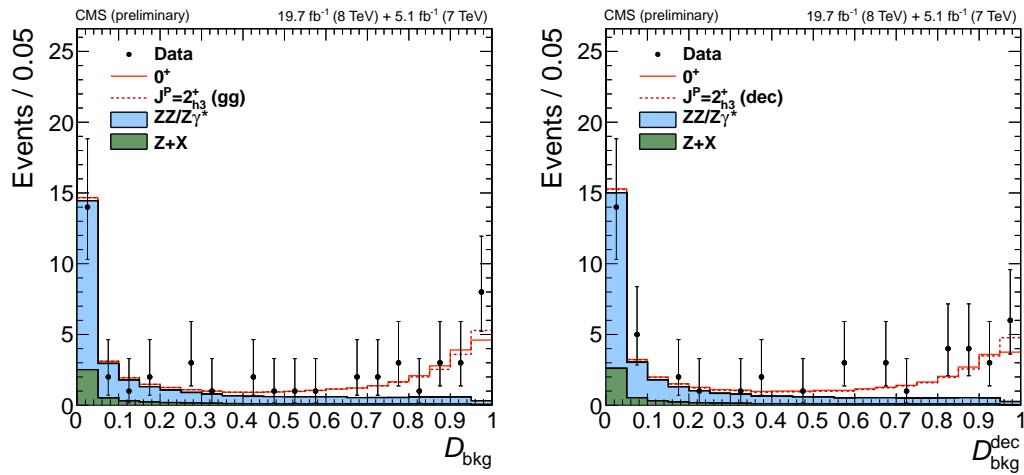


Figure 4.5: Distribution of D_{bkg} in data and MC expectations for the background

and for a signal resonance consistent with SM Higgs boson at $m_H = 125.6$ GeV (left).

The D_{bkg}^{dec} distribution for the production independent scenario (right).

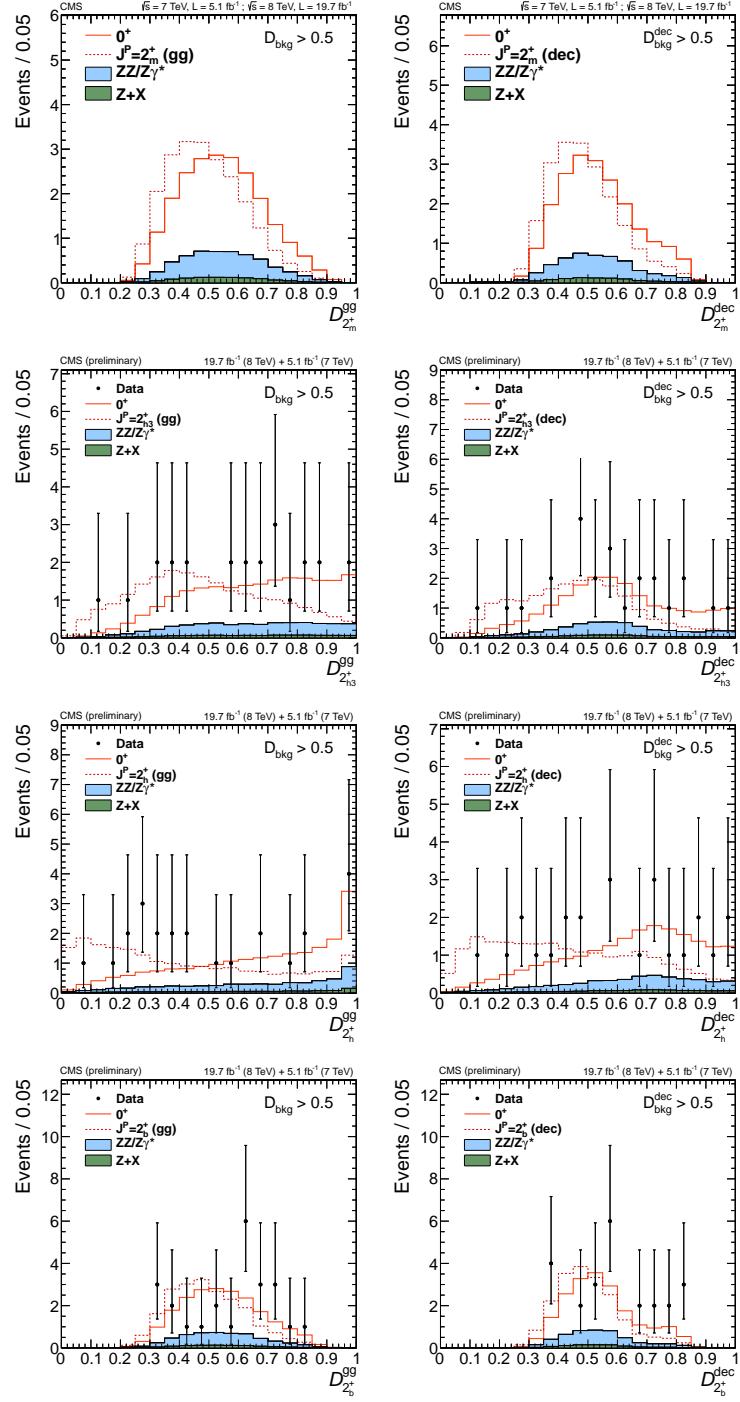
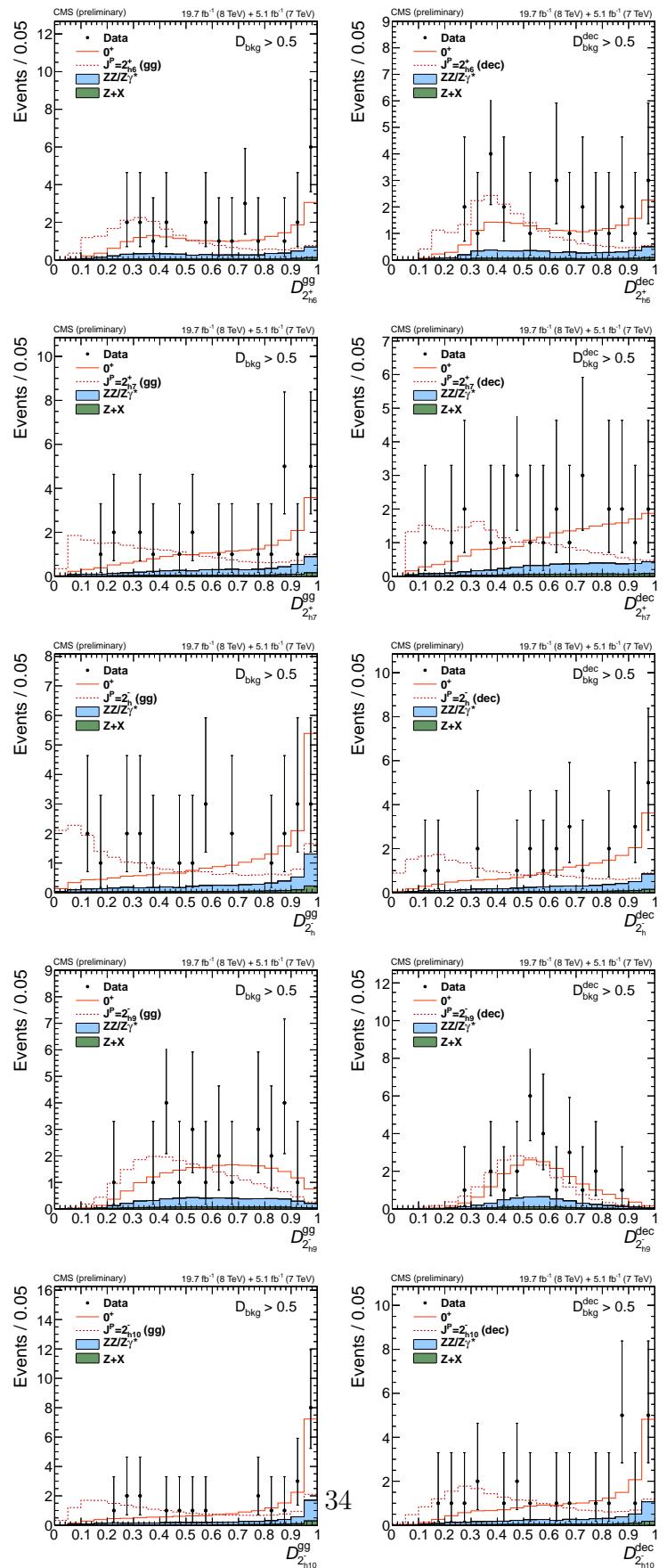


Figure 4.6: Distributions of D_{JP} with a requirement $D_{bkg} > 0.5$. Distributions for background and signal are shown.



4.6 Systematic Uncertainties

4.6.1 Theoretical uncertainties

The theoretical systematics considered as effects on the overall yield are the uncertainties related to the Higgs Branching Ratio to 4ℓ , QCD scale, and Parton Distribution Functions (PDF). All of these come from the latest calculations from the LHC Higgs Cross Section Working Group.⁵⁸ They are listed in Table 4.6.

4.6.2 Experimental uncertainties

The experimental systematics considered as effects on the overall yield are the uncertainties related to the luminosity, muon efficiency, electron efficiency, and the combination of various $Z + X$ uncertainties. While other systematics like lepton momentum scale, lepton resolution, production dependent MC, and $Z + X$ shape uncertainties are implemented as alternative shapes.

The muon efficiency systematic is applied to the 4μ and $2e2\mu$ final states (both signal and background), this uncertainty comes from the “tag-and-probe” techniques used to compute the overall efficiency in the muons. Similarly, the electron efficiency systematic is applied to the 4μ and $2e2\mu$ final states (both signal and background), and arises from a similar “tag-and-probe” technique for the electrons.

The $Z + X$ uncertainties include both the uncertainty on the expected yields and

the uncertainty on the shape. Yield uncertainties are estimated to be 20%, 25%, and 40% for the $4e$, $2e2\mu$, and 4μ decay channels, respectively. The shape uncertainty is taken into account by considering a difference between the $Z + X$ shape and the $q\bar{q} \rightarrow ZZ$ shape for a particular final state.

To account for the lepton momentum scale, and resolution uncertainty on the $m_{4\ell}$ distribution alternative signal shapes are taken from the variations of both of these contributions. Additionally, when performing the hypothesis testing for a production Independent hypothesis, another alternative shape is used to account for the MC coming from a specific production sample. For example, if a gg MC is used to produce the default templates then a $q\bar{q}$ sample is added as a systematic variation.

They are listed in Table 4.6. The "Production dependency" systematic is only applied when the hypothesis testing is used in a production-independent scenario.

Table 4.6: Summary of the systematics.

Systematic	$4e$ 8TeV(7TeV)	4μ 8TeV(7TeV)	$2e2\mu$ 8TeV(7TeV)
Electron efficiency	10%(11%)	N/A	5%(11%)
Muon efficiency	N/A	4%(4%)	3%(3%)
Control region	20%(20%)	40%(40%)	25%(25%)
Luminosity		2.6%(2.2%)	
$\text{BR}(H \rightarrow ZZ)$		2%(2%)	
Missing higher-orders $q\bar{q} \rightarrow ZZ$		3%(3%)	
Missing higher-orders $gg \rightarrow ZZ$		24%(24%)	
Missing higher-orders $gg \rightarrow H$		8%(7%)	
$\alpha_S + \text{PDF}(q\bar{q})$		3%(3%)	
$\alpha_S + \text{PDF}(gg)$		7%(7%)	
Signal acceptance		2%(2%)	
Lepton resolution & scale $2e2\mu$		Shape	
Lepton resolution & scale $4e$		Shape	
Lepton resolution & scale 4μ		Shape	
$Z + X$ shape		Shape	
Production dependency*		Shape	

4.7 Results

4.7.1 Testing of mixed spin-1 hypotheses

The expected separations for spin-one hypotheses using the test statistic distributions are shown in Figs. 4.8 and 4.9 and are summarized in Table 4.7 and in Fig. 4.10.

Table 4.7: A summary of models used in the analysis of spin-one hypotheses. Events are generated with SM expectation for the signal cross section ($\mu=1$).

J^P model	J^P production	Expected ($\mu=1$)	Obs.	0^+	Obs. J^P	CL_s	$f(J^P)$ CL=95% Obs(Exp)	$f(J^P)$ Best-Fit
1^-	$q\bar{q}$	2.9σ (2.8σ)		-1.4σ	$+5.0\sigma$	$<0.001\%$	<0.46 (0.78)	$0.00^{+0.16}_{-0.00}$
$f_{b2} = 0.2$	$q\bar{q}$	2.6σ (2.6σ)		-1.4σ	$+4.6\sigma$	0.002%	<0.49 (0.81)	$0.00^{+0.17}_{-0.00}$
$f_{b2} = 0.4$	$q\bar{q}$	2.5σ (2.4σ)		-1.3σ	$+4.4\sigma$	0.005%	<0.51 (0.83)	$0.00^{+0.19}_{-0.00}$
$f_{b2} = 0.6$	$q\bar{q}$	2.4σ (2.4σ)		-1.2σ	$+4.1\sigma$	0.015%	<0.53 (0.83)	$0.00^{+0.20}_{-0.00}$
$f_{b2} = 0.8$	$q\bar{q}$	2.4σ (2.4σ)		-1.0σ	$+4.0\sigma$	0.021%	<0.55 (0.83)	$0.00^{+0.21}_{-0.00}$
1^+	$q\bar{q}$	2.4σ (2.4σ)		-0.8σ	$+3.8\sigma$	0.031%	<0.57 (0.81)	$0.00^{+0.22}_{-0.00}$
1^-	any	2.9σ (2.7σ)		-2.0σ	$>5.0\sigma$	$<0.001\%$	<0.37 (0.79)	$0.00^{+0.12}_{-0.00}$
$f_{b2} = 0.2$	any	2.7σ (2.5σ)		-2.2σ	$>5.0\sigma$	$<0.001\%$	<0.38 (0.82)	$0.00^{+0.12}_{-0.00}$
$f_{b2} = 0.4$	any	2.5σ (2.4σ)		-2.3σ	$>5.0\sigma$	$<0.001\%$	<0.39 (0.84)	$0.00^{+0.13}_{-0.00}$
$f_{b2} = 0.6$	any	2.5σ (2.3σ)		-2.4σ	$>5.0\sigma$	$<0.001\%$	<0.39 (0.86)	$0.00^{+0.13}_{-0.00}$
$f_{b2} = 0.8$	any	2.4σ (2.3σ)		-2.3σ	$>5.0\sigma$	$<0.001\%$	<0.40 (0.86)	$0.00^{+0.13}_{-0.00}$
1^+	any	2.5σ (2.3σ)		-2.3σ	$>5.0\sigma$	$<0.001\%$	<0.41 (0.85)	$0.00^{+0.13}_{-0.00}$

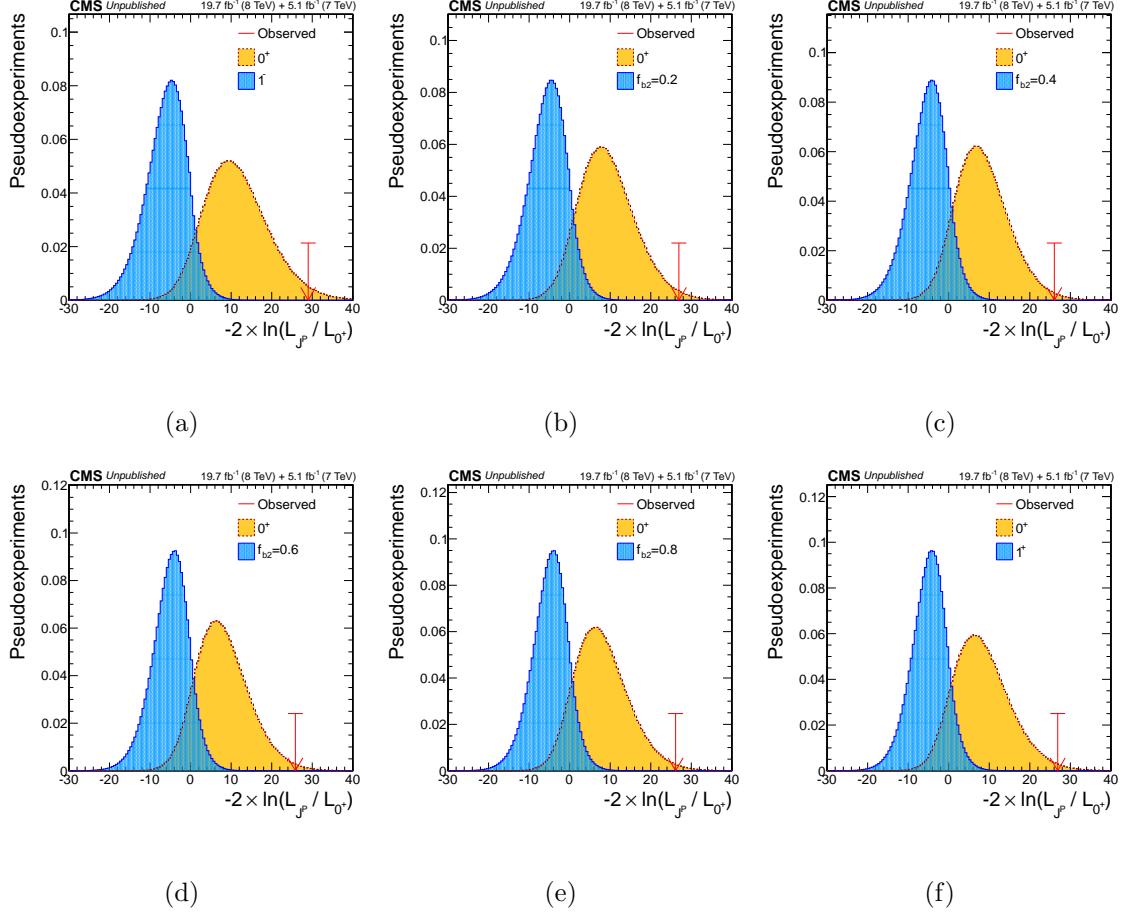


Figure 4.8: Distribution of the test-statistic $q = -2 \ln(\mathcal{L}_{JP}/\mathcal{L}_{0+})$ of the spin-1 hypothesis tested against the SM Higgs boson hypothesis, for the production independent scenario. $J^P = \text{pure } 1^-$ (top left), mixture with $f_{b2}=0.2$ (top middle), $f_{b2}=0.4$ (top right), $f_{b2}=0.6$ (bottom left), $f_{b2}=0.8$ (bottom middle), and pure 1^+ (bottom right). Distributions for the SM Higgs boson are represented by the yellow histogram and for the alternative J^P hypotheses by the blue histogram.

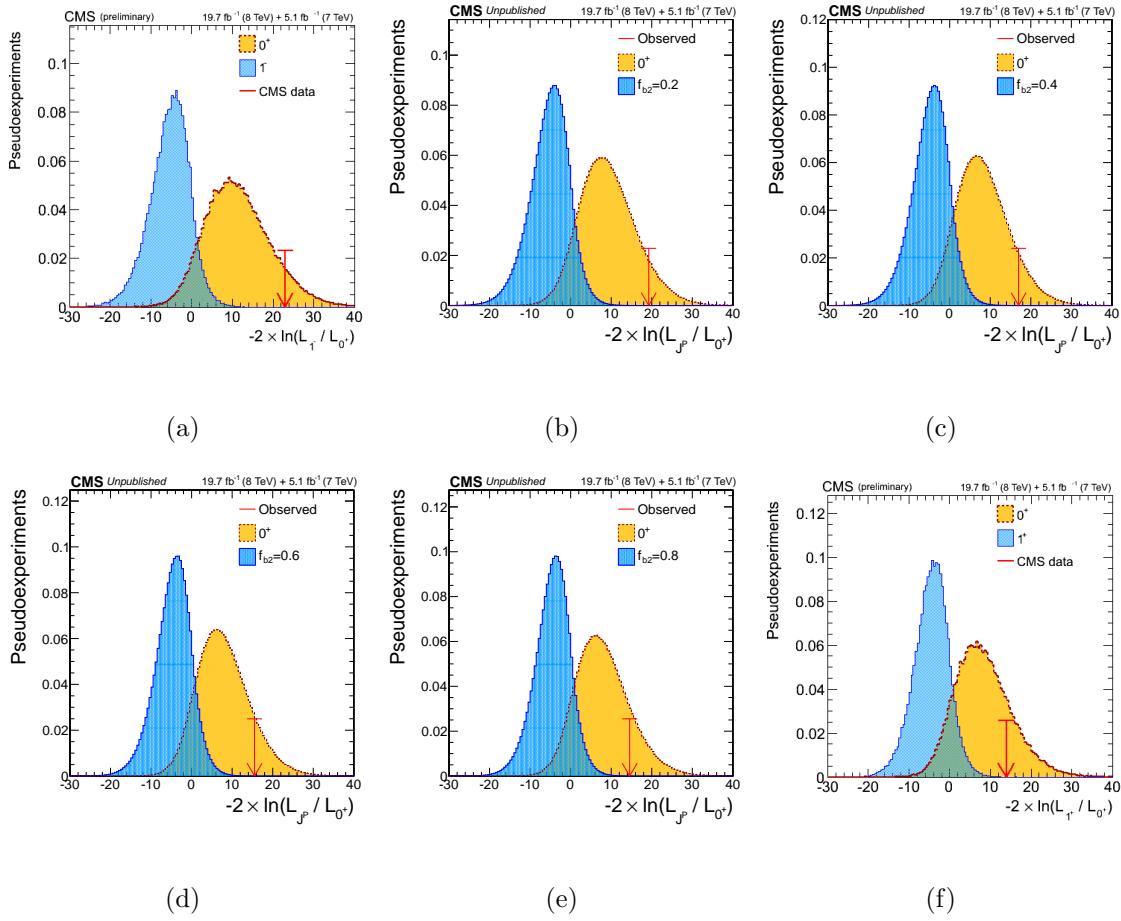


Figure 4.9: Distribution of the test-statistic $q = -2 \ln(\mathcal{L}_{JP}/\mathcal{L}_{0+})$ of the spin-1 hypothesis tested against the SM Higgs boson hypothesis, for the production dependent scenario. $J^P = \text{pure } 1^-$ (top left), mixture with $f_{b2}=0.2$ (top middle), $f_{b2}=0.4$ (top right), $f_{b2}=0.6$ (bottom left), $f_{b2}=0.8$ (bottom middle), and pure 1^+ (bottom right). Distributions for the SM Higgs boson are represented by the yellow histogram and for the alternative J^P hypotheses by the blue histogram.

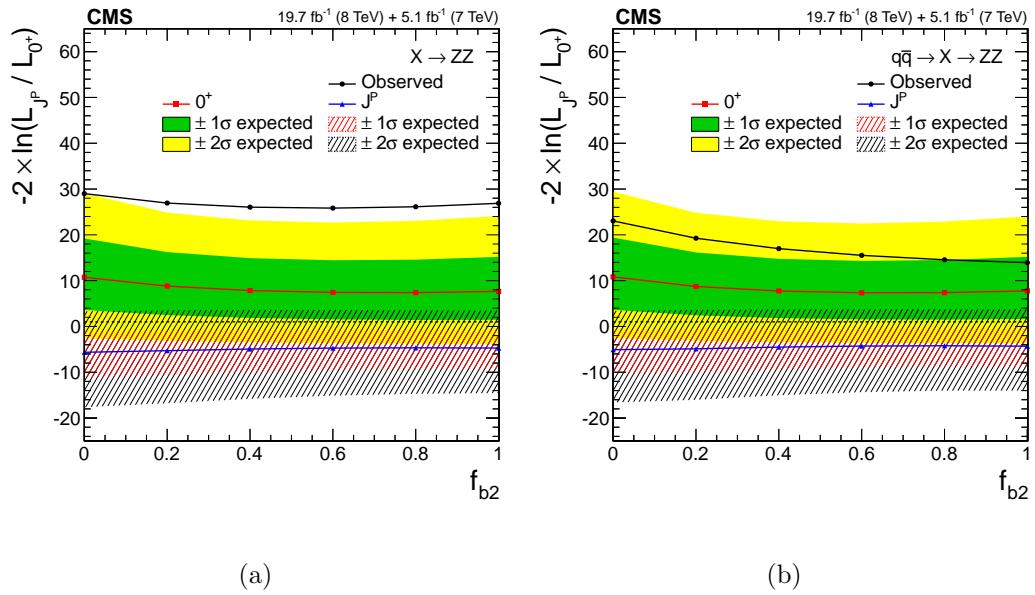


Figure 4.10: Distribution of median test-statistic q , as a function of f_{b2} . The green(red) and blue(black) bands represent the 1σ and 2σ around the median expected value for the SM(spin-1) Higgs boson hypothesis.

4.7.2 Testing of spin-2 hypotheses

The expected and observed test statistics q is shown in Figs. 4.12 for all the hypotheses tested. The results are summarized in the Table 4.8. The expected separation is quoted for two scenarios, where the signal strength for each hypothesis is predetermined from the fit to data and where events are generated with SM expectations for the signal cross section ($\mu=1$). The observed separation quotes consistency of the observation with the 0^+ model or J^P model and corresponds to the scenario where the signal strength is floated in the fit to data. Table also quotes the CL_s value for the J^P model. The last three columns quote the expected and observed limit at 95%C.L. on the fractional presence of J^P model as a state nearly degenerate with the 0^+ state.

The expected separations from the test statistic distributions for all the considered models are also summarized in Fig. 4.13.

