

**DISCOVERY OF THE HIGGS BOSON, MEASUREMENTS OF HIGGS BOSON
PROPERTIES, AND SEARCH FOR HIGH MASS BEYOND THE STANDARD MODEL
SCALAR PARTICLE IN THE DIPHOTON FINAL STATE WITH THE ATLAS
DETECTOR AT THE LARGE HADRON COLLIDER**

by

Hongtao Yang

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Physics)

at the

UNIVERSITY OF WISCONSIN–MADISON

2016

Date of final oral examination: 09/22/2016

The dissertation is approved by the following members of the Final Oral Committee:

Lisa Everett, Professor, Physics

Gary Shiu, Professor, Physics

Wesley Smith, Professor, Physics

Sau Lan Wu, Professor, Physics

Michael Winokur, Professor, Physics



To my family.

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my advisor Prof. Sau Lan Wu. Prof. Wu is a great mentor who truly cares about students' education and welfare. She taught me how to become a good physicist and helped me countless times. Her dedication, perseverance and vision have inspired me in both research and life over the past six years, and will continue inspiring me in the future.

The days I was based at University of Wisconsin-Madison are among the happiest in my life. I was very fortunate to take courses from the outstanding Wisconsin professors, including Prof. Baha Balantekin, Prof. Ludwig Bruch, Prof. Daniel Chung, Prof. Lisa Everett, Prof. Karsten Heeger and many others. They have equipped me with the knowledge needed for future research, for which I really owe great thanks to them. I would like to also thank Prof. Lisa Everett, Prof. Gary Shiu, Prof. Wesley Smith and Prof. Michael Winokur for kindly reading this thesis and providing very helpful feedback.

The Wisconsin ATLAS group is like a big family. I would like to express my sincere appreciations to my colleagues, in particular to those who have shared with me selflessly their knowledge and experience in research. I thank German Carrillo-Montoya for teaching me physics analysis basics and supervising me on the $H \rightarrow ZZ \rightarrow \ell\ell\nu\nu$ and $H \rightarrow ZZ \rightarrow \ell\ell qq$ analyses. I also thank Swagato Banerjee for helping me on the E_T^{miss} trigger project. Since 2012 Haichen Wang had been training me on $H \rightarrow \gamma\gamma$ analyses and guiding me in many other aspects until and even after he graduated, for which I am really grateful. I am also indebted to Haoshuang Ji for the valuable training on statistical combination. At different stages of my PhD program, I have worked closely with Andrew Hard, Xiangyang Ju, Laser Kaplan, Lashkar Kashif, Manuel Silva, Fuquan Wang, Fangzhou Zhang and Chen Zhou on various topics. I value the pleasant time working with them, and I thank them for all the support and understanding they gave. In addition, I would like to thank Werner Wiedenmann for being such a nice office mate who shared knowledge and stories with me.

And I thank Neng Xu, Wen Guan and Shaojun Sun for their support on computing. My thanks also go to Luis Roberto Flores Castillo, Yaquan Fang, Haifeng Li, LianLiang Ma, Yao Ming, Haiping Peng, Ximo Poveda, Bill Quayle, Tapas Sarangi and Haimo Zobernig for their friendship and help.

As a member of the ATLAS Collaboration, I enjoy the privilege of working with outstanding physicists from all over the world. I thank Konstantinos Nikolopoulos for the joyful days working in the HSG2 (now HZZ) working group in 2011. I also thank Junichi Tanaka for being a great HSG1 (now HGam) convener during the Higgs boson discovery time. The follow-up HSG1 conveners, including Kerstin Tackmann, Krisztian Peters, Nicolas Berger, Sandrine Laplace, Dag Gillberg, Elisabeth Petit, Bruno Lenzi and Marco Delmastro have generously provided their guidance and support to me on various topics, which I really appreciate. I also thank many nice colleagues in the HSG1 working group for discussions that help me improve.

I am grateful to Paul Tipton and his Yale team, including Jahred Adelman, Johannes Erdmann, Andrey Longinov and Jared Vasquez for the very fruitful and pleasant collaboration on the search for $t\bar{t}H$ production process in the diphoton decay channel. On the same topic I am thankful for the strong support from HSG8 (now HTop) conveners Aurelio Juste and Peter Onyisi, and also from our editorial board chaired by Stathes Paganis.

I thank Eilam Gross for his guidance and support from Higgs boson search time all the way to the LHC Higgs combination. My thanks also go to my colleagues in the LHC Higgs Combination Group, in particular to Tim Adye, Andrea Gabrielli, Stefan Gadatsch, Rei Tanaka and Guillaume Unal from ATLAS, and to Mingshui Chen, André David, Giovanni Petrucciani and Marco Pieri from CMS. Moreover, I would like to thank the ATLAS editorial board chaired by Kevin Einsweiler for ensuring the quality of the publication with admirable amount of effort. In the HSG7 (now HComb) working group I thank Fabio Cerutti, Michael Duehrssen-Debling, Bruno Mansoulie, Kirill Prokofiev and Wouter Verkerke for their coordination and guidance, and I thank many outstanding colleagues I have worked with on this forum from different areas.

I thank Aurelio Juste for inviting me to the combined search for flavor-changing neutral current $t \rightarrow Hq$ decay. I thank Enrique Kajomovitz, Mike Hance, Alex Martyniuk, Bill Murray, Attilio Picazio, Reina Camacho Toro and other colleagues in the DBL working group for the collaboration on high mass diboson resonance search in all-hadronic channel and in combination with other diboson decay channels.

In the Run 2 high mass diphoton resonance search effort I give my special thanks to Tancredi Carli, Leonardo Carminati and Marco Delmastro for their organization and support, and to our editorial board chaired by Karl Jacobs and later also by Fabio Cerutti for their dedication. I also want to thank Liron Barak, Nicolas Berger, Quentin Buat, Marcello Fanti, Kirill Grevtsov, Giovanni Marchiori, Simone Mazza, Thomas Meideck, Lydia Roos, Jan Stark, Ruggero Turra, Guillaume Unal, Yee Chinn Yap and other colleagues in the analysis team who have helped me.

I would like to thank Marumi Kado for kindly following the analyses I have worked on as first Higgs Convener and later ATLAS Physics Coordinator, and providing very helpful inputs.

Besides colleagues working on physics analyses, I am deeply grateful to Maurice Garcia-Sciveres and also Ian Hinchliffe for arranging me to visit LBNL and work with Maurice on the exciting Pixel Detector Phase 2 upgrade project. During my stay at LBNL I sincerely appreciate the help from Rebecca Carney, Niklaus Lehmann, Manuel Silva, Simon Viel and other nice Berkeley colleagues. I also want to thank Allen Mincer and colleagues in the E_T^{miss} signature group for their patient instructions on my E_T^{miss} trigger project.

My research cannot go smoothly without the excellent administrative support, so I would like to thank Rita Knox, Aimee Lefkow, Renne Lefkow and Sylvie Padlewski for all their kind help. I should give additional thanks to Sylvie for taking care of me in France.

Finally, I would like to thank my parents, my grandparents and my girlfriend Nan Lu. Their love gives me the momentum to continue pursuing science while being able to appreciate all the other wonderful things in life. This thesis is dedicated to them.

TABLE OF CONTENTS

	Page
LIST OF TABLES	viii
LIST OF FIGURES	xii
ABSTRACT	xxii
1 Introduction	1
2 Phenomenology	5
2.1 Higgs boson production at Large Hadron Collider	5
2.2 Higgs boson decay branching ratios and total width	11
2.3 Background processes in diphoton final state at Large Hadron Collider	12
3 ATLAS detector	15
4 Data and simulation samples	18
4.1 Data samples	18
4.2 Simulation samples for Standard Model Higgs boson signals	18
4.3 Simulation samples for high mass scalar signals	21
4.4 Simulation samples for background processes	21
5 Physics object definitions	23
5.1 Photons	23
5.1.1 Photon reconstruction	23
5.1.2 Photon energy calibration	24
5.1.3 Photon identification	25
5.1.4 Photon isolation	26
5.1.5 Diphoton vertex selection	26
5.2 Other physics objects	28
5.2.1 Leptons	29
5.2.2 Jets	30
5.2.3 Missing transverse momentum	31

	Page
6 Diphoton event selections	32
6.1 Event selection used for discovery of Higgs boson	32
6.2 Event selection used for measurements of Higgs boson properties	33
6.3 Event selection used for search of high mass scalar particle	35
7 Signal and background modeling	37
7.1 Modeling of signal diphoton invariant mass shape and yield	37
7.1.1 Modeling of Standard Model Higgs boson decaying into two photons . . .	37
7.1.2 Modeling of high mass scalar particle decaying into two photons	38
7.2 Modeling of background diphoton invariant mass shape and normalization	43
8 Statistical procedure	45
8.1 Likelihood construction for diphoton analyses	45
8.2 Statistical tests	47
8.2.1 Test statistic for discovery of a positive signal	48
8.2.2 Test statistic for upper limits	48
8.2.3 Test statistic for measurements and compatibility tests	50
9 Discovery of Higgs boson in diphoton decay channel	52
9.1 Event categorization	52
9.2 Systematic uncertainties	57
9.3 Results	61
9.3.1 Diphoton invariant mass spectra	61
9.3.2 Statistical interpretations	67
9.3.3 Combination with other decay channels	71
10 Measurement of Higgs boson couplings in diphoton decay channel	73
10.1 Event categorization	73
10.2 Systematic uncertainties	86
10.2.1 Uncertainties on integrated signal yield	86
10.2.2 Uncertainties on signal events migration between categories	90
10.2.3 Uncertainties on signal mass resolution and mass scale	93
10.2.4 Uncertainties on background model	96
10.3 Results	99
10.3.1 Diphoton invariant mass spectra	99
10.3.2 Signal strength measurements	105
10.3.3 Search for $t\bar{t}H$ production process and constraints on Yukawa coupling between top quark and Higgs boson	112

	Page
10.3.4 Combination with other decay channels	116
11 Measurement of Higgs boson mass	119
11.1 Measurement in diphoton decay channel by ATLAS	119
11.2 ATLAS–CMS combined measurement	124
12 Search for high mass scalar resonance in diphoton decay channel	134
12.1 Systematic uncertainties	134
12.2 Results	136
12.2.1 Diphoton invariant mass spectra	136
12.2.2 Compatibility with the background-only hypothesis	136
12.2.3 Limits on fiducial cross section	139
13 Conclusion	142
APPENDIX Coupling modifiers	144
LIST OF REFERENCES	146

LIST OF TABLES

Table	Page
2.1 Standard Model predictions for the Higgs boson production cross sections together with their theoretical uncertainties at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV. The value of the Higgs boson mass is assumed to be $m_H = 125.09$ GeV. The uncertainties on the cross sections are evaluated as the sum in quadrature of the uncertainties resulting from variations of the QCD scales, parton distribution functions, and α_s . The order of the theoretical calculations is also indicated. In the case of the $b\bar{b}H$ production, the values are given for the mixture of five-flavor (5FS) and four-flavor (4FS) schemes.	10
2.2 Standard Model predictions for the decay branching fractions of a Higgs boson with a mass of 125.09 GeV, together with their theoretical uncertainties.	12
9.1 Number of events in the data (N_D) and expected number of signal events (N_S) with $m_H = 126.5$ GeV for each category of the <i>Discovery Analysis</i> and total for the 7 TeV and 8 TeV datasets in the mass range 100 – 160 GeV. The mass resolution quantified by full width at half maximum (FWHM) is also given for the 8 TeV data.	55
9.2 Summary of systematic uncertainties on the expected signal considered in the <i>Discovery Analysis</i> . The values listed in the table are the relative uncertainties (in %) on given quantities from the various sources investigated for a Higgs boson mass of 125 GeV. The sign in the front of values for each systematic uncertainty indicates correlations among categories and processes. Experimental and theoretical uncertainties are separately marked.	59
9.3 List of the functions chosen to model the background distributions of $m_{\gamma\gamma}$ in the <i>Discovery Analysis</i> , and the associated systematic uncertainties on the signal amplitudes in terms of spurious signal (N_{spur}) for the ten categories and the 7 TeV and 8 TeV datasets.	60
10.1 Signal efficiencies ϵ , which include geometrical and kinematic acceptances, and expected signal event fractions f per production mode in each category of the <i>Coupling Analysis</i> for $\sqrt{s} = 7$ TeV and $m_H = 125.4$ GeV. The second-to-last row shows the total efficiency per production process summed over the categories and the overall average efficiency in the far right column. The total number of selected signal events expected in each category N_S is reported in the last column while the total number of selected events expected from each production mode is given in the last row.	83

Table	Page
10.2 Signal efficiencies ϵ , which include geometrical and kinematic acceptances, and expected signal event fractions f per production mode in each category of the <i>Coupling Analysis</i> for $\sqrt{s} = 8$ TeV and $m_H = 125.4$ GeV. The second-to-last row shows the total efficiency per production process summed over the categories and the overall average efficiency in the far right column. The total number of selected signal events expected in each category N_s is reported in the last column while the total number of selected events expected from each production mode is given in the last row.	84
10.3 Number of selected events in each category of the <i>Coupling Analysis</i> and total for the 7 TeV and 8 TeV datasets in the mass range 105 – 160 GeV.	85
10.4 Theoretical uncertainties (in %) on cross sections for Higgs boson production processes at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV for $m_H = 125.4$ GeV in the <i>Coupling Analysis</i>	87
10.5 Relative systematic uncertainties on the inclusive yields (in %) for the 7 TeV and 8 TeV datasets. The numbers in parentheses refer to the uncertainties applied to $t\bar{t}H$ and VH categories. The ranges of the category-dependent uncertainties due to the isolation efficiency are reported.	89
10.6 Relative uncertainties (in %) on the Higgs boson signal yield in each category of the <i>Coupling Analysis</i> and for each production process induced by the combined effects of the systematic uncertainties on the jet energy scale, jet energy resolution and jet vertex fraction. These uncertainties are approximately the same for the 7 TeV and the 8 TeV data.	93
10.7 Relative uncertainties (in %) on the Higgs boson signal yield in each category of the <i>Coupling Analysis</i> and for each production process induced by systematic uncertainty on the E_T^{miss} energy scale and resolution. The uncertainties, which are approximately the same for the 7 TeV and 8 TeV data, are obtained by summing in quadrature the impacts on the signal yield of the variation of each component of the E_T^{miss} energy scale within its uncertainty.	94
10.8 Systematic uncertainties on the diphoton mass resolution for the 8 TeV data (in %) due to the four contributions described in the text. For each category of the <i>Coupling Analysis</i> , the uncertainty is estimated by using a simulation of the Higgs boson production process which makes the largest contribution to the signal yield.	95
10.9 List of the functions chosen to model the background distributions of $m_{\gamma\gamma}$ in the <i>Coupling Analysis</i> , and the associated systematic uncertainties on the signal amplitudes in terms of spurious signal (N_{spur}) and its ratio to the predicted number of signal events in each category (μ_{spur}) for the twelve categories and the 7 TeV and 8 TeV datasets.	99

Table	Page
10.10 Main systematic uncertainties $\sigma_{\mu}^{\text{syst.}}$ on the combined signal strength parameter μ measured from the <i>Coupling Analysis</i> . The values for each group of uncertainties are determined by subtracting in quadrature from the total uncertainty the change in the 68% confidence level range on μ when the corresponding nuisance parameters are fixed to their best fit values. The experimental uncertainty on the yield does not include the luminosity contribution, which is accounted for separately.	107
10.11 Observed and expected 95% confidence level upper limits on the $t\bar{t}H$ production cross section times $\text{BR}(H \rightarrow \gamma\gamma)$ relative to the Standard Model prediction at $m_H = 125.4$ GeV. All other Higgs boson production cross sections, including the cross section for tH production, are set to their respective SM expectations. In addition, the expected limits corresponding to $+2\sigma$, $+1\sigma$, -1σ , and -2σ variations are shown. The expected limits are calculated for the case where $t\bar{t}H$ production is not present. The results are given for the combination of leptonic and hadronic categories with all systematic uncertainties included, and also for leptonic and hadronic categories separately. Expected limits are also derived for the case of statistical uncertainties only.	112
11.1 Summary of the expected number of signal events in the 105 – 160 GeV mass range (n_{sig}), the full width at half maximum (FWHM) of mass resolution, half of the smallest range containing 68% of the signal events (σ_{eff}), number of background events b in the smallest mass window containing 90% of the signal ($\sigma_{\text{eff}90}$), and the ratio s/b and $s/\sqrt{b}$ with s being the expected number of signal events in the window containing 90% of signal events for the <i>Mass Analysis</i> . b is derived from the fit of the data in the 105 – 160 GeV mass range. The value of m_H is taken to be 126 GeV and the signal yield is assumed to be the expected Standard Model value. The estimates are shown separately for the 7 TeV and 8 TeV datasets and for the inclusive sample as well as for each of the categories.	122
11.2 Summary of the relative systematic uncertainties (in %) on the $H \rightarrow \gamma\gamma$ mass measurement for the different categories described in the text. The first seven rows give the impact of the photon energy scale systematic uncertainties grouped into seven classes.	123

Table	Page
11.3 Systematic uncertainties δm_H (see text) associated with the indicated effects for each of the four input channels, and the corresponding contributions of ATLAS and CMS to the systematic uncertainties of the combined result. “ECAL” refers to the electromagnetic calorimeters. The numbers in parentheses indicate expected values obtained from the prefit Asimov data set discussed in the text. The uncertainties for the combined result are related to the values of the individual channels through the relative weight of the individual channel in the combination, which is proportional to the inverse of the respective uncertainty squared. The top section of the table divides the sources of systematic uncertainty into three classes, which are discussed in the text. The bottom section of the table shows the total systematic uncertainties estimated by adding the individual contributions in quadrature, the total systematic uncertainties evaluated using the nominal method discussed in the text, the statistical uncertainties, the total uncertainties, and the analysis weights, illustrative of the relative weight of each channel in the combined m_H measurement.	133
12.1 Summary of systematic uncertainties on the signal and background considered in the <i>Search Analysis</i>	135

LIST OF FIGURES

Figure	Page
1.1 Summary of elementary particles in the Standard Model.	2
2.1 Standard Model Higgs boson production cross sections (a) as a function of Higgs boson mass at $\sqrt{s} = 8$ TeV and (b) as a function of center-of-mass energy with hypothesized Higgs boson mass of 125 GeV. The central values are shown as solid lines, and the theoretical uncertainties are shown as color bands. The $b\bar{b}H$ and tH processes are not included in the left plot.	7
2.2 Examples of leading-order Feynman diagrams for Higgs boson production via the (a) ggF and (b) VBF production processes.	7
2.3 Examples of leading-order Feynman diagrams for Higgs boson production via the (a) VH and (b, c) $ggZH$ production processes.	8
2.4 Examples of leading-order Feynman diagrams for Higgs boson production via the $t\bar{t}H$ and $b\bar{b}H$ processes.	8
2.5 Examples of leading-order Feynman diagrams for Higgs boson production in association with a single top quark via the (a, b) $tHqb$ and (c, d) tHW production processes shown in four-flavor and five-flavor schemes, respectively.	9
2.6 Examples of leading-order Feynman diagrams for Higgs boson decays to diphoton. . .	11
2.7 Examples of leading-order Feynman diagrams for Higgs boson decays (a) to W and Z bosons (one of the two bosons is off-shell for m_H near 125 GeV) and (b) to fermions. . .	12
2.8 Examples of leading-order Feynman diagrams for non-resonant Standard Model diphoton production process.	13
2.9 Examples of leading-order Feynman diagrams for non-resonant Standard Model photon–jet production process.	14
3.1 Sketch of a barrel module (located at $\eta = 0$) of the ATLAS electromagnetic calorimeter. The different longitudinal layers (one presampler and three layers in the accordion calorimeter) are depicted. The granularity in η and ϕ of the cells of each layer and of the trigger towers is also shown.	17

Figure	Page	
5.1	Efficiency to select a diphoton vertex within 0.3 mm of the production vertex (ϵ_{PV}) as a function of the number of primary vertices in the event in the <i>Measurement Analyses</i> . The plot shows ϵ_{PV} for simulated ggF events ($m_H = 125$ GeV) with two unconverted photons (empty blue squares), for $Z \rightarrow e^+e^-$ events with the electron tracks removed for the neural-network-based identification of the vertex, both in data (black triangles) and simulation (red triangles), and the same simulated $Z \rightarrow e^+e^-$ events re-weighted to reproduce the p_T spectrum of simulated ggF events (red circles).	28
6.1	Diphoton sample composition as a function of the invariant mass for the 7 TeV (a) and the 8 TeV (b) dataset used in the <i>Discovery Analysis</i> . The small contribution from Drell-Yan events is included in the diphoton component. The error bars on each point represent the statistical uncertainty on the measurement while the colored bands represent the total uncertainty.	33
6.2	Efficiency to fulfill the isolation requirement (ϵ_{iso}) as a function of the number of primary vertices in each event in the <i>Measurement Analyses</i> , determined with a simulation sample of Higgs bosons decaying into two photons with $m_H = 125$ GeV and $\sqrt{s} = 8$ TeV. Events are required to satisfy the kinematic selection described in the text. The efficiency of the event selection obtained with a tight calorimetric isolation requirement (4 GeV) is compared with the case in which a looser calorimetric isolation (6 GeV) is combined with a track isolation (2.6 GeV) selection.	34
6.3	Diphoton sample composition as a function of the invariant mass for the 7 TeV (a) and the 8 TeV (b) dataset used in the <i>Measurement Analyses</i> . The small contribution from Drell-Yan events is included in the diphoton component. The error bars on each point represent the statistical uncertainty on the measurement while the colored bands represent the total uncertainty.	35
6.4	Diphoton sample composition as a function of the invariant mass for the 13 TeV dataset used in the <i>Search Analysis</i> . The bottom panel shows the purity of diphoton events as determined from two independent methods (matrix and $2 \times 2D$ sidebands) with good agreement achieved. The total uncertainties including statistical and systematic components are shown by error bars.	36
7.1	Simulated diphoton invariant mass distribution for Standard Model Higgs boson signal with $m_H = 125$ GeV in one category of the <i>Coupling Analysis</i> superimposed with parameterization determined from the procedure described in the text.	39

Figure	Page	
7.2	The true $m_{\gamma\gamma}$ distributions of the resonance with $m_X = 750$ GeV and width of 6% of the m_X value in the <i>Search Analysis</i> . The dashed red line is the gluon-gluon luminosity, the dashed green line is the functional form $m_{\gamma\gamma}^7$ (arising from the numerator of the squared matrix element, multiplied by the Jacobian factor of the variable transformation $\hat{s} \rightarrow m_{\gamma\gamma}$), and the dashed blue line is the Breit–Wigner distribution with the same mass and width as generated in the sample. The product of these three is represented by the solid purple line and agrees well with the true invariant mass distribution, shown as the black histogram. No selection cuts have been applied to the true photons.	41
7.3	The $m_{\gamma\gamma}$ distributions for a scalar resonance in the <i>Search Analysis</i> with a mass of 800 GeV with (a) a narrow decay width ($\Gamma_X = 4$ MeV) or with (b) $\Gamma_X/m_X = 6\%$. The parameterization as the convolution of the theoretical mass line shape with the detector resolution is superimposed.	42
9.1	Distribution of p_{Tt} in simulated events with Higgs boson productions and in background events. The signal distribution is shown separately for ggF (blue), and VBF together with associated productions (red). The background distribution and the two signal distributions are normalized to unit area.	54
9.2	Invariant mass distributions for a Higgs boson with $m_H = 125$ GeV, for the best-resolution category (Unconverted central high p_{Tt}) of the <i>Discovery Analysis</i> shown in blue, and for a category with lower resolution (Converted rest low p_{Tt}) shown in red, for the $\sqrt{s} = 8$ TeV simulation. The invariant mass distribution is parametrized by the sum of a Crystal Ball function and a broad Gaussian based on the procedure discussed in Chapter 7.1.1.	56
9.3	Distribution of the invariant mass of diphoton candidates after all selections in the <i>Discovery Analysis</i> for the combined 7 TeV and 8 TeV data sample. The result of a fit to the data of the sum of a signal component fixed to $m_H = 126.5$ GeV and a background component described by a fourth-order Bernstein polynomial is superimposed. The bottom panel shows the data relative to the background component of the fitted model.	61
9.4	Background-only fits to the diphoton invariant mass spectra for (a) Unconverted central low p_{Tt} , (b) Unconverted central high p_{Tt} , (c) Unconverted rest low p_{Tt} , and (d) Unconverted rest high p_{Tt} categories of the <i>Discovery Analysis</i> correspond to the 7 TeV data sample. The bottom panel displays the residual of the data with respect to the background fit. The Higgs boson expectation for a mass hypothesis of 126.5 GeV corresponding to the Standard Model cross section is also shown.	62

Figure	Page
9.5 Background-only fits to the diphoton invariant mass spectra for (a) Unconverted central low p_{Tt} , (b) Unconverted central high p_{Tt} , (c) Unconverted rest low p_{Tt} , and (d) Unconverted rest high p_{Tt} categories of the <i>Discovery Analysis</i> correspond to the 8 TeV data sample. The bottom panel displays the residual of the data with respect to the background fit. The Higgs boson expectation for a mass hypothesis of 126.5 GeV corresponding to the Standard Model cross section is also shown.	63
9.6 Background-only fits to the diphoton invariant mass spectra for (a) Converted central low p_{Tt} , (b) Converted central high p_{Tt} , (c) Converted rest low p_{Tt} , and (d) Converted rest high p_{Tt} categories of the <i>Discovery Analysis</i> correspond to the 7 TeV data sample. The bottom panel displays the residual of the data with respect to the background fit. The Higgs boson expectation for a mass hypothesis of 126.5 GeV corresponding to the Standard Model cross section is also shown.	64
9.7 Background-only fits to the diphoton invariant mass spectra for (a) Converted central low p_{Tt} , (b) Converted central high p_{Tt} , (c) Converted rest low p_{Tt} , and (d) Converted rest high p_{Tt} categories of the <i>Discovery Analysis</i> correspond to the 8 TeV data sample. The bottom panel displays the residual of the data with respect to the background fit. The Higgs boson expectation for a mass hypothesis of 126.5 GeV corresponding to the Standard Model cross section is also shown.	65
9.8 Background-only fits to the diphoton invariant mass spectra for the Two-jet category of the <i>Discovery Analysis</i> correspond to the 7 TeV data (a) and the 8 TeV data (b), and Converted transition categories correspond to the 7 TeV data (c) and the 8 TeV data (d). The bottom panel displays the residual of the data with respect to the background fit. The Higgs boson expectation for a mass hypothesis of 126.5 GeV corresponding to the Standard Model cross section is also shown.	66
9.9 Expected and observed local p_0 values for a Standard Model Higgs boson as a function of the hypothesized Higgs boson mass m_H for the combined analysis and for the $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV data samples separately. The observed p_0 including the effect of the photon energy scale uncertainty on the mass position is included via pseudo-experiments and shown as open circles.	68
9.10 Best fit value for the signal strength as a function of the assumed Higgs boson mass m_H from the <i>Discovery Analysis</i>	69
9.11 Best fit value for the signal strength in the different categories of the <i>Discovery Analysis</i> at $m_H = 126.5$ GeV for the combined $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV data samples.	69
9.12 The two-dimensional best-fit value of $(\mu_{ggF+t\bar{t}H}, \mu_{VBF+VH})$ from the <i>Discovery Analysis</i> . The 68% and 95% confidence level contours are shown with the solid and dashed lines, respectively.	70

Figure	Page
9.13 Confidence intervals in the (μ, m_H) plane for the $H \rightarrow \gamma\gamma$ (<i>Discovery Analysis</i>), $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$, and $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$ channels, including all systematic uncertainties. The markers indicate the maximum likelihood estimates in the corresponding channels	71
9.14 The observed (solid) local p_0 as a function of m_H . The dashed curve shows the expected local p_0 under the hypothesis of a Standard Model Higgs boson signal at that mass with its $\pm 1\sigma$ band. The horizontal dashed lines indicate the p -values corresponding to significances of 1 to 6σ	72
10.1 Illustration of the order in which the criteria for the exclusive event categories in the <i>Coupling Analysis</i> are applied to the selected diphoton events. The division of the last category, which is dominated by ggF production, into four sub-categories is described in the text.	74
10.2 Normalized distributions of the variables described in the text used to sort diphoton events with at least two reconstructed jets into the VH hadronic category of the <i>Coupling Analysis</i> for the data in the sidebands (points), the predicted sum of the WH and ZH signals (red histograms), the predicted signal feed-through from ggF , VBF, and $t\bar{t}H$ production modes (blue histograms), and the simulation of the $\gamma\gamma$, γj , and $j\bar{j}$ background processes (green histograms). The arrows indicate the selection criteria applied to these observables. The mass of the Higgs boson in all signal samples is $m_H = 125$ GeV.	77
10.3 Normalized kinematic distributions of the six variables describe in the text used to build the Boost Decision Tree that assigns events to the VBF categories of the <i>Coupling Analysis</i> , for diphoton candidates with two well-separated jets ($\Delta\eta_{jj} \geq 2.0$ and $ \eta^* < 5.0$). Distributions are shown for data sidebands (points) and simulation of the VBF signal (blue histograms), feed-through from ggF production (red histograms), and the continuum QCD background predicted by MC simulation and data control regions (green histograms) as described in the text. The signal VBF and ggF samples are generated with a Higgs boson mass $m_H = 125$ GeV.	79
10.4 Probability distributions of the output of the Boost Decision Tree (BDT) O_{BDT} for the VBF signal (blue), ggF feed-through (red), continuum QCD background predicted by MC samples and data control regions (green) as described in the text, and data sidebands (points). The two vertical dashed lines indicate the cuts on O_{BDT} that define the VBF loose and tight categories in the <i>Coupling Analysis</i> . The signal VBF and ggF samples are generated with a Higgs boson mass $m_H = 125$ GeV.	80

Figure	Page
10.5 Distributions of p_{Tt} for diphoton candidates in the sidebands in the untagged (a) Central and (b) Forward categories for $\sqrt{s} = 8$ TeV for predicted Higgs boson production processes (solid histograms), the predicted sum of $\gamma\gamma$, γj and $j j$ background processes (green histogram), and data (points). The vertical dashed lines indicate the value used to classify events into the low or high p_{Tt} categories in the <i>Coupling Analysis</i> . The mass for all Higgs boson signal samples is $m_H = 125$ GeV.	81
10.6 (a) The distributions of diphoton invariant mass $m_{\gamma\gamma}$ in the untagged Central low p_{Tt} category in data (points), and simulation samples for the $\gamma\gamma$, γj and $j j$ components of the continuum background (shaded cumulative histograms). The lower plot shows the ratio of data to simulation. (b) Ratio of the fitted number of signal events to the number expected for the Standard Model $\mu_{sp}(m_H)$ as a function of the test mass m_H for the untagged Central low p_{Tt} category. A single fit per value of m_H is performed on the representative pure MC background sample described in the text with signal plus a variety of background parameterizations (exp1, exp2, exp3 for the exponentials of first, second or third-order polynomials, respectively, and bern3, bern4, bern5 for third, fourth and fifth-order Bernstein polynomials, respectively). The bias criteria discussed in Section 7.2 are indicated by the dashed lines.	97
10.7 Comparison of the $m_{\gamma\gamma}$ distributions in data in the signal and control regions for $t\bar{t}H$ hadronic and $t\bar{t}H$ leptonic categories. The background shapes are extracted from data in control regions obtained by removing b-tagging requirement, relaxing the number of jets requirement and replacing leptons with jets (in $t\bar{t}H$ leptonic category only).	98
10.8 Distribution of the invariant mass of diphoton candidates after all selections in the <i>Coupling Analysis</i> for the combined 7 TeV and 8 TeV data sample. The solid red curve shows the fitted signal plus background model where the Higgs boson mass is fixed at 125.4 GeV. The background component of the fit is shown with the dotted blue curve. The signal component of the fit is shown with the solid black curve. Both the signal plus background and background-only curves reported here are obtained from the sum of the individual curves in 7 TeV and 8 TeV categories. The bottom panel shows the data relative to the background component of the fitted model.	100
10.9 Diphoton invariant mass spectra observed in the 7 TeV and 8 TeV data in the untagged categories of the <i>Coupling Analysis</i> : Central low p_{Tt} (a), Central high p_{Tt} (b), Forward low p_{Tt} (c), and Central high p_{Tt} (d).	101
10.10 Diphoton invariant mass spectra observed in the 7 TeV and 8 TeV data in the VBF categories of the <i>Coupling Analysis</i> : VBF loose (a) and VBF tight (b).	102
10.11 Diphoton invariant mass spectra observed in the 7 TeV and 8 TeV data in the VH categories of the <i>Coupling Analysis</i> : VH hadronic (a), $VH E_T^{\text{miss}}$ (b), VH one-lepton (c) and VH dilepton (d).	103

Figure	Page
10.12 Diphoton invariant mass spectra observed in the 7 TeV and 8 TeV data in the $t\bar{t}H$ categories of the <i>Coupling Analysis</i> : $t\bar{t}H$ leptonic (a) and $t\bar{t}H$ hadronic.	104
10.13 The profile of the negative log-likelihood ratio $\lambda(\mu)$ of the combined signal strength μ for $m_H = 125.4$ GeV. The observed result is shown by the solid curve, the expectation from the Standard Model by the dashed curve. The intersections of the solid and dashed curves with the horizontal dashed line at $\lambda(\mu) = 1$ indicate the 68% confidence intervals of the observed and expected results, respectively.	105
10.14 The signal strength for a Higgs boson of mass $m_H = 125.4$ GeV decaying via $H \rightarrow \gamma\gamma$ as measured (a) in the individual categories of the <i>Coupling Analysis</i> , and (b) in groups of categories sensitive to individual production modes for the combination of the 7 TeV and 8 TeV data together with the combined signal strength. The vertical hatched band indicates the 68% confidence interval of the combined signal strength. The vertical dashed line at unity indicates the Standard Model expectation. The vertical dashed red line indicates the limit below which the fitted signal plus background mass distribution for the $t\bar{t}H$ hadronic category becomes negative for some mass in the fit range. The VH dilepton category is not shown because with only two events in the combined sample, the fit results are not meaningful.	106
10.15 Measured signal strengths, for a Higgs boson of mass $m_H = 125.4$ GeV decaying via $H \rightarrow \gamma\gamma$, of the different Higgs boson production modes and the combined signal strength μ obtained with the combination of the 7 TeV and 8 TeV data in the <i>Coupling Analysis</i> . The vertical dashed line at unity indicates the Standard Model expectation. The vertical dashed line at the left end of the μ_{ZH} result indicates the limit below which the fitted signal plus background mass distribution becomes negative for some mass in the fit range.	109
10.16 The two-dimensional best-fit value of (μ_{VBF}, μ_{ggF}) for a Higgs boson of mass $m_H = 125.4$ GeV decaying via $H \rightarrow \gamma\gamma$ when fixing both μ_{tH} and $\mu_{b\bar{b}H}$ to 1 and profiling all the other signal strength parameters in the <i>Coupling Analysis</i> . The 68% and 95% confidence level contours are shown with the solid and dashed lines, respectively. The result is obtained for $m_H = 125.4$ GeV and the combination of the 7 TeV and 8 TeV data. . . .	110
10.17 Measurements of the μ_{VBF}/μ_{ggF} , μ_{VH}/μ_{ggF} and $\mu_{t\bar{t}H}/\mu_{ggF}$ ratios and their total errors for a Higgs boson mass $m_H = 125.4$ GeV in the <i>Coupling Analysis</i> . For a more complete illustration, the log-likelihood curves from which the total uncertainties are extracted are also shown: the best fit values are represented by the solid vertical lines, with the total $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties indicated by the dark- and light-shaded band, respectively. The likelihood curve and uncertainty bands for μ_{VH}/μ_{ggF} stop at zero because below this the hypothesized signal plus background mass distribution in the VH dilepton channel becomes negative (unphysical) for some mass in the fit range.	111

- 10.18 Observed and expected 95% confidence level upper limits on the $t\bar{t}H$ production cross section times $\text{BR}(H \rightarrow \gamma\gamma)$. All other Higgs boson production cross sections, including the cross section for tH production, are set to their respective Standard Model expectations. While the expected limits are calculated for the case where $t\bar{t}H$ production is not present, the lines denoted by “SM signal injected” show the expected 95% confidence level limits for a dataset corresponding to continuum background plus SM Higgs boson production. The limits are given relative to the SM expectations and at $m_H = 125.4$ GeV. 113
- 10.19 Production cross sections for $t\bar{t}H$ and tH divided by their Standard Model expectations as a function of the scale factor to the top quark-Higgs boson Yukawa coupling, κ_t . Production of tH comprises the $tHqb$ and tHW processes. Also shown is the dependence of the $\text{BR}(H \rightarrow \gamma\gamma)$ with respect to its SM expectation on κ_t 114
- 10.20 Observed and expected 95% confidence level upper limits on the inclusive Higgs production cross section with respect to the Standard Model cross section times $\text{BR}(H \rightarrow \gamma\gamma)$ for different values of κ_t at $m_H = 125.4$ GeV, where κ_t is the strength parameter for the top quark-Higgs boson Yukawa coupling. All Higgs boson production processes are considered for the inclusive production cross section. The expected limits are calculated for the case where $\kappa_t = +1$. The CL_s alternative hypothesis is given by continuum background plus Standard Model Higgs boson production. 115
- 10.21 Negative log-likelihood scan of κ_t at $m_H = 125.4$ GeV, where κ_t is the strength parameter for the top quark-Higgs boson Yukawa coupling. 116
- 10.22 Negative log-likelihood contours at 68% and 95% confidence level in the $(\kappa_\gamma, \kappa_g)$ plane for the combination of ATLAS and CMS and for each experiment separately, as obtained from the fit to the parameterization constraining all the other coupling modifiers to their Standard Model values and assuming $\text{BR}_{\text{BSM}} = 0$ 117
- 10.23 (a) Negative log-likelihood contours at 68% and 95% confidence level in the (κ_F^f, κ_V^f) plane for the combination of ATLAS and CMS and for the individual decay channels, as well as for their combination (κ_F versus κ_V shown in black), without any assumption about the relative sign of the coupling modifiers. (b) Observed (solid line) and expected (dashed line) negative log-likelihood scans for the global κ_F parameter, corresponding to the combination of all decay channels. The red (green) horizontal lines at the $-2\Delta \ln \Lambda$ value of 1 (4) indicate the value of the profile likelihood ratio corresponding to a 1σ (2σ) confidence level interval for the parameter of interest. . . 118

Figure	Page
11.1 Distribution of the four-lepton invariant mass for the selected candidates in the $m_{4\ell}$ range from 80 GeV to 170 GeV for the combined 7 TeV and 8 TeV data samples. Superimposed are the expected distributions of a SM Higgs boson signal for $m_H = 124.5$ GeV normalized to the measured signal strength, as well as the expected ZZ^* and reducible backgrounds.	121
11.2 Invariant mass spectra from CMS $H \rightarrow \gamma\gamma$ (a) and $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ (b) analyses based on Run 1 dataset.	124
11.3 Scans of twice the negative log-likelihood ratio $-2 \ln \Lambda(m_H)$ as functions of the Higgs boson mass m_H for the ATLAS and CMS combination of the $H \rightarrow \gamma\gamma$ (red), $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ (blue), and combined (black) channels. The dashed curves show the results accounting for statistical uncertainties only, with all nuisance parameters associated with systematic uncertainties fixed to their best-fit values. The 1 and 2 standard deviation limits are indicated by the intersections of the horizontal lines at 1 and 4, respectively, with the log-likelihood scan curves.	126
11.4 Summary of Higgs boson mass measurements from the individual analyses of ATLAS and CMS and from the combined analysis presented here. The systematic (narrower, magenta-shaded bands), statistical (wider, yellow-shaded bands), and total (black error bars) uncertainties are indicated. The (red) vertical line and corresponding (gray) shaded column indicate the central value and the total uncertainty of the combined measurement, respectively.	127
11.5 Summary of likelihood scans in the two-dimensional plane of signal strength μ versus Higgs boson mass m_H for the ATLAS and CMS experiments. The 68% confidence level contours of the individual measurements are shown by the dashed curves and of the overall combination by the solid curve. The markers indicate the respective best-fit values. The Standard Model signal strength is indicated by the horizontal line at $\mu = 1$	128
11.6 The impacts δm_H (see text) of the nuisance parameter groups in Table 11.3 on the ATLAS (left), CMS (center), and combined (right) mass measurement uncertainty. The observed (expected) results are shown by the solid (empty) bars.	132
12.1 Invariant mass distribution of the selected diphoton candidates in the <i>Search Analysis</i> , with the background-only fit overlaid, for (a) 2015 data and (b) 2016 data. The difference between the data and this fit is shown in the bottom panel. The arrow shown in the lower panel indicates a values outside the range with more than one standard deviation. There is no data event with $m_{\gamma\gamma} > 2500$ GeV.	137

Figure	Page
12.2 Distribution of the diphoton invariant mass of the selected events in the <i>Search Analysis</i> , with the background-only fit. The difference between the data and this fit is shown in the bottom panel. The arrow shown in the lower panel indicates a values outside the range with more than one standard deviation. There is no data event with $m_{\gamma\gamma} > 2500$ GeV.	138
12.3 Compatibility, in terms of local significance σ , with the background-only hypothesis as a function of the assumed signal mass and width for a scalar resonance.	139
12.4 Observed local p_0 values as a function of the assumed scalar resonance mass m_X , for different values of decay widths Γ_X : (a) narrow-width ($\Gamma_X = 4$ MeV), (b) $\Gamma_X/m_X = 2\%$, (c) $\Gamma_X/m_X = 6\%$, and (d) $\Gamma_X/m_X = 10\%$	140
12.5 Upper limits on the fiducial cross section times branching ratio to two photons of a scalar particle produced at $\sqrt{s} = 13$ TeV as a function of its mass m_X , for different values of decay widths Γ_X : (a) narrow-width ($\Gamma_X = 4$ MeV), (b) $\Gamma_X/m_X = 2\%$, (c) $\Gamma_X/m_X = 6\%$, and (d) $\Gamma_X/m_X = 10\%$	141

ABSTRACT

With 4.8 fb^{-1} of proton-proton collision data collected at $\sqrt{s} = 7 \text{ TeV}$ in 2011, and 5.9 fb^{-1} collected at $\sqrt{s} = 8 \text{ TeV}$ in 2012 by the ATLAS detector at the Large Hadron Collider, an excess of 4.5 standard deviations from the background-only hypothesis is observed near 126.5 GeV in the diphoton invariant mass spectra. Along with the excesses observed in the $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ and $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$ channels, the observation of a Higgs-like particle is established at 6.0 standard deviations level.

With more data accumulated during LHC Run 1, the measurements of Higgs boson couplings and mass in the $H \rightarrow \gamma\gamma$ channel are conducted by the ATLAS experiment based on 4.5 fb^{-1} of proton-proton collisions at $\sqrt{s} = 7 \text{ TeV}$ collected in 2011, and 20.3 fb^{-1} at $\sqrt{s} = 8 \text{ TeV}$ collected in 2012. The combined signal strength, defined as number of observed Higgs boson decays to diphoton divided by the corresponding Standard Model prediction, is measured to be $1.17^{+0.28}_{-0.26}$ assuming the Higgs boson mass being 125.4 GeV. The signal strengths for individual Higgs boson production processes are also measured, and are found to be in good consistency with the Standard Model. The mass of the Higgs boson is measured in $H \rightarrow \gamma\gamma$ channel by the ATLAS experiment to be $125.98 \pm 0.50 \text{ GeV}$. This measurement is combined with the ones from ATLAS $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ as well as CMS $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$. The Higgs boson mass measured from the combination is $125.09 \pm 0.24 \text{ GeV}$.

With LHC center-of-mass energy increased to 13 TeV, a search for high mass Beyond the Standard Model scalar resonance is performed in the diphoton decay channel based on 15.4 fb^{-1} of proton-proton collision data collected by the ATLAS detector during 2015 and 2016. While a notable wide excess was first observed in the diphoton invariant mass spectrum from the 2015 data (3.2 fb^{-1}) with mass near 750 GeV, it is not confirmed by the 2016 data with much higher statistics (12.4 fb^{-1}). Limits on the production cross section times branching ratio of such resonances are set.

Chapter 1

Introduction

The Standard Model (SM) of particle physics is the theory that describes elementary particles and the interactions (not including gravity for the moment) between them. It has successfully explained most of the experimental observations so far. According to the SM, all known matter is built up with spin $\frac{1}{2}$ elementary particles including quarks (u, d, c, s, t, b) and leptons (e, μ, τ and corresponding neutrinos). The interactions, including strong (mediated by gluon g) and electroweak (mediated by photon γ and weak bosons $W^\pm$ and Z), are carried by spin 1 particles. All the SM elementary particles are summarized in Figure 1.1

The SM is a gauge field theory based on the symmetry group $SU(3) \otimes SU(2) \otimes U(1)$, which has in total twelve generators. Each generator is corresponding to a *gauge boson* as a force mediator. Among the twelve gauge bosons, four of them ($\gamma, W^\pm$, and Z) are associated with the electroweak (EW) symmetry group $SU(2) \otimes U(1)$, and the remaining eight (gluons with different “color” quantum number pair permutations) are associated with the QCD symmetry group $SU(3)$. To preserve the gauge symmetry, all these gauge bosons should have vanishing masses. However, the fact that the weak interaction is short-ranged suggests that the $W^\pm$ and Z bosons are actually massive. This was directly confirmed in 1983, when both the $W^\pm$ [1] and Z [2] bosons were observed at the CERN Super Proton Synchrotron (SPS) with quite heavy masses¹.

The Englert–Brout–Higgs (EBH) mechanism [4–9] provides a general framework to keep untouched the gauge symmetry of the SM, while still generate the observed masses of $W^\pm$ and Z bosons through electroweak symmetry breaking (EWSB). In the simplest case, a scalar field (Higgs field) is incorporated into the SM as a $SU(2)$ complex doublet with four degrees of freedom. Consequently, three massless Goldstone bosons will be generated and absorbed as the longitudinal

¹ The latest measured value of $W^\pm$ and Z masses are $m_{W^\pm} = 80.385 \pm 0.015$ GeV, $m_Z = 91.1876 \pm 0.0021$ GeV [3].

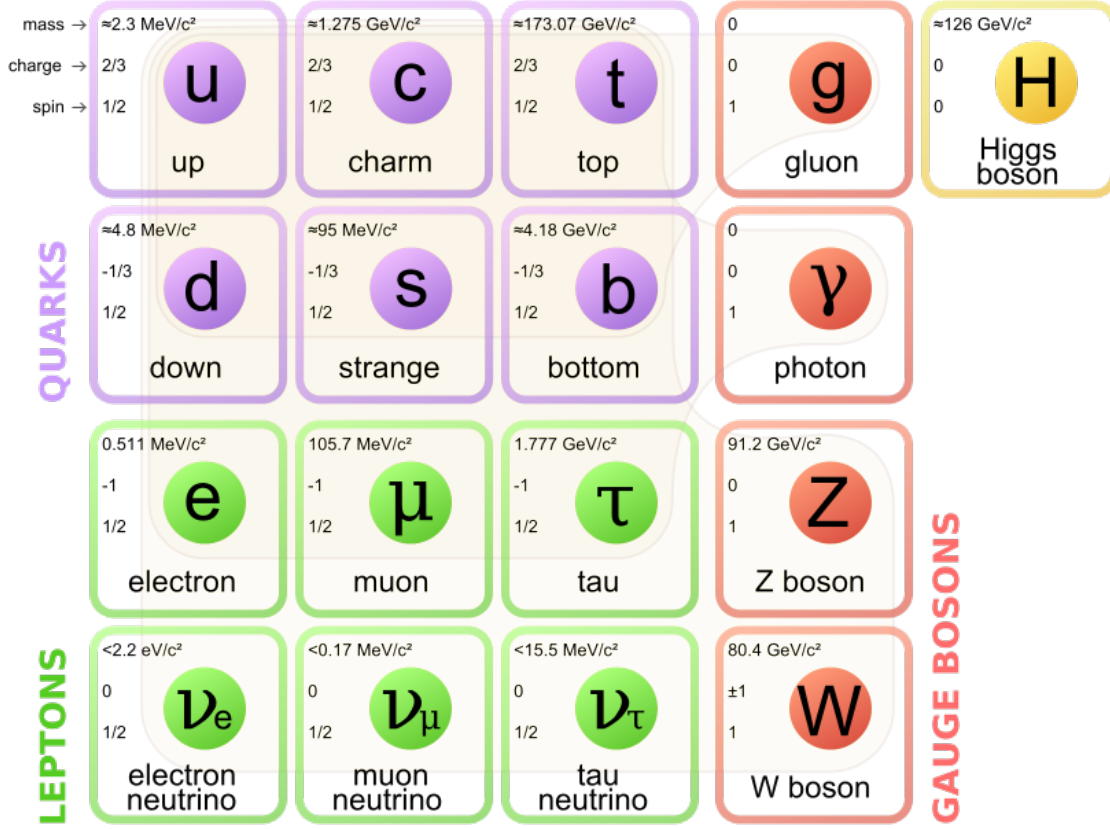


Figure 1.1: Summary of elementary particles in the Standard Model.

components of $W^\pm$ and Z bosons to give them masses, whereas the remaining component of the complex doublet will become a new elementary scalar particle, the Higgs boson, which could be observed by experiments. Once EWSB occurs, the elementary fermions (spin $\frac{1}{2}$ particles) in the SM can also acquire their masses through the Yukawa interactions with the Higgs field.

In spite of the intriguing physics case, searching for the Higgs boson is a challenging task because the mass of the Higgs boson is not predicted by the SM. Direct searches for the Higgs boson were conducted by the experiments at CERN Large Electron-Positron Collider (LEP) [10] and Fermilab Tevatron [11]. Though none of these experiments managed to establish an observation², they provided guidance to future experiments by setting a lower limit of 114.4 GeV on the Higgs boson mass (m_H) and excluding additional region at higher masses.

² A broad excess in data was seen at Tevatron in the Higgs boson mass range $115 \text{ GeV} < m_H < 140 \text{ GeV}$, with a local significance of 3 standard deviations at $m_H = 125 \text{ GeV}$.

The Large Hadron Collider (LHC) at CERN opens a grand new possibility for the search of the Higgs boson. LHC is a circular collider with a circumference of 27 kilometers. It can accelerate and collide two proton beams up to center-of-mass energy $\sqrt{s} = 14$ TeV with an instantaneous luminosity up to more than $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. On July 4th, 2012, as the finale of the decades-long experimental endeavor, the ATLAS [12] and CMS [13] Collaborations announced the observation of a Higgs-like boson around 125 GeV, based on proton–proton (pp) collision data collected in 2011 and first half of 2012. Since then further studies based on more LHC data have shown that the properties of the new particle are consistent with the Higgs boson, and the tag “-like” is hence officially removed from the name of the particle.

In the SM, the Higgs boson is a short-lived particle which decays immediately (in $O(10^{-22} \text{ s})$) into other particles once produced. Among all the Higgs boson decay channels, the diphoton decay channel is playing a pivotal role at the LHC experiments. During the establishment of the Higgs boson discovery, the diphoton decay channel contributed the most prominent excess in both ATLAS (4.5 standard deviations from background-only hypothesis) and CMS (4.1 standard deviations) results. At the measurement stage, the diphoton decay channel remains crucial for understanding the Higgs boson couplings because of its excellent sensitivity and unique loop-induced decay mechanism. It is also one of the only two channels at the LHC (the other is $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$) which can provide measurement of the Higgs boson mass with minimum model dependency.

While the discovery of the Higgs boson is a great success of the SM, it is known that the current theory cannot explain some established phenomena, such as non-zero neutrino mass or the presence of dark matter. Therefore, the need for searching physics Beyond the Standard Model (BSM) is compelling in general. Because many BSM scenarios predict resonance decaying into two photons, the diphoton decay channel can again serve as an instrumental tool in this effort.

This thesis will present the discovery of the Higgs boson and the measurements of its properties in the diphoton decay channel based on the LHC Run 1 data, as well as the search for high mass BSM scalar resonance in the diphoton final state based on the LHC Run 2 data. The rest of the thesis is organized as follows: Chapter 2 briefly reviews the production and decay of the Higgs boson, and also background processes in the diphoton final state at the LHC; Chapter 3 introduces the ATLAS detector; Chapter 4 describes data and Monte Carlo (MC) simulation samples used in

the analyses to be presented in this thesis; Chapter 5 details the definition of photons and other physics objects; Chapter 6 reports diphoton event selection criteria used in each analysis; Chapter 7 reviews the signal and background modeling strategies in the diphoton analyses; Chapter 8 presents the methodology used for the statistical interpretations of the data; Chapter 9 discusses the discovery of the Higgs boson in the diphoton decay channel based on 4.8 fb^{-1} of pp collision data collected at $\sqrt{s} = 7 \text{ TeV}$ in 2011, and 5.9 fb^{-1} collected at $\sqrt{s} = 8 \text{ TeV}$ in 2012 by the ATLAS detector (denoted as *Discovery Analysis* hereinafter); Chapter 10 reports the measurements of the Higgs boson couplings in the diphoton decay channel based on 4.5 fb^{-1} of pp collision data collected at $\sqrt{s} = 7 \text{ TeV}$ in 2011, and 20.3 fb^{-1} collected at $\sqrt{s} = 8 \text{ TeV}$ in 2012 by the ATLAS detector (denoted as *Coupling Analysis* hereinafter); Chapter 11 presents the measurement of the Higgs boson mass in the diphoton decay channel based on the same dataset used for the *Coupling Analysis* (denoted as *Mass Analysis* hereinafter; *Coupling* and *Mass Analyses* will be collectively denoted as *Measurement Analyses*), and the combined measurement of Higgs boson mass by ATLAS and CMS experiments in $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ channels based on full LHC Run 1 dataset collected by both experiments; Chapter 12 details the search of high mass BSM scalar particle in the diphoton decay channel based on 15.4 fb^{-1} of pp collision data collected at $\sqrt{s} = 13 \text{ TeV}$ in 2015 and first half of 2016 by the ATLAS detector (denoted as *Search Analysis* hereinafter); Chapter 13 provides a summary of all the discussions and an outlook to future prospects.

