

**Cross section measurements of the Higgs boson in the
diphoton decay channel using proton-proton collision data
recorded by the ATLAS detector at centre-of-mass
energies of 7 TeV and 8 TeV**

A dissertation presented

by

Jim Lacey

to

The Department of Physics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Physics

Carleton University

Ottawa, Ontario

December 2015

Thesis advisor

Alain Bellerive

Author

Jim Lacey

Abstract

In July of 2012, both the ATLAS and CMS experiments announced the discovery of a new particle consistent with the Standard Model (SM) Higgs boson. This has ushered in a new era of precision physics: measuring the Higgs boson's properties. Precisely measuring all the properties of the newly discovered Higgs boson will tell us if this is the particle predicted by the SM or another particle entirely. This thesis focuses on cross section measurements of the Higgs boson in the diphoton decay channel. The analysis is performed using 4.5 fb^{-1} and 20.3 fb^{-1} of proton-proton collision data recorded by the ATLAS detector at CERN's Large Hadron Collider in 2011 and 2012, at centre-of-mass energies of 7 TeV and 8 TeV, respectively.

Fiducial cross sections are measured and corrected for experimental effects such as detector acceptance and resolution. The results are compared to the latest state-of-the-art theoretical prediction for SM Higgs boson cross sections. The Higgs boson signal is extracted and the background subtracted by performing a signal plus background fit to the invariant diphoton mass spectrum. The $pp \rightarrow H \rightarrow \gamma\gamma$ inclusive fiducial cross section is measured to be 45.3 ± 16.9 (stat.) $^{+5.3}_{-5.3}$ (syst.) ± 0.8 (lumi.) fb at $\sqrt{s} = 7$ TeV and 43.2 ± 9.4 (stat.) $^{+3.2}_{-2.9}$ (syst.) ± 1.2 (lumi.) fb at $\sqrt{s} = 8$ TeV, for a Higgs boson of mass 125.36 GeV. Using the larger 8 TeV dataset, additional fiducial and differential cross section measurements of the Higgs boson are also reported.

This thesis will particularly focus on the following aspects of the Higgs boson cross

section measurements: (i) jet calibration and performance, (ii) a data-driven method for obtaining the diphoton background rate estimations required for deriving the background parameterization used in the signal extraction, (iii) extracting the Higgs boson signal by performing a simultaneous signal plus background likelihood fit to the diphoton invariant mass spectrum, and (iv) deriving the ‘unfolding’ factors used to translate the measurements to a set of detector corrected model-independent observables.

Acknowledgments

Firstly, I would like to express my sincere gratitude to my advisor, Prof. Alain Bellerive, for his patience and enthusiasm, and for his constant support throughout my doctoral studies and related research. He has been a constant source of motivation and encouragement, providing guidance when needed but always allowing me the freedom to forge my own path towards becoming an independent researcher. I could not imagine having had a better advisor and mentor for my doctoral studies.

My sincere thanks also goes to Dr. Dag Gillberg, for helping guide my progression throughout the duration of my doctoral studies, and always finding the time to answer my many questions and share his expertise in jet calibration, Higgs physics, and general data analysis. He played an integral role in my development as a researcher and is an inspirational physicist. It has been a pleasure to learn from him, and work with him.

I would also like to sincerely thank Dr. Andy Pilkington, for all his advice and support, particularly in the last year of my doctoral studies. His insight and guidance have contributed greatly to my progression as a physicist.

A sincere thank you also goes to Dr. Florian Bernlochner, for the many insightful discussions on Higgs physics and for sharing his expertise in statistical fitting procedures.

To my colleagues at Carleton University (Graham Cree, David Di Valentino, Tom McCarthy, Rick Ueno, and Kate Whalen), thank you all for many insightful conversations, both work and non-work related. Working with all of you has been an enriching experience and has helped make the last five years a truly memorable journey.

Last but not the least, I would like to thank my family: my parents, John Lacey and Terry Thomson, for their unwavering support and belief in my ability to achieve any goal I have ever set out to accomplish; and my wonderful partner, Deanna MacLennan, who has been

a unfailing source of support and encouragement throughout my graduate studies, putting up with me through the ups and downs, and all the long hours of work.

Statement of Originality

I hereby certify that I am the sole author of this thesis and that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights. The following is a summary of the author's specific contributions to the measurements presented in this thesis and furthermore to the ATLAS experiment in general:

Higgs Physics

2014-2015: Measurement of Higgs boson cross sections

I was the principal analyzer for the 7 TeV result, where I was responsible for all analysis elements. Main analyzer on the recent 13 TeV fiducial cross section analysis using 3.2 fb^{-1} of data recorded in 2015. For the 13 TeV analysis, I was responsible for providing fiducial correction factors, evaluating experimental uncertainties, and providing data-driven background rate estimations. See the conference notes ATLAS-CONF-2015-060 and ATLAS-CONF-2015-069.

2013-2014: Fiducial and differential Higgs boson cross section measurements

I played a primary role in the 8 TeV analysis that reported the first measurements of fiducial and differential cross sections of the Higgs boson. I provided data-driven background rate estimations in fiducial regions characterized by the level of hadronic activity. Refer to the publication JHEP 09 (2014) 112.

2013-present: diphoton background rate determination

I was responsible for providing the diphoton background rate estimates for the 8 TeV and 13 TeV analyses. Diphoton background rate estimations are crucial for parameterizing the background functions used in extracting the Higgs boson signal. Additionally, these background estimations benefit the whole $H \rightarrow \gamma\gamma$

analysis group, consisting of about 100 physicists from around the world. See the publication JHEP 09 (2014) 112 and the conference note ATLAS-CONF-2015-060.

Jet Energy Scale Calibration

2011-present: η -intercalibration

I led the effort towards measuring the forward Jet Energy Scale (JES) and resolution using Run-I ATLAS data. I developed the *in-situ* forward jet calibration method for obtaining residual correction factors for forward jets, and also the procedures used to assess the associated systematic uncertainties. The efforts led to the first ATLAS measurement of the jet energy resolution using dijet events and also resulted in a reduction (of about a factor of two) on the JES uncertainties on forward jets. I have been the primary contributor and co-editor of two public conference notes on this topic (ATLAS-CONF-2015-017 and ATLAS-CONF-2012-124), with the results forming a significant contribution to other publications and conference notes, for example see EPJC (2015) 75:17 and ATLAS-CONF-2015-037.

2011-2013: Pileup suppression and subtraction for hadronic jets

I contributed significantly to the development of the jet area based pileup subtraction technique that is currently used in ATLAS. I provided detailed performance studies of the method and data validation of the primary inputs. Additionally, I provided the dijet based topology uncertainty on the procedure. Refer to conference note ATLAS-CONF-2013-083.

Table of Contents

Cover Page	i
Abstract	ii
Acknowledgments	iv
Statement of Originality	vi
Table of Contents	vii
List of Figures	5
List of Tables	9
1 Introduction	11
2 Theoretical context	16
2.1 The Standard Model	17
2.2 The Higgs mechanism	21
2.3 Higgs boson production at the LHC	28
2.4 Higgs decay modes	32
2.5 The Higgs boson discovery	34
2.6 Summary	35
3 The Large Hadron Collider and the ATLAS Detector	36
3.1 The Large Hadron Collider	37
3.2 The ATLAS detector	42
3.2.1 The ATLAS coordinate system	43
3.2.2 The inner detector	45
3.2.3 The calorimeter system	48
3.2.4 The muon spectrometer	52
3.3 The trigger system	53
3.4 Monte Carlo simulation of ATLAS data	54
4 Event reconstruction in ATLAS	56
4.1 Jet reconstruction	58

4.2	Electron reconstruction	60
4.3	Photon reconstruction	61
4.3.1	Offline photon identification	62
4.4	Muon reconstruction	62
4.5	Missing E_T reconstruction	64
4.6	Summary	64
5	Hadronic jet calibration in ATLAS	67
5.1	Forward jet energy scale calibration using dijet events: η -intercalibration	70
5.1.1	Monte Carlo simulation	70
5.1.2	Event selection	71
5.1.3	Dijet balance	72
5.1.4	The matrix method	74
5.1.5	Comparison of matrix and central reference methods	76
5.1.6	Comparison of data with simulation	77
5.1.7	Derivation of a residual Jet Energy Scale correction	77
5.1.8	Systematic uncertainties	80
5.2	Summary	83
6	Event and physics object selection for the $H \rightarrow \gamma\gamma$ analysis	85
6.1	Dataset and simulated signal samples	86
6.2	Trigger, data quality, and primary vertex selection applied to data and reconstructed MC	87
6.3	Reconstruction-level physics object selection	88
6.3.1	Photons	88
6.3.2	Electrons and muons	89
6.3.3	Jets and missing transverse energy (E_T^{miss})	89
6.4	Particle-level physics object selection	90
6.4.1	Particle-level photons	91
6.4.2	Particle-level electrons and muons	91
6.4.3	Particle-level jets and missing transverse energy (E_T^{miss})	92
6.4.4	Particle-level object overlap removal	92
6.5	Corrections to the simulation	93
6.6	Definition of fiducial regions	94
6.7	Dalitz decays	96
7	Data-driven background rate estimates in the $H \rightarrow \gamma\gamma$ channel	97
7.1	The $2 \times 2D$ Sideband Method	98
7.2	Systematic uncertainties	105
7.3	Results	107
7.4	Summary	113

8	Higgs boson signal extraction in the $H \rightarrow \gamma\gamma$ channel	114
8.1	Background and signal modelling of $m_{\gamma\gamma}$	115
8.1.1	Non-resonant background modelling of $m_{\gamma\gamma}$	116
8.1.1.1	Choice of non-resonant background $m_{\gamma\gamma}$ parametrization	116
8.1.2	Signal modelling of $m_{\gamma\gamma}$	120
8.1.2.1	Global resolution model	122
8.1.2.2	Signal parameter uncertainties	123
8.2	Higgs boson signal extraction via a simultaneous s+b fit to the $m_{\gamma\gamma}$ spectrum	127
8.3	Systematic uncertainties on the signal extraction	128
8.3.1	Mass resolution uncertainties	129
8.3.2	Mass energy scale uncertainties	130
8.3.3	Background modelling	131
8.3.4	Summary of systematic uncertainties affecting the signal PDF	131
8.3.5	Separation of statistical and systematic uncertainties on the yield	133
8.4	Choice of the Higgs boson mass used in the signal extraction	133
8.5	Fit results	134
8.5.1	Fit results: impact of the chosen Higgs boson mass on the extracted signal yield	141
8.5.1.1	Scan across fixed Higgs boson masses	141
8.5.1.2	Signal yields corresponding to the Run-I ATLAS+CMS combined Higgs boson mass	142
8.6	Summary of the extracted signal yields and uncertainties	142
9	Correction factors	144
9.1	Higgs boson cross section definition	146
9.2	Correcting for detector effects ("unfolding factors")	146
9.2.1	Particle-level photon isolation	148
9.3	Systematic uncertainties on the unfolding factors	153
9.3.1	Photon selection	153
9.3.2	Detector modelling	154
9.3.2.1	Photon and electron energy scale and resolution	155
9.3.2.2	Jet energy scale and resolution uncertainties	155
9.3.2.3	Jet vertex fraction (JVF)	157
9.3.2.4	Missing transverse energy uncertainties	157
9.3.3	Theoretical modelling	157
9.3.3.1	Signal composition	158
9.3.3.2	Generator dependence: matrix element (ME), parton-shower (PS), and underlying event (UE)	159
9.3.3.3	Data-driven Higgs boson p_T and rapidity weighting	160
9.3.3.4	Combination of generator modelling uncertainties	161
9.4	Unfolding factors and uncertainties	161

9.5	Fiducial acceptance factors and non-perturbative correction factors	162
9.5.1	Fiducial acceptance factors	162
9.5.2	Non-Perturbative correction factor	163
10	Results	165
10.1	Theoretical predictions	165
10.2	Higgs boson cross section measurements	166
11	Summary	176
	Appendix	182
A	Simulated samples used in the $H \rightarrow \gamma\gamma$ analysis	182
A.1	Nominal simulated signal samples	182
A.2	Additional simulated signal samples	184
A.3	Simulated background samples	185
	Bibliography	186

List of Abbreviations and Acronyms

ALICE A Large Ion Collider Experiment

CSC Cathode Strip Chamber

EM Electromagnetic

EWSB Spontaneous Electroweak Symmetry Breaking

FCal The Forward LAr Calorimeter

GeV Giga electron volt

ID Inner Detector

LAr Liquid Argon

LCW Local Cell Energy Weighting

LHC Large Hadron Collider

MDT Monitored Drift Tube

MS Muon Spectrometer

PS Proton Synchrotron

PSB Proton Synchrotron Booster

PV Primary Vertex

QCD Quantum Chromodynamics

QED Quantum Electrodynamics

QFT Quantum Field Theory

RF Radio-Frequency

RPC Resistive Plate Chamber

SCT Semiconductor Tracker

SPS Super Proton Synchrotron

TeV Tera electron volt

TGC Thin Gap Chamber

TR Transition Radiation

TRT Transition Radiation Tracker

VEV Vacuum Expectation Value

JER Jet Energy Resolution

ATLAS A Toroidal LHC Apparatus

CERN European Organization for Nuclear Research

CMS Compact Muon Solenoid

EW Electroweak

fb Femtobarn

FSR Final State Radiation

GF Gluon Fusion

GSC Global Sequential Calibration

JES Jet Energy Scale

JVF Jet Vertex Fraction

LO Leading Order

MC Monte Carlo

MPI Multiple Parton Interactions

NLO Next-to-Leading Order

NNLL Next-to-Next-to-Leading Logarithmic Order

NNLO Next-to-Next-to-Leading Order

NP Non-Perturbative

PDF Parton Distribution Functions

SM Standard Model of particle physics

VBF Vector Boson Fusion

ttH Associated production of a Higgs boson with a top anti-top pair

WH Associated production of a Higgs boson with a W boson

ZH Associated production of a Higgs boson with a Z boson

List of Figures

1.1	An cartoon illustrating the signal and background composition of the $m_{\gamma\gamma}$ spectrum used to extract the Higgs boson signal. The $H \rightarrow \gamma\gamma$ signal is depicted in yellow as a resonance over a smooth non-resonant background composed of generic $\gamma\gamma$ (grey), γ +jet (blue), and jet+jet (red) events. . . .	14
2.1	Higgs boson production processes at the LHC: (a) $pp \rightarrow H$ via gluon fusion (GF), (b) $pp \rightarrow qqH$ via weak vector boson fusion (VBF), (c) associated production with a W^\pm or Z boson ($pp \rightarrow WH$ or $pp \rightarrow ZH$), and (d) associated production with a $t\bar{t}$ pair ($pp \rightarrow t\bar{t}H$). Note: time increases from left to right.	30
2.2	Theoretical predictions for the dominant Higgs boson production processes at the LHC, corresponding to a centre-of-mass energies of (a) 7 TeV and (b) 8 TeV.	31
2.3	Theoretical predictions for the branching ratio of the various Higgs boson decay modes as a function of the Higgs boson mass. The predicted branching ration for a Higgs boson decaying to two photons is $0.228 \pm 0.011 \%$	33
2.4	Leading-order Feynman diagrams for the Higgs boson decaying to two photons.	34
3.1	Illustration of the LHC accelerator complex.	39
3.2	Integrated luminosity as a function of time delivered to (green), recorded by ATLAS (yellow), and verified to be good quality data (blue) during stable beam conditions for proton-proton collisions at centre-of-mass energies of 7 and 8 TeV in 2011 and 2012, respectively.	41
3.3	Schematic view of the ATLAS detector.	44
3.4	Cartoon illustrating the ATLAS coordinate system.	45
3.5	A cut-away view of the ATLAS inner detector.	46
3.6	A cut-away view of the ATLAS calorimeter system.	49
3.7	A cut-away view of the ATLAS muon system.	53