It can be appreciated that the data has disfavored all tested spin-two hypotheses in favor of SM hypothesis 0^+ with CL_s value larger than 95%C.L.

J^P model	J^P prod.	Expected ($\mu=1$)	Obs. 0 ⁺	Obs. J^P	CL_s	$f(J^P)$ obs. (exp.)	95% CL	$f(J^P)$ best fit
2_m^+ (legacy)	gg	1.9σ (1.8σ)	-1.1σ	$+3.0\sigma$	0.90%	<0.71 (1.00)	$0.00^{+0.30}_{-0.00}$	
2_{h2}^+	gg	2.0σ (2.1σ)	-0.3σ	$+2.4\sigma$	2.0%	<0.85 (0.89)	$0.09^{+0.39}_{-0.09}$	
2_{h3}^+	gg	3.2σ (3.4σ)	$+0.3\sigma$	$+3.0\sigma$	0.17%	<0.72 (0.58)	$0.13^{+0.29}_{-0.13}$	
2_h^+ (legacy)	gg	3.8σ (4.0σ)	$+1.8\sigma$	$+2.0\sigma$	2.3%	<1.00 (0.52)	$0.48^{+0.24}_{-0.29}$	
2_b^+ (legacy)	gg	1.6σ (1.8σ)	-1.4σ	$+3.4\sigma$	0.50%	<0.64 (1.00)	$0.00^{+0.24}_{-0.00}$	
2_{h6}^+	gg	3.4σ (3.7σ)	-0.6σ	$+4.9\sigma$	$<0.001\%$	<0.38 (0.58)	$0.00^{+0.13}_{-0.00}$	
2_{h7}^+	gg	3.8σ (4.5σ)	-0.3σ	$+4.5\sigma$	$<0.001\%$	<0.44 (0.43)	$0.00^{+0.19}_{-0.00}$	
2_h^- (legacy)	gg	4.2σ (4.5σ)	$+1.0\sigma$	$+3.2\sigma$	0.090%	<0.77 (0.47)	$0.29^{+0.21}_{-0.23}$	
2_{h9}^-	gg	2.5σ (2.6σ)	-1.1σ	$+4.0\sigma$	0.029%	<0.46 (0.76)	$0.00^{+0.15}_{-0.00}$	
2_{h10}^-	gg	4.2σ (4.3σ)	-0.1σ	$+4.8\sigma$	$<0.001\%$	<0.57 (0.50)	$0.06^{+0.27}_{-0.06}$	
2_m^+ (legacy)	$\bar{q}q$	1.7σ (1.7σ)	-1.7σ	$+3.8\sigma$	0.17%	<0.56 (1.00)	$0.00^{+0.19}_{-0.00}$	
2_{h2}^+	$\bar{q}q$	2.2σ (2.2σ)	-0.8σ	$+3.3\sigma$	0.26%	<0.61 (0.86)	$0.00^{+0.23}_{-0.00}$	
2_{h3}^+	$\bar{q}q$	3.1σ (3.0σ)	$+0.2\sigma$	$+3.0\sigma$	0.21%	<0.81 (0.70)	$0.13^{+0.40}_{-0.13}$	
2_h^+	$\bar{q}q$	4.0σ (3.9σ)	$+0.2\sigma$	$+3.9\sigma$	0.008%	<0.71 (0.53)	$0.21^{+0.28}_{-0.21}$	
2_b^+	$\bar{q}q$	1.7σ (1.7σ)	-1.9σ	$+4.1\sigma$	0.062%	<0.45 (1.00)	$0.00^{+0.14}_{-0.00}$	
2_{h6}^+	$\bar{q}q$	3.4σ (3.3σ)	-0.2σ	$+4.0\sigma$	0.008%	<0.74 (0.71)	$0.04^{+0.45}_{-0.04}$	
2_{h7}^+	$\bar{q}q$	4.1σ (3.9σ)	$+0.4\sigma$	$+3.8\sigma$	0.010%	<0.77 (0.55)	$0.35^{+0.23}_{-0.28}$	
2_h^-	$\bar{q}q$	4.3σ (4.4σ)	$+0.0\sigma$	$+4.6\sigma$	$<0.001\%$	<0.57 (0.48)	$0.01^{+0.31}_{-0.01}$	
2_{h9}^-	$\bar{q}q$	2.4σ (2.2σ)	$+0.5\sigma$	$+2.0\sigma$	3.1%	<0.99 (0.86)	$0.31^{+0.43}_{-0.31}$	
2_{h10}^-	$\bar{q}q$	4.0σ (3.9σ)	$+0.4\sigma$	$+4.0\sigma$	0.006%	<0.75 (0.59)	$0.30^{+0.26}_{-0.30}$	
2_m^+ (legacy)	any	1.5σ (1.5σ)	-1.6σ	$+3.4\sigma$	0.71%	<0.63 (1.00)	$0.00^{+0.22}_{-0.00}$	
2_{h2}^+	any	1.9σ (2.0σ)	-0.9σ	$+3.0\sigma$	0.74%	<0.66 (0.95)	$0.00^{+0.27}_{-0.00}$	
2_{h3}^+	any	3.0σ (3.1σ)	$+0.0\sigma$	$+3.1\sigma$	0.18%	<0.69 (0.64)	$0.00^{+0.35}_{-0.00}$	
2_h^+	any	3.8σ (4.0σ)	$+0.3\sigma$	$+3.6\sigma$	0.025%	<0.64 (0.49)	$0.07^{+0.30}_{-0.07}$	
2_b^+	any	1.7σ (1.7σ)	-1.6σ	$+3.6\sigma$	0.29%	<0.55 (1.00)	$0.00^{+0.19}_{-0.00}$	
2_{h6}^+	any	3.3σ (3.4σ)	-0.3σ	$+4.2\sigma$	0.003%	<0.54 (0.62)	$0.00^{+0.23}_{-0.00}$	
2_{h7}^+	any	4.0σ (4.2σ)	$+0.6\sigma$	$+3.5\sigma$	0.032%	<0.70 (0.47)	$0.17^{+0.28}_{-0.17}$	
2_h^-	any	4.2σ (4.6σ)	-0.2σ	$+4.8\sigma$	$<0.001\%$	<0.48 (0.43)	$0.04^{+0.21}_{-0.04}$	
2_{h9}^-	any	2.2σ (2.1σ)	-0.6σ	$+2.9\sigma$	0.57%	<0.69 (0.89)	$0.00^{+0.27}_{-0.00}$	
2_{h10}^-	any	3.9σ (4.0σ)	$+0.1\sigma$	$+4.3\sigma$	0.002%	<0.61 (0.54)	$0.08^{+0.30}_{-0.08}$	

Table 4.8: List of models used in the analysis of the spin-two hypotheses corresponding to the pure states of the type noted.

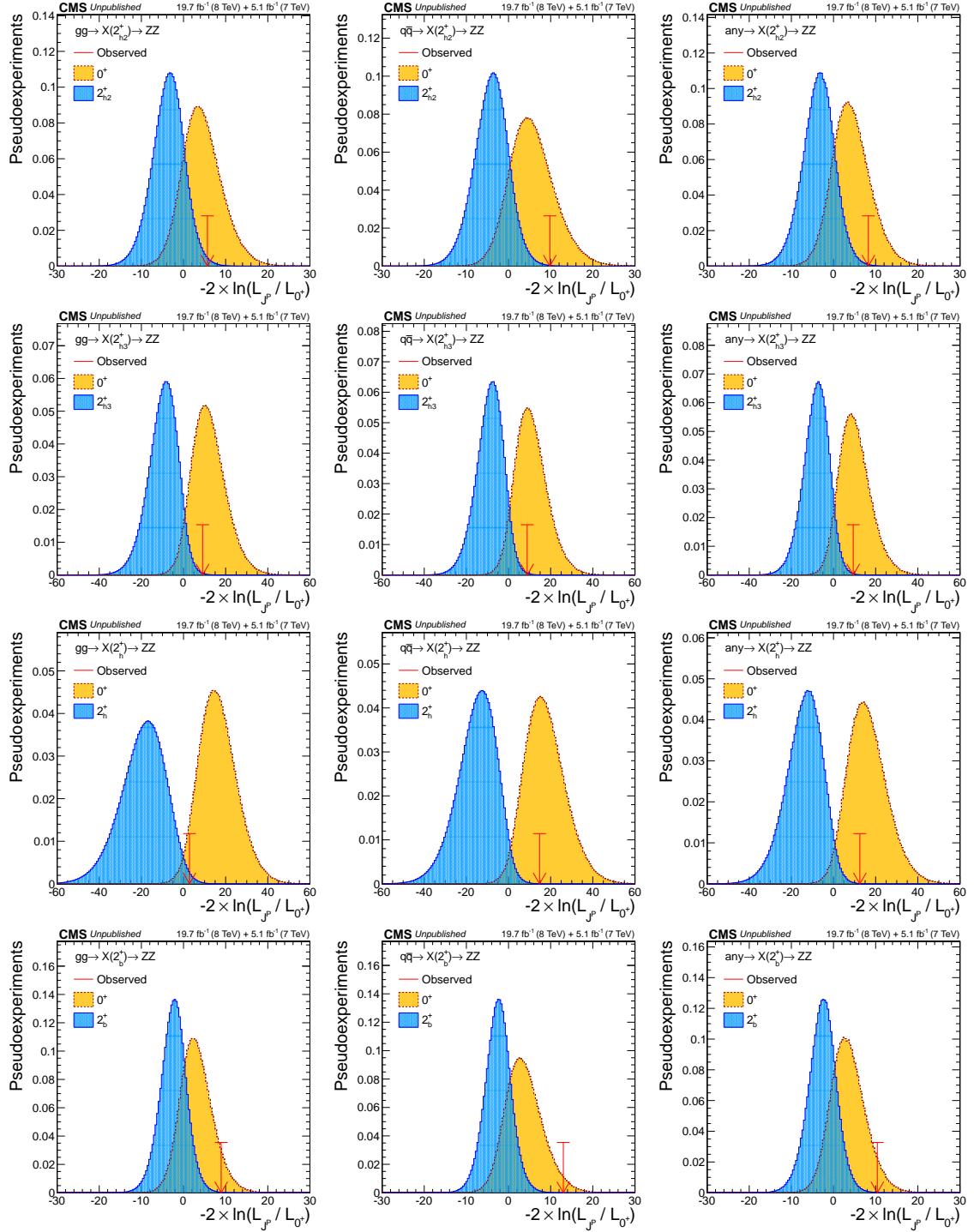


Figure 4.11: Distribution of test-statistics $q = -2 \ln(\mathcal{L}_{JP}/\mathcal{L}_{0+})$ for two signal types (0^+ represented by the yellow histogram and alternative hypothesis by the blue histogram) for $m_H = 125.6$ GeV shown with a large number of generated experiments.

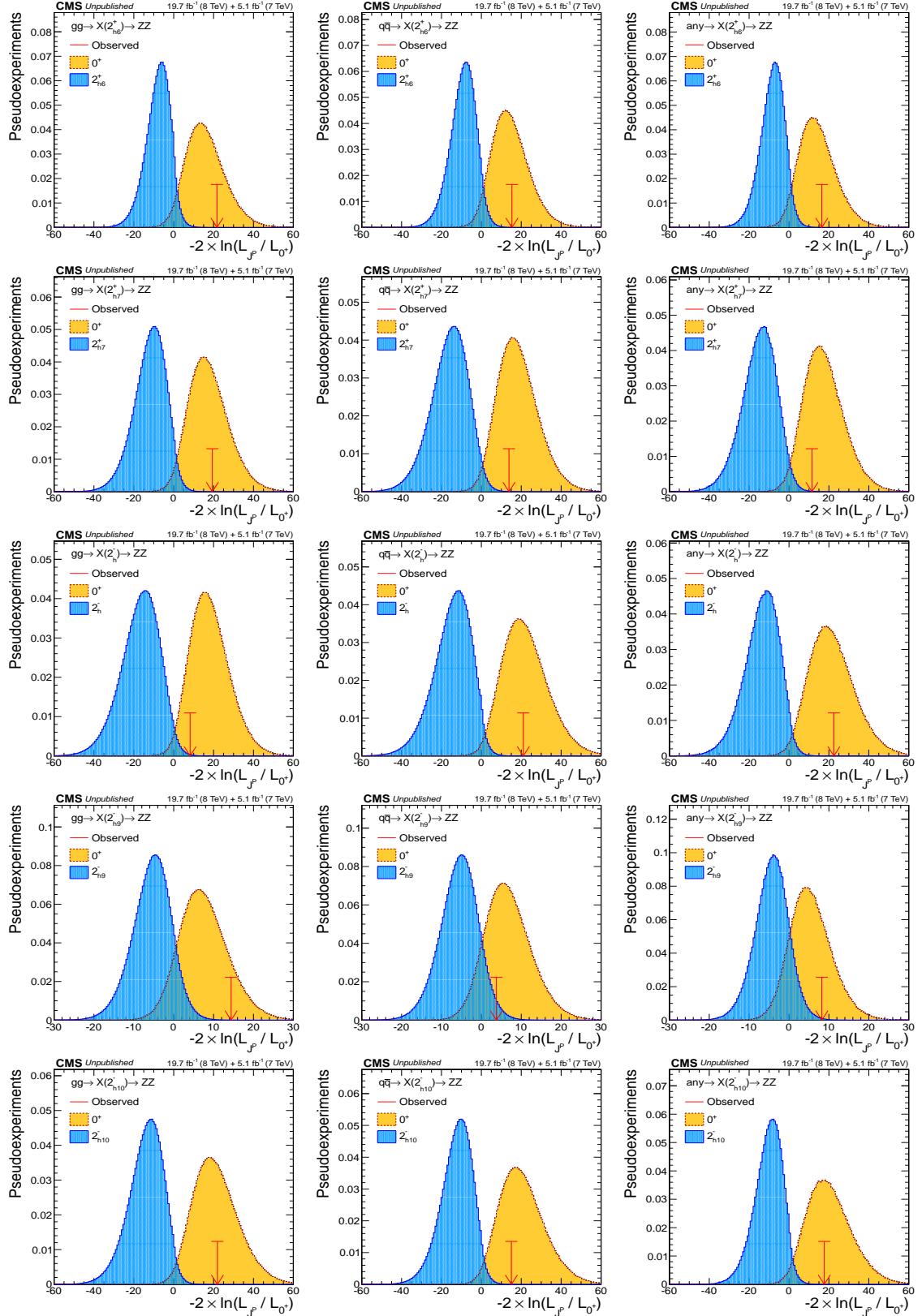


Figure 4.12: Distribution of $q = -2 \ln(\mathcal{L}_{J^F} / \mathcal{L}_{0+})$ for two signal types (0^+ represented

by the yellow histogram and alternative hypothesis by the blue histogram) for m_H

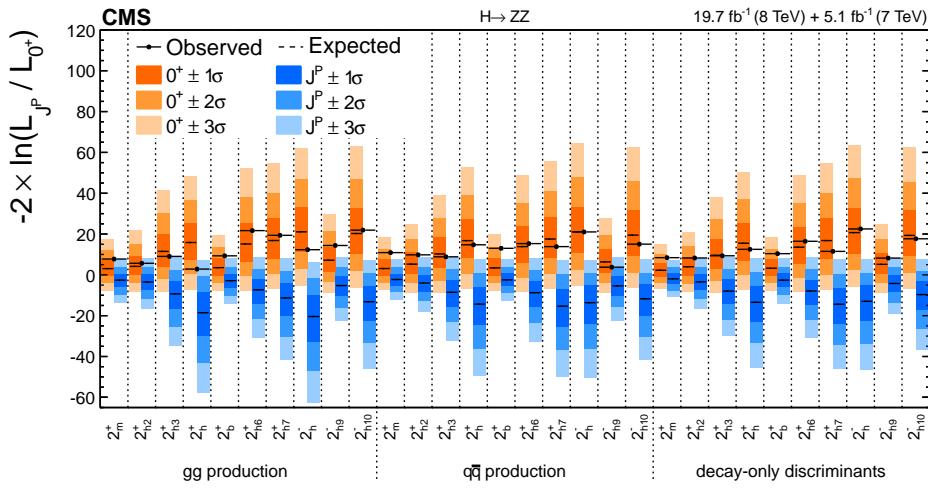


Figure 4.13: Summary of the expected and observed values for the test-statistic q distributions for the twelve alternative hypotheses tested with respect to the SM Higgs boson. The orange (blue) bands represent the 1σ , 2σ , and 3σ around the median expected value for the SM Higgs boson hypothesis (alternative hypothesis). The black point represents the observed value.

4.7.3 Constrains on mixed non-interfering states

Figures 4.14 and 4.15 show the likelihood scans of $f(J^P)$ for the spin-1 hypotheses, in the production-independent and $q\bar{q}$ cases, respectively. The results are also summarize in Table 4.7.

Figures 4.17, 4.18, and 4.19 show the likelihood scans of $f(J^P)$ for the spin-2 hypotheses, in the production-independent, gg , and $\bar{q}q$ cases, respectively.

The observed non-interfering fraction measurements are also summarized in Fig. 4.16 and Fig. 4.20.

These results are consistent with the expected SM contribution to the signal. Each of these fractions is tested and reported independently of the other hypotheses, but the reader should note that there are correlations between the various alternate hypotheses.

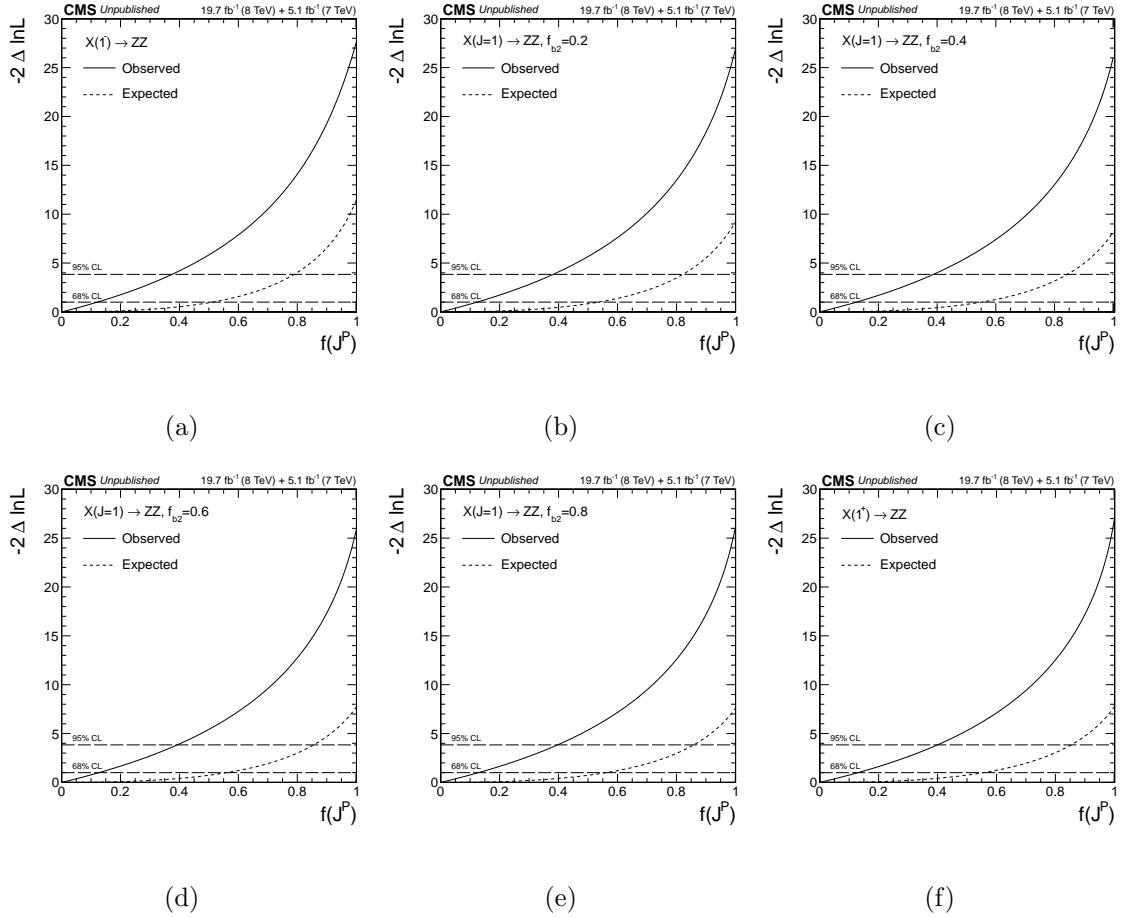


Figure 4.14: Expected and observed distributions of $-2\Delta \ln \mathcal{L}$ as a function of $f(J^P)$, in the production independent (any $\rightarrow J^P$) case. $J^P =$ pure 1^- (top left), mixture with $f_{b2}=0.2$ (top middle), $f_{b2}=0.4$ (top right), $f_{b2}=0.6$ (bottom left), $f_{b2}=0.8$ (bottom middle), and pure 1^+ (bottom right). The horizontal lines at $-2\Delta \ln \mathcal{L} = 1$ and 3.84 represent the 68% and 95% C.L., respectively.

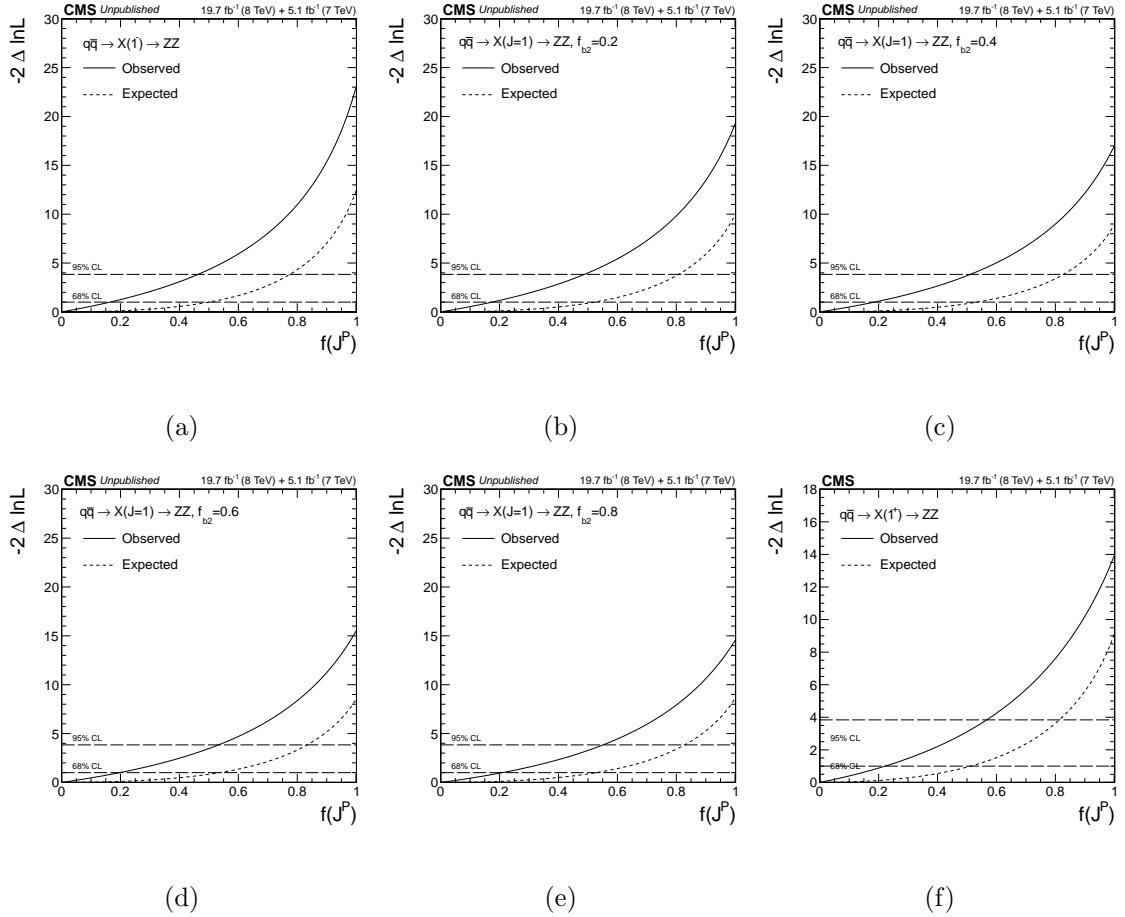


Figure 4.15: Expected and observed distributions of $-2\Delta \ln \mathcal{L}$ as a function of $f(J^P)$, in the $\bar{q}q$ case. $J^P = \text{pure } 1^-$ (top left), mixture with $f_{b2}=0.2$ (top middle), $f_{b2}=0.4$ (top right), $f_{b2}=0.6$ (bottom left), $f_{b2}=0.8$ (bottom middle), and pure 1^+ (bottom right). The horizontal lines at $-2\Delta \ln \mathcal{L} = 1$ and 3.84 represent the 68% and 95% C.L., respectively.