Chapter 2

Phenomenology

The SM Higgs boson is a CP -even spin-0 neutral particle. Its mass is given by $m_H = \sqrt{2\lambda}v$, where λ characterizes the Higgs self-coupling, and $v = (\sqrt{2}G_F)^{-1/2} \approx 246$ GeV is the vacuum expectation value of the Higgs field. Since λ is a free parameter, the Higgs boson mass is not predicted by theory a priori as mentioned in Chapter 1. Once m_H is known, all the other properties of the SM Higgs boson can be fixed in principle.

The SM Higgs boson couplings to other particles are set by their masses. They are summarized in the following Lagrangian:

$$\mathcal{L} = -g_{Hff}f\bar{f}H + \frac{g_{HHH}}{6}H^3 + \frac{g_{HHHH}}{24}H^4 + \delta_V V_\mu V^\mu (g_{HVV}H + \frac{g_{HHVV}}{2}H^2), \quad (2.1)$$

where f stands for elementary fermions, and V stands for $W^\pm$ ($\delta_W = 1$) or Z ($\delta_Z = 1/2$) bosons.

The couplings can be more explicitly written as:

$$\begin{aligned} g_{Hff} &= m_f/v, \quad g_{HVV} = 2m_V^2/v, \quad g_{HHVV} = 2m_V^2/v^2, \\ g_{HHH} &= 3m_H^2/v, \quad g_{HHHH} = 3m_H^2/v^2. \end{aligned} \quad (2.2)$$

As one can see the coupling strengths to elementary fermions are proportional to the fermion masses, whereas the coupling strengths to bosons are proportional to the square of the boson masses.

The rest of the chapter will provide a briefly review of the production and decay of the Higgs boson at the LHC [14–16], and also discuss the background processes entering the LHC diphoton analyses.

2.1 Higgs boson production at Large Hadron Collider

The SM Higgs boson can be produced at the LHC mainly via the following processes:

- gluon fusion production (ggF) $gg \rightarrow H$;
- vector boson fusion production (VBF) $qq' \rightarrow qq'H$;
- associated production with a W boson (WH) or a Z boson (ZH) $q\bar{q} \rightarrow WH/ZH$, including a small contribution (around 8%) from $gg \rightarrow ZH$ ($ggZH$);
- associated production with a pair of bottom quarks ($b\bar{b}H$) $q\bar{q}/gg \rightarrow b\bar{b}H$, or a pair of top quarks ($t\bar{t}H$) $q\bar{q}/gg \rightarrow t\bar{t}H$;
- associated production with a single top quark (tH) through t -channel ($tHqb$) $qg \rightarrow tHq'b$ processes (four-flavor scheme), or in association with a W boson (tHW) $gb \rightarrow tHW$ (five-flavor scheme); s -channel production is negligible.

The production cross sections of different processes changing as a function of Higgs boson mass at $\sqrt{s} = 8$ TeV, and as a function of center-of-mass energy with hypothesized Higgs boson mass of 125 GeV are summarized in Figure 2.1 (a) and (b), respectively.

The ggF process is expected to be the dominant Higgs boson production process at the LHC, accounting for about 90% of the total production cross section around $m_H=125$ GeV. It is hence fundamental for the discovery of the Higgs boson and measurement of its properties. When searching for new heavy scalar particles (Chapter 12), the ggF process with modified coupling structure is also commonly assumed as the major production mechanism. Since the gluon is massless, its interaction with the Higgs boson has to be intermediated by heavy quark loop. The leading-order (LO) Feynman diagram of ggF is shown in Figure 2.2 (a). The inclusive cross section of the ggF process used to estimate the expected event rate in this thesis is taken from a calculation at next-to-next-to-leading order (NNLO) [17–22] in QCD¹. Next-to-leading-order (NLO) EW corrections are also included [23, 24].

The VBF process has the second largest cross section at the LHC. Its LO Feynman diagram is shown in Figure 2.2 (b). This production mode is featured by two forward jets, which can be exploited to separate events produced by VBF from those by ggF . The VBF process can be used to probe the Higgs boson couplings to the W/Z bosons. Its cross section is calculated with full NLO QCD and EW corrections [25–27] with an approximate NNLO QCD correction applied [28].

¹ Recently the next-to-next-to-next-to-leading order (N³LO) calculation became available.

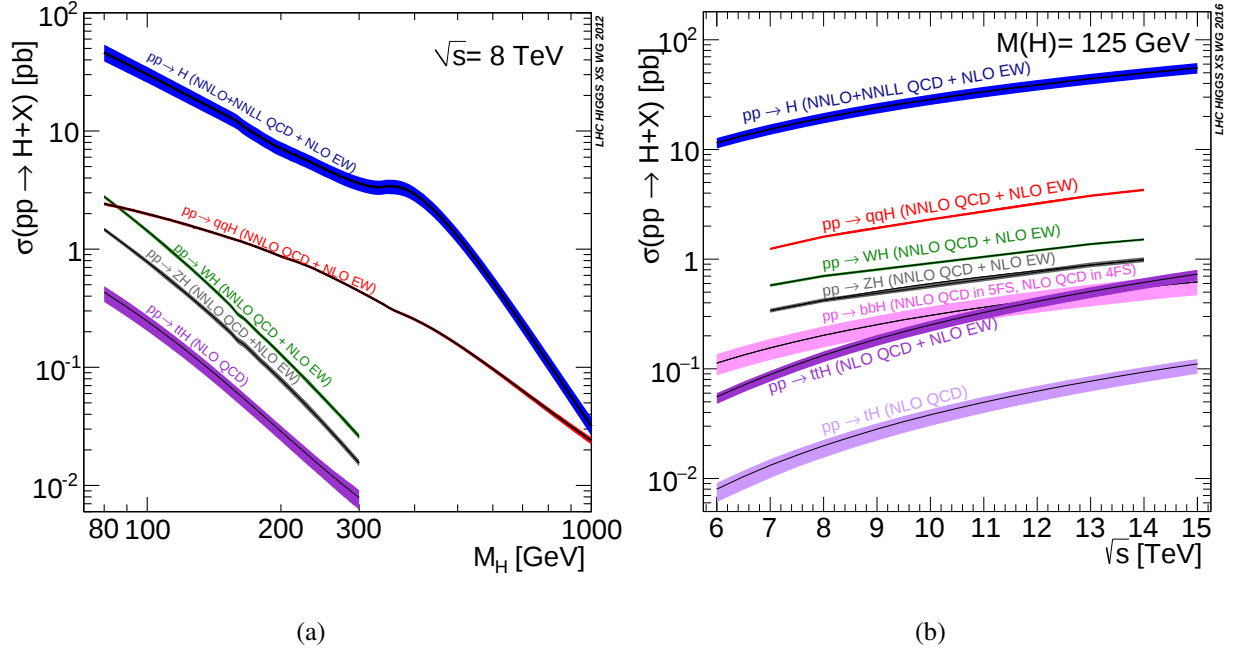


Figure 2.1: Standard Model Higgs boson production cross sections (a) as a function of Higgs boson mass at $\sqrt{s} = 8$ TeV and (b) as a function of center-of-mass energy with hypothesized Higgs boson mass of 125 GeV. The central values are shown as solid lines, and the theoretical uncertainties are shown as color bands. The $b\bar{b}H$ and tH processes are not included in the left plot.

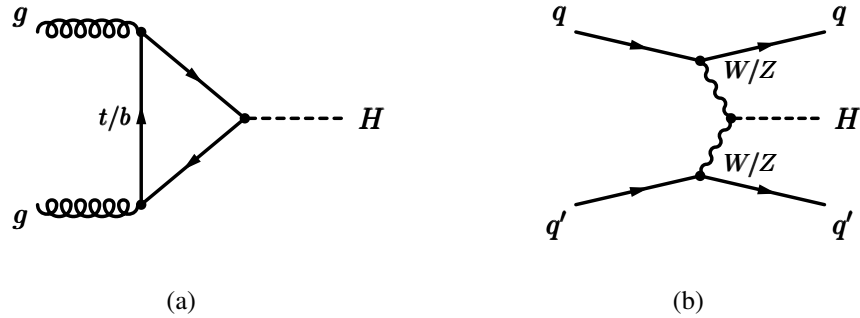


Figure 2.2: Examples of leading-order Feynman diagrams for Higgs boson production via the (a) ggF and (b) VBF production processes.

As shown in Figure 2.3, the WH and ZH processes (collectively denoted as VH process) also provide probes to the couplings to the W/Z bosons. Analyses on these two processes benefit from their distinct event topologies due to the presence of W/Z bosons decay products (leptons, quarks, and/or missing transverse momentum) in the final states. The $ggZH$ process has very small cross

section compared with $qqZH$, and is hence not considered in any diphoton analyses detailed in this thesis due to the lack of sensitivity. The VH inclusive cross sections are calculated at NNLO [29] with NLO EW radiative corrections [30] applied.

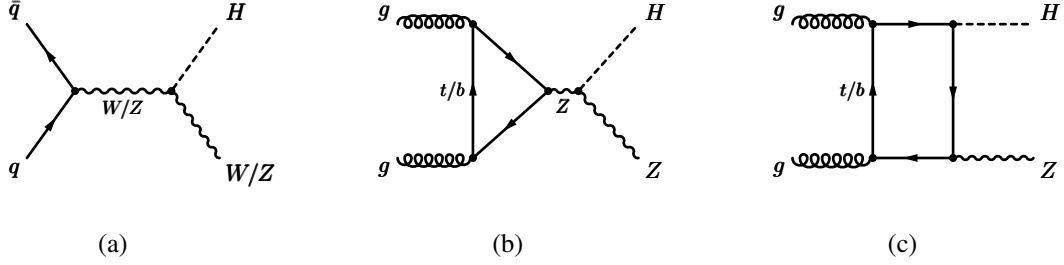


Figure 2.3: Examples of leading-order Feynman diagrams for Higgs boson production via the (a) VH and (b, c) $ggZH$ production processes.

The $b\bar{b}H$ and $t\bar{t}H$ productions (Figure 2.4) have relatively small cross section compared with other processes. The $b\bar{b}H$ production is very difficult to study at LHC because the associated b quarks are too soft to be efficiently tagged. This process has been considered in the measurements of Higgs boson properties for completeness, but was not considered during the Higgs boson discovery time. Its cross section is calculated in a four-flavor scheme at NLO QCD [31–33] and a five-flavor scheme at NNLO QCD [34]. These two calculations are combined using the Santander matching procedure [15, 35]. The $t\bar{t}H$ production, on the other hand, is more extensively explored at the LHC to probe of the large Yukawa coupling between the top quark and the Higgs boson. The full NLO QCD corrections are included [36–39] in the $t\bar{t}H$ cross section calculation.

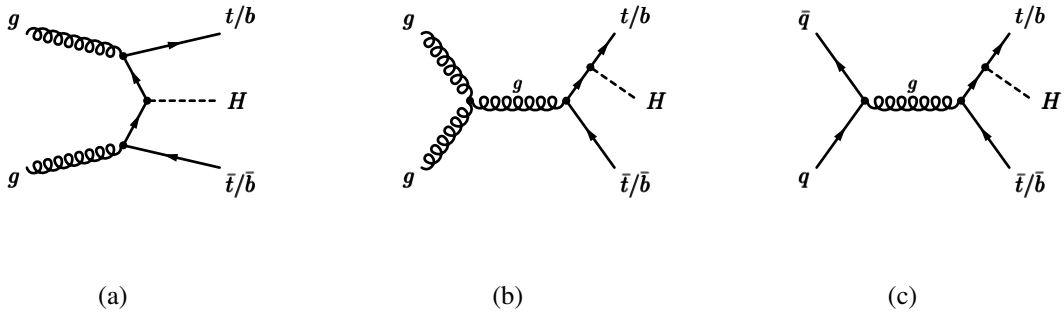


Figure 2.4: Examples of leading-order Feynman diagrams for Higgs boson production via the $t\bar{t}H$ and $b\bar{b}H$ processes.

Finally, the tH production (Figure 2.5) has smallest cross section among all processes, but it can be exploited together with $t\bar{t}H$ to put direct constraints on both strength and sign of the Yukawa coupling between top quark and the Higgs boson, because it is sensitive to altered Yukawa couplings. The cross sections for $tHqb$ production are calculated for SM as well as altered Yukawa couplings at LO using MADGRAPH [40, 41]. LO-to-NLO K-factors are obtained by comparing to NLO cross sections calculated using MADGRAPH5_AMC@NLO. The cross sections for tHW production are calculated for different Yukawa couplings at NLO in QCD using MADGRAPH5_AMC@NLO with dynamic renormalization and factorization scales [41].

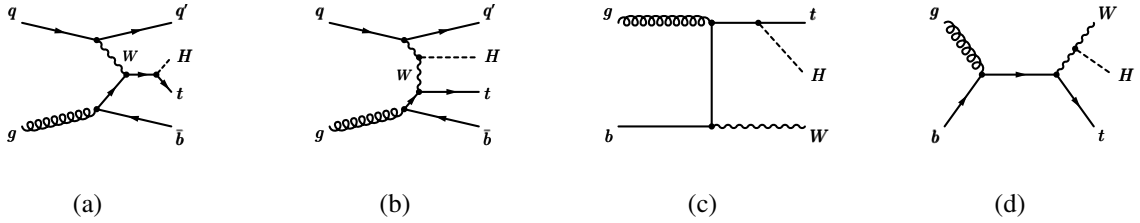


Figure 2.5: Examples of leading-order Feynman diagrams for Higgs boson production in association with a single top quark via the (a, b) $tHqb$ and (c, d) tHW production processes shown in four-flavor and five-flavor schemes, respectively.

The numerical values of the cross sections for each process at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV with $m_H = 125.09$ GeV² are summarized in Table 2.1 together with theoretical uncertainties and orders of calculations. The uncertainties due to missing higher-order terms in the perturbative calculations of QCD processes are estimated by varying the factorization and renormalization scales. There are additional uncertainties related to the parton distribution functions (PDFs) and the strong coupling constant α_S .

² The combined measurement of m_H by ATLAS and CMS Experiments based on LHC Run 1 data gives $m_H = 125.09 \pm 0.24$ GeV. More details can be found in Chapter 11.

Production process	Cross section [pb]		Order of calculation
	$\sqrt{s} = 7 \text{ TeV}$	$\sqrt{s} = 8 \text{ TeV}$	
ggF	15.0 ± 1.6	19.2 ± 2.0	NNLO(QCD) + NLO(EW)
VBF	1.22 ± 0.03	1.58 ± 0.04	NLO(QCD+EW) + APPROX. NNLO(QCD)
WH	0.577 ± 0.016	0.703 ± 0.018	NNLO(QCD) + NLO(EW)
ZH	0.334 ± 0.013	0.414 ± 0.016	NNLO(QCD) + NLO(EW)
$[ggZH]$	0.023 ± 0.007	0.032 ± 0.010	NLO(QCD)
$t\bar{t}H$	0.086 ± 0.009	0.129 ± 0.014	NLO(QCD)
tH	0.012 ± 0.001	0.018 ± 0.001	NLO(QCD)
$b\bar{b}H$	0.156 ± 0.021	0.203 ± 0.028	5FS NNLO(QCD) + 4FS NLO(QCD)
Total	17.4 ± 1.6	22.3 ± 2.0	

Table 2.1: Standard Model predictions for the Higgs boson production cross sections together with their theoretical uncertainties at $\sqrt{s} = 7 \text{ TeV}$ and $\sqrt{s} = 8 \text{ TeV}$. The value of the Higgs boson mass is assumed to be $m_H = 125.09 \text{ GeV}$. The uncertainties on the cross sections are evaluated as the sum in quadrature of the uncertainties resulting from variations of the QCD scales, parton distribution functions, and α_s . The order of the theoretical calculations is also indicated. In the case of the $b\bar{b}H$ production, the values are given for the mixture of five-flavor (5FS) and four-flavor (4FS) schemes.

2.2 Higgs boson decay branching ratios and total width

The SM Higgs boson branching ratios at $m_H = 125.09$ GeV are listed in Table 2.2. The dominant decay mode is $H \rightarrow b\bar{b}$ (57.5%), whereas the $H \rightarrow \gamma\gamma$ branching ratio is only 0.228%. Nevertheless, as discussed in Chapter 1 the $H \rightarrow \gamma\gamma$ channel remains one of the most competitive Higgs boson decay channels at the LHC. From experimental point of view, the power of this channel is mainly from good diphoton mass resolution and clean signature, which ensure excellent sensitivity and straightforward analysis strategy. From theoretical point of view, this channel is particularly interesting because of the loop-induced decay mechanism as shown in Figure 2.6. Within the SM, the destructive interference between the W boson and the top quark in the loop allows lifting the sign degeneracy between Higgs boson couplings to these two particles (or to bosons versus to fermions in general). Potential new charged particles from BSM scenarios can also contribute to the loop, and thus shift the decay rate from the SM expectation.

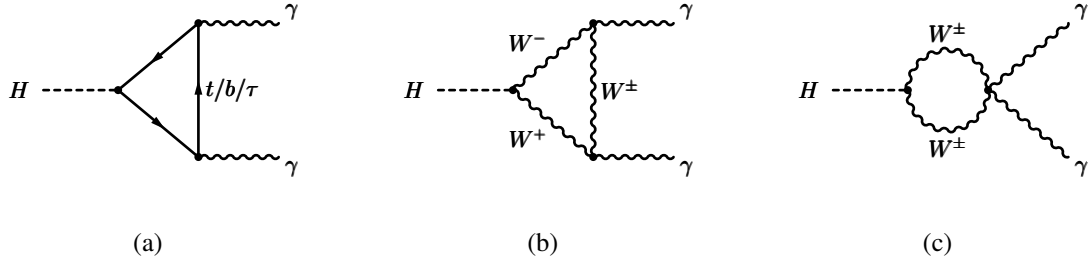


Figure 2.6: Examples of leading-order Feynman diagrams for Higgs boson decays to diphoton.

Examples of LO Feynman diagrams for the Higgs boson decays to W/Z bosons and fermions are shown in Figure 2.7. The total width of the SM Higgs boson with m_H near 125 GeV is predicted to be around 4 MeV, which is negligible compared with current detector resolution ($O(1$ GeV)).

Finally, it is worth pointing out that the natural widths of scalar particles are not necessarily narrow, in particular, when the particle is heavy and/or new physics are involved. Hence in the search of a heavy scalar diphoton resonance (Chapter 12) a scan on the width of the potential new particle is performed.

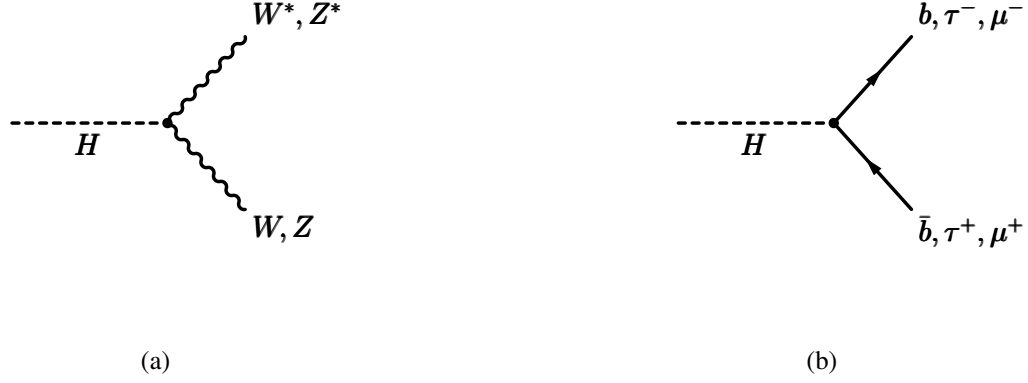


Figure 2.7: Examples of leading-order Feynman diagrams for Higgs boson decays (a) to W and Z bosons (one of the two bosons is off-shell for m_H near 125 GeV) and (b) to fermions.

Decay mode	Branching fraction [%]
$H \rightarrow b\bar{b}$	57.5 ± 1.9
$H \rightarrow WW^{(*)}$	21.6 ± 0.9
$H \rightarrow gg$	8.56 ± 0.86
$H \rightarrow \tau\tau$	6.30 ± 0.36
$H \rightarrow cc$	2.90 ± 0.35
$H \rightarrow ZZ^{(*)}$	2.67 ± 0.11
$H \rightarrow \gamma\gamma$	0.228 ± 0.011
$H \rightarrow Z\gamma$	0.155 ± 0.014
$H \rightarrow \mu\mu$	0.022 ± 0.001

Table 2.2: Standard Model predictions for the decay branching fractions of a Higgs boson with a mass of 125.09 GeV, together with their theoretical uncertainties.

2.3 Background processes in diphoton final state at Large Hadron Collider

In the SM, the diphoton events can also be produced promptly through several non-resonant processes at the LHC, including Born process $q\bar{q} \rightarrow \gamma\gamma$ (Figure 2.8 (a)), box process $gg \rightarrow \gamma\gamma$

(Figure 2.8 (b)), and Bremsstrahlung process $qg \rightarrow \gamma\gamma q$ (Figure 2.8 (c)). The total production cross section of these SM non-resonant diphoton ($\gamma\gamma$) processes within the fiducial acceptance of typical $H \rightarrow \gamma\gamma$ analyses is $O(100)$ larger than the SM Higgs boson production cross section times diphoton branching fraction [42]. They build up the *irreducible background* in the diphoton analyses, usually accounting for about 80% to 90% of the total background depending on detailed selection criteria.

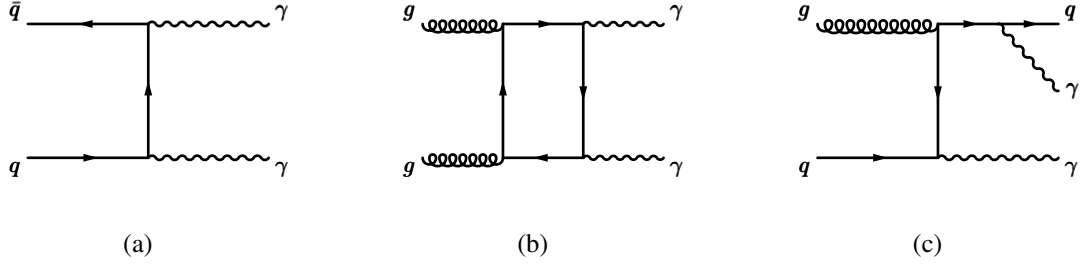


Figure 2.8: Examples of leading-order Feynman diagrams for non-resonant Standard Model diphoton production process.

The second largest contribution to the background comes from the SM photon–jet (γj) processes, where a jet is mis-identified as a photon. Figure 2.9 provides selected LO Feynman diagrams for such processes. While the total cross section of γj processes is $O(10^6)$ times larger than $H \rightarrow \gamma\gamma$ within fiducial acceptance of the analyses to be discussed [42], it is typically controlled to be 10 to 20% of the total background due to effective photon quality requirements to be discussed in Chapter 5. In addition, di-jet (jj) processes ($O(10^9)$ times larger cross section than $H \rightarrow \gamma\gamma$ [42]) will also contribute up to a few percent. All these background components that contain at least one fake photon are collectively classified as the *reducible background* in the diphoton analyses.

In the end, the typically signature of the diphoton analyses, as to be shown in coming chapters, is a narrow resonant signal peak sitting on top of smoothly decaying SM continuum background in the diphoton invariant mass spectrum. One thing worth pointing out is that in practice, analyses in the diphoton decay channel are usually not relying on accurate theoretical knowledge of background processes: the composition of background can be directly measured from data, and

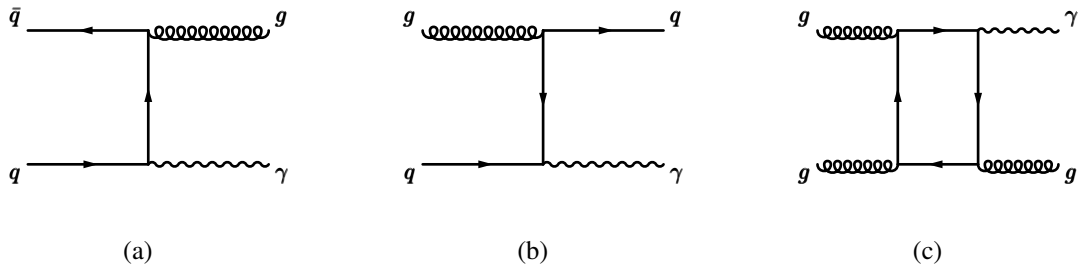


Figure 2.9: Examples of leading-order Feynman diagrams for non-resonant Standard Model photon-jet production process.

background estimates can usually be derived from parameterizing the data sideband and then interpolating into the signal region³. More details on these topics will be discussed in Chapter 6 and Chapter 7.

³ There are actually also diphoton analyses which use pre-defined background template to provide background estimation. For example please see “spin-2” analysis in Reference [43].

Chapter 3

ATLAS detector

The ATLAS detector [44] is a multipurpose detector with a forward-backward symmetric cylindrical geometry and nearly 4π coverage in solid angle.¹

The inner tracking detector (ID) covers the pseudorapidity range $|\eta| < 2.5$ and consists of a silicon pixel detector, a silicon microstrip detector, and a transition radiation tracker (TRT) in the range $|\eta| < 2.0$. The ID is immersed in a 2 T magnetic field produced by a thin superconducting solenoid located in front of the calorimeter. It allows an accurate reconstruction of charged-particle tracks originating from the proton–proton collision region as well as from secondary vertices, which permits an efficient reconstruction of photons interacting in the ID through e^+e^- pair production up to a radius in the transverse plane of about 80 cm.

The solenoid is surrounded by a high-granularity lead/liquid-argon (LAr) sampling electromagnetic (EM) calorimeter with an accordion geometry. The EM calorimeter measures the energy and the position of electromagnetic showers with $|\eta| < 3.2$. It is divided into a barrel section, covering the pseudorapidity region $|\eta| < 1.475$, and two end-cap sections, covering the pseudorapidity regions $1.375 < |\eta| < 3.2$. The transition region between the barrel and the end-caps, $1.37 < |\eta| < 1.52$, has a large amount of material upstream of the first active calorimeter layer. The EM calorimeter is composed, for $|\eta| < 2.5$, of three sampling layers longitudinal in shower depth. The first layer has a thickness of about 4.4 radiation lengths (X_0). In the ranges $|\eta| < 1.4$ and $1.5 < |\eta| < 2.4$, the first layer is segmented into high-granularity strips in the η direction, with a typical cell size of 0.003×0.0982 in $\Delta\eta \times \Delta\phi$ in the barrel. For $1.4 < |\eta| < 1.5$ and $2.4 < |\eta| < 2.5$

¹ ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the z -axis along the beam pipe. The x -axis points from the IP to the center of the LHC ring, and the y -axis points upward. Cylindrical coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar θ angle as $\eta = -\ln [\tan(\theta/2)]$.

the η -segmentation of the first layer is coarser, and the cell size is $\Delta\eta \times \Delta\phi = 0.025 \times 0.0982$. The fine η granularity of the strips is sufficient to provide, for transverse momenta up to $\mathcal{O}(100 \text{ GeV})$, an event-by-event discrimination between single photon showers and two overlapping showers coming from the decays of neutral hadrons, mostly π^0 and η mesons, in jets in the fiducial pseudorapidity region $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$. The second layer has a thickness of about $17 X_0$ and a granularity of 0.025×0.0245 in $\Delta\eta \times \Delta\phi$. It collects most of the energy deposited in the calorimeter by photon and electron showers. The third layer has a granularity of 0.05×0.0245 in $\Delta\eta \times \Delta\phi$ and a depth of about $2 X_0$. It is used to correct for leakage beyond the EM calorimeter of high-energy showers. In front of the accordion calorimeter, a thin presampler layer, covering the pseudorapidity interval $|\eta| < 1.8$, is used to correct for energy loss upstream of the calorimeter. The presampler consists of an active LAr layer with a thickness of 1.1 cm (0.5 cm) in the barrel (end-cap) and has a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.0982$. The material upstream of the presampler has a thickness of about $2 X_0$ for $|\eta| < 0.6$. In the region $0.6 < |\eta| < 0.8$ this thickness increases linearly from $2 X_0$ to $3 X_0$. For $0.8 < |\eta| < 1.8$ the material thickness is about or slightly larger than $3 X_0$, with the exception of the transition region between the barrel and the end-caps and the region near $|\eta| = 1.7$, where it reaches 5–6 X_0 . A sketch of a barrel module of the electromagnetic calorimeter is shown in Figure 3.1.

The hadronic calorimeter, which surrounds the EM calorimeter, consists of a steel/scintillator-tile calorimeter in the range $|\eta| < 1.7$ and two copper/LAr calorimeters spanning $1.5 < |\eta| < 3.2$. The acceptance is extended to $|\eta| = 4.9$ by two sampling calorimeters longitudinally segmented in shower depth into three sections using LAr as active material and copper (first section) or tungsten (second and third sections) as absorber.

The muon spectrometer (MS), located outside the calorimeters, consists of three large air-core superconducting toroid systems with precision tracking chambers that provide accurate muon tracking for $|\eta| < 2.7$ and fast detectors for triggering for $|\eta| < 2.4$.

A multi-level (three-level for Run 1, two-level for Run 2) trigger system is used to select events to be recorded for offline analyses. The first-level trigger is hardware-based using a subset of detector information. It reduces the event rate to 100 kHz level. High-level trigger(s) is software-based and has access to the full detector information. It further reduces the event rate to 1 kHz level.

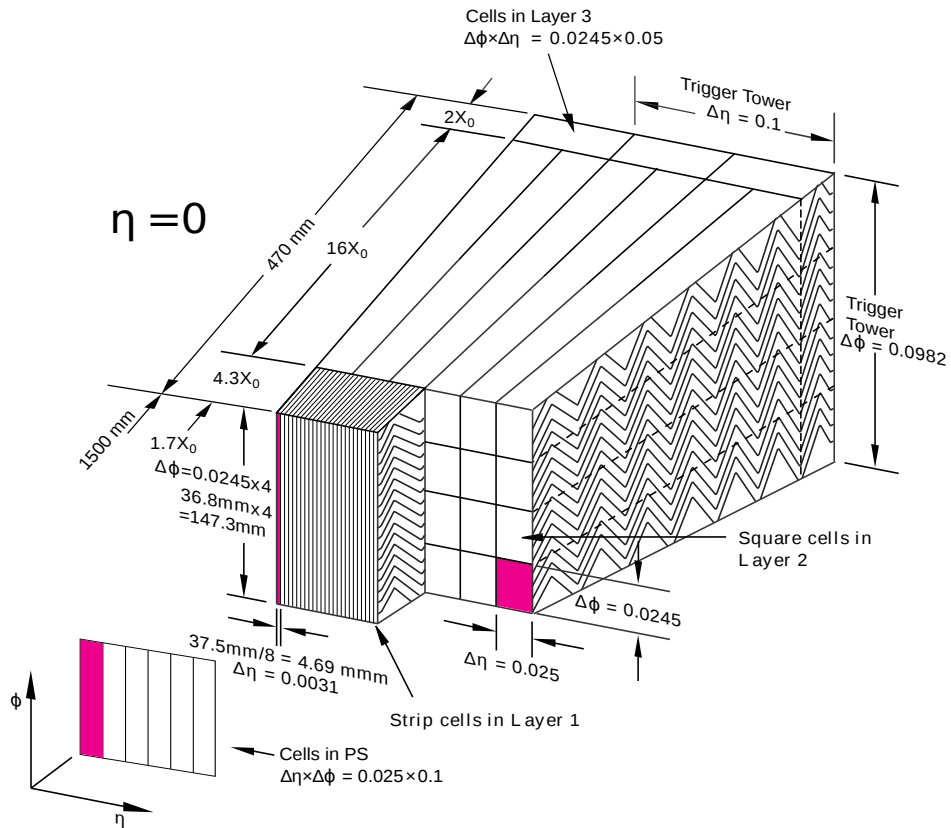


Figure 3.1: Sketch of a barrel module (located at $\eta = 0$) of the ATLAS electromagnetic calorimeter. The different longitudinal layers (one presampler and three layers in the accordion calorimeter) are depicted. The granularity in η and ϕ of the cells of each layer and of the trigger towers is also shown.

Run 2 triggers have improved capacities compared with Run 1 at both first-level and high-level due to infrastructure improvements [45].

Chapter 4

Data and simulation samples

4.1 Data samples

All the analyses reported in this thesis use pp collision data collected by the ATLAS detector in 2011, 2012, 2015 and 2016. Only events taken in stable beam conditions, and in which the trigger system, the tracking devices and the calorimeters were fully operational, are considered.

The dataset used for the discovery of Higgs boson (*Discovery Analysis*) corresponds to 4.8 fb^{-1} of integrated luminosity at $\sqrt{s} = 7 \text{ TeV}$ collected in 2011, and 5.9 fb^{-1} of integrated luminosity at $\sqrt{s} = 8 \text{ TeV}$ collected in 2012.

The dataset used for the measurements of Higgs boson properties (*Measurement Analyses*) corresponds to 4.5 fb^{-1} of integrated luminosity at $\sqrt{s} = 7 \text{ TeV}$ collected in 2011, and 20.3 fb^{-1} of integrated luminosity at $\sqrt{s} = 8 \text{ TeV}$ collected in 2012. In the ATLAS–CMS combined measurement of Higgs boson properties the CMS Run 1 dataset is also involved. It is corresponding to 5.1 fb^{-1} of 7 TeV data and 19.7 fb^{-1} of 8 TeV data.

The dataset used for the search of heavy scalar diphoton resonance (*Search Analysis*) corresponds to 15.4 fb^{-1} of integrated luminosity at $\sqrt{s} = 13 \text{ TeV}$ collected in 2015 and 2016.

4.2 Simulation samples for Standard Model Higgs boson signals

While the SM Higgs boson signal MC samples used for the *Discovery* and *Measurement Analyses* are largely unchanged, differences such as versions of theoretical programs and treatment of Higgs boson transverse momentum (p_T) spectra have been introduced over time between the two analyses. Since the prescriptions used in the *Measurement Analyses* are more advanced and harmonious in general, and the differences are essentially secondary for establishing the observation,

the discussion below will focus on the samples used for the *Measurement Analyses*. Readers can refer to Reference [12] for details of MC samples used for the *Discovery Analysis*.

Simulated samples of Higgs boson decaying into two photons are generated separately for different production modes. The ggF samples are generated with POWHEG-BOX [46–50] interfaced with PYTHIA8 [51] for the underlying event, parton showering and hadronization. The VBF samples are also generated using POWHEG-BOX [52] interfaced with PYTHIA8. The WH and ZH samples are generated with PYTHIA8. The $t\bar{t}H$ samples are generated using the POWHEG generator, a combination of the POWHEG-BOX and HELAC-NLO [53] generators, interfaced with PYTHIA8. Events from $tHqb$ production with different Yukawa couplings are generated with MADGRAPH [40] interfaced with PYTHIA8, whereas tHW events are generated using MADGRAPH5_AMC@NLO [54] interfaced to HERWIG++ [55]. The CT10 [56] PDF set is used for the POWHEG-BOX samples while CTEQ6L1 [57] is used for the PYTHIA8 samples. The AU2 [58] tuning of PYTHIA8 is used to simulate the minimum-bias events and the underlying event.

Samples mentioned above are generated at $m_H = 125$ GeV for $tHqb$ and tHW , from $m_H = 115$ GeV to $m_H = 135$ GeV in 5 GeV steps for $t\bar{t}H$, and from 110 GeV to 160 GeV in 5 GeV steps for the rest. The normalizations of samples from different Higgs boson production processes at different masses are determined according to the cross sections and branching ratios mentioned in Chapter 2. In the measurement of Higgs boson couplings (Chapter 10), interference of $gg \rightarrow H \rightarrow \gamma\gamma$ with the continuum $gg \rightarrow \gamma\gamma$ background induced by quark loops is taken into account using an averaging procedure that combines LO [59] and NLO corrections [60]. The destructive interference causes about 1% reduction of the ggF cross section. In addition, the interference between the Higgs boson signal and the continuum background is expected to produce a downward shift of the signal peak relative to the true value of Higgs boson mass (Chapter 11). The overall effect in the $H \rightarrow \gamma\gamma$ channel [59–61] is expected to be a few tens of MeV for a Higgs boson with a width near the SM value, which is small compared to the current experimental precision and is hence neglected.

Additional corrections to the shape of the generated p_T distribution of Higgs bosons produced by ggF are applied to match the distribution from a calculation at NNLO+NNLL provided by HRES2.1, which includes exact calculations of the effects of the top and bottom quark masses [62, 63] as well as dynamical renormalization and factorization scales. Calculations based

on HRES predict softer ggF Higgs boson p_T spectrum compared with the nominal POWHEG-BOX one, and thus the contribution from events with two or more jets, which mostly populate the high- p_T region, is affected. To simultaneously reproduce the inclusive ggF Higgs boson p_T distribution as well as the jet multiplicity $N_{\text{jet}} \geq 2$ component, the ggF events with two or more jets are first normalized to a NLO calculation [64]. Then, Higgs boson p_T -dependent weighting functions are determined using an iterative procedure. First, the events with two or more jets are weighted in order to match the Higgs boson p_T distribution from MINLO HJJ predictions [65]. As a second step, the inclusive spectrum is weighted to match the HRES distribution. These two steps are iteratively repeated until the inclusive Higgs p_T spectrum agrees well with the HRES prediction while preserving the normalization of the $N_{\text{jet}} \geq 2$ component. The events simulated for VBF, WH , and ZH production are re-weighted so that the p_T distributions of the Higgs bosons match the ones predicted by HAWK [25, 26, 66].

The $b\bar{b}H$ production is only considered in the *Measurement Analyses*. Since the p_T spectrum of the associated b -jets is expected to be rather soft, and the kinematic distributions between the ggF and $b\bar{b}H$ samples are checked to be quite similar, the selection efficiencies for $b\bar{b}H$ samples are assumed to be always the same as ggF ones, including cases where the presence of b -jets is required.

The stable particles, defined as the particles with a lifetime longer than 10 ps, are passed through a full detector simulation [67] based on GEANT4 [68]. Pileup effects are simulated by overlaying each MC event with a variable number of MC inelastic pp collisions generated using PYTHIA8, taking into account in-time pileup (collisions in the same bunch crossing as the signal), out-of-time pileup (collisions in other bunch crossings within the time-window of the detector sensitivity), and the LHC bunch train structure. The resulting detector signals are passed through the same event reconstruction algorithms as used for the data. The MC events are weighted to reproduce the distribution of the average number of interactions per bunch crossing observed in the data. Other known discrepancies between data and MC are also accounted for by applying corrections or assigning systematic uncertainties.

4.3 Simulation samples for high mass scalar signals

The scalar signal MC samples used in the *Search Analysis* are generated using the effective field theory approach implemented in MADGRAPH5_AMC@NLO at NLO in QCD. From the Higgs characterization framework [69], CP-even dimension five operators coupling the new resonance to gluons and photons are included. Samples are generated with the NNPDF3.0 NLO [70] PDF set, using the A14 tune [71] of PYTHIA8 for the parton shower. Samples are produced for fixed values of hypothesized signal masses (from 200 GeV to 2.4 TeV) and decay widths (from 4 MeV to 15% of the signal mass value). Interference effect between signal and background processes are neglected because it is expected to have limited impact [72, 73]. The generated events are passed through full detector simulation, with pileup simulated using PYTHIA8 with the AZNLO set of tuned parameters [74]. The simulated events are corrected for pileup and other discrepancies between data and MC.

4.4 Simulation samples for background processes

As mentioned in Chapter 2, the background estimates in the diphoton analyses discussed in this thesis can be derived from data-driven methods. Nevertheless, MC simulations are necessary in the procedure of selecting background fit models and evaluating associated systematic uncertainties to be discussed in Chapter 7.

In the *Discovery* and *Measurement Analyses*, $\gamma\gamma$ and γj background samples with large statistics ($O(10^8)$ events) are generated with SHERPA [75, 76], and the jj samples are generated with PYTHIA8. Since it is impossible to pass such large MC samples through full detector simulation, which is very computationally intensive, they are instead processed by a fast smearing procedure to account for major detector effects including photon and jet energy resolutions, photon identification efficiencies and fake rates, and photon conversion status etc., which are derived from full simulation samples. Due to the intrinsic limitation of the procedure, the *smear*ed MC samples do not contain b -jets, leptons, or missing transverse momentum (E_T^{miss}) measurement. When such objects are involved (in some categories of the *Coupling Analysis*), data control samples are usually used as replacement to study background models.

In the *Discovery Analysis*, the $\gamma\gamma$ events are also produced with RESBOS [77] and DIPHOX [78] as cross checks.

In the *Search Analysis* the $\gamma\gamma$ sample is also simulated with SHERPA, but it has relatively limited statistics ($O(10^6)$ events) and is passed through full detector simulation. The γj and jj background contributions are estimated from data control samples instead of MC simulations.

Chapter 5

Physics object definitions

5.1 Photons

5.1.1 Photon reconstruction

The photon reconstruction in the ATLAS detector is seeded by energy deposits (clusters) in the EM calorimeter with transverse energy $E_T > 2.5$ GeV in projective towers of size 0.075×0.123 in the $\eta \times \phi$ plane formed by a sliding-window algorithm [79]. The cluster reconstruction efficiency for photons and electrons with $E_T > 25$ GeV is estimated from simulation [80] to be close to 100%. The reconstruction algorithm looks for possible matches between energy clusters and tracks reconstructed in the inner detector and extrapolated to the calorimeter. Clusters matched to a well reconstructed track originating from a vertex found in the beam interaction region are classified as electron candidates. If the matched track(s) is consistent with originating from a photon conversion process $\gamma \rightarrow e^+e^-$, and if in addition a conversion vertex is reconstructed, the corresponding candidates are considered as *converted photons*. Converted photons are classified as single-track or double-track conversions depending on the number of assigned tracks. Clusters without matching tracks are classified as *unconverted photon* candidates. The efficiency to correctly reconstruct photons from the clusters and tracks is 96% for prompt photons with $E_T > 25$ GeV, while the remaining 4% are incorrectly reconstructed as electron candidates. The probability for a real electron with $E_T > 25$ GeV to be reconstructed as a photon fulfilling the *tight* identification criteria described below is measured in data to vary between 3% and 10%, depending on the pseudorapidity and the conversion class of the candidate [81]. The efficiency to reconstruct photon conversions decreases at high E_T (larger than 150 GeV) as it becomes more difficult to separate the two conversion tracks. The overall photon reconstruction efficiency is thus reduced to about 90% for E_T around 1 TeV.

5.1.2 Photon energy calibration

The understanding on photon energy calibration is evolving over time as more data are accumulated. In this subsection the focus will be on the calibration scheme used in the *Measurement Analyses* based on the Run 1 data [82]. Compared with the photon energy calibration used in the *Discovery Analysis* [12], the updated calibration scheme to be discussed gives about 10% improvement in the expected invariant mass resolution for $H \rightarrow \gamma\gamma$ events. The scheme has also been adopted for the calibration of the 13 TeV data sample in Run 2 used by the *Search Analysis* [83].

The energy measurement of photon starts from summing the energies measured in the EM calorimeter cells belonging to the candidate cluster. The size of the cluster depends on the photon classification: in the barrel, a $\Delta\eta \times \Delta\phi = 0.075 \times 0.123$ cluster is used for unconverted photons, and 0.075×0.172 for converted photons to account for the opening of the e^+e^- pair in the ϕ direction due to the magnetic field. In the end-cap, a cluster size of $\Delta\eta \times \Delta\phi = 0.125 \times 0.123$ is used for both types of candidates. The cluster energy is corrected for energy losses in the inactive materials in front of the calorimeter, for the fraction of energy deposited outside the cluster defined in the η - ϕ plane, and into the hadronic calorimeter in the direction of the shower propagation. Finally, due to the finite cluster size in η and ϕ coordinates, and the variation of the amount of absorber material crossed by incident particles as a function of ϕ , a correction is introduced account for the variation of the energy response as a function of the impact point on the calorimeter. The calibration coefficients used to make this correction are obtained from a detailed simulation of the detector geometry and are optimized with a Boosted Decision Tree (BDT) [84]. The response is calibrated separately for converted and unconverted photon candidates. The inputs to the energy calibration algorithm are the measured energy per calorimeter layer (including the presampler), the η position of the cluster, and the local position of the shower within the second-layer cell corresponding to the cluster centroid. In addition, the track transverse momenta and the conversion radius for converted photons are used as input to the regression algorithm to further improve the energy resolution, especially at low energy. The energy scales of the data and simulation are equalized by applying η -dependent correction factors to match the invariant mass distributions of $Z \rightarrow e^+e^-$ events. In this procedure, the simulated width of the Z boson resonance is matched to the one observed in data by adding a contribution to the constant term of the electron energy resolution.

The photon energy scales are confirmed by an independent analysis of radiative Z boson decays within uncertainties. The energy response of the calorimeter in data varies by less than 0.1% over time. The simulation is found to describe the dependence of the response on pileup conditions at the same accuracy level.

5.1.3 Photon identification

The photon identification algorithm is based on the lateral and longitudinal energy profiles of the shower measured in the EM calorimeter and the leakage into the hadronic calorimeter [85]. Prompt photons typically produce narrower energy profiles in the EM calorimeter and have smaller leakage to the hadronic calorimeter compared with background jets. In addition, neutral hadrons from jet fragmentation are often characterized by two separated energy maxima in the first strip layer of the EM calorimeter. Based on these discriminating variables, two reference selections, a *loose* one and a *tight* one, are defined. The *loose* selection is based on information only from the second layer of the EM calorimeter and the hadronic calorimeter, and is typically used in photon triggers and background studies. The *tight* selection adds information from the fine segmentation of the first layer of the EM calorimeter, and is typically used as part of the photon quality requirement in offline analyses. The photon identification criteria are optimized separately for unconverted and converted photon candidates. They are also tuned separately for different center-of-mass energy with different pileup conditions: the criteria used for the 8 TeV and 13 TeV are based on rectangular cuts optimized on MC simulations, while for the 7 TeV data the discriminating variables are combined into a single discriminant by a neural-network algorithm [84]. Small corrections are applied to the discriminating variables in the simulation to account for the observed mis-modeling of lateral shower profiles in the calorimeter. Correction factors as a function of η , E_T and conversion class are derived to correct for the residual mismatch between the efficiency in the simulation and the one measured in the data. The identification efficiency of the *tight* criteria is typically around 85% for photon candidates with $E_T > 40$ GeV [85, 86]. The efficiency is in general increasing with photon E_T and plateaus at high E_T . The efficiency for converted photons are typically a few percent higher than the one for unconverted photon. Typical rejection of background from the *tight* criteria is about 5000 [42].

5.1.4 Photon isolation

Two complementary isolation variables can be used to further suppress the jet background in the photon candidate samples. The first variable is the calorimetric isolation (E_T^{iso}), which is calculated by summing up the transverse energies of positive-energy topological clusters [79] deposited in the calorimeter within a cone of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$ around each photon. The energy sum excludes the contribution due to the photon cluster and an estimate of the energy deposited by the photon candidate outside its associated cluster. The median E_T density for the event in question, caused by the underlying event and additional minimum-bias interactions occurring in the same or neighboring bunch crossings (in-time and out-of-time pileup, respectively), is subtracted on an event-by-event basis using an algorithm described in Reference [87] and implemented as described in Reference [88].

In spite of the corrections described above, the calorimetric isolation is observed to have decreasing signal selection efficiency at high pileup. To improve the efficiency of the isolation selection for events with high pileup, the calorimetric isolation can be complemented by a track isolation (p_T^{iso}) defined as the scalar sum of the transverse momenta of all tracks with $p_T > 1$ GeV (0.4 GeV) for the 8 TeV and 13 TeV (7 TeV) data within a cone of size $\Delta R = 0.2$ around each photon. The track isolation efficiency is insensitive to out-of-time pileup, and its dependence on the in-time pileup is reduced by selecting only tracks consistent with originating from the diphoton production vertex (see next subsection) and not associated with converted photon candidates with requirements on impact parameters [81, 89]. The efficiency of the isolation cuts in the simulation is corrected by a small p_T -dependent factor extracted from measurements in data performed with a pure sample of photons from radiative $Z \rightarrow e^+e^-\gamma$ decays and $Z \rightarrow e^+e^-$ events.

5.1.5 Diphoton vertex selection

The invariant mass of the two photons $m_{\gamma\gamma}$ is given by

$$m_{\gamma\gamma} = \sqrt{2E_1E_2(1 - \cos \alpha)}, \quad (5.1)$$

where E_1 and E_2 are the energies of the two photons, and α is the opening angle between them with respect to their production vertex. The selection of the correct diphoton production vertex is hence important for the resolution of the α measurement, and thus for the precise measurement of

$m_{\gamma\gamma}$. Since the ATLAS EM calorimeter is longitudinally segmented, it can provide measurement of photon trajectories (photon pointing), which can then be extrapolated back to the beam interaction region. A position resolution on the diphoton production vertex of about 15 mm in the z direction with the photon pointing technique alone is achieved [42], which is sufficient to keep the contribution from the opening angle to the mass resolution smaller than the contribution from the energy resolution. Nevertheless, an efficient procedure to select the diphoton production vertex among the primary vertex candidates reconstructed with the ID is still necessary, because it allows the computation of the track-based quantities used in the quality requirements of physics objects, such as the photon track isolation discussed in the last subsection and the jet vertex fraction to be detailed in the next subsection.

The diphoton production vertex is selected using multivariate analysis (MVA) technique combining the position and recoil information of each candidate vertex. In the *Measurement and Search Analyses*, the following input variables are used¹: the combined z -position of the intersections of the extrapolated photon trajectories with the beam axis; the sum of the squared transverse momenta $\sum p_T^2$ and the scalar sum of the transverse momenta $\sum p_T$ of the tracks associated with the vertex; the difference in azimuthal angle $\Delta\phi$ between the direction defined by the vector sum of the track momenta and that of the diphoton system. For converted photons, the position of the conversion vertex is also used if tracks from the conversion have hits in the silicon detectors.

The production vertex selection is studied with $Z \rightarrow e^+e^-$ events in data and simulation by removing the electron tracks from the events and then measuring the efficiency for finding the vertex associated with the Z boson production. The MC simulation is found to accurately describe the efficiency measured in data, as shown in Figure 5.1 from the *Measurement Analyses*. The efficiency for finding the reconstructed diphoton primary vertex (ϵ_{PV}) in simulated $H \rightarrow gg$ events from ggF production within 0.3 mm (15 mm) of the true vertex is around 85% (93%) over the typical range of the number of collision vertices per event observed in the 8 TeV data. The efficiency ϵ_{PV} increases for large diphoton p_T as the hadron system recoiling against the diphoton evolves into one or more jets, which in turn contain additional higher p_T tracks. These additional tracks make it more likely to reconstruct the diphoton vertex as a primary vertex. Therefore, by re-weighting the simulated $Z \rightarrow e^+e^-$ events to approximate the harder p_T spectrum of the simulated Higgs boson

¹ The *Discovery Analysis* uses similar algorithm, but with fewer input variables and different MVA method.

signal, ϵ_{PV} is well reproduced. The selection efficiencies for the 7 TeV data and MC samples are slightly higher due to lower pileup. In the 13 TeV data the MVA algorithm is re-trained, and the corresponding efficiency ϵ_{PV} is about 88%.

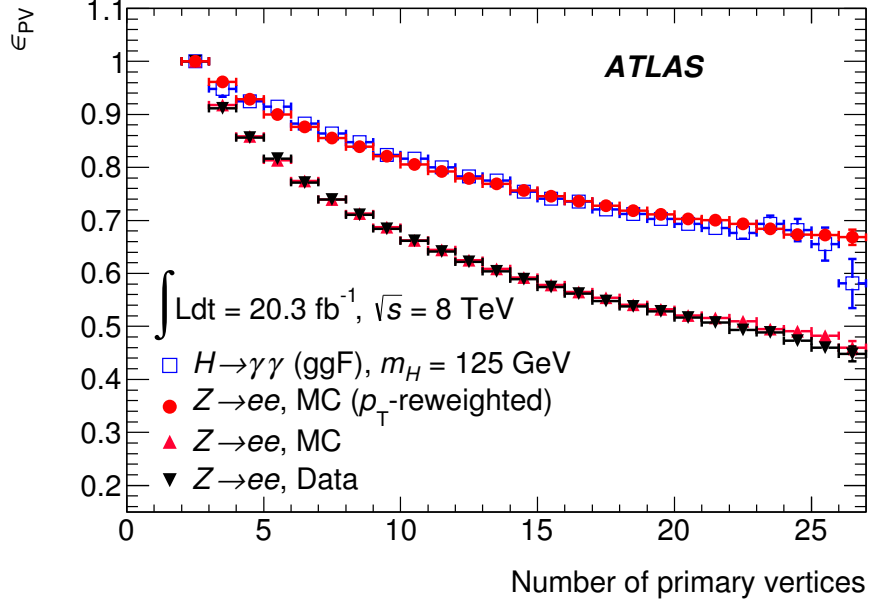


Figure 5.1: Efficiency to select a diphoton vertex within 0.3 mm of the production vertex (ϵ_{PV}) as a function of the number of primary vertices in the event in the *Measurement Analyses*. The plot shows ϵ_{PV} for simulated ggF events ($m_H = 125$ GeV) with two unconverted photons (empty blue squares), for $Z \rightarrow e^+e^-$ events with the electron tracks removed for the neural-network-based identification of the vertex, both in data (black triangles) and simulation (red triangles), and the same simulated $Z \rightarrow e^+e^-$ events re-weighted to reproduce the p_T spectrum of simulated ggF events (red circles).

5.2 Other physics objects

Leptons (muons and electrons), jets, and missing transverse momentum are essential for separating different production mechanisms of the Higgs boson, which is the key to the measurements of Higgs boson couplings. Below the definitions of these objects used in the *Coupling Analysis* will be briefly discussed. Jets are actually also used in the *Discovery Analysis* to select candidate VBF events, but with slightly different (and obsolete) definition which will be covered in Section 9.1.