- 5.1 Two representative asymmetry distributions corresponding to the central reference method, obtained for jets with $25 \text{ GeV} \leq p_{\text{T}}^{\text{avg}} < 40 \text{ GeV}$ and a probe jet pseudo rapidity in the range of (a) $0.3 \leq \eta_{\text{probe}} < 0.8$ and (b) $3.2 \leq \eta_{\text{probe}} < 3.6$. The red and blue histograms correspond to data collected by central and forward triggers, respectively, with the black points being the sum of the two. The thick magenta line is the Gaussian function used to obtain the mean and width of the asymmetry distribution. N_{eff} is the number of events in the distribution, while A, R, and σ (left side, black text) are the mean asymmetry, the response, and the root mean square (RMS) of the asymmetry distribution, respectively. The results of the Gaussian fit are given on the left (magenta text). 73
- 5.2 Relative response measured using the matrix and the central reference methods for anti- k_t jets with $R = 0.4$ calibrated with the EM+JES scheme as a function of the probe jet η . Results are presented for $65 \text{ GeV} \leq p_{\text{T}}^{\text{avg}} < 85 \text{ GeV}$ and $270 \text{ GeV} \leq p_{\text{T}}^{\text{avg}} < 330 \text{ GeV}$. The green and red dashed lines indicate $1 \pm 2\%$ and $1 \pm 5\%$, respectively. 76
- 5.3 Relative jet response, $1/c$, as a function of the jet η for anti- k_t jets with $R = 0.4$ calibrated with the EM+JES scheme, separately for $25 \text{ GeV} < p_{\text{T}}^{\text{avg}} < 40 \text{ GeV}$, $85 \text{ GeV} < p_{\text{T}}^{\text{avg}} < 115 \text{ GeV}$, $220 \text{ GeV} < p_{\text{T}}^{\text{avg}} < 270 \text{ GeV}$ and $760 \text{ GeV} < p_{\text{T}}^{\text{avg}} < 1200 \text{ GeV}$. The lower parts of the figures show the ratios between the data and MC relative response. These measurements are performed using the matrix method. The green and red dashed lines indicate $1 \pm 2\%$ and $1 \pm 5\%$, respectively. 78
- 5.4 Relative jet response, $1/c$, as a function of the jet p_{T} for anti- k_t jets with $R = 0.4$ calibrated with the EM+JES scheme, separately for $-1.5 \leq \eta_{\text{det}} < -1.2$ and $2.1 \leq \eta_{\text{det}} < 2.4$. The lower parts of the figures show the ratios between the data and MC relative response. The green and red dashed lines indicate $1 \pm 2\%$ and $1 \pm 5\%$, respectively. 79
- 5.5 Summary of uncertainties on the intercalibration as a function of the jet η_{det} for anti- k_t jets with $R = 0.4$ calibrated with the EM+JES scheme, separately for (a) $p_{\text{T}} = 35 \text{ GeV}$ and (b) $p_{\text{T}} = 300 \text{ GeV}$. The individual components are added in quadrature to obtain the total uncertainty. The modelling uncertainty is the dominant component. 82
- 5.6 Final $\sqrt{s} = 8 \text{ TeV}$ η -intercalibration and associated uncertainties as a function of the jet pseudorapidity for anti- k_t jets with $R = 0.4$ calibrated with the EM+JES scheme, separately for (a) $25 \leq p_{\text{T}}^{\text{avg}} < 40 \text{ GeV}$ and (b) $220 \leq p_{\text{T}}^{\text{avg}} < 270 \text{ GeV}$. The green and red dashed lines indicate $1 \pm 2\%$ and $1 \pm 5\%$, respectively. 83

7.1	Illustration of the 1 signal and 15 background control regions of the 2x2D sideband method, with the leading (sub-leading) photon candidate in red (blue). All regions are enriched with various compositions of signal and background contributions ($\gamma\gamma$, γ +jet, jet+ γ , and jet+jet). Regions enriched with a particular background type are indicated by the text in the upper left corner of the ‘cells’ representing each region. Text in lower right of each cell gives each photon candidates identification. See the text for the meaning of the grey-shaded regions.	100
7.2	(a) and (b) illustrate the 2D sidebands corresponding to the leading and sub-leading photon candidates, respectively.	101
7.3	Photon efficiencies from the SHERPA simulation (Jet $p_T > 30$ GeV & $ y_{\text{jet}} < 4.4$) as a function of $m_{\gamma\gamma}$ for (a) the inclusive selection, (b) $N_{\text{jet}} = 0$, (c) $N_{\text{jet}} = 1$, (d) $N_{\text{jet}} = 2$, and (e) $N_{\text{jet}} \geq 3$	110
7.4	Data-driven fractional background decomposition (Jet $p_T > 30$ GeV & $ y_{\text{jet}} < 4.4$) as a function of $m_{\gamma\gamma}$ for (a) the inclusive selection, (b) $N_{\text{jet}} = 0$, (c) $N_{\text{jet}} = 1$, (d) $N_{\text{jet}} = 2$, and (e) $N_{\text{jet}} \geq 3$. The left y-axis is the number of events per bin and corresponds to the stacked histogram. The dashed lines represent the purity of each type of event and corresponds to the right y-axis.	111
7.5	(a) gives the average photon efficiencies as a function of N_{jet} . (b) gives the background decomposition as a function of N_{jet} . The left vertical axis is the number of events per bin and the right vertical axis is the purity of each background process per bin.	112
8.1	Comparison of the diphoton mass spectra between data and smeared simulation sample for the inclusive selection (label: CP0) and in the different N_{jet} bins: zero, one, two and three or more jets, labelled CP1-CP4, respectively.	119
8.2	Results from simultaneous fits to the $m_{\gamma\gamma}$ spectrum over the ten ‘‘mass points’’ spanning Higgs boson masses from 105 GeV to 150 GeV at (a) 7 TeV and (b) 8 TeV. The functional form is given in Equation 8.2 where the parameters are constrained as a function of $m_{\gamma\gamma}$ according to Equation 8.5.	124
8.3	Results from simultaneous fits to the $m_{\gamma\gamma}$ spectrum over the ten ‘‘mass points’’ spanning Higgs boson masses from 105 GeV to 150 GeV in at 8 TeV. The functional form is given in Equation 8.2 where the parameters are constrained as a function of $m_{\gamma\gamma}$ according to Equation 8.5. The fit is performed separately for each N_{jet} bin.	125
8.4	Signal plus background fit to the inclusive $m_{\gamma\gamma}$ spectrum for data recorded at (a) $\sqrt{s} = 7$ TeV and (b) $\sqrt{s} = 8$ TeV.	137
8.5	Signal plus background fit to the $m_{\gamma\gamma}$ spectrum corresponding to (a) $N_{\text{jet}} = 0$, (b) $N_{\text{jet}} = 1$, (c) $N_{\text{jet}} = 2$, and (d) $N_{\text{jet}} \geq 0$, for $\sqrt{s} = 8$ TeV.	138

8.6	The fitted signal yields in data compared to the expected number of reconstructed signal events from simulation corresponding to the (a) 7 TeV and (b) 8 TeV datasets for the inclusive selection, and (c) in bins of N_{jet} for the 8 TeV dataset. The signal contribution due to non-GF processes is denoted as XH ($XH \equiv VBF + WH + ZH + ttH$).	139
8.7	Pulls on the nuisance parameters for (a) the 7 TeV inclusive fit, (b) the 8 TeV inclusive fit, and (c) the simultaneous fit of the N_{jet} spectrum at 8 TeV. See Table 8.2 for the definitions of the various nuisance parameters in the figures.	140
8.8	Scan across fixed Higgs boson masses for the inclusive selection at (a) 7 TeV and (b) 8 TeV. The black solid line is the central value of the signal yield and the red dashed lines are the 1σ error bands.	141
9.1	An illustration of the application of the different correction factors. Unfolding (correcting for detector effects) takes the data from the reconstruction-level to the particle-level within a fiducial region or bin of a differential distribution. The kinematic acceptance and isolation correction factors translate inclusive theoretical predictions at the parton-level to fiducial theoretical predictions at the parton-level. The non-perturbative (NP) correction factor translates fiducial parton-level theoretical predictions to fiducial particle-level theoretical predictions.	145
9.2	Calorimeter isolation energy against particle-level isolation energy for simulated events at (a) 7 TeV and (b) 8 TeV. The calorimeter isolation energy was profiled against truth isolation energy (black points) and a linear fit (red line) was performed over the range $5 \text{ GeV} \leq \text{particle-level } E_{\text{T}}^{\text{iso}} \leq 19 \text{ GeV}$. These results are independent of the track isolation transverse momentum selection.	151
9.3	Fiducial unfolding factors as a function of the Higgs boson production mode with and without the particle-level photon isolation: (a) 7 TeV and (b) 8 TeV.	152
10.1	Fiducial cross sections for $pp \rightarrow H \rightarrow \gamma\gamma$ as a function of \sqrt{s}	168
10.2	Fiducial cross sections (and limits) for $pp \rightarrow H \rightarrow \gamma\gamma$ measured at $\sqrt{s} = 8 \text{ TeV}$ compared to theoretical predictions.	170
10.3	The differential cross section for $pp \rightarrow H \rightarrow \gamma\gamma$ as a function of the jet multiplicity for sets satisfying $p_{\text{T}} > 30 \text{ GeV}$	173
10.4	The differential cross section for $pp \rightarrow H \rightarrow \gamma\gamma$ as a function of the diphoton transverse momentum, $p_{\text{T}}^{\gamma\gamma}$	174
10.5	The differential cross section for $pp \rightarrow H \rightarrow \gamma\gamma$ as a function of the absolute rapidity of the diphoton system, $ y_{\gamma\gamma} $	175
11.1	Total $pp \rightarrow H$ cross sections measured at different center-of-mass energies compared to LHC-XS theory predictions at up to NNLO in QCD.	181

List of Tables

4.1	Calorimeter shower profile variables used for Loose, Loose', and Tight photon identification selection.	66
7.1	Notations and descriptions used for the photon candidate categorization. . .	99
7.2	Extracted background yields for each event type in the signal region, including statistical and systematic uncertainties.	109
7.3	8 TeV jet fake rates (f_1, f_2) obtained from the minimization. The fake rates (Equation 7.6) are defined as the efficiency for a not Tight identified jet to pass the isolation selection.	109
7.4	Fractional decomposition of dominant backgrounds in the signal region, in %. Here γ -jet $\equiv \gamma$ -jet + jet- γ	109
8.1	Signal PDF nuisance parameters obtained from the global resolution model corresponding to the 7 TeV inclusive fiducial region, for a Higgs boson mass of 125.36 GeV.	123
8.2	Naming convention and definitions for the systematic uncertainty sources affecting the signal PDF and included in the fit.	132
8.3	Presented in the table are the event yields in each fiducial region: the total number of events selected in data (N_{data}), the signal yield obtained from the s+b fit to data (ν_i^{sig}), and the expected yield from the MC simulation ($N_{\text{sim}}^{\text{sig}}$). The extracted signal yields (ν_i^{sig}) are always larger than the expected number of signal events ($N_{\text{sim}}^{\text{sig}}$), reflecting the fact that the measured signal strength is larger than the expected signal strength in the $H \rightarrow \gamma\gamma$ decay channel, though consistent with expectations for a Standard Model Higgs boson within the statistical precision of the measurements. The quoted uncertainty on the fitted yield represents both the statistical and systematic uncertainties on the signal extraction. The uncertainty on the expected yields takes into account both theoretical and experimental uncertainties. . .	136

8.4	Percent difference in the extracted signal yields due to using the Run-I ATLAS+CMS combined Higgs boson mass instead of the Run-I ATLAS $H \rightarrow ZZ^* \rightarrow 4\ell + H \rightarrow \gamma\gamma$ combined Higgs boson mass value used in the current measurements.	142
8.5	Signal yield associated uncertainties measured using the 7 TeV dataset, for $m_H = 125.36$ GeV.	143
8.6	Signal yields measured in data and their associated uncertainties corresponding to the 8 TeV analysis.	143
9.1	Production mode dependence of the particle-level photon isolation.	149
9.2	Unfolding factors (c) for the inclusive fiducial region, for each SM $H \rightarrow \gamma\gamma$ production process and the total. Unfolding factors are given for each particle-level isolation definition. The 8 TeV unfolding factors are also included for comparison.	150
9.3	Signal composition uncertainties on the 7 TeV unfolding factor.	158
9.4	Uncertainties on the 7 TeV unfolding factor due to generator modelling, in percent.	160
9.5	Unfolding factors and uncertainties for the 7 TeV and 8 TeV analyses, corresponding to a particle-level photon isolation cut of $E_T^{\text{iso}} < 14$ GeV.	161
9.6	Diphoton fiducial acceptance factors in percent. The kinematic acceptance factor is defined as the probability to fulfil the diphoton kinematic acceptance criteria: $ y < 2.37$ and $p_T/m_{\gamma\gamma} < 0.35(0.25)$ for the leading (sub-leading) photon. The isolation efficiency factor is defined as the probability for both photons to pass the particle-level isolation requirement ($E_T^{\text{iso}} < 14$ GeV) for events that fulfil the diphoton kinematic criteria.	163
9.7	Non-perturbative correction factors that account for the impact of hadronization and underlying event, in percent.	164
10.1	Measured cross sections and 95 % cross-section limits in the various fiducial regions for the 7 TeV and 8 TeV analyses. The fiducial regions are defined in Section 6.6. The uncertainty associated with the luminosity (lumi.) is quoted separately along with the statistical (stat.) and systematic (syst.) uncertainties.	169
10.2	Predicted cross sections and 95 % cross-section limits in the various fiducial regions for the 7 TeV and 8 TeV analyses. The fiducial regions are defined in Section 6.6.	169
10.3	Uncertainties on the measured cross sections, expressed as percentages, corresponding to the inclusive fiducial region, for the 7 TeV analysis.	171
10.4	Uncertainties on the measured cross sections, expressed as percentages, corresponding to the inclusive, $N_{\text{jet}} \geq 1$, $N_{\text{jet}} \geq 2$, $N_{\text{jet}} \geq 3$ and VBF-enhanced fiducial regions, for the 8 TeV analysis.	171

Chapter 1

Introduction

The goal of particle physics is to understand the fundamental forces between the elementary particles that make up all of the matter in the universe. The Standard Model (SM) of particle physics is a theoretical framework that provides an extremely precise and accurate description of the elementary particles and their interactions, with nineteen free parameters whose numerical values are established by experiment. It does an outstanding job at describing all experimental data. A key feature of the SM is the Higgs mechanism, which was postulated in the 1960's to provide an explanation of how the elementary particles acquire mass. The Higgs mechanism predicts the existence of a new particle, which is known as the Higgs boson. The Higgs boson has been the focus of numerous dedicated experimental searches over the past half century.