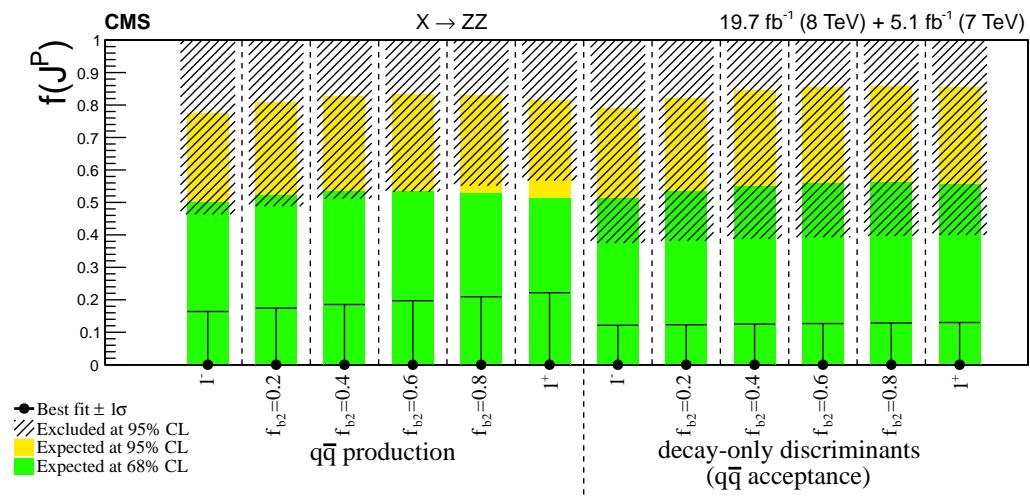


Figure 4.16: Summary of expected and observed constraints on the non-interfering fraction measurements for spin-one.

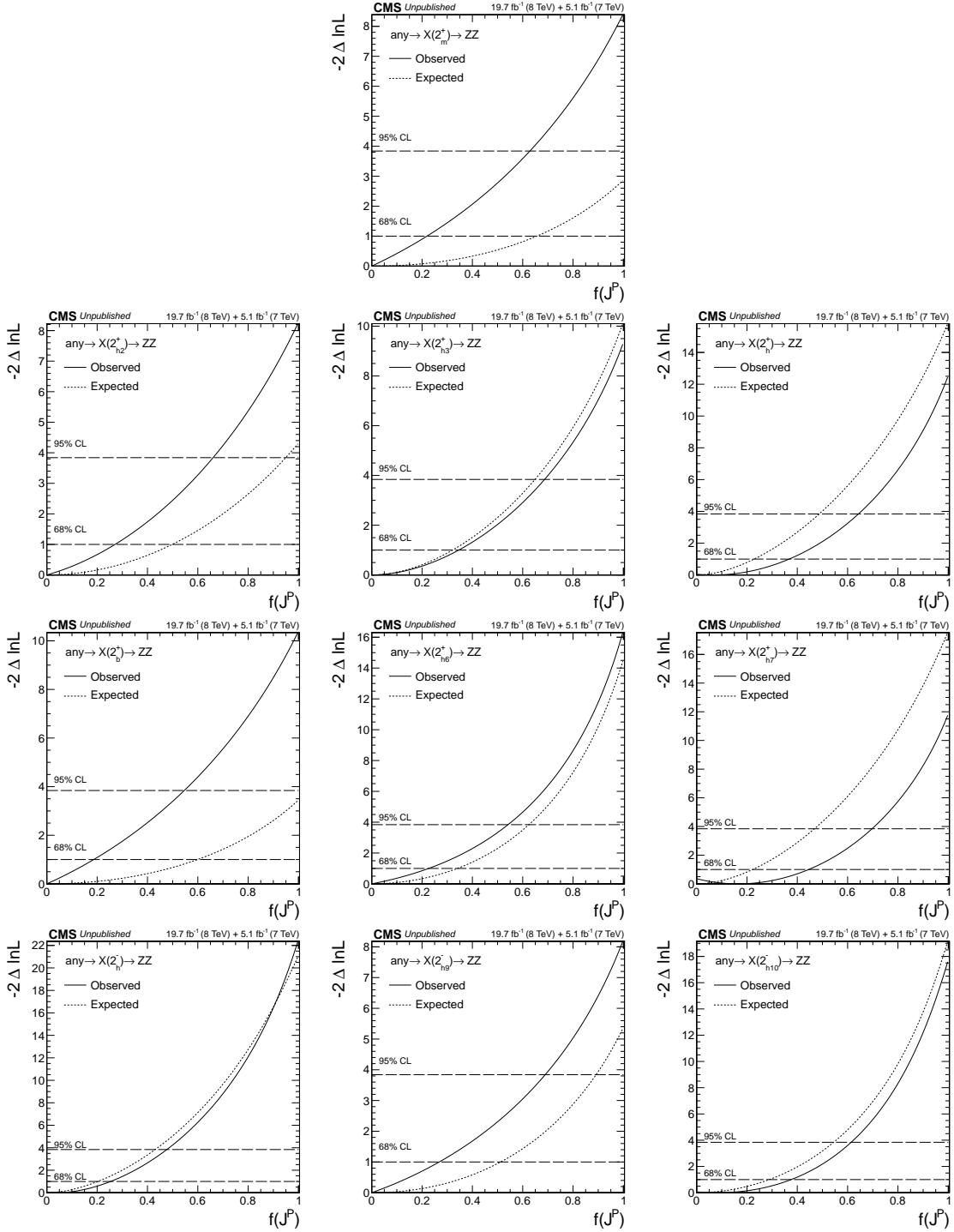


Figure 4.17: Non-interfering fraction expectations for various spin-2 models. The

above results are production-independent.

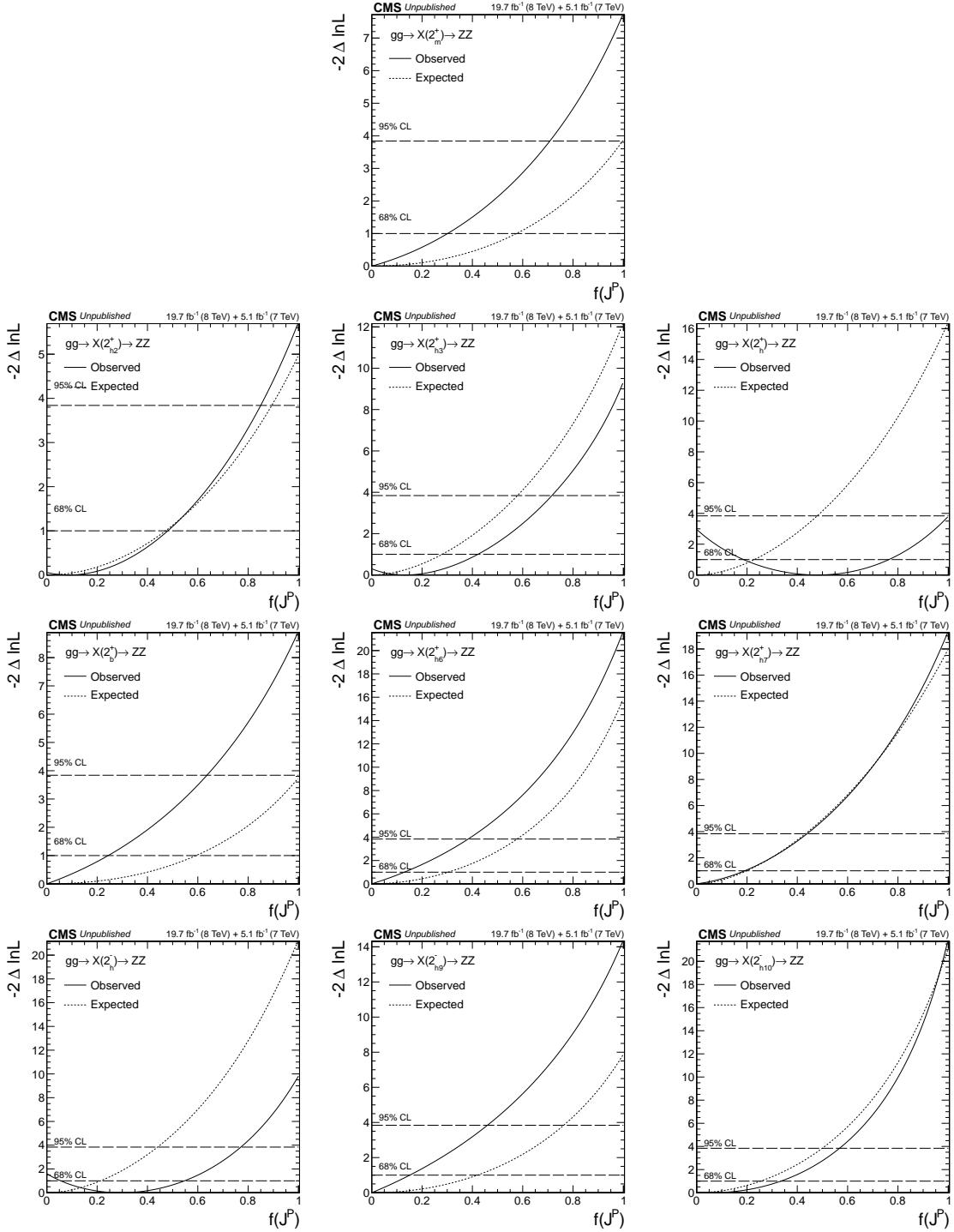


Figure 4.18: Non-interfering fraction expectations for various spin-2 models ($gg \rightarrow J^P$).

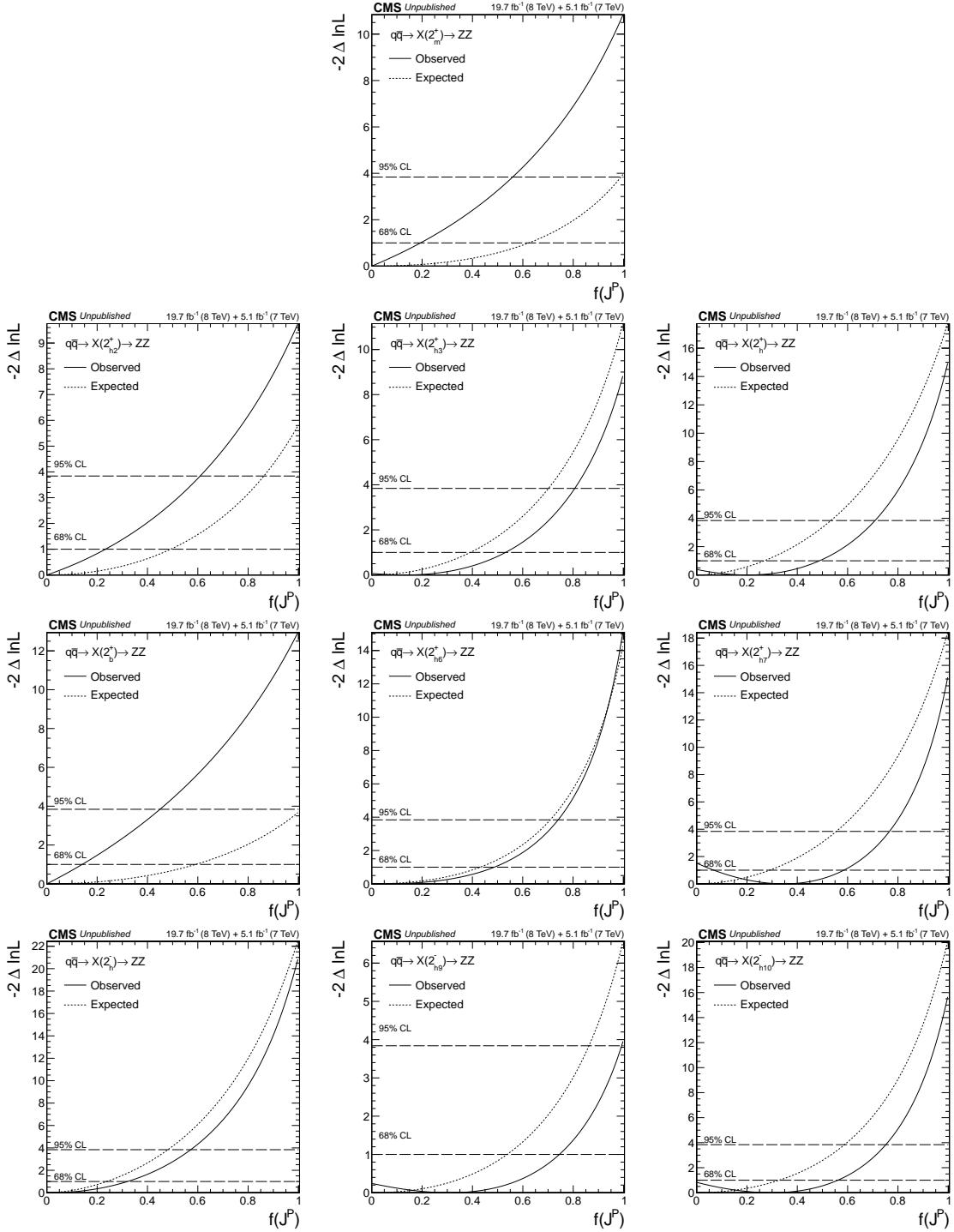


Figure 4.19: Non-interfering fraction expectations for various spin-2 models ($\bar{q}q \rightarrow J^P$).

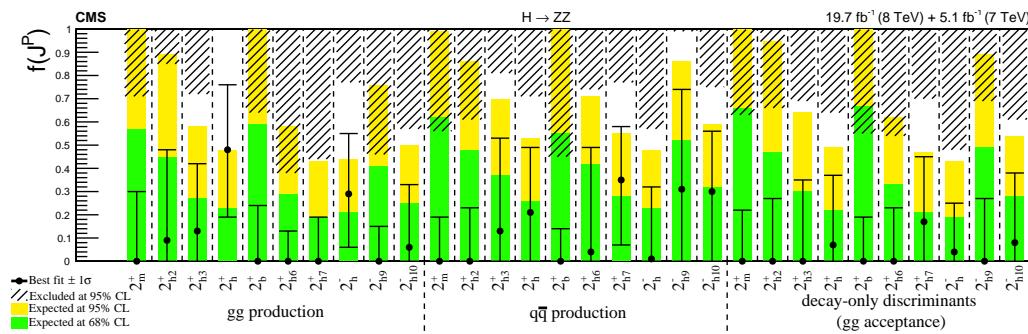


Figure 4.20: Summary of expected and observed constraints on the non-interfering fraction measurements for spin-two.

4.8 Summary

The tested spin-two boson hypotheses are excluded at a 95% C.L. or higher, while any mixture of a vector and a pseudo-vector state is excluded at 99% C.L. or higher.

4.9 $H \rightarrow ZZ \rightarrow 4l$

4.9.1 Datasets

4.9.1.1 Experimental data

The dataset used are pp collision data recorded in 2016 by the CMS detector corresponding to an integrated luminosity of 12.9fb^{-1} .

Collision events are selected by high-level trigger algorithms that require the presence of leptons passing loose identification and isolation requirements. The main triggers of this analysis select either a pair of electrons or muons, or an electron and a muon. The minimal transverse momentum of the leading electron(muon) is 23(17) GeV, while that of the subleading lepton is 12(8) GeV. To maximize the coverage of the $H \rightarrow 4\ell$ phase space, triggers requiring three leptons with relaxed p_T thresholds and no isolation requirement are also used, as are isolated single-electron and single-muon triggers with thresholds of 27GeV and 22GeV, respectively.

The overall trigger efficiency for simulated signal events that pass the full selection chain of this analysis is larger than 99%. The trigger efficiency is measured in data using a sample of 4ℓ events collected by the single-lepton triggers and is found to be in agreement with the expectation from simulation.

4.9.1.2 Simulated samples

Monte Carlo (MC) simulation samples for the signals and the relevant background processes are used to estimate backgrounds, optimize the event selection, and evaluate the acceptance and systematic uncertainties.

The SM Higgs boson signals are generated at next-to-leading order (NLO) in perturbative QCD (pQCD) with the POWHEG2.0^{10,39,62} generator for the five main production modes: gluon fusion ($gg \rightarrow H$), vector boson fusion ($qq \rightarrow qqH$) and associated production (WH , ZH and $t\bar{t}H$). For WH and ZH the MINLO HVJ⁶⁰ extension of POWHEG2.0 is used. The cross section for the dominant gluon fusion production mode is taken from Ref.¹³ The default parton distribution function (PDF) used in all simulations is NNPDF30_nlo_as_0118.¹⁸ The decay of the Higgs boson to four leptons is modeled with JHUGEN.^{21,42} In the case of ZH and $t\bar{t}H$, the Higgs boson is also allowed to decay as $H \rightarrow ZZ \rightarrow 2\ell 2X$ where X stands either for a quark or a neutrino, thus accounting for four-lepton events where two leptons originate from the decay of the associated Z boson or top quarks. The list of signal samples and their cross sections are shown in Table 4.9.

The SM ZZ background contribution from quark-antiquark annihilation is generated at NLO pQCD with POWHEG2.0, while the $gg \rightarrow ZZ$ process is generated at leading order (LO) with MCFM.²⁵ We discuss scale factors that account for missing higher-order corrections (K factors) in Section ???. All signal and background generators are interfaced with PYTHIA8⁶⁵ to simulate the multi-parton interaction and

Process	Dataset Name	$\sigma \times BR(\times \epsilon_{\text{filter}})$
$gg \rightarrow H \rightarrow ZZ \rightarrow 4\ell$	/GluGluHToZZTo4L_M125_13TeV_powheg2_JHUGenV6_pythia8/[1]	12.18 fb
$qq \rightarrow Hqq \rightarrow ZZqq \rightarrow 4\ell qq$	/VBF_HToZZTo4L_M125_13TeV_powheg2_JHUGenV6_pythia8/[1]	1.044 fb
$q\bar{q} \rightarrow W^+H \rightarrow W^+ZZ \rightarrow 4\ell + X$	/WplusH_HToZZTo4L_M125_13TeV_powheg2-minlo-HWJ_JHUGenV6_pythia8/[1]	0.232 fb
$q\bar{q} \rightarrow W^-H \rightarrow W^-ZZ \rightarrow 4\ell + X$	/WminusH_HToZZTo4L_M125_13TeV_powheg2-minlo-HWJ_JHUGenV6_pythia8/[1]	0.147 fb
$q\bar{q} \rightarrow ZH \rightarrow ZZZ \rightarrow 4\ell + X$	/ZH_HToZZ_4LFilter_M125_13TeV_powheg2-minlo-HZJ_JHUGenV6_pythia8/[1]	0.668 fb
$gg \rightarrow ttH \rightarrow ttZZ \rightarrow 4\ell + X$	/ttH_HToZZ_4LFilter_M125_13TeV_powheg_JHUGen_pythia8/[1]	0.393 fb

[1] RunIISpring16MiniAODv*-PUSpring16_80X_mcRun2_asymptotic_2016_v3-v*

Table 4.9: Signal Monte Carlo samples and cross sections.

hadronization effects. The generated events are processed through a detailed simulation of the CMS detector based on GEANTfour^{9,12} and are reconstructed with the same algorithms that are used for data. The simulated events include overlapping pp interactions (pileup) and have been reweighted so that the distribution of the number of interactions per LHC bunch crossing in simulation matches that observed in data. The list of background samples and their cross sections are shown in Table 4.10.

Process	Dataset Name	$\sigma \cdot BR$
$qq \rightarrow ZZ \rightarrow 4\ell$	/ZZTo4L_13TeV_powheg_pythia8/[1]	1.256pb
$qq \rightarrow ZZ \rightarrow 4\ell$	/ZZTo4L_13TeV-amcatnloFXFX-pythia8/[1]	1.212pb
$gg \rightarrow ZZ \rightarrow 4e$	/GluGluToContinToZZTo4e_13TeV_MCFM701/[1]	0.00159pb
$gg \rightarrow ZZ \rightarrow 4\mu$	/GluGluToContinToZZTo4mu_13TeV_MCFM701/[1]	0.00159pb
$gg \rightarrow ZZ \rightarrow 4\tau$	/GluGluToContinToZZTo4tau_13TeV_MCFM701/[1]	0.00159pb
$gg \rightarrow ZZ \rightarrow 2e2\mu$	/GluGluToContinToZZTo2e2mu_13TeV_MCFM701/[1]	0.00319pb
$gg \rightarrow ZZ \rightarrow 2e2\tau$	/GluGluToContinToZZTo2e2tau_13TeV_MCFM701/[1]	0.00319pb
$gg \rightarrow ZZ \rightarrow 2\mu2\tau$	/GluGluToContinToZZTo2mu2tau_13TeV_MCFM701/[1]	0.00319pb
$Z \rightarrow \ell\ell + \text{jets}$	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/[1]	6104pb
$Z \rightarrow \ell\ell + \text{jets}$	/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/[1]	18610pb
$WZ \rightarrow 3\ell\nu$	/WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8/[1]	4.430pb
$t\bar{t}$	/TTJets_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/[1]	815.96pb
$t\bar{t} \rightarrow 2\ell 2\nu 2b$	/TTTo2L2Nu_13TeV-powheg/[1]	87.31pb

[1] RunIISpring16MiniAODv*-PUSpring16RAWAODSIM_80X_mcRun2_asymptotic_2016_v3-v*

Table 4.10: Background Monte Carlo samples and cross sections.

4.9.2 Event Selection

The event selection is designed to extract signal candidates from events containing at least four well-identified and isolated leptons, each originating from the primary vertex and possibly accompanied by an FSR photon candidate, as described in Section ???. In what follows, unless otherwise stated, FSR photons are included in invariant mass computations.

First, Z candidates are formed with pairs of leptons of the same flavor and opposite-charge (e^+e^- , $\mu^+\mu^-$) and required to pass $12 < m_{\ell^+\ell^-} < 120$ GeV. They are then combined into ZZ candidates, wherein we denote as Z_1 the Z candidate with an invariant mass closest to the nominal Z boson mass, and as Z_2 the other one. The flavors of involved leptons define three mutually exclusive subchannels: $4e$, 4μ and $2e2\mu$.

To be considered for the analysis, ZZ candidates have to pass a set of kinematic requirements that improve the sensitivity to Higgs boson decays. The Z_1 invariant mass must be larger than 40 GeV. All leptons must be separated in angular space by at least $\Delta R(\ell_i, \ell_j) > 0.02$. At least two leptons are required to have $pT > 10$ GeV at least one is required to have $pT > 20$ GeV. In the 4μ and $4e$ subchannels, where an alternative Z_aZ_b candidate can be built out of the same four leptons, we discard candidates with $m_{Z_b} < 12$ GeV if Z_a is closer to the nominal Z boson mass than Z_1 is. This protects against events that contain an on-shell Z and a low-mass dilepton resonance. To further suppress events with leptons originating from hadron decays

in jet fragmentation or from the decay of low-mass hadronic resonances, all four opposite-charge lepton pairs that can be built with the four leptons (irrespective of flavor) are required to satisfy $m_{\ell^+\ell'^-} > 4$ GeV, where selected. FSR photons are disregarded in the invariant mass computation. Finally, the four-lepton invariant mass $m_{4\ell}$ must be larger than 70 GeV, which defines the mass range of interest for the subsequent steps of the analysis.

Up to one RSE is allowed to enter the definition of selected four-leptons events if both Z candidates satisfy $60 < m_{\ell\ell} < 120$ GeV. In addition, for these cases the opposite-charge requirement for the pair of leptons forming the Z candidate that includes the RSE is relaxed. This complementary set of events constitutes an additional gain in efficiency of about 10 (5%) for the $4e$ ($2e2\mu$) channel and is treated as a separated category in the analysis.

In events where more than one ZZ candidate passes the above selection, the candidate with the highest value of $\mathcal{D}_{\text{bkg}}^{\text{kin}}$ (defined in Section ??) is retained, except if two candidates consist of the same four leptons in which case the candidate with the Z_1 mass closest to the nominal Z boson mass is retained. This arbitration is different from Refs.^{30,34} and leads to a better assignment of the ZZ candidate in the case of events with more than 4 leptons, and subsequently a more pure event categorization, described later in Section ??.

4.9.3 Event Categorization

In order to improve the sensitivity to the Higgs boson production mechanisms, the selected events are classified into mutually exclusive categories. Category definitions exploit the multiplicity of jets, b-tagged jets and additional leptons (defined as leptons that are not involved in the best ZZ candidate and pass tight ID, SIP and isolation requirements), and requirements on the kinematic discriminants described in Section ??.

Six categories are defined, using the following criteria applied in this exact order (i.e. an event is considered for the subsequent category only if it does not satisfy the requirements of the previous category):

- **VBF-2jet-tagged category** requires exactly 4 leptons. In addition there must be either 2 or 3 jets of which at most 1 is b-tagged, or at least 4 jets and no b-tagged jets. Finally, a high $\mathcal{D}_{2\text{jet}}$ value is required.
- **VH-hadronic-tagged category** requires exactly 4 leptons. In addition there must be 2 or 3 jets of which at most 1 is b-tagged, or at least 4 jets and no b-tagged jets. Finally, a high \mathcal{D}_{ZH} or \mathcal{D}_{WH} value is required. In order to select $Z \rightarrow b\bar{b}$ decays, this category also includes events with exactly 4 leptons and 2 or 3 jets, of which at least 2 are b-tagged.
- **VH-leptonic-tagged category** requires no more than 3 jets and no b-tagged jets in the event, and exactly 1 additional lepton or 1 pair of additional opposite-

sign leptons. This category also includes events with no jets and at least 1 additional lepton.

- **ttH-tagged category** requires at least 4 jets of which at least 1 is b-tagged, or at least 1 additional lepton.
- **VBF-1jet-tagged category** requires exactly 4 leptons, exactly 1 jet and a high $\mathcal{D}_{1\text{jet}}$ value.
- **Untagged category** consists of the remaining events.

Figure 4.21 shows the composition of the six categories in terms of Higgs production processes. This categorization is applied for analysis of events around the on-shell peak of $H(125)$ in the mass window $105 < m_{4\ell} < 140\text{GeV}$. The working points of the discriminants used for the event categorization are shown in Fig. ???. For the analysis of the Higgs boson width including the offshell region, and for the search for a high-mass resonance, all categories other than the VBF-2jet-tagged category are combined into one, because at high mass, the contribution of production mechanisms other than VBF and gluon fusion is expected to be negligible.