5.2.1 Leptons

The reconstruction of electron candidates has been discussed together with photon candidates in Section 5.1.1. Electron candidates are required to satisfy the loose identification criterion of a likelihood-based discriminating variable [90]. A cut-based identification selection is used in the 7 TeV analysis and the electrons are required to fulfill the medium criteria defined in Reference [91]. The determination of the energy of the electron candidate is performed using a $\Delta\eta \times \Delta\phi = 0.075 \times 0.172$ cluster in the barrel to recover the energy spread in ϕ from bremsstrahlung photons, while a 0.125×0.123 cluster is used in the end-cap. The cluster energy is calibrated using similar procedure as discussed in Section 5.1.2 with a dedicated set of calibration coefficients optimized for electrons. The transverse momentum p_T of an electron is computed from the cluster energy and the track direction at the interaction point. Electrons are required to be in the region $|\eta| < 2.47$ and to satisfy $E_T > 15$ GeV. The combined electron reconstruction and identification efficiency for the analysis of the 8 TeV (7 TeV) data ranges from 86% (68%) to 93% (89%) for electron transverse energies between 15 GeV and 50 GeV [90,91]. Finally, the electron candidates must satisfy both the track-based and calorimetric isolation criteria relative to the E_T of the candidate. The calorimetric transverse isolation energy within a $\Delta R = 0.4$ cone is required to be less than 20% of the electron E_T , whereas the sum of the transverse momenta of the tracks within a cone of $\Delta R = 0.2$ around the track of the electron candidate is required to be less than 15% of the electron E_T .

Muon candidates are built from tracks reconstructed independently in the MS and in the ID and from energy deposits measured in the calorimeters [92]. Different types of muon candidates are built depending on the available information from the different sub-detector systems: the main algorithm combines tracks reconstructed separately by the ID and the MS. To extend the acceptance region beyond the ID limit to include $2.5 < |\eta| < 2.7$, tracks reconstructed in the MS standalone are used. Finally, to increase the acceptance for low- p_T muons or for muons that pass through uninstrumented regions of the MS, muon candidates are reconstructed from tracks in the ID associated with a track segment in the MS or to a calorimetric energy deposition compatible with the one from a minimum-ionizing particle. Muons from all different algorithms mentioned above are used and required to have $|\eta| < 2.7$ and $p_T > 10$ GeV. The combination of the different algorithms

ensure around 99% efficiency to detect a muon over the full acceptance range. A candidate is also required to satisfy exactly the same isolation criteria (relative to its p_T) as for electrons.

5.2.2 Jets

Jets are reconstructed using the anti- k_t algorithm [93] with radius parameter $R = 0.4$, and are required to have $|\eta| < 4.4$ and satisfy (unless stated otherwise) $p_T > 25$ GeV. Jets are discarded if they are within $\Delta R = 0.2$ of an isolated electron or within $\Delta R = 0.4$ of an isolated photon. The inputs to the jet-finding are topological calorimeter clusters [94] formed with the energy calibration appropriate for electromagnetic showers. The jet energy is calibrated using scale factors extracted from simulated dijet events by matching the energies of the generator-level and reconstructed jets. In addition, for the 8 TeV data, the pileup dependence of the jet response is suppressed by subtracting the median E_T density for the event multiplied by the transverse area of the jet [95, 96]. A residual pileup correction that is proportional to the number of reconstructed primary vertices and to the average number of interactions per bunch crossing further reduces the pileup dependence, in particular, in the forward region. Finally, the jet energy is corrected by an absolute scale factor determined using γ +jet, Z +jet and multijet events in data, and a relative η -dependent factor measured with dijet events in data. In order to suppress jets produced by pileup, jets within the tracking acceptance ($|\eta_j| < 2.4$) are required to have a jet vertex fraction (JVF) [96], defined as the sum of p_T of the tracks associated with the jet that are produced at the diphoton primary vertex, divided by the sum of p_T of the tracks associated with the jet from all collision vertices, larger than 0.5 (0.25) for the 7 TeV (8 TeV) data, respectively.

In order to identify jets containing a b -hadron (b -jets), a neural-network-based algorithm is used to combine information from the tracks in a jet. Since b -hadrons are long-lived (life time is approximately 10^{-12} s), the neural-network exploits the measurements of the impact parameters of the tracks, any secondary vertices, and the outputs of decay topology algorithms as discussed in References [97, 98]. Four different working points with efficiencies for identifying b -jets (rejection factors for light jets) of 60% (450), 70% (140), 80% (29), and 85% (11) are used in the analysis. The efficiencies and rejection factors at the working points are calibrated using control samples of data.

5.2.3 Missing transverse momentum

The measurement of the magnitude of the missing transverse momentum E_T^{miss} is based on the transverse energy of all photon, electron and muon candidates, all jets sufficiently isolated from these candidates, and all calorimeter energy clusters not associated with these candidates nor jets (*soft term*) [99]. In order to improve the discrimination of multi-jet events, where E_T^{miss} arises mainly from energy resolution effects, from events with a large fraction of E_T^{miss} due to non-interacting particles, an E_T^{miss} -significance is defined as $E_T^{\text{miss}}/\sigma_{E_T^{\text{miss}}}$, where the square root of the scalar sum of the transverse energies of all objects ΣE_T is used in the estimator of the E_T^{miss} resolution $\sigma_{E_T^{\text{miss}}} = 0.67 [\text{GeV}^{1/2}] \sqrt{\Sigma E_T}$. The proportionality factor $0.67 [\text{GeV}^{1/2}]$ is determined with fully reconstructed $Z \rightarrow \ell\ell$ events by removing the leptons in the measurement of E_T^{miss} [100].

Chapter 6

Diphoton event selections

Diphoton event selections in various analyses to be covered in this thesis are quite similar. In order to avoid duplicated discussions, this chapter will give a compact summary of their common aspects as well as differences.

6.1 Event selection used for discovery of Higgs boson

Samples used in the *Discovery Analysis* are selected using a diphoton trigger [101], which requires two clusters formed from energy depositions in the EM calorimeter. An E_T threshold of 20 GeV is applied to each cluster for the 7 TeV data, while for the 8 TeV data the thresholds are increased to 35 GeV for the leading (the highest E_T) cluster and to 25 GeV for the sub-leading (the next-highest E_T) cluster. In addition, *loose* criteria are applied to the shapes of the clusters to match the expectations for electromagnetic showers initiated by photons. The efficiency of the trigger is greater than 99% for events passing the final event selection.

Photon candidates are reconstructed in the fiducial region of the EM calorimeter defined by $|\eta| < 2.37$ ¹, excluding the calorimeter barrel/end-cap transition region $1.37 < |\eta| < 1.52$. Photon candidates in the fiducial region are ordered according to their E_T , and only the leading two photon candidates are considered. The leading and sub-leading photon candidates are required to have $E_T > 40$ GeV and 30 GeV, respectively. Photon candidates are required to pass *tight* identification criteria as discussed in Section 5.1.3. In addition, the calorimetric isolation of photon candidates E_T^{iso} is required to be less than 4 GeV. Finally, the diphoton invariant mass is required to be between 100 GeV and 160 GeV.

¹ Pseudorapidity defining the fiducial region of the EM calorimeter is measured from the second layer of the EM calorimeter.

With all the selections applied, 23788 events are selected from the 7 TeV data, and 35271 events are selected from the 8 TeV data. The composition of the dataset is estimated using data-driven methods detailed in Reference [102] for $\gamma\gamma$, γj and jj contributions. The fraction of $\gamma\gamma$ events in the selected sample is estimated to be $(80 \pm 4)\%$ in the 7 TeV data and $(75^{+3}_{-2})\%$ in the 8 TeV data. The fractions of γj and jj events are estimated to be $(19 \pm 3)\%$ ($(22 \pm 2)\%$) and $(1.8 \pm 0.5)\%$ ($(2.6 \pm 0.5)\%$) in the 7 TeV (8 TeV) data sample. Contribution from Drell-Yan process is also estimated to be only about 1%. The number of events for each component in the selected diphoton events sample, obtained independently in each bin of $m_{\gamma\gamma}$, is shown in Figure 6.1.

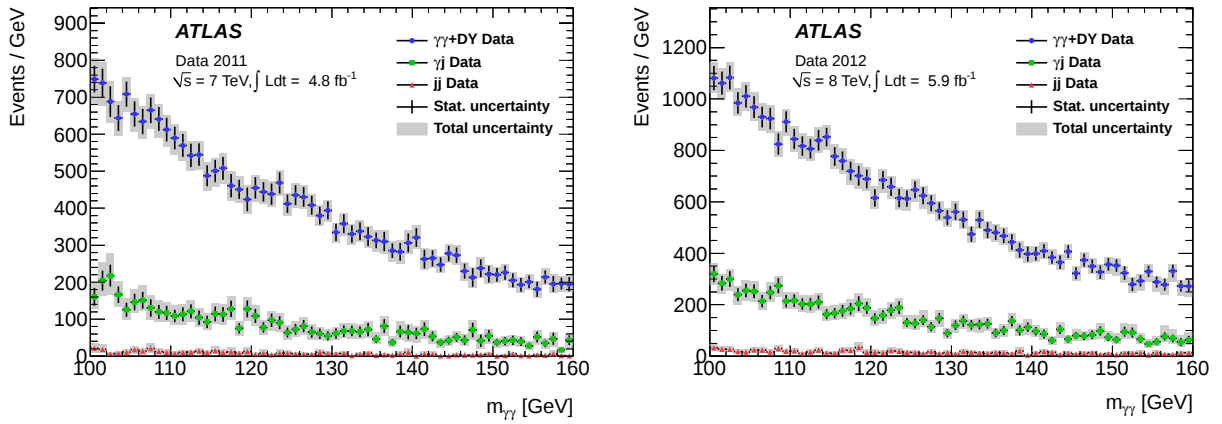


Figure 6.1: Diphoton sample composition as a function of the invariant mass for the 7 TeV (a) and the 8 TeV (b) dataset used in the *Discovery Analysis*. The small contribution from Drell-Yan events is included in the diphoton component. The error bars on each point represent the statistical uncertainty on the measurement while the colored bands represent the total uncertainty.

6.2 Event selection used for measurements of Higgs boson properties

Changes introduced to the diphoton event selection of the *Measurement Analyses* with respect to the *Discovery Analysis* are summarized below.

Fiducial region of the EM calorimeter: the barrel/end-cap transition region to be excluded is extended from $1.37 < |\eta| < 1.52$ to $1.37 < |\eta| < 1.56$ due to deteriorated performance of photon identification in the data sample.

E_T requirements: the E_T requirements for leading and sub-leading photon candidates change to $E_T/m_{\gamma\gamma} > 0.35$ and 0.25 , respectively. While the new requirements do not affect the sensitivity to

a SM Higgs boson with $m_H = 125$ GeV, compared with the constant E_T requirements used in the *Discovery Analysis* they make $m_{\gamma\gamma}$ spectra easier to model. The new criteria also allow harmonized selection with other $H \rightarrow \gamma\gamma$ analyses such as spin [103] and differential cross sections [104].

Isolation requirements: calorimetric isolation is loosened from 4 GeV to 6 GeV, and a track isolation requirement of less than 2.6 GeV is added to mitigate the increase of background and also reduce the dependency on pileup, as shown in Figure 6.2. In 7 TeV data sample the combination of calorimetric ($E_T^{\text{iso}} < 5.5$ GeV) and track isolation ($p_T^{\text{iso}} < 2.2$ GeV) requirements are tuned to achieve similar performance as those used for 8 TeV.

Mass range: because at $m_{\gamma\gamma} = 100$ GeV the offline E_T selection of 35 GeV (25 GeV) for leading (sub-leading) photon is right at the 8 TeV trigger threshold, the lower boundary of mass range is raised from 100 GeV to 105 GeV to avoid trigger inefficiency.

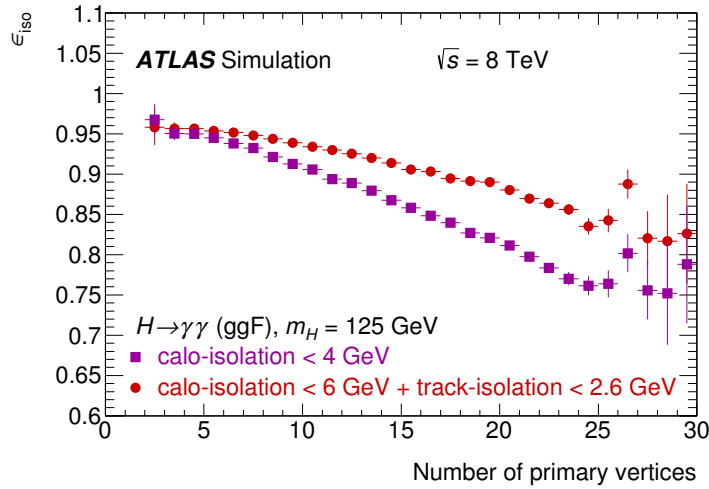


Figure 6.2: Efficiency to fulfill the isolation requirement (ϵ_{iso}) as a function of the number of primary vertices in each event in the *Measurement Analyses*, determined with a simulation sample of Higgs bosons decaying into two photons with $m_H = 125$ GeV and $\sqrt{s} = 8$ TeV. Events are required to satisfy the kinematic selection described in the text. The efficiency of the event selection obtained with a tight calorimetric isolation requirement (4 GeV) is compared with the case in which a looser calorimetric isolation (6 GeV) is combined with a track isolation (2.6 GeV) selection.

A total of 94566 and 17225 events are selected from the 8 TeV and 7 TeV data, respectively. The contributions from $\gamma\gamma$, γj and jj are estimated to be $84 \pm 8\%$ ($77 \pm 3\%$), $15 \pm 8\%$ ($20 \pm 2\%$),

and $1 \pm 1\%$ ($3 \pm 1\%$) for the 7 TeV ($\sqrt{s} = 8$ TeV) data, respectively based on method described in References [81]. The number of events for each component in the selected diphoton events sample, obtained independently in each bin of $m_{\gamma\gamma}$, is shown in Figure 6.3.

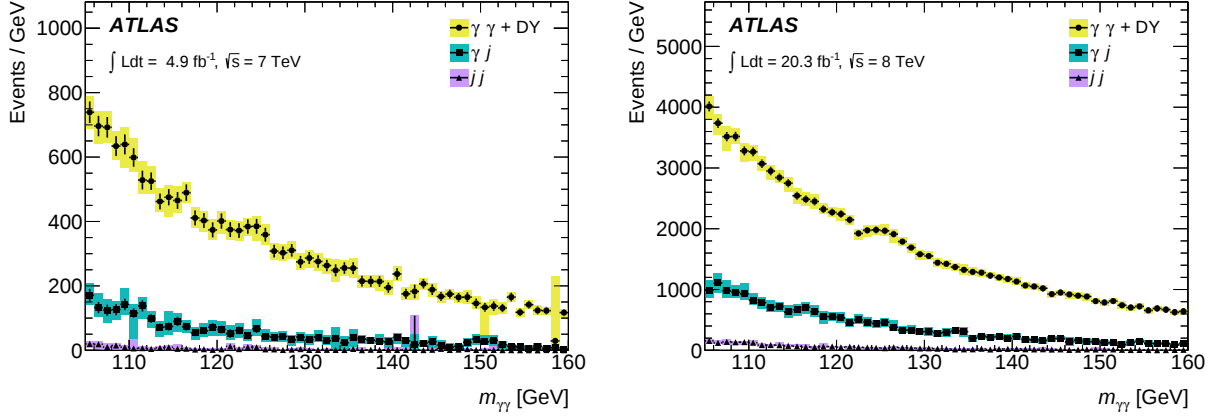


Figure 6.3: Diphoton sample composition as a function of the invariant mass for the 7 TeV (a) and the 8 TeV (b) dataset used in the *Measurement Analyses*. The small contribution from Drell-Yan events is included in the diphoton component. The error bars on each point represent the statistical uncertainty on the measurement while the colored bands represent the total uncertainty.

6.3 Event selection used for search of high mass scalar particle

Selections used in the *Search Analysis* have been optimized for searching new scalar particle with mass heavier than the SM Higgs boson. More details are summarized below.

Trigger: trigger threshold (35 GeV for leading photon candidate, 25 GeV for sub-leading) is unchanged with respect to the one used for the 8 TeV data. Also similar to 8 TeV trigger the *loose* criteria is applied on the shower shape.

Fiducial region of the EM calorimeter: same as the *Discovery Analysis* ($|\eta| < 2.37$, excluding $1.37 < |\eta| < 1.52$).

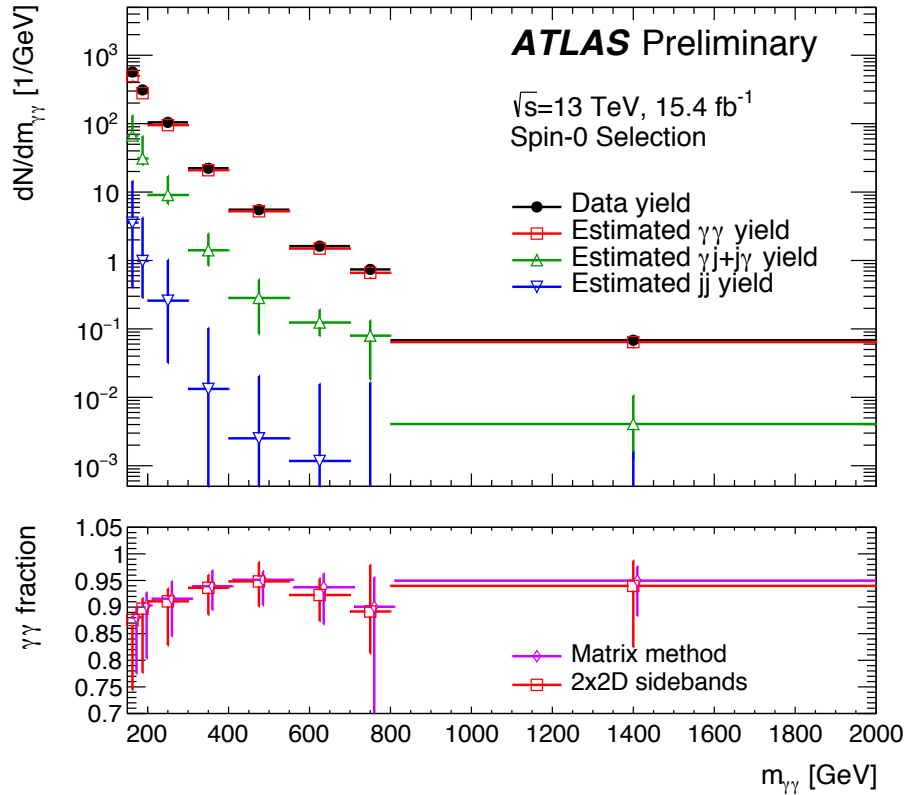
E_T requirements: the *Search Analysis* use similar mass dependent E_T requirements similar to the *Measurement Analyses*, but the coefficients have been re-optimized for high mass scalar signal to be $E_T/m_{\gamma\gamma} > 0.4$ and 0.3 for leading and sub-leading photon candidates, respectively.

Photon identification: *tight* identification criteria tuned for LHC Run 2 data taking condition are applied.

Isolation requirements: the isolation criteria in the *Search Analysis* are mass dependent. The calorimetric isolation requirement for photon candidates is $E_T^{\text{iso}} < 0.022E_T + 2.45 \text{ GeV}$. The track isolation requirement is $p_T^{\text{iso}} < 0.05E_T$.

Mass range: the diphoton invariant mass is required to be larger than 150 GeV.

In total 7765 events are selected from the 2015 data, and 28126 events are selected from the 2016 data with selections detailed above. The average purity of the $\gamma\gamma$ component is $(90^{+3}_{-10})\%$. The composition of the data sample is measured with more than two methods [43], with good agreement achieved between their results. The differential decomposition is shown in Figure 6.4.



(a)

Figure 6.4: Diphoton sample composition as a function of the invariant mass for the 13 TeV dataset used in the *Search Analysis*. The bottom panel shows the purity of diphoton events as determined from two independent methods (matrix and 2×2 D sidebands) with good agreement achieved. The total uncertainties including statistical and systematic components are shown by error bars.

Chapter 7

Signal and background modeling

The elegance of the diphoton analyses lies in the fact that both the signal and background invariant mass distributions can be modeled by analytic functions. This chapter is devoted to describing how the fit models for signal (Section 7.1) and background (Section 7.2) are determined.

7.1 Modeling of signal diphoton invariant mass shape and yield

7.1.1 Modeling of Standard Model Higgs boson decaying into two photons

Both *Discovery* and *Measurement Analyses* involve modeling of the SM Higgs boson signal. In this subsection the *Measurement Analyses* will be used as an example to demonstrate the procedure, which is essentially identical to the one used in the *Discovery Analyses* except for a few minor technical differences that have almost no effect on the results.

The signal modeling of the Higgs boson involves two aspects: modeling of the signal shape and modeling of the expected signal yields. The expected signal yield as a function of Higgs boson mass m_H can be well described by a third order polynomial for each production mode in each analysis category. The modeling of the signal diphoton invariant mass shape, on the other hand, is more sophisticated and will be detailed below.

As discussed in Chapter 2, the SM Higgs boson has only about 4 MeV natural width near 125 GeV. Therefore the diphoton invariant mass shape of the signal is expected to be dominated by detector resolution ($O(1 \text{ GeV})$), which can be well described by the sum of a Crystal Ball function [105] (f_{CB}) and a Gaussian (f_{GA}):

$$f_S(m_{\gamma\gamma}, \mu_{CB}, \sigma_{CB}, \alpha_{CB}, n_{CB}, \phi_{CB}, \mu_{GA}, \sigma_{GA}) \\ = \phi_{CB} f_{CB}(m_{\gamma\gamma}, \mu_{CB}, \sigma_{CB}, \alpha_{CB}, n_{CB}) + (1 - \phi_{CB}) f_{GA}(m_{\gamma\gamma}, \mu_{GA}, \sigma_{GA}). \quad (7.1)$$

Here μ_{CB} and σ_{CB} are the peak position and the width of the (Gaussian) core of the Crystal Ball function, respectively, and μ_{GA} and σ_{GA} are the peak position and the width of the Gaussian function, respectively. The power-law tail of f_{CB} is parameterized by α_{CB} and n_{CB} . The fraction of f_{CB} in the composite model is controlled by ϕ_{CB} . Typically, the bulk part of the signal shape is described by f_{CB} , while f_{GA} describes the outliers, therefore ϕ_{CB} is usually more close to unity than zero.

Except for n_{CB} , which is fixed to $n_{\text{CB}} = 10$, the rest of the parameters in f_{S} are parameterized as polynomials of m_H . The coefficients of the polynomials are determined by a simultaneous maximum likelihood fit to signal MC simulations at multiple mass points discussed in Section 4.2. MC sample at a specific mass point is weighted by the corresponding expected number of selected signal events in the fit. Since the signal $m_{\gamma\gamma}$ shape has been checked to be close enough among different Higgs boson production processes, only one signal model corresponding to the inclusive case will be derived and then applied to all production modes. This procedure is repeated for each analysis category to provide a categorized analytic signal model as a function of m_H for the analysis.

Figure 7.1 shows the signal parameterization overlaid with signal MC simulation at $m_H = 125$ GeV in one actual category of the *Coupling Analysis* to be discussed in Chapter 10. Good agreement between the parameterization and MC simulation is achieved.

7.1.2 Modeling of high mass scalar particle decaying into two photons

The *Search Analysis* is looking for not only narrow resonances (i.e. natural width of the particle is negligible compared with detector resolution), but also wide resonances with non-trivial natural widths. Hence the procedure discussed in the previous subsection needs to be expanded to consider both detector resolution and theoretical line-shape.

In the *Search Analysis*, the detector resolution is parameterized with double-sided Crystal Ball (DSCB) function, which is adopted in Run 2 diphoton analyses to replace the Crystal Ball plus

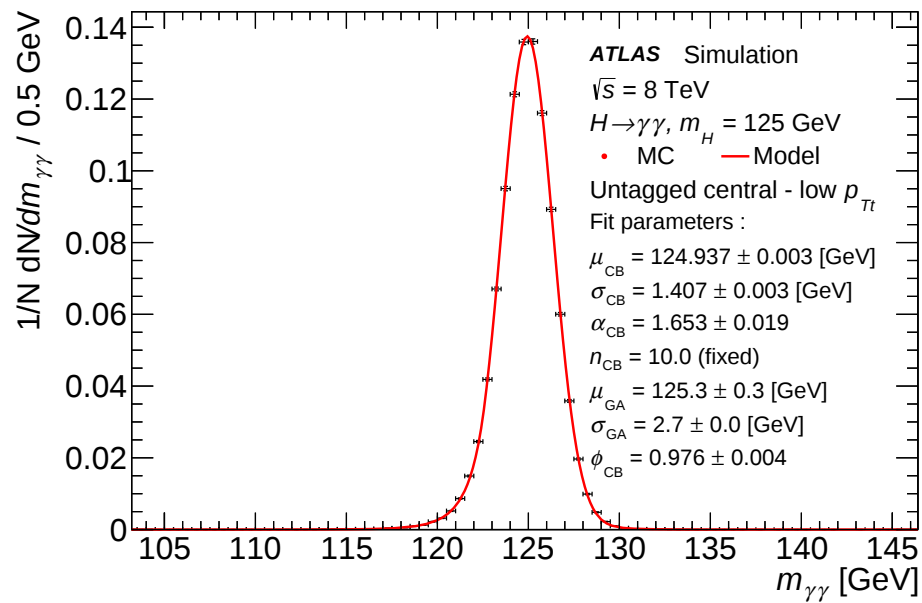


Figure 7.1: Simulated diphoton invariant mass distribution for Standard Model Higgs boson signal with $m_H = 125 \text{ GeV}$ in one category of the *Coupling Analysis* superimposed with parameterization determined from the procedure described in the text.

Gaussian prescription previously discussed. It is defined as:

$$N \cdot \begin{cases} e^{-\frac{1}{2}t^2} & \text{if } -\alpha_{\text{low}} \geq t \geq \alpha_{\text{high}} \\ e^{-\frac{1}{2}\alpha_{\text{low}}^2} \left[\frac{\alpha_{\text{low}}}{n_{\text{low}}} \left(\frac{n_{\text{low}}}{\alpha_{\text{low}}} - \alpha_{\text{low}} - t \right) \right]^{-n_{\text{low}}} & \text{if } t < -\alpha_{\text{low}} \\ e^{-\frac{1}{2}\alpha_{\text{high}}^2} \left[\frac{\alpha_{\text{high}}}{n_{\text{high}}} \left(\frac{n_{\text{high}}}{\alpha_{\text{high}}} - \alpha_{\text{high}} + t \right) \right]^{-n_{\text{high}}} & \text{if } t > \alpha_{\text{high}}, \end{cases} \quad (7.2)$$

where $t = (m_{\gamma\gamma} - \mu_{\text{CB}})/\sigma_{\text{CB}}$, N is a normalization parameter, μ_{CB} and σ_{CB} are again the peak position and the width of the Gaussian core, respectively. Unlike the Crystal Ball function, which only has power-law tail on one side, DSCB has power-law tails at both low-mass and high-mass sides, which are controlled by $\alpha_{\text{low/high}}$ and $n_{\text{low/high}}$. Very similar to the procedure discussed in last subsection, the parameters of the DSCB are determined as function of the particle mass m_X using narrow width ($\Gamma_X = 4$ MeV) simulations at different m_X points.

The theoretical line-shape of the signal is the multiplication of the following three constituents:

- the Breit–Wigner form used in MADGRAPH5_AMC@NLO [54];
- analytic parameterization of the gg parton luminosity provided by NNPDF3.0 NLO [70];
- the product of the matrix element of the effective field theory by the flux factor ($\propto m_{\gamma\gamma}^6$) and the phase space ($\propto m_{\gamma\gamma}$) [69].

The above prescription has been validated on the particle-level invariant mass distribution of the signal MC simulation, as shown in Figure 7.2.

The final signal parameterization used in the *Search Analysis* is the convolution of the detector resolution and theoretical line-shape. The good agreement between the parameterization and simulation at both narrow width and large width is shown in Figure 7.3.

As for the signal yield, it can be expressed as the product of three terms: the production cross section times branching ratio to two photons, the acceptance (A) of the kinematic requirements, and the reconstruction and identification efficiency (C). The acceptance A is defined as the fraction of events satisfying the fiducial acceptance at the generator level. The factor C is defined as the ratio of the number of events fulfilling all the selections placed on reconstructed quantities to the number of events in the fiducial acceptance. The fiducial acceptance in the *Search Analysis* closely follows the selection criteria applied to the reconstructed data as discussed in Chapter 6:

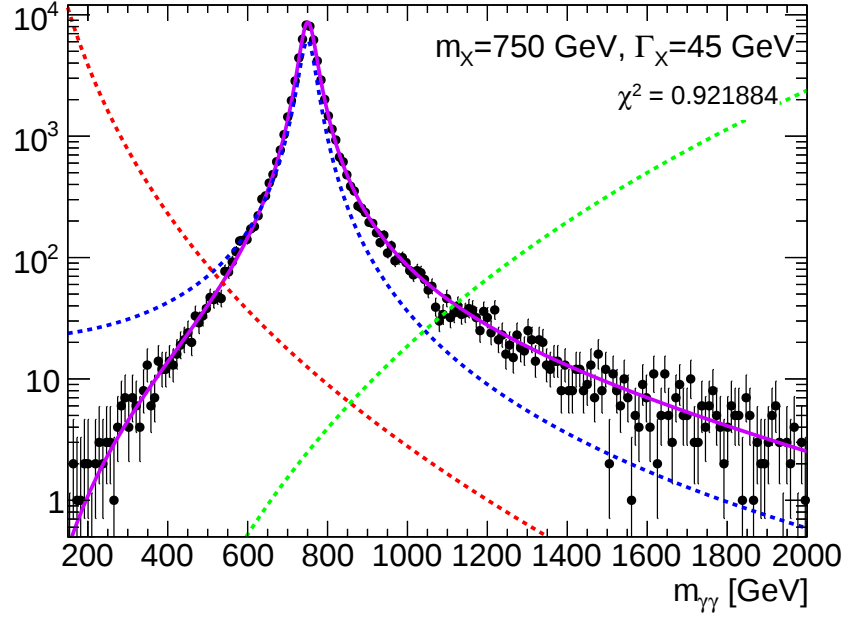


Figure 7.2: The true $m_{\gamma\gamma}$ distributions of the resonance with $m_X = 750$ GeV and width of 6% of the m_X value in the *Search Analysis*. The dashed red line is the gluon-gluon luminosity, the dashed green line is the functional form $m_{\gamma\gamma}^7$ (arising from the numerator of the squared matrix element, multiplied by the Jacobian factor of the variable transformation $\hat{s} \rightarrow m_{\gamma\gamma}$), and the dashed blue line is the Breit–Wigner distribution with the same mass and width as generated in the sample. The product of these three is represented by the solid purple line and agrees well with the true invariant mass distribution, shown as the black histogram. No selection cuts have been applied to the true photons.

$|\eta_\gamma| < 2.37$, $E_T > 0.4m_{\gamma\gamma}$ (leading γ), $E_T > 0.3m_{\gamma\gamma}$ (sub-leading γ). The isolation requirement $E_T^{\text{iso}} < 0.05E_T^\gamma + 6$ GeV is applied, where E_T^{iso} is computed using all particles with lifetime greater than 10 ps at the generator level in a cone of $\Delta R = 0.4$ around the photon direction. The value of the isolation requirement applied at the particle level is adjusted to reproduce the selection applied at the reconstruction level.

To reduce the dependence on the production mechanism, the limit results of the *Search Analysis* to be discussed in Chapter 12 are thus quoted in terms of the fiducial cross section, defined as the product of the cross section times the branching ratio to two photons within the fiducial acceptance. The simulation of a narrow-width signal produced by ggF mentioned in Chapter 4 is used to

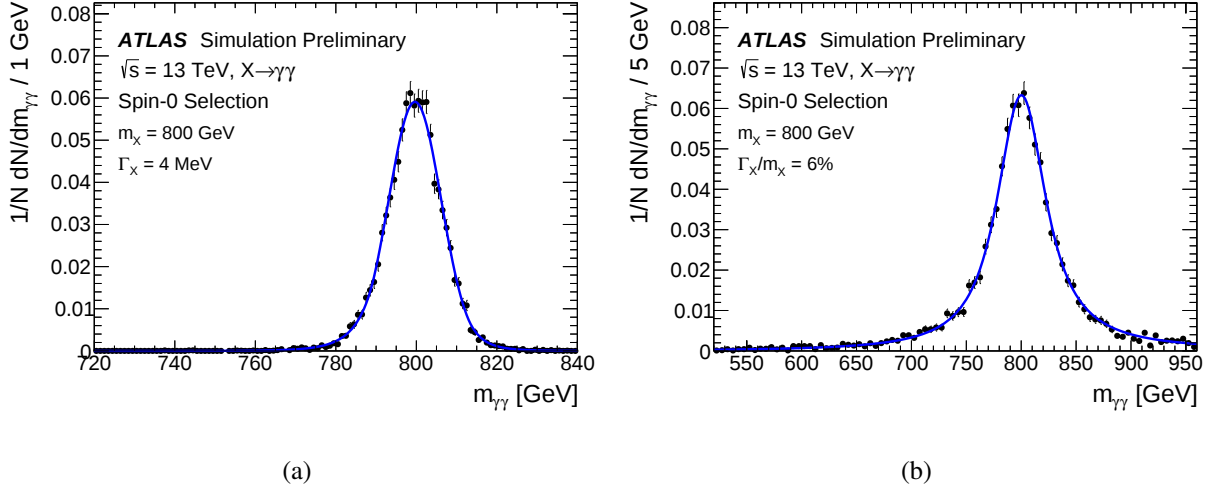


Figure 7.3: The $m_{\gamma\gamma}$ distributions for a scalar resonance in the *Search Analysis* with a mass of 800 GeV with (a) a narrow decay width ($\Gamma_X = 4$ MeV) or with (b) $\Gamma_X/m_X = 6\%$. The parameterization as the convolution of the theoretical mass line shape with the detector resolution is superimposed.

compute the nominal value of C , which ranges from 66% for a particle of mass 200 GeV to 74% at 700 GeV and is almost constant above 700 GeV.

7.2 Modeling of background diphoton invariant mass shape and normalization

Though it seems rather convenient to parameterize data sideband to provide background estimates, given typical signal over background ratio is only a few percent in the diphoton decay channel, it is critical to have a robust background modeling procedure that balances statistical power and potential modeling bias. Below a general description of the background modeling procedure used in this thesis will be discussed, assuming that there has been a background template (prepared from MC simulation or data control region) with reasonable description of data sideband as well as decent statistics ready for studies. Preparation of such background templates and background modeling results will be detailed for each analysis in coming chapters.

A *fit model* is the complete set of functional form, fit range, and possible additional constraints on the model parameters. As technical prerequisites, the background fit model should be stable and proved to be robust under the conditions of the data sample for which it is tested. In addition, it should not show evident systematic bias when fitting the background template and possibly also the data sideband. On the background template side, it should be validated without unexpected features, and should have decent enough statistics to discern potential bias from the fit model.

As the first step, the a signal component is added to the background fit model to form an extended signal plus background probability density function (pdf) ¹. This pdf will be fitted to the background template (usually normalized to data sideband), and the number of apparent signal events obtained from the fit, or the so-called *spurious signal*, will quantify the potential bias from the current choice of background fit model on the hypothesized signal tested. In the context of *Discovery and Measurement Analyses*, the absolute value of the spurious signal is required to be smaller than either of the following two criteria:

1. 10% of the expected number of signal events;
2. 20% of the uncertainty on the signal yield.

It is worth noting that as a desirable feature, the second criterion will become more stringent with increased luminosity L , because the spurious signal scales with L while this criterion only scales

¹ Extended pdf means that the expected sum of signal and background events can fluctuate around the normalization of the sample being fitted.

with $\sqrt{L}$. In the *Search Analysis*, since there is no well-defined signal expectation, only the second criterion is kept and loosened from 20% to 30% of the uncertainty on the signal yield due to known caveats of the background templates used.

The spurious signal criteria are still easy to circumvent: one could simply add more and more degrees of freedom to the functional form until it can pass. By doing so, however, the sensitivity of the analysis is spoiled, and the functional form with too many degrees of freedom may over-fit the data by picking up fluctuations. To avoid such a situation, for all the background fit models passing the criteria, only the one yields best sensitivity should be selected.

In the search of a potential new particle (e.g. in the *Discovery* and *Search Analyses*), the spurious criteria should be satisfied in all phase space to be explored. In the measurement of particle properties (e.g. in the *Measurement Analysis*), the spurious signal also needs to be evaluated in a reasonable span around the signal region to ensure the quality of background model. When there are multiple background templates available in the test, the most conservative spurious signal among all should be used.

The normalization of background is always a free parameter which will be determined when fitting the extended signal plus background pdf to the data.

Chapter 8

Statistical procedure

8.1 Likelihood construction for diphoton analyses

The extended likelihood function for a given category c of a diphoton analysis discussed in this thesis can be written as follows:

$$\mathcal{L}_c = \text{Pois}(n_c | N_c(\boldsymbol{\alpha}, \boldsymbol{\theta})) \cdot \prod_{i=1}^{n_c} f_c(m_{\gamma\gamma}^i, \boldsymbol{\alpha}, \boldsymbol{\theta}) \cdot G(\boldsymbol{\theta}'), \quad (8.1)$$

where n_c is the observed number of events, N_c is the expected number of events, $f_c(m_{\gamma\gamma}^i)$ is the value of the pdf of the $m_{\gamma\gamma}$ distribution evaluated for each event i , $\boldsymbol{\alpha}$ are *parameters of interest* (parameters to be studied), and $\boldsymbol{\theta}$ are *nuisance parameters* (other parameters which can be adjusted in the maximum likelihood fit). Those nuisance parameters corresponding to the implementation of systematic uncertainties $\boldsymbol{\theta}'$ are constrained by a set of pdfs $G(\boldsymbol{\theta}')$ representing the knowledge from auxiliary measurements (e.g. measurement of integrated luminosity) together with the response functions to be discussed later. The mean values of the constraint pdfs in $G(\boldsymbol{\theta}')$, which represent the assumed central values of auxiliary measurements (usually scaled to 0), are called *global observables*. The exact scope of parameters of interest and nuisance parameters need to be determined on a case-by-case basis. For instance, in the measurement of Higgs boson couplings (Chapter 10) the signal strength parameter μ (to be defined later) is the parameter of interest, whereas in the measurement of Higgs boson mass (Chapter 11) m_H is the parameter of interest instead, and μ now becomes a nuisance parameter which is floating in the fit to reduce model dependency.

The number of expected events is the sum of the hypothesized number of signal events, $N_{\text{sig},c}^{\text{hypo}}$, plus the fitted number of background candidates, $N_{\text{bkg},c}$, and the fitted spurious signal, $N_{\text{spur},c}$.

$\theta_{\text{spur},c}$,

$$N_c = N_{\text{sig},c}^{\text{hypo}} + N_{\text{bkg},c} + N_{\text{spur},c} \cdot \theta_{\text{spur},c}. \quad (8.2)$$

In the *Discovery* and *Measurement Analyses*, the SM prediction of the Higgs boson as a function of m_H is well defined, hence the hypothesized number of signal events, $N_{\text{sig},c}^{\text{hypo}}$, can be further expanded as

$$N_{\text{sig},c}^{\text{hypo}} = \mu \cdot N_{\text{S},c}(\theta_c^{\text{yield}}, \theta_c^{\text{migr}}, m_H) \quad (8.3)$$

$$= \sum_p \mu_p N_{p,c}(\theta_{p,c}^{\text{yield}}, \theta_{p,c}^{\text{migr}}, m_H), \quad (8.4)$$

where μ is the combined *signal strength* parameter, $N_{\text{S},c}(\theta_c^{\text{yield}}, \theta_c^{\text{migr}}, m_H)$ is the number of signal events predicted by the SM from all production processes, θ_c^{yield} and θ_c^{migr} are the nuisance parameters that implement the systematic uncertainties affecting the yields of the Higgs boson production in the current category and migration from other categories, respectively. $N_{\text{sig},c}^{\text{hypo}}$ can be further decomposed to the sum of signal strength times expected number of signal events for each production process p (ggF , VBF ...) as shown in Equation (8.4).

In the *Search Analysis* the theoretical prediction is less sharp. In particular, the production cross section of the new physics is essentially unknown a priori. Hence instead of introducing signal strength, as discussed in Section 7.1.2 the fiducial production cross section times branching ratio $(\sigma \times BR)^{\text{fid.}}$ is used as parameter of interest:

$$N_{\text{sig},c}^{\text{hypo}} = L(\theta_{\text{lumi}}) \cdot C(\theta_c^{\text{yield}}, m_X) \cdot (\sigma \times BR)^{\text{fid.}}, \quad (8.5)$$

where L is the integrated luminosity, θ_{lumi} is the nuisance parameter for uncertainty on luminosity measurement, C is the correction factor derived for the fiducial acceptance defined in Section 7.1.2. Due to the way the dataset is categorized (to be detailed in Chapter 12), only uncertainties on signal yields are relevant in the *Search Analysis*.

The signal and background pdfs in each category are introduced as follows:

$$f_c(m_{\gamma\gamma}^i) = [(N_{\text{sig},c}^{\text{hypo}} + N_{\text{spur},c} \cdot \theta_{\text{spur},c}) \cdot f_{\text{sig},c}(m_{\gamma\gamma}^i, \theta_{\text{sig},c}^{\text{shape}}) + N_{\text{bkg},c} \cdot f_{\text{bkg},c}(m_{\gamma\gamma}^i, \theta_{\text{bkg},c}^{\text{shape}})] / N_c,$$

where $\theta_{\text{S},c}^{\text{shape}}$ and $\theta_{\text{bkg},c}^{\text{shape}}$ are nuisance parameters affecting the shapes of the invariant mass distributions of the signal $f_{\text{S},c}$ and background $f_{\text{bkg},c}$, respectively.

Apart from the spurious signal, systematic uncertainties are incorporated into the likelihood by multiplying the relevant parameter of the statistical model by a response function. In the case of a Gaussian pdf for the effect of an uncertainty of size σ , it can be written as

$$F_G(\sigma, \theta) = (1 + \sigma \cdot \theta), \quad (8.6)$$

and for cases where a negative model parameter does not make physical sense, the log-normal pdf is used instead:

$$F_{LN}(\sigma, \theta) = e^{\ln(1+\sigma \cdot \theta)}. \quad (8.7)$$

In both cases the corresponding component of the constraint product $G(\theta)$ is a unit Gaussian centered at zero for θ . The systematic uncertainties affecting the yield (such as luminosity, photon identification etc.) and mass resolution use the log-normal form while a Gaussian form is used for the rest.

The combined likelihood function is essentially the multiplication of the likelihood of individual categories (in the context of categorized analysis) or channels (in the context of statistical combination). When two systematic uncertainties are considered fully correlated, they share the same nuisance parameter and also the same constraint pdf in the combined likelihood, although the magnitude and sign of the uncertainties are still allowed to differ. Systematic uncertainties (e.g. luminosity uncertainties from ATLAS and CMS experiments) with partial correlations are usually first decomposed into their uncorrelated and fully correlated components before being assigned to nuisance parameters.

8.2 Statistical tests

Profiled likelihood ratio is the foundation of statistical interpretations of physics results presented in this thesis. Take the combined signal strength μ as parameter of interest (the discussion can be easily extended to other cases), the profile likelihood ratio can be constructed as

$$\Lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\boldsymbol{\theta}}(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})}, \quad (8.8)$$

where $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the values of the combined signal strength and nuisance parameters that unconditionally maximize the likelihood, while $\hat{\boldsymbol{\theta}}(\mu)$ are the values of the nuisance parameters that

maximize the likelihood on the condition that μ is fixed to a given value. In the following subsections the constructions of test statistics for different purposes are detailed. For all these test statistics, the analytic asymptotic approximations [106] are valid for producing results reported in this thesis and will also be provided.

8.2.1 Test statistic for discovery of a positive signal

To establish the discovery of a *positive signal*, the test statistic q_0 is used:

$$q_0 = \begin{cases} -2 \ln \Lambda(0) & \text{if } \hat{\mu} > 0 \\ 0 & \text{otherwise.} \end{cases} \quad (8.9)$$

In the definition above a deficit in the data is considered compatible with background-only hypothesis (i.e. $\mu = 0$). The asymptotic sampling distribution $f(q_0|0)$ is half delta function plus half chi-squared function with unit degree of freedom:

$$f(q_0|0) = \frac{1}{2} \delta(q_0) + \frac{1}{2} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{q_0}} e^{-q_0/2}. \quad (8.10)$$

The cumulative distribution is found to be simply:

$$F(q_0|0) = \Phi(\sqrt{q_0}). \quad (8.11)$$

where Φ is the cumulative distribution of the unit Gaussian. The p -value of the $\mu = 0$ hypothesis therefore can be written as:

$$p_0 = 1 - F(q_0|0). \quad (8.12)$$

8.2.2 Test statistic for upper limits

To establish an upper limit on the signal strength parameter μ of a potential *positive signal*, the definition of $\Lambda(\mu)$ can be slightly modified, such that a deficit in the data is considered compatible with background-only hypothesis:

$$\tilde{\Lambda}(\mu) = \begin{cases} \Lambda(\mu) & \text{if } \hat{\mu} > 0 \\ \frac{\mathcal{L}(\mu, \hat{\theta}(\mu))}{\mathcal{L}(0, \hat{\theta}(0))} & \text{otherwise.} \end{cases} \quad (8.13)$$

The test statistic used for limit setting, $\tilde{q}_\mu$, is defined as:

$$\tilde{q}_\mu = \begin{cases} -2 \ln \tilde{\Lambda}(\mu) & \text{if } \hat{\mu} \leq \mu \\ 0 & \text{otherwise.} \end{cases} \quad (8.14)$$

The reason for setting $\tilde{q}_\mu = 0$ for $\hat{\mu} > \mu$ is that when setting an upper limit, one would not regard data with $\hat{\mu} > \mu$ as representing less compatibility with μ , and therefore this is not taken as part of the rejection region of the test.

As with the case of discovery, the p -value quantifying the level of agreement between the data and hypothesized μ is defined as

$$p_\mu = \int_{\tilde{q}_{\mu, \text{obs}}}^{\infty} f(\tilde{q}_\mu | \mu) d\tilde{q}_\mu \quad (8.15)$$

The asymptotic sampling distribution of is found to be:

$$f(\tilde{q}_\mu | \mu') = \Phi\left(\frac{\mu' - \mu}{\sigma}\right) \delta(\tilde{q}_\mu) + \begin{cases} \frac{1}{2} \frac{1}{\sqrt{2\pi}} \frac{1}{\tilde{q}_\mu} \exp\left[-\frac{1}{2} \left(\sqrt{\tilde{q}_\mu} - \frac{\mu - \mu'}{\sigma}\right)^2\right] & \text{if } 0 < \tilde{q}_\mu \leq \mu^2 / \sigma^2 \\ \frac{1}{\sqrt{2\pi} (2\mu/\sigma)} \exp\left[-\frac{1}{2} \frac{(\tilde{q}_\mu - (\mu^2 - 2\mu\mu')/\sigma^2)^2}{(2\mu/\sigma)^2}\right] & \text{otherwise.} \end{cases} \quad (8.16)$$

The corresponding cumulative distribution is

$$F(\tilde{q}_\mu | \mu') = \begin{cases} \Phi\left(\sqrt{\tilde{q}_\mu} - \frac{\mu - \mu'}{\sigma}\right) & \text{if } 0 < \tilde{q}_\mu < \mu^2 / \sigma^2 \\ \Phi\left(\frac{\tilde{q}_\mu - (\mu^2 - 2\mu\mu')/\sigma^2}{2\mu/\sigma}\right) & \text{otherwise,} \end{cases} \quad (8.17)$$

and the p -value of the hypothesized μ is

$$p_\mu = 1 - F(\tilde{q}_\mu | \mu). \quad (8.18)$$

The upper limits presented in this thesis are calculated using the CL_s procedure [107]. The quantity CL_s is defined as

$$CL_s(\mu) = \frac{p_\mu}{1 - p_b}, \quad (8.19)$$

where p_b is the p -value corresponding to the background-only hypothesis

$$p_b = F(\tilde{q}_\mu | 0). \quad (8.20)$$

The CL_s upper limit on μ (μ_{up}) is obtained by solving for $CL_s(\mu_{\text{up}}) = 5\%$. The CL_s procedure is used to ensure a signal hypothesis is not excluded by the experiment when there is little sensitivity.

As for the bands around the expected limit, a more advanced procedure compared with the one presented in Reference [106] has been used to improve the agreement between asymptotic approximation and pseudo-experiments. Readers can refer to Reference [108] for more technical details.

8.2.3 Test statistic for measurements and compatibility tests

Many measurements discussed in this thesis only involve a single parameter of interest. The test statistic used in such cases (t_μ) is defined as follows:

$$t_\mu = -2 \ln \Lambda(\mu). \quad (8.21)$$

The asymptotic sampling distribution t_μ follows is simply the chi-squared distribution with one degree of freedom:

$$f(t_\mu|\mu) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{t_\mu}} e^{-t_\mu/2}, \quad (8.22)$$

and the corresponding cumulative distribution is

$$F(t_\mu|\mu) = 2\Phi(\sqrt{t_\mu}) - 1. \quad (8.23)$$

The total uncertainty ${}^{+\delta\mu}_{-\delta\mu}$ at the 68% confidence level (CL) of measured signal strength μ with best fit value $\hat{\mu}$ is estimated by finding the points which make $t_{\hat{\mu}+\delta\mu} = t_{\hat{\mu}-\delta\mu} = 1$. The calculation is usually done through scanning twice of the negative log-likelihood as a function of parameter of interest as to be shown in coming chapters. The statistical component of the total uncertainty is estimated by fixing all the nuisance parameters associated with systematic uncertainties (θ') to their values from the unconditional maximum likelihood fit, and finding the new points which set $t_{\mu+\delta\mu} = 1$. The total systematic uncertainty is given by the quadratic difference between the total and statistical uncertainties. A more detailed split of the systematic uncertainty can be carried out similarly by fixing corresponding nuisance parameters and then performing quadratic subtractions.

When more than one parameter of interest is evaluated at the same time, the construction of the test statistic can be generalized to:

$$T(\alpha) = -2 \ln \frac{\mathcal{L}(\alpha, \hat{\theta}(\alpha))}{\mathcal{L}(\hat{\alpha}, \hat{\theta})}. \quad (8.24)$$

The asymptotic sampling distribution $T(\boldsymbol{\alpha})$ follows is a chi-squared distribution with n degrees of freedom, n being the size of $\boldsymbol{\alpha}$.

The technical implementations of the statistical procedure is based on the ROOFIT [109], ROOSTATS [110], and HISTFACTORY [111] data modeling and handling packages.

Chapter 9

Discovery of Higgs boson in diphoton decay channel

This chapter details the observation of the Higgs boson in the diphoton decay channel by the ATLAS experiment [12] (*Discovery Analysis*). The focus will be on event categorization, systematic uncertainties (including background modeling study), and statistical results (including combination with other decay channels). Discussions on other aspects of the analysis have been covered previous chapters (same for other analyses to be reported).

9.1 Event categorization

To increase the sensitivity to a SM Higgs boson signal, in the *Discovery Analysis* the events are classified into ten mutually exclusive categories with different mass resolutions and signal-to-background ratios.

First, an exclusive *Two-jet* category concentrating the VBF events is defined. It requires at least two jets with $|\eta| < 4.5$ and $p_T > 25$ GeV. In the analysis of the 8 TeV data, the p_T threshold is raised to 30 GeV for jets with $2.5 < |\eta| < 4.5$. For jets in the ID acceptance ($|\eta| < 2.5$) JVF (defined in Section 5.2.2) is required to be at least 0.75. Motivated by the VBF topology, three additional selections are applied:

1. the difference of the pseudorapidity between the leading and sub-leading jets (ranked by p_T) is required to be larger than 2.8;
2. the invariant mass of these two jets needs to be larger than 400 GeV;
3. the azimuthal angle difference between the diphoton system and the system of these two jets has to be larger than 2.6.

With above selections, about 70% of the signal events in the category is expected to be from VBF process.

The rest of the events are divided into nine categories using photon pseudorapidity, photon conversion status and diphoton p_{Tt} ¹, which is defined as:

$$p_{Tt} = |(\mathbf{p}_T^{\gamma_1} + \mathbf{p}_T^{\gamma_2}) \times (\mathbf{p}_T^{\gamma_1} - \mathbf{p}_T^{\gamma_2})| / |\mathbf{p}_T^{\gamma_1} - \mathbf{p}_T^{\gamma_2}|. \quad (9.1)$$

Here $\mathbf{p}_T^{\gamma_1}$ and $\mathbf{p}_T^{\gamma_2}$ are the transverse momenta of the two photons. Events with both photons unconverted are separated into *Unconverted central* ($|\eta| < 0.75$ for both candidates) and *Unconverted rest* (all other events) categories. Events with at least one converted photon are separated into *Converted central* ($|\eta| < 0.75$ for both candidates), *Converted transition* (at least one photon with $1.3 < |\eta| < 1.75$) and *Converted rest* (all other events) categories. Except for the *Converted transition* category, each category is further divided into *low* p_{Tt} ($p_{Tt} < 60$ GeV) and *high* p_{Tt} (otherwise) categories to differentiate signal over background ratio between the two, because the Higgs boson signals, in particular those from non- ggF processes, have more boosted p_{Tt} distributions compared with background processes as shown in Figure 9.1.

The same categorization is applied to the 7 TeV data and the 8 TeV data. The number of data events in each category, as well as the sum of all the categories are given in Table 9.1. The signal resolution difference between the best-resolution category (*Unconverted central high* p_{Tt}) and a category with worse resolution (*Converted rest low* p_{Tt}) is shown in Figure 9.2.

¹ p_{Tt} is preferred over the p_T of diphoton system because it has better resolution and is less correlated with the diphoton invariant mass.