In order to produce Higgs bosons, particle collisions at extremely high energies are required. The Large Hadron Collider (LHC), located at the European Organization for Nuclear Research (CERN), is the world's largest and most powerful particle collider, and is

able to accelerate protons up to approximately the speed of light, providing proton-proton collisions at the energy frontier. Positioned around the LHC are two experiments, ATLAS and CMS, capable of identifying the Higgs boson through the unique detector signatures produced by its decay products. In 2012, both ATLAS and CMS announced the discovery of a new particle consistent with the expectations for a SM Higgs boson [1, 2]. This has ushered in a new era of precision physics: measuring the Higgs boson's properties and testing the consistency of the data with theoretical predictions. Precisely investigating all the properties of the newly discovered Higgs boson will firmly establish whether the newly discovered particle is exactly the Higgs boson predicted by the SM or another particle entirely.

Cross section measurements provide an ideal probe for testing the consistency of the data with theoretical predictions, and also provide an alternative method to study particle properties. Through unfolding the data to the particle-level, the measurements are designed to be as model-independent as possible, allowing direct comparison with various theoretical predictions. The cross-section is defined as:

$$\sigma_i = \frac{\nu_i^{\text{data}}}{c_i \cdot \mathcal{L}_{\text{int}}}, \quad (1.1)$$

where i denotes the particular fiducial region or bin of a differential distribution being considered, ν_i^{data} being the number of extracted signal events in region (or bin) i , c_i is a correction factor accounting for differences in the event yield between detector-level and particle-level due to experimental effects, and \mathcal{L}_{int} is the integrated luminosity of the recorded dataset. The integrated luminosity used in this analysis was provided by the

ATLAS group dedicated to measuring the luminosity recorded by the ATLAS detector and no further luminosity investigations are performed in the context of this analysis. This thesis focuses on measuring fiducial cross sections of the Higgs boson using a sample of events containing a Higgs boson that decays to a pair of photons. The key analysis ingredients are a precise extraction of the Higgs boson signal yield (N_i^{data}) and the determination of the unfolding factors (c_i). The Higgs boson signal is extracted from a simultaneous signal plus background (s+b) fit to the diphoton invariant mass spectrum ($m_{\gamma\gamma}$). The $m_{\gamma\gamma}$ distribution is background dominated, consisting primarily of three SM background processes with the Higgs boson signal appearing as an excess of events in the vicinity of the Higgs boson mass. A cartoon depicting the composition of the $m_{\gamma\gamma}$ spectrum is shown in Figure 1.1. The unfolding factors are obtained from simulation by comparing the event yields obtained at the reconstruction-level to those obtained at the particle-level. Cross sections obtained at centre-of-mass energies of 7 TeV and 8 TeV will be reported and compared to the latest state-of-the-art theoretical predictions for SM Higgs boson production. Naturally, the thesis focuses on areas in which the author made substantial contributions.

Chapter 2 provides a brief description of the SM, followed by a more detailed account of the Higgs mechanism and its role in elementary particle mass generation through the process of spontaneous electroweak symmetry breaking. Additionally, a description of the dominant Higgs boson production processes at the LHC is provided, followed by a description of the Higgs boson decay to a photon pair. Chapter 2 concludes with a brief discussion of the 2012 discovery of a SM like Higgs boson by both the ATLAS and CMS experiments.

Chapter 3 provides an overview of the Large Hadron Collider and a description of

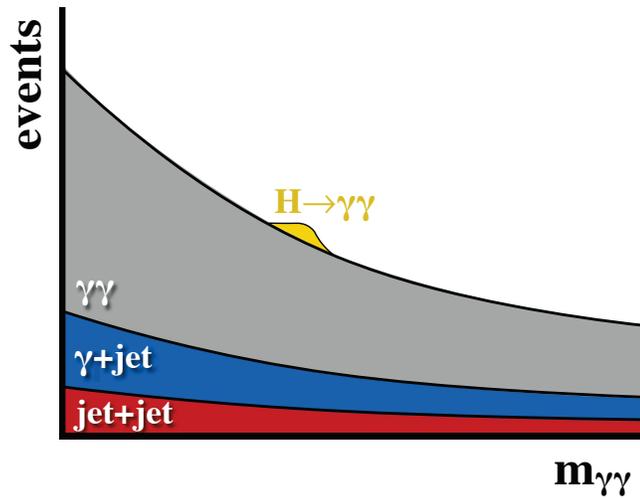


Figure 1.1: An cartoon illustrating the signal and background composition of the $m_{\gamma\gamma}$ spectrum used to extract the Higgs boson signal. The $H \rightarrow \gamma\gamma$ signal is depicted in yellow as a resonance over a smooth non-resonant background composed of generic $\gamma\gamma$ (grey), γ +jet (blue), and jet+jet (red) events.

the ATLAS detector. The ATLAS coordinate system is introduced, followed by detailed descriptions of each of the ATLAS sub-detectors and the data acquisition system. Chapter 4 describes the event and physics object reconstruction used in the analysis.

Chapter 5 provides an overview of hadronic jet energy calibration. Particular focus is paid to the *in-situ* energy calibration of jets that are produced at a low angle with respect to the proton beam-line (so called ‘forward jets’). The details of the *in-situ* calibration procedure and the assessment of the systematic uncertainties will be provided, followed by a summary of the results.

Chapter 6 provides the details of the event and object selection for both reconstruction-level and particle-level physics objects, as well as the definition of the fiducial regions that

define the measurements.

Chapter 7 provides the details of the data-driven method used to obtain the background rate estimates in the baseline inclusive fiducial region and also for fiducial regions characterized by the amount of hadronic jet activity.

Chapter 8 gives an overview of the Higgs boson signal extraction via a maximum likelihood fit to the $m_{\gamma\gamma}$ spectrum. A detailed description of the signal and background modelling of the $m_{\gamma\gamma}$ spectrum is provided, followed by a description of the fitting procedure used to extract the signal yields and a summary of the fit results.

Chapter 9 describes the derivation of various correction factors. First, a discussion of the ‘unfolding’ procedure used to correct for experimental effects such as detector resolution and inefficiencies is presented. This is followed by a description of the derivation of fiducial acceptance factors and non-perturbative correction factors used to translate inclusive theoretical predictions at the parton-level to fiducial predictions at the particle-level in order to facilitate comparison with the measurements.

Chapter 10 presents the results, and finally, Chapter 11 gives a summary and a conclusion. Additionally, a description of the simulated samples used in the $H \rightarrow \gamma\gamma$ analysis is given in Appendix A.

Chapter 2

Theoretical context

The Standard Model (SM) of particle physics is a mathematical framework that is formulated in the language of Quantum Field Theory (QFT) [3, 4]. It describes all the fundamental properties of the universe (excluding gravity) in terms of interacting quantum fields, where individual elementary particles are representations of quantum field excitations that are localized in spacetime, and the forces between particles are due to interactions of their governing fields. Through this framework the SM describes all of the elementary particles that make up matter, along with the forces that mediate their interactions. The SM does an outstanding job at describing all experimental observations of elementary particles and their interactions, making unique predictions that have been extensively tested over the past half century. This chapter provides a brief overview of the general structure of the SM with an emphasis on the parts that are particularly relevant to this body of work; namely the role of the Higgs mechanism in the SM and the expectations for Higgs boson production at the LHC. A complete description of the SM is far beyond the scope of this document, and

the interested reader is referred to References [3, 4]. Throughout this work natural units are assumed, which means that all physical units are redefined by setting the fundamental constants to unity, for example $\hbar = c = 1$. Therefore, masses are reported in units of electron volts, typically giga electron volts (GeV) or tera electron volts (TeV) for the energy scales of the LHC .

2.1 The Standard Model

The three fundamental forces that are described by the SM are: the electromagnetic force, the weak force, and the strong force. Each force is mediated by the exchange of a “force carrying” particle, referred to as a vector boson. The strong force (Quantum Chromodynamics, QCD) is responsible for interactions between the fundamental building blocks of atomic nuclei, quarks and gluons, binding them together to form protons and neutrons. The electromagnetic (EM) force (Quantum Electrodynamics, QED) mediates the interaction of electrically charged particles, for example the interaction of an electron and proton in a hydrogen atom. The weak force mediates radioactive decay processes and also plays a major role in the nuclear reactions that power stars. The fourth force is gravity, which is responsible for the macroscopic structure of the universe. However, there is currently no accepted quantum theory of gravity and it is not included in the SM. Furthermore, gravity’s effect is completely negligible at the subatomic scale and therefore it is ignored in the context of this discussion, and the work that follows.

The SM contains twelve matter particles collectively referred to as fermions: there are six leptons and six quarks, each with half integer spin. In addition, each fermion has a

corresponding antiparticle with the same mass, opposite charge, and a reversed sign for all other quantum numbers. The six leptons consist of three massive charged particles, the electron (e^-), the muon (μ^-), and the tau (τ^-), along with three corresponding massless neutral particles, the electron-neutrino (ν_e), the muon-neutrino (ν_μ), and the tau-neutrino (ν_τ). There are six quarks, listed here in order of increasing mass: u , d , s , c , b , and t , referred to as ‘up’, ‘down’, ‘strange’, ‘charm’, ‘bottom’, and ‘top’, respectively. Three of the quarks (u, c, t) have electric charges of $+2/3$, while the other three quarks (d, s, b) have electric charges of $-1/3$. Quarks also carry an additional quantum number termed ‘colour charge’, with each quark possessing one of three possible colours, arbitrarily named: red, green, and blue. Colour confinement is a defining feature of the strong interaction and is a direct consequence of its fundamental behaviour: the strength of the strong force increases with separation and for separations greater than approximately a femtometer it becomes energetically favourable to “pull” quark-antiquark pairs from the vacuum. Therefore isolated quarks cannot be observed and are only observable in bound states consisting of two or three quarks forming a colour singlet, with these bound states referred to as mesons and hadrons, respectively. Of these bound states the most familiar are the proton (uud) and neutron (udd).

The quark and lepton families are each grouped into three generations sharing identical properties except for their mass. The fermionic fields have both left (L) and right (R) chiral components, transforming as $SU(2)$ doublets and singlets, respectively [3, 4].

$$\text{leptons :} \quad L_\ell = \begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}_L, \quad \ell_R, \quad \text{with } \ell = e, \mu, \tau.$$

$$\text{quarks :} \quad Q_q = \begin{pmatrix} q_i \\ q_j \end{pmatrix}_L, \quad q_{iR}, \quad q_{jR}, \quad \text{with } q_i = u, c, t \text{ and } q_j = d, s, b.$$

The weak interaction only couples to left-handed fermions, and it is this feature that is responsible for parity (P) violation in the weak interaction (also CP violation, with C being charge conjugation). The SM assumes neutrinos are massless and hence right chiral neutrinos do not appear; the situation is somewhat complicated by the experimental observation that neutrinos do indeed have mass. Extensions of the SM have been formulated that account for non-zero neutrino masses, however within the scope of this discussion and the work that follows the zero mass assumption is suitable, with potential effects due to non-zero mass neutrinos being irrelevant.

Gauge theories are used to describe the field interactions responsible for the forces, with each force mediated by the exchange of a spin-1 gauge boson. The SM contains twelve gauge bosons responsible for mediating the three forces. The strong force is mediated by an octet of electrically neutral and massless gluons, transforming according to the $SU(3)$ symmetry group. Gluons also carry colour charge and obey the restrictions of confinement discussed above in the context of quarks. The EM force is mediated by electrically neutral and massless photons (γ 's), transforming according to the $U(1)$ symmetry group, while the weak force is mediated by the massive weak vector bosons (W^\pm and Z), transforming according to the $SU(2)$ symmetry group.

The full theory is represented by the combined $SU(3)_c \times SU(2)_L \times U(1)_Y$ symmetry group, where C denotes colour, L represents the chiral nature of the $SU(2)$ couplings, and Y denotes the weak hypercharge. The unified electroweak interactions of the SM are represented by the $SU(2)_L \times U(1)_Y$ portion of the full symmetry group. In order to insure the theory is renormalizable, the SM Lagrangian is required to be invariant under local transformations of the $SU(3)_c \times SU(2)_L \times U(1)_Y$ gauge group. It is this feature that is

responsible for much of the SM's theoretical structure. The preservation of the Lagrangian's local gauge invariance is achieved through the introduction of a covariant derivative, with the structure constants of the Lie Algebra defining the form of the covariant derivative and the properties of the gauge field interactions [3, 4].

The requirement that the SM Lagrangian must be invariant under local $SU(3)_c \times SU(2)_L \times U(1)_Y$ gauge transformations imposes the constraint that the gauge fields in the Lagrangian must be massless, therefore in its basic form the theory predicts that all the fermions and gauge bosons should be massless. This prediction is in direct conflict with experimental observation, where it is a well established fact that the fermions as well as the W and Z gauge bosons are massive, with the gauge bosons actually being roughly 85 times the mass of the proton. Therefore a procedure for introducing mass terms to the SM Lagrangian is required, but any attempt to modify the Lagrangian by introducing mass terms by hand explicitly breaks the $SU(2)_L \times U(1)_Y$ gauge symmetry, making the theory non-renormalizable. However, if the $SU(2)_L \times U(1)_Y$ symmetry is broken spontaneously during the quantization of the theory, the Lagrangian remains invariant under the local gauge transformations. Furthermore, a consequence of this type of spontaneous symmetry breaking is that the gauge bosons of the spontaneously broken symmetry acquire a mass. The 'Higgs mechanism' provides exactly this type of prescription for spontaneously breaking the $SU(2)_L \times U(1)_Y$ symmetry, thereby generating the masses for the weak interaction gauge bosons while still preserving the Lagrangian's gauge invariance. The following section presents the details of the Higgs mechanism and the generation of mass for the elementary particles.

2.2 The Higgs mechanism

The electroweak interaction represents the unified description of the EM and weak forces. At low energies these two forces appear very distinct from one another, however above the unification threshold, $\mathcal{O}(100 \text{ GeV})$, these two forces merge into a single electroweak force, which is described by the $SU(2)_L \times U(1)_Y$ gauge group. The Lagrangian representing the $SU(2)_L \times U(1)_Y$ portion of the SM can be expressed as:

$$\mathcal{L}_{EW}^{SM} = -\frac{1}{4} (W_{\mu\nu}^a W_a^{\mu\nu} + B_{\mu\nu} B^{\mu\nu}) + i\bar{L}_\ell D_\mu \gamma^\mu L_\ell + i\bar{\ell}_R D_\mu \gamma^\mu \ell_R, \quad (2.1)$$

where $W_{\mu\nu}^a$ and $B_{\mu\nu}$ are the gauge field strengths of the weak and EM interactions, respectively, L_ℓ are the left-chiral lepton doublets, ℓ_R are the right-chiral leptons, γ^μ are the Dirac matrices, and D_μ is the covariant derivative.