4.9.4 Kinematic discriminants

The full kinematic information from each event using either the Higgs boson decay products or associated particles in its production is extracted using matrix element calculations and used to form several kinematic discriminants.

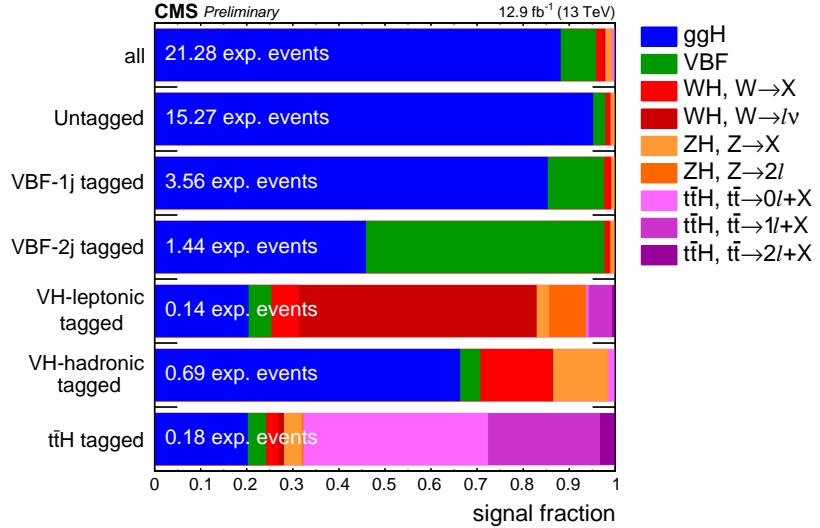


Figure 4.21: Signal relative purity of the six event categories in terms of the 5 main production mechanisms of the $H(125)$ boson in a $118 < m_{4\ell} < 130$ GeV window. The WH , ZH and $t\bar{t}H$ processes are split according to the decay of associated objects, whereby X denotes anything else than a lepton.

Both H boson decay kinematics and kinematics of associated production of $H+\text{jet}$, $H+2$ jets, VBF, ZH , WH are explored in this analysis. The full event kinematics is described by decay observables $\vec{\Omega}^{H \rightarrow 4\ell}$ or observables describing associated production $\vec{\Omega}^{H+\text{JJ}}$.

$$\mathcal{D}_{\text{bkg}}^{\text{kin}} = \left[1 + \frac{\mathcal{P}_{\text{bkg}}^{\bar{q}q}(\vec{\Omega}^{H \rightarrow 4\ell}|m_{4\ell})}{\mathcal{P}_{\text{sig}}^{gg}(\vec{\Omega}^{H \rightarrow 4\ell}|m_{4\ell})} \right]^{-1}, \quad (4.6)$$

where the denominator contains the probability for the signal and the numerator includes the probability for the dominant $q\bar{q} \rightarrow 4\ell$ background process, all calculated

either with the JHUGEN or MCFMmatrix elements within the MELA framework.

The discriminant sensitive to the VBF signal topology with two associated jets is calculated as^{53,55}

$$\mathcal{D}_{2\text{jet}} = \left[1 + \frac{\mathcal{P}_{\text{HJJ}}(\vec{\Omega}^{\text{H+JJ}} | m_{4\ell})}{\mathcal{P}_{\text{VBF}}(\vec{\Omega}^{\text{H+JJ}} | m_{4\ell})} \right]^{-1}, \quad (4.7)$$

where \mathcal{P}_{VBF} and \mathcal{P}_{XJJ} are probabilities obtained from the JHUGEN matrix elements for the VBF process and gluon fusion (technically combination of $gg/\bar{q}g/\bar{q}\bar{q}'$ parton collisions) in association with two jets (X+2 jets) within the MELA framework. This discriminant is equally efficient in separating VBF from either $gg \rightarrow X + 2$ jets signal or gg or $\bar{q}\bar{q} \rightarrow 2\ell 2\bar{q} + 2$ jets background because jet correlations in these processes are distinct from the VBF process.

The discriminant sensitive to the VBF signal topology with one associated jet is calculated as^{53,55}

$$\mathcal{D}_{1\text{jet}} = \left[1 + \frac{\mathcal{P}_{\text{HJ}}(\vec{\Omega}^{\text{H+J}} | m_{4\ell})}{\int d\eta_J \mathcal{P}_{\text{VBF}}(\vec{\Omega}^{\text{H+JJ}} | m_{4\ell})} \right]^{-1}, \quad (4.8)$$

where \mathcal{P}_{HJ} is the probability obtained from the JHUGEN matrix element for gluon fusion in association with one jet ($H + 1$ jet). The $\int d\eta_J \mathcal{P}_{\text{VBF}}$ is the integral of the two-jet VBF matrix element probability discussed above over the η_J values of the unobserved jet with the constraint that the total transverse momentum of the HJJ system is zero.

The discriminant sensitive to the VH signal topology with two associated jets is

calculated as^{53,55}

$$\mathcal{D}_{\text{WH}} = \left[1 + \frac{\mathcal{P}_{\text{HJJ}}(\vec{\Omega}^{\text{H+JJ}} | m_{4\ell})}{\mathcal{P}_{\text{WH}}(\vec{\Omega}^{\text{H+JJ}} | m_{4\ell})} \right]^{-1}, \quad (4.9)$$

$$\mathcal{D}_{\text{ZH}} = \left[1 + \frac{\mathcal{P}_{\text{HJJ}}(\vec{\Omega}^{\text{H+JJ}} | m_{4\ell})}{\mathcal{P}_{\text{ZH}}(\vec{\Omega}^{\text{H+JJ}} | m_{4\ell})} \right]^{-1}, \quad (4.10)$$

where \mathcal{P}_{VH} and \mathcal{P}_{HJJ} are probabilities obtained from the JHUGEN matrix elements for the VH (either ZH or WH) process and gluon fusion in association with two jets ($H + 2$ jets).

The four discriminants In Eqs. (4.16,4.8,4.10,4.9) are used to enhance the purity of event categories.

4.9.5 Signal Modeling

The $H \rightarrow 4\ell$ final state offers excellent efficiency and resolution in reconstruction of the final state particles. Therefore, parameterization of the signal can be separated in two steps: modeling of the ideal distributions, both the shape and yield, and introducing detector effect. The ideal distribution for a narrow $H(125)$ or a high-mass resonance is simple, which can be taken essentially as a delta function when convoluting with a resolution function. Parameterization of a broad resonance or of the off-shell tail of an $H(125)$ signal, or any other signal below $2m_Z$ threshold, requires special care of modeling the line-shape and interference with background. We therefore split description of the signal model into two parts: first discussion

of detector effects and then discussion of a more complex model for a broad signal distribution.

4.9.5.1 Signal model: reconstruction efficiency

Parameterization of efficiency function $\mathcal{E}(m_{4\ell})$ is extracted for each final state ($2e2\mu$, $4e$, 4μ) separately from the full simulation of each production mechanism and each signal type with POWHEG+JHUGen simulation for the wide-resonance with $\Gamma_X = \Gamma_H^{\text{SM}}(m_H)$. For efficiency, a total of 57 mass points are used for fits: 115, 120, 124, 125, 126, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180 within 1 GeV of m_H ; 190, 200, 210, 230, 250, 270, 300 within 5 GeV of m_H ; and at 400 GeV and above increments of 50 GeV up to 3 TeV. The efficiency parameterization is shown in Fig. 4.34 for a SM-like signal with $J^P = 0^+$ in the gluon fusion production, for a SM-like signal with $J^P = 0^+$ in the VBF production, and for a spin-2 signal with $J^P = 2_b^+$ in the model of KK graviton propagating in the bulk in the gluon fusion production. Parameterization of spin-2 model is performed using spin-0 simulation re-weighted with the MELA weights to account for different kinematics. In general, efficiency for the 2_b^+ model is higher than for isotropic spin-0 decay because of higher acceptance efficiency in the spin-2 model with the Z bosons predominantly in the central region of the detector.

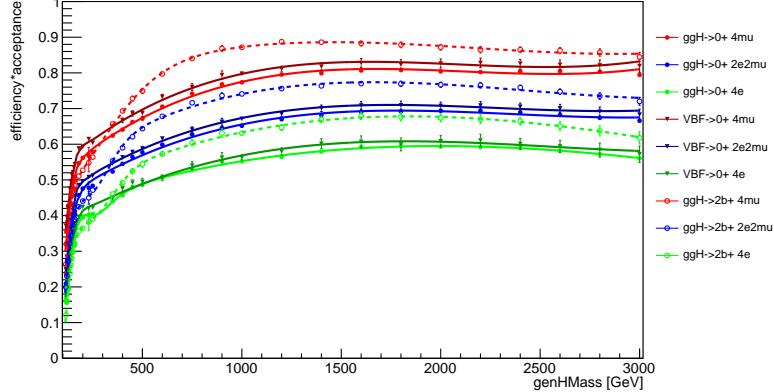


Figure 4.22: Parameterization of reconstruction efficiency $\mathcal{E}(m_{4\ell})$ (including acceptance) for a SM-like signal with $J^P = 0^+$ in the gluon fusion production (solid circles); for a SM-like signal with $J^P = 0^+$ in the VBF production (solid triangles); and for a spin-2 signal with $J^P = 2_b^+$ in the model of KK graviton propagating in the bulk in the gluon fusion production (open circles). Dependence as a function of mass $m_{4\ell}$ is shown in three channels: $X \rightarrow 4\mu$ (red, top); $2e2\mu$ (blue, middle); $X \rightarrow 4e$ (green, bottom). Points show results from full simulation, curves show analytical fits.

4.9.5.2 Signal mass model: resolution

Parameterization of the resolution function $\mathcal{R}(m_{4\ell}|m_{4\ell}^{\text{ideal}})$ is extracted for each final state ($2e2\mu$, $4e$, 4μ), separately from the full simulation of the $gg \rightarrow X \rightarrow 4\ell$ process with POWHEG+JHUGen for the wide-resonance with $\Gamma_X = \Gamma_H^{\text{SM}}(m_H)$.

The resolution in other production mechanisms in the spin-2 model are generally very similar to the SM gluon fusion production. Therefore, a common parameterization of resolution can be used. Nonetheless, we provide the option to use separate

resolution function for each production mechanism.

A search for a broad resonance with mass up to 2.5 TeV requires $m_{4\ell}$ parameterization and convolution up to about 3 TeV. Since simulation is performed for broad resonances and parameterization is done as a function of a given value of $m_{4\ell}$, the following procedure is used. For resolution, a total of 54 mass points are used for fits: from 120, increments of 20 GeV up to 900 GeV, increments of 50 GeV up to 1600 GeV, with the selection window within 5 GeV below 1 TeV and within 5% above. The distribution of resolution ($(m_{4\ell} - m_{4\ell}^{\text{ideal}})$) is parameterized, which does not vary much within the mass windows quoted above. Several representative examples are shown in Fig. 4.37. Dependence of six parameters of the double-CrystalBall function is parameterized with the glued polynomial functions in several mass regions, which preserve the value and derivative of the function at the junctions. After initializing the parameters from individual fits in each mass window, a joint fit is performed all mass distributions at the same time, leading to a smooth and optimal dependence of all parameter on $m_{4\ell}$, see Fig. 4.38 for such dependence.

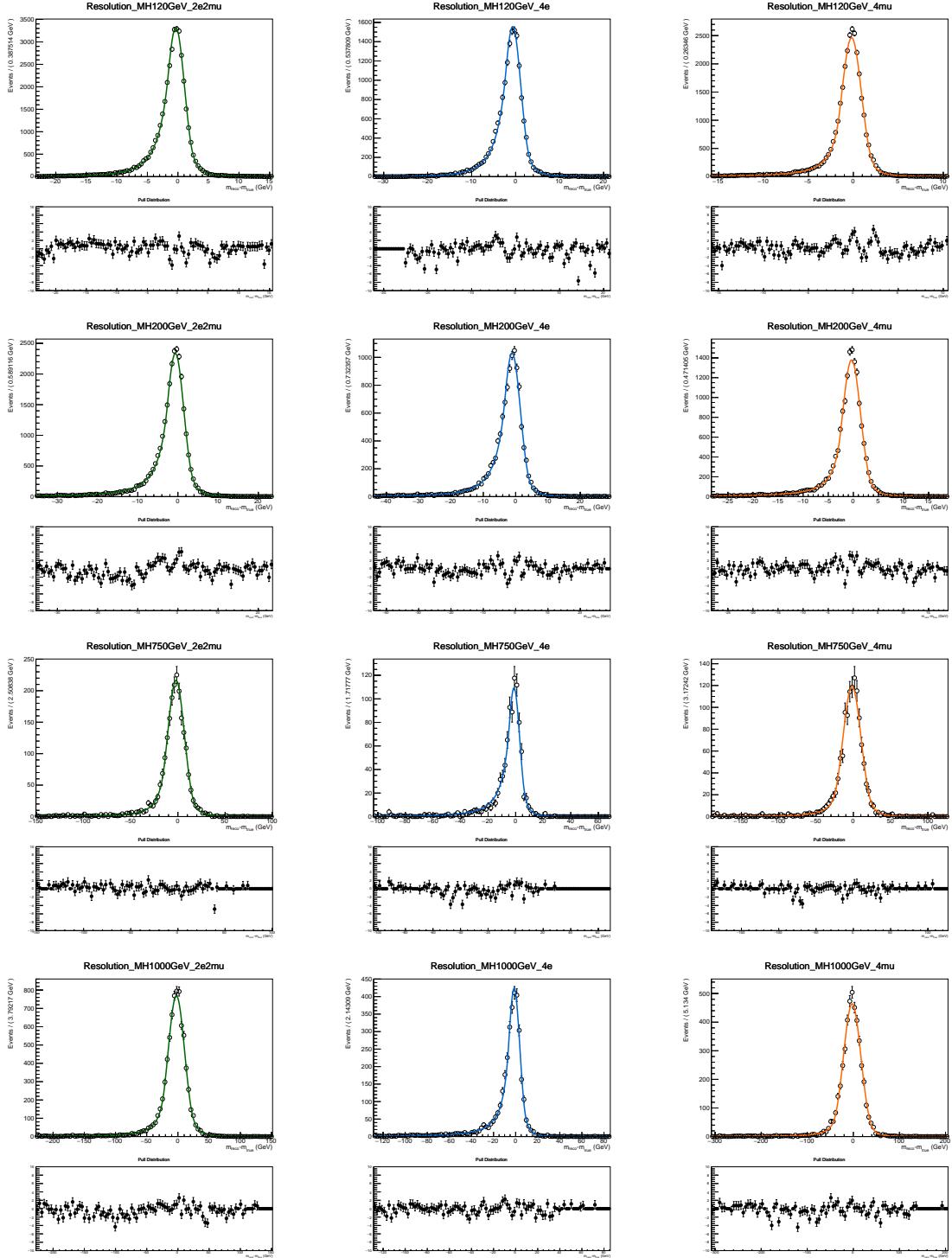


Figure 4.23: Representative $m_{4\ell}$ distributions of a resolution function $\mathcal{R}(m_{4\ell}|m_{4\ell}^{\text{ideal}})$

in $2e2\mu$ (left), $4e$ (middle), and 4μ (right) channels. Several examples are shown from top to bottom: $m_{4\ell} = 120, 200, 740, 1000$ GeV .

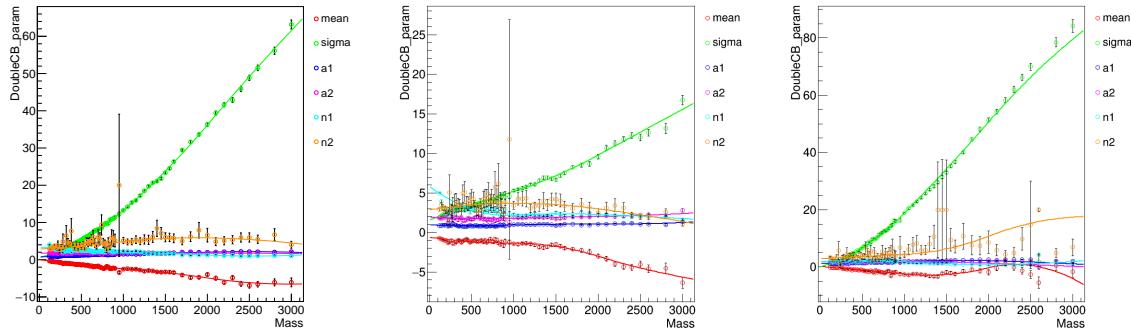


Figure 4.24: Dependence of six parameters of the resolution function in the $2e2\mu$ (left), $4e$ (middle), and 4μ (right) channels. Point show results of individual fits, smooth curves show parameterization from a simultaneous fit of all distributions.

4.9.5.3 Signal mass model: lineshape prior to detector effects

Parameterization of the signal lineshape for a off-shell tail (above $2m_Z$ threshold for $m_X < 2m_Z$) or a broad resonance requires special care to include interference with background.

The following production mechanisms are considered: gluon fusion, VBF, VH . The cross section for the $t\bar{t}H$ production mechanism is highly suppressed for $m_X > 2m_Z$ and we therefore do not consider it.

A general parameterization of the signal and background $gg \rightarrow 4\ell$ (or similarly VBF/VH) process which includes all contributing amplitudes for $gg \rightarrow \text{bkg} + H(125)^* + X(m_X) \rightarrow 4\ell$ and their interference is used.

The mass and width of both $H(m_H, \Gamma_H)$ and $X(m_X, \Gamma_X)$ are also included as general parameters of the model. This is supported by the MC simulation with the MCFM+JHUGen framework, and including NNLO corrections for cross section using HNNLO program based on MCFM. The matrix element package MELA supports all of these processes and allows fast MC re-weighting and optimal discriminant calculation. This work has been developed as part of the LHC Higgs Cross Section Group Yellow Report 4 preparation.

The probability density function $\mathcal{P}^{\text{ideal}}(m_{4\ell}^{\text{ideal}} | m_X, \Gamma_X, \sigma_X)$, as implemented in RooFit and CMS combine tool, is based on the matrix element MCFM+JHUGen+HNNLO implemented within the MELA framework and has the most general parameterization of $H(m_H, \Gamma_H)$ and $X(m_X, \Gamma_X)$ for any width and mass of the resonances. As can be

seen in Fig. 4.25, interference of X with both background and off-shell tail of $H(125)$ is an important effect for a resonance with a non-negligible width. This parameterization naturally scales the size of interference with cross section of the signal peak σ_X . Proper accounting for interference depending on the signal strength, parametric dependence on the mass distribution (without the need for histogram morphing and excessive MC simulation), and including off-shell $H(125)^*$ tail are the particularly new important features introduced since Run-I. Detector effects in the probability density function $\mathcal{P}^{\text{reco}}(m_{4\ell})$ are implemented with convolution for the mass resolution $\mathcal{R}(m_{4\ell}|m_{4\ell}^{\text{ideal}})$ and multiplicative efficiency function $\mathcal{E}(m_{4\ell})$, both extracted from full simulation with the POWHEG+JHUGen samples. The proper re-weighting of the mass distribution in these samples has been applied with the MELA approach in order to remove the Complex Pole Scheme corrections, since those are no longer applicable for a new non-SM resonance. Figure 4.25 shows more detailed breakdown of interference components shown at generator level.

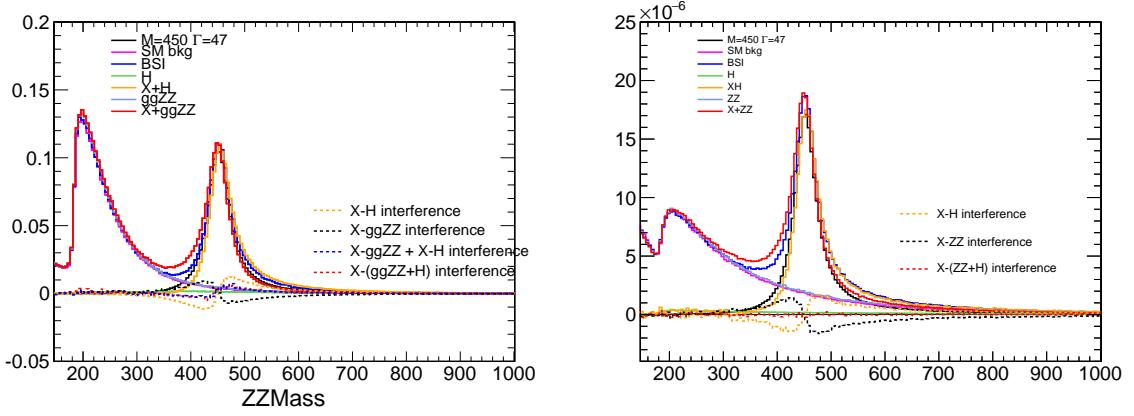


Figure 4.25: Distribution of events (arbitrary units) in the process $gg \rightarrow 4\ell$ (left) and $qq \rightarrow 4\ell$ (right) as a function of invariant mass $m_{4\ell}$ generated with the MCFM+JHUGen and Phantom framework. Various components in the $X(450) + H(125)^*$ + bkg interference are shown and explained in the legend.

4.9.5.4 Signal mass model: closure tests

The final signal mass parameterization is built from efficiency, resolution, and ideal distributions as follows:

$$\mathcal{P}_{\text{sig}}^{\text{reco}}(m_{4\ell}) = \mathcal{E}(m_{4\ell}) \times (\mathcal{R}(m_{4\ell}|m_{4\ell}^{\text{ideal}}) \otimes \mathcal{P}_{\text{sig}}(m_{4\ell}^{\text{ideal}}|m_X, \Gamma_X, m_H, \Gamma_H)) \quad (4.11)$$

In the following we provide a bit more details on parameterization and closure tests of the procedure. show closure test of the full probability density function $\mathcal{P}^{\text{reco}}(m_{4\ell})$ against full simulation for both signal and background separately. There is an excellent agreement between full simulation and the model for an arbitrary $X(m_X, \Gamma_X)$ resonance as well as for the SM background. Figure 4.26 shows signal parameterization in VBF. Figure 4.27 shows application to the mass and width measurement of $H(125)$ with off-shell tail (gluon fusion is shown).

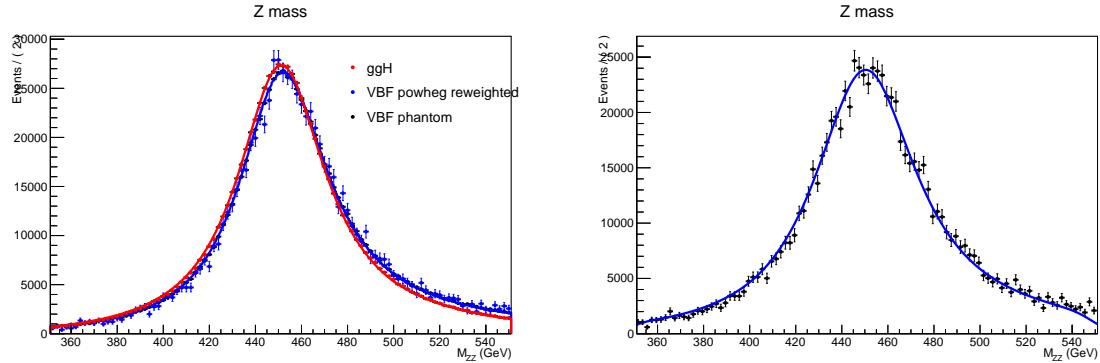


Figure 4.26: Left: comparison of POWHEG, Phantom simulation, analytical model in VBF and gluon fusion model. Right: closure test on VBF full simulation.

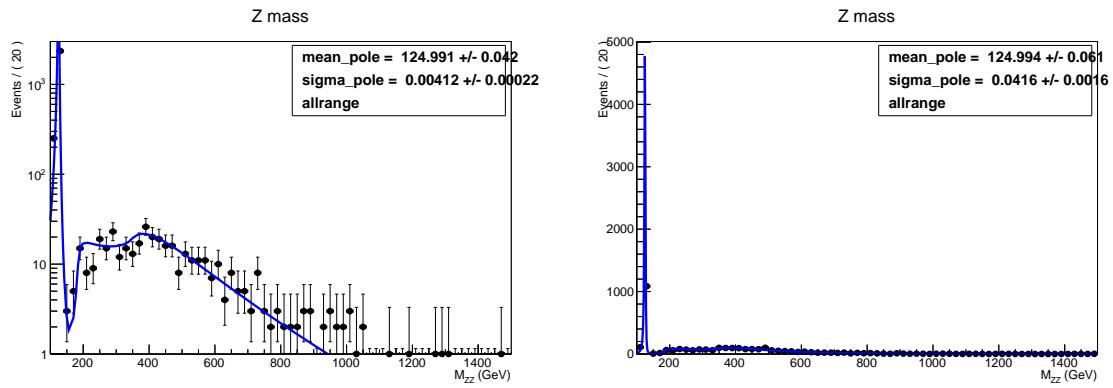


Figure 4.27: Closure test if $H(125)$ width parameterization with off-shell tail. Left: SM width, right: 10 times SM width. Statistics box shows fit results for the mass and width, in agreement with the input.