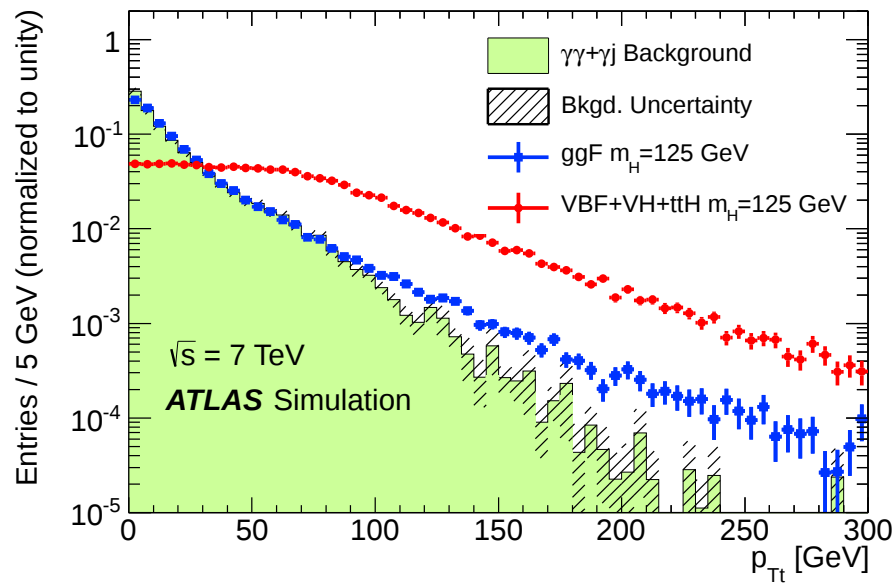


Figure 9.1: Distribution of p_{Tt} in simulated events with Higgs boson productions and in background events. The signal distribution is shown separately for ggF (blue), and VBF together with associated productions (red). The background distribution and the two signal distributions are normalized to unit area.

$\sqrt{s}$	7 TeV		8 TeV		
Category	N_D	N_S	N_D	N_S	FWHM [GeV]
Unconverted central low p_{Tt}	2054	10.5	2945	14.2	3.4
Unconverted central high p_{Tt}	97	1.5	173	2.5	3.2
Unconverted rest low p_{Tt}	7129	21.6	12136	30.9	3.7
Unconverted rest high p_{Tt}	444	2.8	785	5.2	3.6
Converted central low p_{Tt}	1493	6.7	2015	8.9	3.9
Converted central high p_{Tt}	77	1.0	113	1.6	3.5
Converted rest low p_{Tt}	8313	21.1	11099	26.9	4.5
Converted rest high p_{Tt}	501	2.7	706	4.5	3.9
Converted transition	3591	9.5	5140	12.8	6.1
Two-jet	89	2.2	139	3.0	3.7
All categories (inclusive)	23788	79.6	35251	110.5	3.9

Table 9.1: Number of events in the data (N_D) and expected number of signal events (N_S) with $m_H = 126.5$ GeV for each category of the *Discovery Analysis* and total for the 7 TeV and 8 TeV datasets in the mass range 100 – 160 GeV. The mass resolution quantified by full width at half maximum (FWHM) is also given for the 8 TeV data.

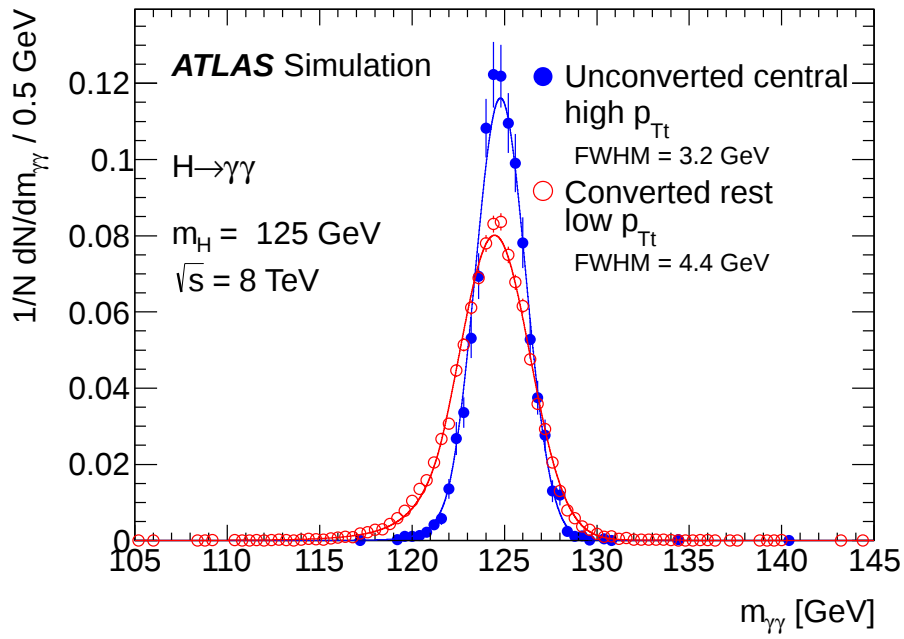


Figure 9.2: Invariant mass distributions for a Higgs boson with $m_H = 125 \text{ GeV}$, for the best-resolution category (Unconverted central high p_{Tt}) of the *Discovery Analysis* shown in blue, and for a category with lower resolution (Converted rest low p_{Tt}) shown in red, for the $\sqrt{s} = 8 \text{ TeV}$ simulation. The invariant mass distribution is parametrized by the sum of a Crystal Ball function and a broad Gaussian based on the procedure discussed in Chapter 7.1.1.

9.2 Systematic uncertainties

The systematic uncertainties included in the *Discovery Analysis* are summarized in Table 9.2. The dominant experimental uncertainty on the signal yield (8% for 7 TeV, 11% for 8 TeV) comes from the photon reconstruction and identification efficiency, which is estimated with data using electrons from Z decays and photons from $Z \rightarrow \ell^+ \ell^- \gamma$ events. Pileup modelling also affects the expected yields and contributes to the uncertainty (4%). Further uncertainties on the integrated signal yield are related to the trigger (1%), photon isolation (about 0.5%) and luminosity (1.8% for 7 TeV, 3.6% for 8 TeV) [112]. Uncertainties on the predicted cross sections and branching ratio have been discussed in Chapter 2. The uncertainty values used are different here due to different hypothesized Higgs boson mass (125 GeV) and also slightly obsolete calculation. All the uncertainties on the signal event yield have no impact on the evaluation of excess significance to be discussed in Section 9.3. They will only affect the measurement of overall signal rate and the limit setting.

As for the migration of signal events between categories, the uncertainty on the knowledge of the material in front of the calorimeter on the migration between converted and unconverted categories is 4%. The uncertainty from pileup on the population of the converted and unconverted categories is 2%. The uncertainty from the jet energy scale (JES) amounts to up to 19% for the Two-jet category, and up to 4% for the other categories. Uncertainty from the JVF modelling is 12% (for the 8 TeV data) for the Two-jet category, estimated from $Z+2$ -jets events by comparing data and MC. Uncertainties due to the modelling of the underlying event are 6% for VBF and 30% for other production processes in the Two-jet category. Different PDFs and scale variations in the HQT [15] calculations are used to derive possible event migration among categories (9%) due to the modelling of the Higgs boson kinematics. The theoretical uncertainty associated with the exclusive Higgs boson production process with additional jets is estimated based on the Stewart–Tackmann method [15, 113], with the noticeable difference that an explicit calculation of the ggF process at NLO using MCFM [65] in the Two-jet category reduces the uncertainty on this non-negligible contribution to 25%.

The total uncertainty on the mass resolution is 14%. The dominant contribution (12%) comes from the uncertainty on the energy resolution of the calorimeter, which is determined from $Z \rightarrow e^+ e^-$

events. Smaller contributions come from the imperfect knowledge of the material in front of the calorimeter, which affects the extrapolation of the calibration from electrons to photons (6%), and from pileup (4%). As for the mass position of the signal, the uncertainty from photon energy scale is estimated to be 0.6%.

Background fit models and associated potential biases are studied with the procedure discussed in Section 7.2 using three different sets of high statistics background templates which only differ in $\gamma\gamma$ component as discussed in Section 4.4. In addition, the Drell-Yan background component is taken into account from data-driven approach. These components are mixed according to the proportions estimated from data (Section 6.1), and the overall normalization is fixed to the number of data events per category.

A variety of functional forms are considered for the background parametrization: single and double exponential functions, Bernstein polynomials [114] up to seventh order, exponentials of second and third-order polynomials, and exponentials with modified turn-on behavior. The fit range of the background model is fixed to be between 100 GeV and 160 GeV. The spurious signal is evaluated between 110 GeV and 150 GeV, which is the same as the search range. For categories with low statistics, an exponential function is found to have sufficiently small bias, while polynomials and exponentials of polynomials, respectively, are needed for limiting the potential bias to stay within the predefined requirements for the higher-statistics categories.

For the chosen parametrization, the largest absolute spurious signal over the full mass range studied (from 110 GeV to 150 GeV) is used as the systematic uncertainty from background model. The selected parametrizations along with their systematic uncertainties are shown in Table 9.3.

The invariant mass distributions and their composition obtained from the high-statistics simulation model based on SHERPA for the diphoton component have been cross checked against data in different categories, using the same background decomposition method as used for the inclusive sample. Within the statistical uncertainties of data, a good agreement is found for the shapes of the invariant mass distributions.

Systematic uncertainties			$\sqrt{s} = 7 \text{ TeV} \text{ [%]}$	$\sqrt{s} = 8 \text{ TeV} \text{ [%]}$
Integrated signal yield				
Experiment	Photon identification		± 8.4	± 10.8
	Effect of pileup on photon rec/ID			± 4
	Luminosity		± 1.8	± 3.6
	Trigger			± 1
	Photon energy scale			± 0.3
	Photon Isolation		± 0.4	± 0.5
Theory	Scales		$ggF: {}^{+12}_{-8}, \text{VBF: } \pm 0.3,$ $WH: {}^{+0.2}_{-0.8}, ZH: {}^{+1.4}_{-1.6}, t\bar{t}H: {}^{+3}_{-9}$	$ggF: {}^{+7}_{-8}, \text{VBF: } \pm 0.2,$ $WH: {}^{+0.2}_{-0.6}, ZH: {}^{+1.6}_{-1.5}, t\bar{t}H: {}^{+4}_{-9}$
	PDF+ α_S		$ggF: {}^{+8}_{-7}, \text{VBF: } {}^{+2.5}_{-2.1},$ $VH: \pm 3.5, t\bar{t}H: \pm 9$	$ggF: {}^{+8}_{-7}, \text{VBF: } {}^{+2.6}_{-2.8},$ $VH: \pm 3.5, t\bar{t}H: \pm 8$
	Branching ratio			± 5
Signal category migration				
Experiment	Material		Unconv: ± 4 , Conv: ∓ 3.5	
	Effect of pileup on photon rec/ID		Unconv: ± 3 , Conv: ∓ 2 ,	Unconv: ± 2 , Conv: ∓ 2 ,
			Two-jet: ± 2	Two-jet: ± 12
	Jet energy scale	low p_{Tt}	$ggF: \pm 0.1, \text{VBF: } \pm 2.6,$	$ggF: \pm 0.1, \text{VBF: } \pm 2.3,$
			Others: ± 0.1	Others: ± 0.1
		high p_{Tt}	$ggF: \pm 0.1, \text{VBF: } \pm 4,$	$ggF: \pm 0.1, \text{VBF: } \pm 4,$
			Others: ± 0.1	Others: ± 0.1
		Two-jet	$ggF: \mp 19, \text{VBF: } \mp 8,$	$ggF: \mp 18, \text{VBF: } \mp 9,$
			Others: ∓ 15	Others: ∓ 13
		Jet-vertex-fraction		Two-jet: ± 13 , Others: ∓ 0.3
Theory	Primary vertex selection		negligible	
	Underlying Event (Two-jet)		VBF: ± 6 , Others: ± 30	
	Two-jet bin		$gg \rightarrow H + 2 \text{ jets: } \pm 25$	
	Higgs boson p_T modeling		low p_{Tt} : ± 1.1 , high p_{Tt} : ∓ 12.5 , Two-jet: ∓ 9	
Signal mass resolution (Exp.)				
Calorimeter energy resolution			± 12	
Electron to photon extrapolation			± 6	
Effect of pileup on energy resolution			± 4	
Primary vertex selection			negligible	
Signal mass position (Exp.)				
Photon energy scale			± 0.6	

Table 9.2: Summary of systematic uncertainties on the expected signal considered in the *Discovery Analysis*. The values listed in the table are the relative uncertainties (in %) on given quantities from the various sources investigated for a Higgs boson mass of 125 GeV. The sign in the front of values for each systematic uncertainty indicates correlations among categories and processes. Experimental and theoretical uncertainties are separately marked.

Category	Model	N_{spur}	
		$\sqrt{s} = 7 \text{ TeV}$	$\sqrt{s} = 8 \text{ TeV}$
Inclusive	4th order pol.	7.3	10.6
Unconverted central low p_{Tt}	Exp. of 2nd order pol.	2.1	3.0
Unconverted central high p_{Tt}	Exponential	0.2	0.3
Unconverted rest low p_{Tt}	4th order pol.	2.2	3.3
Unconverted rest high p_{Tt}	Exponential	0.5	0.8
Converted central low p_{Tt}	Exp. of 2nd order pol.	1.6	2.3
Converted central high p_{Tt}	Exponential	0.3	0.4
Converted rest low p_{Tt}	4th order pol.	4.6	6.8
Converted rest high p_{Tt}	Exponential	0.5	0.7
Converted transition	Exp. of 2nd order pol.	3.2	4.6
Two-jet	Exponential	0.4	0.6

Table 9.3: List of the functions chosen to model the background distributions of $m_{\gamma\gamma}$ in the *Discovery Analysis*, and the associated systematic uncertainties on the signal amplitudes in terms of spurious signal (N_{spur}) for the ten categories and the 7 TeV and 8 TeV datasets.

9.3 Results

9.3.1 Diphoton invariant mass spectra

The inclusive $m_{\gamma\gamma}$ spectrum is shown in Figure 9.3. The result of a fit including a signal component fixed to $m_H = 126.5$ GeV and a background component described by a fourth-order Bernstein polynomial is superimposed. The $m_{\gamma\gamma}$ spectra correspond to the $\sqrt{s} = 7$ TeV ($\sqrt{s} = 8$ TeV) data are shown in Figure 9.4 (9.5) and Figure 9.6 (9.7) for unconverted and converted categories, respectively, except for the Converted transition category which is shown together with the Two-jet category in Figure 9.8

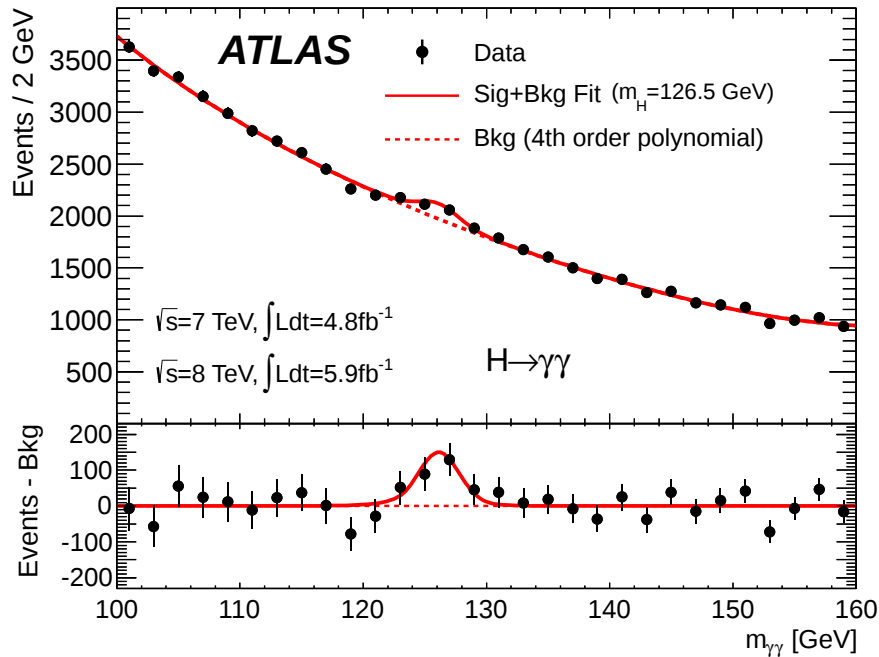


Figure 9.3: Distribution of the invariant mass of diphoton candidates after all selections in the *Discovery Analysis* for the combined 7 TeV and 8 TeV data sample. The result of a fit to the data of the sum of a signal component fixed to $m_H = 126.5$ GeV and a background component described by a fourth-order Bernstein polynomial is superimposed. The bottom panel shows the data relative to the background component of the fitted model.

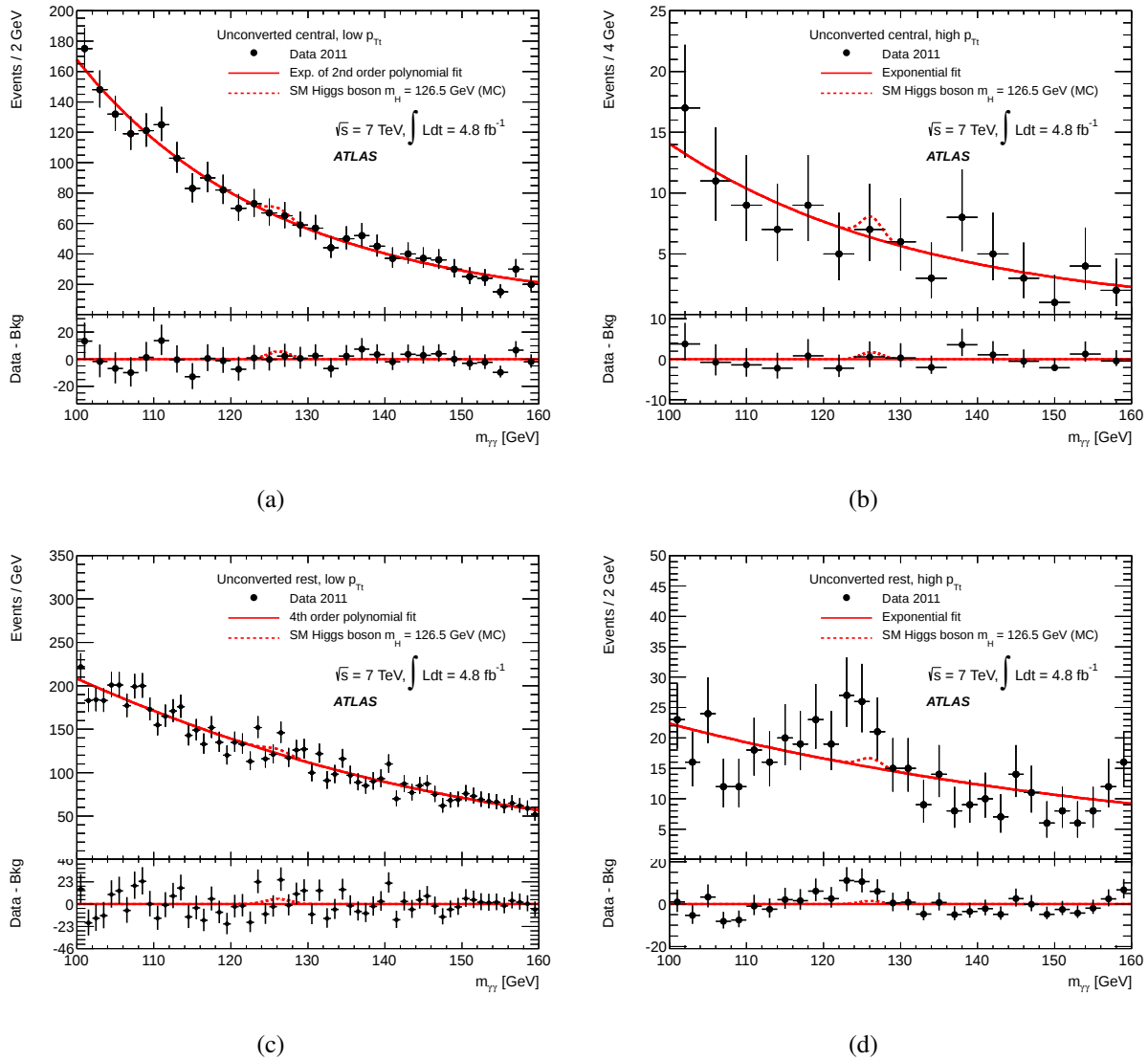


Figure 9.4: Background-only fits to the diphoton invariant mass spectra for (a) Unconverted central low p_{Tt} , (b) Unconverted central high p_{Tt} , (c) Unconverted rest low p_{Tt} , and (d) Unconverted rest high p_{Tt} categories of the *Discovery Analysis* correspond to the 7 TeV data sample. The bottom panel displays the residual of the data with respect to the background fit. The Higgs boson expectation for a mass hypothesis of 126.5 GeV corresponding to the Standard Model cross section is also shown.

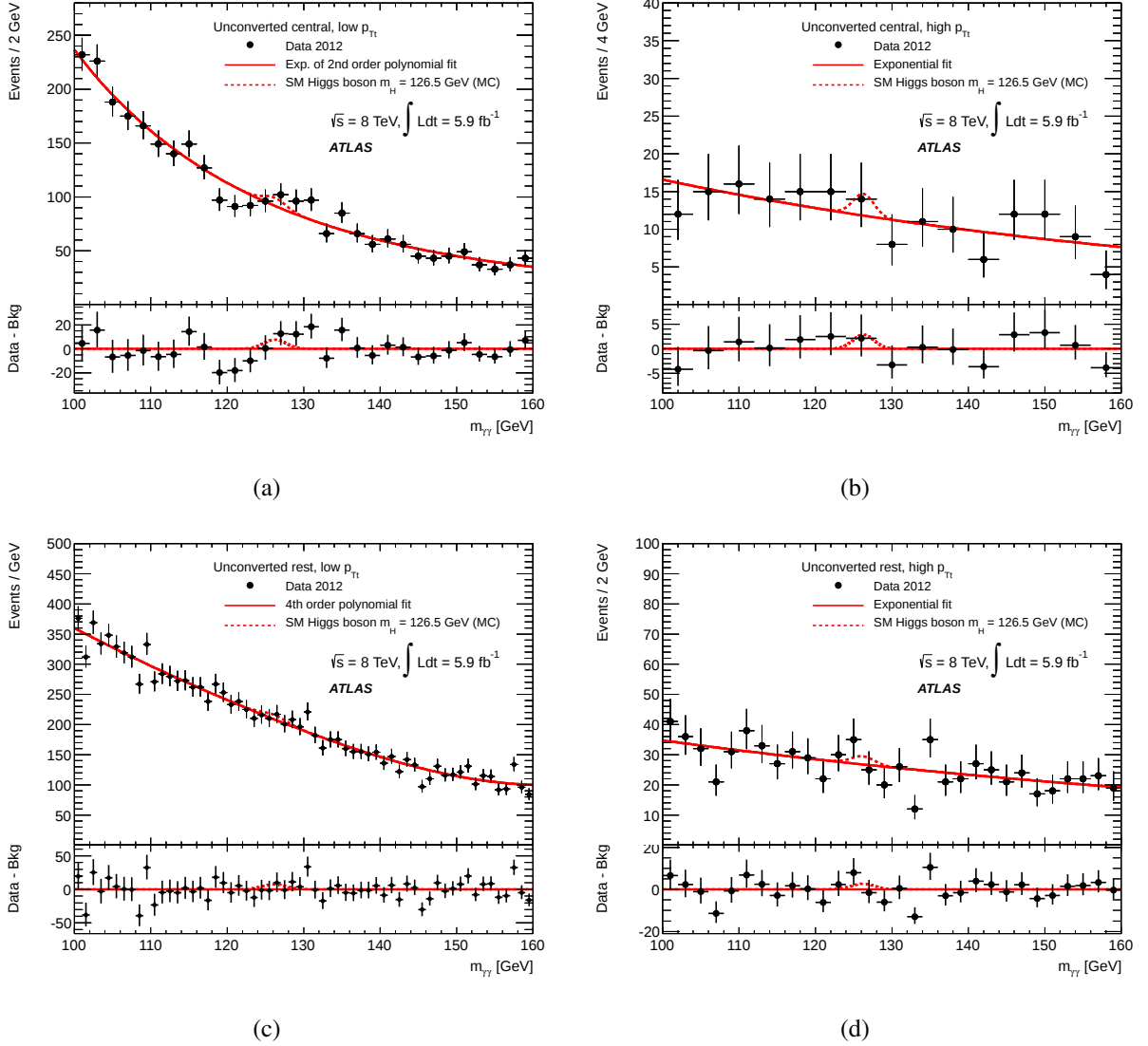


Figure 9.5: Background-only fits to the diphoton invariant mass spectra for (a) Unconverted central low p_{Tt} , (b) Unconverted central high p_{Tt} , (c) Unconverted rest low p_{Tt} , and (d) Unconverted rest high p_{Tt} categories of the *Discovery Analysis* correspond to the 8 TeV data sample. The bottom panel displays the residual of the data with respect to the background fit. The Higgs boson expectation for a mass hypothesis of 126.5 GeV corresponding to the Standard Model cross section is also shown.

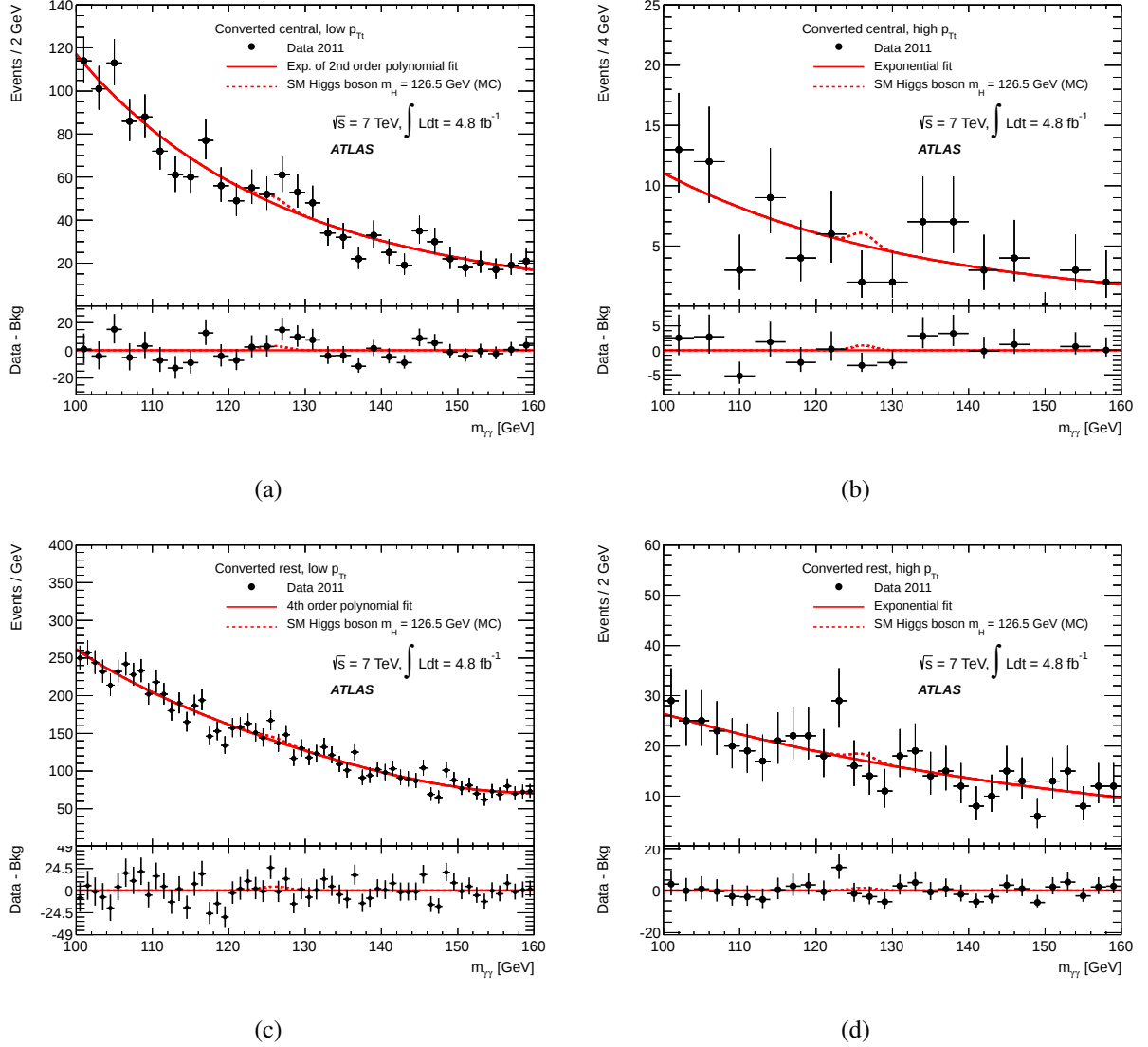


Figure 9.6: Background-only fits to the diphoton invariant mass spectra for (a) Converted central low p_{Tt} , (b) Converted central high p_{Tt} , (c) Converted rest low p_{Tt} , and (d) Converted rest high p_{Tt} categories of the *Discovery Analysis* correspond to the 7 TeV data sample. The bottom panel displays the residual of the data with respect to the background fit. The Higgs boson expectation for a mass hypothesis of 126.5 GeV corresponding to the Standard Model cross section is also shown.

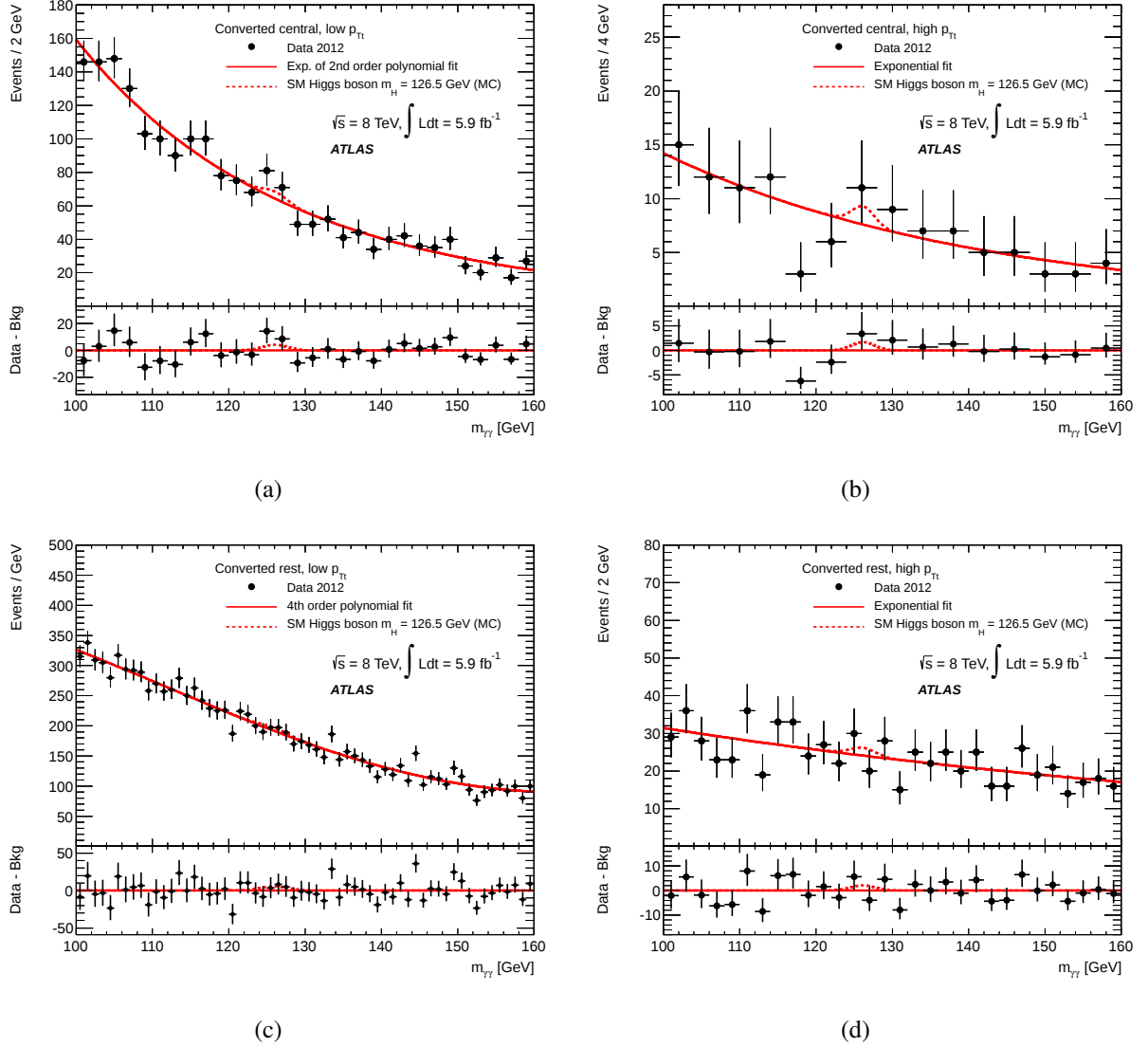


Figure 9.7: Background-only fits to the diphoton invariant mass spectra for (a) Converted central low p_{Tt} , (b) Converted central high p_{Tt} , (c) Converted rest low p_{Tt} , and (d) Converted rest high p_{Tt} categories of the *Discovery Analysis* correspond to the 8 TeV data sample. The bottom panel displays the residual of the data with respect to the background fit. The Higgs boson expectation for a mass hypothesis of 126.5 GeV corresponding to the Standard Model cross section is also shown.

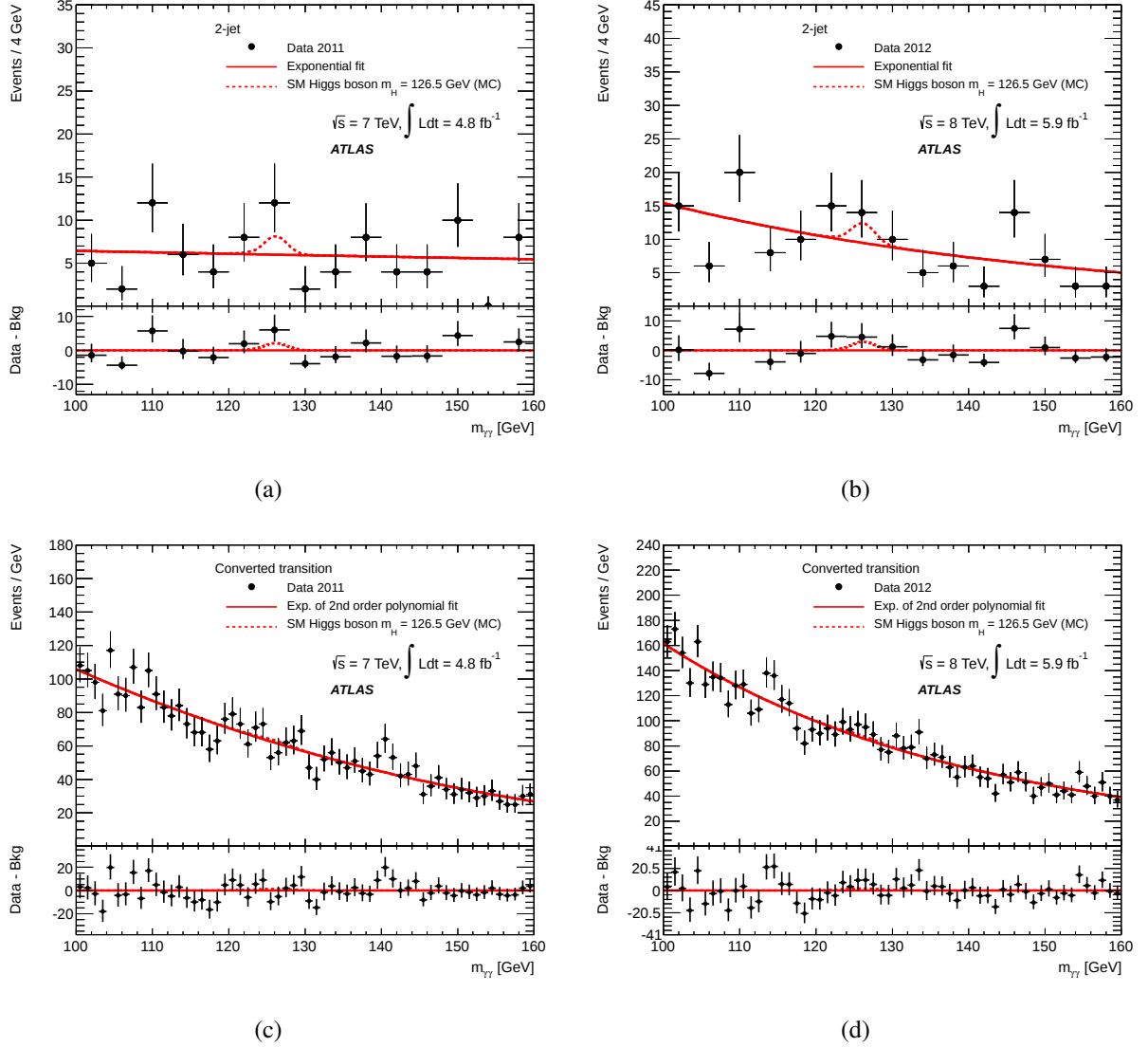


Figure 9.8: Background-only fits to the diphoton invariant mass spectra for the Two-jet category of the *Discovery Analysis* correspond to the 7 TeV data (a) and the 8 TeV data (b), and Converted transition categories correspond to the 7 TeV data (c) and the 8 TeV data (d). The bottom panel displays the residual of the data with respect to the background fit. The Higgs boson expectation for a mass hypothesis of 126.5 GeV corresponding to the Standard Model cross section is also shown.

9.3.2 Statistical interpretations

The statistical procedures used to test the background-only hypothesis and to set exclusion limits are described in detail in Chapter 8. The compatibility of the selected events with the background-only hypothesis is quantified by the background-only p_0 . Scanning the mass range 110 – 150 GeV with 0.5 GeV step size, the largest deviation observed for the 7 TeV and the 8 TeV data samples are 3.5 standard deviations (σ) at $m_H = 126$ GeV and 3.4σ at $m_H = 127$ GeV, respectively. For a SM Higgs boson, the expected significances would be 1.6σ and 1.9σ at these hypothesized mass values, respectively. The positions of the two minima are compatible within their uncertainties.

Combining the $\sqrt{s} = 7$ TeV data sample and the $\sqrt{s} = 8$ TeV data sample, the combined p_0 -value results is shown in Figure 9.9, along with the p_0 for the $\sqrt{s} = 7$ TeV and the $\sqrt{s} = 8$ TeV analyses. The largest deviation from background-only hypothesis observed is 4.7σ at $m_H = 126.5$ GeV. This is reduced to 4.5σ when taking the energy scale systematic uncertainty into account as shown by circles in the plot. After correction for the look-elsewhere effect [115], a global significance of 3.6σ is found. At this hypothesized mass, the expected p_0 value for a SM Higgs boson is 7×10^{-3} (2.4σ local significance).

The best fit value for the combined signal strength μ is obtained from a simultaneous fit to all categories in the $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV data samples. The result is shown in Figure 9.10. As the first step towards understanding the properties of the new particle, at $m_H = 126.5$ GeV, the best fit value is found to be $\mu = 1.9 \pm 0.5$, which correspond to about 360 signal events. The best fit signal strengths at $m_H = 126.5$ GeV obtained from fits to the individual categories are in good agreement with the combined fit as shown in Figure 9.11.

Since the analysis is able to separate VBF events from other production modes, a first look into the signal strengths separately for fermion induced production processes (ggF and $t\bar{t}H$) and vector-boson induced production processes (VBF and VH) is shown by the likelihood contour in the $\mu_{ggF+t\bar{t}H}$ and μ_{VBF+VH} plane in Figure 9.12

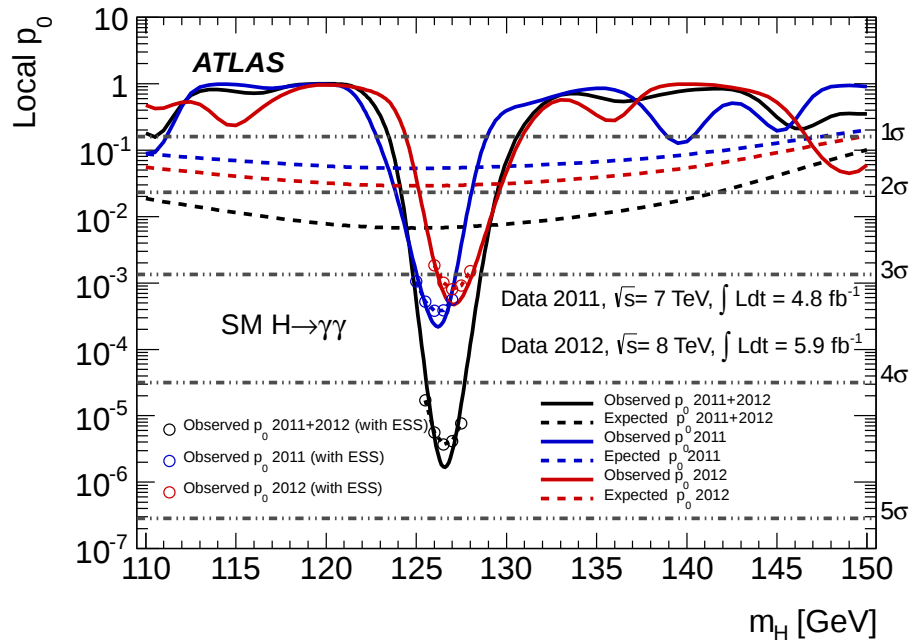


Figure 9.9: Expected and observed local p_0 values for a Standard Model Higgs boson as a function of the hypothesized Higgs boson mass m_H for the combined analysis and for the $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV data samples separately. The observed p_0 including the effect of the photon energy scale uncertainty on the mass position is included via pseudo-experiments and shown as open circles.

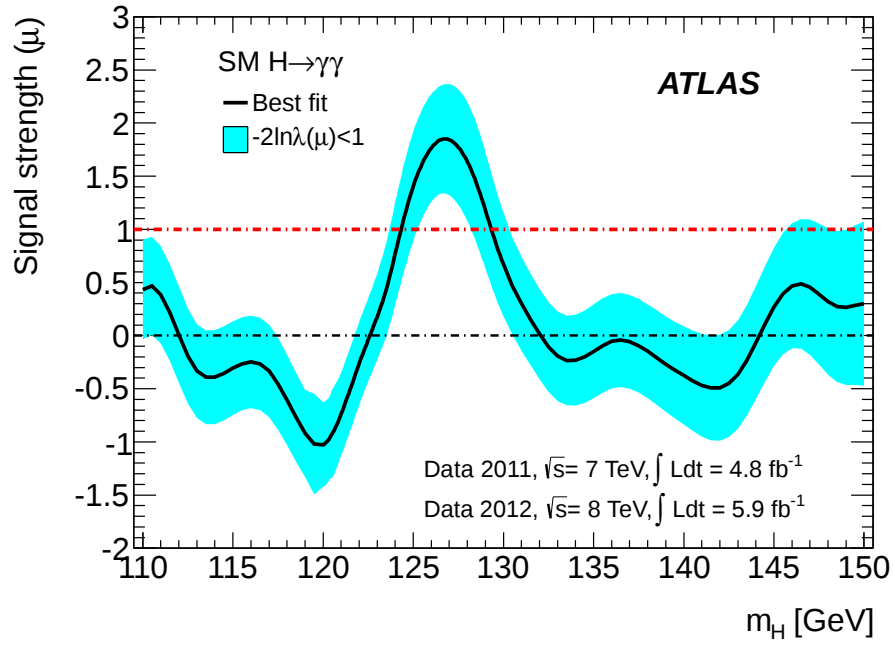


Figure 9.10: Best fit value for the signal strength as a function of the assumed Higgs boson mass m_H from the *Discovery Analysis*.

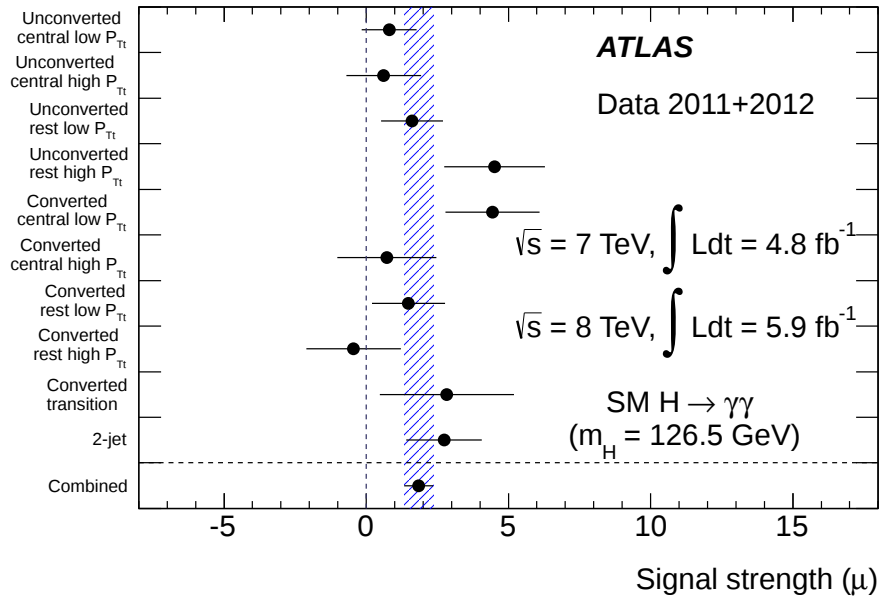


Figure 9.11: Best fit value for the signal strength in the different categories of the *Discovery Analysis* at $m_H = 126.5$ GeV for the combined $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV data samples.

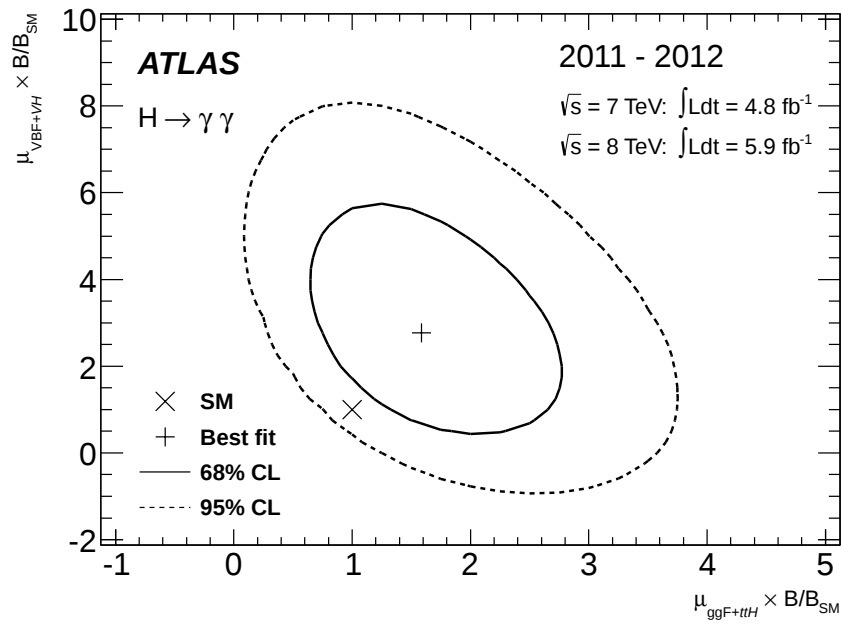


Figure 9.12: The two-dimensional best-fit value of $(\mu_{ggF+ttH}, \mu_{\text{VBF+VH}})$ from the *Discovery Analysis*. The 68% and 95% confidence level contours are shown with the solid and dashed lines, respectively.

9.3.3 Combination with other decay channels

To maximize the statistical power for the search of the Higgs boson, the $H \rightarrow \gamma\gamma$ channel is combined with $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ and $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$ channels following the procedure discussed in Reference [12]. A summary of the excess positions (quantified by m_H) and sizes (quantified by μ) in each channel is shown in Figure 9.13. As one can see they are in very good compatibility.

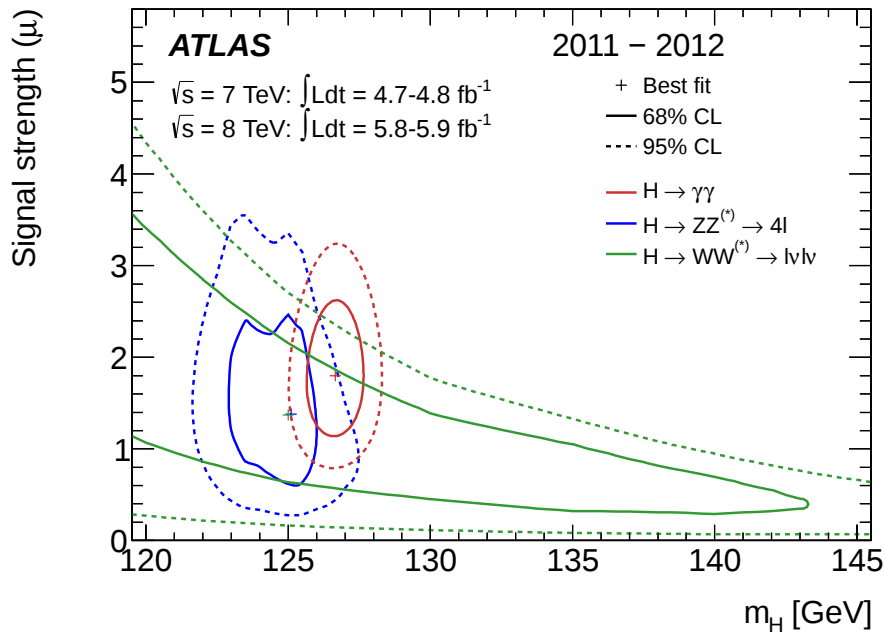


Figure 9.13: Confidence intervals in the (μ, m_H) plane for the $H \rightarrow \gamma\gamma$ (*Discovery Analysis*), $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$, and $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$ channels, including all systematic uncertainties. The markers indicate the maximum likelihood estimates in the corresponding channels .

The observed local p_0 values from the combination of three channels, using the asymptotic approximation as discussed in Section 8.2, are shown as a function of m_H in Figure 9.14. The largest local significance for the combination of the 7 and 8 TeV data is found for a SM Higgs boson mass hypothesis of $m_H = 126.5$ GeV, where it reaches 6.0σ , with an expected value in the presence of a SM Higgs boson signal at that mass of 4.9σ . This is conclusive evidence for the discovery of Higgs(-like) boson with mass around 125 GeV. In the next two chapters, properties of this new particle will be measured with larger data sample and better analyses.

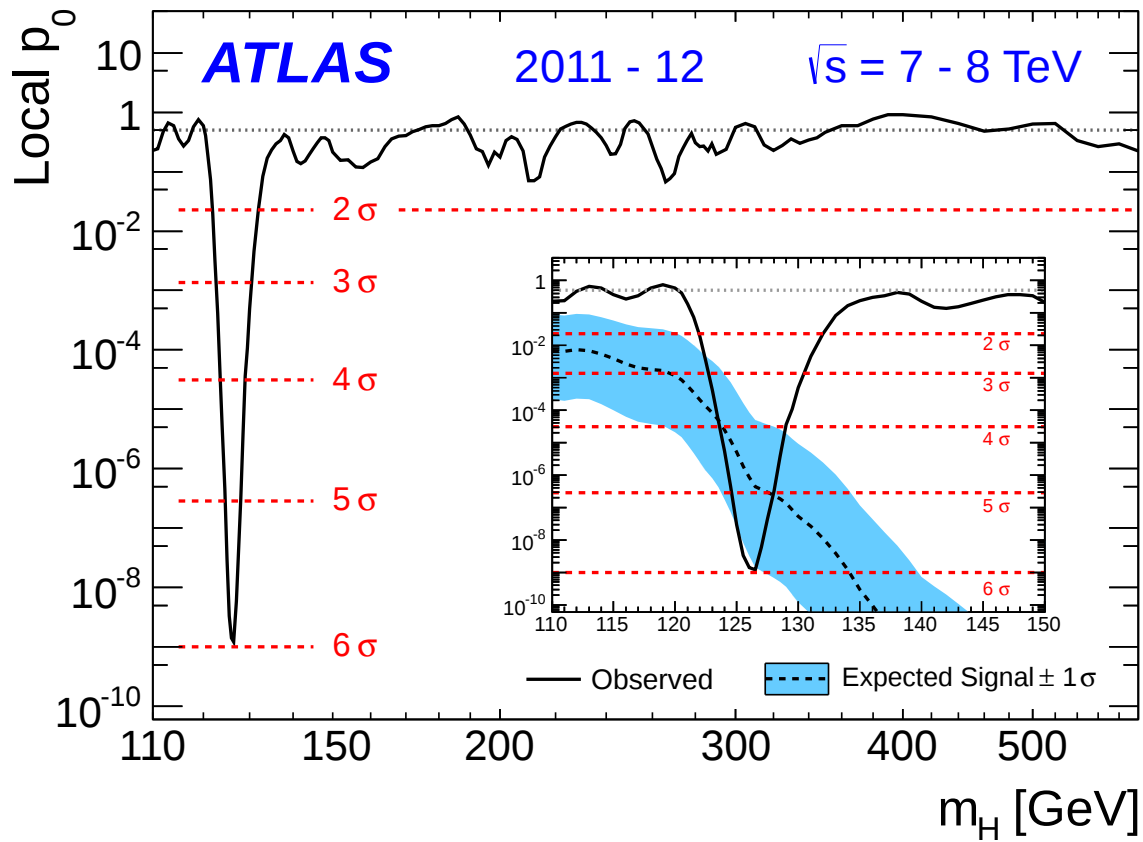


Figure 9.14: The observed (solid) local p_0 as a function of m_H . The dashed curve shows the expected local p_0 under the hypothesis of a Standard Model Higgs boson signal at that mass with its $\pm 1\sigma$ band. The horizontal dashed lines indicate the p -values corresponding to significances of 1 to 6σ .

Chapter 10

Measurement of Higgs boson couplings in diphoton decay channel

The measurements of Higgs boson couplings based on LHC Run 1 data starts from measuring the rates of different Higgs boson production processes in different decay channels. In $H \rightarrow \gamma\gamma$ channel, these rates are extracted from resonance signals in the diphoton invariant mass spectra of independent categories of events targeting different production modes (*Coupling Analysis*).