The theoretical requirements of renormalizability and local gauge invariance of the SM Lagrangian, combined with the necessity to include mass terms for the fermions and weak interaction gauge bosons precipitated the need to develop the theoretical framework of spontaneous electroweak symmetry breaking (EWSB). The framework of spontaneous symmetry breaking was developed in the 1960's by three independent research groups; Peter Higgs [5]; Francois Englert and Robert Brout [6]; and Gerald Guralnik, Carl Hagen, and Tom Kibble [7]. The full development of spontaneous symmetry breaking in the context of the unified electroweak interaction and the generation of elementary particle masses was later developed by Steven Weinberg and Abdus Salam. By exploiting the Higgs mechanism and incorporating previous suggestions from Sheldon Glashow for formulating a gauge theory that described the weak force as being mediated by the gauge bosons,

Steven Wienberg and Abdus Salam independently proposed unified theory of electroweak interactions in 1967 and 1968, respectively. The main idea was to introduce to the SM an additional scalar field with a nonzero ground state, referred to as a vacuum expectation value (VEV). The presence of the scalar field's VEV spontaneously breaks the $SU(2)_L \times U(1)_Y$ electroweak symmetry down to the $U(1)_{EM}$ symmetry, generating masses for the W^\pm and Z bosons while leaving the photon massless. The scalar field is required to have at least three degrees of freedom given that three masses are generated, *i.e.* one degree of freedom for each mass generated. The simplest choice for such a field is a complex doublet field belonging to $SU(2)$. Such a field can be expressed in terms of four real valued scalar fields:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^1 + i\phi^2 \\ \phi^3 + i\phi^4 \end{pmatrix}. \quad (2.2)$$

The Lagrangian of the scalar field is given by:

$$\begin{aligned} \mathcal{L}_{\text{Scalar}} &= (D^\mu \Phi)^\dagger (D_\mu \Phi) - \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2 \\ &= |D^\mu \Phi|^2 - V(\Phi), \end{aligned} \quad (2.3)$$

where the first term represents the kinetic interaction of the field and the second term represent the field's potential. The covariant derivative acting on a scalar field Φ with hyper-charge $\frac{1}{2}$ is defined as:

$$D_\mu \Phi = \left(\partial_\mu + \frac{ig_2 \tau_a W_\mu^a}{2} + \frac{ig_1 B_\mu}{2} \right) \Phi, \quad (2.4)$$

where W_μ^a and B_μ are the massless gauge bosons associated to $SU(2)_L$ and $U(1)_Y$, respec-

tively, τ represents the Pauli matrices, and g_1 (g_2) is the EM (weak) interaction coupling constant. The λ term describes quartic self-interactions among the scalar fields. Vacuum stability requires λ to be greater than zero, then for $\mu^2 < 0$ the field develops a VEV which spontaneously breaks the symmetry. This represents the canonical example of the "Mexican hat" potential for a complex scalar field. Due to the symmetry of $V(\Phi)$ there are an infinite number of degenerate minimum energy states satisfying $\Phi^\dagger\Phi = \nu^2/2$, where $\nu = (-\mu^2/\lambda)^{1/2}$. Given that the potential depends solely on $\Phi^\dagger\Phi$ the choice of the direction for the VEV is arbitrary, with a standard choice being: $\phi_1 = \phi_2 = \phi_4 = 0$ with $\phi_3 = \nu$. In this case, the real part of the neutral component of the complex doublet field (ϕ^3) obtains a VEV, and the ground state of the scalar field is nonzero and can be expressed as:

$$\Phi_0 = \begin{pmatrix} 0 \\ \frac{\nu}{\sqrt{2}} \end{pmatrix}. \quad (2.5)$$

.

Since the ground state of Φ is non-zero, it is not possible to perform a perturbative expansion around Φ . Instead, Φ is expanded about its VEV using a shifted field, h , that is small for perturbations around the VEV. In the unitary gauge, Φ can be expressed as:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu + h \end{pmatrix}. \quad (2.6)$$

The kinetic term of scalar field Lagrangian can be expanded in terms of Higgs field, h , and

its ground state energy using Equation 2.6:

$$\begin{aligned} |D^\mu \Phi|^2 &= \left| \left(\partial_\mu + \frac{ig_2 \tau_a W_\mu^a}{2} + \frac{ig_1 B_\mu}{2} \right) \Phi \right|^2 \\ &= \frac{1}{2} (\partial_\mu h)^2 + \frac{1}{8} g_2^2 (\nu + h)^2 |W_\mu^1 + iW_\mu^2|^2 + \frac{1}{8} (\nu + h)^2 |g_2 W_\mu^3 - g_1 B_\mu|^2. \end{aligned} \quad (2.7)$$

The first term is the kinetic term of the Higgs field, terms proportional to ν^2 represent the gauge boson mass terms, and terms describing the interaction between the Higgs field and the gauge fields are proportional to νh and h^2 .

Four new fields can be constructed from the linear superposition of the electroweak fields (W_μ^a and B_μ):

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}} (W_\mu^1 \mp iW_\mu^2) \\ \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} &= \begin{pmatrix} \cos(\theta_W) & -\sin(\theta_W) \\ \sin(\theta_W) & \cos(\theta_W) \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}, \end{aligned} \quad (2.8)$$

where θ_W is the weak mixing angle defined as $\tan(\theta_W) = g_1/g_2$. These four new fields, W^\pm , Z , and A , are the physical fields of the weak interaction gauge bosons and the photon, respectively. Rewriting Equation 2.7 in terms of the physical fields and ignoring terms describing the interaction between the Higgs field and the gauge fields (νh and h^2), gives:

$$|D^\mu \Phi|^2 = \frac{1}{2} (\partial_\mu h)^2 + \frac{(g_2 \nu)^2}{4} W_\mu^+ W^{-\mu} + \frac{(g_2 \nu)^2}{8 \cos^2(\theta_W)} Z_\mu Z^\mu + 0 \cdot A_\mu A^\mu. \quad (2.9)$$

By comparing the above to the required form for the boson mass terms, $M_{W^\pm}^2 W_\mu^+ W^{-\mu}$, $\frac{1}{2} M_Z^2 Z_\mu Z^\mu$, and $\frac{1}{2} M_A^2 A_\mu A^\mu$, the gauge boson masses can be immediately identified as the

coefficients of the bilinear field terms, giving:

$$M_{W^\pm} = \frac{1}{2}g_2\nu, \quad M_Z = \frac{1}{2}\frac{g_2\nu}{\cos(\theta_W)}, \quad M_A = 0. \quad (2.10)$$

Therefore through the Higgs mechanism's introduction of an additional scalar field with a nonzero ground state, the $SU(2)_L \times U(1)_Y$ electroweak symmetry gets spontaneously broken down to $U(1)_{EM}$. A direct consequence of spontaneously breaking the electroweak symmetry is the appearance of the physical W^\pm and Z bosons which acquire their mass by absorbing the three liberated degrees of freedom ($\phi_{1,2,4}$) into their longitudinal polarizations, leaving one remaining real valued scalar field ($\phi_3 = h$), the Higgs field, with the photon (A_μ) remaining massless. Equation 2.9 also elucidates the way in which the weak and EM interactions are unified; the neutral weak gauge boson that acquires its mass through EWSB is a superposition of a gauge boson from $SU(2)_L$ and a gauge boson from $U(1)_Y$.

The SM does not predict exact values for the masses of the W and Z bosons, it does however provide a precise prediction for the ratio of their masses, expressed in terms of the so called ρ parameter. The ρ parameter, defined below, provides a powerful experimental test of the SM, and also for testing the consistency of other hypothetical models with experimental data. The ρ parameter is defined by:

$$\rho = \frac{M_W^2}{M_Z^2 \cos^2(\theta_W)}. \quad (2.11)$$

For the SM $\rho = 1$, this relation has been precisely confirmed using current electroweak measurements of M_W , M_Z , and θ_W [8].

Expanding the scalar field Lagrangian's (Equation 2.3) kinetic term leads to the weak gauge boson masses. To obtain the Higgs boson mass the potential term of Equation 2.3 must be expanded about the VEV in a similar manner:

$$\begin{aligned} V(\Phi) &= \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \\ &= -\frac{1}{2} \lambda \nu^2 (\nu + h)^2 + \frac{1}{4} \lambda (\nu + h)^4 . \end{aligned} \quad (2.12)$$

After expanding the potential term and only considering terms which involve the Higgs field, the Lagrangian for the Higgs field, h , reads:

$$\mathcal{L}_h = \frac{1}{2} (\partial_\mu h)^2 - \lambda \nu^2 h^2 - \lambda \nu h^3 - \frac{1}{4} \lambda h^4. \quad (2.13)$$

By comparing the term that is quadratic in the Higgs field with the expected form of the mass term of a scalar particle, $\frac{1}{2} M_h^2$, the mass of the Higgs boson is obtained:

$$M_h = \sqrt{2\lambda\nu} = \sqrt{-2\mu^2}. \quad (2.14)$$

Fermi's theory of beta decay relates the W boson mass (M_W) to the weak coupling constant (g_2) through Fermi's constant (G_F). Using the Fermi constant and the expression for M_W given in Equation 2.10 an estimation of the value of the vacuum expectation value of the Higgs field, ν , can be obtained: $G_F = \sqrt{2}g_2^2/8M_W^2 = 1/\sqrt{2}\nu^2 \rightarrow \nu \approx 246$ GeV. This is known as the electroweak energy scale, however since λ is a completely free parameter the SM does not provide a prediction for the Higgs mass and therefore it must be determined from experiment.

The discussion up to this point has only addressed the mass generation of the electroweak gauge bosons. Of course the Higgs mechanism is also responsible for fermion mass generation via a three point Yukawa interaction between the fermion field and the ground state of the Higgs field. Once the Higgs field is introduced, gauge invariant interaction terms appear that behave like masses when the Higgs field gets its VEV. The $SU(2)_L \times U(1)_Y$ invariant Yukawa Lagrangian for the fermion fields is:

$$\mathcal{L}_{\text{Yukawa}} = -\lambda_\ell \bar{L}_\ell \Phi \ell_R - \lambda_{q_j} \bar{Q}_q \Phi q_{j,R} - \lambda_{q_i} \bar{Q}_q \tilde{\Phi} q_{i,R} + h.c. \quad (2.15)$$

where L_ℓ and Q_q are the left-chiral lepton and quark doublets, Φ is the same scalar field, $\tilde{\Phi} = i\tau_2 \Phi^*$ is the conjugate representation of the Higgs doublet, λ_ℓ , λ_{q_i} , and λ_{q_j} are the Yukawa couplings for the leptons, up-type quarks (q_i), and down-type quarks (q_j) respectively, and $h.c.$ denotes the hermitian conjugate. The introduction of $\tilde{\Phi}$ is required to generate mass terms for the up-type quarks. Expanding $\mathcal{L}_{\text{Yukawa}}$ about the ground state of the Higgs field results in fermion mass terms.

As an explicit example consider the electron. The relevant term from the Yukawa Lagrangian is:

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}}^{\text{electron}} &= -\frac{1}{\sqrt{2}} \lambda_e (\bar{\nu}_e, \bar{e}_L) \begin{pmatrix} 0 \\ \nu + h \end{pmatrix} e_R + h.c. \\ &= -\frac{1}{\sqrt{2}} \lambda_e (\nu + h) \bar{e}_L e_R + h.c. \end{aligned} \quad (2.16)$$

The electron mass is the constant coefficient of the $\bar{e}_L e_R$ term that includes ν . The masses

of the remaining fermions are obtained in a similar manner, giving:

$$m_\ell = \frac{\lambda_\ell \nu}{\sqrt{2}} \quad \text{and} \quad m_q = \frac{\lambda_q \nu}{\sqrt{2}}. \quad (2.17)$$

Thus, through the Higgs mechanism's introduction of a complex doublet of scalar fields, Φ , the $SU(2)_L \times U(1)_Y$ gauge symmetry is preserved but spontaneously broken down to $U(1)_{EM}$, generating masses for the weak vector bosons (W^\pm and Z) and also for all fermion generations. The $SU(3)$ symmetry remains intact throughout the entire process.

2.3 Higgs boson production at the LHC

In order to produce on-shell Higgs bosons, high energy particle collisions are required. The Large Hadron Collider (LHC) provides high energy proton-proton collisions capable of producing a Higgs boson. Positioned around the LHC are two experiments, ATLAS and CMS, capable of identifying the Higgs boson through the unique detector signatures produced by its decay products.

The proton is a composite object, with the fundamental constituents being quarks and gluons. Each proton is made up of three "valence" quarks (uud) along with a "sea" of gluons and virtual quarks and antiquarks. In the high energy proton-proton collisions at the LHC the principal interactions occur between the proton constituents, and the production of a Higgs boson proceeds through the following dominant production mechanisms, given in decreasing order of their cross section.

- **Gluon Fusion (GF):**

A Higgs boson is produced by the ‘fusion’ of two gluons. Gluons are massless and do not directly couple to the Higgs field, therefore Higgs boson production proceeds primarily via a heavy-quark loop. The quark loop is dominated by top quarks since the Higgs coupling is proportional to the particle’s mass and the top is the most massive quark.

- **Weak Vector Boson Fusion (VBF):**

VBF Higgs production involves the emission of a weak vector boson off a quark in each proton that is involved in the interaction. The weak bosons fuse to produce the Higgs boson. The VBF process is characterized by two jets that are at a low angle with respect to the proton beam-line (‘forward jets’), one on each side of the Higgs boson.

- **W^\pm and Z associated production (WH, ZH):**

A Higgs boson is radiated off a weak vector boson produced from the interaction of a quark and anti-quark. The vector boson goes on to decay either leptonically or hadronically producing characteristic W/Z decay signatures.

- **$t\bar{t}$ associated production ($t\bar{t}H$):**

A Higgs boson is produced via the ‘fusion’ of a top quark from a $t\bar{t}$ pair with an anti-quark from a second $t\bar{t}$ pair. The leftover top and anti-top quarks go on to hadronize, generally producing a characteristic multi-jet signature.