4.9.5.5 Signal model: kinematics

The final signal parameterization in 2D is built with conditional template $\mathcal{T}(D_{\text{bkg}}^{\text{kin}}|m_{4\ell})$ which describes $D_{\text{bkg}}^{\text{kin}}$ discriminant for each value of $m_{4\ell}$.

$$\mathcal{P}_{\text{sig}}^{\text{reco}}(m_{4\ell}, D_{\text{bkg}}^{\text{kin}}) = \mathcal{P}_{\text{sig}}^{\text{reco}}(m_{4\ell}) \times \mathcal{T}(D_{\text{bkg}}^{\text{kin}}|m_{4\ell}) \quad (4.12)$$

The template $\mathcal{T}(D_{\text{bkg}}^{\text{kin}}|m_{4\ell})$ parameterization includes all detector effects affecting the $D_{\text{bkg}}^{\text{kin}}$ distribution. Definition of $D_{\text{bkg}}^{\text{kin}}$ depends on the signal spin model. Examples of $\mathcal{T}(D_{\text{bkg}}^{\text{kin}}|m_{4\ell})$ distributions are shown in Section on kinematics.

4.9.6 Background Estimation

The irreducible background to the Higgs signal in the 4ℓ channel, which come from the production of ZZ via $\bar{q}q$ annihilation or gluon fusion, is estimated using simulation. The fully differential cross section for the $q\bar{q} \rightarrow ZZ$ process has been computed at NNLO,⁴⁵ and the NNLO/NLO K factor as a function of m_{ZZ} has been applied to the POWHEG sample. This K factor varies from 1.0 to 1.2 and is 1.1 at $m_{ZZ} = 125$ GeV. Additional NLO electroweak corrections which depend on the initial state quark flavor and kinematics are also applied in the region $m_{ZZ} > 2m_Z$ where the corrections have been computed.

The production of ZZ via gluon fusion contributes at NNLO in pQCD. It has been shown²² that the soft collinear approximation is able to describe the background cross section and the interference term at NNLO. Further calculations also show that the

K factors are very similar at NLO for the signal and background⁶¹ and at NNLO for the signal and interference terms.⁵⁹ Therefore, the same K factor is used for the signal and background.⁶⁴ The NNLO K factor for the signal is obtained as a function of m_{ZZ} using the HNNLO v2 program^{26,44,46} by calculating the NNLO and LO $gg \rightarrow H \rightarrow 2\ell 2\ell'$ cross sections at the small H boson decay width of 4.07 MeV and taking their ratios. The NNLO/LO K factor for $gg \rightarrow ZZ$ varies from 2.0 to 2.6 and is 2.27 at $m_{ZZ} = 125$ GeV, and a systematic uncertainty of 10% on its determination when applied to the background process is used in the analysis.

Additional backgrounds to the Higgs signal in the 4ℓ channel arise from processes in which heavy-flavor jets produce secondary leptons, and also from processes in which decays of heavy-flavor hadrons, in-flight decays of light mesons within jets, or (for electrons) the decay of charged hadrons overlapping with π^0 decays are misidentified as leptons. The main processes producing these backgrounds are $Z + \text{jets}$, $t\bar{t} + \text{jets}$, $Z\gamma + \text{jets}$, $WW + \text{jets}$ and $WZ + \text{jets}$. We denote these reducible backgrounds as “Z+X” since they are dominated by the $Z + \text{jets}$ process.

The contribution from the reducible background is estimated using two independent control regions in data. The control regions are formed by requiring the presence of two leptons which satisfy the “tight” identification criteria, plus two additional opposite-sign (OS) or same-sign (SS) leptons satisfying relaxed (“loose”) identification requirements. These four leptons are then required to pass the same ZZ candidate selection as described in Section 4.9.2. The event yield in the signal region is obtained

by weighting the control region events by the lepton misidentification probability (or fake rate) f , defined as the fraction of non-signal leptons which are identified by the analysis selection criteria.

The lepton fake rates f_e and f_μ are measured by forming a sample dominated by events which include a Z boson and a misidentified lepton. This sample requires the presence of two same flavor, opposite charge leptons with $p_T > 20/10$ GeV passing the tight selection criteria, thus forming a Z_1 candidate. In addition, the presence of exactly one additional lepton passing the loose identification criteria is required. This additional lepton is used as the probe to measure the misidentification rate. The invariant mass of the probe lepton and the opposite sign lepton from the reconstructed Z candidate is required to satisfy $m_{2\ell} > 4$ GeV, consistent with the candidate selection in the signal region. For the OS method, a requirement on the invariant mass of the two tight leptons $|m_{\ell_1\ell_2} - m_Z| < 7$ GeV is applied in order to reduce the contribution of (asymmetric) photon conversions which are estimated separately. In the SS method, the contribution from photon conversions is estimated by determining an average fake rate which will be described in Section ???. Furthermore, the E_T^{miss} , defined as the modulus of the vector sum of the transverse momenta of all reconstructed PF candidates (charged or neutral) in the event, is required to be less than 25 GeV in order to suppress contamination from WZ and $t\bar{t}$ processes.

The control region for the OS method consists of events with a Z_1 candidate and two additional leptons with the same flavor and opposite charge. The expected yield

in the signal region is obtained from two categories of events.

The first category is composed of events with two leptons which pass (P) the lepton tight identification requirements and two leptons which pass the loose identification but fail (F) the tight identification, and is denoted as the 2P2F region. Backgrounds which intrinsically have only two prompt leptons, such as $Z + \text{jets}$ and $t\bar{t}$, are estimated by this control region. To obtain the expected yield in the signal region, each event i in the 2P2F region is weighted by a factor $\frac{f_3^i}{1-f_3^i} \frac{f_4^i}{1-f_4^i}$, where f_3^i and f_4^i are the misidentification rates for the third and fourth lepton.

The second category consists of events where exactly one of the two additional leptons passes the analysis selection, and is referred to as the 3P1F region. Backgrounds with three prompt leptons, such as $WZ + \text{jets}$ and $Z\gamma + \text{jets}$ with the photon converting to e^+e^- , are estimated using this region. To obtain the expected yield in the signal region, each event j in the 3P1F region is weighted by a factor $\frac{f_a^j}{1-f_a^j}$, where f_a^j is the misidentification rate for the fourth lepton. The contribution from ZZ events to the 3P1F region (N_{3P1F}^{ZZ}) which arises from events where a prompt lepton fails the identification requirements, is estimated from simulation.

The contamination of 2P2F-type processes in the 3P1F region is estimated as $\sum_i (\frac{f_3^j}{1-f_3^j} + \frac{f_4^j}{1-f_4^j})$ and contributes to the final weighted sum of the 3P1F events with the component $\sum_i (2 \frac{f_3^j}{1-f_3^j} \frac{f_4^j}{1-f_4^j})$, which is subtracted from the background estimate.

The final expected yield from this method for the reducible background in the signal region, $N_{\text{SR}}^{\text{reducible}}$, can be written as:

$$N_{\text{SR}}^{\text{reducible}} = \sum_j^{N_{\text{3P1F}}} \frac{f_a^j}{1 - f_a^j} - \sum_j^{N_{\text{3P1F}}^{\text{ZZ}}} \frac{f_a^j}{1 - f_a^j} - \sum_i^{N_{\text{2P2F}}} \frac{f_3^i}{1 - f_3^i} \frac{f_4^i}{1 - f_4^i}. \quad (4.13)$$

where N_{3P1F} and N_{2P2F} are the number of events in the 3P1F and 2P2F regions, respectively. The same method is used for the background estimation of events with an RSE.

The control region for the same-sign (SS) method, referred to as the 2P2L_{SS} region, consists of events with a Z_1 candidate and two additional leptons with the same flavor and same charge. The two additional leptons are required to pass the loose selection requirements for leptons.

To estimate the contribution of photon conversions to the electron misidentification probability f , the linear dependence of f on the fraction of loose electrons with tracks having one missing hit in the pixel detector, r_{miss} , is used to derive a corrected misidentification rate \tilde{f} . The dependence of f on the fraction r_{miss} is determined by varying the requirement on $|m_{\ell_1 \ell_2} - m_Z|$ and $|m_{\ell_1 \ell_2 e_{\text{loose}}} - m_Z|$ to obtain samples with different FSR contributions. The value of r_{miss} measured in the control region is then used to obtain the corrected misidentification rate \tilde{f} .

The expected number of reducible background events in the signal region can then be written as:

$$N_{\text{SR}}^{\text{reducible}} = r_{\text{OS/SS}} \cdot \sum_i^{N_{\text{2P2LSS}}} \tilde{f}_3^i \cdot \tilde{f}_4^i, \quad (4.14)$$

where the ratio $r_{\text{OS/SS}}$ between the number of events in the 2P2L_{OS} and 2P2L_{SS} control regions is obtained from simulation. The 2P2L_{OS} region is defined analogously to the 2P2L_{SS} region but with an OS requirement for the additional pair of loose leptons.

The predicted yield in the signal region of the reducible background from the two methods are in agreement within their statistical uncertainties, and since they are mutually independent, the results of the two methods are combined. The shape of the $m_{4\ell}$ distribution for the reducible background is obtained by combining the prediction from the OS and SS methods and fitting the distributions with empirical functional forms built from Landau⁵⁷ and exponential distributions.

The dominant systematic uncertainty on the reducible background estimation arises from the limited number of events in the control regions as well as in the region where the misidentification rates are computed. Additional sources of systematic uncertainty arise from the difference in the composition of the sample from which the misidentification rate is computed and the control regions of the two methods where the fake rate is applied. The systematic uncertainty on the $m_{4\ell}$ shape is determined by taking the envelope of differences among the shapes from the OS and SS methods in the three different final states. The combined systematic uncertainty is estimated to be about 40 to 55%.

4.9.7 Systematic uncertainties

The main experimental uncertainties which affect both signal and background are the uncertainty on the integrated luminosity (6.2%) and the uncertainty on the lepton identification and reconstruction efficiency (6–11%, depending on the final state). Experimental uncertainties for the reducible background estimation, described in Section ??, vary between 40–55%. The uncertainty on the lepton energy scale is 0.04% for muons and 0.3% for electrons, and the uncertainty on the 4ℓ mass resolution coming from the uncertainty on the per-lepton energy resolution is 20%. Theoretical uncertainties which affect both the background signal and background estimation include uncertainties from the renormalization and factorization scale and choice of PDF set. The uncertainty from the renormalization and factorization scale is determined by varying these scales between 0.5 and 2 times their nominal value while keeping their ratio between 0.5 and 2. The uncertainty from the PDF set is determined by taking the root mean square of the variation when using different replicas of the default NNPDF set. An additional uncertainty of the 10% on the K factor used for the $gg \rightarrow ZZ$ prediction is applied as described in Section ??.

A systematic uncertainty of 2% on the branching ratio of $H \rightarrow ZZ \rightarrow 4\ell$ only affects the signal yield. In the case of event categorization, experimental and theoretical uncertainties which account for possible migration of signal and background events between categories are included. The main sources of uncertainty on the event categorization include the QCD scale, PDF set, and the modeling of hadronization and the under-

lying event. These uncertainties amount to between 8–20% for the signal and 3–20% for the background depending on the category, and are largest for the prediction of the $gg \rightarrow H$ yield in the VBF-2jet-tagged category. Additional uncertainties come from the imprecise knowledge of the jet energy scale (2–18%) and b-tagging efficiency and mistag rate (about 6% in the $t\bar{t}H$ -tagged category). The systematic uncertainties which affect the $H \rightarrow ZZ \rightarrow 4\ell$ measurements are summarized in Table ??.

Summary of the systematic uncertainties in the $H \rightarrow 4\ell$ measurements.

Summary of relative systematic uncertainties	
Common experimental uncertainties	
Luminosity	6.2 %
Lepton identification/reconstruction efficiencies	6 – 11 %
Background related uncertainties	
QCD scale ($q\bar{q} \rightarrow ZZ, gg \rightarrow ZZ$)	3 – 10 %
PDF set ($q\bar{q} \rightarrow ZZ, gg \rightarrow ZZ$)	3 – 5 %
Electroweak corrections ($q\bar{q} \rightarrow ZZ$)	1 – 15 %
$gg \rightarrow ZZ$ K factor	10 %
Reducible background (Z+X)	40 – 55 %
Event categorization (experimental)	2 – 18 %
Event categorization (theoretical)	3 – 20 %
Signal related uncertainties	
QCD scale ($q\bar{q} \rightarrow \text{VBF/VH}, gg \rightarrow H/t\bar{t}H$)	3 – 10 %
PDF set ($q\bar{q} \rightarrow \text{VBF/VH}, gg \rightarrow H/t\bar{t}H$)	3 – 4 %
$\text{BR}(H \rightarrow ZZ \rightarrow 4\ell)$	2 %
Lepton energy scale	0.04 – 0.3 %
Lepton energy resolution	20 %
Event categorization (experimental)	2 – 15 %
Event categorization (theoretical)	8 – 20 %

4.9.8 Results

The search for a scalar resonance X decaying to $ZZ \rightarrow 4\ell$ is performed for the range of masses m_X between 130 and 2500 GeV. The width of the resonance Γ_X is allowed to have any value, starting from the narrow-width approximation (denoted as $\Gamma_X = 0$) to an arbitrarily large width, generally $\Gamma_X < m_X$. Production of the X resonance is considered to be either in gluon fusion or vector boson fusion, where VH production is included according to expectation of the relative VH and VBF cross sections. The fraction of VBF and VH production is parameterized with f_{VBF} , so that the gluon fusion cross section is $\sigma_X \times (1 - f_{\text{VBF}})$.

An unbinned maximum likelihood fit of the $m_{4\ell}$ distribution is performed over the full range of selected events between 100 and 3000 GeV. Constraints on the resonance production cross section σ_X are reported for a scan of the m_X and Γ_X values, where the value of f_{VBF} can be either fixed to a certain value or left unconstrained in the fit. Constraints on f_{VBF} are possible due to splitting events into two categories with VBF-like two-jet topology and the rest of the events. The only exception is the RSE category of $4e$ and $2e2\mu$ events which are treated inclusively in the mass range $m_{4\ell} > 300$ GeV.

The general parameterization of the probability density function $\mathcal{P}(m_{4\ell}|m_X, \Gamma_X, \sigma_X)$ for the VBF or $gg \rightarrow 4\ell$ via background, $H(125)^*$ and $X(m_X)$, is based on the framework of MCFM+ JHUGEN + HNNLO within MELA, allows inclusion of interference between various components, including the off-shell tail of $H(125)$, in a parametric

way with proper dependence of the differential signal cross section $d\sigma_X/dm_{4\ell}$. This way, both signal and background are treated as one process, for either gluon fusion or vector boson fusion. The cross section σ_X is reported for the pure signal process only without including effects of interference. For a wide resonance with the pole mass m_X below the $2m_Z$ threshold, σ_X includes the off-shell production above the threshold.

Figure 4.39 shows observed and expected 95% CL limits on the $pp \rightarrow X \rightarrow ZZ \rightarrow 4\ell$ cross section including the four-lepton branching fraction, as a function of m_X and Γ_X , floating f_{VBF} . The systematic uncertainties are similar to those used in the other measurements. The uncertainties on the background predictions affect the shape of the $m_{4\ell}$ spectrum at high mass and these effects are included. Under assumption of pure electroweak production ($f_{VBF} = 1$) constraints are tighter due to lower background in the VBF-like category of events.

Upper limits at a 95% CL are set on the production of an additional Higgs boson for masses up to 2.5 TeV and for various widths. The details of the exclusions depend on the assumptions of width and production, but in general no significant excesses appear under any of the scenarios considered.

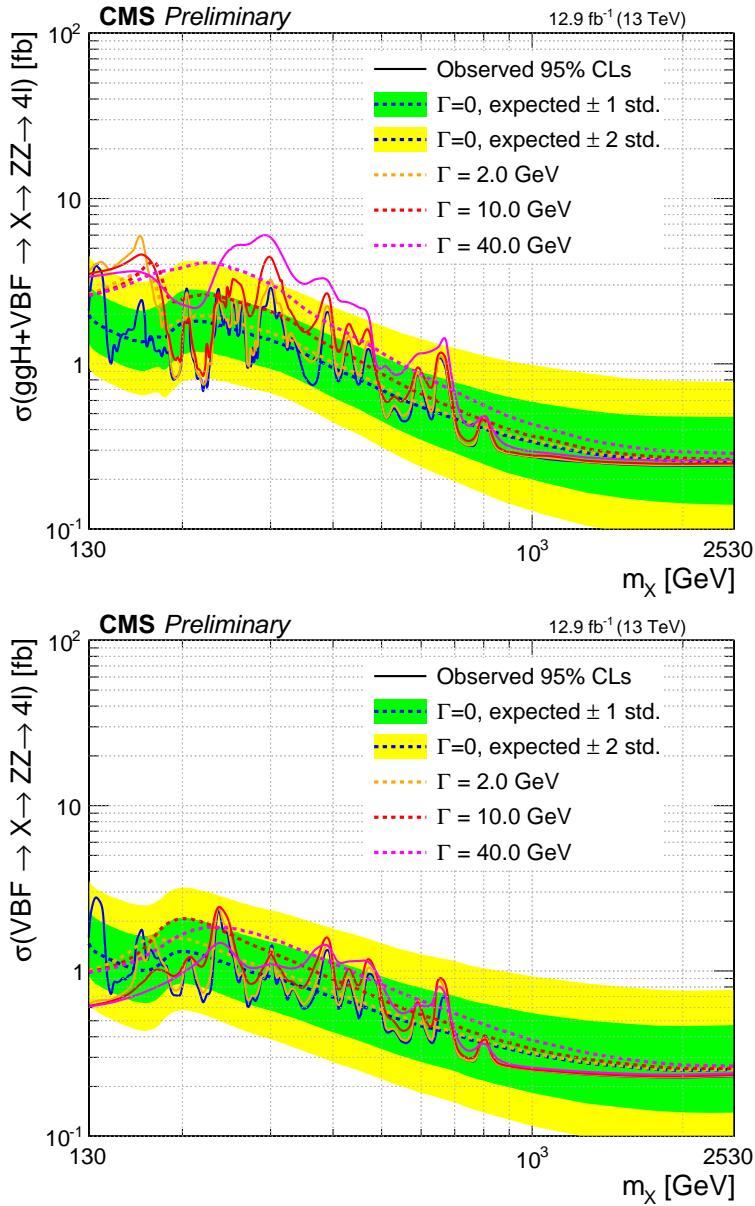


Figure 4.28: Top: Observed and expected upper limits at the 95% CL on the $X \rightarrow ZZ \rightarrow 4\ell$ cross section σ_X (including four-lepton branching fraction) as a function of m_X at several Γ_X values with f_{VBF} floated, with 12.9 fb^{-1} of data at 13 TeV. Bottom: Same as above but with $f_{VBF} = 1$ fixed.

4.10 $H \rightarrow ZZ \rightarrow 2l2q$

The search for a Higgs-like boson in $H \rightarrow ZZ$ when one of Z decays to jets, using Run-I data, has contributed to the combination of CMS Run-I high mass search in di-boson channel. The same final state is interesting in some models predicting Kaluza-Klein excitations of spin-2 particles, like the bulk graviton.

4.10.1 Data and Simulation

4.10.1.1 Experimental Data

Data sample recorded by the CMS experiment during 2016, corresponding to $\mathcal{L} = 12.9 \text{ fb}^{-1}$ of 25 ns data, runs B to D (“ICHEP2016 dataset”) is used.

The High Level Trigger (HLT) selections require one or two opposite-sign and same flavor electrons or muons, with asymmetric p_T thresholds. The analysis relies on four different primary datasets (PDs), *DoubleEG*, *DoubleMuon*, *SingleElectron* and *SingleMuon*, each of which combines a certain collections of HLT paths, where *MuEG* dataset is used for $t\bar{t}$ background estimation. To avoid duplicate events from different primary datasets, events are selected, in a specific trigger path only from one of these datasets and vetoed in others.

4.10.1.2 Simulated Samples

Signals with spin-0 hypothesis are generated using the SM Higgs boson production obtained by the POWHEG ^{10,39,62} generator for the two main production modes: gluon fusion (ggH) including quark mass effects ¹⁷, vector boson fusion (VBF).⁶³ The description of the decay of the Higgs boson to two quarks and two leptons is obtained using the JHUGEN generator.⁴²

Signals with spin-2 hypothesis are generated using the JHUGEN generator at LO. JHUGEN is able to model correctly the kinematic distributions of outgoing particles, given a set of effective coefficients describing production and decay. These coefficients are set to describe the interactions of a bulk graviton signal.⁷ Widths range from narrow to 30% of the mass.

The mass ranges from 400GeV to 2TeV. All samples are generated with the NNPDF 3.0 NLO parton distribution functions (PDFs).¹⁹

The background samples used to simulate Z+jets, $t\bar{t}$, di-boson processed are listed in Tab. ??.

A large set of jet-binned DY+jets samples produced at LO are used. The number of b-partons at generator level (both at LHE and parton shower level) is restricted to make the samples mutually exclusive when combining all the LO samples. In order to match the combination of LO Z+jets samples to NLO, a constant K factor derived (about 1.23) from MCFM is used.

Process	Dataset Name	$\sigma \cdot BR$
DY LO	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/	4895.0 pb
DY LO	/DY1JetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/	1016.0 pb
DY LO	/DY2JetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/	331.4 pb
DY LO	/DY3JetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/	96.36 pb
DY LO	/DY4JetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/	51.4 pb
DY LO	/DYBJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/	88.2771 pb
DY LO	/DYBJetsToLL_M-50_BGenFilter_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/	11.57 pb
$t\bar{t}$	/TT_TuneCUETP8M1_13TeV-powheg-pythia8	831.76 pb
$WZ \rightarrow 2\ell 2q$	/WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8/	5.595 pb
$ZZ \rightarrow 3\ell 2q$	/ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8/	3.22 pb
$ZZ \rightarrow 2\ell 2\nu$	/ZZTo2L2Nu_13TeV_powheg_pythia8/	0.564 pb
$WW \rightarrow 2\ell 2\nu$	/WWTo2L2Nu_13TeV-powheg/	10.481 pb
$WZ \rightarrow 3\ell\nu$	/WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8/	5.29 pb

4.10.2 Event Selection

4.10.2.1 Leptonic Z selection

Pairs of opposite sign and same flavour electrons and muons, selected as described in the previous section, are selected if the dilepton mass is between 60 and 120 GeV, rejecting backgrounds without a resonant Z, for example $t\bar{t}$. The leading and subleading leptons must have minimal pT of 40 and 24 GeV, and a minimum dilepton pT of 100 GeV is also required to reject low-HT Drell-Yan. The leptonic Z mass and pT is shown in Fig. ???. The two leptons must have a minimum separation of $\Delta R > 0.02$ to remove ghost tracks.

Here and in the following a slight data-over-MC excess is visible in the resolved

jet topology (see next paragraph). This is due to a mismodeling of the dilepton pT which can be seen in the lower row of Fig. ???. In the standard background estimation we neglect this small effect, since the normalisation is data-driven; in an alternative estimation, we correct for this effect (see Background Estimation section).

4.10.2.2 Hadronic Z selection

In resolved jet selection, hadronic Z boson candidates are constructed from pairs of AK4 jets selected as explained in the previous section. The dijet pT is required to be larger than 100 GeV and the dijet mass m_{jj} has a loose cut between 40 and 180 GeV. Among this, the region 70–105 GeV is considered to be the signal region, while the regions 40–70 GeV 135–180 GeV form the sideband region. The region with possible $H \rightarrow b\bar{b}$ signals, 105–135 GeV, is removed from any selection. Jets must not overlap with leptons so a cut $\Delta R(\ell, jet) > 0.4$ is applied to each of them. Finally, we just consider events with $m(\ell\ell jj) > 300$ GeV. Fig. ??? (left) shows these quantities.

In merged jet selection, hadronic Z boson candidates are constructed from a single AK8 jet selected as explained in the previous section. The jet pT is required to be larger than 170 GeV and the pruned jet mass m_J has a loose cut between 40 and 180 GeV. Among this, the region 70–105 GeV is considered to be the signal region, while the regions 40–70 GeV 135–180 GeV form the sideband region. The region with possible $H \rightarrow b\bar{b}$ signals, 105–135 GeV, is removed from any selection. Jets must not overlap with leptons so a cut $\Delta R(\ell, jet) > 0.8$ is applied to each of them. Finally, we

just consider events with $m(\ell\ell J) > 300$ GeV. Fig. ?? (right) shows these quantities.

In order to minimise contamination from standard model DY + jets production, we further require that the hadronic Z boson candidate from a merged selection has substructure. We exploit the techniques that are already standard in searches with merged jets coming from boosted bosons.[?] The N-subjettiness, τ_N is defined as

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

where the index k runs over the jet constituents and the distances $\Delta R_{N,k}$ are calculated with respect to the axis of the n^{th} subjet. The normalisation factor d_0 is calculated as $d_0 = \sum_k pT_{,k} R_0$, setting R_0 to the jet radius of the original jet. Jets with smaller τ_N are more compatible with the a N subjets configuration. We use the ratio of 2-subjettiness over 1-subjettiness, $\tau_{21} = \tau_2/\tau_1$, as the discriminating variable for the jet substructure, and impose a $\tau_{21} < 0.6$ requirement on merged hadronic Z candidates.

Many events have candidates passing both selections above or they have multiple candidates in one category. To arbitrate the events passing both resolved and merged jet selection, the two possible options are

- option 1: select resolved jets in the signal region, if not present select merged jet in the signal region
- option 2: select merged jet in the signal region, if not present select resolved jets in the signal region

Preliminary sensitivity studies showed that giving the resolved-jet selection higher priority (*option 1*) than merged jet leads to similar limit result. Fig. ?? compares the likelihood scan on signal cross section between the two options.