10.1 Event categorization

Compared with the *Discovery Analysis*, the *Coupling Analysis* has updated event categorization to better perform measurements of Higgs boson couplings. There are twelve categories introduced in total, which are applied to the 7 TeV and 8 TeV data each. The flowchart showing the sequence the twelve categories is in Figure 10.1. Only events that fail the selections of the previous category can be considered as candidates for the next category, to ensure that the categories are mutually exclusive. The sequence is chosen to prioritize the production modes that are expected to have the lowest signal yields. As to be described in the following subsections, the twelve categories can be classified into four groups depending on the production processes they are targeting at.

10.1.0.1 $t\bar{t}H$ categories

The first group consists of two categories that are designed to select data samples enriched with $t\bar{t}H$ events. $t\bar{t}H$ *leptonic* category exploits the products of the (semi-)leptonic decay of the top quark pair in a $t\bar{t}H$ event, while the $t\bar{t}H$ *hadronic* category captures the hadronic decay products of the top quark pair. The selections are optimized for the $t\bar{t}H$ signal, while maintaining good acceptance for the tH signal [41].

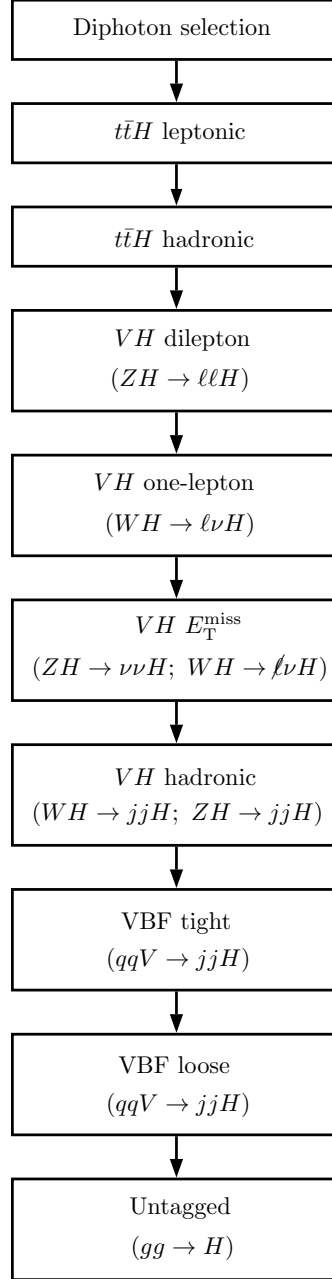


Figure 10.1: Illustration of the order in which the criteria for the exclusive event categories in the *Coupling Analysis* are applied to the selected diphoton events. The division of the last category, which is dominated by gg F production, into four sub-categories is described in the text.

Events in the $t\bar{t}H$ leptonic category are required to contain at least one electron or muon. Events are retained if either two or more b -jets are found or a single b -jet is found together with $E_T^{\text{miss}} > 20$ GeV. The b -jets are required to have $p_T > 25$ GeV and to be tagged using the 80%

(85%) efficiency working point of the b -tagging algorithm [98] in the 8 TeV (7 TeV) data. In order to suppress the background contribution from Z +jets with $Z \rightarrow e^+e^-$, where an electron is misidentified as a photon, events with an invariant electron-photon mass between 84 GeV and 94 GeV are rejected.

The $t\bar{t}H$ hadronic category vetos well-reconstructed and identified electron or muon passing the kinematic cuts described in Chapter 5. The events are then required to fulfill at least one of the following sets of criteria, which take into account the inefficiency of jet reconstruction or flavor tagging while still maintaining low level contamination of non- $t\bar{t}H$ and non- tH Higgs boson signals by tightening b -tagging working point and/or increasing jet p_T selection:

1. at least six jets with $p_T > 25$ GeV out of which two are b -tagged using the 80% working point;
2. at least six jets with $p_T > 30$ GeV out of which one is b -tagged using the 60% working point;
3. at least five jets with $p_T > 30$ GeV out of which two are b -tagged using the 70% working point.

Only the first set of criteria above are applied to the 7 TeV data but with a working point efficiency of 85%.

The fraction of $t\bar{t}H$ events relative to all signal production passing this selection in the $t\bar{t}H$ hadronic category is larger than 80%. whereas in the $t\bar{t}H$ leptonic category it ranges from 73% to 84% depending on the center-of-mass energy. The fractions of tH events in both categories are about 10%. The dominant contamination in $t\bar{t}H$ hadronic category is from ggF , whereas in the $t\bar{t}H$ leptonic category it is from WH .

10.1.0.2 VH categories

The second group of categories are optimized to identify VH events. The VH dilepton category is dedicated to the ZH production mode by exploiting the dilepton decay of the Z boson. It requires a pair of opposite-sign electrons or muons, with the dilepton invariant mass required to be in the range between 70 GeV and 110 GeV. These requirements lead to a 99% signal-only purity for ZH production, the remaining 1% coming from $t\bar{t}H$ production.

The VH *one-lepton* category is optimized mainly to select events with a leptonic decay of the W boson by requiring the presence of one electron or muon. In order to exploit the E_T^{miss} signature of the W boson decay, the E_T^{miss} significance, as defined in Section 5.2, is required to be larger than 1.5. Approximately 90% of the signal events in this category are predicted to come from WH production, about 6% from ZH production, and 1 – 2% from $t\bar{t}H$ production.

The VH E_T^{miss} category is mainly aiming for ZH events with the Z boson decays to two neutrinos, or WH events with the W decays leptonically, but the lepton escapes detection or does not pass the selections. It requires the E_T^{miss} significance to be larger than 5, roughly equivalent to a direct requirement of $E_T^{\text{miss}} > 70 - 100$ GeV, depending on the value of $\sum E_T$. A further enrichment is obtained by requiring p_{Tt} greater than 20 GeV. After the event selection approximately 50% of the signal events in this category are predicted to come from ZH production, 40% from WH production, and the remaining 10% mainly from $t\bar{t}H$ production.

The VH *hadronic* category consists of events that include the signature of a hadronically decaying vector boson. They are selected by requiring the presence of two reconstructed jets with a dijet invariant mass m_{jj} in the range between 60 GeV and 110 GeV. The sensitivity is further enhanced by requiring the difference between the pseudorapidities of the diphoton and the dijet systems $|\eta_{\gamma\gamma} - \eta_{jj}|$ to be less than one, and the diphoton p_{Tt} greater than 70 GeV. The distributions of the discriminating variables used to define the VH hadronic category are shown in Figure 10.2 for signal events from different production modes and for events from data and MC background. The MC background is composed of a mixture of $\gamma\gamma$, γj and jj samples. Approximately 30% (20%) of the events in the VH hadronic category come from WH (ZH) production after the selection, while the remaining fraction is mainly accounted for by ggF events surviving the selections.

10.1.0.3 VBF categories

The selection of VBF-like events starts from requiring at least two reconstructed jets. The two leading jets (ranked by jet p_T) are required to satisfy $|\eta^*| < 5.0$ and $\Delta\eta_{jj} \geq 2.0$, where η^* is the pseudorapidity of the diphoton system relative to the average rapidity of the two leading jets $\eta^* \equiv \eta_{\gamma\gamma} - (\eta_{j1} + \eta_{j2})/2$ [116] and $\Delta\eta_{jj}$ is the pseudorapidity separation between the two leading jets. In order to optimize the sensitivity to VBF production mode, after the pre-selection

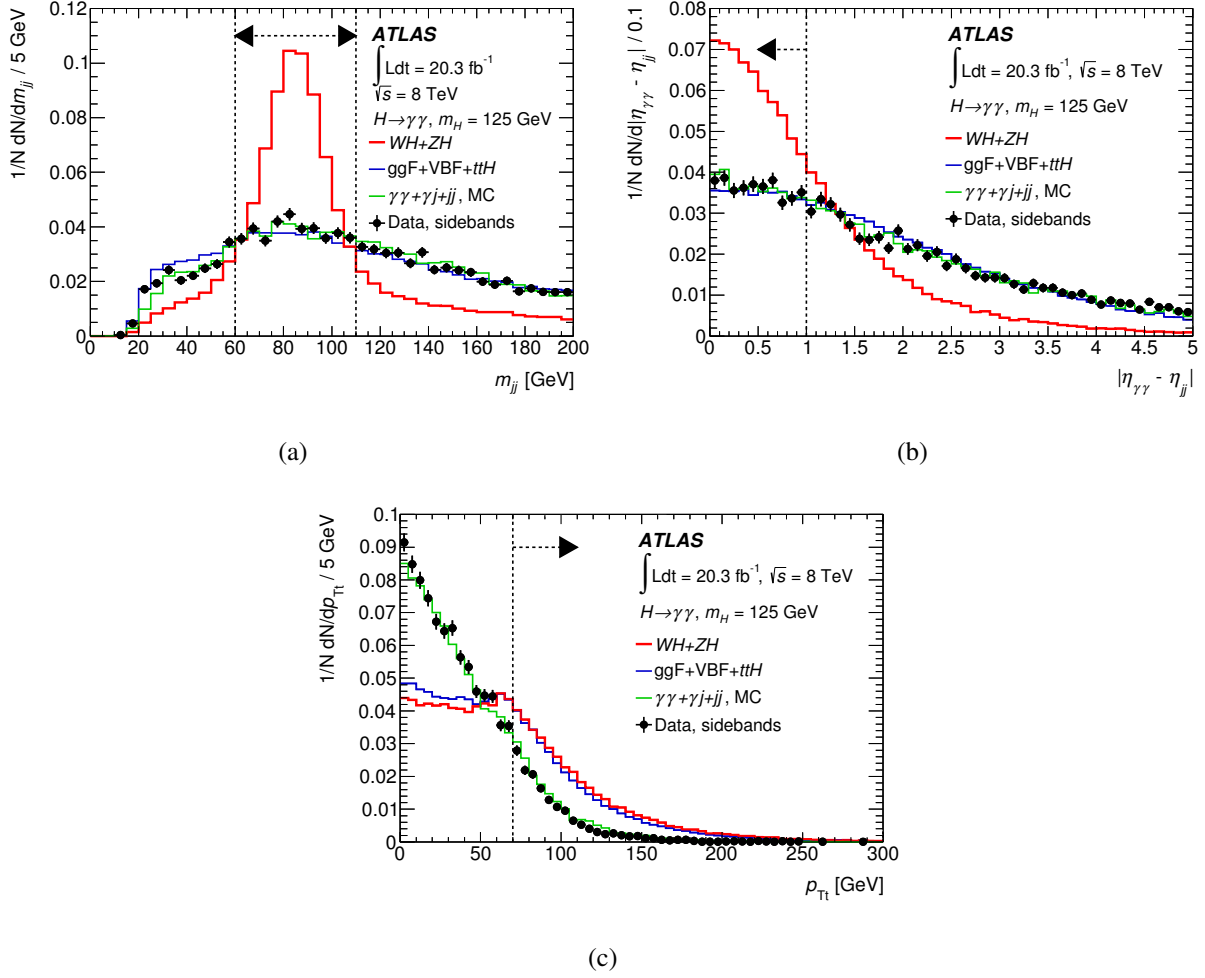


Figure 10.2: Normalized distributions of the variables described in the text used to sort diphoton events with at least two reconstructed jets into the VH hadronic category of the *Coupling Analysis* for the data in the sidebands (points), the predicted sum of the WH and ZH signals (red histograms), the predicted signal feed-through from ggF , VBF , and $t\bar{t}H$ production modes (blue histograms), and the simulation of the $\gamma\gamma$, γj , and jj background processes (green histograms). The arrows indicate the selection criteria applied to these observables. The mass of the Higgs boson in all signal samples is $m_H = 125 \text{ GeV}$.

described above, a BDT algorithm then combines following six discriminating variables into a single discriminant that takes into account the correlations among them:

1. m_{jj} , the invariant mass of the two leading jets j_1 and j_2 ;

2. $\Delta\eta_{jj}$;
3. p_{Tt} ;
4. $\Delta\phi_{\gamma\gamma,jj}$, the azimuthal angle between the diphoton and the dijet systems;
5. $\Delta R_{\gamma,j}^{\min}$, the minimum separation between the leading/sub-leading photon and the leading/sub-leading jet;
6. η^* .

After the pre-selection, these variables are found to have little or no correlation to $m_{\gamma\gamma}$, thus ensuring that no biases in the final diphoton mass fit are introduced. The individual separation power between VBF and ggF and prompt $\gamma\gamma$, γj and jj background events is illustrated in Figure 10.3 for each discriminating variable.

The signal sample used to train the BDT is composed of simulated VBF events, while a mixture of samples is used for the background, including simulated ggF events, simulation of the $\gamma\gamma$ background component with SHERPA, and events from data in which one or both photon candidates fail to satisfy the isolation criteria for the reducible γj and jj components. The contribution from ggF to the background sample is normalized to the rate predicted by the SM. The other background components are weighted in order to reproduce the background composition measured in the data.

Events are sorted into two categories with different VBF purities according to the output value of the BDT, O_{BDT} :

1. VBF *tight*: $O_{\text{BDT}} \geq 0.83$;
2. VBF *loose*: $0.3 < O_{\text{BDT}} < 0.83$.

Figure 10.4 shows the distributions of O_{BDT} for the VBF signal, feed-through from ggF production, the simulated continuum background, and data from the sidebands. The O_{BDT} distributions of the background MC prediction and the data in the sidebands are in good agreement. As an additional cross-check, the BDT is applied to a large sample of $Z \rightarrow e^+e^- + \text{jets}$ in data and MC samples. The resulting O_{BDT} distributions are found to be in excellent agreement between data

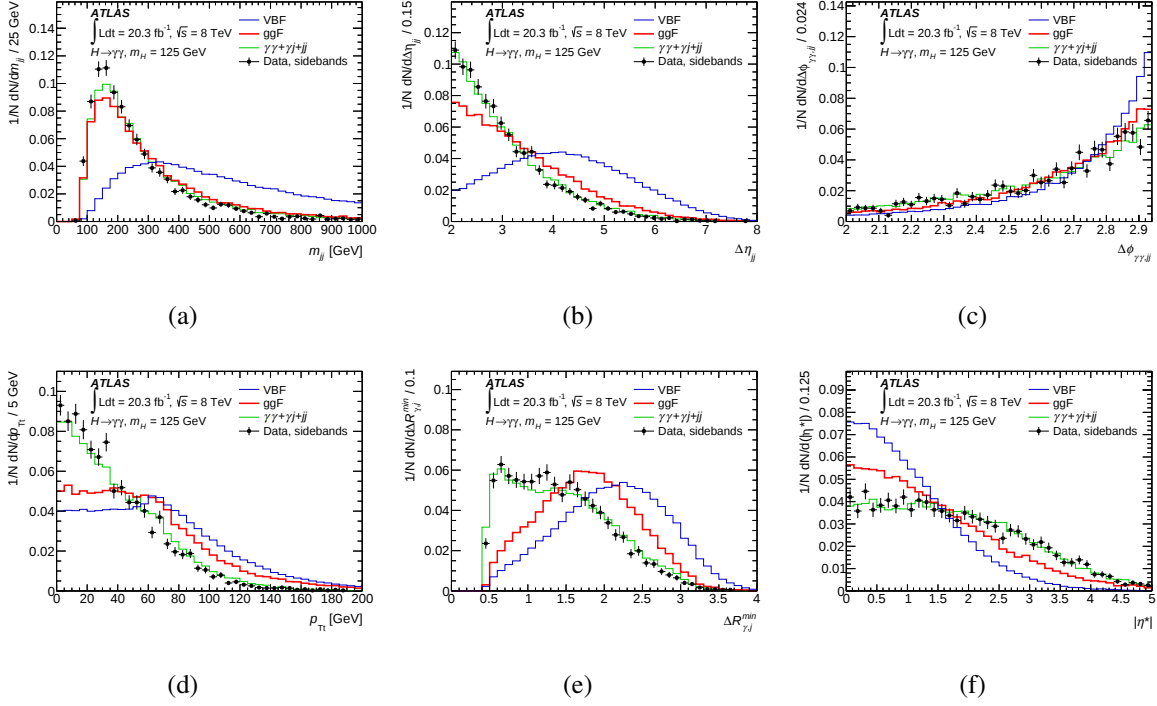


Figure 10.3: Normalized kinematic distributions of the six variables describe in the text used to build the Boost Decision Tree that assigns events to the VBF categories of the *Coupling Analysis*, for diphoton candidates with two well-separated jets ($\Delta\eta_{jj} \geq 2.0$ and $|\eta^*| < 5.0$). Distributions are shown for data sidebands (points) and simulation of the VBF signal (blue histograms), feed-through from ggF production (red histograms), and the continuum QCD background predicted by MC simulation and data control regions (green histograms) as described in the text. The signal VBF and ggF samples are generated with a Higgs boson mass $m_H = 125$ GeV.

and MC. The fraction of VBF events in the VBF tight (loose) category is approximately 80% (60%), the remaining 20% (40%) being contributed by ggF events.

10.1.0.4 Untagged (ggF) categories

Compared with the *Discovery Analysis*, the categorization for the bulk part of the dataset enriched with ggF events (untagged categories) has been simplified from nine (Section 9.1) to four defined by p_{Tt} and η . The loss of sensitivity from not using conversion information and reducing η segmentation is mitigated by increasing the p_{Tt} threshold and re-tuning η selection as shown below:

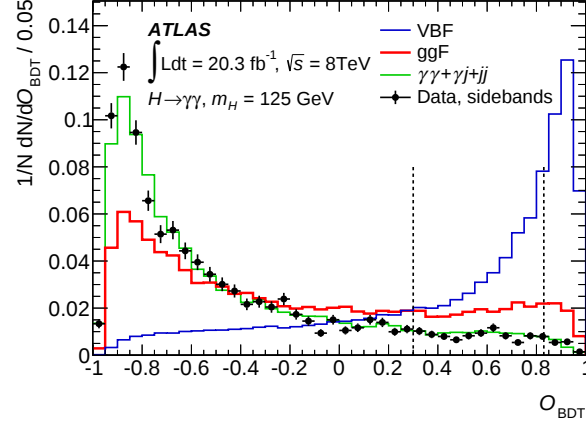


Figure 10.4: Probability distributions of the output of the Boost Decision Tree (BDT) O_{BDT} for the VBF signal (blue), $gg\text{F}$ feed-through (red), continuum QCD background predicted by MC samples and data control regions (green) as described in the text, and data sidebands (points). The two vertical dashed lines indicate the cuts on O_{BDT} that define the VBF loose and tight categories in the *Coupling Analysis*. The signal VBF and $gg\text{F}$ samples are generated with a Higgs boson mass $m_H = 125$ GeV.

1. *Central low* p_{Tt} : $p_{\text{Tt}} \leq 70$ GeV and both photons have $|\eta| < 0.95$;
2. *Central high* p_{Tt} : $p_{\text{Tt}} > 70$ GeV and both photons have $|\eta| < 0.95$;
3. *Forward low* p_{Tt} : $p_{\text{Tt}} \leq 70$ GeV and at least one photon has $|\eta| \geq 0.95$;
4. *Forward high* p_{Tt} : $p_{\text{Tt}} > 70$ GeV and at least one photon has $|\eta| \geq 0.95$.

This categorization of the *untagged* events increases the signal-to-background ratio of the events with high p_{Tt} with a gain of about a factor of three (two) for Central (Forward) categories with respect to low p_{Tt} events as shown in Figure 10.5.

The optimization of the categorization is studied using MC background sample and cross checked using data sideband. Good agreement has been shown between the results from the two approaches. The typical fraction of $gg\text{F}$ events in the low (high) p_{Tt} categories is 90% (70%). The remaining 10% (30%) is equally accounted for by the contribution from VBF events and the sum of all the remaining processes.

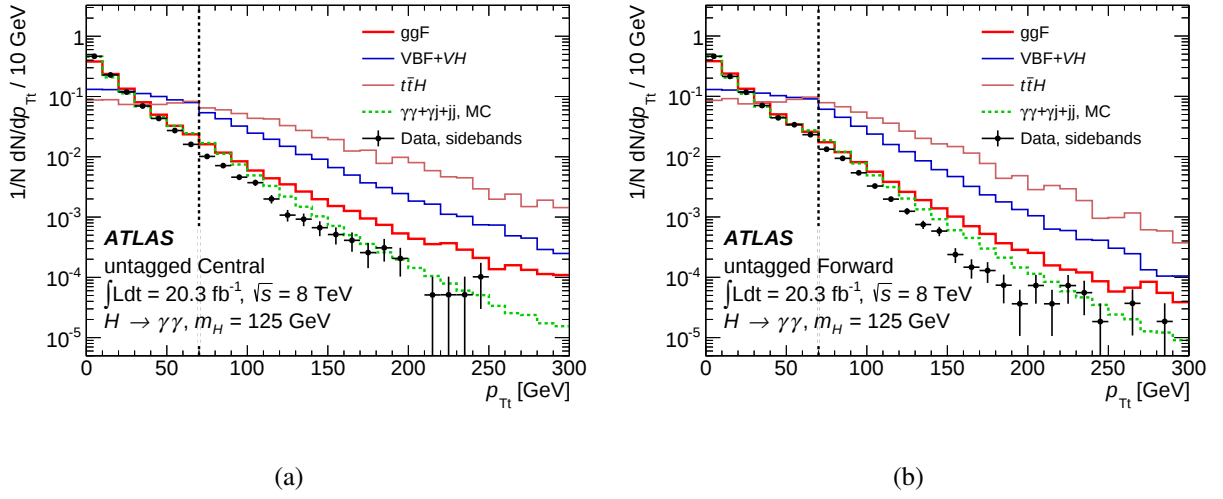


Figure 10.5: Distributions of p_{Tt} for diphoton candidates in the sidebands in the untagged (a) Central and (b) Forward categories for $\sqrt{s} = 8$ TeV for predicted Higgs boson production processes (solid histograms), the predicted sum of $\gamma\gamma$, γj and $j j$ background processes (green histogram), and data (points). The vertical dashed lines indicate the value used to classify events into the low or high p_{Tt} categories in the *Coupling Analysis*. The mass for all Higgs boson signal samples is $m_H = 125$ GeV.

10.1.0.5 Summary of categories

The predicted signal efficiencies, which include geometrical and kinematic acceptances, and event fractions per production mode in each category for $m_H = 125.4 \text{ GeV}$ ¹ are listed in Tables 10.1 and 10.2 for the 7 TeV and 8 TeV data, respectively. The total expected numbers of signal events per category N_S are also shown. The expected contamination of ggF and VBF in the $VH E_T^{\text{miss}}$ category is larger in 7 TeV data than in 8 TeV data due to the poorer resolution of the E_T^{miss} reconstruction algorithm used in the 7 TeV analysis.

The number of events observed in data in each category is reported in Table 10.3 separately for the 7 TeV and 8 TeV data. The impact of the event categorization described in the previous sections on the uncertainty in the combined signal strength is estimated on a representative signal plus MC background sample generated under the SM hypothesis. The event categorization is found to provide a 20% reduction of the total uncertainty with respect to an inclusive analysis.

¹ $m_H = 125.4 \text{ GeV}$ is measured from the combination of ATLAS $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ channels. More details will be given in Chapter 11.

Category	ggF		VBF		WH		ZH		$t\bar{t}H$		$b\bar{b}H$		$tHq\bar{b}$		tHW	
	$\epsilon(\%)$	$f(\%)$	$\epsilon(\%)$	$f(\%)$	$\epsilon(\%)$	$f(\%)$	$\epsilon(\%)$	$f(\%)$	$\epsilon(\%)$	$f(\%)$	$\epsilon(\%)$	$f(\%)$	$\epsilon(\%)$	$f(\%)$	$\epsilon(\%)$	$f(\%)$
Central low p_{Tt}	15.5	92.2	8.5	4.1	7.2	1.6	7.9	1.0	3.4	0.1	15.5	1.0	-	-	-	26.0
Central high p_{Tt}	1.0	71.8	2.7	16.4	2.1	6.1	2.3	3.7	2.9	1.2	1.0	0.7	-	-	-	2.1
Forward low p_{Tt}	23.3	91.5	13.2	4.2	13.5	2.0	14.3	1.2	4.3	0.1	23.3	0.9	-	-	-	39.5
Forward high p_{Tt}	1.3	70.6	4.0	16.7	3.5	6.9	3.6	4.1	2.9	0.9	1.3	0.7	-	-	-	3.0
VBF loose	0.4	38.6	7.9	60.0	0.2	0.6	0.2	0.3	0.2	0.1	0.4	0.4	-	-	-	1.7
VBF tight	0.1	18.1	6.3	81.5	<0.1	0.1	<0.1	0.1	0.1	<0.1	0.1	0.2	-	-	-	1.0
VH hadronic	0.2	43.5	0.1	3.3	3.2	31.8	3.4	19.8	0.9	1.3	0.2	0.4	-	-	-	0.6
$VH E_T^{\text{miss}}$	<0.1	8.7	0.1	3.7	1.7	35.7	3.6	44.8	2.3	7.1	<0.1	0.1	-	-	-	0.3
VH one-lepton	<0.1	0.7	<0.1	0.2	5.0	91.4	0.6	5.9	0.7	1.8	<0.1	<0.1	-	-	-	0.3
VH dilepton	<0.1	<0.1	<0.1	<0.1	<0.1	<0.1	1.3	99.3	<0.1	0.6	<0.1	<0.1	-	-	-	0.1
$t\bar{t}H$ hadronic	<0.1	10.5	<0.1	1.3	<0.1	1.3	<0.1	1.4	6.1	81.0	<0.1	0.1	1.5	2.6	4.3	1.9
$t\bar{t}H$ leptonic	<0.1	0.6	<0.1	0.1	0.3	14.9	0.1	4.0	8.5	72.6	<0.1	<0.1	4.8	5.3	8.7	2.5
Total efficiency (%)	41.8	-	42.9	-	36.7	-	37.3	-	32.2	-	41.8	-	-	-	-	41.6%
Events	64.8		5.4		2.2		1.3		0.3		0.7		<0.1		<0.1	74.5

Table 10.1: Signal efficiencies ϵ , which include geometrical and kinematic acceptances, and expected signal event fractions f per production mode in each category of the *Coupling Analysis* for $\sqrt{s} = 7$ TeV and $m_H = 125.4$ GeV. The second-to-last row shows the total efficiency per production process summed over the categories and the overall average efficiency in the far right column. The total number of selected signal events expected in each category N_S is reported in the last column while the total number of selected events expected from each production mode is given in the last row.

Category	ggF		VBF		WH		ZH		$t\bar{t}H$		$b\bar{b}H$		$tHq\bar{b}$		tHW		N_S
	$\epsilon(\%)$	$f(\%)$	$\epsilon(\%)$	$f(\%)$	$\epsilon(\%)$	$f(\%)$	$\epsilon(\%)$	$f(\%)$	$\epsilon(\%)$	$f(\%)$	$\epsilon(\%)$	$f(\%)$	$\epsilon(\%)$	$f(\%)$	$\epsilon(\%)$	$f(\%)$	
Central low p_{Tt}	14.1	92.3	7.5	4.0	6.5	1.5	7.2	1.0	2.9	0.1	14.1	1.0	-	-	-	-	135.5
Central high p_{Tt}	0.9	73.3	2.5	15.7	1.9	5.5	2.0	3.4	2.4	1.3	0.9	0.8	-	-	-	-	11.3
Forward low p_{Tt}	21.6	91.7	11.9	4.1	12.3	1.9	13.0	1.2	3.8	0.1	21.6	1.0	-	-	-	-	208.6
Forward high p_{Tt}	1.3	71.9	3.6	16.2	3.2	6.4	3.3	3.9	2.5	0.9	1.3	0.8	-	-	-	-	16.1
VBF loose	0.4	41.9	7.2	56.5	0.2	0.6	0.2	0.4	0.2	0.1	0.4	0.4	-	-	-	-	9.3
VBF tight	0.1	19.0	6.4	80.5	< 0.1	0.2	< 0.1	0.1	0.1	0.1	0.1	0.2	-	-	-	-	5.7
VH hadronic	0.2	45.9	0.1	3.2	3.0	30.3	3.1	18.8	0.7	1.3	0.2	0.5	-	-	-	-	3.2
$VH E_{T}^{miss}$	< 0.1	2.3	< 0.1	0.3	1.3	36.9	3.0	51.0	1.8	9.5	< 0.1	< 0.1	-	-	-	-	1.1
VH one-lepton	< 0.1	0.5	< 0.1	0.2	4.8	89.8	0.6	6.3	1.0	3.3	< 0.1	< 0.1	-	-	-	-	1.7
VH dilepton	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	1.3	99.1	< 0.1	0.9	< 0.1	< 0.1	-	-	-	-	0.3
$t\bar{t}H$ hadronic	< 0.1	7.3	< 0.1	1.0	< 0.1	0.7	< 0.1	1.3	6.9	84.1	< 0.1	< 0.1	2.1	3.4	4.8	2.1	0.5
$t\bar{t}H$ leptonic	< 0.1	1.0	< 0.1	0.2	0.1	8.1	0.1	2.3	7.9	80.3	< 0.1	< 0.1	4.1	5.5	7.1	2.6	0.6
Total efficiency (%)	38.7	-	39.1	-	33.3	-	33.8	-	30.2	-	38.7	-	< 0.1	< 0.1	< 0.1	< 0.1	38.5%
Events	342.8		28.4		10.7		6.4		1.8		3.6		< 0.1	< 0.1	< 0.1	< 0.1	393.8

Table 10.2: Signal efficiencies ϵ , which include geometrical and kinematic acceptances, and expected signal event fractions f per production mode in each category of the *Coupling Analysis* for $\sqrt{s} = 8$ TeV and $m_H = 125.4$ GeV. The second-to-last row shows the total efficiency per production process summed over the categories and the overall average efficiency in the far right column. The total number of selected signal events expected in each category N_S is reported in the last column while the total number of selected events expected from each production mode is given in the last row.

Category	$\sqrt{s} = 7 \text{ TeV}$	$\sqrt{s} = 8 \text{ TeV}$
Central low p_{Tt}	4400	24080
Central high p_{Tt}	141	806
Forward low p_{Tt}	12131	66394
Forward high p_{Tt}	429	2528
VBF loose	58	411
VBF tight	7	67
VH hadronic	34	185
$VH E_T^{\text{miss}}$	14	35
VH one-lepton	5	38
VH dilepton	0	2
$t\bar{t}H$ hadronic	3	15
$t\bar{t}H$ leptonic	3	5
Total	17225	94566

Table 10.3: Number of selected events in each category of the *Coupling Analysis* and total for the 7 TeV and 8 TeV datasets in the mass range 105 – 160 GeV.

10.2 Systematic uncertainties

The various types of systematic uncertainties are presented in this section according to the way they affect the determination of the signal strengths. The theoretical and experimental uncertainties on the yields of diphoton events from Higgs boson decays are discussed in Section 10.2.1. The systematic uncertainties affecting the event categorization due to migrations of signal events from or to other categories are presented in Section 10.2.2. The systematic uncertainties related to the photon energy scale and resolution are reported in Section 10.2.3. The systematic uncertainties due to potential spurious signals induced by systematic differences between the background parameterization and the background component of the data are obtained with the technique described in Section 7.2 and reported in Section 10.2.4.

10.2.1 Uncertainties on integrated signal yield

10.2.1.1 Theoretical uncertainties

The predicted total cross sections for the signal processes have uncertainties as listed in Table 10.4 for $m_H = 125.4$ GeV, separately for $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV. The uncertainties estimated by varying the QCD scales affect the production processes independently, apart from WH and ZH uncertainties, which are treated as fully correlated. For the $tHqb$ and tHW production processes, the scale uncertainties are obtained by varying the renormalization and factorization scales by factors of 1/2 and 2 in the event generators (Section 4.2) and the PDF uncertainties are estimated by studying the impact of the variations within the CT10 [56] PDF set. For the other processes these uncertainties are taken from Reference [16]. The combined uncertainties on the effective luminosities for gg - and qq -initiated processes due to PDF and α_s uncertainties are independent, but they affect the relevant processes coherently. Both of these sets of uncertainties affect the 7 TeV and the 8 TeV cross sections coherently. The impact of scale and PDF uncertainties on the kinematic acceptance for signal events is found to be negligible relative to the impact of the uncertainties on the cross sections. The uncertainty on the $H \rightarrow \gamma\gamma$ branching ratio for $m_H = 125.4$ GeV is 5%.

Process	QCD scale		PDF+ α_s	
	$\sqrt{s} = 7 \text{ TeV}$	$\sqrt{s} = 8 \text{ TeV}$	$\sqrt{s} = 7 \text{ TeV}$	$\sqrt{s} = 8 \text{ TeV}$
ggF	+7.1	+7.2	+7.6	+7.5
	-7.8	-7.8	-7.1	-6.9
VBF	+0.3	+0.2	+2.5	+2.6
	-0.3	-0.2	-2.1	-2.8
WH	+1.0	+1.0	+2.6	+2.4
	-1.0	-1.0	-2.6	-2.4
ZH	+2.9	+3.1	+2.6	+2.5
	-2.9	-3.1	-2.6	-2.5
$t\bar{t}H$	+3.2	+3.8	+8.4	+8.1
	-9.3	-9.3	-8.4	-8.1
$b\bar{b}H$	+10	+10	+6.2	+6.1
	-15	-15	-6.2	-6.1
$tHqb$	+7	+6	+4	+4
	-6	-5	-4	-4
tHW	+7	+9	+10	+10
	-6	-7	-10	-10

Table 10.4: Theoretical uncertainties (in %) on cross sections for Higgs boson production processes at $\sqrt{s} = 7 \text{ TeV}$ and $\sqrt{s} = 8 \text{ TeV}$ for $m_H = 125.4 \text{ GeV}$ in the *Coupling Analysis*.

10.2.1.2 Experimental uncertainties

The expected signal yields are affected by the experimental systematic uncertainties listed below.

1. The uncertainties on the integrated luminosities are 1.8% for the 7 TeV data and 2.8% for the 8 TeV data [112]. They are treated as uncorrelated.
2. The trigger efficiencies in data are determined by combining the results from two different measurements. The first measurement is performed with photons in $Z \rightarrow \ell\ell\gamma$ events, where

ℓ is an electron or a muon. These events are collected with lepton-based triggers, making the photon candidates in these samples unbiased with respect to the trigger. The second measurement, based on the bootstrap technique described in Reference [101], is performed on a background-corrected photon sample selected only by a first-level trigger, which has an efficiency of 100% for signal-like photons in events that pass the diphoton selection criteria. Both measurements are dominated by statistical uncertainties. The uncertainties on the trigger efficiencies based on these measurements are estimated to be 0.2% for both the 7 TeV and 8 TeV data and are fully uncorrelated.

3. The uncertainty on the photon identification efficiency for the 8 TeV data is derived from measurements performed with data using three different methods [85] that cover the full E_T spectrum relevant for this analysis. In the first method, the efficiency is measured in a pure and unbiased sample of photons obtained by selecting radiative $Z \rightarrow \ell\ell\gamma$ decays without using the photon identification to select the photon, and where ℓ is an electron or a muon. In the second method, the photon efficiency is measured using $Z \rightarrow e^+e^-$ data by extrapolating the properties of electron showers to photon showers using MC events [85]. In the third method, the photon efficiency is determined from a data sample of isolated photon candidates from prompt γj production after subtracting the measured fraction of $j j$ background events. The combined uncertainty on the photon identification efficiency in data relative to MC simulation ranges between 0.5% and 2.0% depending on the E_T and η of the photon and on whether the photon is unconverted or converted and reconstructed with one or two tracks. For the 7 TeV data, more conservative uncertainties, ranging from 4% to 7%, are used because of the stronger correlation of the neural-network-based identification algorithm with the photon isolation, and because it relies more strongly on the correlations between the individual shower shape variables. Because these two effects complicate the measurement of the identification performance in data, conservative uncertainties, taken as the full difference between the efficiencies measured in data and the ones predicted by simulation, are used. The uncertainties on the signal yield due to the uncertainty on the photon identification efficiency are 8.4% for the 7 TeV data and 1.0% for the 8 TeV data and are treated as uncorrelated.

4. The uncertainty on the isolation efficiency is conservatively taken as the full size of the applied correction described in Section 5.1.4. The effect on the signal yield varies among categories (depending on their photon E_T spectrum). These uncertainties, which range between 1.3% and 2.3%, are estimated with the 8 TeV dataset and are assumed to be the same in the 7 TeV data but uncorrelated between the two datasets.
5. Finally, uncertainties on the signal yields due to the photon energy scale and primary vertex selection are found to be negligible relative to the ones discussed above.

The estimated values of the experimental uncertainties for both datasets are summarized in Table 10.5. Larger uncertainties, also shown in the table, on the photon identification and isolation selection efficiencies are assigned to $t\bar{t}H$ categories and VH hadronic category. The presence of large hadronic activity (high jet multiplicity) in these events, which is partially correlated with the photon selection and isolation efficiency, makes it difficult to measure the efficiencies precisely.

Uncertainty source	$\sqrt{s} = 7 \text{ TeV}$	$\sqrt{s} = 8 \text{ TeV}$
Luminosity	1.8	2.8
Trigger	0.2	0.2
Photon Id.	8.4(9.3)	1.0(4.1)
Isolation	1.3–2.3(3.8)	1.3–2.3(3.8)

Table 10.5: Relative systematic uncertainties on the inclusive yields (in %) for the 7 TeV and 8 TeV datasets. The numbers in parentheses refer to the uncertainties applied to $t\bar{t}H$ and VH categories. The ranges of the category-dependent uncertainties due to the isolation efficiency are reported.

10.2.1.3 Sizes of MC samples

While the MC statistics are decent in general, in some categories requiring distinct topologies (for example the presence of lepton(s) or high multiplicity of jets) the MC statistics are significantly reduced for certain production processes, resulting in non-negligible statistical uncertainties. The impact of these uncertainties on the individual signal strength parameters is estimated for each event category by analyzing representative signal MC samples containing both Higgs boson signal

and continuum background: the signal the sample size is fixed to the one expected in data by the SM predictions, and for the background to the observed numbers of events. The uncertainties that contribute more than 0.1% in quadrature to the total expected uncertainties are retained. The 14 uncertainties that contribute more than 0.1% in quadrature to the total expected uncertainties are propagated, but their contribution is at the level of 1% or less, which is much smaller than the expected statistical uncertainties on the individual signal strengths.

10.2.2 Uncertainties on signal events migration between categories

The impacts of theoretical and experimental uncertainties on the predicted contributions from the various Higgs boson production processes to each event category are summarized in the following.

10.2.2.1 Theory uncertainties

1. The uncertainty on the Higgs boson production cross section through gluon fusion in association with two or more jets is estimated by applying an extension of the Stewart–Tackmann method [113] to predictions made by the MCFM [65] generator. An uncertainty of 20% is assigned to the ggF component in the VBF loose, VBF tight, and VH hadronic categories. Since the VBF categories make use of the azimuthal angle between the diphoton and dijet systems, which is sensitive to the presence of a third jet, additional uncertainties are introduced for the ggF contribution in these categories using a technique described in Reference [16]. These uncertainties are found to be 25% and 52% for the VBF loose and tight categories, respectively.
2. The presence of additional hadronic activity from the underlying event may produce significant migrations of ggF events to the VBF and $t\bar{t}H$ hadronic categories. The uncertainties on the UE modeling are conservatively estimated as the full change in signal migration in MC simulation with and without the UE. The uncertainties are 5–6% of the 18–41% component of ggF in the VBF categories, and 60% of the 8–11% ggF contribution in the $t\bar{t}H$ hadronic category. In addition, the presence of the UE directly affects the $t\bar{t}H$ yield in the $t\bar{t}H$ hadronic and $t\bar{t}H$ leptonic categories by 11% and 3%, respectively. The differences

between the uncertainties for the 7 TeV and 8 TeV data are small. Tables 10.1 and 10.2 show details of the nominal yields of the signal processes in the event categories. The impacts of these uncertainties are small compared with the statistical uncertainties on the signal strengths for these categories.

3. The uncertainty on the modeling of the p_T spectrum of the Higgs boson for the ggF process can cause migrations of events between the low and the high p_{Tt} categories. The size of the effect has been checked using the HRES2.1 prediction by varying the renormalization, factorization, and two resummation scales. The uncertainties for the high p_{Tt} categories are estimated from the absolute values of the largest changes in the event categorization caused by the scale variations. Events in the low p_{Tt} categories are assigned an uncertainty according to the Stewart–Tackmann procedure. The size of the effect varies among categories; it is as large as about 24% in the high p_{Tt} categories.
4. The VBF selection uses angular variables $\Delta\phi_{jj}$ and η^* that involve the two leading jets, as discussed in Section 10.1.0.3. The second jet in the generation of ggF events by POWHEG-BOX+PYTHIA8 predominantly comes from the parton shower generated by PYTHIA8; therefore, the angular correlation between the two jets is not well modeled. The uncertainty due to this modeling is taken to be the difference in the event categorization caused by re-weighting the events in the POWHEG-BOX sample to reproduce the p_T spectrum of the Higgs boson predicted by MINLO HJJ, which models the angular correlation between the first and second jet produced in gluon fusion to NLO accuracy. The mis-modeling of $\Delta\phi_{jj}$ (η^*) typically changes the number of ggF events in the VBF tight and loose categories by at most 11.2% (6.6%) and 8.9% (4.8%), respectively.
5. Additional uncertainties are estimated for production processes contributing significantly to the $t\bar{t}H$ categories due to acceptance changes observed when varying the renormalization and factorization scales. The uncertainty on $t\bar{t}H$ production itself is 2% (1%) in the $t\bar{t}H$ leptonic (hadronic) category. An uncertainty of 50% is attributed to the ggF contribution in the $t\bar{t}H$ sensitive categories while an uncertainty of 4–8% is attributed to the WH , $tHqb$ and tHW contributions to account for the sensitivity of the acceptance to scale changes. The impact is independent for the three $t\bar{t}H$ and tH production processes, but coherent in

the two $t\bar{t}H$ event categories and for $\sqrt{s} = 7$ TeV and 8 TeV. In addition, the uncertainties on the ggF , VBF and WH contributions to the $t\bar{t}H$ categories are assumed to be 100% to account for the uncertainty on the heavy flavor (HF) fraction in these production processes. The overall impact of these large uncertainties on $\mu_{t\bar{t}H}$ is about 10% (and much less for the other signal strength measurements), due to the small contributions from ggF , VBF and WH production to the $t\bar{t}H$ categories (Tables 10.1 and 10.2).

10.2.2.2 Experimental uncertainties

The following potential sources of signal migration between categories caused by experimental effects are investigated.

1. Uncertainties related to jet and E_T^{miss} reconstruction affect the predicted distributions of signal events from the various production modes among the categories. The effect of the uncertainty on the jet energy scale, jet energy resolution and jet vertex fraction is estimated by varying individually each component of the uncertainties [117]. The effect of the E_T^{miss} energy scale and resolution uncertainty is estimated by varying independently the uncertainty in the energy scale and resolution of each type of physics object entering the calculation of E_T^{miss} as well as the uncertainty on the scale and resolution of the soft term [99]. There are 20 and 5 uncorrelated components that account for the jet- and E_T^{miss} -related uncertainties, respectively. Tables 10.6 and 10.7 show the impact of the jet and E_T^{miss} uncertainties. To simplify the presentation of the results, categories and processes for which each source of uncertainty has a similar impact are merged. These uncertainties are fully correlated between the 7 TeV and 8 TeV datasets.
2. The impact of the uncertainty in the b -tagging efficiency on the migration of events to and from the $t\bar{t}H$ categories is decomposed into 10 (3) independent contributions in the 8 TeV (7 TeV) data analysis. The uncertainty on the $t\bar{t}H$ yield in the $t\bar{t}H$ categories from the uncertainty on the b -tagging efficiency ranges from 1 to 3%. The uncertainties affecting other production processes that have the largest impact on the yield in the $t\bar{t}H$ categories are 20-30% of the ggF component in the hadronic category and 6-7% to the WH contribution in the leptonic channel.

3. The total impact of the lepton reconstruction, identification and isolation uncertainties on any of the selection efficiencies and event fractions of the signal production processes for the event categories in Tables 10.1 and 10.2 is found to be below 1%.

Category	ggF	VBF	$t\bar{t}H$	$WH+ZH$
Central+Forward low p_{Tt}	0.1	2.9	4.0	0.1
Central+Forward high p_{Tt}	1.1	4.5	3.5	1.4
VBF loose	12	4.4	7.6	13
VBF tight	13	9.1	6.3	17
VH hadronic	2.8	4.1	9.5	2.5
VH E_T^{miss}	2.6	9.0	1.2	0.2
VH one-lepton	4.9	6.2	2.8	0.5
VH dilepton	0	0	5.1	1.0
$t\bar{t}H$ hadronic	11	21	7.3	22
$t\bar{t}H$ leptonic	37	7.7	0.5	7.4

Table 10.6: Relative uncertainties (in %) on the Higgs boson signal yield in each category of the *Coupling Analysis* and for each production process induced by the combined effects of the systematic uncertainties on the jet energy scale, jet energy resolution and jet vertex fraction. These uncertainties are approximately the same for the 7 TeV and the 8 TeV data.

10.2.3 Uncertainties on signal mass resolution and mass scale

10.2.3.1 Diphoton mass resolution

The precise determination of the uncertainty on the signal strengths due to the diphoton mass resolution is a key point in this analysis. It defines the range over which the signal model width is allowed to change in the fit, thus directly affecting the estimation of the number of signal events. The energy resolution and its uncertainty for photons are estimated by extrapolating from the ones for electrons. The electron energy resolution and its uncertainty are measured in data using $Z \rightarrow e^+e^-$ events that, however, can only provide constraints for electrons with $p_T \simeq 40$ GeV. The extrapolation from electrons to photons and to different energy ranges relies on an accurate

Category	$ggF+VBF$	$t\bar{t}H$	WH	ZH
Untagged	0.0	0.2	0.1	0.2
VBF loose	0.0	1.0	0.1	0.2
VBF tight	0.0	2.7	1.1	0.0
VH hadronic	0.0	0.7	0.0	0.1
VH E_T^{miss}	35	1.1	1.3	0.9
VH one-lepton	4.5	0.6	0.4	4.0
VH dilepton	0.0	2.0	0.0	0.1
$t\bar{t}H$ hadronic	0.0	0.0	0.0	0.0
$t\bar{t}H$ leptonic	1.9	0.1	1.0	3.0

Table 10.7: Relative uncertainties (in %) on the Higgs boson signal yield in each category of the *Coupling Analysis* and for each production process induced by systematic uncertainty on the E_T^{miss} energy scale and resolution. The uncertainties, which are approximately the same for the 7 TeV and 8 TeV data, are obtained by summing in quadrature the impacts on the signal yield of the variation of each component of the E_T^{miss} energy scale within its uncertainty.

modeling of the resolution in the detector simulation. In the model used in this analysis, the total resolution is described in terms of four energy-dependent contributions [82]:

1. the asymptotic resolution at high energy, i.e. the constant term;
2. the intrinsic sampling fluctuations of the calorimeter;
3. the effect of passive material upstream of the calorimeter;
4. the electronic and pileup noise.

The effects on the various categories due to the four contributions to the uncertainty in the mass resolution are summarized in Table 10.8 for the 8 TeV data: the typical relative uncertainty on the diphoton mass resolution obtained from the sum in quadrature of these contributions is 10 – 15% for $m_H \simeq 125$ GeV. The uncertainties for the 7 TeV data are very similar except for the reduced size of the pileup contribution, which ranges from 0.9% to 1.4%. These four contributions are

uncorrelated while each contribution affects both of the parametric width parameters σ_{CB} and σ_{GA} in the signal model (Section 7.1) for all the categories and for both the 7 TeV and 8 TeV data coherently.

Category	Constant term	Sampling term	Material modeling	Noise term
Central low p_{Tt}	7.5	2.6	4.9	2.6
Central high p_{Tt}	9.6	5.6	6.2	1.7
Forward low p_{Tt}	9.9	1.3	6.0	2.1
Forward high p_{Tt}	12	2.8	7.8	1.9
VBF loose	9.4	2.6	6.0	2.1
VBF tight	10	3.8	6.5	2.1
VH hadronic	11	4.0	7.2	1.6
VH $E_{\text{T}}^{\text{miss}}$	11	3.6	7.4	1.7
VH one-lepton	9.8	2.8	6.3	2.1
VH dilepton	9.5	2.7	6.2	2.1
$t\bar{t}H$ hadronic	9.6	3.6	6.3	1.9
$t\bar{t}H$ leptonic	9.5	3.4	6.2	2.1

Table 10.8: Systematic uncertainties on the diphoton mass resolution for the 8 TeV data (in %) due to the four contributions described in the text. For each category of the *Coupling Analysis*, the uncertainty is estimated by using a simulation of the Higgs boson production process which makes the largest contribution to the signal yield.

10.2.3.2 Diphoton mass scale

The uncertainties on the diphoton mass scale affect the position of the signal mass peak through variations of the peak of the Crystal Ball (μ_{CB}) and Gaussian (μ_{GA}) components of the signal model (Section 7.1). The dominant systematic uncertainties on the position of the mass peak arise from uncertainties on the photon energy scale. These uncertainties, discussed in detail in References. [82, 118], are propagated to the diphoton mass distribution in the signal model for each

of the twelve categories. The total uncertainty on the position of the mass peak from the photon energy scale systematic uncertainties ranges from 0.18% to 0.31% depending on the category. A second contribution, varying from 0.02% to 0.31%, comes from the choice of the background model and is evaluated using the technique presented in Reference [118]. Finally, the systematic uncertainty on the mass scale related to the reconstruction of the diphoton vertex is estimated to be 0.03% for all the categories. As to be discussed in Section 10.3, the uncertainty on the diphoton mass scale is expected to flatten the dependence of μ as a function of m_H in the region around the true value of m_H .

10.2.4 Uncertainties on background model

The background parameterizations are selected using MC samples or control samples of data as described in the following.

For the four untagged, the two VBF, the VH hadronic and E_T^{miss} categories, the background parameterizations are tested with a mixture of $\gamma\gamma$, γj and γj samples with detector effect smearing as mentioned in Chapter 4. They are mixed based on composition studies based on a double two-dimensional sideband method as discussed in Section 6.2, and the overall normalization of each template is determined to be the observed number of events summarized in Table 10.3.

For the VH E_T^{miss} category, since the effect of the E_T^{miss} cut on the background shape is found to be negligible, it is not applied to the MC events. The background samples for the VH one-lepton category are also obtained from the MC $\gamma\gamma$ and γj events introduced previously, where one jet is treated as a lepton for the category selection.

An example of the diphoton invariant mass distributions in data and a MC background sample is shown in Figure 10.6 (a) for the Central low p_{Tt} category. For each category, the simulation describes the distributions of the data sufficiently well (apart from the signal region $m_{\gamma\gamma} \sim 125$ GeV) to be used to select the parameterization of the background model and to assess the corresponding systematic uncertainty on the signal yield.

A sample of fully simulated $Z\gamma\gamma$ events is used for the VH dilepton category since the contributions from $Z\gamma$ +jets and Z +jets events are estimated to be negligible after the event selection.

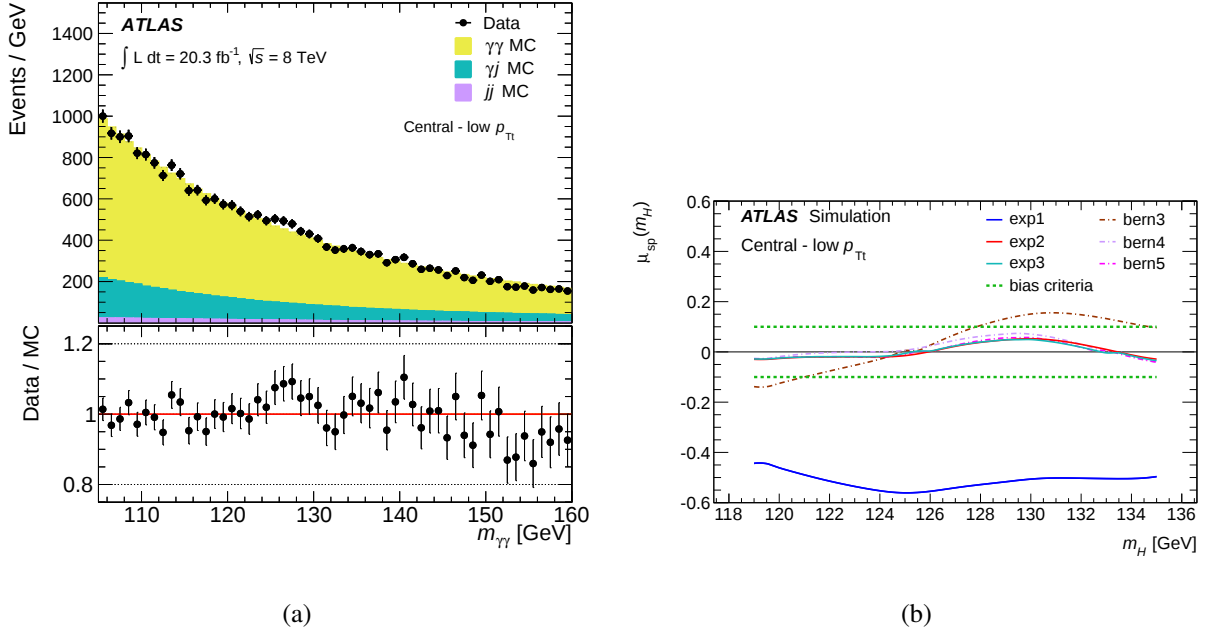


Figure 10.6: (a) The distributions of diphoton invariant mass $m_{\gamma\gamma}$ in the untagged Central low p_{Tt} category in data (points), and simulation samples for the $\gamma\gamma$, γj and jj components of the continuum background (shaded cumulative histograms). The lower plot shows the ratio of data to simulation. (b) Ratio of the fitted number of signal events to the number expected for the Standard Model $\mu_{sp}(m_H)$ as a function of the test mass m_H for the untagged Central low p_{Tt} category. A single fit per value of m_H is performed on the representative pure MC background sample described in the text with signal plus a variety of background parameterizations (exp1, exp2, exp3 for the exponentials of first, second or third-order polynomials, respectively, and bern3, bern4, bern5 for third, fourth and fifth-order Bernstein polynomials, respectively). The bias criteria discussed in Section 7.2 are indicated by the dashed lines.