Perturbation theory is used to calculate the predictions for each production process. Leading order (LO) predictions account for the tree-level Feynman diagrams, depicted in

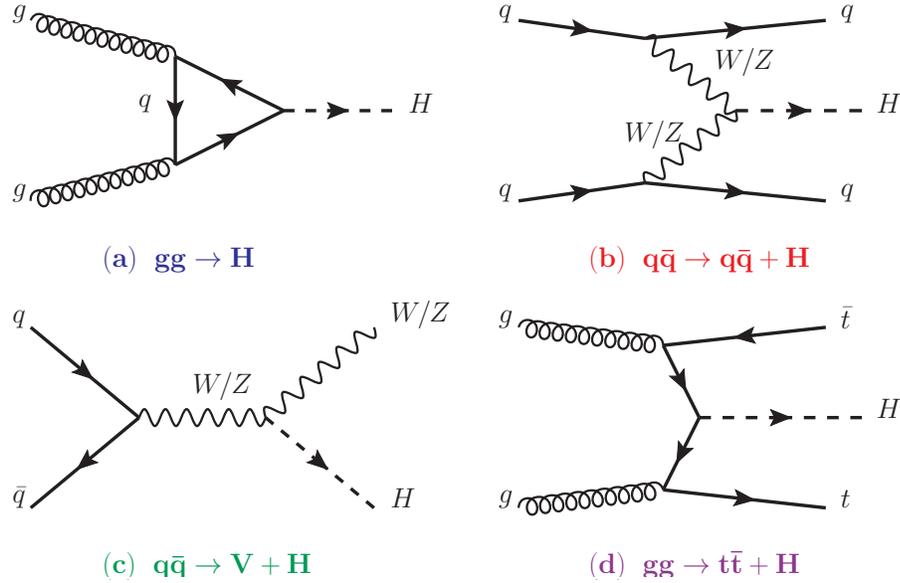
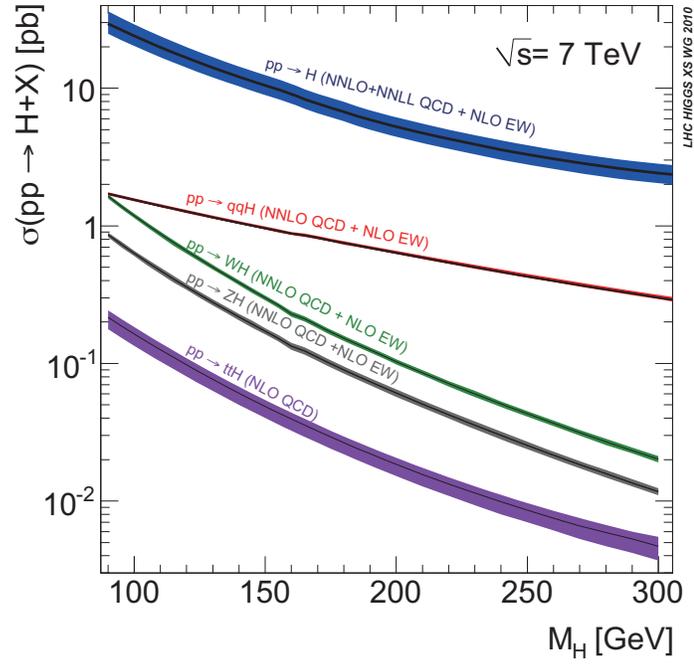
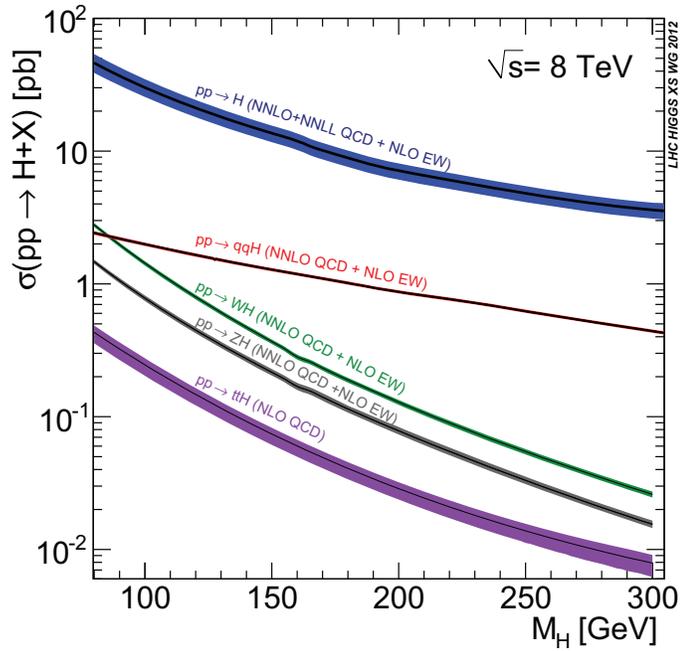


Figure 2.1: Higgs boson production processes at the LHC: (a) $pp \rightarrow H$ via gluon fusion (GF), (b) $pp \rightarrow qqH$ via weak vector boson fusion (VBF), (c) associated production with a W^\pm or Z boson ($pp \rightarrow WH$ or $pp \rightarrow ZH$), and (d) associated production with a $t\bar{t}$ pair ($pp \rightarrow ttH$). Note: time increases from left to right.

Figure 2.1. Next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) predictions take into account QCD corrections for gluon emission and electroweak corrections for W or Z boson emission. In general, preference is given to the highest order theoretical predictions available since they are the most accurate. Theoretical predictions for the Higgs boson cross section for each individual production process are given as a function of the Higgs mass in Figure 2.2(a), and Figure 2.2(b), for proton-proton collisions at centre-of-mass energies of 7 TeV and 8 TeV, respectively. The cross section for GF Higgs production is by far the most dominant, followed by VBF Higgs boson production, with both these cross sections falling off gradually with increasing Higgs boson mass. The cross section for the three types of associated production drop off rapidly with increasing Higgs mass.



(a)



(b)

Figure 2.2: Theoretical predictions for the dominant Higgs boson production processes at the LHC [9], corresponding to a centre-of-mass energies of (a) 7 TeV and (b) 8 TeV. Cross sections are provided for gluon fusion in blue ($pp \rightarrow H$), weak vector boson fusion in red ($pp \rightarrow qqH$), and associated production with a W boson ($pp \rightarrow WH$) in green or a Z boson ($pp \rightarrow ZH$) in grey, and $t\bar{t}$ pair ($pp \rightarrow t\bar{t}H$) in purple.

2.4 Higgs decay modes

The Higgs boson is heavy and therefore unstable, decaying almost immediately into lighter stable particles that interact with the detector leaving unique detector signatures. Within the SM the Higgs mass is a free parameter, meaning the theory does not predict its exact value, but for a fixed Higgs mass the SM has definite predictions for the Higgs boson's coupling to the massive bosons and fermions. Given that the Higgs boson interacts with all massive SM particles there are many different allowable decay modes. Each decay mode has a unique probability, expressed in terms of its branching ratio: the fraction of Higgs bosons decaying via the particular decay mode with respect to the total number of Higgs boson decays. Theoretical predictions for the Higgs boson's branching ratio for each of the possible decay modes as a function of the Higgs mass are shown in Figure 2.3.

The diphoton decay channel ($H \rightarrow \gamma\gamma$) is the focus of the body of this work. The Higgs boson does not couple directly to photons and therefore the decay proceeds via heavy-particle loops, typically top quark or W boson loops, as depicted in Figure 2.4. Despite its rather small branching ratio, $BR = 0.00228$ for a Higgs boson with a mass of 125 GeV, it served as a highly important discovery channel thanks to a very clean decay signature and narrow mass resolution, allowing efficient event selection and sufficient background discrimination. These same features also make the $H \rightarrow \gamma\gamma$ channel ideal for studying the coupling properties of the Higgs boson. Furthermore, the clean signature and the large dataset recorded also provides the ability to measure fiducial and differential cross sections for Higgs boson production, which is the topic of this thesis.

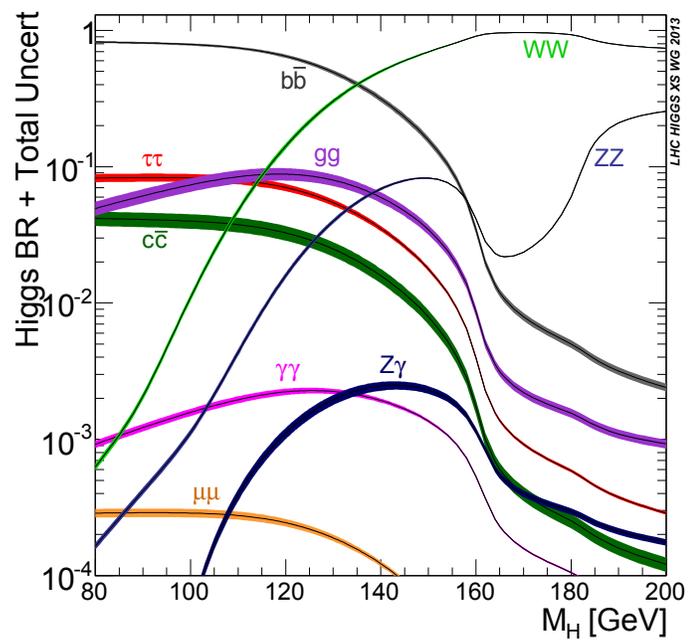


Figure 2.3: Theoretical predictions [9] for the branching ratio of the various Higgs boson decay modes as a function of the Higgs boson mass. The predicted branching ratio for a Higgs boson, with a mass of 125 GeV, decaying to two photons is $(0.228 \pm 0.011) \%$.

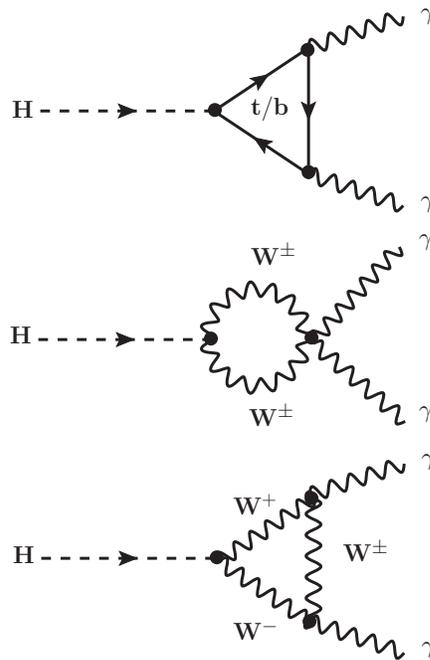


Figure 2.4: Leading-order Feynman diagrams for the Higgs boson decaying to two photons.

2.5 The Higgs boson discovery

In 2012 both ATLAS and CMS announced the discovery of a new particle consistent with the expectation for a SM Higgs boson. By the end of the first data-taking period at the LHC (Run I, 2010-2012), both ATLAS and CMS reported a discovery of a SM like Higgs boson decaying to photon, Z and W pairs [1, 2]. Observational evidence was reported for the decay to tau-lepton pairs, providing direct evidence for the Higgs coupling to fermions [10, 11]. The mass of the Higgs boson has been measured to be 125.09 GeV with a precision of 0.2 % by combining the data from the ATLAS and CMS experiments [12]. The spin of the Higgs boson has been found to be consistent with the SM prediction of a spin-0 particle, with spin-2 models ruled out at the 95 % confidence level [13]. The strengths of the Higgs boson's interaction with SM particles have been studied using global fits to all

decay channels with no significant deviation from the SM observed [14].

2.6 Summary

The process of spontaneous electroweak symmetry breaking via the Higgs mechanism forms an integral part of the Standard Model of particle physics, providing an explanation of how elementary particles get their mass. Through the introduction of a massive scalar field to the Standard Model, the Higgs mechanism predicts the existence of a massive scalar particle, the Higgs boson. However, the mass of the Higgs boson is left as a free parameter of the theory and must be determined by experiment. The search for the Higgs boson comprised the principal motivations for the physics programme of the LHC, with the subsequent discovery of a SM like Higgs boson representing a huge achievement in the field of particle physics. Following the discovery the focus has shifted to precisely measuring the particle's properties and testing the consistency of the data with theoretical predictions in order to determine whether it is truly the Higgs boson of the SM or a different particle altogether. This work focuses on performing measurements of the Higgs boson cross section in the $H \rightarrow \gamma\gamma$ decay channel; these measurements provide a powerful test of the newly discovered particle's consistency with theoretical predictions, helping to further elucidate its exact role in the SM.

Chapter 3

The Large Hadron Collider and the ATLAS Detector

In order to produce on-shell Higgs bosons, particle collisions at extremely high energies are required. The Large Hadron Collider (LHC) is the only machine in the world capable of providing the high energy particle collisions needed to produce Higgs bosons with a high enough rate to allow detailed investigations. The ATLAS detector is one of two general-purpose detectors positioned around the LHC capable of studying the Higgs boson through precise measurement of the final state particles produced by its decay. This chapter introduces and describes the LHC [15] and the ATLAS (A Toroidal LHC Apparatus) detector [16]. An introduction to the simulation of ATLAS data follows the descriptions of the LHC and ATLAS detector.

3.1 The Large Hadron Collider

The LHC is a circular particle accelerator and collider located at the European Organization for Nuclear Research (CERN), straddling the Swiss and French border near Geneva. The LHC is the world's largest and most powerful particle collider, and is able to accelerate protons up to approximately the speed of light, providing proton-proton collisions at the energy frontier. The accelerator ring has a 26.7 km circumference and is placed approximately 100 m below ground. The LHC can be essentially divided into eight segments, with each segment consisting of a 528 m straight section with two bending sections at either end. The eight straight sections serve as insertion points for experiments or beam utilities. The LHC houses four experiments, ATLAS [16], CMS (Compact Muon Solenoid) [17], LHCb [18], and ALICE (A Large Ion Collider Experiment) [19], leaving four remaining insertion points. Two of the remaining insertion points are used for housing beam collimators, one is used for the beam dump, and one contains the radio-frequency (RF) accelerator cavities. The LHC is designed to collide two proton beams at a maximum centre-of-mass energy of $\sqrt{s} = 14$ TeV, with each beam composed of 2808 bunches of protons with 10^{11} protons per bunch. Operating at its full design specification the bunches will cross every 25 ns, however, for the collisions recorded at 7 and 8 TeV the shortest bunch spacing was 50 ns.

The LHC ring is the last stage of a multi-stage acceleration chain, where the proton beams are first accelerated to an energy of 450 GeV before being injected into the LHC. The acceleration chain starts with a plasmatron that is used to extract protons from hydrogen molecules and inject the protons into a linear accelerator, the Linac2. The Linac2 focusses

the proton beam and accelerates it to an energy of 50 MeV, then feeds the beam into the Proton Synchrotron Booster (PSB). The PSB accelerates the protons up to 1.4 GeV and then passes the beam to the Proton Synchrotron (PS) where the protons are accelerated to 25 GeV. The PS then feeds the proton beam into the Super Proton Synchrotron (SPS). The SPS further accelerates the protons up to an energy of 450 GeV before injection into the LHC. From the SPS two proton beams are injected into separate LHC beam pipes, circulating in opposite directions, and brought up to their final collision energy. A system of 1232 superconducting dipole magnets is used to steer the high energy proton beams around the LHC ring. Additionally, a system of quadrupole, sextupole, and octopole magnets are used to focus the beams. The final stage of the acceleration chain is provided by the RF cavities, which boost the beam energies up to the 7 TeV design energy. The total time required to obtain 7 TeV beams is approximately 25 minutes. The full accelerator complex is illustrated in Figure 3.1.