Another complication to be taken into account is that this arbitration introduces the possibility of rejecting a good event because it falls in the sideband region of the preferred topology. Finally we arbitrate like this:

- If two or more candidates are found in the resolved jet category, we take the one with the largest $p_T(jj)$ (vector sum of jet pT);
- If two or more candidates are found in the merged jet category, we take the one with the largest $p_T(J)$;
- If there is one in merged and one in resolved jet category, we take the following decision. We define a “good merging region” by the selections $pT(\ell\ell) > 200$ GeV, $pT(J) > 300$ GeV and $\tau_{21}(J) < 0.6$. If we have a merged jet + dilepton candidate with these requirements, the merged jet is preferred *regardless* of its pruned mass being in signal or sideband region. Otherwise resolved is chosen.

Notice that this selection of a high signal-purity region minimizes the possibilities of rejecting a good event because it falls in the sideband region of the preferred (merged) topology.

4.10.3 Kinematic fitting and Kinematic discriminants

4.10.3.1 Kinematic fitting

In order to improve the $2\ell 2q$ invariant mass resolution a kinematic fit is performed using a mass constraint on the intermediate $Z \rightarrow q\bar{q}$. The basic idea is to re-evaluate the kinematics of final state particles (here the pT of two jets forming the Z bosons of the Higgs candidate with a constraint on the reconstructed Z mass to follow the Z boson true lineshape. For each event, the likelihood is maximised and pT information of the refitted leptons are updated. After this kinematic constraint/refitting, the mass of the Z candidate and $m_{2\ell 2q}$ are re-calculated.

This framework is also being updated to fully model $H \rightarrow ZZ \rightarrow \ell^+\ell^- q\bar{q}$ kinematics, as discussed in the next section. There is an option to constrain full kinematics including all the angles and masses, characterising the decay, and obtain refitted values of the momenta and angles of the leptons and quarks. However, in reduced form only constraint to the two-dimensional correlated distribution of (m_1, m_2) is applied and the transverse momenta of the four fermions are refitted. This implementation is based on analytical MELA probability distribution $\mathcal{P}(m_1, m_2, \vec{\Omega} | m_{4f})$, which is conditional on the m_{4f} value.

4.10.3.2 Kinematic discriminants

The full kinematic information from each event using either the H boson decay or associated particles in its production is extracted using the matrix element calculations. These discriminants use a complete set of mass and angular input observables $\vec{\Omega}^{14,42}$ to describe kinematics at LO in QCD, see Fig. ???. The pT of either the combined H boson and 2 jets system for the production discriminant or the H boson itself for the decay discriminants is not included in the input observables in order to reduce associated QCD uncertainties.

The kinematic discriminants used in this study are computed using the MELA package,^{?,28,42} which provides the full set of processes studied in this paper and uses JHUGEN matrix elements for the signal and MCFM matrix elements for the background.

The discriminant sensitive to the $Z + JJ$ kinematics is calculated as

$$\mathcal{D}_{Zjj} = \left[1 + \frac{c_{Zjj} \times \mathcal{P}_{Zjj}(\vec{\Omega}^{H \rightarrow 2\ell 2q} | m_{2\ell 2q})}{\mathcal{P}_{H \rightarrow 2\ell 2q}(\vec{\Omega}^{H \rightarrow 2\ell 2q} | m_{2\ell 2q})} \right]^{-1}, \quad (4.15)$$

where the denominator contains the probability for the signal and the numerator includes the probability for the dominant $Z+jj$ background process, all calculated either with the JHUGEN (signal) or MCFM (background) matrix elements within the MELA framework. The value of c_{Zjj} is tuned as a function of m_{ZZ} in order to preserve good population of events in the range [0,1].

Figure 4.29 shows the distribution of \mathcal{D}_{Zjj} for several representative examples. Figures 4.30 to 4.33 show the distribution of \mathcal{D}_{Zjj} vs $m_{2\ell 2q}$ with conditional normalisation for signal and background. The template $\mathcal{T}(\mathcal{D}_{Zjj}|m_{2\ell 2q})$ used in analysis is conditionally normalised such that each slice of \mathcal{D}_{Zjj} is normalised to unit area for a given value of $m_{2\ell 2q}$. An analogous discriminant is defined for spin-2 signals, where the denominator probability is based on bulk graviton production and decay.

To take the different kinematics in resolved jets and in subjets (of merged jet) into account, different templates on \mathcal{D}_{Zjj} vs m_{ZZ} are derived for events passing resolved-jet based selection and merged-jet based selection. Templates for spin-0 case are given in Fig. 4.30 and 4.31, while templates for spin-2 case are given in Fig. 4.32 and 4.33.

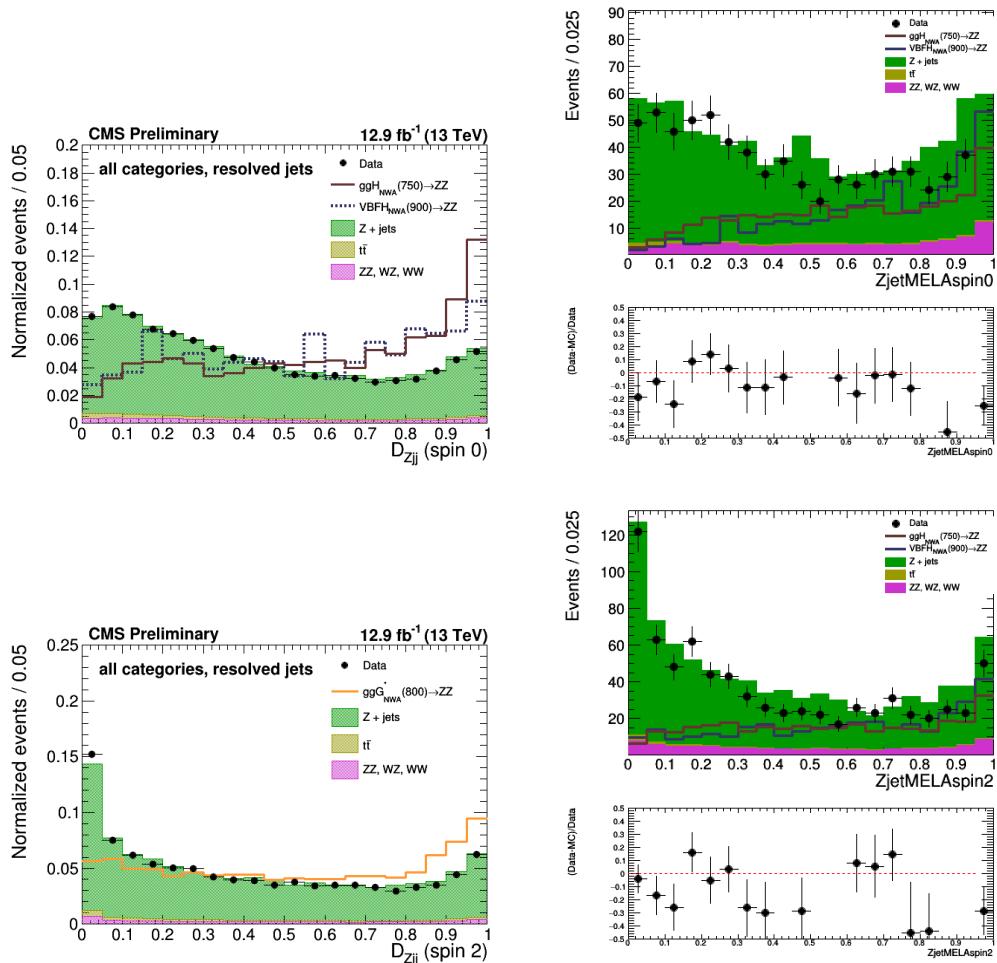


Figure 4.29: Distribution of \mathcal{D}_{Zjj} with events with two resolved jets (left) and with a merged jet (right). Spin-0 at top, spin-2 at bottom.

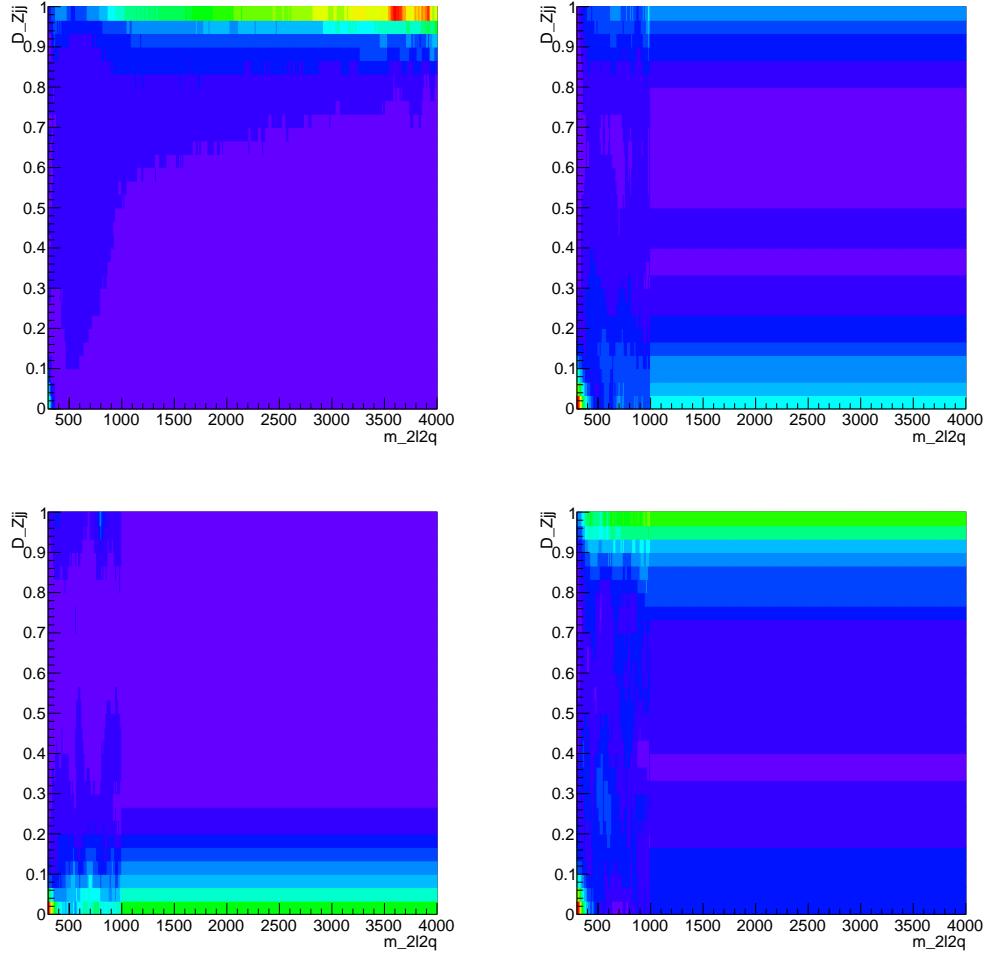


Figure 4.30: Distribution of \mathcal{D}_{Zjj} vs $m_{2\ell 2q}$ with conditional normalisation for spin-0 signal (top left), DY + jets (top right), $t\bar{t} + WW$ (bottom left), ZW, ZZ diboson backgrounds (bottom right), for the merged selection. When conditionally normalised, each slice of \mathcal{D}_{Zjj} is normalised to unit area for a given value of $m_{2\ell 2q}$.

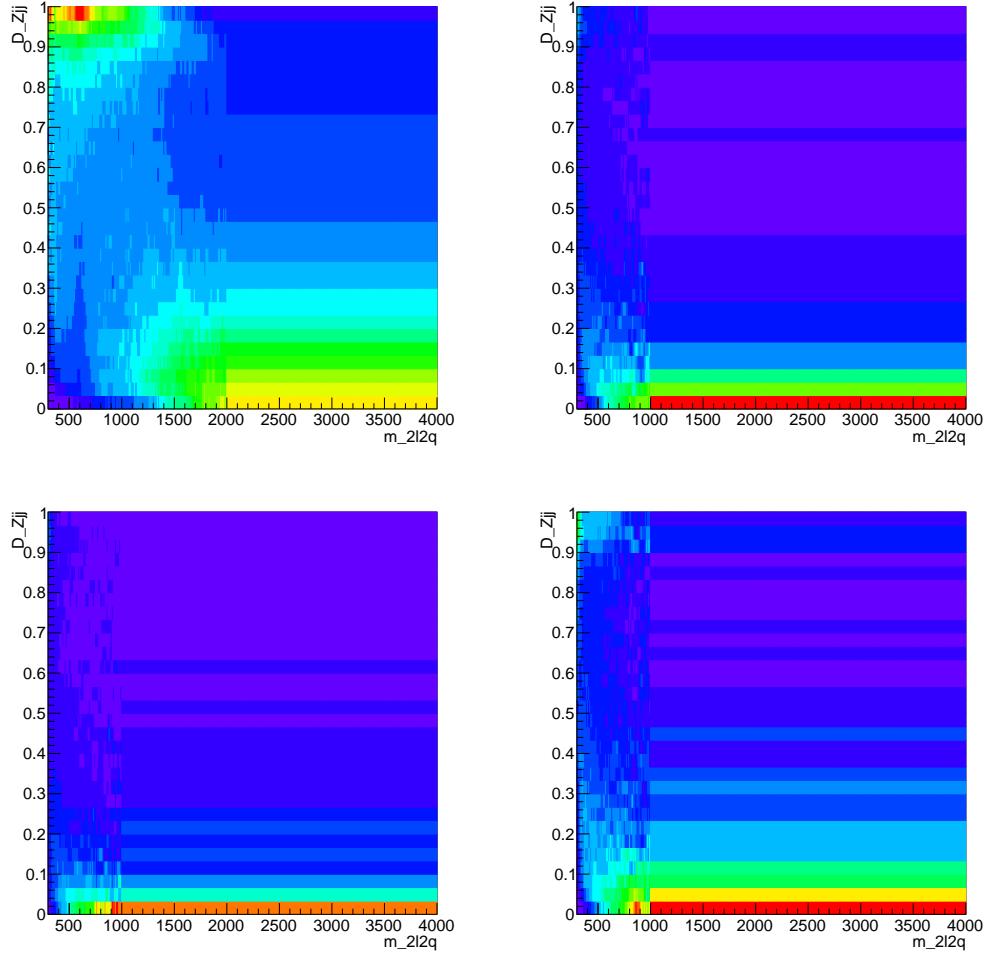


Figure 4.31: Distribution of \mathcal{D}_{Zjj} vs $m_{2\ell 2q}$ with conditional normalisation for spin-0 signal (top left), DY + jets (top right), $t\bar{t} + WW$ (bottom left), ZW, ZZ diboson backgrounds (bottom right), for the resolved selection. When conditionally normalised, each slice of \mathcal{D}_{Zjj} is normalised to unit area for a given value of $m_{2\ell 2q}$.

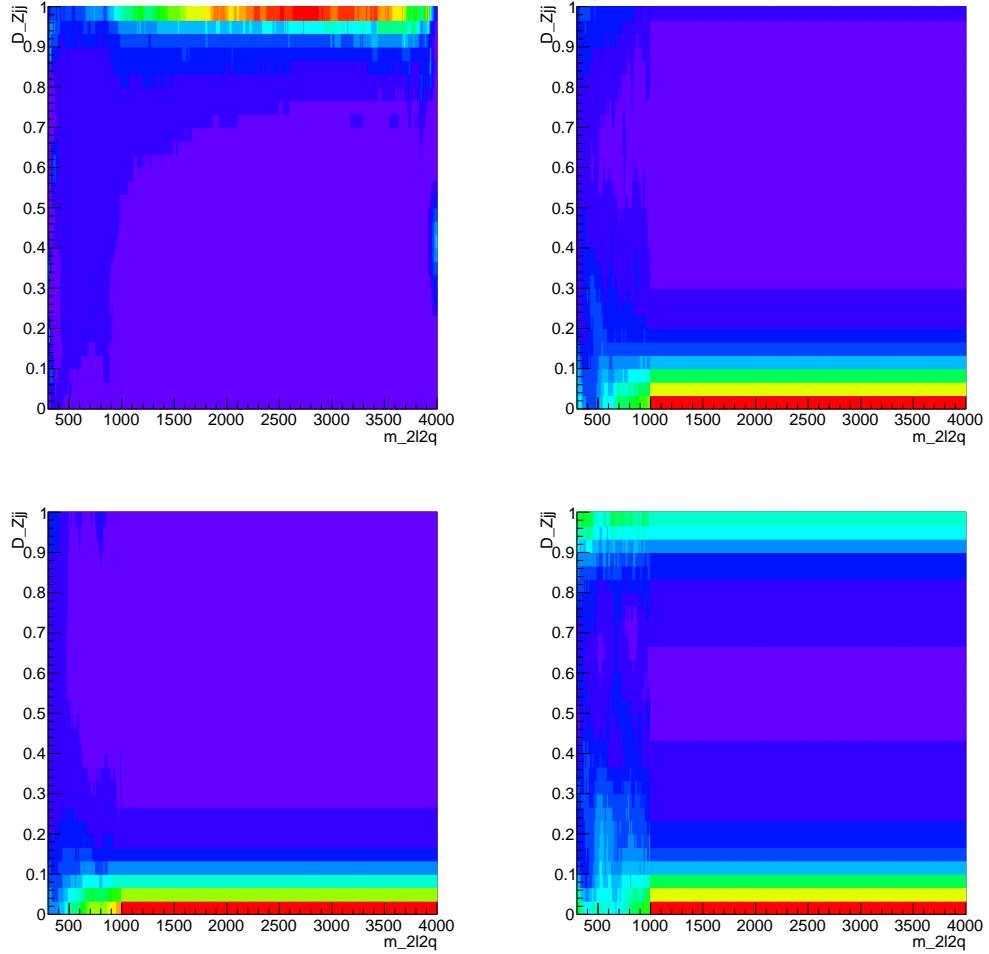


Figure 4.32: Distribution of \mathcal{D}_{Zjj} vs $m_{2\ell 2q}$ with conditional normalisation for spin-2 signal (top left), DY + jets (top right), $t\bar{t} + WW$ (bottom left), ZW, ZZ diboson backgrounds (bottom right), for the merged selection. When conditionally normalised, each slice of \mathcal{D}_{Zjj} is normalised to unit area for a given value of $m_{2\ell 2q}$.

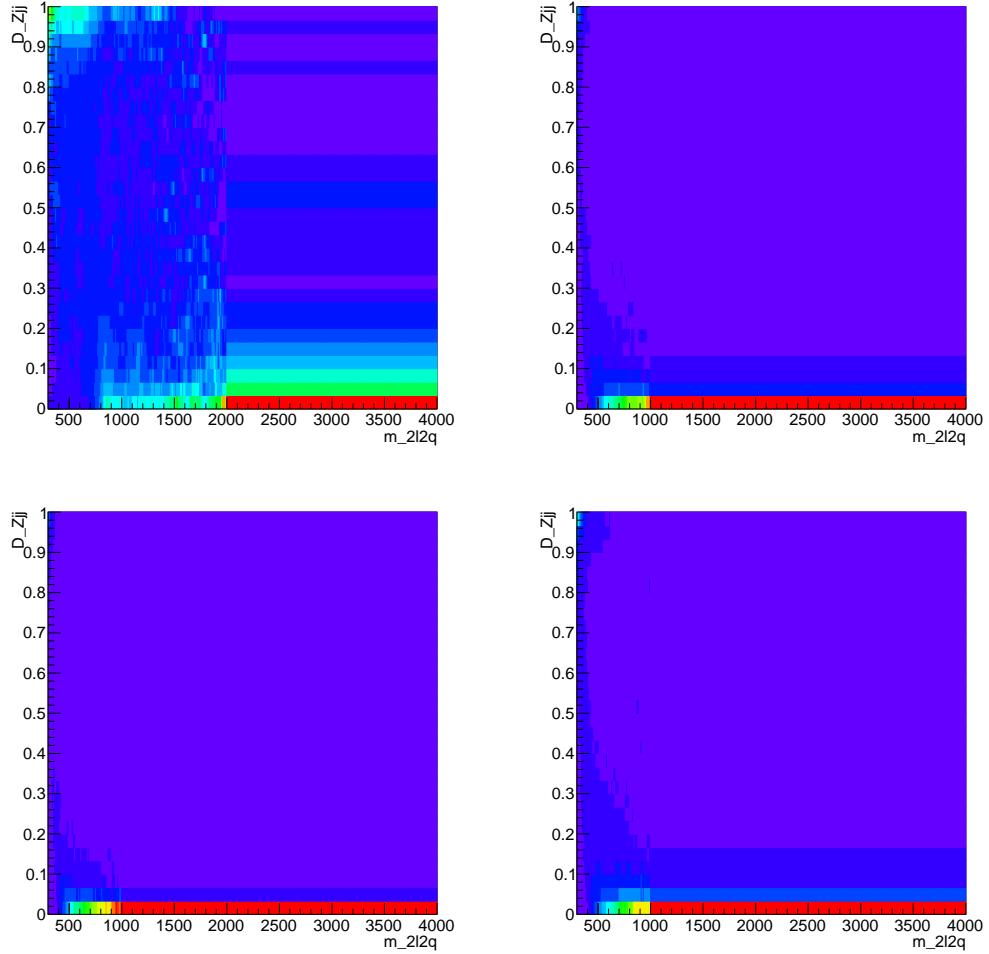


Figure 4.33: Distribution of \mathcal{D}_{Zjj} vs $m_{2\ell 2q}$ with conditional normalisation for spin-2 signal (top left), DY + jets (top right), $t\bar{t} + WW$ (bottom left), ZW, ZZ diboson backgrounds (bottom right), for the resolved selection. When conditionally normalised, each slice of \mathcal{D}_{Zjj} is normalised to unit area for a given value of $m_{2\ell 2q}$.

The discriminant sensitive to the VBF signal topology with two associated jets is calculated as^{7,55}

$$D_{2\text{jet}} = \left[1 + \frac{c_{\text{HJJ}} \times \mathcal{P}_{\text{HJJ}}(\vec{\Omega}^{\text{H+JJ}} | m_{2\ell 2q})}{\mathcal{P}_{\text{VBF}}(\vec{\Omega}^{\text{H+JJ}} | m_{2\ell 2q})} \right]^{-1}, \quad (4.16)$$

where \mathcal{P}_{VBF} and \mathcal{P}_{HJJ} are probabilities obtained from the JHUGEN matrix elements for the VBF process and gluon fusion (technically combination of $gg/qg/qq'$ parton collisions) in association with two jets ($H + 2$ jets) within the MELA framework. The value of c_{HJJ} is tuned as a function of mass in order to preserve good population of events in the range [0,1].

4.10.4 Event categorisation and final selection

Categorisation of events is based on topology of associated jets. The categorisation is done separately for the merged and resolved selections.

Events with at least two jets that meet the requirement $D_{2\text{jet}} > 0.5$, using the two highest-pT jets not belonging to signal, are assigned to the *VBF-tagged category*.

The rest of the events are further categorised as follows.

To further improve the signal-over-background ratio, we apply further categorisation only for events that are not VBF-tagged based on (sub)jet b-tagging. A resolved jet (or subjet of a merged jet) is considered to be b-tagged if it passes the loose working point of the CSVv2 algorithm. Events are categorised into *b-tagged category* where both resolved jets (or subjets of the merged jet) from hadronic Z decay are

considered to be b-tagged. Events that are not selected into the b-tagged category form the *untagged category*.

Table 4.11 summarises the final event selection, including the threefold categorisation.

Table 4.11: Summary of selection requirements and categorisation. Signal and sideband regions are defined by ranges in the hadronic Z boson candidate mass Z_{had} , after all other selection criteria are applied.

Observable	Selection
	same lepton flavour, opposite charge
$pT(\ell)(\text{GeV})$	$> 24/40$
$ \eta(\ell) $	$(e) < 2.5, (\mu) < 2.4$
$M(\ell^+\ell^-)(\text{GeV})$	$[60, 120]$
$pT(\ell^+\ell^-)(\text{GeV})$	> 100
$pT(\text{jet}) (\text{GeV})$	> 30
$ \eta(\text{jet}) $	< 2.4
$\Delta R(\ell, \ell)$	> 0.02
$\Delta R(\ell, \text{jet})$	> 0.4 (AK4 jets) > 0.8 (AK8 jets)
$M(Z_{\text{had}})(\text{GeV})$	$[40, 70]$ (lower sideband) $[70, 105]$ (signal region) $[105, 135]$ (Higgs boson region, not used in this analysis) $[135, 180]$ (upper sideband)
$pT(Z_{\text{had}})(\text{GeV})$	> 100 (resolved) > 170 (merged)
τ_{21}	< 0.6 (merged only)
VBF-tagged	m_{ZZ} -dependant selection on $\mathcal{D}_{2\text{jet}}$
b-tagged	not VBF-tagged and two b-tagged components of Z_{had}
untagged	not VBF-tagged and not b-tagged

4.10.5 Signal Modeling

Parameterization of the signal can be separated in two steps: modeling of the ideal distributions, both the shape and yield, and introducing detector effects.

The ideal distribution for a narrow resonance is simple, which can be taken essentially as a delta function when convoluting with a resolution function. Parameterization of a broad resonance or any other signal below $2m_Z$ threshold, requires special care of modeling the line-shape and interference with background. We will not cover these cases here. A narrow resonance is the limiting case in this model of the width going to zero.

The main observable in this analysis is the di-boson mass, generically called $m_{2\ell 2q}$ even for the case of a merged jet generated by the quark-antiquark pair. Therefore, parameterization of $m_{2\ell 2q}$ is of primary interest here.