For the $t\bar{t}H$ categories, the background parameterizations are tested on data control samples obtained by inverting photon identification criteria, isolation and the b -tagging, replacing the electron(s) with jet(s) and/or loosening the requirement on the number of jets. Multiple background templates are prepared and two typical ones are shown in Figure 10.7.

The selection of the parameterization for the background model follows the procedure discussed in Section 7.2. The fit range for the distributions of $m_{\gamma\gamma}$ is from 105 GeV to 160 GeV for all categories. The spurious signals are evaluated in the mass range between 119 GeV and

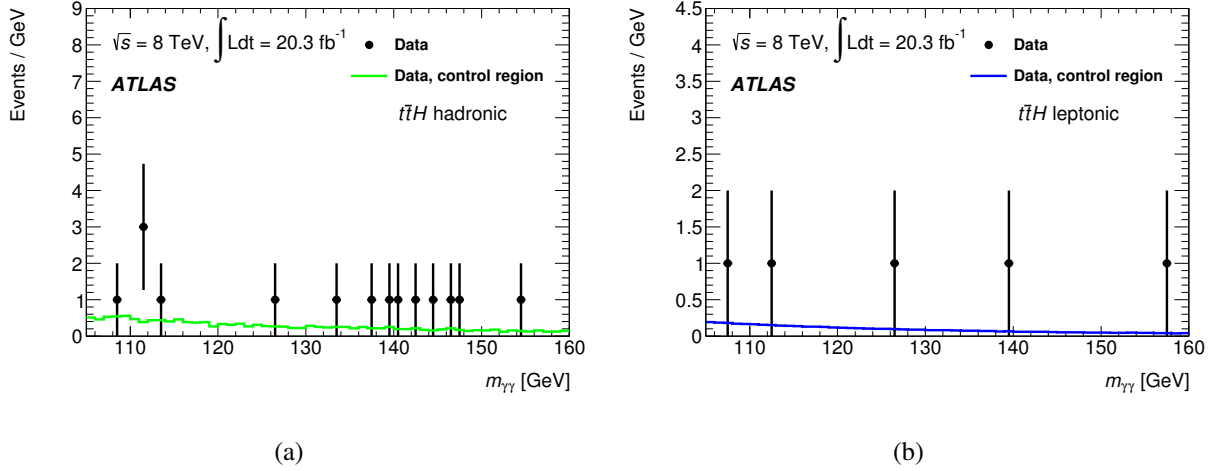


Figure 10.7: Comparison of the $m_{\gamma\gamma}$ distributions in data in the signal and control regions for $t\bar{t}H$ hadronic and $t\bar{t}H$ leptonic categories. The background shapes are extracted from data in control regions obtained by removing b-tagging requirement, relaxing the number of jets requirement and replacing leptons with jets (in $t\bar{t}H$ leptonic category only).

135 GeV, which was decided a priori to cover a region of approximately five times the expected signal mass resolution on either side of the value of m_H measured by ATLAS in the $H \rightarrow \gamma\gamma$ channel [119]. The spurious signal normalized to the expected number of signal events is shown in Figure 10.6 (b) for different candidate background models as functions of the test mass m_H for the Central low p_{Tt} category. The candidate parameterizations include polynomials on the index of the exponential and Bernstein polynomials [114]. The bands representing the criteria are also shown.

As shown in Table 10.9, all the categories in the *Coupling Analysis* can be modeled either by exponential or exponential of second-order polynomial. The former works for categories with low statistics (all the non- ggF categories and Central high p_{Tt} category), while the latter is only need for high statistics categories. The largest spurious signal in the mass range between 119 GeV and 135 GeV of a chosen parameterization is assigned as the systematic uncertainty for background modeling.

For $VH E_T^{\text{miss}}$, VH one-lepton, VH dilepton and both $t\bar{t}H$ categories, the expected number of data events in the 7 TeV data sideband has very high chance to be less than two (an exponential function needs to be constrained by at least two events). To ensure the 7 TeV background modeling

is working properly, the exponential slopes in these categories are correlated between the 7 TeV and 8 TeV. In $t\bar{t}H$ categories additional systematic uncertainty is assigned on this assumption.

Category	Model	$\sqrt{s} = 7 \text{ TeV}$		$\sqrt{s} = 8 \text{ TeV}$	
		N_{spur}	μ_{spur}	N_{spur}	μ_{spur}
Central low p_{Tt}	Exp. of 2nd order pol.	1.1	0.041	6.7	0.050
Central high p_{Tt}	Exponential	0.1	0.029	0.4	0.036
Forward low p_{Tt}	Exp. of 2nd order pol.	0.6	0.016	7.0	0.034
Forward high p_{Tt}	Exp. of 2nd order pol.	0.3	0.088	1.2	0.073
VBF loose	Exponential	0.2	0.091	1.3	0.14
VBF tight	Exponential	< 0.1	0.031	0.3	0.054
VH hadronic	Exponential	0.1	0.14	0.5	0.14
VH E_T^{miss}	Exponential	0.1	0.18	0.1	0.11
VH one-lepton	Exponential	< 0.1	0.094	0.1	0.064
VH dilepton	Exponential	< 0.1	0.080	< 0.1	0.08
$t\bar{t}H$ hadronic	Exponential	0.1	0.86	0.2	0.49
$t\bar{t}H$ leptonic	Exponential	< 0.1	0.10	0.2	0.28

Table 10.9: List of the functions chosen to model the background distributions of $m_{\gamma\gamma}$ in the *Coupling Analysis*, and the associated systematic uncertainties on the signal amplitudes in terms of spurious signal (N_{spur}) and its ratio to the predicted number of signal events in each category (μ_{spur}) for the twelve categories and the 7 TeV and 8 TeV datasets.

10.3 Results

10.3.1 Diphoton invariant mass spectra

The observed diphoton invariant mass $m_{\gamma\gamma}$ distribution for the sum of the 7 TeV and 8 TeV data is shown in Figure 10.8 for the inclusive case. It is also shown for untagged (Figure 10.9), VBF (Figure 10.10), VH (Figure 10.11) and $t\bar{t}H$ (Figure 10.12) categories. The results of signal plus background fits to these spectra with m_H set to 125.4 GeV are shown together with the separate

signal and background components. Both the signal plus background and background-only curves reported here are obtained from the sum of the individual curves in 7 TeV and 8 TeV categories.

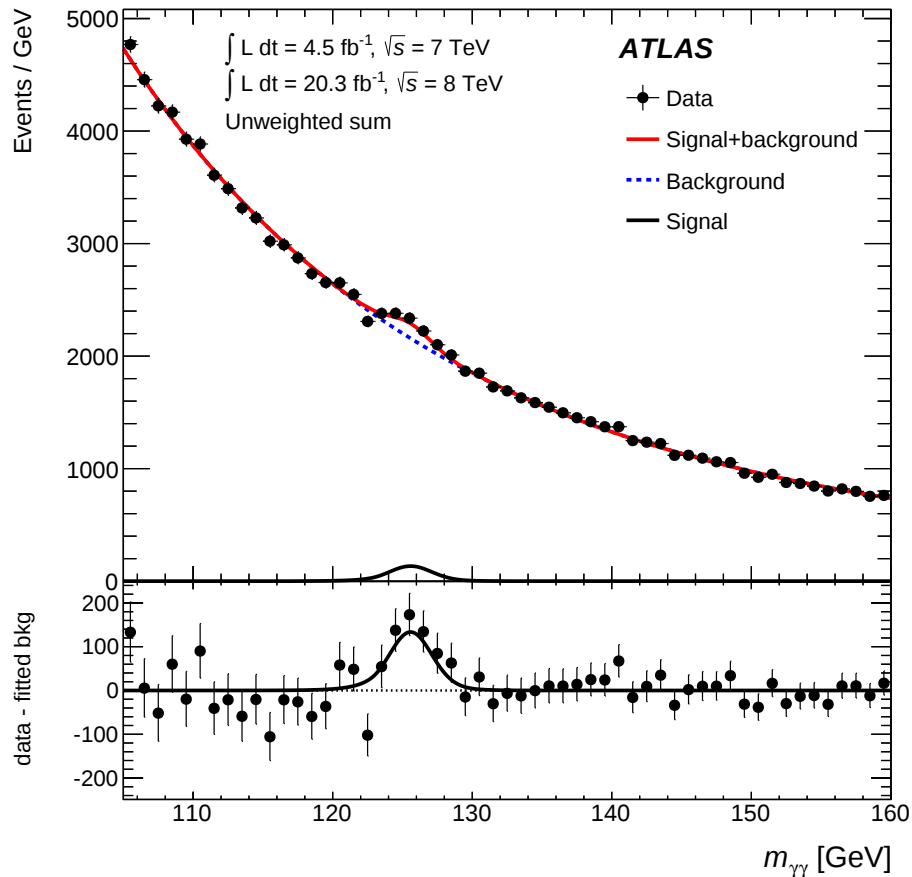


Figure 10.8: Distribution of the invariant mass of diphoton candidates after all selections in the *Coupling Analysis* for the combined 7 TeV and 8 TeV data sample. The solid red curve shows the fitted signal plus background model where the Higgs boson mass is fixed at 125.4 GeV. The background component of the fit is shown with the dotted blue curve. The signal component of the fit is shown with the solid black curve. Both the signal plus background and background-only curves reported here are obtained from the sum of the individual curves in 7 TeV and 8 TeV categories. The bottom panel shows the data relative to the background component of the fitted model.

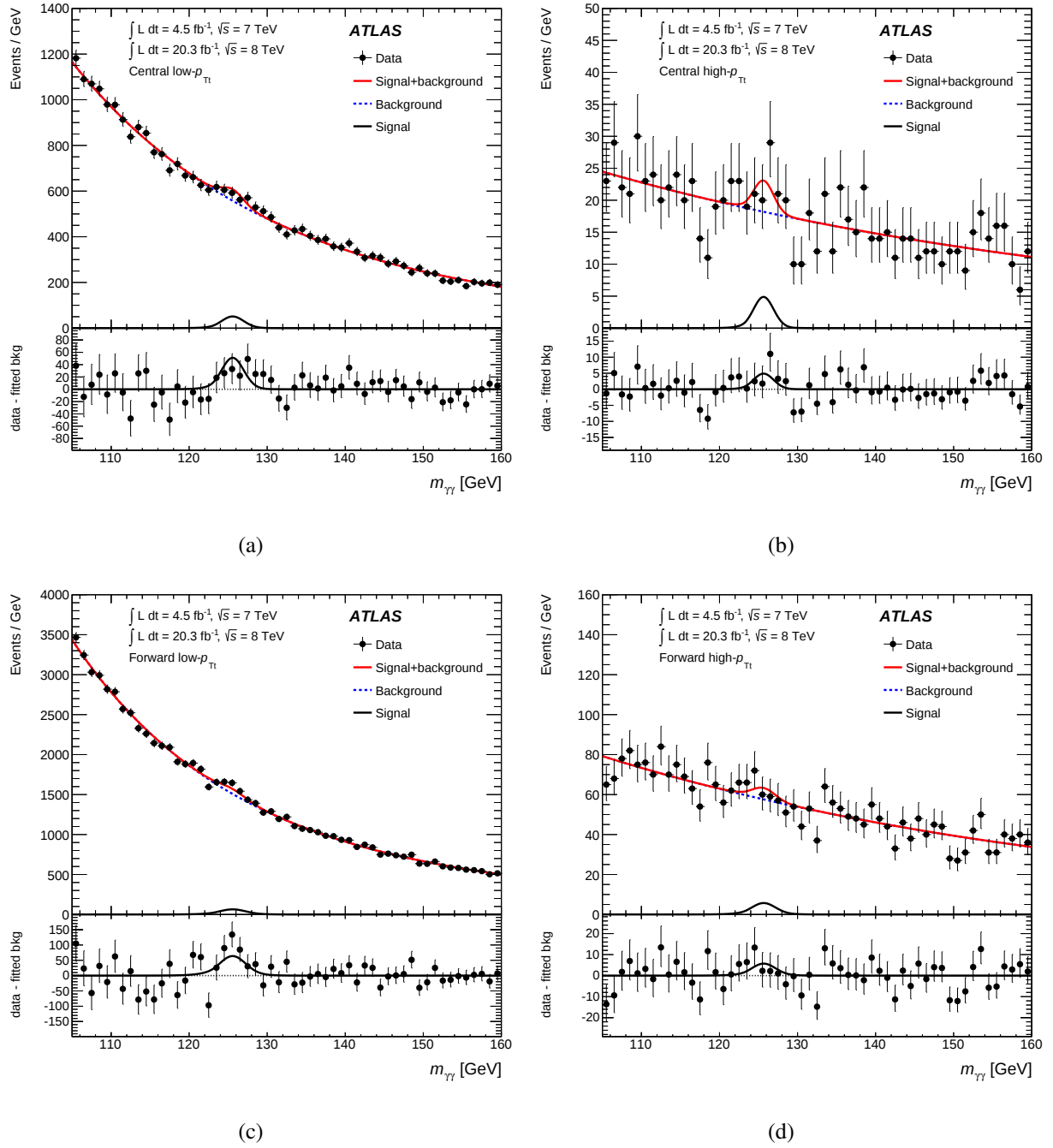


Figure 10.9: Diphoton invariant mass spectra observed in the 7 TeV and 8 TeV data in the untagged categories of the *Coupling Analysis*: Central low p_{Tt} (a), Central high p_{Tt} (b), Forward low p_{Tt} (c), and Central high p_{Tt} (d).

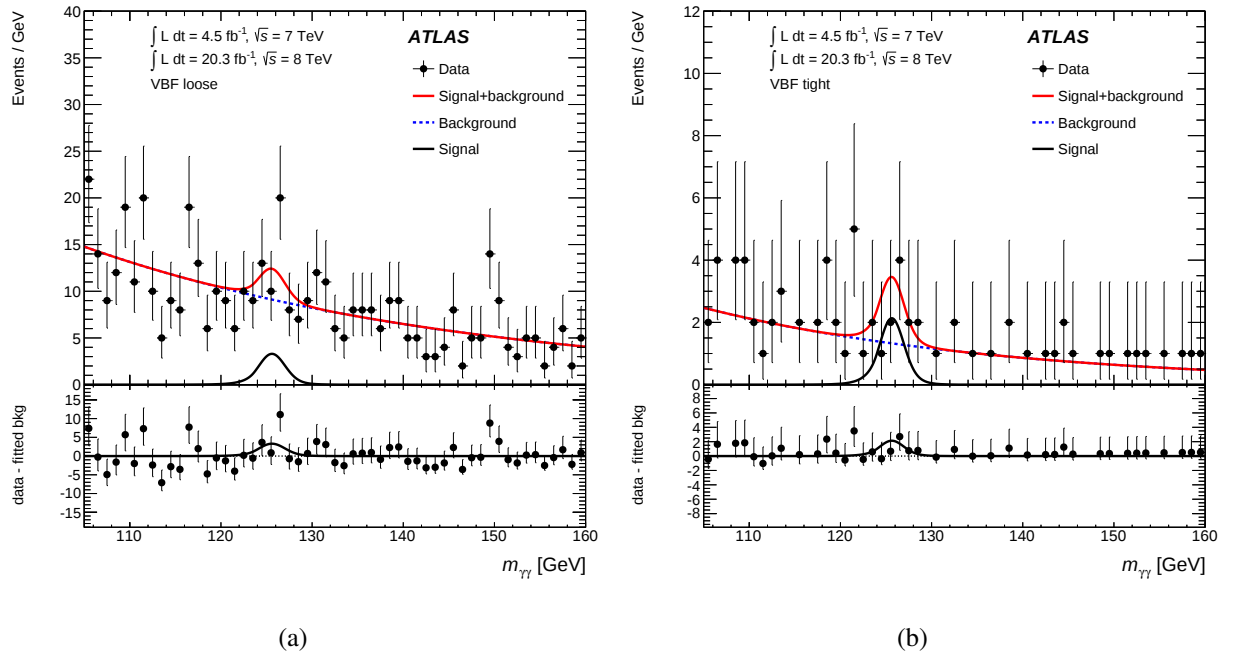


Figure 10.10: Diphoton invariant mass spectra observed in the 7 TeV and 8 TeV data in the VBF categories of the *Coupling Analysis*: VBF loose (a) and VBF tight (b).

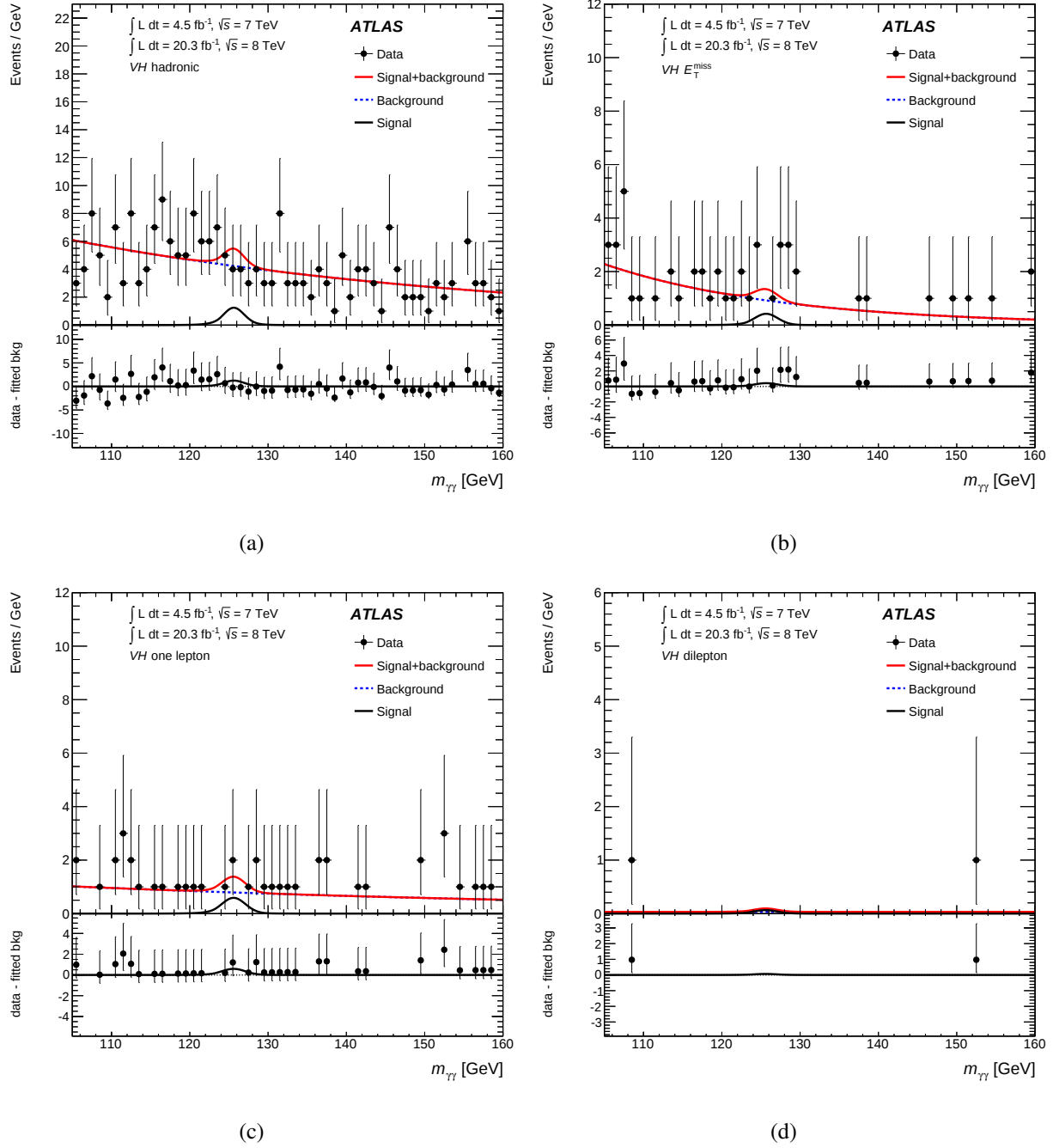


Figure 10.11: Diphoton invariant mass spectra observed in the 7 TeV and 8 TeV data in the VH categories of the *Coupling Analysis*: VH hadronic (a), $VH E_T^{\text{miss}}$ (b), VH one-lepton (c) and VH dilepton (d).

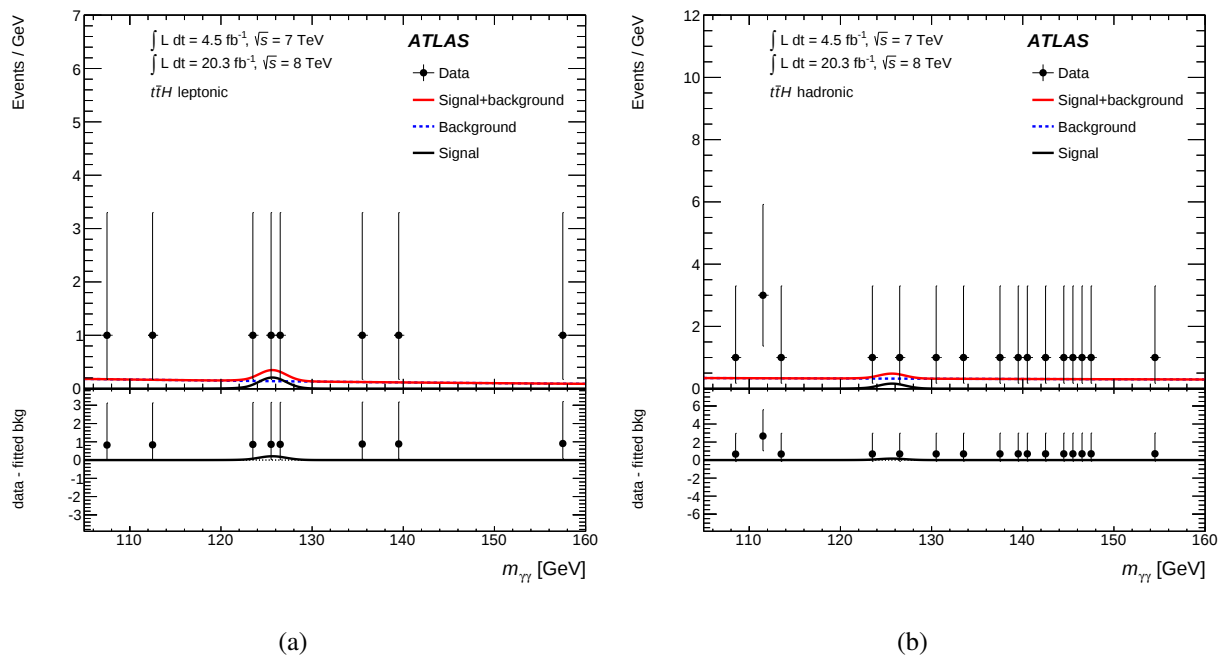


Figure 10.12: Diphoton invariant mass spectra observed in the 7 TeV and 8 TeV data in the $t\bar{t}H$ categories of the *Coupling Analysis*: $t\bar{t}H$ leptonic (a) and $t\bar{t}H$ hadronic.

10.3.2 Signal strength measurements

The signal strengths are measured with the procedure described in Chapter 8. The profile of the negative log-likelihood ratio of the combined signal strength μ for $m_H = 125.4$ GeV is shown in Figure 10.13.

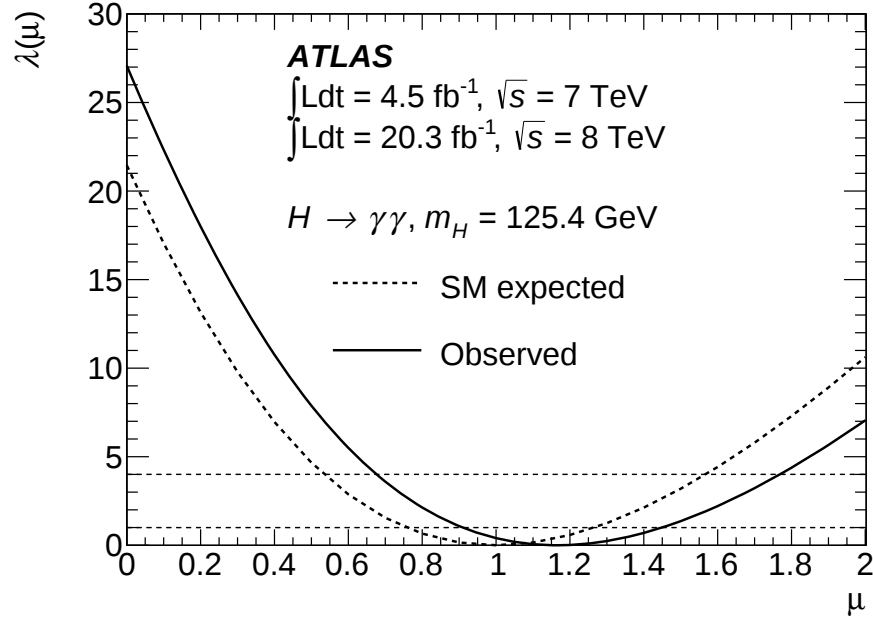


Figure 10.13: The profile of the negative log-likelihood ratio $\lambda(\mu)$ of the combined signal strength μ for $m_H = 125.4$ GeV. The observed result is shown by the solid curve, the expectation from the Standard Model by the dashed curve. The intersections of the solid and dashed curves with the horizontal dashed line at $\lambda(\mu) = 1$ indicate the 68% confidence intervals of the observed and expected results, respectively.

The local significance of the observed combined excess of events is 5.2σ (4.6σ expected). The combined signal strength is measured to be

$$\begin{aligned} \mu &= 1.17 \pm 0.27 \\ &= 1.17 \pm 0.23 \text{ (stat.) } {}^{+0.10}_{-0.08} \text{ (syst.) } {}^{+0.12}_{-0.08} \text{ (theory)} \end{aligned} \tag{10.1}$$

The signal strengths measured in the individual event categories are shown in Figure 10.14 (a). The signal strengths measured in the four groups of categories based on production modes described in Section 10.1 are presented in Figure 10.14 (b). All of the individual and grouped signal strengths are compatible with the combined signal strength.

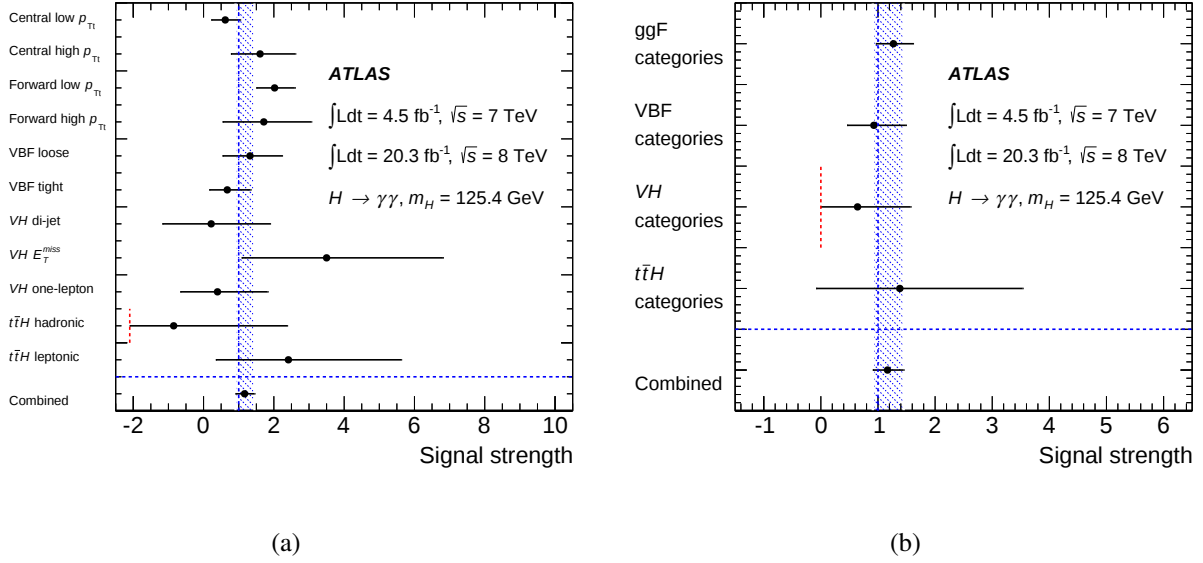


Figure 10.14: The signal strength for a Higgs boson of mass $m_H = 125.4$ GeV decaying via $H \rightarrow \gamma\gamma$ as measured (a) in the individual categories of the *Coupling Analysis*, and (b) in groups of categories sensitive to individual production modes for the combination of the 7 TeV and 8 TeV data together with the combined signal strength. The vertical hatched band indicates the 68% confidence interval of the combined signal strength. The vertical dashed line at unity indicates the Standard Model expectation. The vertical dashed red line indicates the limit below which the fitted signal plus background mass distribution for the $t\bar{t}H$ hadronic category becomes negative for some mass in the fit range. The VH dilepton category is not shown because with only two events in the combined sample, the fit results are not meaningful.

The impacts of the main sources of systematic uncertainty presented in Section 10.2 on the combined signal strength parameter measurement are presented in Table 10.10. They are determined from the difference in quadrature between the nominal uncertainty and change in the 68% CL range on μ when the corresponding nuisance parameters are fixed to their best fit values. The sums of the squares of the theoretical uncertainties linked to the QCD scales, PDFs, and $H \rightarrow \gamma\gamma$

branching ratio account for approximately 50% of the square of the total systematic uncertainty. The dominant experimental uncertainty is from the photon energy resolution, which represents approximately 30% of the total systematic uncertainty (as above in terms of its contribution to the square of the total systematic uncertainty). In the fit to extract the signal strengths, the post-fit values of the most relevant nuisance parameters (those apart from the ones of the background model) do not show significant deviations from their prefit input values.

Uncertainty group	$\sigma_{\mu}^{\text{syst.}}$
Theory (yield)	0.09
Experimental (yield)	0.02
Luminosity	0.03
MC statistics	< 0.01
Theory (migrations)	0.03
Experimental (migrations)	0.02
Resolution	0.07
Mass scale	0.02
Background shape	0.02

Table 10.10: Main systematic uncertainties $\sigma_{\mu}^{\text{syst.}}$ on the combined signal strength parameter μ measured from the *Coupling Analysis*. The values for each group of uncertainties are determined by subtracting in quadrature from the total uncertainty the change in the 68% confidence level range on μ when the corresponding nuisance parameters are fixed to their best fit values. The experimental uncertainty on the yield does not include the luminosity contribution, which is accounted for separately.

In addition to the combined signal strength, the signal strengths of the primary production processes are determined by exploiting the sensitivities of the analysis categories to specific production processes, and found to be (see also Figure 10.15):

$$\begin{aligned}\mu_{ggF} &= 1.32 \pm 0.38 \\ &= 1.32 \pm 0.32 \text{ (stat.) } {}^{+0.13}_{-0.09} \text{ (syst.) } {}^{+0.19}_{-0.11} \text{ (theory)}\end{aligned}\quad (10.2)$$

$$\begin{aligned}\mu_{VBF} &= 0.8 \pm 0.7 \\ &= 0.8 \pm 0.7 \text{ (stat.) } {}^{+0.2}_{-0.1} \text{ (syst.) } {}^{+0.2}_{-0.3} \text{ (theory)}\end{aligned}\quad (10.3)$$

$$\begin{aligned}\mu_{WH} &= 1.0 \pm 1.6 \\ &= 1.0 \pm 1.5 \text{ (stat.) } {}^{+0.3}_{-0.1} \text{ (syst.) } {}^{+0.2}_{-0.1} \text{ (theory)}\end{aligned}\quad (10.4)$$

$$\begin{aligned}\mu_{ZH} &= 0.1 {}^{+3.7}_{-0.1} \\ &= 0.1 {}^{+3.6}_{-0.1} \text{ (stat.) } {}^{+0.7}_{-0.0} \text{ (syst.) } {}^{+0.1}_{-0.0} \text{ (theory)}\end{aligned}\quad (10.5)$$

$$\begin{aligned}\mu_{t\bar{t}H} &= 1.6 {}^{+2.7}_{-1.8} \\ &= 1.6 {}^{+2.6}_{-1.8} \text{ (stat.) } {}^{+0.6}_{-0.4} \text{ (syst.) } {}^{+0.5}_{-0.2} \text{ (theory)}\end{aligned}\quad (10.6)$$

In this measurement, both μ_{tH} and $\mu_{b\bar{b}H}$ are fixed to the SM expectations at unity. The correlation between the fitted values of μ_{ggF} and μ_{VBF} has been studied by still fixing both μ_{tH} and $\mu_{b\bar{b}H}$ to 1 and profiling² the remaining signal strengths μ_{ZH} , μ_{WH} , and $\mu_{t\bar{t}H}$. The best-fit values of μ_{ggF} and μ_{VBF} and the 68% and 95% CL contours are shown in Figure 10.16.

In order to test the production through VBF, VH and $t\bar{t}H$ processes independent of the $H \rightarrow \gamma\gamma$ branching ratio, the ratios μ_{VBF}/μ_{ggF} , μ_{VH}/μ_{ggF} , and $\mu_{t\bar{t}H}/\mu_{ggF}$ are fitted separately by fixing μ_{tH} and $\mu_{b\bar{b}H}$ to unit and profiling the remaining signal strengths. The measured ratios are

$$\mu_{VBF}/\mu_{ggF} = 0.6 {}^{+0.8}_{-0.5}, \quad (10.7)$$

$$\mu_{VH}/\mu_{ggF} = 0.6 {}^{+1.1}_{-0.6}, \quad (10.8)$$

$$\mu_{t\bar{t}H}/\mu_{ggF} = 1.2 {}^{+2.2}_{-1.4}. \quad (10.9)$$

They are not significantly different from zero, and are consistent with the SM predictions of unity. Likelihood scans of these ratios are presented in Figure 10.17.

² Profiling here means maximizing the likelihood with respect to all parameters apart from the parameters of interest μ_{ggF} and μ_{VBF} .

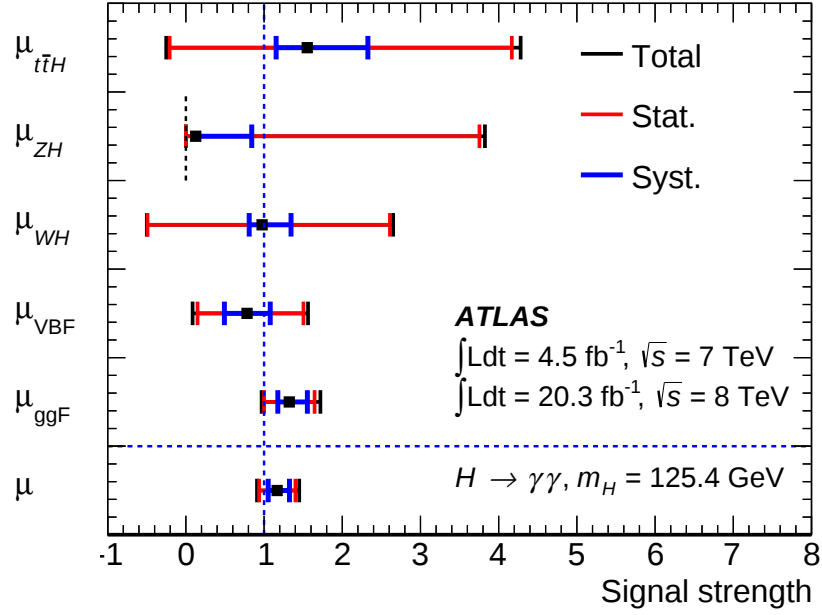


Figure 10.15: Measured signal strengths, for a Higgs boson of mass $m_H = 125.4$ GeV decaying via $H \rightarrow \gamma\gamma$, of the different Higgs boson production modes and the combined signal strength μ obtained with the combination of the 7 TeV and 8 TeV data in the *Coupling Analysis*. The vertical dashed line at unity indicates the Standard Model expectation. The vertical dashed line at the left end of the μ_{ZH} result indicates the limit below which the fitted signal plus background mass distribution becomes negative for some mass in the fit range.

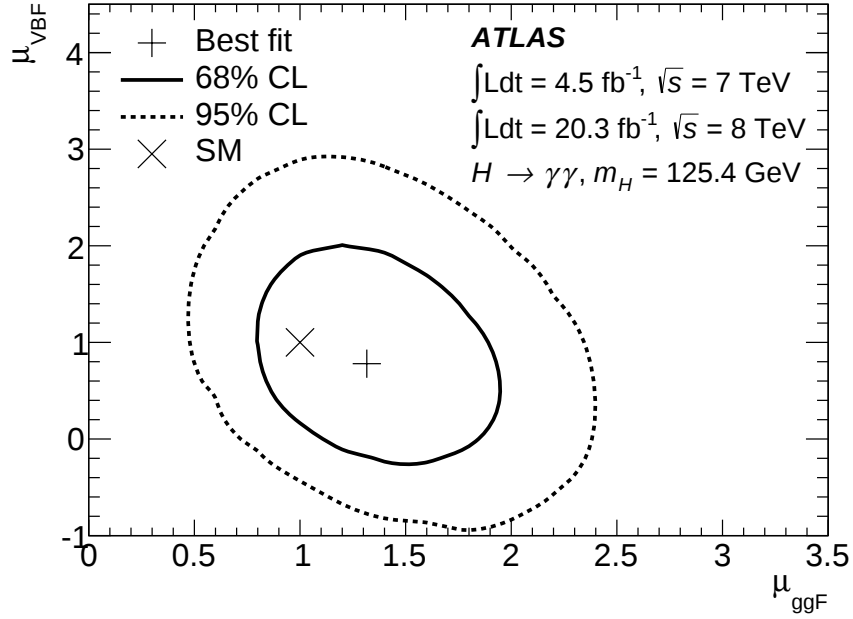


Figure 10.16: The two-dimensional best-fit value of $(\mu_{\text{VBF}}, \mu_{\text{ggF}})$ for a Higgs boson of mass $m_H = 125.4 \text{ GeV}$ decaying via $H \rightarrow \gamma\gamma$ when fixing both μ_{tH} and $\mu_{b\bar{b}H}$ to 1 and profiling all the other signal strength parameters in the *Coupling Analysis*. The 68% and 95% confidence level contours are shown with the solid and dashed lines, respectively. The result is obtained for $m_H = 125.4 \text{ GeV}$ and the combination of the 7 TeV and 8 TeV data.

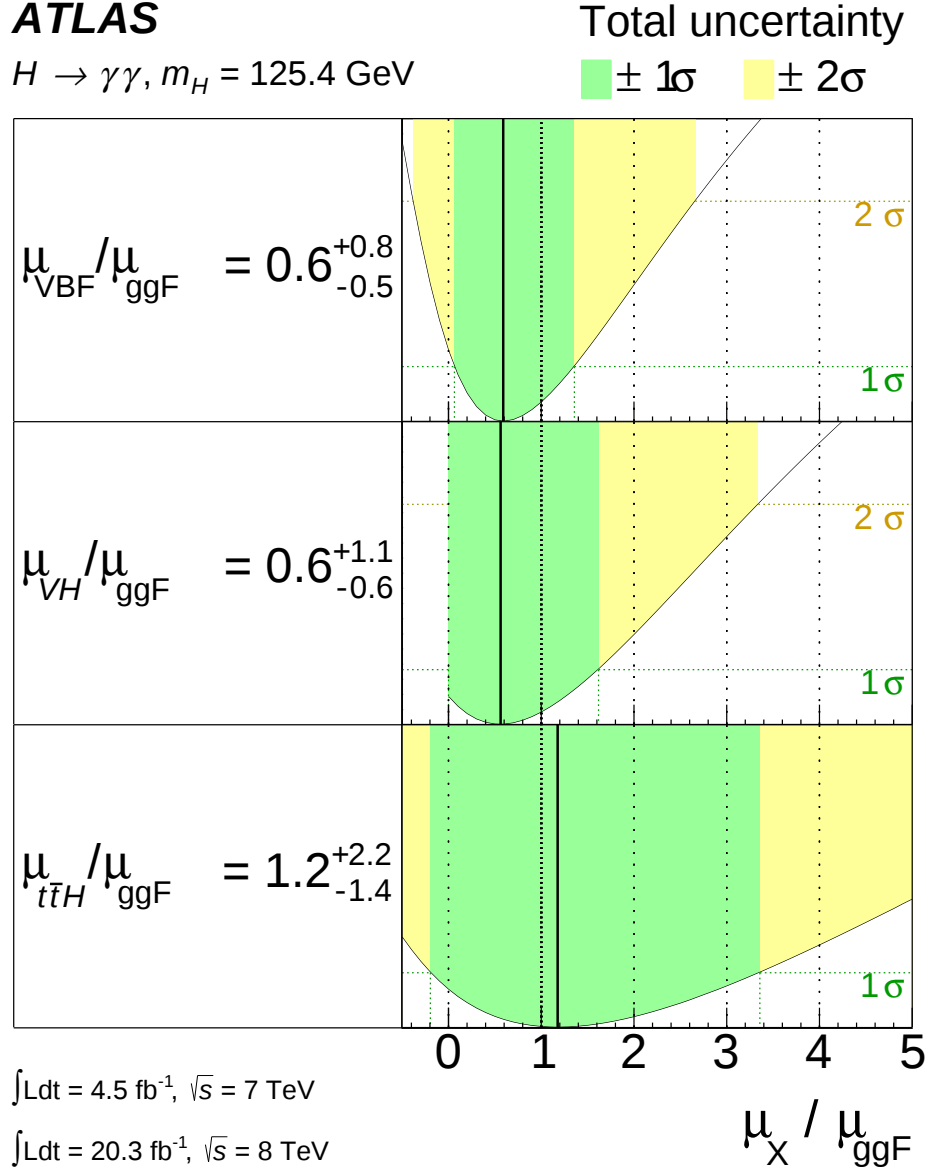


Figure 10.17: Measurements of the $\mu_{\text{VBF}}/\mu_{\text{ggF}}$, $\mu_{\text{VH}}/\mu_{\text{ggF}}$ and $\mu_{t\bar{t}H}/\mu_{\text{ggF}}$ ratios and their total errors for a Higgs boson mass $m_H = 125.4 \text{ GeV}$ in the *Coupling Analysis*. For a more complete illustration, the log-likelihood curves from which the total uncertainties are extracted are also shown: the best fit values are represented by the solid vertical lines, with the total $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties indicated by the dark- and light-shaded band, respectively. The likelihood curve and uncertainty bands for $\mu_{\text{VH}}/\mu_{\text{ggF}}$ stop at zero because below this the hypothesized signal plus background mass distribution in the VH dilepton channel becomes negative (unphysical) for some mass in the fit range.

10.3.3 Search for $t\bar{t}H$ production process and constraints on Yukawa coupling between top quark and Higgs boson

A search for the $t\bar{t}H$ production can be conducted in the $t\bar{t}H$ categories [41]. The CL_s -based 95% CL exclusion upper limits have been set for $t\bar{t}H$ production times $H \rightarrow \gamma\gamma$ branching ratio following the procedure discussed in Section 8.2. The results based on asymptotic assumptions are found to be consistent with limits derived from ensembles of pseudo-experiments. The observed and expected upper limits for $\mu_{t\bar{t}H}$ at $m_H = 125.4$ GeV are summarized in Figure 10.18 as well as in Table 10.11, where the expected limits assume $\mu_{t\bar{t}H} = 0$. The non- $t\bar{t}H$ Higgs boson production modes, including tH , are fixed to their SM expectations with corresponding theory and experimental uncertainties assigned. An upper limit of 6.7 times the SM cross section times $\text{BR}(H \rightarrow \gamma\gamma)$ is observed. Upper limits at 95% CL are also set on the signal strength of the sum of all $H \rightarrow \gamma\gamma$ processes, μ , and the observed (expected) limit is 5.7 (3.8).

	Observed limit	Expected limit	+2 σ	+1 σ	-1 σ	-2 σ
Combined (with systematics)	6.7	4.9	11.9	7.5	3.5	2.6
Combined (statistics only)		4.7	10.5	7.0	3.4	2.5
Leptonic (with systematics)	10.7	6.6	16.5	10.1	4.7	3.5
Leptonic (statistics only)		6.4	15.1	9.6	4.6	3.4
Hadronic (with systematics)	9.0	10.1	25.4	15.6	7.3	5.4
Hadronic (statistics only)		9.5	21.4	14.1	6.8	5.1

Table 10.11: Observed and expected 95% confidence level upper limits on the $t\bar{t}H$ production cross section times $\text{BR}(H \rightarrow \gamma\gamma)$ relative to the Standard Model prediction at $m_H = 125.4$ GeV. All other Higgs boson production cross sections, including the cross section for tH production, are set to their respective SM expectations. In addition, the expected limits corresponding to +2 σ , +1 σ , -1 σ , and -2 σ variations are shown. The expected limits are calculated for the case where $t\bar{t}H$ production is not present. The results are given for the combination of leptonic and hadronic categories with all systematic uncertainties included, and also for leptonic and hadronic categories separately. Expected limits are also derived for the case of statistical uncertainties only.

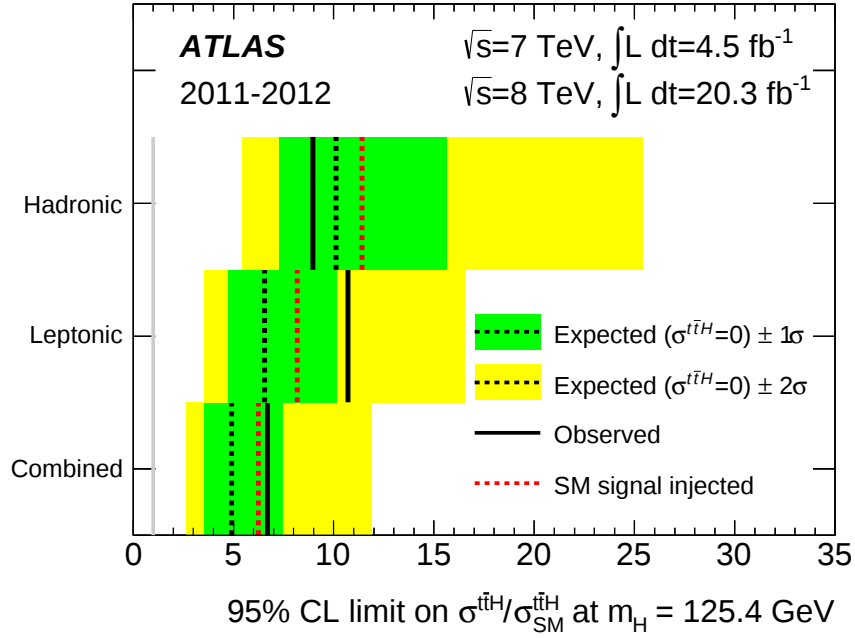


Figure 10.18: Observed and expected 95% confidence level upper limits on the $t\bar{t}H$ production cross section times $\text{BR}(H \rightarrow \gamma\gamma)$. All other Higgs boson production cross sections, including the cross section for tH production, are set to their respective Standard Model expectations. While the expected limits are calculated for the case where $t\bar{t}H$ production is not present, the lines denoted by “SM signal injected” show the expected 95% confidence level limits for a dataset corresponding to continuum background plus SM Higgs boson production. The limits are given relative to the SM expectations and at $m_H = 125.4 \text{ GeV}$.

These results are also interpreted as 95% CL limits on the coupling modifier κ_t ³ of the top–Higgs Yukawa coupling. Variations in κ_t not only change the production cross sections of the $t\bar{t}H$ and tH processes, but also affect the $\text{BR}(H \rightarrow \gamma\gamma)$, and the cross sections of the other Higgs production processes [16]. Figure 10.19 illustrates the dependence of the $t\bar{t}H$ and tH cross sections and of the $\text{BR}(H \rightarrow \gamma\gamma)$ on κ_t . For $\kappa_t = 0$, the $t\bar{t}H$ process is turned off, and the top quark contribution to tH production and to the loop-induced $H \rightarrow \gamma\gamma$ decay is removed, leaving mainly the contribution from W bosons. For values of $\kappa_t < 0$, on the other hand, the interference between contributions from W bosons and top quarks to tH production and to the $\text{BR}(H \rightarrow \gamma\gamma)$ becomes constructive, thus enhancing the two processes with respect to their respective SM expectations.

³ More details of the κ -framework can be found in Appendix A.

Accidental cancellations of the contributions of top quarks and W bosons to the loop-induced $H \rightarrow \gamma\gamma$ decay lead to a minimum of the $\text{BR}(H \rightarrow \gamma\gamma)$ around a value of $\kappa_t = +4.7$. The combined selection efficiency differs slightly for the three values of κ_t for which $tHqb$ and tHW MC samples were generated. From these, the efficiency at different values of κ_t in the range $[-3, +10]$ is calculated by combining re-weighted MC samples with $\kappa_t = +1, 0$ and -1 . The weight for each sample is assigned in such a way that the cross-section value from the combination follows the prediction shown in Figure 10.19. The largest relative difference with respect to the efficiency at $\kappa_t = +1$ over the entire range is found to be 14% (20%) for $tHqb$ (tHW) production.

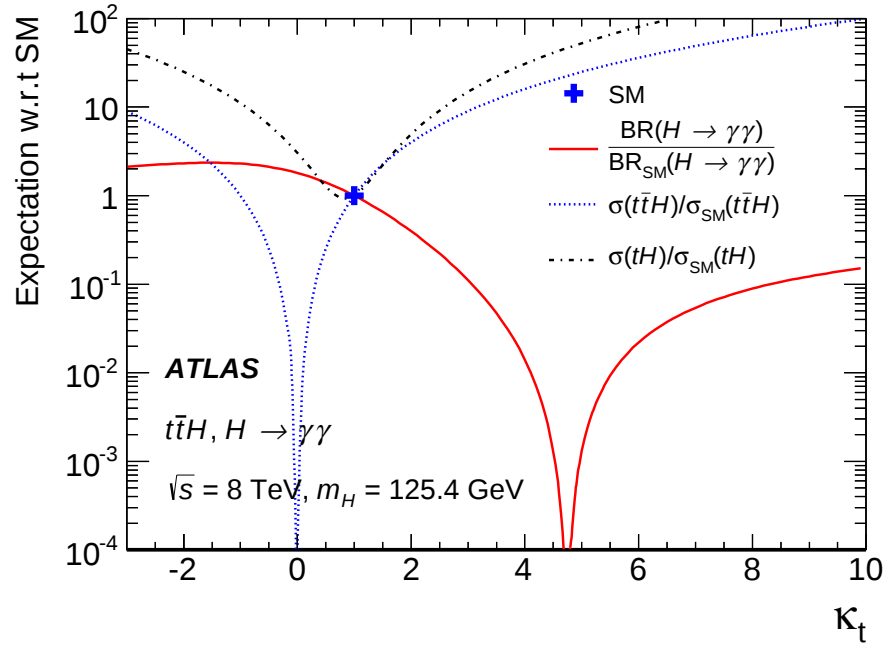


Figure 10.19: Production cross sections for $t\bar{t}H$ and tH divided by their Standard Model expectations as a function of the scale factor to the top quark-Higgs boson Yukawa coupling, κ_t . Production of tH comprises the $tHqb$ and tHW processes. Also shown is the dependence of the $\text{BR}(H \rightarrow \gamma\gamma)$ with respect to its SM expectation on κ_t .

All $H \rightarrow \gamma\gamma$ processes are considered and 95% CL limits are set on the total Higgs boson production cross section times $\text{BR}(H \rightarrow \gamma\gamma)$ with respect to the SM cross section for different values of κ_t . Coupling strengths other than κ_t are set to their respective SM values. The continuum background plus SM Higgs boson production ($\kappa_t = +1$) is taken as alternative hypothesis.

The observed and expected limits on κ_t at $m_H = 125.4$ GeV are summarized in Figure 10.20, where the observed (expected) lower and upper limits on κ_t at 95% CL are -1.3 and $+8.0$ (-1.2 and $+7.8$). The expected limits assume $\kappa_t = +1$. The form of the limit curve shown in Figure 10.20 is the result of the different dependencies of the different Higgs boson production processes as well as the $\text{BR}(H \rightarrow \gamma\gamma)$ on κ_t . The negative log-likelihood scan of κ_t is shown in Figure 10.21 and it is consistent with the SM expectation of unity. Although two different values of κ_t exist with the same total number of expected events, there are no double minima at zero shown in Figure 10.20 because different relative contributions from the Higgs boson production processes in different categories have lifted the degeneracy of the likelihood.

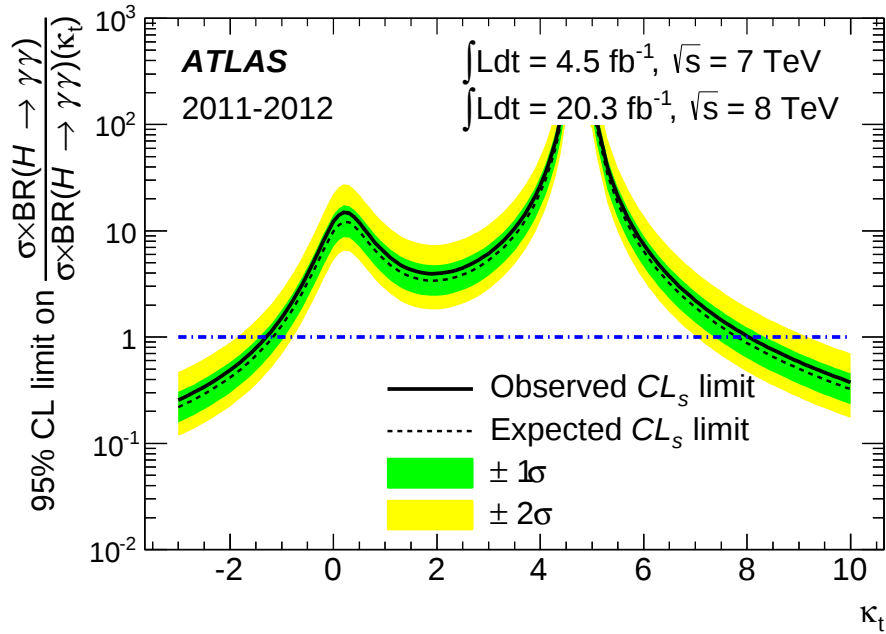


Figure 10.20: Observed and expected 95% confidence level upper limits on the inclusive Higgs production cross section with respect to the Standard Model cross section times $\text{BR}(H \rightarrow \gamma\gamma)$ for different values of κ_t at $m_H = 125.4$ GeV, where κ_t is the strength parameter for the top quark-Higgs boson Yukawa coupling. All Higgs boson production processes are considered for the inclusive production cross section. The expected limits are calculated for the case where $\kappa_t = +1$. The CL_s alternative hypothesis is given by continuum background plus Standard Model Higgs boson production.