The LHC began running in September 2008, but suddenly stopped after just nine days of operation when a faulty connection between two superconducting magnets caused a sudden quench that ended up damaging several magnets [21]. The machine restarted in 2009 and operated throughout 2010 and 2011, however it was decided to operate at a reduced centre-of-mass energy of $\sqrt{s} = 7$ TeV until further machine maintenance could be done. In 2012 the centre-of-mass energy was increased to $\sqrt{s} = 8$ TeV, with the majority of the recorded ATLAS dataset obtained at this energy. In 2013 the LHC was shut down to perform the required upgrades in order to allow operation at the designed centre-of-mass energy of $\sqrt{s} = 14$ TeV. In June 2015 the LHC restarted, delivering proton-proton collisions at centre-of-mass energy of $\sqrt{s} = 13$ TeV. The combination of all LHC runs that occurred during 2010,

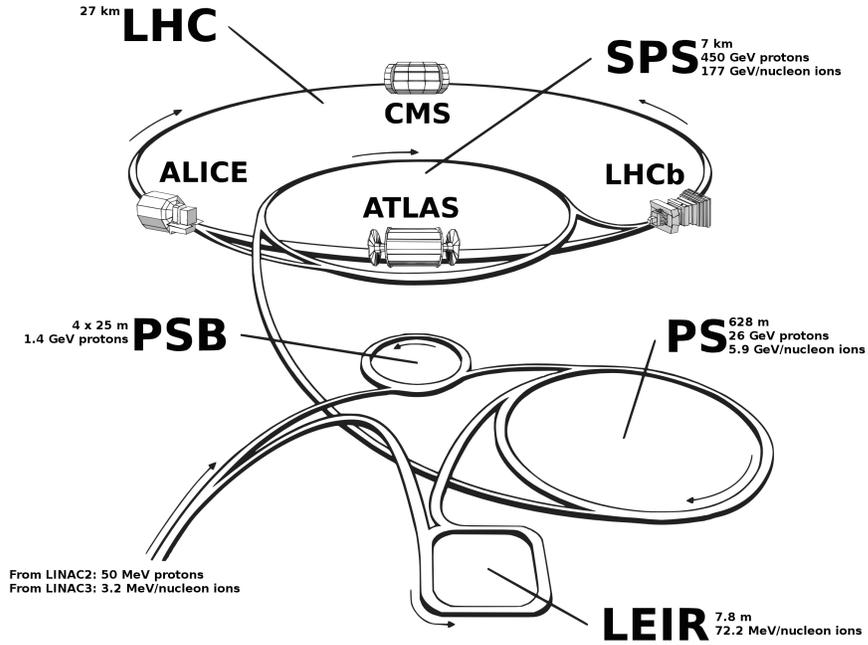


Figure 3.1: Illustration of the LHC accelerator complex (not to scale) [20].

2011 and 2012 is collectively referred to as Run-I, while Run-II refers to LHC data that is recorded after the 2013 long shutdown.

After the collision energy, the most important parameter is the luminosity \mathcal{L} . Instantaneous luminosity is a direct measure of the rate of collisions (events) delivered by the accelerator. Higher instantaneous luminosities result in larger event production rates. The instantaneous luminosity is the ratio of the total number of events produced (per second) to the total inelastic cross section, and has units of $\text{cm}^{-2}\text{s}^{-1}$. The total instantaneous luminosity can be expressed as a function of the beam and collision parameters:

$$\mathcal{L} = \frac{N_b N_1 N_2 f_{\text{rev}} \gamma}{4\pi \beta^* \epsilon_n} S, \quad (3.1)$$

where:

- N_b : the number of colliding bunches per beam.
- $N_{1,2}$: the number of protons per bunch in beam one and beam two.
- f_{rev} : the frequency of revolution.
- γ : Lorentz correction factor (E/m_p).
- β^* : "squeeze" parameter accounting for the degree of beam focus at the interaction point, in units of meters.
- ϵ_n : beam emittance in units of meters; viewed as the normalized momentum-position phase-space occupied by the beam.
- S : geometrical correction factor related to the beam crossing angle.

Complementary to the instantaneous luminosity is the integrated luminosity, which measures the total number of collisions (events) delivered over time. Integrated luminosity is generally reported in units of inverse barns, where a barn is defined as 10^{-28} m^2 (100 fm^2), which is approximately the cross sectional area of a uranium nucleus. Figure 3.2 shows the total integrated luminosity collected during the 2011 and 2012 LHC operation. The total integrated luminosity recorded by ATLAS is approximately 5 fb^{-1} in 2011 and 20 fb^{-1} in 2012. These large integrated luminosities correspond to large datasets, providing the required statistics in order to study rare processes such as the production of a Higgs boson and measure its properties. Figure 3.2 also illustrates the very high instantaneous luminosity in 2012 compared to 2011, and with this increased instantaneous luminosity

comes additional challenges. The higher rate results in an increased number of proton-proton collisions within the same bunch crossing. The average number of proton-proton collisions per bunch crossing in 2011 was 9.1, increasing up to 20.7 in 2012. These many simultaneous proton-proton interactions are referred to as pileup. Pileup complicates the event structure and requires sophisticated techniques to identify the interaction of interest and reject and suppress the effects of the additional (uninteresting) interactions. However, the enlarged dataset is well worth the additional complexity due to the pileup associated with higher luminosity.

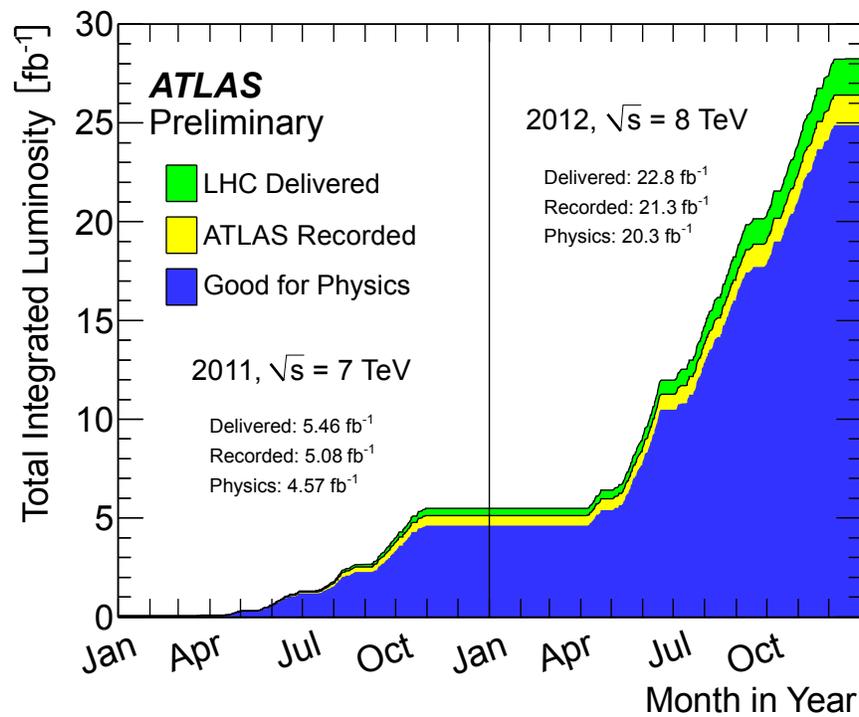


Figure 3.2: Integrated luminosity as a function of time delivered to (green), recorded by ATLAS (yellow), and verified to be good quality data (blue) during stable beam conditions and for proton-proton collisions at centre-of-mass energies of 7 TeV and 8 TeV in 2011 and 2012, respectively [22].

3.2 The ATLAS detector

ATLAS is a general purpose particle detector designed to cover the largest possible range of physics opportunities offered by the high energy proton-proton collisions provided by the LHC. The ATLAS detector is depicted in Figure 3.3, it is 25 m high, 44 m long, weighs approximately 7,000 tonnes, and consists of multiple layers of sub-detectors designed to provide an essentially full 4π coverage around the primary interaction point. Each sub-detector has a dedicated design in order to measure the energies and momenta of specific particle types. The design specifications and performance requirements for the ATLAS detector were largely motivated by Higgs boson and Supersymmetry searches, but always with a full and rich physics program in mind. For a 125 GeV Higgs boson the principal decay channels are; $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^* \rightarrow 4l$, $H \rightarrow WW$, $H \rightarrow \tau\tau$, and $H \rightarrow bb$, with many of these decays also involving hadronic jets in the final state. Supersymmetry searches can involve many final state jets, leptons and often missing transverse energy. Precision Standard Model measurements generally require excellent spatial and energy resolution. In order to accurately identify the different decay signatures and individual decay products, precisely measure the properties of the various particles, and disentangle the complex final states produced by the high energy and luminosity collisions at the LHC, the ATLAS detector was designed to meet the following fundamental requirements:

- Fast readout electronics.
- Radiation resistant detector components.
- Efficient charged particle reconstruction and identification combined with good mo-

momentum resolution in the tracking detector.

- Good primary and secondary vertex reconstruction, both needed for tau-lepton identification, hadronic jet b-tagging, and pileup rejection/suppression.
- Excellent electromagnetic calorimetry for efficient electron and photon identification and precision measurements.
- Full 4π hadronic calorimeter coverage for jet and missing transverse energy reconstruction.
- Efficient muon identification combined with good momentum resolution.
- A sophisticated data acquisition ('triggering') system.

3.2.1 The ATLAS coordinate system

ATLAS uses a right-handed coordinate system, with the x-axis pointing to the centre of the LHC ring, the y-axis pointing upwards, the z-axis following the beam direction, and the origin location coinciding with the nominal interaction point at the centre of the detector (0, 0, 0). The azimuthal angle $\Phi = 0$ aligns with the positive x-axis and Φ increases clockwise with respect to the positive z direction. The x-y plane defines the transverse momentum plane for particles emerging from the interaction point. In place of the polar angle θ , two analogous variables are defined; for massive particles one defines the rapidity, $y = \frac{1}{2} \ln [(E - p_z) / (E + p_z)]$, while for massless particles an equivalent variable is defined, the pseudorapidity (η). Pseudorapidity is directly related to the polar angle with

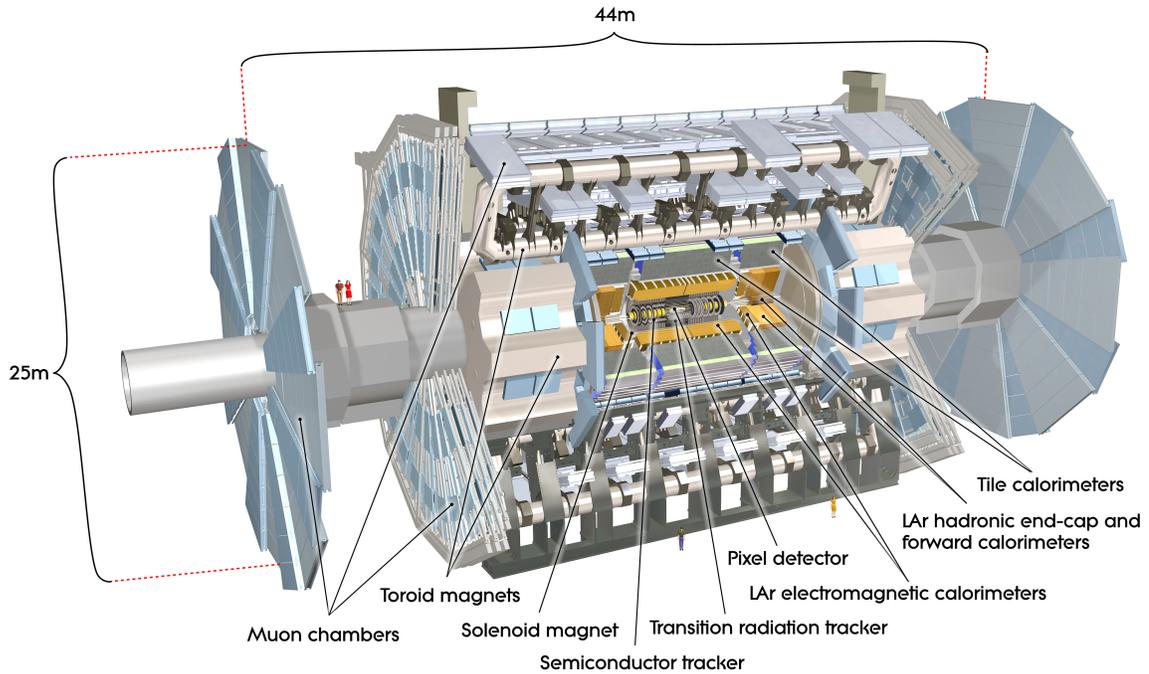


Figure 3.3: Schematic view of the ATLAS detector [16].

respect to the beam line by: $\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right]$. Generally pseudorapidity is preferred when describing the detector as it is well defined in the detector frame of reference and it is independent of a particles mass.

From the definition of pseudorapidity we have that $\eta = 0$ along the positive y -direction and $\eta = \infty$ along the beam (z) direction. The ATLAS detector provides coverage over the region $|\eta| < 4.9$. Furthermore, angular separation between two objects is defined by: $\Delta R = \sqrt{\Delta\eta^2 + \Delta\Phi^2}$. A cartoon illustrating the ATLAS coordinate system is given in Figure 3.4.

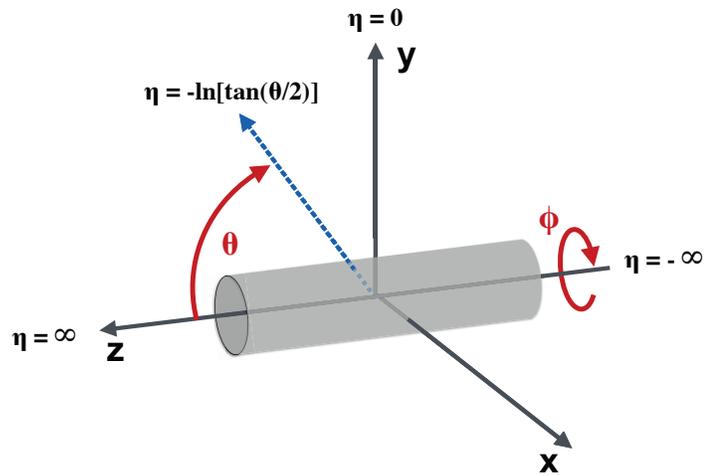


Figure 3.4: Cartoon illustrating the ATLAS coordinate system.

3.2.2 The inner detector

The ATLAS inner detector (ID) [23] is designed to accurately reconstruct the positions, momenta and origins of charged particles within the region defined by $|\eta| < 2.5$. A large solenoid magnet surrounds the ID, providing a uniform magnetic field of 2 T which enables charge identification and transverse momenta measurements of charged particles based on the curvature of the ‘tracks’ they leave in the ID. The ID is composed of three sub-detector systems which utilize two types of detector technology: the Pixel Detector [23] and the Semiconductor Tracker (SCT) [23] are both silicon pixel based detectors, while the Transition Radiation Tracker (TRT) [23] is a straw tube based detector that utilizes transition radiation for particle identification. An overview of the ID is shown in Figure 3.5.