4.10.5.1 Signal selection efficiency

Parameterization of efficiency function $\mathcal{E}(m_{2\ell 2q})$ is extracted for each final state ($2e2J$, $2e1J$, $2\mu 2J$, $2\mu 1J$) separately from the full simulation of each production mechanism (gluon fusion or VBF) and each signal type for the wide-resonance with $\Gamma_X = \Gamma_H^{\text{SM}}(m_H)$. To have a smooth estimation of efficiency in our search region, a total of 35 mass points are used for fits, overcovering this region: starting at 300 GeV and above in increments of 50 GeV up to 3 TeV. The efficiency parameterization is shown in Fig. 4.34 for a SM-like signal with $J^P = 0^+$ in the gluon fusion production,

in Fig. 4.35 for a SM-like signal with $J^P = 0^+$ in the VBF production, in Fig. 4.36 for a spin-2 signal with $J^P = 2_b^+$ in the model of KK graviton propagating in the bulk in the gluon fusion production. Parameterization of spin-2 model is performed using spin-0 simulation re-weighted with the MELA weights to account for different kinematics. In general, efficiency for the 2_b^+ model is higher than for isotropic spin-0 decay because of higher acceptance efficiency in the spin-2 model with the Z bosons predominantly in the central region of the detector.

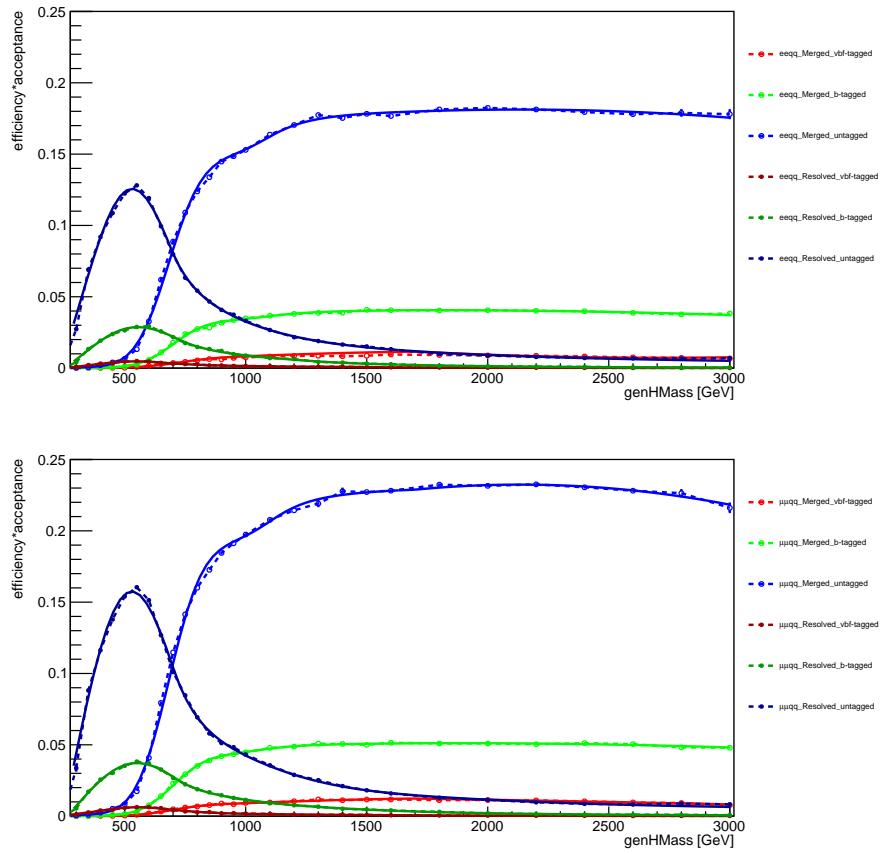


Figure 4.34: Parameterization of reconstruction efficiency $\mathcal{E}(m_{2\ell 2q})$ (including acceptance) (left: $Z \rightarrow ee$, right: $Z \rightarrow \mu\mu$) for a SM-like signal with $J^P = 0^+$ in the gluon fusion production. Lines refer to the 6 jet categories.

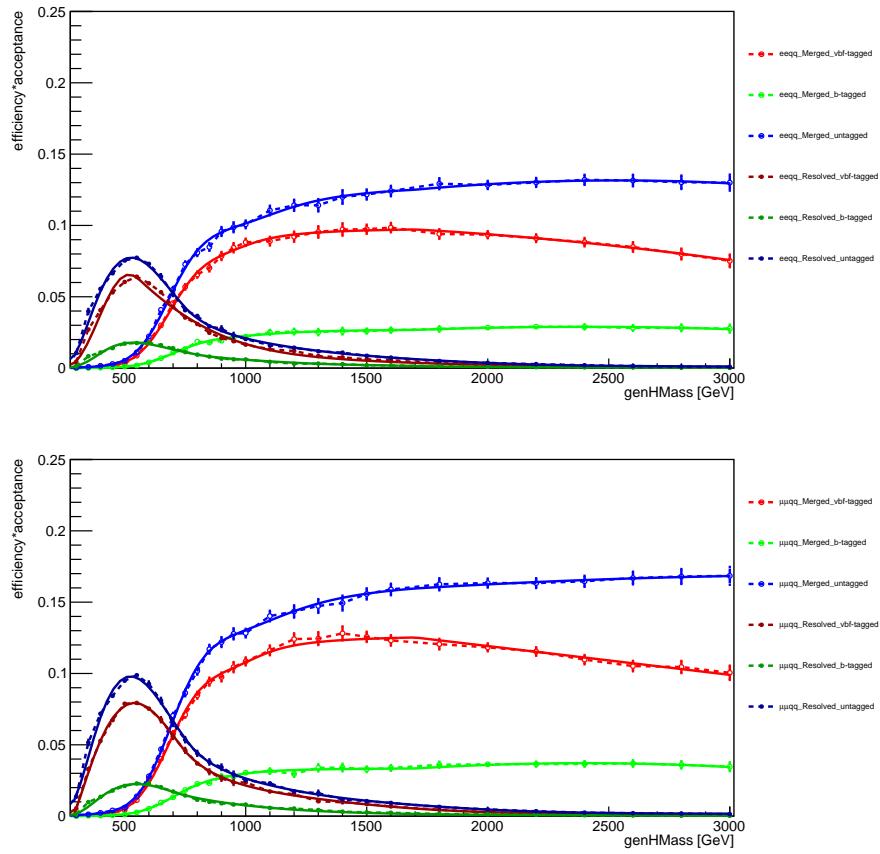


Figure 4.35: Parameterization of reconstruction efficiency $\mathcal{E}(m_{2\ell 2q})$ (including acceptance) (left: $Z \rightarrow ee$, right: $Z \rightarrow \mu\mu$) for a SM-like signal with $J^P = 0^+$ in the VBF production. Lines refer to the 6 jet categories.

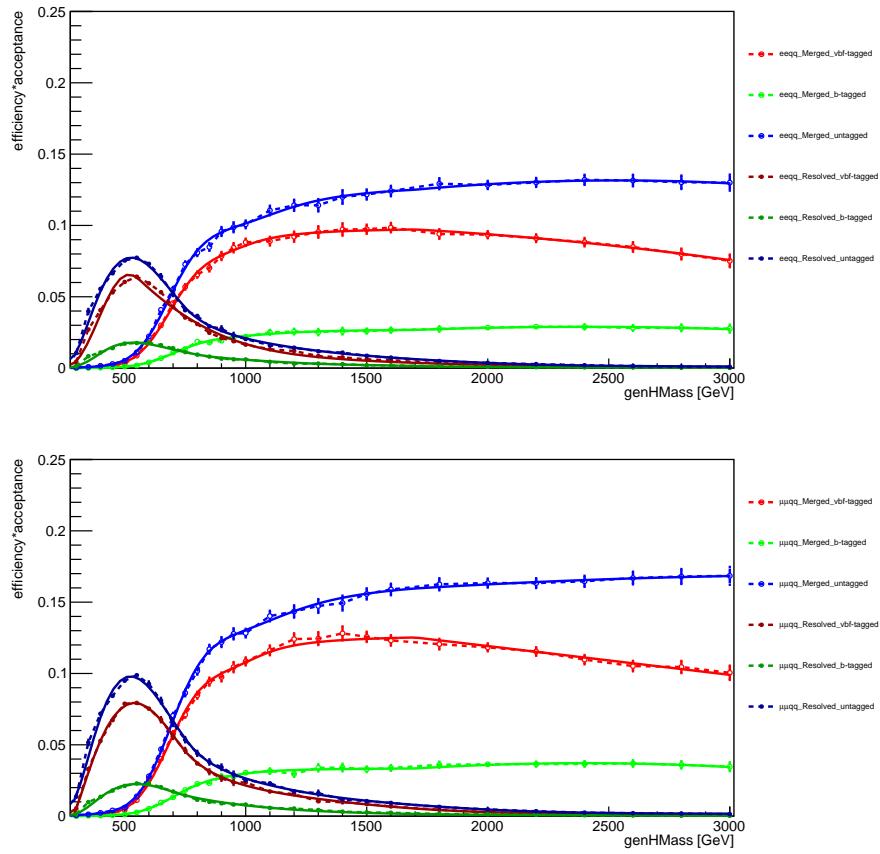


Figure 4.36: Parameterization of reconstruction efficiency $\mathcal{E}(m_{2\ell 2q})$ (including acceptance) (left: $Z \rightarrow ee$, right: $Z \rightarrow \mu\mu$) for a spin-2 signal with $J^P = 2_b^+$ in the model of KK graviton . Lines refer to the 6 jet categories.

4.10.5.2 Signal mass model: resolution

Parameterization of the resolution function $\mathcal{R}(m_{2\ell 2q} | m_{2\ell 2q}^{\text{ideal}})$ is extracted for each final state ($2e2J$, $2e1J$, $2\mu 2J$, $2\mu 1J$), but with resolution effects dominated by jet reconstruction, the electron and muon channels are merged. Parameterization is obtained from the full simulation of the $gg \rightarrow X \rightarrow 2\ell 2q$ process with POWHEG+JHUGen for the wide-resonance with $\Gamma_X = \Gamma_H^{\text{SM}}(m_H)$. We find that resolution in other production mechanisms and in the spin-2 model are generally very similar to the SM gluon fusion production. Therefore, a common parameterization of resolution can be used.

Several representative examples are shown in Fig. 4.37. Dependence of six parameters of the double CrystalBall function is parameterized with the glued polynomial functions in several mass regions, which preserve the height and derivative of the function at the junctions. After initializing the parameters from individual fits in each mass window, a joint fit is performed all mass distributions at the same time, leading to a smooth and optimal dependence of all parameter on $m_{2\ell 2q}$, see Fig. 4.38 for such dependence.

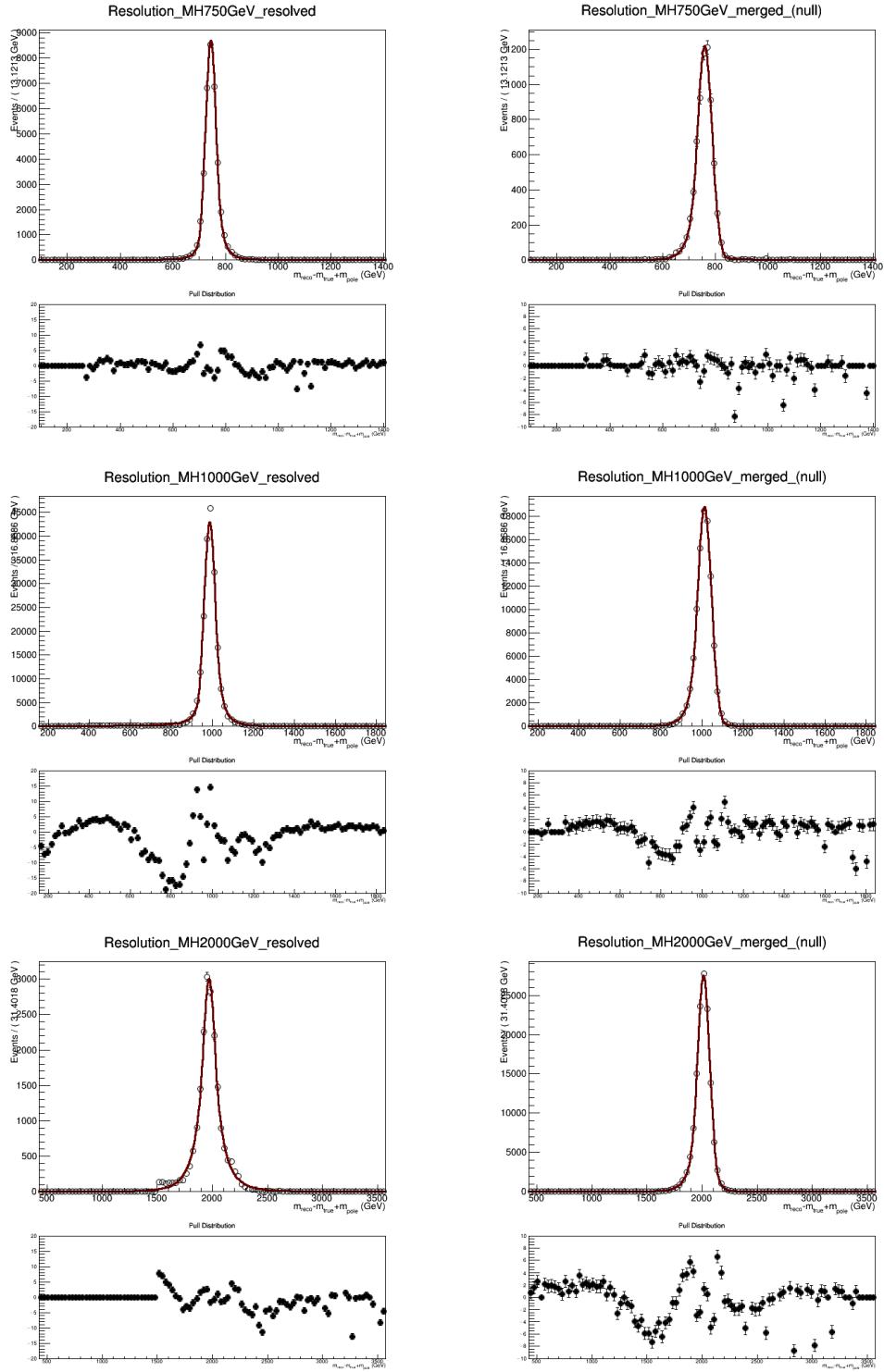


Figure 4.37: Representative $m_{2\ell 2q}$ distributions of a resolution function

$\mathcal{R}(m_{2\ell 2q} | m_{2\ell 2q}^{\text{ideal}})$ two-jet (left) and merged-jet (right) categories. Several examples

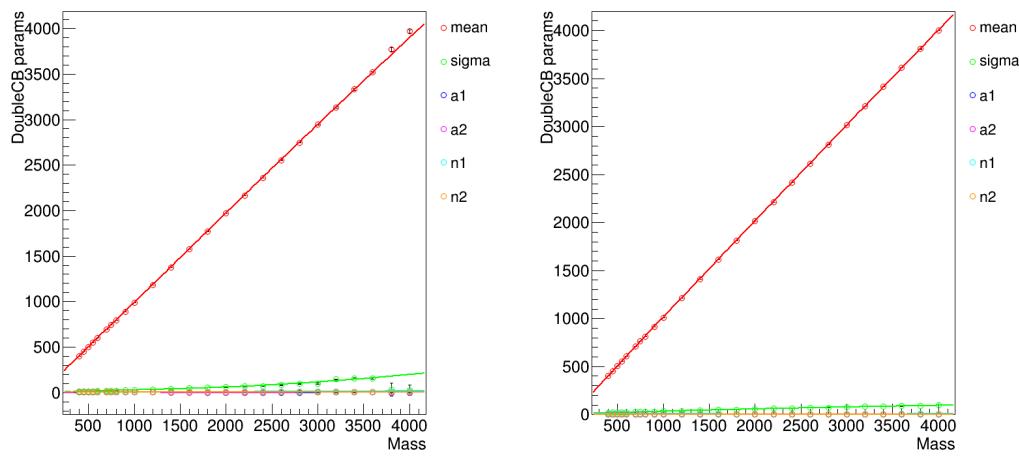


Figure 4.38: Dependence of six parameters of the resolution function in the two-jet (left) and merged-jet (right) categories. Point show results of individual fits, smooth curves show parameterization from a simultaneous fit of all distributions.

4.10.5.3 Signal model with kinematics

The final signal parameterization in 2D is built with conditional template $\mathcal{T}(D_{Zjj}|m_{2\ell 2q})$ which describes D_{Zjj} discriminant for each value of $m_{2\ell 2q}$.

$$\mathcal{P}_{\text{sig}}^{\text{reco}}(m_{2\ell 2q}, D_{Zjj}) = \mathcal{P}_{\text{sig}}^{\text{reco}}(m_{2\ell 2q}) \times \mathcal{T}(D_{Zjj}|m_{2\ell 2q}) \quad (4.17)$$

The template $\mathcal{T}(D_{Zjj}^{\text{kin}}|m_{2\ell 2q})$ parameterization includes all detector effects affecting the D_{Zjj} distribution. Definition of D_{Zjj} depends on the signal spin model.

4.10.6 Background Estimation

The following SM processes are considered as background in this analysis: diboson, Z+jets, and tt production.

The diboson production (mainly ZZ and WZ) is simulated using MC while the other two contributions are estimated either using MC simulation corrected to data in control regions (Z+jets), or extracted directly from data in control regions (tt). The latter two cases are described in detail in the following subsections.

4.10.6.1 Estimation of ttbar from data

The tt background is an important source of contamination in the btagged category. It is estimated from the data using in control region built with events in $e^\pm \mu^\mp$ final state passing the same kinematic cuts as the signal region. The control region also accounts for other small backgrounds (as WW + jets, Z $\rightarrow \tau^+\tau^-$ + jets,

single top, fakes) where the lepton flavor symmetry can be invoked as well. The Run-I $H \rightarrow ZZ \rightarrow \ell^+ \ell^- q\bar{q}$ uses the similar technique to estimate this background component, where specific checks and studies were performed in order to validate the method and quantify the related uncertainties.[?]

MC studies show that the $e^\pm \mu^\mp$ v.s. $e^+ e^- + \mu^+ \mu^-$ symmetry works at the level of the shapes of the distributions of all considered variables and of distributions in all btagging-based categories as shown in figure. ?? for resolved jet based selection and figure. ?? for merged jet based selection.

Due to the limit statistic in data with $e^\pm \mu^\mp$ final state, the di-boson mass shape is taken from $t\bar{t}$ and WW MC simulations, while the normalizatin is control by the number of event in $e^\pm \mu^\mp$ data control region as

$$N_{tt+WW}^{SR}(m_{ZZ}) = N_{tt+WW \text{ MC}}^{SR}(m_{ZZ}) \times \frac{N_{Data}^{e^\pm \mu^\mp CR}}{N_{tt+WW \text{ MC}}^{e^\pm \mu^\mp CR}}, \quad (4.18)$$

where $N_{tt+WW}^{SR}(m_{ZZ})$ is the mass spectrum from $t\bar{t}$ and WW MC simulations, and $\frac{N_{Data}^{e^\pm \mu^\mp CR}}{N_{tt+WW \text{ MC}}^{e^\pm \mu^\mp CR}}$, the ration of number events between data and MC in $e^\pm \mu^\mp$ control region, is used to weight the normalization.

4.10.6.2 Estimation of Z+jets

Two methodology to predict Z+jets background are performed, one of which use the control region data for prediction, the other extract the shape from MC simulation and normalization for data in sideband region.

While the event selection described in Sec. ?? is extremely effective, there is still a sizeable background from SM processes. The $Z + \text{jets}$ background is the dominant SM background to this search. Events from this background are characterised by having a real $Z \rightarrow ll$ decay, while associated jets from regular q/g emission are misidentified as coming from a $Z \rightarrow q\bar{q}$ decay, and are combined to generate a misidentified hadronic V boson.

The misidentified V comes either from the combinatoric background for the resolved category – $Z + 2$ jets events where the dijet system that happens to have an invariant mass in the range compatible with that of the Z boson – or from an unusual parton-shower + hadronization development for a single jet, leading to a configuration similar to that of the boosted $V \rightarrow q\bar{q}$ decay in the merged case. In both cases, a sideband region with a misidentified hadronic V mass close to that of the signal region can be used to estimate the contribution of this background. To address the correlation between the hadronic V mass and the m_{ZZ} in these configurations, a correction factor is estimated from simulation. This is the essence of the *alpha method*, which we proceed to describe in detail now.

The sideband region is defined as: the set of events that pass the nominal selection **except** the requirement on the hadronic V mass. Instead, candidate events for the sideband region are required to have a hadronic V mass in the range 40-70 GeV or 135-180 GeV. Notice that every category in the nominal selection has its accompanying sideband region, i.e., the background estimation is independent for every signal region.

The *alpha transfer function* $\alpha(m_{ZZ})$ is defined as

$$\alpha(m_{ZZ}) = \frac{N_{\text{SIG}}^{\text{MC}}(m_{ZZ})}{N_{\text{SB}}^{\text{MC}}(m_{ZZ})}, \quad (4.19)$$

calculated from the ratio of the m_{ZZ} distributions in the sideband and signal regions calculated in the Z+jets simulation. The expectation is that, by multiplying the alpha function by the the Z+jets m_{ZZ} distribution in the data sideband region, one can obtain an estimate of the Z+jets m_{ZZ} distribution in the data signal region. Naturally, it is impossible to obtain a pure Z+jets distribution in data; we estimate that object by subtracting the subdominant backgrounds from the data sideband region.

The method as described above has a small flaw; it relies on simulation to predict the ratio of the yields of the Z+jets process in the sideband and signal regions, in addition to the ratio of the shapes. In order to make the method more robust, we use a second method to predict the Z+jets yield in the signal region from the sidebands of the hadronic V mass distribution. Again, we estimate a pure Z+jets distribution in the hadronic V mass by subtracting the subdominant backgrounds from the data. We then estimate the total yield of Z+jets events in the signal region by fitting the hadronic V mass distribution in the lower sideband region ($40 - 70 GeV$) and upper sideband region ($135 - 180 GeV$). The fit is then used to estimate the Z+jets background yield in the signal region.

To summarise, the method to estimate the Z+jets background comprises the following steps:

- Estimate the Z+jets m_{ZZ} distribution in the data sideband region by subtracting the subdominant backgrounds from the data.
- Estimate the sideband to signal alpha transfer function by the ratio of the m_{ZZ} distributions in the sideband and signal regions in Z+jets simulation.
- Estimate the shape of Z+jets m_{ZZ} distribution in the data signal region by the product of the alpha function and the Z+jets m_{ZZ} distribution in the data sideband region.

4.10.7 Statistical Analysis and Systematics

The search is performed for a narrow scalar resonance X(m_X) produced in gluon fusion and VBF process and decaying to $ZZ \rightarrow 2\ell 2q$ through a tree-level process. We also consider a spin-two resonance with gravity-like XVV couplings. In that case, we do not use VBF tagging, as this is expected to be predominantly produced.

The analysis of two observables is performed where the total probability density function also includes template parameterization of the \mathcal{D}_{Zjj} as $\mathcal{P}(m_{2\ell 2q}|m_X, \sigma_X) \times \mathcal{T}(\mathcal{D}_{Zjj}|m_{2\ell 2q})$. Both signal and background in either gluon fusion or VBF are treated as one process. The cross section σ_X is reported for the pure signal process only without including effects of interference. For a scalar X boson, the fraction of VBF production is parameterized with f_{VBF} , so that VBF cross section is $\sigma_X \times f_{VBF}$ and gluon fusion cross section is $\sigma_X \times (1 - f_{VBF})$. This number is varied from 0 to 1.

The final mass parameterization is built from efficiency, resolution, and ideal distributions as follows:

$$\mathcal{P}_{\text{sig}}^{\text{reco}}(m_{2\ell 2q}) = \mathcal{E}(m_{2\ell 2q}) \times (\mathcal{R}(m_{2\ell 2q}|m_{2\ell 2q}^{\text{ideal}}) \otimes \mathcal{P}_{\text{sig}}(m_{2\ell 2q}^{\text{ideal}}|m_X)) \times \mathcal{T}(\mathcal{D}_{Zjj}|m_{2\ell 2q}) \quad (4.20)$$

where the efficiency $\mathcal{E}(m_{2\ell 2q})$ and resolution $\mathcal{R}(m_{2\ell 2q}|m_{2\ell 2q}^{\text{ideal}})$ functions are obtained from full simulation of events with appropriate spin-parity properties and are parameterized analytically over a wide range of masses $200 < m_{2\ell 2q} < 3000$ GeV. For narrow width this simplifies to:

$$\mathcal{P}_{\text{sig}}^{\text{reco}}(m_{2\ell 2q}) = \mathcal{E}(m_{2\ell 2q}) \times \mathcal{R}(m_{2\ell 2q}|m_{2\ell 2q}^{\text{ideal}}) \times \mathcal{T}(\mathcal{D}_{Zjj}|m_{2\ell 2q}) \quad (4.21)$$

The systematic uncertainties influence both the normalization and shape of the backgrounds and signal.

The theoretical uncertainties summarized in Table. ?? come from the QCD scale uncertainty and uncertainty on the PDF set which apply on both signal and background. In addition, acceptance uncertainty and di-Z boson decay branch ratio are included and the uncertainty on the NNLO-to-NLO K factor of ZZ and WZ cross section is about 10%.