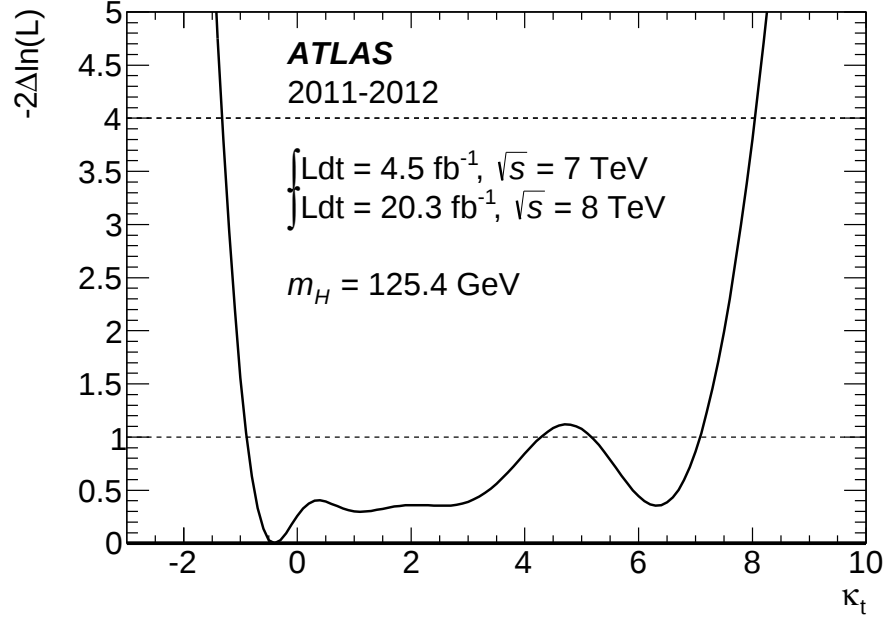


Figure 10.21: Negative log-likelihood scan of κ_t at $m_H = 125.4$ GeV, where κ_t is the strength parameter for the top quark–Higgs boson Yukawa coupling.

10.3.4 Combination with other decay channels

While more details from ATLAS–CMS combined measurement of Higgs boson couplings based on Run 1 data can be found in Reference [120], here two sets of results will be highlighted to demonstrate the power of the $H \rightarrow \gamma\gamma$ channel in combined Higgs boson coupling measurements. The κ -framework is intensively involved in the following results, and readers can find in Appendix A more discussions.

The first set of results are the 2D likelihood contours in the $(\kappa_\gamma, \kappa_g)$ plane, where κ_γ and κ_g are the coupling modifiers to the effective vertices $H \rightarrow \gamma\gamma$ and $gg \rightarrow H$, respectively. Since these two processes are loop-induced, in some BSM scenarios new particles may contribute to the loop processes. If it is assumed that all the couplings to SM particles are the same as in the SM, and the new particles are heavy such that there are no BSM decays of the Higgs boson, (i.e. only the ggF production and $H \rightarrow \gamma\gamma$ decay loops may be affected by the presence of additional particles), the results are shown in Figure 10.22. The data is in very good compatibility with the SM prediction at $\kappa_\gamma = 1$ and $\kappa_g = 1$.

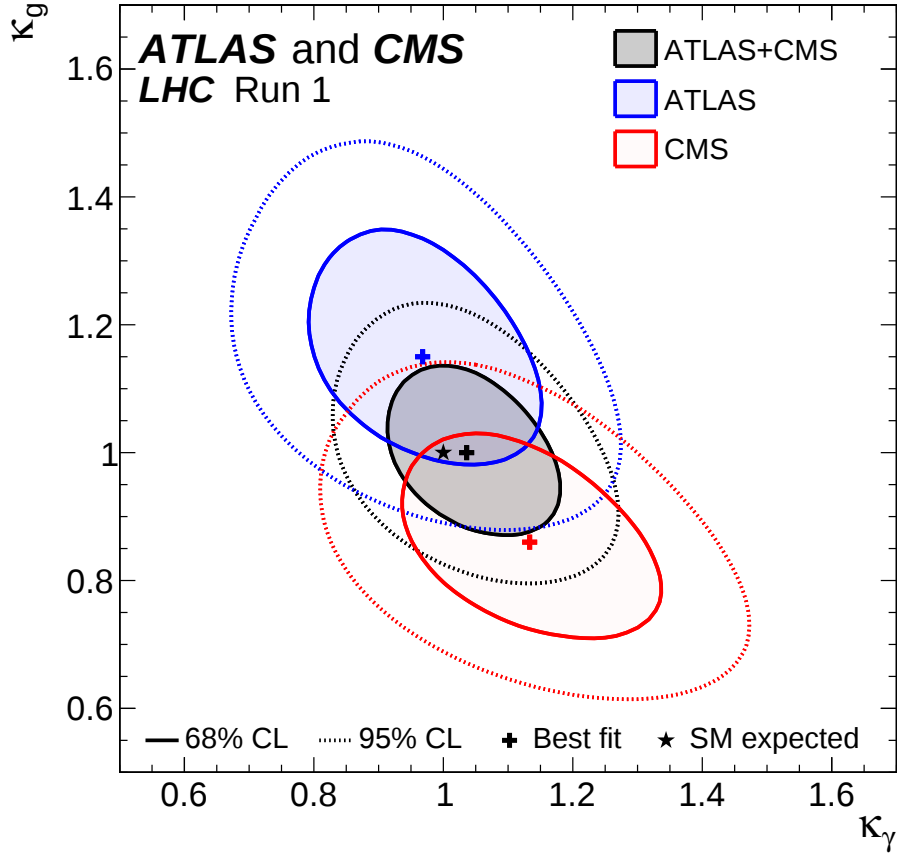


Figure 10.22: Negative log-likelihood contours at 68% and 95% confidence level in the $(\kappa_\gamma, \kappa_g)$ plane for the combination of ATLAS and CMS and for each experiment separately, as obtained from the fit to the parameterization constraining all the other coupling modifiers to their Standard Model values and assuming $\text{BR}_{\text{BSM}} = 0$.

The second set of results are the 2D likelihood contours in the (k_V, k_F) plane, where k_V and k_F are the coupling modifiers which are assumed to be uniformly applied to W/Z bosons (couplings originated from EWSB) and fermions (Yukawa couplings), respectively. Though the LHC experiments are not sensitive to the absolute sign of k_V and k_F , the ambiguity of their relative sign can be resolved through interference effects in $H \rightarrow \gamma\gamma$ decay, tH production (as discussed in Section 10.3.3), and $ggZH$ production (mentioned in Section 2.1). Assuming in addition that only SM particles couples to the Higgs sector, the results are shown in Figure 10.23. As one can see from Figure 10.23 (a), while other decay channels shown almost symmetric contour pattern between the k_F positive and negative branches, the $H \rightarrow \gamma\gamma$ channel uniquely shows obvious asymmetry,

suggesting that most of the sensitivity for lifting the sign degeneracy between k_V and k_F at 5σ level shown in Figure 10.23 (b) comes from this process.

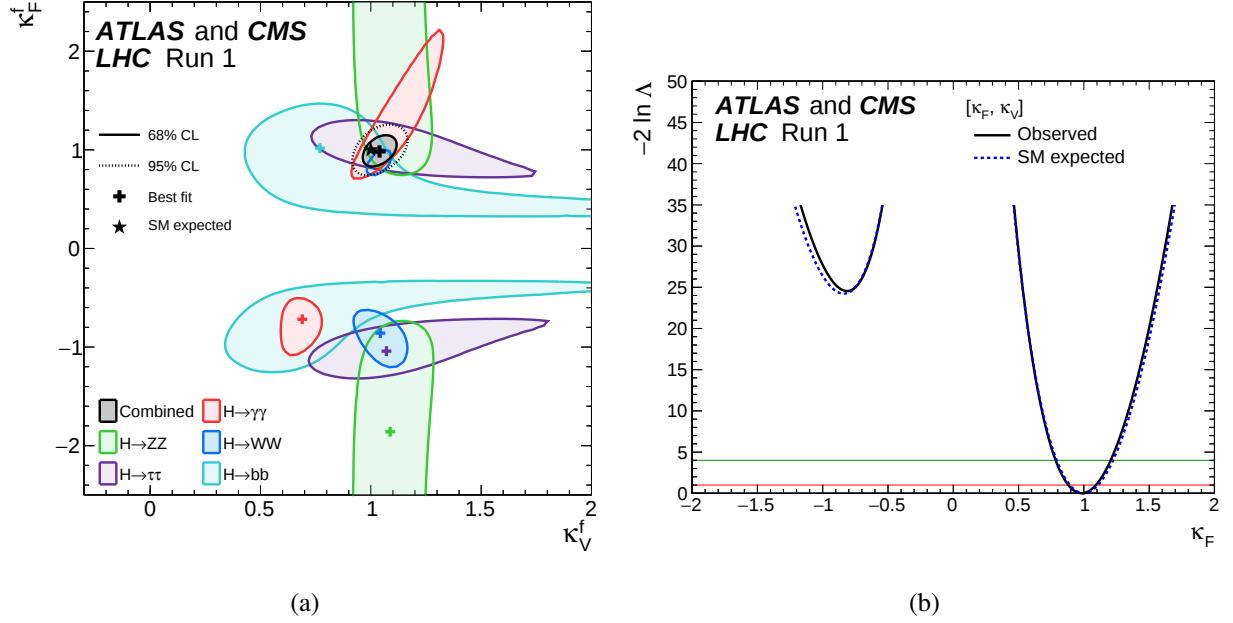


Figure 10.23: (a) Negative log-likelihood contours at 68% and 95% confidence level in the (κ_F^f, κ_V^f) plane for the combination of ATLAS and CMS and for the individual decay channels, as well as for their combination (κ_F versus κ_V shown in black), without any assumption about the relative sign of the coupling modifiers. (b) Observed (solid line) and expected (dashed line) negative log-likelihood scans for the global κ_F parameter, corresponding to the combination of all decay channels. The red (green) horizontal lines at the $-2\Delta \ln \Lambda$ value of 1 (4) indicate the value of the profile likelihood ratio corresponding to a 1σ (2σ) confidence level interval for the parameter of interest.

Chapter 11

Measurement of Higgs boson mass

11.1 Measurement in diphoton decay channel by ATLAS

In this section, a brief summary of the Higgs boson mass measurement in the $H \rightarrow \gamma\gamma$ channel based on LHC Run 1 dataset by the ATLAS experiment (*Mass Analysis*) will be given. Readers can refer to Reference [118] for more details.

The *Mass Analysis* is based on the same selected diphoton data sample used for the *Coupling Analysis* (Chapter 10). The categorization applied is optimized to reduce systematic uncertainty on the measurement of m_H . First of all, the events are divided into two groups, one for events with both photons unconverted, and the other for events with at least one converted photon. Each group is then classified according to the η of two photons into *Central* (both photon with $|\eta| < 0.75$), *Transition* (at least one photon with $1.3 < |\eta| < 1.75$), and *Rest* categories. Finally, the *Central* and *Rest* categories are further split into low p_{Tt} (< 70 GeV) and high p_{Tt} (otherwise) ones. In total ten categories are introduced to the 7 TeV and the 8 TeV data each.

The signal modeling has been discussed in Section 7.1. The background models are selected based on the same smeared MC used for the untagged categories of the *Coupling Analysis*. Exponential functional form has been chosen for the four high p_{Tt} categories with low statistics, whereas the rest six categories with high statistics are modeled by exponential of second-order polynomial.

The expected signal rate, mass resolution and background in each category is summarized in Table 11.1. Small differences in mass resolution arise from the differences in the effective constant term measured with $Z \rightarrow e^+e^-$ events and from the lower pile-up level in the 7 TeV data. In the 8 TeV data which dominate the sensitivity of the measurement, the expected resolution of the signal peak quantified by half of the smallest mass window which contains 68% of the total signal varies from 1.2 GeV to 2.4 GeV.

The systematic uncertainties affecting the mass measurement are summarized in Table 11.2. The total systematic uncertainty on the measured mass is 0.22%, dominated by the uncertainty on the photon energy scale.

The measured Higgs boson mass in the $H \rightarrow \gamma\gamma$ channel is

$$\begin{aligned} m_H &= 125.98 \pm 0.42(\text{stat}) \pm 0.28(\text{syst}) \text{ GeV} \\ &= 125.98 \pm 0.50 \text{ GeV}. \end{aligned} \tag{11.1}$$

The $H \rightarrow \gamma\gamma$ channel is combined with $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ channel (invariant mass spectrum shown in Figure 11.1) by ATLAS to give the ATLAS Run 1 Higgs boson mass measurement result used in the *Coupling Analysis*:

$$\begin{aligned} m_H &= 125.36 \pm 0.37(\text{stat}) \pm 0.18(\text{syst}) \text{ GeV} \\ &= 125.36 \pm 0.41 \text{ GeV}. \end{aligned} \tag{11.2}$$

In the combined measurement of the Higgs boson mass, the signal strength scale factors in each channel is floating independently as a nuisance parameter to reducing the dependence of the results on assumptions about the Higgs boson production cross section times decay branching ratio with the mass. Only $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ channels are included in the combined measurement because these two channels are the only ones which have both excellent mass resolution and sensitivity (same for the CMS measurements to be discussed). The interference effect between signal and continuum background in the $H \rightarrow \gamma\gamma$ channel has been discussed in Section 4.2 to be small. The effect in the $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ channel is expected to be even much smaller [121]. These interference effects are thus neglected in the combined measurement and also in the ATLAS–CMS combined measurement to be discussed.

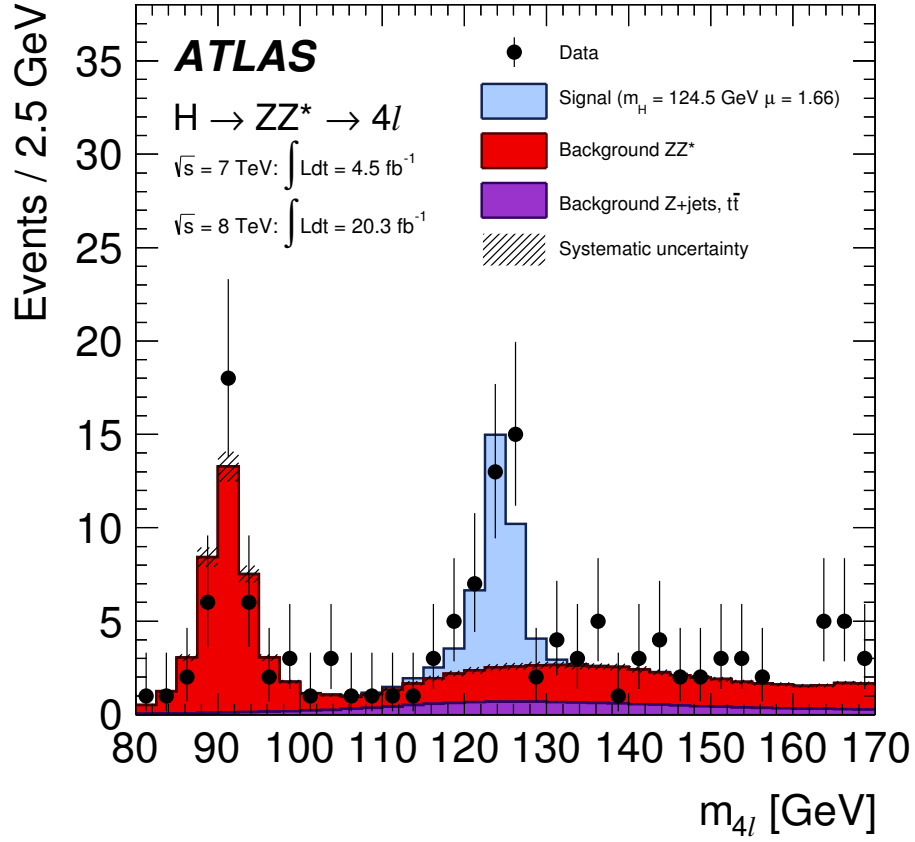


Figure 11.1: Distribution of the four-lepton invariant mass for the selected candidates in the $m_{4\ell}$ range from 80 GeV to 170 GeV for the combined 7 TeV and 8 TeV data samples. Superimposed are the expected distributions of a SM Higgs boson signal for $m_H = 124.5 \text{ GeV}$ normalized to the measured signal strength, as well as the expected ZZ^* and reducible backgrounds.

Category	n_{sig}	FWHM [GeV]	σ_{eff} [GeV]	b in $\pm\sigma_{\text{eff}90}$	s/b [%]	$s/\sqrt{b}$
$\sqrt{s}=8$ TeV						
Inclusive	402.	3.69	1.67	10670	3.39	3.50
Unconverted central low p_{Tt}	59.3	3.13	1.35	801	6.66	1.88
Unconverted central high p_{Tt}	7.1	2.81	1.21	26.0	24.6	1.26
Unconverted rest low p_{Tt}	96.2	3.49	1.53	2624	3.30	1.69
Unconverted rest high p_{Tt}	10.4	3.11	1.36	93.9	9.95	0.96
Unconverted transition	26.0	4.24	1.86	910	2.57	0.78
Converted central low p_{Tt}	37.2	3.47	1.52	589	5.69	1.38
Converted central high p_{Tt}	4.5	3.07	1.35	20.9	19.4	0.88
Converted rest low p_{Tt}	107.2	4.23	1.88	3834	2.52	1.56
Converted rest high p_{Tt}	11.9	3.71	1.64	144.2	7.44	0.89
Converted transition	42.1	5.31	2.41	1977	1.92	0.85
$\sqrt{s}=7$ TeV						
Inclusive	73.9	3.38	1.54	1752	3.80	1.59
Unconverted central low p_{Tt}	10.8	2.89	1.24	128	7.55	0.85
Unconverted central high p_{Tt}	1.2	2.59	1.11	3.7	30.0	0.58
Unconverted rest low p_{Tt}	16.5	3.09	1.35	363	4.08	0.78
Unconverted rest high p_{Tt}	1.8	2.78	1.21	13.6	11.6	0.43
Unconverted transition	4.5	3.65	1.61	125	3.21	0.36
Converted central low p_{Tt}	7.1	3.28	1.44	105	6.06	0.62
Converted central high p_{Tt}	0.8	2.87	1.25	3.5	21.6	0.40
Converted rest low p_{Tt}	21.0	3.93	1.75	695	2.72	0.72
Converted rest high p_{Tt}	2.2	3.43	1.51	24.7	7.98	0.40
Converted transition	8.1	4.81	2.23	365	2.00	0.38

Table 11.1: Summary of the expected number of signal events in the 105 – 160 GeV mass range (n_{sig}), the full width at half maximum (FWHM) of mass resolution, half of the smallest range containing 68% of the signal events (σ_{eff}), number of background events b in the smallest mass window containing 90% of the signal ($\sigma_{\text{eff}90}$), and the ratio s/b and $s/\sqrt{b}$ with s being the expected number of signal events in the window containing 90% of signal events for the *Mass Analysis*. b is derived from the fit of the data in the 105 – 160 GeV mass range. The value of m_H is taken to be 126 GeV and the signal yield is assumed to be the expected Standard Model value. The estimates are shown separately for the 7 TeV and 8 TeV datasets and for the inclusive sample as well as for each of the categories.

Class	Unconverted						Converted					
	Central		Rest		Transition		Central		Rest		Transition	
	low p_{Tt}	high p_{Tt}	low p_{Tt}	high p_{Tt}	low p_{Tt}	high p_{Tt}	low p_{Tt}	high p_{Tt}	low p_{Tt}	high p_{Tt}	low p_{Tt}	high p_{Tt}
$Z \rightarrow e^+e^-$ calibration	0.02	0.03	0.04	0.04	0.11	0.04	0.02	0.02	0.05	0.05	0.11	0.11
LAr cell nonlinearity	0.12	0.19	0.09	0.16	0.39	0.16	0.09	0.19	0.06	0.14	0.29	0.29
Layer calibration	0.13	0.16	0.11	0.13	0.13	0.13	0.07	0.10	0.05	0.07	0.07	0.07
ID material	0.06	0.06	0.08	0.08	0.10	0.08	0.05	0.05	0.06	0.06	0.06	0.06
Other material	0.07	0.08	0.14	0.15	0.35	0.15	0.04	0.04	0.07	0.08	0.20	0.20
Conversion reconstruction	0.02	0.02	0.03	0.03	0.05	0.03	0.03	0.02	0.05	0.04	0.06	0.06
Lateral shower shape	0.04	0.04	0.07	0.07	0.06	0.07	0.09	0.09	0.18	0.19	0.16	0.16
Background modeling	0.10	0.06	0.05	0.11	0.16	0.11	0.13	0.06	0.14	0.18	0.20	0.20
Vertex measurement	0.03											
Total	0.23	0.28	0.24	0.30	0.59	0.30	0.21	0.25	0.27	0.33	0.47	0.47

Table 11.2: Summary of the relative systematic uncertainties (in %) on the $H \rightarrow \gamma\gamma$ mass measurement for the different categories described in the text. The first seven rows give the impact of the photon energy scale systematic uncertainties grouped into seven classes.

11.2 ATLAS–CMS combined measurement

The ATLAS and CMS Collaborations have independently measured the Higgs boson mass m_H using the samples of pp collision data collected in LHC Run 1 during 2011 and 2012. Combined results in the context of the separate experiments, as well as those in the individual channels, are presented in References [118, 122–124]. As m_H is a unconstrained fundamental parameter of the SM and is essential for the measurement of other Higgs boson properties, it is paramount to further enhance the precision of the measurement of m_H by combining the datasets from the two experiments.

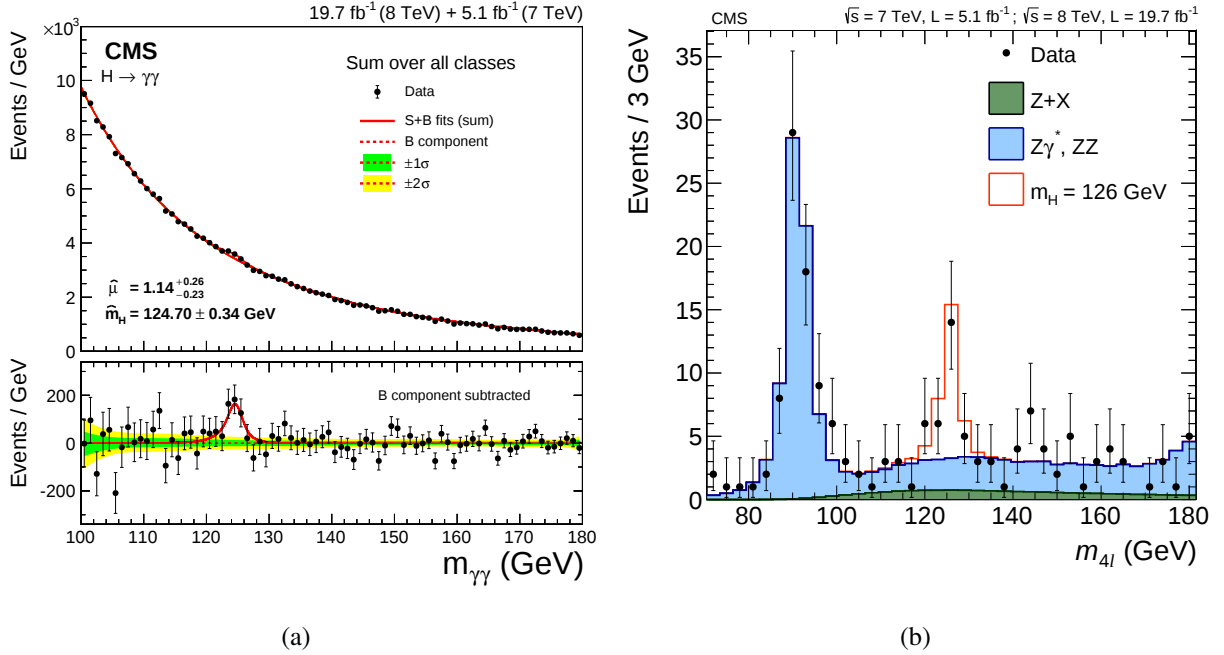


Figure 11.2: Invariant mass spectra from CMS $H \rightarrow \gamma\gamma$ (a) and $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ (b) analyses based on Run 1 dataset.

The invariant mass spectra from CMS analyses are shown in Figure 11.2. The analysis strategies used by ATLAS and CMS are quite similar in both $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ channels. In the combined fit, the signal strength parameters are also floating independently between two decay channels. In addition, the signal strength in the $H \rightarrow \gamma\gamma$ channel is split into two independent free parameters, $\mu_{ggF+t\bar{t}H}^{\gamma\gamma}$ and $\mu_{VBF+VH}^{\gamma\gamma}$. The production processes involving Higgs

boson couplings to fermions, namely ggF and $t\bar{t}H$, are scaled with the $\mu_{ggF+t\bar{t}H}^{\gamma\gamma}$ factor. The production processes involving couplings to vector bosons, namely VBF and VH , are scaled with the $\mu_{VBF+VH}^{\gamma\gamma}$ factor. This split is not introduced in ATLAS combined measurement discussed in the last section, because ATLAS $H \rightarrow \gamma\gamma$ analysis has little sensitivity to separate different production modes. Nevertheless, since the CMS $H \rightarrow \gamma\gamma$ analysis contains very similar categorization as used for the ATLAS coupling analysis discussed in Chapter 10¹, it is pertinent to further reduce the dependence of the results on assumptions about the Higgs boson couplings. The signal strength parameter in the $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ channel ($\mu^{4\ell}$) is not split for $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ analyses because the m_H measurement in this case is found to exhibit almost no sensitivity to the different production mechanisms. The procedure based on the two scale factors $\mu_{ggF+t\bar{t}H}^{\gamma\gamma}$ and $\mu_{VBF+VH}^{\gamma\gamma}$ for the $H \rightarrow \gamma\gamma$ channel results in a shift of about 40 MeV in the ATLAS $H \rightarrow \gamma\gamma$ result reported in Equation (11.1), leading to a shift of 20 MeV in the ATLAS combined mass measurement result report in Equation (11.2).

Combining the ATLAS and CMS data for the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ channels, the mass of the Higgs boson is determined to be

$$\begin{aligned} m_H &= 125.09 \pm 0.24 \text{ GeV} \\ &= 125.09 \pm 0.21 \text{ (stat.)} \pm 0.11 \text{ (syst.) GeV.} \end{aligned} \tag{11.3}$$

By the time this thesis is written, Equation (11.3) still represents the best knowledge on the Higgs boson mass, which is consistent with the values of m_H derived from the less precise WW and $\tau\tau$ Higgs boson decay modes [125–128]. The result shows that the uncertainties in the m_H measurement are dominated by the statistical term, even when the Run 1 data sets of ATLAS and CMS are combined. Figure 11.3 shows the negative log-likelihood ratio scans as a function of m_H , with all nuisance parameters profiled (solid curves), and with the nuisance parameters fixed to their best-fit values (dashed curves).

The combined ATLAS and CMS results for m_H measured in the separate $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ channels are

$$\begin{aligned} m_H^{\gamma\gamma} &= 125.07 \pm 0.29 \text{ GeV} \\ &= 125.07 \pm 0.25 \text{ (stat.)} \pm 0.14 \text{ (syst.) GeV} \end{aligned} \tag{11.4}$$

¹ Unlike ATLAS, CMS uses the same $H \rightarrow \gamma\gamma$ analysis to study both Higgs boson mass and couplings.

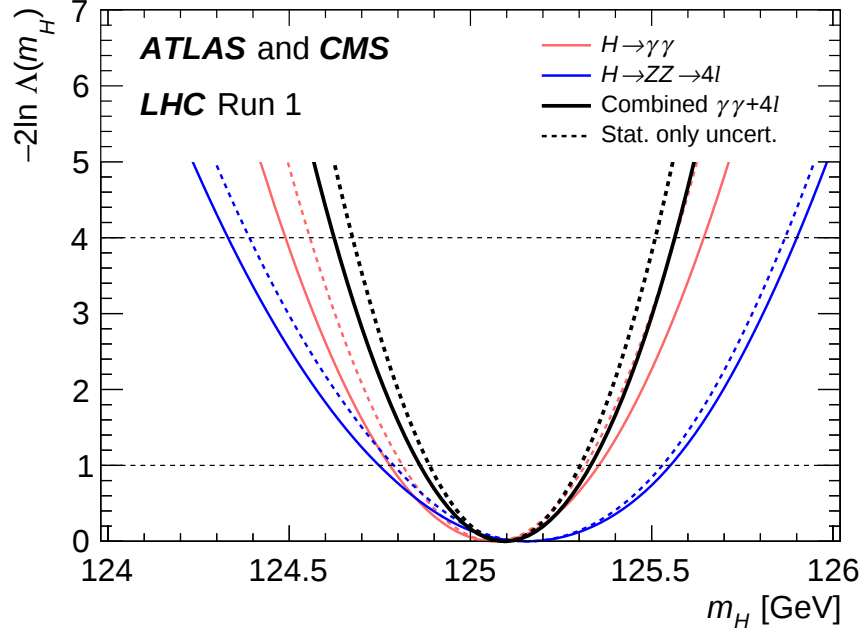


Figure 11.3: Scans of twice the negative log-likelihood ratio $-2 \ln \Lambda(m_H)$ as functions of the Higgs boson mass m_H for the ATLAS and CMS combination of the $H \rightarrow \gamma\gamma$ (red), $H \rightarrow ZZ^{(*)} \rightarrow llll$ (blue), and combined (black) channels. The dashed curves show the results accounting for statistical uncertainties only, with all nuisance parameters associated with systematic uncertainties fixed to their best-fit values. The 1 and 2 standard deviation limits are indicated by the intersections of the horizontal lines at 1 and 4, respectively, with the log-likelihood scan curves.

and

$$\begin{aligned}
 m_H^{4\ell} &= 125.15 \pm 0.40 \text{ GeV} \\
 &= 125.15 \pm 0.37 \text{ (stat.)} \pm 0.15 \text{ (syst.) GeV.}
 \end{aligned}
 \tag{11.5}$$

The corresponding likelihood ratio scans are also shown in Figure 11.3.

A summary of the results from the individual analyses and their combinations is presented in Figure 11.4. The mutual compatibility of the m_H results from the four individual channels is tested using a likelihood ratio with four masses in the numerator and a common mass in the denominator, and thus three degrees of freedom. The three signal strengths are profiled in both the numerator and denominator of the likelihood ratio. The resulting compatibility, defined as the asymptotic p -value (Section 8.2) of the fit, is 10%. Allowing the ATLAS and CMS signal strengths to vary

independently yields a compatibility of 7%. This latter fit results in an m_H value that is 40 MeV larger than the nominal result.

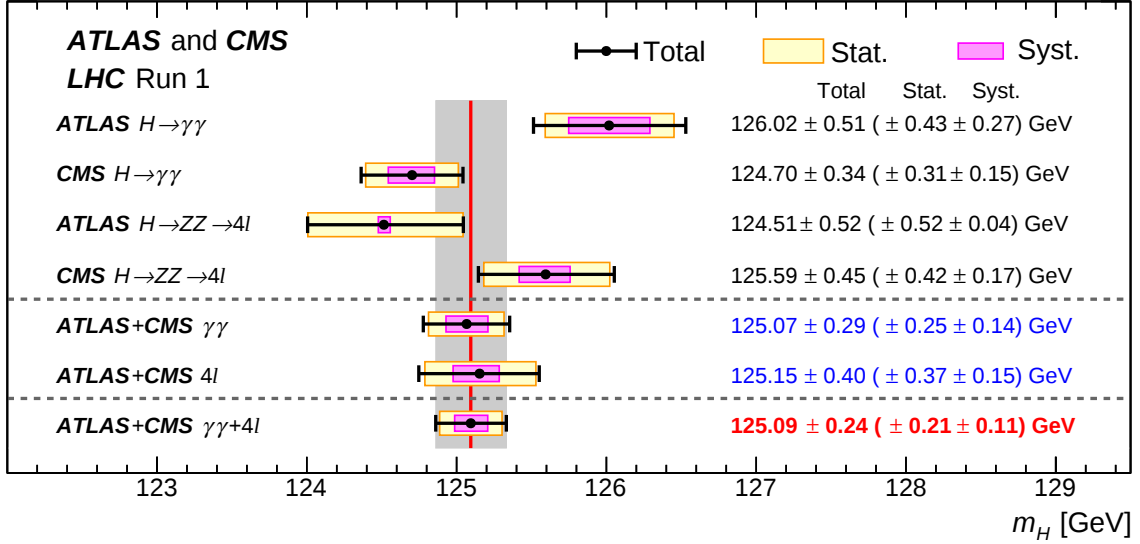


Figure 11.4: Summary of Higgs boson mass measurements from the individual analyses of ATLAS and CMS and from the combined analysis presented here. The systematic (narrower, magenta-shaded bands), statistical (wider, yellow-shaded bands), and total (black error bars) uncertainties are indicated. The (red) vertical line and corresponding (gray) shaded column indicate the central value and the total uncertainty of the combined measurement, respectively.

The compatibility of the combined ATLAS and CMS mass measurement in the $H \rightarrow \gamma\gamma$ channel with the combined measurement in the $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ channel is evaluated using the variable $\Delta m_{\gamma Z} \equiv m_H^{\gamma\gamma} - m_H^{4\ell}$ as the parameter of interest, with all other parameters, including m_H , profiled. Similarly, the compatibility of the ATLAS combined mass measurement in the two channels with the CMS combined measurement in the two channels is evaluated using the variable $\Delta m^{exp.} \equiv m_H^{ATLAS} - m_H^{CMS}$. The observed results, $\Delta m_{\gamma Z} = -0.1 \pm 0.5$ GeV and $\Delta m^{exp.} = 0.4 \pm 0.5$ GeV, are both consistent with zero within 1σ . The difference between the mass values in the two experiments is $\Delta m_{\gamma\gamma}^{exp.} = 1.3 \pm 0.6$ GeV (2.1σ) for the $H \rightarrow \gamma\gamma$ channel and $\Delta m_{4\ell}^{exp.} = -0.9 \pm 0.7$ GeV (1.3σ) for the $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ channel. The combined results exhibit a greater degree of compatibility than the results from the individual decay channels because the $\Delta m^{exp.}$ value has opposite signs in the two channels.

The correlation between the signal strength and the measured mass is explored with two-dimensional likelihood scans as functions of μ and m_H . The three signal strengths are assumed to be the same: $\mu_{ggF+t\bar{t}H}^{\gamma\gamma} = \mu_{VBF+VH}^{\gamma\gamma} = \mu^{4\ell} \equiv \mu$, and thus the ratios of the production cross sections times branching fractions are constrained to the SM predictions. The 68% CL confidence regions derived from asymptotic assumptions (Section 8.2) are shown in Figure 11.5 for each individual measurement, as well as for the combined result.

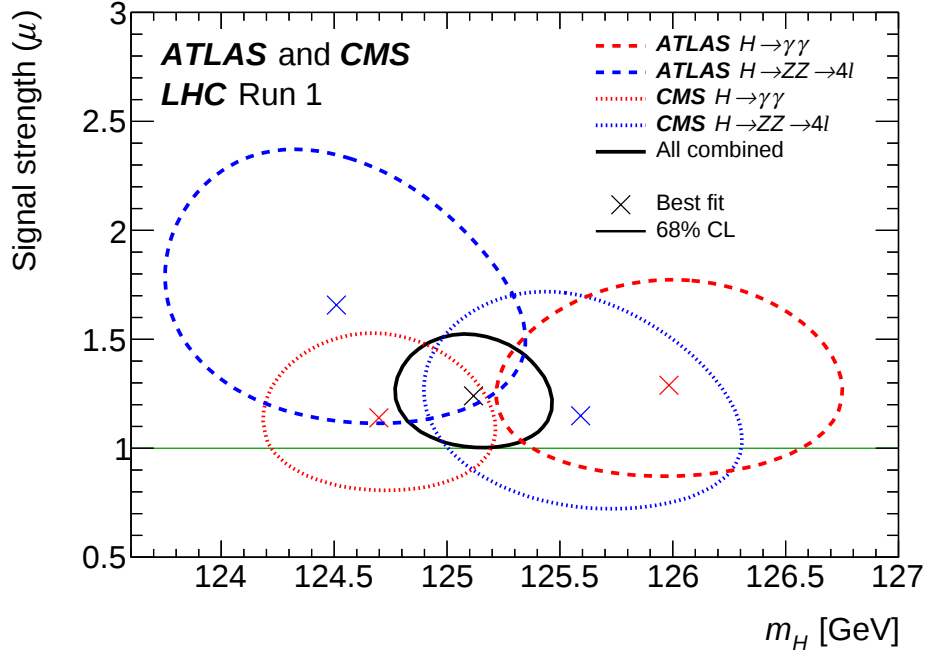


Figure 11.5: Summary of likelihood scans in the two-dimensional plane of signal strength μ versus Higgs boson mass m_H for the ATLAS and CMS experiments. The 68% confidence level contours of the individual measurements are shown by the dashed curves and of the overall combination by the solid curve. The markers indicate the respective best-fit values. The Standard Model signal strength is indicated by the horizontal line at $\mu = 1$.

Constraining all signal yields to their SM predictions results in an m_H value that is about 70 MeV larger than the nominal result with a comparable uncertainty. The increase in the central value reflects the combined effect of the higher-than-expected $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ measured signal strength and the increase of the $H \rightarrow ZZ$ branching fraction with m_H . Thus, the fit assuming SM couplings forces the mass to a higher value in order to accommodate the value $\mu = 1$ expected in the SM.

The observed uncertainties in the combined measurement can be compared with expectations. The latter are evaluated by generating two Asimov data sets [106]². The first Asimov data set is a “prefit” sample, generated using $m_H = 125.0$ GeV and the SM predictions for the couplings, with all nuisance parameters fixed to their nominal values. The second Asimov data set is a “postfit” sample, in which m_H , the three signal strengths $\mu_{ggF+t\bar{t}H}^{\gamma\gamma}$, $\mu_{VBF+VH}^{\gamma\gamma}$, and $\mu^{4\ell}$, and all nuisance parameters are fixed to their best-fit estimates from the data. The expected uncertainties for the combined mass are

$$\delta m_{H_{\text{prefit}}} = \pm 0.24 \text{ GeV} = \pm 0.22 \text{ (stat.)} \pm 0.10 \text{ (syst.) GeV} \quad (11.6)$$

for the prefit case and

$$\delta m_{H_{\text{postfit}}} = \pm 0.22 \text{ GeV} = \pm 0.19 \text{ (stat.)} \pm 0.10 \text{ (syst.) GeV} \quad (11.7)$$

for the postfit case. Both are very similar to the observed uncertainties reported in Equation (11.3).

Since the discovery of the Higgs boson, both experiments have improved their understanding of the electron, photon, and muon measurements [82, 92, 123, 129–131], leading to a significant reduction of the systematic uncertainties in the mass measurement. Nevertheless, the treatment and understanding of systematic uncertainties is an important aspect of the individual measurements and their combination. Based on the results from the individual experiments, the dominant systematic uncertainties for the combined m_H result are expected to be those associated with the energy or momentum scale and its resolution: for the photons in the $H \rightarrow \gamma\gamma$ channel and for the electrons and muons in the $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ channel [118, 122, 123]. These uncertainties are assumed to be uncorrelated between the two experiments since they are related to the specific characteristics of the detectors as well as to the calibration procedures, which are fully independent except for negligible effects due to the use of the common Z boson mass to specify the absolute energy and momentum scales. Other experimental systematic uncertainties [118, 122, 123] are similarly assumed to be uncorrelated between the two experiments. Uncertainties in the theoretical predictions and in the measured integrated luminosities are treated as fully and partially correlated, respectively.

² An Asimov data set is a representative event sample that provides both the median expectation for an experimental result and its expected statistical variation, in the asymptotic approximation, without the need for an extensive MC-based calculation.

To evaluate the relative importance of the different sources of systematic uncertainty, the nuisance parameters are first grouped according to their correspondence to three broad classes of systematic uncertainty:

- uncertainties in the energy or momentum scale and resolution for photons, electrons, and muons (“scale”),
- theoretical uncertainties, e.g. uncertainties in the Higgs boson cross section and branching fractions, and in the normalization of SM background processes (“theory”),
- other experimental uncertainties (“other”).

The uncertainties associated with the different classes of nuisance parameters are defined by the difference in quadrature between the uncertainties resulting from consecutive scans, where iterative fits are performed with the different classes of nuisance parameters cumulatively held fixed to their best-fit values as discussed in Section 8.2. The statistical uncertainty is determined from the final scan, with all nuisance parameters associated with systematic terms held fixed. The result is

$$m_H = 125.09 \pm 0.21 \text{ (stat.)} \pm 0.11 \text{ (scale)} \pm 0.02 \text{ (other)} \pm 0.01 \text{ (theory)} \text{ GeV}, \quad (11.8)$$

from which it is seen that the systematic uncertainty is indeed dominated by the energy and momentum scale terms.

The relative importance of the various sources of systematic uncertainty is further investigated by dividing the nuisance parameters into yet-finer groups, with each group associated with a specific underlying effect, and evaluating the impact of each group on the overall mass uncertainty. The matching of nuisance parameters to an effect is not strictly rigorous because nuisance parameters in the two experiments do not always represent exactly the same effect and in some cases multiple effects are related to the same nuisance parameter. A few experiment-specific groups of nuisance parameters are defined. For example, ATLAS includes a group of nuisance parameters to account for the inaccuracy of the background modeling for the $H \rightarrow \gamma\gamma$ channel. To model the background, ATLAS uses specific analytic functions in each category as discussed in the last section, while CMS simultaneously considers different background parameterizations [132]. The systematic uncertainty in m_H related to the background modeling in CMS is estimated to be negligible [122].

The impact of groups of nuisance parameters is evaluated starting from the contribution of each individual nuisance parameter to the total uncertainty. This contribution is defined as the mass shift δm_H observed when re-performing the maximum likelihood fit after fixing the nuisance parameter in question to its best-fit value increased or decreased by 1 standard deviation (σ) in its distribution, which corresponds to the contribution of that particular nuisance parameter to the final uncertainty. The impact of a group of nuisance parameters is estimated by summing in quadrature the contributions from the individual parameters.

The impacts δm_H due to each of the considered effects are listed in Table 11.3. The results are reported for the four individual channels, both for the data and (in parentheses) the prefit Asimov data set. The row labeled “Systematic uncertainty (sum in quadrature)” shows the total sums in quadrature of the individual terms in the table. The row labeled “Systematic uncertainty (nominal)” shows the corresponding total systematic uncertainties derived using the subtraction in quadrature method discussed in connection with Equation (11.3). The two methods to evaluate the total systematic uncertainty are seen to agree within 10 MeV, which is comparable with the precision of the estimates. The two rightmost columns of Table 11.3 list the contribution of each group of nuisance parameters to the uncertainties in the combined mass measurement, for ATLAS and CMS separately.

The statistical and total uncertainties are summarized in the bottom section of Table 11.3. Since the weight of a channel in the final combination is approximately determined by the inverse of the squared uncertainty, the relative weights for the combined result are estimated to be 19% ($H \rightarrow \gamma\gamma$) and 18% ($H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$) for ATLAS, and 40% ($H \rightarrow \gamma\gamma$) and 23% ($H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$) for CMS. These weights are reported in the last row of Table 11.3, along with the expected values.

Figure 11.6 presents the impact of each group of nuisance parameters on the total systematic uncertainty in the mass measurement of ATLAS, CMS, and their combination. For the individual ATLAS and CMS measurements, the results in Figure 11.6 are approximately equivalent to the sum in quadrature of the respective δm_H terms in Table 11.3 multiplied by their analysis weights, after normalizing these weights to correspond to either ATLAS only or CMS only. The ATLAS and CMS combined results in Figure 11.6 are the sum in quadrature of the combined results in Table 11.3.

The results in Table 11.3 and Figure 11.6 establish that the largest systematic effects for the mass uncertainty are those related to the determination of the energy scale of the photons, followed by those associated with the determination of the electron and muon momentum scales. Since the CMS $H \rightarrow \gamma\gamma$ channel has the largest weight in the combination, its impact on the systematic uncertainty of the combined result is largest.

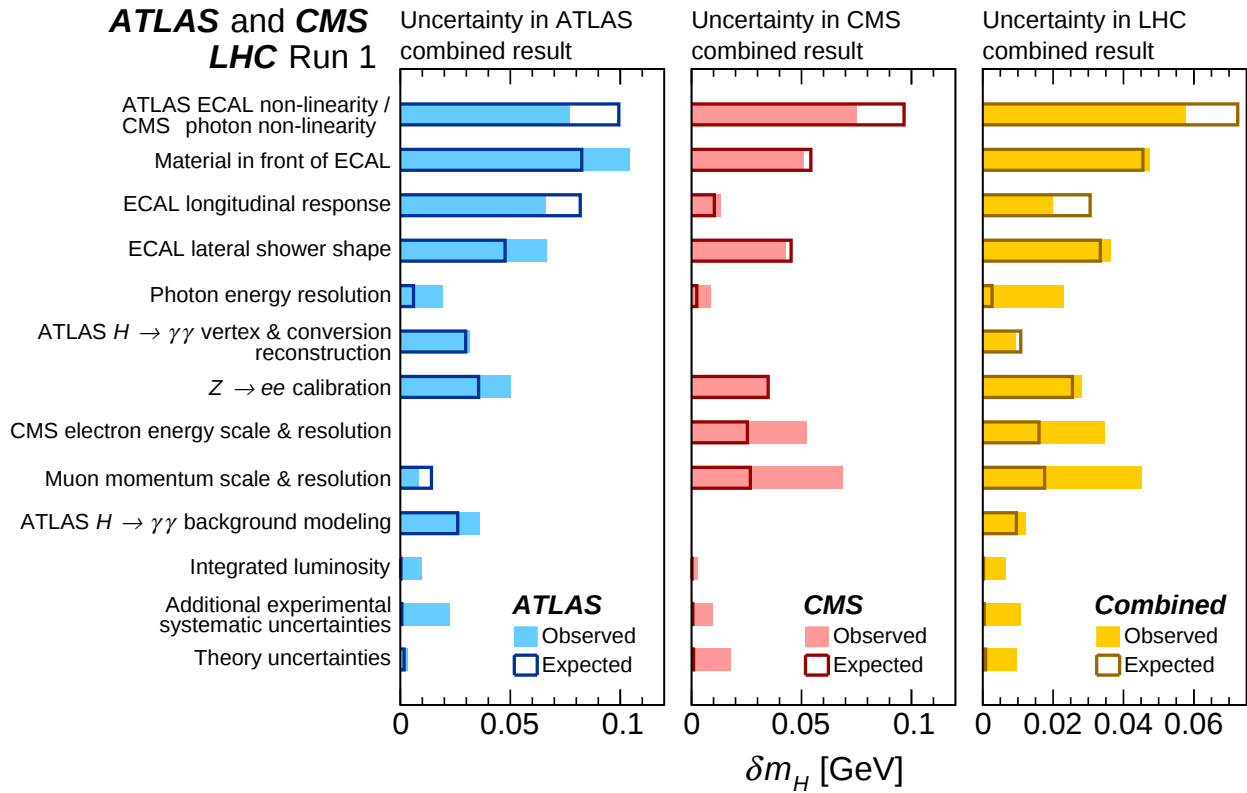


Figure 11.6: The impacts δm_H (see text) of the nuisance parameter groups in Table 11.3 on the ATLAS (left), CMS (center), and combined (right) mass measurement uncertainty. The observed (expected) results are shown by the solid (empty) bars.

	Uncertainty in ATLAS		Uncertainty in CMS		Uncertainty in	
	results [GeV]:		results [GeV]:		combined result [GeV]:	
	observed (expected)		observed (expected)		observed (expected)	
	$H \rightarrow \gamma\gamma$	$H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell$	$H \rightarrow \gamma\gamma$	$H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell$	ATLAS	CMS
Scale uncertainties:						
ATLAS ECAL non-linearity or CMS photon non-linearity	0.14 (0.16)	–	0.10 (0.13)	–	0.02 (0.04)	0.05 (0.06)
Material in front of ECAL	0.15 (0.13)	–	0.07 (0.07)	–	0.03 (0.03)	0.04 (0.03)
ECAL longitudinal response	0.12 (0.13)	–	0.02 (0.01)	–	0.02 (0.03)	0.01 (0.01)
ECAL lateral shower shape	0.09 (0.08)	–	0.06 (0.06)	–	0.02 (0.02)	0.03 (0.03)
Photon energy resolution	0.03 (0.01)	–	0.01 (<0.01)	–	0.02 (<0.01)	<0.01 (<0.01)
ATLAS $H \rightarrow \gamma\gamma$ vertex and conversion reconstruction	0.05 (0.05)	–	–	–	0.01 (0.01)	–
$Z \rightarrow ee$ calibration	0.05 (0.04)	0.03 (0.02)	0.05 (0.05)	–	0.02 (0.01)	0.02 (0.02)
CMS electron energy scale and resolution	–	–	–	0.12 (0.09)	–	0.03 (0.02)
Muon momentum scale and resolution	–	0.03 (0.04)	–	0.11 (0.10)	<0.01 (0.01)	0.05 (0.02)
Other uncertainties:						
ATLAS $H \rightarrow \gamma\gamma$ background modeling	0.04 (0.03)	–	–	–	0.01 (0.01)	–
Integrated luminosity	0.01 (<0.01)	<0.01 (<0.01)	0.01 (<0.01)	<0.01 (<0.01)	0.01 (<0.01)	
Additional experimental systematic uncertainties	0.03 (<0.01)	<0.01 (<0.01)	0.02 (<0.01)	0.01 (<0.01)	0.01 (<0.01)	0.01 (<0.01)
Theory uncertainties	<0.01 (<0.01)	<0.01 (<0.01)	0.02 (<0.01)	<0.01 (<0.01)	0.01 (<0.01)	
Systematic uncertainty (sum in quadrature)	0.27 (0.27)	0.04 (0.04)	0.15 (0.17)	0.16 (0.13)	0.11 (0.10)	
Systematic uncertainty (nominal)	0.27 (0.27)	0.04 (0.05)	0.15 (0.17)	0.17 (0.14)	0.11 (0.10)	
Statistical uncertainty	0.43 (0.45)	0.52 (0.66)	0.31 (0.32)	0.42 (0.57)	0.21 (0.22)	
Total uncertainty	0.51 (0.52)	0.52 (0.66)	0.34 (0.36)	0.45 (0.59)	0.24 (0.24)	
Analysis weights	19% (22%)	18% (14%)	40% (46%)	23% (17%)	–	

Table 11.3: Systematic uncertainties δm_H (see text) associated with the indicated effects for each of the four input channels, and the corresponding contributions of ATLAS and CMS to the systematic uncertainties of the combined result. “ECAL” refers to the electromagnetic calorimeters. The numbers in parentheses indicate expected values obtained from the prefit Asimov data set discussed in the text. The uncertainties for the combined result are related to the values of the individual channels through the relative weight of the individual channel in the combination, which is proportional to the inverse of the respective uncertainty squared. The top section of the table divides the sources of systematic uncertainty into three classes, which are discussed in the text. The bottom section of the table shows the total systematic uncertainties estimated by adding the individual contributions in quadrature, the total systematic uncertainties evaluated using the nominal method discussed in the text, the statistical uncertainties, the total uncertainties, and the analysis weights, illustrative of the relative weight of each channel in the combined m_H measurement.

Chapter 12

Search for high mass scalar resonance in diphoton decay channel

The search for a new scalar resonance in the diphoton final state has been conducted by the ATLAS Experiment based on the LHC Run 1 dataset [133]. The increase of center-of-mass energy in LHC Run 2 allows to push the search range to even higher mass. To reduce model dependency, the search is conducted on the inclusive $m_{\gamma\gamma}$ spectra of data collected during 2015 and 2016 separately. The reason not to merge the two years' datasets is due to the fact that a notable excess has been observed in the 2015 dataset [43], and 2016 dataset thus need to be separated out to provide an independent cross check of the excess.

12.1 Systematic uncertainties

The systematic uncertainties used in the search analyses are summarized in Table 12.1. The experimental uncertainties on the integrated signal yield are similar to the ones used in the *Measurement Analyses*, and have only marginal impact on the analysis results. The uncertainty on the correction factor C (2.8%) is estimated based on variations from using different production modes (VBF and associated production with vector boson or top quark pair) and also different decay width for the calculation. The relative uncertainty in the signal mass resolution is mostly driven by the uncertainty in the constant term of the energy resolution, which is the dominant contribution at high energy, and varies as a function of the mass from 17% at a mass of 200 GeV to 40% at a mass of 2 TeV. The photon energy scale uncertainty is about 0.5%, and is not used for most of the results since the hypothesized new particle mass is scanned.