The Pixel Detector

The Pixel Detector is the closest detector to the interaction point, consisting of three layers

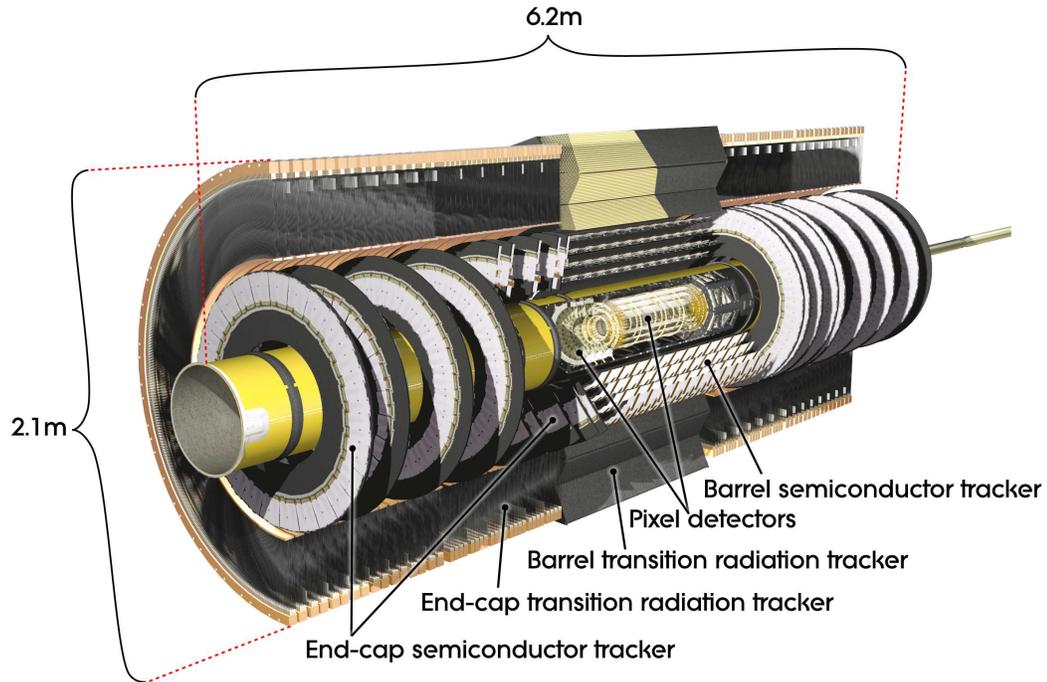


Figure 3.5: A cut-away view of the ATLAS inner detector [23].

of silicon pixels that are positioned concentric to the beam axis at distances of 50.5 mm, 88.5 mm, and 122.5 mm, and three end-cap wheels that are positioned perpendicular to the z-axis, located between 50 cm and 65 cm from $z=0$. The pixel detector contributes 80.4 million readout channels and provides precision reconstruction of particle tracks and of primary and secondary vertices. The first barrel layer is referred to as the vertexing-layer or alternatively the B-Layer and is crucial for identifying hadronic jets originating from b-quark decays. It also plays an important role in distinguishing electrons from photons that undergo conversion into electron-positron pairs.

The Semiconductor Tracker (SCT)

The SCT is located outside of the pixel detector and consists of silicon micro-strip detectors that provide additional precision track measurements. It contributes an additional 6.3

million readout channels. The SCT consists of four layers of silicon micro-strip detectors concentric to the beam axis at distances of 30 cm, 37 cm, 44 cm, and 51 cm, along with nine end-cap wheels located between 85 cm and 272 cm from $z=0$. The SCT plays an invaluable role in providing high-precision measurements of vertex positioning, track momentum and impact parameters. The impact parameter is the perpendicular distance between the interaction point and the trajectory of a track.

The Transition Radiation Tracker (TRT)

The TRT is positioned outside of the SCT and consists of a straw tracking and transition radiation detectors. The TRT has full azimuthal coverage and extends to $|\eta| < 2.0$, providing trajectory measurements of charged particles, with an average of 35 ‘hits’ per track. The SCT is constructed from 350,000, 4 mm diameter straw drift tubes made of Kapton, each tube has a conductive coating and is filled with a 70% Xe , 27% CO_2 , and 10% O_2 gas mixture. A gold-tungsten wire runs through the centre of each straw drift tube and serves as the anode, while the tube itself serves as the cathode. Straw drift tubes surrounding the barrel are aligned parallel to the beam pipe, while straw drift tubes in the end-caps are positioned radially outward from the beam pipe in a fan shape. Between the drift tubes are radiator foils that produce transition radiation (TR) with the TR thresholds set to help identify different types of particles. For instance, the energy produced from TR for electrons is generally larger than the energy produced by hadrons, thereby TR provides efficient discrimination between electrons and hadrons.

3.2.3 The calorimeter system

Particles that pass through matter lose the majority of their energy through the production of electromagnetic or hadronic ‘showers’. A shower is a secondary cascade of particles produced by the passing particle’s interactions with the material of the detector. Through these interactions additional lower energy particles are produced which then also interact with the material and produce further new particles, and so on. This process terminates when the initiating particle and all the shower particles have an energy below the threshold required to produce new particles. The particles are then absorbed into the material.

Electromagnetic and hadronic showers have very different shower (energy deposition) profiles. The energy deposited in the form of electromagnetic and hadronic showers is measured using calorimetry, with the calorimeters also being able to measure many properties of the shower profile which aids particle identification. The material of a calorimeter is characterized in terms of its radiation length X_0 . A radiation length for photons is defined as the distance required to reach $7/9$ of the mean free path for the photon to convert into an electron-positron pair, while for electrons it is defined as the required distance in a material to lose $\sim 37\%$ ($1/e$) of its initial energy through Bremsstrahlung.

Information from the calorimeters provides the principal source of identification for photons, electrons and all hadronically decaying particles. By including information from the muon spectrometer, neutrinos and potentially other types of exotic particles can be inferred through the missing transverse energy in an event. The ATLAS calorimeter system completely surrounds the inner detector, utilizing two different types of calorimeters: (i) an electromagnetic calorimeter system which provides accurate identification of photons

and electrons, and (ii) a hadronic calorimeter system for measuring the energy deposited by hadronically decaying particles. Each calorimeter type consists of a number of subsystems, with each subsystem having full azimuthal coverage. The total calorimeter system provides coverage up to $|\eta| < 4.9$. A cut-away view of the full ATLAS calorimeter system is given in Figure 3.6.

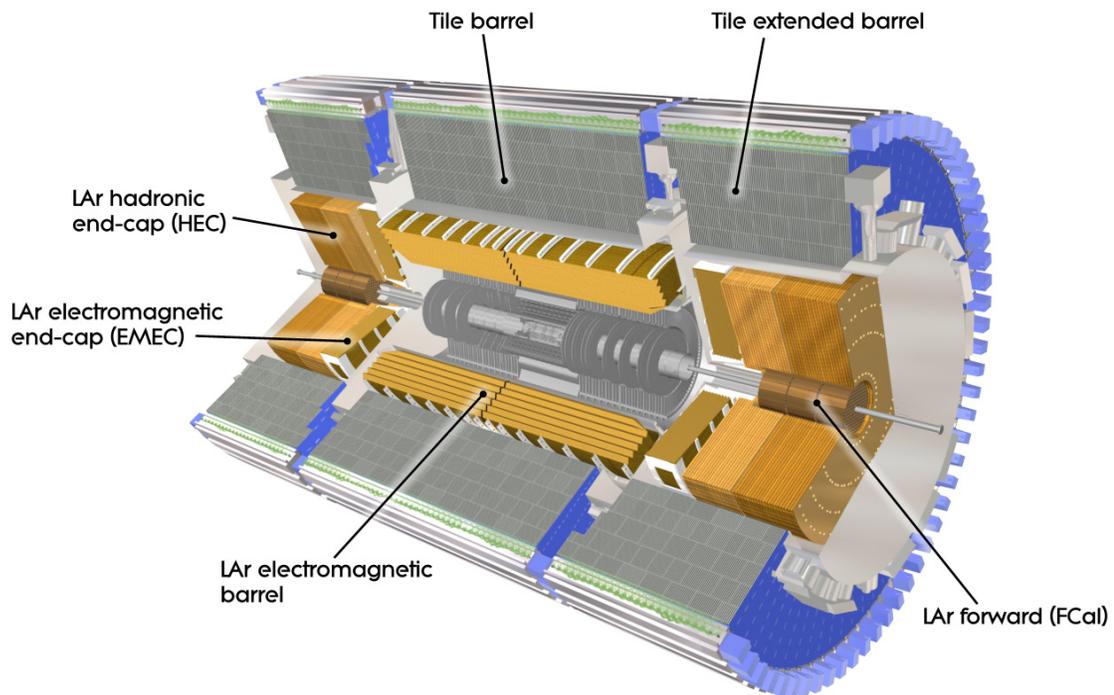


Figure 3.6: A cut-away view of the ATLAS calorimeter system [16].

The electromagnetic calorimeter

The electromagnetic (EM) calorimeter [16] is a sampling instrument arranged in an accordion geometry, composed of alternating layers of active liquid argon (LAr) and inactive lead. The accordion geometry provides full azimuthal coverage without any dead regions.

In order to optimize the material usage, a common vacuum chamber is shared between the LAr calorimeter and the solenoid magnet system; this reduces the required number of vacuum walls by two. The EM calorimeter is separated into three sections, a central barrel section ($|\eta| < 1.475$) and two end-cap sections ($1.375 < |\eta| < 3.2$), each with a dedicated cryostat. A LAr pre-sampler is installed in front of the first calorimeter layer and before the material used for detector services. The pre-sampler is used to sample the energy of EM showers and estimate energy losses for photons and electrons due to material upstream of the calorimeter. The pre-sampler extends over the range $|\eta| < 1.8$ and has fine η - Φ granularity.

The central barrel section is composed of two identical half barrels which are separated by a 4 mm gap at $\eta = 0.0$. The total thickness of the central barrel calorimeters is required to be at least $22 X_0$, ensuring the full containment of EM showers. There are three EM calorimeter layers, each segmented in Φ ($\Delta\Phi=0.025$). The first layer has a depth of $4.3 X_0$, provides pseudorapidity coverage out to $|\eta| < 3.2$, and is very finely segmented in η ($\Delta\eta=0.0031$) in order to provide efficient π^0 identification. The second layer is the thickest, with a depth of $16 X_0$, and measures the majority of the energy deposited by EM showers. It also provides pseudorapidity coverage out to $|\eta| < 3.2$ but with a reduced η -segmentation ($\Delta\eta=0.025$). The third layer supplies the last energy measurement before the hadronic calorimeter, capturing the energy tails of EM showers and provides discrimination against hadronic showers. It has a depth of $2X_0$ and provides pseudorapidity coverage out to $|\eta| < 2.5$ with a further reduced η -segmentation ($\Delta\eta=0.05$). The total energy resolution of the EM barrel calorimeter has been measured to be $\sigma_E/E = 10\% \cdot \sqrt{\text{GeV}}/\sqrt{E} \oplus 0.17\%$, where E is in GeV and σ_E is the root-mean-square (RMS) spread of the energy measurements.

The two end-cap sections are composed of two coaxial segments ('wheels') on either side of the central barrel, with the inner (outer) wheel providing pseudorapidity coverage over the range $1.375 < |\eta| < 2.5$ ($2.5 < |\eta| < 3.2$). The total active depth of each end-cap is greater than $24 X_0$ for $|\eta| > 1.475$. Similar to the EM barrel calorimeter, the end-cap calorimeters are divided into three longitudinal layers for the region covering $1.5 < |\eta| < 2.5$. The first layer has a depth of $4.4 X_0$ and is segmented along the η direction. The transverse segmentation of the second layer is the same as in the second layer of the barrel calorimeter, $\Delta\eta \times \Delta\Phi = 0.025 \times 0.025$. The η -segmentation of the third layer is the $\Delta\eta=0.05$. The outermost of the end-caps ($2.5 < |\eta| < 3.2$) are sub-divided into two longitudinal layers, each with a coarser transverse segmentation.

The hadronic calorimeter

The hadronic calorimeter [16] measures the energy deposited by showers initiated by hadrons (*i.e.* hadronic showers) and is composed of several calorimeter subsystems: the tile calorimeter, the LAr hadronic end-cap calorimeter, and the LAr forward calorimeter.

The tile calorimeter is composed of a low-carbon steel absorber medium and a scintillating tile sampling medium. It is positioned directly outside the EM calorimeter, with its central barrel covering the pseudorapidity region $|\eta| < 1.0$ and its two extended barrels providing coverage over the pseudorapidity range $0.8 < |\eta| < 1.7$. The tile calorimeter has a measured energy resolution of $\sigma_E/E = 56\% \cdot \sqrt{\text{GeV}}/\sqrt{E} \oplus 6\%$, where E is in GeV. Two LAr end-cap calorimeters are positioned directly behind the EM calorimeter end-caps, sharing the same cryostat, and provide coverage over the pseudorapidity range $1.5 < |\eta| < 3.2$. The forward LAr calorimeter (FCal) further extends this coverage over

the region $3.1 < |\eta| < 4.9$. The FCal is composed of three modules, the first module is made of copper and primarily provides electromagnetic measurements, while the following two modules made of tungsten measure hadronic energy deposits. The hadronic energy resolution of the FCal has been measured to be $\sigma_E/E = 70\% \cdot \sqrt{GeV}/\sqrt{E} \oplus 3\%$, where E is in GeV.

3.2.4 The muon spectrometer

The muon spectrometer (MS) [16] is the outermost layer of the ATLAS detector, arranged with three concentric layers around the barrel and two wheels at each end-cap. The MS is depicted in Figure 3.7. The MS is composed of two tracking and two trigger systems, spanning the region $|\eta| < 2.7$, and is immersed in 0.5-1.5 T magnetic field generated by the 1300-ton superconducting air-core toroidal magnet after which ATLAS was named. The toroidal field bends muon trajectories providing muon transverse momentum measurements within a few percent and also providing charge identification. Precision tracking is achieved using Monitored Drift Tubes (MDTs) [16] and the Cathode Strip Chambers (CSCs) [16]. The MDTs provide measurements in the region $|\eta| < 2.7$. The CSCs are used in order to obtain present-level precision in the region $2.0 < |\eta| < 2.7$, where particle flux is too high for the drift tubes. The muon triggering system has very different requirements from the tracking system and therefore utilizes different types of sub-detectors: Resistive Plate Chambers (RPCs) [16] and Thin Gap Chambers (TGCs) [16]. The RPCs cover the region $|\eta| < 1.05$, while the TGCs cover the region $1.05 < |\eta| < 2.4$.