The experimental uncertainties originate from lepton/jet reconstruction and identification, lepton/jet energy scale and resolution, b-tagging and the background estimation methodology and are summarized in Table.4.12.

Summary of theoretical systematic uncertainties	
QCD scale (Z/WZ)	3 – 10 %
PDF set (Z/WZ)	3 – 5 %
NNLO/NLO K factor (Z/WZ)	10 %
QCD scale ($q\bar{q} \rightarrow VBF/VH, gg \rightarrow H/ttH$)	3 – 10 %
PDF set ($q\bar{q} \rightarrow VBF/VH, gg \rightarrow H/ttH$)	3 – 4 %
Acceptance	2 %
$BR(X \rightarrow ZZ \rightarrow \ell^+ \ell^- q\bar{q})$	2 %

The uncertainty on the knowledge of the integrated luminosity of the data sample (6.2%) introduces an uncertainty on the number of signal events and $V(W/Z)Z$ process passing the final selection. This uncertainty is fully correlated in all channels. The systematic uncertainties in the lepton trigger, identification, and isolation efficiencies are derived using tag-and-probe analysis of $Z \rightarrow ll$ events. 4%(8%) uncertainty is applied and fully correlated in electron(muon) channels as the uncertainty of the number of signal and $V(W/Z)Z$ process passing the final selection.

The uncertainties on the signal selection efficiency arise from uncertainties in the efficiency of tagging the hadronic jet as a V in the low-mass and high-mass boosted categories, and from uncertainties in the b-tagging efficiency in the low-mass b-tagged categories. The efficiency of the W tagging selection and its corresponding systematic uncertainty has been measured in data using a sample enriched in semileptonics $t\bar{t}$ events.[?] The b-tagging efficiencies and their corresponding systematic uncertainties have been measured in data using samples enriched in beauty decay events.[?]

Uncertainties on the reconstruction of leptons/jets affect not only the signal ef-

ficiency, but also on the shape of the m_{ZZ} distribution. The momentum and the mass of the reconstructed jets are rescaled according to the uncertainties of the jet energy-momentum scale.⁷ And the uncertainty on the lepton energy scale is 0.3% for electrons and 0.1% for muons and 20% uncertainty lepton resolution is applied. The lepton/jet energy scale and resolution energy are propagated to the uncertainty on the signal di-Z boson mass shape.

The alpha method for the Z+jets background estimation depends on the uncertainty of extrapolation factor and statistics of the di-jet mass or pruned jet mass sideband region. The studies show that the jet energy scale and resolution can effect the extrapolation factor $\alpha(m_{ZZ})$, and the size of the uncertainty is about 3 to 10% for different categories.

Summary of the experimental systematic uncertainties.

4.10.8 Results

4.10.8.1 Observed Data

The following tables summarize the expected signal/background events for the example mass of 800 GeV. The signal assumes to be gluon-fusion produced scalar particle with mass $800GeV$ with cross section $1 pb^{-1}$. The expected yield is calculated in a di-boson mass window from $600GeV$ to $950GeV$.

Table 4.13 lists the numbers of signal and background events expected for 12.9

Table 4.12: Summary of the experimental systematic uncertainties.

Summary experimental systematic uncertainties	
Common experimental uncertainties	
Luminosity	6.2 %
Lepton identification/reconstruction efficiencies	4/8% for electrons/muons
Lepton energy scale	0.3/0.1 % for electrons/muons %
Lepton energy resolution	20 %
Jet energy scale	1%
Jet energy resolution	10%
W-jet(merged jet) tagging efficiecy	3%
VBF tagging efficiency	5–10%
b tagging scale factor	5%–7%
t̄t+WW	statistical unc. of $N_{e\mu}$
Z+Jets	JES/JER on α 3 – 10 %
Z+Jets	binning of α sub-percent – 30 %
Z+Jets	uncertainty from fitting sideband data

fb^{-1} in the m_{jj} range [70, 105]GeV in the two categories, separately for the $\mu^+\mu^-jj$

and $EEjj$ channels (denoted $\mu\mu$ and ee, respectively).

Table 4.13: Expected yields in the categories passing resolved-jet selection.

category	M_H	signal		Z+jets		$t\bar{t}+ww$		diboson		total background	
		$\mu\mu$	ee	$\mu\mu$	ee	$\mu\mu$	ee	$\mu\mu$	ee	$\mu\mu$	ee
un-tagged	800	121.6	94.9	1205.7	920.0	31.7	31.0	52.1	14.6	1306.0	1003.1
b-tagged		5.2	4.0	21.8	14.8	0.4	0.3	1.4	1.1	23.6	16.2
VBF-tagged		5.2	4.1	34.7	26.5	0.106	0.0101	0.0992	0.0742	34.9	26.7

Table 4.14 lists the numbers of signal and background events in expected in the pruned jet mass range [70, 105]GeV in the two categories, separately for the $\mu^+\mu^-jj$ and e^+e^-jj channels (denoted $\mu\mu$ and ee, respectively).

Table 4.14: Expected yields in the categories passing merged-jet selection.

category	M_H	signal		Z+jets		$t\bar{t}+ww$		diboson		total background	
		$\mu\mu$	ee	$\mu\mu$	ee	$\mu\mu$	ee	$\mu\mu$	ee	$\mu\mu$	ee
un-tagged	800	280.2	217.2	152.7	118.6	4.02	4.04	18.4424	13.9	175.1	136.548
btagged		60.67	46.9	7.6	5.9	0.29	0.17	1.7	1.4	9.6	7.0
VBF-tagged		11.9	9.7200	2.1	1.6682	0.0116	0.0882	0.4187	0.34	2.5	2.1

4.10.8.2 Spin-zero Limits

Figure 4.39 shows expected constraints on the $pp \rightarrow X \rightarrow ZZ$ cross section for mass point from $m_X = 550$ GeV to 2 TeV with narrow width approximation. The merge jet category is only used when $m_X > 700$. The fraction of VBF production cross section is allowed to float free for each m_X hypothesis. The limits are improved by a few percent to about 20% due to including kinematic discriminant depending on m_X .

Limits assuming 100% gluon fusion production ($f_{VBF} = 0$) and 100% VBF production ($f_{VBF} = 1$) are shown in fig.4.40 and fig.4.41, respectively.

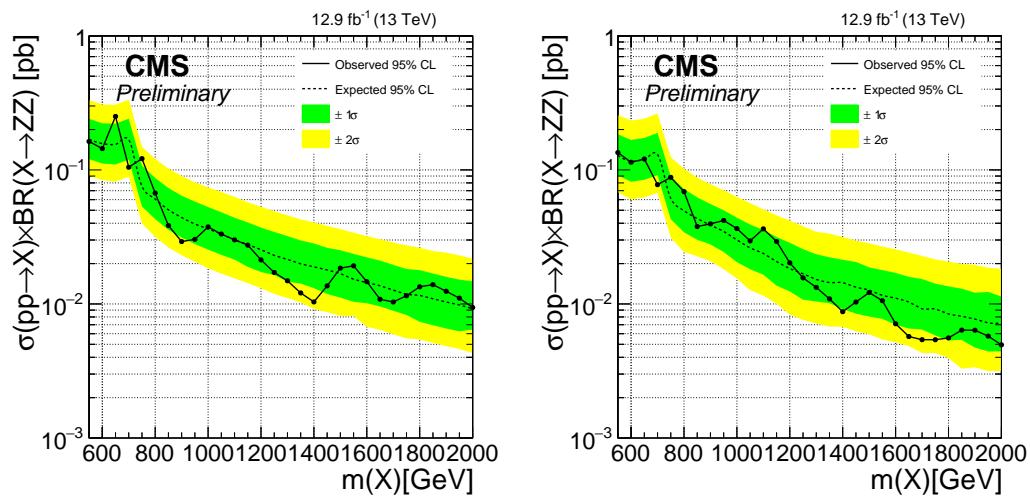


Figure 4.39: Expected and observed upper limits on the $pp \rightarrow X \rightarrow ZZ$ cross section under narrow width resonance assumption for 12.9 fb^{-1} of CMS data at 13 TeV where the fraction of VBF production is allowed to float. Left plot is done using m_{ZZ} as the observable and right plot is done with statistical model including both m_{ZZ} and D_{jet} information.

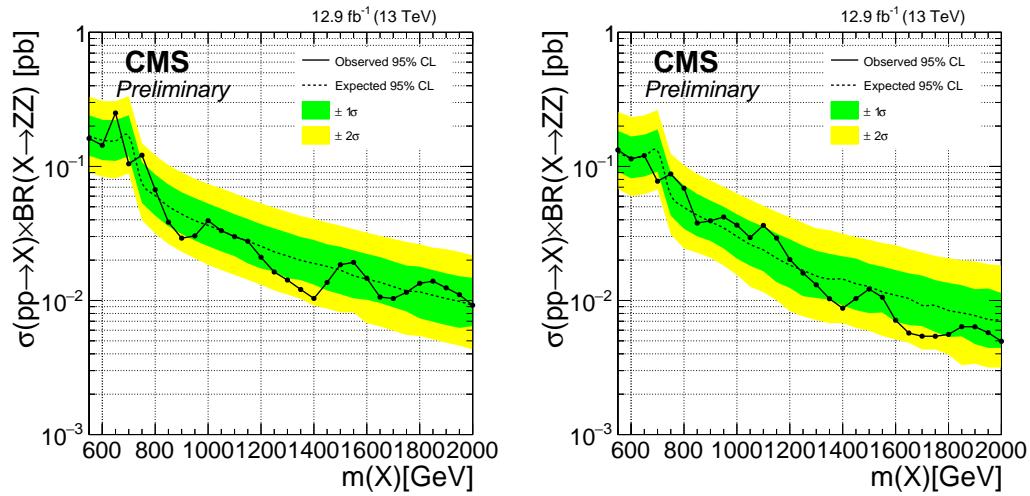


Figure 4.40: Expected upper limits on the $pp \rightarrow X \rightarrow ZZ$ cross section under narrow width resonance assumption for 12.9 fb^{-1} of CMS data at 13 TeV assuming 100% gluon fusion production. Left plot is done using m_{ZZ} as the observable and right plot is done with statistical model including both m_{ZZ} and D_{jet} information.

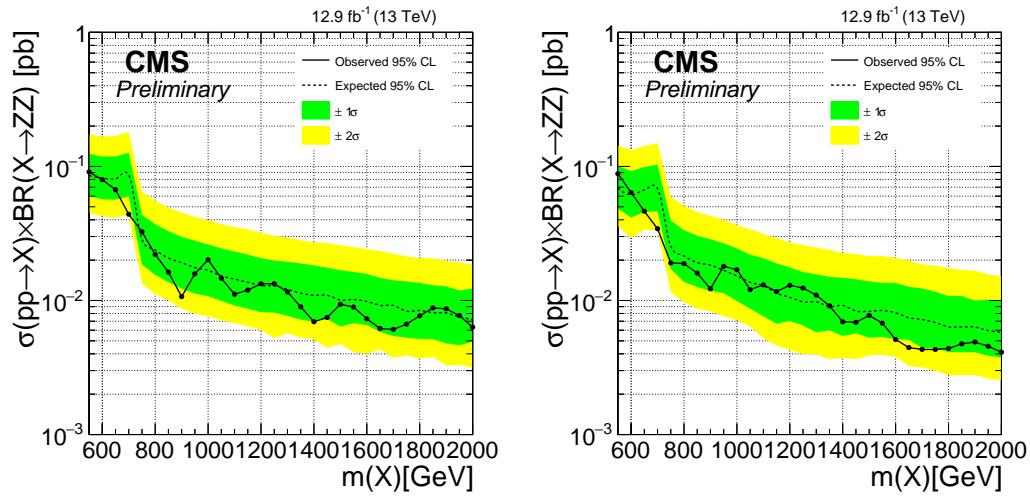


Figure 4.41: Expected upper limits on the $pp \rightarrow X \rightarrow ZZ$ cross section under narrow width resonance assumption for 12.9 fb^{-1} of CMS data at 13 TeV assuming 100% VBF production. Left plot is done using m_{ZZ} as the observable and right plot is done with statistical model including both m_{ZZ} and D_{jet} information.

4.10.8.3 Spin-two Limits

Figure 4.42 shows expected constraints on the $pp \rightarrow X \rightarrow ZZ$ cross section for mass point from $m_X = 550$ GeV to 2 TeV with narrow width approximation, assuming 100% gluon fusion production of graviton.

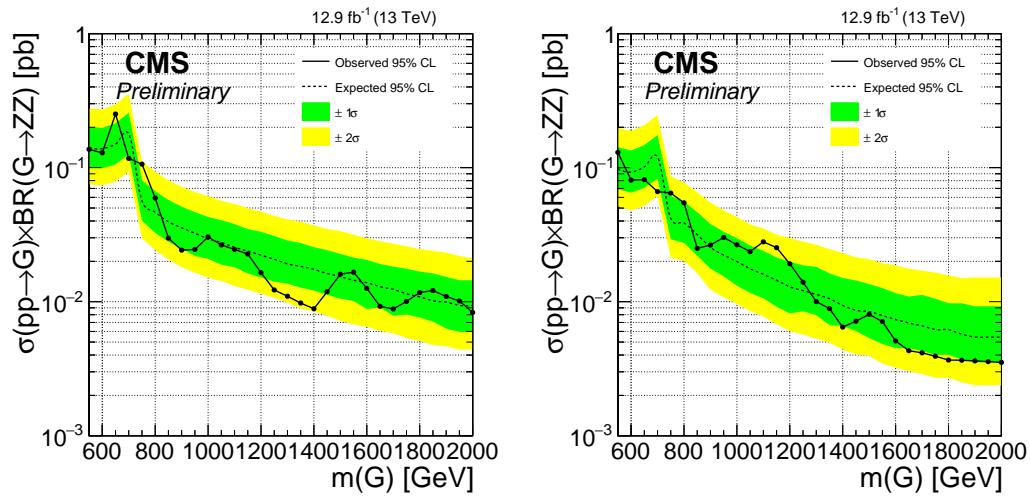


Figure 4.42: Expected and observed upper limits on the $pp \rightarrow X \rightarrow ZZ$ cross section under narrow width resonance assumption for 12.9 fb^{-1} of CMS data at 13 TeV with a spin-2 graviton assumption. Left plot is done using m_{ZZ} as the observable and right plot is done with statistical model including both m_{ZZ} and D_{jet} information.

4.11 Summary

The results obtained are compatible with the standard model prediction in the explored mass range. A set of limits on production cross section times decay branching

fraction of a scalar boson or spin-2 boson in the model with gravity propagating in the bulk of extra dimensions is obtained.

Chapter 5

Conclusion

HL-LHC

Bibliography

- [1] Commissioning of the particle-flow event reconstruction with the first LHC collisions recorded in the CMS detector. CMS Physics Analysis Summary CMS-PAS-PFT-10-001, 2010.
- [2] Measurement of differential and integrated fiducial cross sections for higgs boson production in the four-lepton decay channel in pp collisions at $\sqrt{s} = 7$ and 8 tev. 2015.
- [3] Measurement of differential cross sections for Higgs boson production in the diphoton decay channel in pp collisions at $\sqrt{s}=8$ TeV. *Eur. Phys. J. C*, 76:13, 2015.
- [4] Measurement of the properties of the higgs boson in the four-lepton final state at $\sqrt{s} = 13$ tev. CMS Physics Analysis Note CMS-AN-15-277, 2016.
- [5] W-like measurement of the z boson mass using di-muon events collected in pp collisions at $\sqrt{s} = 7$ tev. CMS Physics Analysis Summary CMS-PAS-SMP-14-007, 2016.

- [6] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys.Lett.*, B716:1–29, 2012.
- [7] Georges Aad et al. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett.*, 114:191803, 2015.
- [8] 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. Chytracek, 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. Gonzlez, 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. Zschiesche. Geant4 - a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250 – 303, 2003.

- [9] S. Agostinelli et al. GEANT4: a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.
- [10] S. Alioli, P. Nason, C. Oleari, and E. Re. NLO vector-boson production matched with shower in POWHEG. *JHEP*, 07:060, 2008.
- [11] B.C. Allanach, K. Odagiri, M.J. Palmer, Michael Andrew Parker, A. Sabetfakhri, et al. Exploring small extra dimensions at the large hadron collider. *JHEP*, 0212:039, 2002.
- [12] John Allison et al. Geant4 developments and applications. *IEEE Trans. Nucl. Sci.*, 53:270, 2006.

- [13] Charalampos Anastasiou et al. High precision determination of the gluon fusion higgs boson cross-section at the lhc. *JHEP*, 2016(5):1, 2016.
- [14] Ian Anderson, Sara Bolognesi, Fabrizio Caola, Yanyan Gao, Andrei V. Gritsan, et al. Constraining anomalous HVV interactions at proton and lepton colliders. *Phys.Rev.*, D89:035007, 2014.
- [15] Oleg Antipin and Amarjit Soni. Towards establishing the spin of warped gravitons. *JHEP*, 0810:018, 2008.
- [16] ATLAS and CMS Collaborations, LHC Higgs Combination Group. Procedure for the lhc higgs boson search combination in summer 2011. ATL-PHYS-PUB/CMS NOTE 2011-11, 2011/005, CERN, 2011.
- [17] E. Bagnaschi, G. Degrassi, P. Slavich, and A. Vicini. Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM. *JHEP*, 1202:088, 2012.
- [18] Richard Ball et al. Unbiased global determination of parton distributions and their uncertainties at nnlo and at lo. *Nucl. Phys. B*, 855(2):153, 2012.
- [19] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.
- [20] Vernon D. Barger, King-man Cheung, A. Djouadi, Bernd A. Kniehl, and P.M.

Zerwas. Higgs bosons: Intermediate mass range at e+ e- colliders. *Phys.Rev.*, D49:79–90, 1994.

- [21] Sara Bolognesi, Yanyan Gao, Andrei V. Gritsan, Kirill Melnikov, Markus Schulze, et al. On the spin and parity of a single-produced resonance at the LHC. *Phys.Rev.*, D86:095031, 2012.
- [22] Marco Bonvini, Fabrizio Caola, Stefano Forte, Kirill Melnikov, and Giovanni Ridolfi. Signal-background interference effects in $gg \rightarrow H \rightarrow WW$ beyond leading order. *Phys. Rev. D*, 88:034032, 2013.
- [23] C.P. Buszello, I. Fleck, P. Marquard, and J.J. van der Bij. Prospective analysis of spin- and CP-sensitive variables in $H \rightarrow ZZ \rightarrow l(1)^+l(1)^-l(2)^+l(2)^-$ at the LHC. *Eur.Phys.J.*, C32:209–219, 2004.
- [24] Nicola Cabibbo and Alexander Maksymowicz. Angular Correlations in K-4 Decays and Determination of Low-Energy pi-pi Phase Shifts. *Phys.Rev.*, 137:B438–B443, 1965.
- [25] John M. Campbell and R. K. Ellis. MCFM for the Tevatron and the LHC. *Nucl. Phys. Proc. Suppl.*, 205:10, 2010.
- [26] Stefano Catani and Massimiliano Grazzini. An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC. *Phys. Rev. Lett.*, 98:222002, 2007.

- [27] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012.
- [28] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012.
- [29] Serguei Chatrchyan et al. Performance of CMS muon reconstruction in pp collision events at $\text{sqrt}(s) = 7$ TeV. *JINST*, 7:P10002, 2012.
- [30] Serguei Chatrchyan et al. Measurement of the properties of a Higgs boson in the four-lepton final state. *Phys. Rev. D*, 89:092007, 2014.
- [31] Yi Chen, Nhan Tran, and Roberto Vega-Morales. Scrutinizing the Higgs Signal and Background in the $2e2\mu$ Golden Channel. *JHEP*, 1301:182, 2013.
- [32] S.Y. Choi, 2 Miller, D.J., M.M. Muhlleitner, and P.M. Zerwas. Identifying the Higgs spin and parity in decays to Z pairs. *Phys.Lett.*, B553:61–71, 2003.
- [33] CMS Collaboration. Particle-flow event reconstruction in cms and performance for jets, taus, and met. CMS Physics Analysis Summary CMS-PAS-PFT-09-001, 2009.
- [34] CMS Collaboration. Studies of Higgs boson production in the four-lepton final state at $\sqrt{s} = 13$ TeV. CMS Physics Analysis Summary CMS-PAS-HIG-15-004, 2016.

- [35] Robert D. Cousins. Why isn't every physicist a Bayesian? *Am. J. Phys.*, 63:398, 1995.
- [36] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71:1554, 2011.
- [37] A. De Rujula, Joseph Lykken, Maurizio Pierini, Christopher Rogan, and Maria Spiropulu. Higgs look-alikes at the LHC. *Phys.Rev.*, D82:013003, 2010.
- [38] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys.Rev.Lett.*, 13:321–322, 1964.
- [39] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 0711:070, 2007.
- [40] James S. Gainer, Kunal Kumar, Ian Low, and Roberto Vega-Morales. Improving the sensitivity of Higgs boson searches in the golden channel. *JHEP*, 1111:027, 2011.
- [41] James S. Gainer, Joseph Lykken, Konstantin T. Matchev, Stephen Mrenna, and Myeonghun Park. Geolocating the Higgs Boson Candidate at the LHC. *Phys.Rev.Lett.*, 111:041801, 2013.
- [42] Yanyan Gao, Andrei V. Gritsan, Zijin Guo, Kirill Melnikov, Markus Schulze,

et al. Spin determination of single-produced resonances at hadron colliders. *Phys.Rev.*, D81:075022, 2010.

[43] Rohini M. Godbole, 2 Miller, D.J., and M. Margarete Muhlleitner. Aspects of CP violation in the H ZZ coupling at the LHC. *JHEP*, 0712:031, 2007.

[44] Massimiliano Grazzini. NNLO predictions for the Higgs boson signal in the H → WW → lnu lnu and H → ZZ → 4l decay channels. *JHEP*, 02:043, 2008.

[45] Massimiliano Grazzini, Stefan Kallweit, and Dirk Rathlev. ZZ production at the LHC: Fiducial cross sections and distributions in NNLO QCD. *Phys. Lett. B*, 750:407 – 410, 2015.

[46] Massimiliano Grazzini and Hayk Sargsyan. Heavy-quark mass effects in Higgs boson production at the LHC. *JHEP*, 09:129, 2013.

[47] G.S. Guralnik, C.R. Hagen, and T.W.B. Kibble. Global Conservation Laws and Massless Particles. *Phys.Rev.Lett.*, 13:585–587, 1964.

[48] Kaoru Hagiwara, Qiang Li, and Kentarou Mawatari. Jet angular correlation in vector-boson fusion processes at hadron colliders. *JHEP*, 0907:101, 2009.

[49] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys.Rev.Lett.*, 13:508–509, 1964.

[50] Peter W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys.Lett.*, 12:132–133, 1964.

- [51] Peter W. Higgs. Spontaneous Symmetry Breakdown without Massless Bosons. *Phys.Rev.*, 145:1156–1163, 1966.
- [52] Wai-Yee Keung, Ian Low, and Jing Shu. Landau-Yang Theorem and Decays of a Z' Boson into Two Z Bosons. *Phys.Rev.Lett.*, 101:091802, 2008.
- [53] Vardan Khachatryan et al. Limits on the Higgs boson lifetime and width from its decay to four charged leptons. *Phys. Rev. D*, 92:072010, 2015.
- [54] Vardan Khachatryan et al. Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8\text{TeV}$. *JINST*, 10:P06005, 2015.
- [55] Vardan Khachatryan et al. Search for a Higgs Boson in the Mass Range from 145 to 1000 GeV Decaying to a Pair of W or Z Bosons. *JHEP*, 10:144, 2015.
- [56] T.W.B. Kibble. Symmetry breaking in nonAbelian gauge theories. *Phys.Rev.*, 155:1554–1561, 1967.
- [57] L. Landau. On the energy loss of fast particles by ionization. *J. Phys. (USSR)*, 8:201, 1944.
- [58] LHC Higgs Cross Section Working Group, S. Heinemeyer, C. Mariotti, G. Passarino, and R. Tanaka (Eds.). Handbook of LHC Higgs Cross Sections: 3. Higgs Properties. *CERN-2013-004*, CERN, Geneva, 2013.

- [59] Chong Sheng Li, Hai Tao Li, Ding Yu Shao, and Jian Wang. Soft gluon resummation in the signal-background interference process of $gg(\rightarrow h^*) \rightarrow ZZ$. 2015.
- [60] Gionata Luisoni, Paolo Nason, Carlo Oleari, and Francesco Tramontano. $Hw^\pm/\text{hz} + 0$ and 1 jet at nlo with the powheg box interfaced to gosam and their merging within minlo. *JHEP*, 10:1, 2013.
- [61] Kirill Melnikov and Matthew Dowling. Production of two Z-bosons in gluon fusion in the heavy top quark approximation. *Phys. Lett. B*, 744:43, 2015.
- [62] Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.
- [63] Paolo Nason and Carlo Oleari. NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG. *JHEP*, 1002:037, 2010.
- [64] Giampiero Passarino. Higgs CAT. *Eur. Phys. J. C*, 74:2866, 2014.
- [65] Torbjörn Sjöstrand et al. An introduction to pythia 8.2. *Computer Physics Communications*, 191:159, 2015.
- [66] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 physics and manual. *JHEP*, 05:026, 2006.
- [67] A. Soni and R.M. Xu. Probing CP violation via Higgs decays to four leptons. *Phys. Rev.*, D48:5259–5263, 1993.