As for the background modeling study, the background template is prepared using MC simulation for the $\gamma\gamma$ background component, and using data control regions for the γj and γj contributions. The data control samples are selected with at least one of the two photons failing the *tight*

Source	Uncertainty
<i>Signal yield</i>	
Luminosity	$\pm 2.1\%$ (2015), $\pm 3.7\%$ (2016)
Trigger	$\pm 0.63\%$
Photon identification	$\pm 1 - 2\%$, mass-dependent
Isolation efficiency	$\pm 1 - 4\%$
Scalar production process	$\pm 2.8\%$
Photon energy scale	negligible
Photon energy resolution	negligible
<i>Signal modeling</i>	
Photon energy resolution	$^{+17\%}_{-17\%}$ (at $m_X = 200$ GeV) $-$ $^{+35\%}_{-38\%}$ (at $m_X = 2$ TeV)
Photon energy scale	$\pm 0.5\%$ $-$ $\pm 0.7\%$
Pile-up	negligible
<i>Background</i>	
Spurious Signal for 2015	$3 \times 10^{-3} (6.4 \times 10^{-3}) - 3.6(7.7)$ events, NWA
(2016 (12.4 fb^{-1}))	$8 \times 10^{-2} (17.2 \times 10^{-2}) - 20(43)$ events, LWA $\alpha = 10\%$

Table 12.1: Summary of systematic uncertainties on the signal and background considered in the *Search Analysis*.

photon identification criteria but still pass the *loose* ones¹. To overcome the problem of having limit statistics in the control samples, the $m_{\gamma\gamma}$ shapes of the control samples are parameterized by various smooth functions. Different components of the background are then ensembled based on the decomposition study mentioned in Section 6.3 for the spurious signal study. The spurious signals are evaluated for all the hypothesized masses (m_X) and widths (Γ_X) to be searched, ranging from 200 GeV to 2400 GeV in m_X and from 4 MeV to 10% of m_X for Γ_X . The background functional form is selected to be

$$f(x; b, a) = N(1 - x^{1/3})^b x^a, \quad (12.1)$$

¹ Multiple sets of *loose* criteria are used in this analysis as cross checks.

where $x = m_{\gamma\gamma}/\sqrt{s}$, b and a are free parameters controlling the shape, and N is the free normalization factor. The fit range is chosen to be starting from 150 GeV when analyzing the 2015 data, and is raised to 180 GeV when analyzing the 2016 or the combined dataset in order to satisfy the spurious signal criterion mentioned in Chapter 7.2. Since the MC statistics are quite limited for the $\gamma\gamma$ component, before they are implemented in the likelihood model the raw spurious signals are parameterized with a smooth envelope function for conservativeness.

12.2 Results

12.2.1 Diphoton invariant mass spectra

Figure 12.1 shows the diphoton invariant mass distributions together with the background-only fits, for events selected in the 2015 and in the 2016 data separately. The invariant mass distribution including events selected from both years' data is shown in Figure 12.2.

12.2.2 Compatibility with the background-only hypothesis

The 2015 data were first published in Reference [43]. Since then the data have been re-analyzed with improved photon reconstruction algorithms. The photon calibration applied has also been updated with 13 TeV data collected. In addition, the signal model has been switched from POWHEG to MADGRAPH5_AMC@NLO to overcome the unphysical shoulder in POWHEG signal model which shows up at very large width. The significance of the largest excess above the background-only hypothesis in 2015 data hence is slightly reduced to 3.4σ . The corresponding signal mass is 730 GeV with a relative width of 8%.

In the 2016 data, no significant deviation from the background-only hypothesis is observed at the value of the mass corresponding to the most significant excess in 2015 data. The compatibility between the 2016 data and the best fit signal associated to the largest excess in the 2015 data is investigated by assuming a signal mass and width determined from the smallest local p_0 in the 2015 data. The compatibility between the signal cross sections extracted from the 2015 and 2016 datasets is at the level of 2.7σ , due to the absence of a significant excess in the 2016 dataset near 730 GeV.

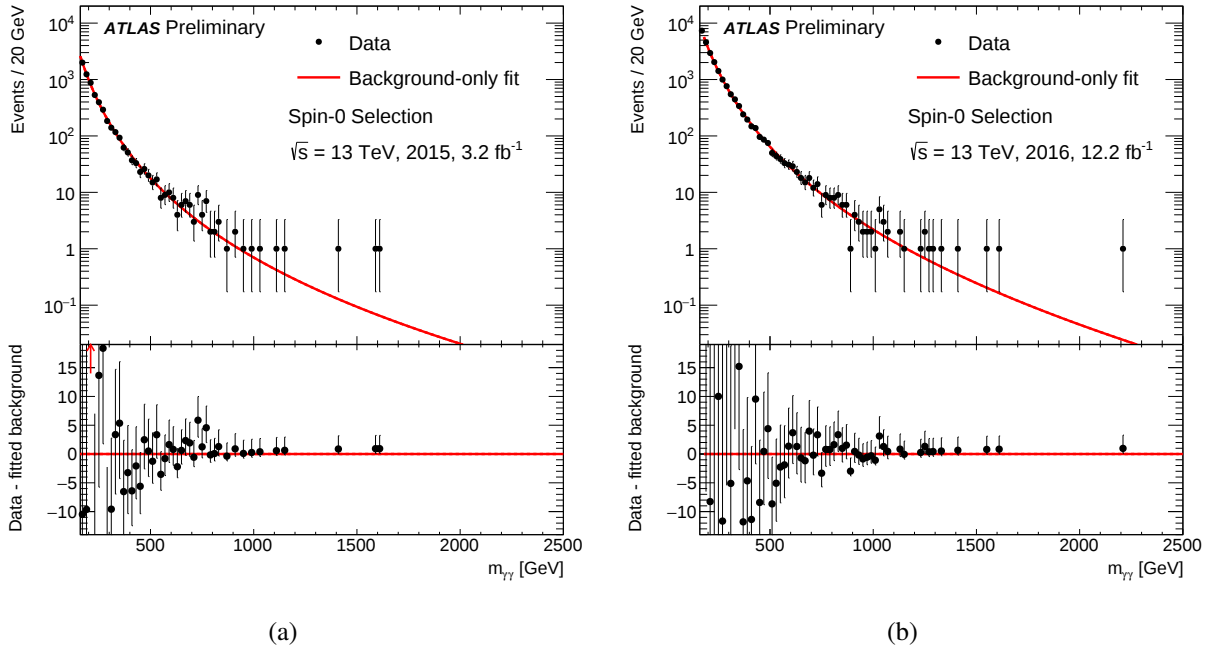


Figure 12.1: Invariant mass distribution of the selected diphoton candidates in the *Search Analysis*, with the background-only fit overlaid, for (a) 2015 data and (b) 2016 data. The difference between the data and this fit is shown in the bottom panel. The arrow shown in the lower panel indicates a values outside the range with more than one standard deviation. There is no data event with $m_{\gamma\gamma} > 2500 \text{ GeV}$.

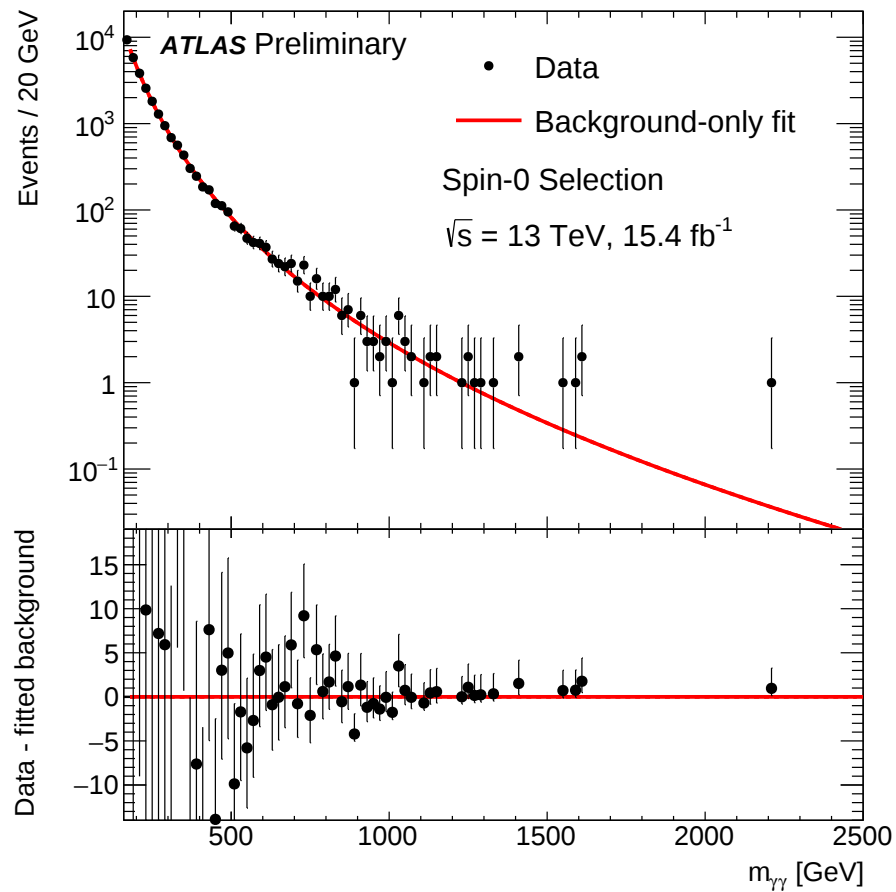


Figure 12.2: Distribution of the diphoton invariant mass of the selected events in the *Search Analysis*, with the background-only fit. The difference between the data and this fit is shown in the bottom panel. The arrow shown in the lower panel indicates a values outside the range with more than one standard deviation. There is no data event with $m_{\gamma\gamma} > 2500 \text{ GeV}$.

The two datasets are then combined to compute local compatibility with the background-only hypothesis as a function of the hypothesized resonance mass and width, as shown in Figure 12.3. Figure 12.4 illustrates the local compatibility as a function of the assumed mass for a few values of the assumed resonance width, comparing the results observed with the 2015 dataset only, the 2016 dataset only and the combined dataset.

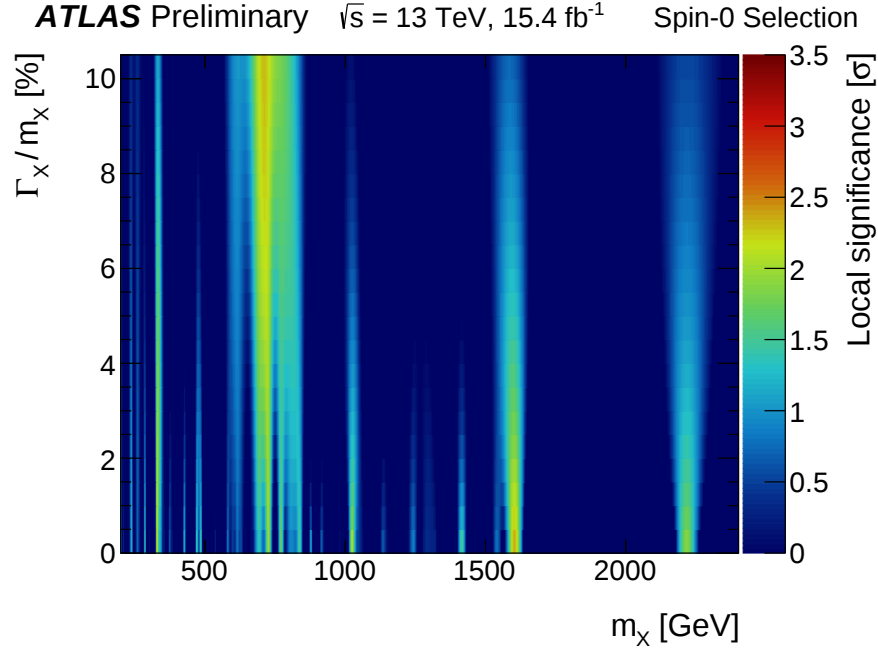
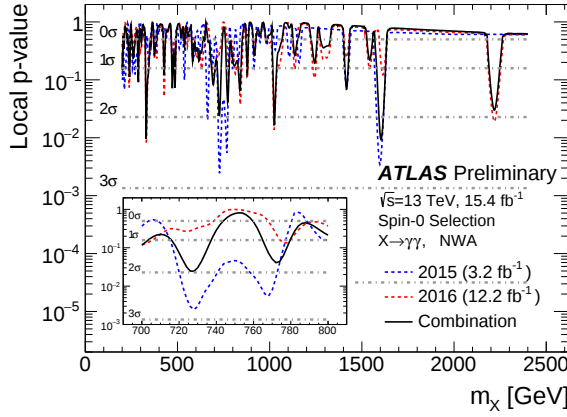


Figure 12.3: Compatibility, in terms of local significance σ , with the background-only hypothesis as a function of the assumed signal mass and width for a scalar resonance.

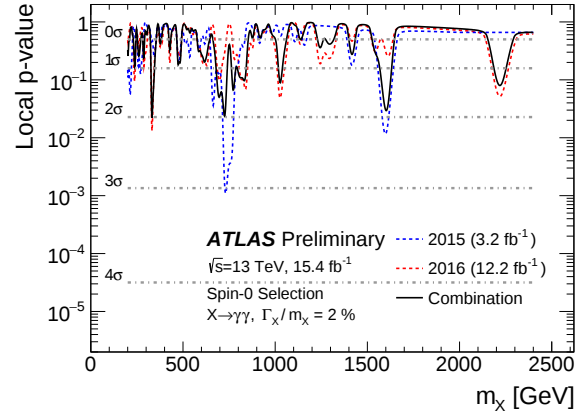
In the combined dataset, the largest deviation over the background-only hypothesis is observed at a mass of 1600 GeV for an assumed narrow width, corresponding to a local significance of 2.4σ . The corresponding global significance is smaller than one standard deviation. In the 700 – 800 GeV mass range the largest local significance is 2.3σ for a mass near 710 GeV and a relative width of 10%.

12.2.3 Limits on fiducial cross section

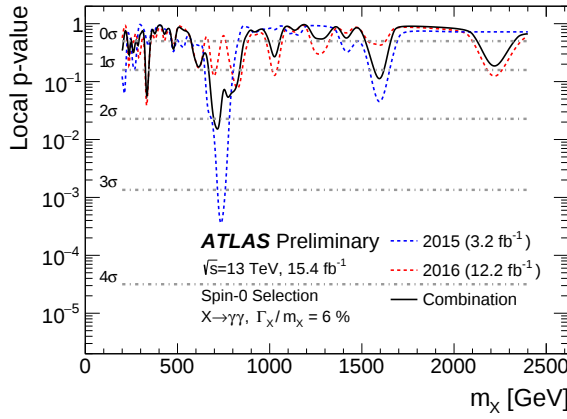
Limits on the cross section times branching ratio to diphoton are derived. The limits are interpreted in a nearly model-independent way in terms of the cross section within the fiducial acceptance defined in Section 7.1.2.



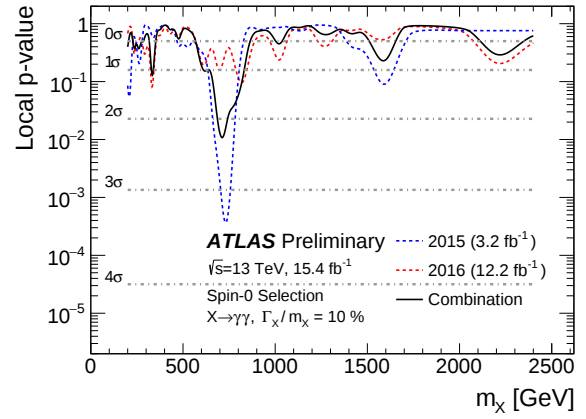
(a)



(b)



(c)



(d)

Figure 12.4: Observed local p_0 values as a function of the assumed scalar resonance mass m_X , for different values of decay widths Γ_X : (a) narrow-width ($\Gamma_X = 4$ MeV), (b) $\Gamma_X/m_X = 2\%$, (c) $\Gamma_X/m_X = 6\%$, and (d) $\Gamma_X/m_X = 10\%$.

Figure 12.5 shows the limits on the signal fiducial cross section times branching ratio to two photons for a scalar particle as a function of the assumed signal mass and for different values of the signal width. For a narrow decay width, the limits on the fiducial cross section times branching ratio range from 15 fb near a mass of 200 GeV to 0.2 fb at 2400 GeV.

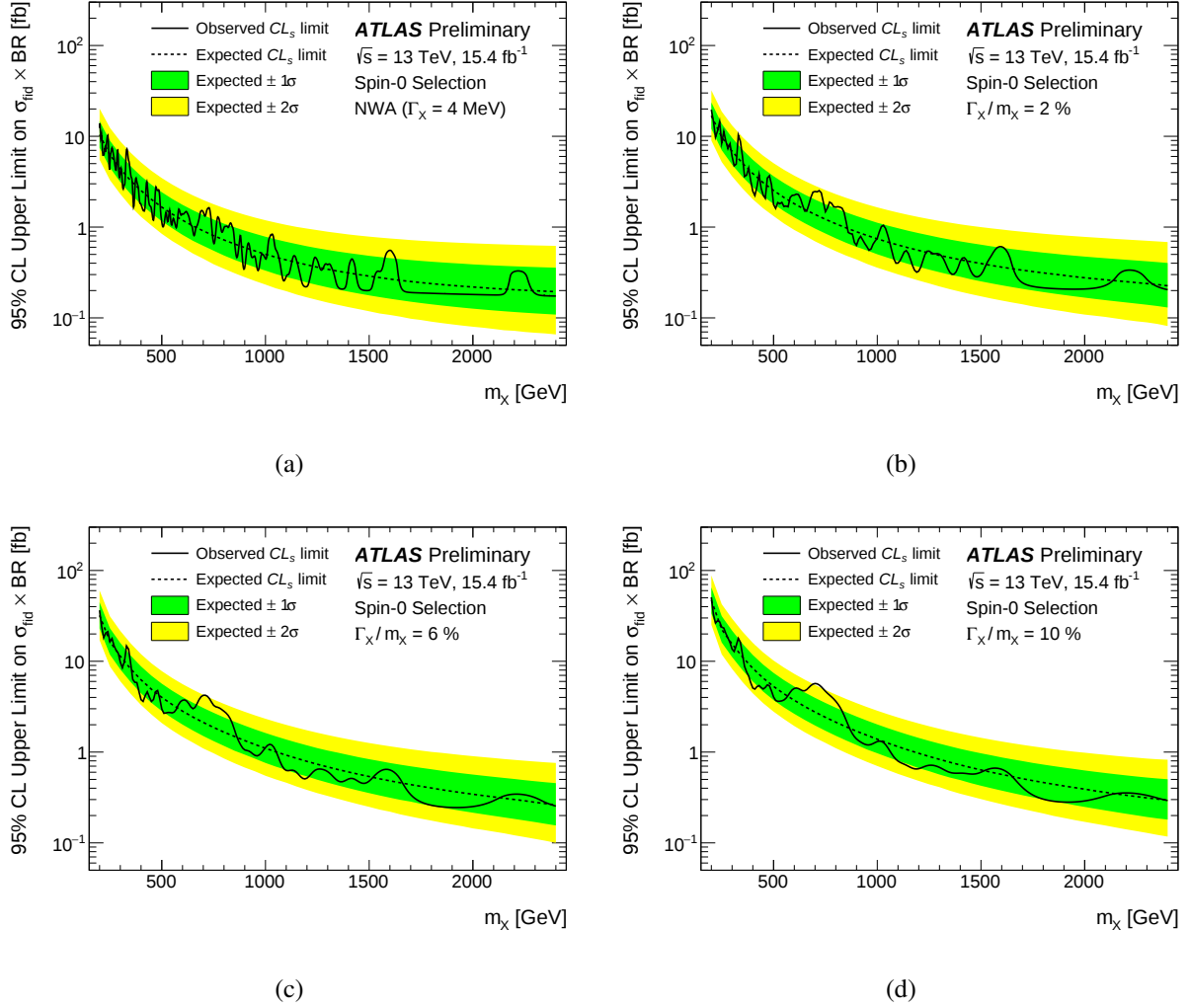


Figure 12.5: Upper limits on the fiducial cross section times branching ratio to two photons of a scalar particle produced at $\sqrt{s} = 13$ TeV as a function of its mass m_X , for different values of decay widths Γ_X : (a) narrow-width ($\Gamma_X = 4 \text{ MeV}$), (b) $\Gamma_X/m_X = 2\%$, (c) $\Gamma_X/m_X = 6\%$, and (d) $\Gamma_X/m_X = 10\%$.

Chapter 13

Conclusion

With 4.8 fb^{-1} of pp collision data collected at $\sqrt{s} = 7 \text{ TeV}$ in 2011, and 5.9 fb^{-1} data collected at 8 TeV in 2012 by the ATLAS detector at the LHC, an excess of 4.5σ deviation from the background-only hypothesis is observed near 126.5 GeV in the diphoton final state. Along with the excesses observed in the $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ and $H \rightarrow WW^{(*)} \rightarrow \ell\nu\ell\nu$ channels, the observation of a Higgs-like particle is established at 6.0σ level.

With more data accumulated during LHC Run 1, the measurements of Higgs boson couplings and mass are conducted in the $H \rightarrow \gamma\gamma$ channel by the ATLAS experiment based on 4.5 fb^{-1} of pp collision data at $\sqrt{s} = 7 \text{ TeV}$ collected in 2011, and 20.3 fb^{-1} data at $\sqrt{s} = 8 \text{ TeV}$ collected in 2012. The combined signal strength is measured to be $\mu = 1.17^{+0.28}_{-0.26}$ at $m_H = 125.4 \text{ GeV}$, which is in good agreement with SM expectation at unity within uncertainty. The signal strength parameters for individual Higgs boson production processes are also measured and found to be consistent with SM predictions. The results from ATLAS $H \rightarrow \gamma\gamma$ analysis are then combined with other decay channels from both ATLAS and CMS experiments, and no obvious deviation from SM expectation is found. The mass of the Higgs boson is first measured in $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^{(*)} \rightarrow \ell\ell\ell\ell$ channels, and then combined between the two channels and also between ATLAS and CMS experiments to maximize the statistical power. The combined mass is measured to be $m_H = 125.09 \pm 0.24 \text{ GeV}$.

With LHC center-of-mass energy increased to 13 TeV , a search for high mass BSM scalar resonance is performed based on 15.4 fb^{-1} of pp collision data collected during 2015 and 2016. While a notable wide excess is first found in the 2015 dataset (3.2 fb^{-1}) near 750 GeV , it is not confirmed by the 2016 data with much higher integrated luminosity (12.4 fb^{-1}). Limits are set on the production cross section times branching ratio of such resonances.

Most of the measurements of the Higgs boson properties reported in this thesis are limited by statistical uncertainties. With about 100 fb^{-1} of pp collision data to be collected in the ongoing LHC Run 2, better precision on the Higgs boson properties measurements will be achieved. The searches conducted at high mass will also lead to more solid understanding of physics at TeV scale.

APPENDIX

Coupling modifiers

Following the LO tree level motivated framework (κ -framework) recommended in Reference [16], measurements of coupling scale factors are implemented in the coupling studies detailed in Section 10.3. This framework is based on the assumptions below.

- The signals observed in the different search channels originate from a single narrow resonance. The case of several, possibly overlapping, resonances in the signal region is not considered.
- The width of the assumed Higgs boson is neglected; i.e. the zero-width approximation is used. Hence the product $\sigma \times BR(i \rightarrow H \rightarrow f)$ can be decomposed in the following way for all channels:

$$\sigma \times BR(i \rightarrow H \rightarrow f) = \frac{\sigma_i \cdot \Gamma_f}{\Gamma_H}, \quad (\text{A.1})$$

where σ_i is the production cross section through the initial state i , Γ_f the partial decay width into the final state f and Γ_H the total width of the Higgs boson.

- Only modifications of couplings strengths, i.e. of absolute values of couplings, are taken into account, while the tensor structure of the couplings is assumed to be the same as in the SM. This means in particular that the observed state is assumed to be a CP-even scalar as in the SM.

The LO-motivated coupling scale factors κ_j are defined in such a way that the cross section σ_j and the partial decay width Γ_j associated with the SM particle j scale with the factor κ_j^2 when compared to the corresponding SM prediction. The effective scale factors κ_γ and κ_g for the loop-induced processes $H \rightarrow \gamma\gamma$ and ggF can be treated as a function of more fundamental coupling scale factors κ_t , κ_W , κ_b , and similarly for all the other particles that contribute to these SM loop

processes. In such cases, the scaled fundamental couplings are propagated through the loop calculations, including all interference effects, using the functional form derived from the SM.

Considering only SM particles, to a very good approximation, the typical expressions for Higgs boson production with $m_H = 125.09$ GeV are:

$$\sigma_{ggF} \sim 1.06 \cdot \kappa_t^2 + 0.01 \cdot \kappa_b^2 - 0.07 \cdot \kappa_t \kappa_b, \quad (\text{A.2})$$

$$\sigma_{\text{VBF}} \sim 0.74 \cdot \kappa_W^2 + 0.26 \cdot \kappa_Z^2, \quad (\text{A.3})$$

$$\sigma_{WH/ZH} \sim \kappa_{W/Z}^2, \quad (\text{A.4})$$

$$\sigma_{t\bar{t}H/b\bar{b}H} \sim \kappa_{t/b}^2, \quad (\text{A.5})$$

$$\sigma_{tHqb} \sim 3.40 \cdot \kappa_t^2 + 3.56 \cdot \kappa_W^2 - 5.96 \cdot \kappa_t \kappa_W, \quad (\text{A.6})$$

$$\sigma_{tHW} \sim 1.84 \cdot \kappa_t^2 + 1.57 \cdot \kappa_W^2 - 2.41 \cdot \kappa_t \kappa_W, \quad (\text{A.7})$$

$$(\text{A.8})$$

For Higgs boson decay partial and total widths, the expressions are:

$$\Gamma_{WW/ZZ} \sim \kappa_{W/Z}^2, \quad (\text{A.9})$$

$$\Gamma_{\tau\tau/b\bar{b}/\mu^+\mu^-} \sim \kappa_{\tau/b/\mu}^2, \quad (\text{A.10})$$

$$\Gamma_{\gamma\gamma} \sim 1.59 \cdot \kappa_W^2 + 0.07 \cdot \kappa_t^2 - 0.66 \cdot \kappa_W \kappa_t, \quad (\text{A.11})$$

$$\begin{aligned} \Gamma_H \sim & 0.57 \cdot \kappa_b^2 + 0.22 \cdot \kappa_W^2 + \\ & 0.09 \cdot \kappa_g^2 + 0.06 \cdot \kappa_\tau^2 + 0.03 \cdot \kappa_Z^2 + 0.03 \cdot \kappa_c^2 + \\ & 0.0023 \cdot \kappa_\gamma^2 + 0.0016 \cdot \kappa_{(Z\gamma)}^2 + 0.0001 \cdot \kappa_s^2 + 0.00022 \cdot \kappa_\mu^2. \end{aligned} \quad (\text{A.12})$$

LIST OF REFERENCES

- [1] UA1 Collaboration, “Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV”, *Phys. Lett.* **B122** (1983) 103–116, doi:10.1016/0370-2693(83)91177-2.
- [2] UA1 Collaboration, “Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/ c^2 at the CERN SPS Collider”, *Phys. Lett.* **B126** (1983) 398–410, doi:10.1016/0370-2693(83)90188-0.
- [3] Particle Data Group Collaboration, “Review of Particle Physics”, *Chin. Phys.* **C38** (2014) 090001, doi:10.1088/1674-1137/38/9/090001.
- [4] F. Englert and R. Brout, “Broken Symmetry and the Mass of Gauge Vector Mesons”, *Phys. Rev. Lett.* **13** (1964) 321–323, doi:10.1103/PhysRevLett.13.321.
- [5] P. W. Higgs, “Broken symmetries, massless particles and gauge fields”, *Phys. Lett.* **12** (1964) 132–133, doi:10.1016/0031-9163(64)91136-9.
- [6] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev. Lett.* **13** (1964) 508–509, doi:10.1103/PhysRevLett.13.508.
- [7] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global Conservation Laws and Massless Particles”, *Phys. Rev. Lett.* **13** (1964) 585–587, doi:10.1103/PhysRevLett.13.585.
- [8] P. W. Higgs, “Spontaneous Symmetry Breakdown without Massless Bosons”, *Phys. Rev.* **145** (1966) 1156–1163, doi:10.1103/PhysRev.145.1156.
- [9] T. W. B. Kibble, “Symmetry breaking in nonAbelian gauge theories”, *Phys. Rev.* **155** (1967) 1554–1561, doi:10.1103/PhysRev.155.1554.
- [10] ALEPH, DELPHI, L3 and OPAL Collaborations, LEP Working Group for Higgs Boson Searches, “Search for the standard model Higgs boson at LEP”, *Phys. Lett.* **B565** (2003) 61–75, doi:10.1016/S0370-2693(03)00614-2, arXiv:hep-ex/0306033.
- [11] CDF and D0 Collaborations, “Higgs Boson Studies at the Tevatron”, *Phys. Rev.* **D88** (2013) 052014, doi:10.1103/PhysRevD.88.052014, arXiv:1303.6346.

- [12] ATLAS Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Phys. Lett.* **B716** (2012) 1–29, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [13] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Phys. Lett.* **B716** (2012) 30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [14] LHC Higgs Cross Section Working Group, “Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables”, *CERN-2011-002* (CERN, Geneva, 2011) doi:10.5170/CERN-2011-002, arXiv:1101.0593.
- [15] LHC Higgs Cross Section Working Group, “Handbook of LHC Higgs Cross Sections: 2. Differential Distributions”, *CERN-2012-002* (CERN, Geneva, 2012) doi:10.5170/CERN-2012-002, arXiv:1201.3084.
- [16] LHC Higgs Cross Section Working Group, “Handbook of LHC Higgs Cross Sections: 3. Higgs Properties”, *CERN-2013-004* (CERN, Geneva, 2013) doi:10.5170/CERN-2013-004, arXiv:1307.1347.
- [17] A. Djouadi, M. Spira, and P. M. Zerwas, “Production of Higgs bosons in proton colliders: QCD corrections”, *Phys. Lett.* **B264** (1991) 440–446, doi:10.1016/0370-2693(91)90375-Z.
- [18] S. Dawson, “Radiative corrections to Higgs boson production”, *Nucl. Phys.* **B359** (1991) 283–300, doi:10.1016/0550-3213(91)90061-2.
- [19] M. Spira, A. Djouadi, D. Graudenz, and P. M. Zerwas, “Higgs boson production at the LHC”, *Nucl. Phys.* **B453** (1995) 17–82, doi:10.1016/0550-3213(95)00379-7, arXiv:hep-ph/9504378.
- [20] R. V. Harlander and W. B. Kilgore, “Next-to-next-to-leading order Higgs production at hadron colliders”, *Phys. Rev. Lett.* **88** (2002) 201801, doi:10.1103/PhysRevLett.88.201801, arXiv:hep-ph/0201206.
- [21] C. Anastasiou and K. Melnikov, “Higgs boson production at hadron colliders in NNLO QCD”, *Nucl. Phys.* **B646** (2002) 220–256, doi:10.1016/S0550-3213(02)00837-4, arXiv:hep-ph/0207004.
- [22] V. Ravindran, J. Smith, and W. L. van Neerven, “NNLO corrections to the total cross-section for Higgs boson production in hadron hadron collisions”, *Nucl. Phys.* **B665** (2003) 325–366, doi:10.1016/S0550-3213(03)00457-7, arXiv:hep-ph/0302135.
- [23] U. Aglietti, R. Bonciani, G. Degrossi, and A. Vicini, “Two loop light fermion contribution to Higgs production and decays”, *Phys. Lett.* **B595** (2004) 432–441, doi:10.1016/j.physletb.2004.06.063, arXiv:hep-ph/0404071.

- [24] S. Actis, G. Passarino, C. Sturm, and S. Uccirati, “NLO Electroweak Corrections to Higgs Boson Production at Hadron Colliders”, *Phys. Lett.* **B670** (2008) 12–17, doi:10.1016/j.physletb.2008.10.018, arXiv:0809.1301.
- [25] M. Ciccolini, A. Denner, and S. Dittmaier, “Strong and electroweak corrections to the production of Higgs + 2jets via weak interactions at the LHC”, *Phys. Rev. Lett.* **99** (2007) 161803, doi:10.1103/PhysRevLett.99.161803, arXiv:0707.0381.
- [26] M. Ciccolini, A. Denner, and S. Dittmaier, “Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC”, *Phys. Rev.* **D77** (2008) 013002, doi:10.1103/PhysRevD.77.013002, arXiv:0710.4749.
- [27] K. Arnold et al., “VBFNLO: A parton level Monte Carlo for processes with electroweak bosons”, *Comput. Phys. Commun.* **180** (2009) 1661–1670, doi:10.1016/j.cpc.2009.03.006, arXiv:0811.4559.
- [28] P. Bolzoni, F. Maltoni, S.-O. Moch, and M. Zaro, “Higgs production via vector-boson fusion at NNLO in QCD”, *Phys. Rev. Lett.* **105** (2010) 011801, doi:10.1103/PhysRevLett.105.011801, arXiv:1003.4451.
- [29] O. Brein, A. Djouadi, and R. Harlander, “NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders”, *Phys. Lett.* **B579** (2004) 149–156, doi:10.1016/j.physletb.2003.10.112, arXiv:hep-ph/0307206.
- [30] M. L. Ciccolini, S. Dittmaier, and M. Kramer, “Electroweak radiative corrections to associated WH and ZH production at hadron colliders”, *Phys. Rev.* **D68** (2003) 073003, doi:10.1103/PhysRevD.68.073003, arXiv:hep-ph/0306234.
- [31] S. Dawson, C. B. Jackson, L. Reina, and D. Wackerroth, “Exclusive Higgs boson production with bottom quarks at hadron colliders”, *Phys. Rev.* **D69** (2004) 074027, doi:10.1103/PhysRevD.69.074027, arXiv:hep-ph/0311067.
- [32] S. Dittmaier, M. Kramer, I., and M. Spira, “Higgs radiation off bottom quarks at the Tevatron and the CERN LHC”, *Phys. Rev.* **D70** (2004) 074010, doi:10.1103/PhysRevD.70.074010, arXiv:hep-ph/0309204.
- [33] S. Dawson, C. B. Jackson, L. Reina, and D. Wackerroth, “Higgs production in association with bottom quarks at hadron colliders”, *Mod. Phys. Lett.* **A21** (2006) 89–110, doi:10.1142/S0217732306019256, arXiv:hep-ph/0508293.
- [34] R. V. Harlander and W. B. Kilgore, “Higgs boson production in bottom quark fusion at next-to-next-to leading order”, *Phys. Rev.* **D68** (2003) 013001, doi:10.1103/PhysRevD.68.013001, arXiv:hep-ph/0304035.
- [35] R. Harlander, M. Krämer, and M. Schumacher, “Bottom-quark associated Higgs-boson production: Reconciling the four- and five-flavour scheme approach”, arXiv:1112.3478.

- [36] W. Beenakker et al., “Higgs radiation off top quarks at the Tevatron and the LHC”, *Phys. Rev. Lett.* **87** (2001) 201805, doi:10.1103/PhysRevLett.87.201805, arXiv:hep-ph/0107081.
- [37] W. Beenakker et al., “NLO QCD corrections to $t\bar{t}H$ production in hadron collisions”, *Nucl. Phys.* **B653** (2003) 151–203, doi:10.1016/S0550-3213(03)00044-0, arXiv:hep-ph/0211352.
- [38] S. Dawson, L. H. Orr, L. Reina, and D. Wackeroth, “Associated top quark Higgs boson production at the LHC”, *Phys. Rev.* **D67** (2003) 071503, doi:10.1103/PhysRevD.67.071503, arXiv:hep-ph/0211438.
- [39] S. Dawson et al., “Associated Higgs production with top quarks at the large hadron collider: NLO QCD corrections”, *Phys. Rev.* **D68** (2003) 034022, doi:10.1103/PhysRevD.68.034022, arXiv:hep-ph/0305087.
- [40] F. Maltoni and T. Stelzer, “MadEvent: Automatic event generation with MadGraph”, *JHEP* **02** (2003) 027, doi:10.1088/1126-6708/2003/02/027, arXiv:hep-ph/0208156.
- [41] ATLAS Collaboration, “Search for $H \rightarrow \gamma\gamma$ produced in association with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs boson using data taken at 7 TeV and 8 TeV with the ATLAS detector”, *Phys. Lett.* **B740** (2015) 222–242, doi:10.1016/j.physletb.2014.11.049, arXiv:1409.3122.
- [42] ATLAS Collaboration, “Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics”, arXiv:0901.0512.
- [43] ATLAS Collaboration, “Search for resonances in diphoton events at $\sqrt{s}=13$ TeV with the ATLAS detector”, *JHEP* **09** (2016) 001, doi:10.1007/JHEP09(2016)001, arXiv:1606.03833.
- [44] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider”, *JINST* **3** (2008) S08003, doi:10.1088/1748-0221/3/08/S08003.
- [45] ATLAS Collaboration, “2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data”, *ATL-DAQ-PUB-2016-001* (Mar, 2016). <https://cdsweb.cern.ch/record/2136007>.
- [46] P. Nason, “A New method for combining NLO QCD with shower Monte Carlo algorithms”, *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
- [47] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”, *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

- [48] S. Alioli, P. Nason, C. Oleari, and E. Re, “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”, *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
- [49] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO Higgs boson production via gluon fusion matched with shower in POWHEG”, *JHEP* **04** (2009) 002, doi:10.1088/1126-6708/2009/04/002, arXiv:0812.0578.
- [50] E. Bagnaschi, G. Degrossi, P. Slavich, and A. Vicini, “Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM”, *JHEP* **02** (2012) 088, doi:10.1007/JHEP02(2012)088, arXiv:1111.2854.
- [51] T. Sjöstrand, S. Mrenna, P. Skands, “A brief introduction to PYTHIA 8.1”, *Comput. Phys. Commun.* **178** (2008) 852–867, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.
- [52] P. Nason and C. Oleari, “NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG”, *JHEP* **02** (2010) 037, doi:10.1007/JHEP02(2010)037, arXiv:0911.5299.
- [53] G. Bevilacqua et al., “HELAC-NLO”, *Comput. Phys. Commun.* **184** (2013) 986–997, doi:10.1016/j.cpc.2012.10.033, arXiv:1110.1499.
- [54] J. Alwall et al., “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.
- [55] M. Bahr et al., “Herwig++ Physics and Manual”, *Eur. Phys. J.* **C58** (2008) 639–707, doi:10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883.
- [56] H.-L. Lai et al., “New parton distributions for collider physics”, *Phys. Rev.* **D82** (2010) 074024, doi:10.1103/PhysRevD.82.074024, arXiv:1007.2241.
- [57] J. Pumplin et al., “New generation of parton distributions with uncertainties from global QCD analysis”, *JHEP* **07** (2002) 012, doi:10.1088/1126-6708/2002/07/012, arXiv:hep-ph/0201195.
- [58] ATLAS Collaboration, “Summary of ATLAS Pythia 8 tunes”, *ATL-PHYS-PUB-2012-003* (Aug, 2012). <http://cdsweb.cern.ch/record/1474107>.
- [59] L. J. Dixon and M. S. Siu, “Resonance continuum interference in the diphoton Higgs signal at the LHC”, *Phys. Rev. Lett.* **90** (2003) 252001, doi:10.1103/PhysRevLett.90.252001, arXiv:hep-ph/0302233.
- [60] L. J. Dixon and Y. Li, “Bounding the Higgs Boson Width Through Interferometry”, *Phys. Rev. Lett.* **111** (2013) 111802, doi:10.1103/PhysRevLett.111.111802, arXiv:1305.3854.
- [61] S. P. Martin, “Shift in the LHC Higgs diphoton mass peak from interference with background”, *Phys. Rev.* **D86** (2012) 073016, doi:10.1103/PhysRevD.86.073016, arXiv:1208.1533.

- [62] D. de Florian, G. Ferrera, M. Grazzini, and D. Tommasini, “Higgs boson production at the LHC: transverse momentum resummation effects in the $H \rightarrow \gamma\gamma$, $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ and $H \rightarrow ZZ \rightarrow 4\ell$ decay modes”, *JHEP* **06** (2012) 132, doi:10.1007/JHEP06(2012)132, arXiv:1203.6321.
- [63] M. Grazzini and H. Sargsyan, “Heavy-quark mass effects in Higgs boson production at the LHC”, *JHEP* **09** (2013) 129, doi:10.1007/JHEP09(2013)129, arXiv:1306.4581.
- [64] R. Boughezal et al., “Combining Resummed Higgs Predictions Across Jet Bins”, *Phys. Rev.* **D89** (2014) 074044, doi:10.1103/PhysRevD.89.074044, arXiv:1312.4535.
- [65] J. M. Campbell, R. K. Ellis, and G. Zanderighi, “Next-to-Leading order Higgs + 2 jet production via gluon fusion”, *JHEP* **10** (2006) 028, doi:10.1088/1126-6708/2006/10/028, arXiv:hep-ph/0608194.
- [66] A. Denner, S. Dittmaier, S. Kallweit, and A. Muck, “Electroweak corrections to Higgs-strahlung off W/Z bosons at the Tevatron and the LHC with HAWK”, *JHEP* **03** (2012) 075, doi:10.1007/JHEP03(2012)075, arXiv:1112.5142.
- [67] ATLAS Collaboration, “The ATLAS Simulation Infrastructure”, *Eur. Phys. J.* **C70** (2010) 823–874, doi:10.1140/epjc/s10052-010-1429-9, arXiv:1005.4568.
- [68] GEANT4 Collaboration, “GEANT4: A Simulation toolkit”, *Nucl. Instrum. Meth.* **A506** (2003) 250–303, doi:10.1016/S0168-9002(03)01368-8.
- [69] P. Artoisenet et al., “A framework for Higgs characterisation”, *JHEP* **11** (2013) 043, doi:10.1007/JHEP11(2013)043, arXiv:1306.6464.
- [70] S. Carrazza, S. Forte, and J. Rojo, “Parton Distributions and Event Generators”, in *Proceedings, 43rd International Symposium on Multiparticle Dynamics (ISMD 13)*, pp. 89–96, 2013. arXiv:1311.5887.
- [71] ATLAS Collaboration, “ATLAS Pythia 8 tunes to 7 TeV data”, *ATL-PHYS-PUB-2014-021* (Nov, 2014). <http://cdsweb.cern.ch/record/1966419>.
- [72] S. Jung, J. Song, and Y. W. Yoon, “How Resonance-Continuum Interference Changes 750 GeV Diphoton Excess: Signal Enhancement and Peak Shift”, *JHEP* **05** (2016) 009, doi:10.1007/JHEP05(2016)009, arXiv:1601.00006.
- [73] A. Djouadi, J. Ellis, and J. Quevillon, “Interference effects in the decays of spin-zero resonances into $\gamma\gamma$ and $t\bar{t}$ ”, *JHEP* **07** (2016) 105, doi:10.1007/JHEP07(2016)105, arXiv:1605.00542.
- [74] ATLAS Collaboration, “Measurement of the Z boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector”, *JHEP* **1409** (2014) 145, doi:10.1007/JHEP09(2014)145, arXiv:1406.3660.
- [75] T. Gleisberg et al., “Event generation with SHERPA 1.1”, *JHEP* **02** (2009) 007, doi:10.1088/1126-6708/2009/02/007, arXiv:0811.4622.

- [76] S. Hoeche, S. Schumann, and F. Siegert, “Hard photon production and matrix-element parton-shower merging”, *Phys. Rev.* **D81** (2010) 034026, doi:10.1103/PhysRevD.81.034026, arXiv:0912.3501.
- [77] C. Balazs, E. L. Berger, P. M. Nadolsky, and C. P. Yuan, “Calculation of prompt diphoton production cross-sections at Tevatron and LHC energies”, *Phys. Rev.* **D76** (2007) 013009, doi:10.1103/PhysRevD.76.013009, arXiv:0704.0001.
- [78] T. Binoth, J. P. Guillet, E. Pilon, and M. Werlen, “A Full next-to-leading order study of direct photon pair production in hadronic collisions”, *Eur. Phys. J.* **C16** (2000) 311–330, doi:10.1007/s100520050024, arXiv:hep-ph/9911340.
- [79] W. Lampl et al., “Calorimeter Clustering Algorithms: Description and Performance”, *ATL-LARG-PUB-2008-002* (Apr, 2008). <http://cdsweb.cern.ch/record/1099735>.
- [80] ATLAS Collaboration, “Expected photon performance in the ATLAS experiment”, *ATL-PHYS-PUB-2011-007* (Apr, 2011). <http://cdsweb.cern.ch/record/1345329>.
- [81] ATLAS Collaboration, “Measurement of Higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector”, *Phys. Rev.* **D90** (2014) 112015, doi:10.1103/PhysRevD.90.112015, arXiv:1408.7084.
- [82] ATLAS Collaboration, “Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data”, *Eur. Phys. J.* **C74** (2014) 3071, doi:10.1140/epjc/s10052-014-3071-4, arXiv:1407.5063.
- [83] ATLAS Collaboration, “Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at $\sqrt{s} = 13$ TeV”, *ATL-PHYS-PUB-2016-015* (Aug, 2016). <https://cdsweb.cern.ch/record/2203514>.
- [84] A. Hoecker et al., “TMVA: Toolkit for Multivariate Data Analysis”, *PoS ACAT* (2007) 040, arXiv:physics/0703039.
- [85] ATLAS Collaboration, “Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data”, arXiv:1606.01813. Submitted to *Eur. Phys. J. C*.
- [86] ATLAS Collaboration, “Photon identification in 2015 ATLAS data”, *ATL-PHYS-PUB-2016-014* (Aug, 2016). <https://cdsweb.cern.ch/record/2203125>.
- [87] M. Cacciari, G. P. Salam, and S. Sapeta, “On the characterisation of the underlying event”, *JHEP* **04** (2010) 065, doi:10.1007/JHEP04(2010)065, arXiv:0912.4926.
- [88] ATLAS Collaboration, “Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector”, *Phys. Rev.* **D83** (2011) 052005, doi:10.1103/PhysRevD.83.052005, arXiv:1012.4389.

- [89] ATLAS Collaboration, “Search for scalar diphoton resonances with 15.4 fb^{-1} of data collected at $\sqrt{s}=13 \text{ TeV}$ in 2015 and 2016 with the ATLAS detector”, *ATLAS-CONF-2016-059* (Aug, 2016). <https://cdsweb.cern.ch/record/2206154>.
- [90] ATLAS Collaboration, “Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data”, *ATLAS-CONF-2014-032* (Jun, 2014). <https://cdsweb.cern.ch/record/1706245>.
- [91] ATLAS Collaboration, “Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data”, *Eur. Phys. J. C* **74** (2014) 2941, doi:10.1140/epjc/s10052-014-2941-0, arXiv:1404.2240.
- [92] ATLAS Collaboration, “Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data”, *Eur. Phys. J. C* **74** (2014) 3130, doi:10.1140/epjc/s10052-014-3130-x, arXiv:1407.3935.
- [93] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm”, *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- [94] ATLAS Collaboration, “Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7 \text{ TeV}$ with the ATLAS detector”, *Eur. Phys. J. C* **75** (2015) 17, doi:10.1140/epjc/s10052-014-3190-y, arXiv:1406.0076.
- [95] M. Cacciari and G. P. Salam, “Pileup subtraction using jet areas”, *Phys. Lett. B* **659** (2008) 119–126, doi:10.1016/j.physletb.2007.09.077, arXiv:0707.1378.
- [96] ATLAS Collaboration, “Pile-up subtraction and suppression for jets in ATLAS”, *ATLAS-CONF-2013-083* (Aug, 2013). <http://cdsweb.cern.ch/record/1570994>.
- [97] ATLAS Collaboration, “Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data”, *ATLAS-CONF-2011-102* (Jul, 2011). <http://cdsweb.cern.ch/record/1369219>.
- [98] ATLAS Collaboration, “Calibration of the performance of b -tagging for c and light-flavour jets in the 2012 ATLAS data”, *ATLAS-CONF-2014-046* (Jul, 2014). <http://cdsweb.cern.ch/record/1741020>.
- [99] ATLAS Collaboration, “Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV”, *ATLAS-CONF-2013-082* (Aug, 2013). <http://cdsweb.cern.ch/record/1570993>.
- [100] ATLAS Collaboration, “Performance of Missing Transverse Momentum Reconstruction in ATLAS with 2011 Proton-Proton Collisions at $\sqrt{s} = 7 \text{ TeV}$ ”, *ATLAS-CONF-2012-101* (Jul, 2012). <http://cdsweb.cern.ch/record/1463915>.
- [101] ATLAS Collaboration, “Performance of the ATLAS Trigger System in 2010”, *Eur. Phys. J. C* **72** (2012) 1849, doi:10.1140/epjc/s10052-011-1849-1.

- [102] ATLAS Collaboration, “Search for the Standard Model Higgs boson in the two photon decay channel with the ATLAS detector at the LHC”, *Phys. Lett.* **B705** (2011) 452–470, doi:10.1016/j.physletb.2011.10.051, arXiv:1108.5895.
- [103] ATLAS Collaboration, “Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector”, *Eur. Phys. J.* **C75** (2015) 476, doi:10.1140/epjc/s10052-015-3685-1, 10.1140/epjc/s10052-016-3934-y, arXiv:1506.05669.
- [104] ATLAS Collaboration, “Measurements of fiducial and differential cross sections for Higgs boson production in the diphoton decay channel at $\sqrt{s} = 8$ TeV with ATLAS”, *JHEP* **09** (2014) 112, doi:10.1007/JHEP09(2014)112, arXiv:1407.4222.
- [105] M. Oreglia, “A Study of the Reactions $\psi' \rightarrow \gamma\gamma\psi$ ”. PhD thesis, SLAC, 1980.
- [106] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics”, *Eur. Phys. J.* **C71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, 10.1140/epjc/s10052-013-2501-z, arXiv:1007.1727.
- [107] A. L. Read, “Presentation of search results: The CL(s) technique”, *J. Phys.* **G28** (2002) 2693–2704, doi:10.1088/0954-3899/28/10/313.
- [108] ATLAS Collaboration, “Combined search for the Standard Model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector”, *Phys. Rev.* **D86** (2012) 032003, doi:10.1103/PhysRevD.86.032003, arXiv:1207.0319.
- [109] W. Verkerke and D. P. Kirkby, “The RooFit toolkit for data modeling”, in *13th International Conference for Computing in High-Energy and Nuclear Physics (CHEP03)*, 2003. arXiv:physics/0306116. CHEP-2003-MOLT007.
- [110] L. Moneta et al., “The RooStats Project”, *PoS ACAT2010* (2010) 057, arXiv:1009.1003.
- [111] ROOT Collaboration, “HistFactory: A tool for creating statistical models for use with RooFit and RooStats”, *CERN-OPEN-2012-016* (Jan, 2012). <https://cdsweb.cern.ch/record/1456844>.
- [112] ATLAS Collaboration, “Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC”, *Eur. Phys. J.* **C73** (2013) 2518, doi:10.1140/epjc/s10052-013-2518-3, arXiv:1302.4393.
- [113] I. W. Stewart and F. J. Tackmann, “Theory Uncertainties for Higgs and Other Searches Using Jet Bins”, *Phys. Rev.* **D85** (2012) 034011, doi:10.1103/PhysRevD.85.034011, arXiv:1107.2117.
- [114] S. N. Bernstein, “Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités”, *Comm. Soc. Math. Kharkov* **13** (1912) 1.

- [115] E. Gross and O. Vitells, “Trial factors or the look elsewhere effect in high energy physics”, *Eur. Phys. J.* **C70** (2010) 525–530, doi:10.1140/epjc/s10052-010-1470-8, arXiv:1005.1891.
- [116] D. L. Rainwater, R. Szalapski, and D. Zeppenfeld, “Probing color singlet exchange in Z + two jet events at the CERN LHC”, *Phys. Rev.* **D54** (1996) 6680–6689, doi:10.1103/PhysRevD.54.6680, arXiv:hep-ph/9605444.
- [117] ATLAS Collaboration, “Properties of Jets and Inputs to Jet Reconstruction and Calibration with the ATLAS Detector Using Proton-Proton Collisions at $\sqrt{s} = 7$ TeV”, *ATLAS-CONF-2010-053* (Jul, 2010). <http://cdsweb.cern.ch/record/1281310>.
- [118] ATLAS Collaboration, “Measurement of the Higgs boson mass from the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ channels with the ATLAS detector using 25 fb^{-1} of pp collision data”, *Phys. Rev.* **D90** (2014) 052004, doi:10.1103/PhysRevD.90.052004, arXiv:1406.3827.
- [119] ATLAS Collaboration, “Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC”, *Phys. Lett.* **B726** (2013) 88–119, doi:10.1016/j.physletb.2013.08.010, arXiv:1307.1427.
- [120] ATLAS, CMS Collaboration, “Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV”, *JHEP* **08** (2016) 045, doi:10.1007/JHEP08(2016)045, arXiv:1606.02266.
- [121] N. Kauer and G. Passarino, “Inadequacy of zero-width approximation for a light Higgs boson signal”, *JHEP* **08** (2012) 116, doi:10.1007/JHEP08(2012)116, arXiv:1206.4803.
- [122] CMS Collaboration, “Observation of the diphoton decay of the Higgs boson and measurement of its properties”, *Eur. Phys. J.* **C74** (2014) 3076, doi:10.1140/epjc/s10052-014-3076-z, arXiv:1407.0558.
- [123] CMS Collaboration, “Measurement of the properties of a Higgs boson in the four-lepton final state”, *Phys. Rev.* **D89** (2014) 092007, doi:10.1103/PhysRevD.89.092007, arXiv:1312.5353.
- [124] CMS Collaboration, “Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV”, *Eur. Phys. J.* **C75** (2015) 212, doi:10.1140/epjc/s10052-015-3351-7, arXiv:1412.8662.
- [125] ATLAS Collaboration, “Observation and measurement of Higgs boson decays to WW^* with the ATLAS detector”, *Phys. Rev.* **D92** (2015) 012006, doi:10.1103/PhysRevD.92.012006, arXiv:1412.2641.

- [126] ATLAS Collaboration, “Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector”, *JHEP* **04** (2015) 117, doi:10.1007/JHEP04(2015)117, arXiv:1501.04943.
- [127] CMS Collaboration, “Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states”, *JHEP* **01** (2014) 096, doi:10.1007/JHEP01(2014)096, arXiv:1312.1129.
- [128] CMS Collaboration, “Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons”, *JHEP* **05** (2014) 104, doi:10.1007/JHEP05(2014)104, arXiv:1401.5041.
- [129] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV”, *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.
- [130] CMS Collaboration, “Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV”, *JINST* **10** (2015) P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.
- [131] CMS Collaboration, “Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV”, *JINST* **10** (2015) P08010, doi:10.1088/1748-0221/10/08/P08010, arXiv:1502.02702.
- [132] P. Dauncey, M. Kenzie, N. Wardle, and G. Davies, “Handling uncertainties in background shapes: the discrete profiling method”, *JINST* **10** (2015) P04015, doi:10.1088/1748-0221/10/04/P04015, arXiv:1408.6865.
- [133] ATLAS Collaboration, “Search for Scalar Diphoton Resonances in the Mass Range 65 – 600 GeV with the ATLAS Detector in pp Collision Data at $\sqrt{s} = 8$ TeV”, *Phys. Rev. Lett.* **113** (2014) 171801, doi:10.1103/PhysRevLett.113.171801, arXiv:1407.6583.