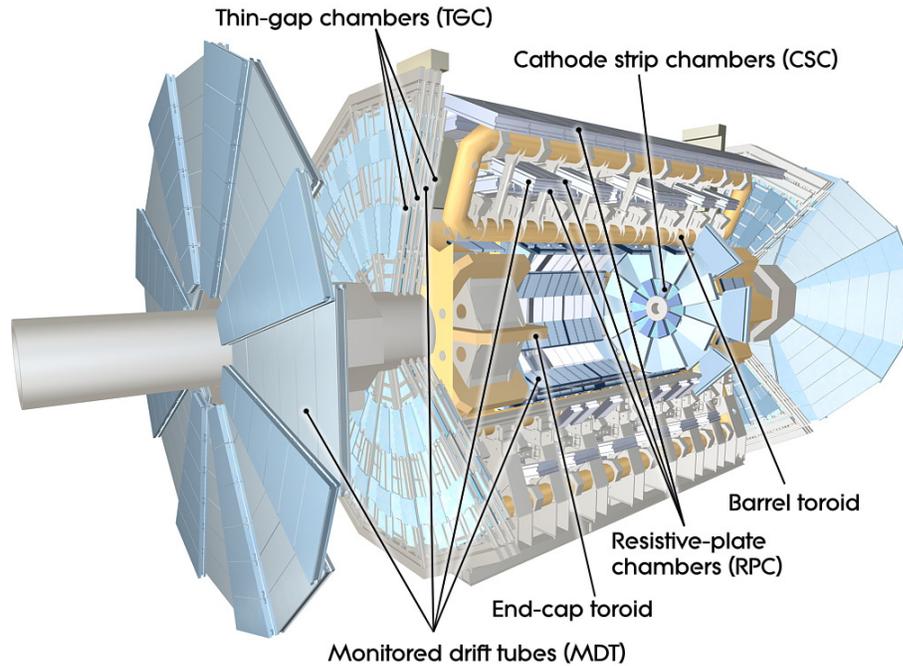


Figure 3.7: A cut-away view of the ATLAS muon spectrometer [16].

3.3 The trigger system

During 2011 and 2012, the LHC operation resulted in a total event rate of 15 MHz. ATLAS implements a sophisticated triggering system [24] in order to selectively record only interesting physics events containing hard scattering (large polar angle) interactions for data analysis, while rejecting less interesting scattering events. Elastic scattering of the protons and diffractive events will not be seen by the detectors as it is only the inelastic scatterings that give rise to particles at sufficiently high angles with respect to the beam axis.

A three-level trigger system is utilized at both the hardware and software levels to perform the rate reduction. The first level, termed Level 1 (L1), is a hardware-based trigger

implementation that uses a sliding window algorithm with coarse granularity calorimeter towers, combined with the RPCs and TGCs of the muon system. The L1 trigger decision is made within $\sim 2.5 \mu\text{s}$ of the collision and is used to trigger the detector readout. The L1 trigger brings the total rate down to $\sim 75 \text{ kHz}$, a reduction by a factor of ~ 200 compared to the total event rate.

The Level 2 (L2) trigger is a software-based implementation using "Regions of Interest" (ROIs) to reconstruct physics quantities (tracks and particles) used in event filtering decisions. The rate output after L2 is around 3.5 kHz , representing a further rate reduction by a factor of ~ 20 .

The final level in the trigger system is the event filter (EF) trigger, which implements offline analysis algorithms. It fully reconstructs the physics objects in order to select the most interesting events and separate them into various physics analysis streams, depending on which physics objects are reconstructed. The EF trigger provides the final factor of 10 reduction, bringing the event rate down to 400 Hz . The combination of the EF and L2 trigger is collectively referred to as the high-level trigger (HLT). After the application of the full trigger chain the event rate is reduced considerably, resulting in more reasonable data rate of 400 Hz , all of which are written to disk.

3.4 Monte Carlo simulation of ATLAS data

Three classifications of theoretical prediction will be used throughout the following work: particle-level, parton-level, and reconstruction-level. Parton-level refers to theoretical predictions calculated using bare quarks and gluons, *i.e.* before fragmentation and

hadronization. Particle-level refers to predictions obtained from stable particles after the fragmentation and hadronization of the bare quarks and gluons. Stable particles are those particles that interact directly with the detector, defined as having a lifetime longer than 10 ps. Reconstruction-level (detector-level) refers to predictions that are built from fully digitized outputs of the detector simulation obtained by propagating all stable particles through a Geant4 [25, 26] simulation of the ATLAS detector [27, 28]. The Geant4 simulation models the interactions of the particles with the detector material, with all outputs fully digitized in order to obtain reconstruction-level predictions that accurately represent the signals that would be experimentally measured. Details on the simulated event samples used in this work are provided in Section 5.1.1, Section 10.1, and Appendix A.

The detector signals from the simulated events are overlaid with additional simulated detector signals from pileup events (additional inelastic proton-proton interactions produced in LHC beam crossings) to reflect the experimental conditions of the 2011 and 2012 data taking periods. The effects of both in-time pileup (pp interactions occurring in the same bunch crossing), and out-of-time pileup (pp interactions occurring in neighbouring bunch crossings) are included in the simulation. The number of pileup events in each bunch crossing is randomly sampled from a Poisson distribution with mean given by $\langle\mu\rangle$, which is directly proportional to the instantaneous luminosity. A weighting function is applied to the simulation so that the $\langle\mu\rangle$ profile in the simulation matches that observed in the data.

Chapter 4

Event reconstruction in ATLAS

This chapter describes the basic event and particle object reconstruction, or simply reconstruction, used in ATLAS. Reconstruction refers to building high-level physics objects from primary detector and electronics outputs. The final output of reconstruction is a collection of measurements representing the general event structure and the properties of all the particles produced in the collision. Reconstruction at ATLAS is essentially accomplished using a two stage procedure: first ‘tracks’ and ‘clusters’ are built from low-level detector signals, then higher level physics objects are built from the information contained in the tracks and clusters.

As a charged particle travels through the ID, or MS, it deposits energy (referred to as ‘hits’) in the various detector layers along its trajectory. Track reconstruction associates collections of hits to the trajectory of individual particles passing through the ID and/or MS. Tracks are reconstructed by performing a three-dimensional fit to the hits left in the ID and/or MS. The fitting algorithms utilize various procedures to account for the presence of

magnetic fields and detector material effects such as multiple scatterings [29]. The ATLAS pattern recognition reconstructs tracks down to a few hundred MeV in transverse momentum ($p_T = \sqrt{p_x^2 + p_y^2}$).

Proton-proton interaction vertices are reconstructed from the collection of tracks in a given event, after implementing a minimal set of track quality criteria to ensure only ‘good’ tracks are considered. An iterative algorithm determines whether multiple tracks point to the same interaction point in the detector, reconstructing a collection of primary vertices (PVs), each with a set of associated tracks. The primary interaction vertex (PV₀) is defined as the one with the highest $\sum (p_T^{\text{track}})^2$, where the scalar sum is taken over the set of tracks associated to the vertex. In the next vertex finding iteration, all tracks not associated to the previously found PV₀ are used as input, and the second primary vertex (PV₁) is identified. The iterative process continues until all tracks are associated to a vertex, resulting in the full set of primary vertices (PV₀, ..., PV_n).

Particles such as photons and electrons typically produce narrow showers that are contained in the EM calorimeters. Hadrons, such as kaons and charged pions, pass through the EM calorimeters and penetrate deep into the hadronic calorimeters before coming to rest, typically producing much broader showers. Significant amounts of energy can also be deposited in the EM calorimeters from hadronic showers due to neutral pions that decay to two photons before reaching the calorimeters.

Energy deposited in the calorimeters by all types of particle is generally spread over many neighbouring cells. An iterative combination algorithm combines individual cell energies into three-dimensional topological clusters [16], referred to as topo-clusters, with the grouping algorithm defining the topology. The algorithm uses a 4-2-0 clustering scheme,

where the cluster seed is the cell containing more energy than 4 times the energy expected from noise (electronic and pileup). Next, all neighbouring cells with an energy greater than 2 times the expected noise contribution are added to the seed, forming a proto-cluster. Finally, all positive energy cells directly neighbouring the proto-cluster are included to form the final topo-cluster.

A sliding window clustering algorithm is also used, primarily for photons and electrons [16]. In the sliding window algorithm, a small rectangular window is scanned across the EM calorimeter in (η, Φ) -space to find significant energy deposit used to seed clusters within the window. The window position is chosen such that the contained transverse energy is a local maximum exceeding 2.5 GeV. This technique works well for photons and electrons since they generally produce narrow showers that are fully contained in the EM calorimeters.

4.1 Jet reconstruction

In a hadron collider such as the LHC, free quarks or gluons are created. Due to the nature of the strong interaction isolated quarks and gluons cannot exist individually. Instead, these particles undergo fragmentation, forming hadrons by combining with quarks and anti-quarks that are spontaneously ‘pulled’ from the vacuum. The number of hadrons formed from fragmentation in a high energy proton-proton collision can be quite large. Instead of trying to reconstruct each hadron individually, collimated sprays of hadrons are clustered together (referred to as hadronic jets). Jets therefore do not represent individual elementary particles, rather they are used as ‘tools’ to help interpret complex hadronic final states and

are among the main tools used for analyzing ATLAS data. Due to the author's high degree of involvement in their study, additional details on jet reconstruction and jet calibration (see Chapter 5) will be included.

In ATLAS jets are reconstructed using sequential recombination algorithms, with the standard being the anti- k_t algorithm [30]. The anti- k_t algorithm clusters objects (particles, calorimeter topo-clusters) together by introducing two distance measures; the distance between two objects i and j , d_{ij} , and the distance between an object i and the beam B , d_{iB} . The distance measures are defined as:

$$d_{ij} = \min(k_{ti}^{-2}, k_{tj}^{-2}) \frac{\Delta r_{ij}}{R} \quad (4.1a)$$

$$d_{iB} = k_{ti}^{-2}, \quad (4.1b)$$

where k_{ti} is the transverse momentum of object i , $\Delta r_{ij} = (y_i - y_j)^2 + (\Phi_i - \Phi_j)^2$ is the distance between objects i and j with y_i and Φ_i the rapidity and azimuthal angle of object i , respectively, and R is the jet clustering radius parameter. As can be inferred from the above equations the anti- k_t algorithm clusters objects into jets starting from the highest transverse momentum constituent. The FASTJET implementation [31, 32] of the anti- k_t algorithm is used.

Calorimeter jets are reconstructed using three-dimensional calorimeter topo-clusters. The basic topo-cluster definition described in the previous section provides energy clusters reconstructed on the electromagnetic scale (EM-scale). Additionally, topo-clusters can be calibrated using a local cell energy weighting (LCW) scheme [33], whereby the average calibrated energy response of clusters identified as originating from hadrons is unity. The

energy response is defined as the ratio of the measured energy to the deposited energy. The LCW calibration is designed to reduce energy fluctuations due to calorimeter non-compensation, thereby improving the energy resolution as compared to jets reconstructed using EM-scale clusters [34].

Simulated events are used to build particle-level jets (excluding muons and neutrinos) from stable particles in the event generator record as defined in Section 3.4. The exclusion of muons and neutrinos ensures that the particle-level jets are built only from particles that deposit a significant amount of energy in the calorimeters.

4.2 Electron reconstruction

Electrons leave charged particle tracks in the ID and also deposit energy in the calorimeters. Electron candidates are reconstructed by exploiting the information from both these sub-detectors. First, topological energy clusters in the EM calorimeter are built using the sliding window technique described previously. Next, the candidate cluster is required to have an associated track reconstructed in the ID. The requirement for cluster to track association (‘matching’) is that the distance between the cluster and a track with significant transverse momentum ($p_T^{\text{track}} > 500 \text{ MeV}$) must satisfy $\Delta\eta < 0.05$ and $\Delta\Phi < 0.1$ in the direction of the track’s curvature, with $\Delta\Phi < 0.05$ on the opposite side of the tracks curvature. The asymmetric Φ treatment is to allow for energy losses due to bremsstrahlung radiation along the electron’s trajectory through the ID. Further offline identification requirements based on shower shapes in the EM calorimeters and reconstructed track quality are implemented, these ensure the electrons entering into a physics analysis are very well

defined.

4.3 Photon reconstruction

Photons candidates are primarily reconstructed from energy clusters in the EM calorimeter, built using the sliding window technique described previously. Full photon identification also requires tracking information, therefore only photons reconstructed within the tracking acceptance of the ID ($|\eta| < 2.5$) are considered. A photon passing through the material of the ID can convert into an electron-positron pair, with photon candidates classified as either unconverted or converted, where the classification is based on cluster to track matching described above. Unconverted photon candidates are required to have no tracks with significant transverse momentum associated to the cluster ($p_T^{\text{track}} < 500 \text{ MeV}$). Converted photon candidates are those clusters that have an associated track, similar to electrons. Converted photons are cleanly separated from electrons by also requiring that the associated track originated from a conversion vertex in one of the layers of the inner tracker. As for electrons, additional offline identification requirements are implemented based on the shower profile in the EM calorimeter, ensuring the photons entering the analysis are very well defined. The fraction of converted photons in a sample depends on the photon's pseudorapidity and transverse energy due to the different amounts of ID material the photon must traverse before reaching the calorimeter.

Overwhelmingly, the majority of all objects reconstructed in ATLAS are jets. Therefore aggressive and efficient procedures are required to properly identify photons and reject jets from QCD dijet (dijet events produced in QCD hard scattering processes) and photon-

jet backgrounds. Calorimeter shower profiles are the primary discriminators used for photon identification. Narrow showers that are well-contained in the EM calorimeter are expected for photons, while showers induced by hadronic particles are broader and deposit a significant portion of their energy in the hadronic calorimeter. Due to the high degree of relevancy of photons to the measurement of the Higgs boson cross sections in the $H \rightarrow \gamma\gamma$ analysis channel, further details of their offline identification will now be discussed.

4.3.1 Offline photon identification

Photon identification in ATLAS is achieved using a combination of nine calorimeter based variables that describe the shower profile. Two main classifications of photons are selected from the full sample of photon candidates, they are: *Tight photons* and *Loose photons*. Tight photons are required to pass selection criteria applied on all nine of the shower profile variables and are considered to be highly photon-like, while Loose photons are only required to pass selection criteria on three of the shower profile variables. An alternative to the standard Loose photon selection (Loose') is also implemented for photon background studies (see Chapter 7). The Loose' photon selection applies a shower profile selection criteria that is intermediate to Loose and Tight. The nine variables, their meaning, and their usage in the various photon definitions are presented in Table 4.1.

4.4 Muon reconstruction

Muons are charged leptons with a mass approximately 207 times the electron mass. Due to their large mass muons typically pass through the detector unimpeded, depositing only

minimal energy in the calorimeters. They play an integral role in many physics analyses, making efficient and accurate muon identification and reconstruction of great importance to the ATLAS physics program. The ATLAS muon reconstruction algorithms [35, 36] are used to build muon candidates in the MS using information from the RPCs, TGCs, CSCs, and MDTs. First, tracks in each individual chamber are reconstructed, then a global fit is performed on the track segments to obtain the final muon track in the MS. Charge identification and transverse momentum measurements are performed using the reconstructed muon tracks. Additionally, independent muon momentum measurements are performed using tracks reconstructed in the inner detector. Accurate muon identification and reconstruction relies on exploiting the available information from the MS, the inner detector, and to a lesser extent the calorimeters. Due to the multiple information sources used in reconstructing muons, there are four different categories of muon candidates:

- **Stand-alone muons:** The MS alone is used to reconstruct the muon trajectory. The muon's direction, track impact parameter, and primary vertex are obtained by extrapolating the MS track back to the beam line, accounting for calorimeter energy losses.
- **Combined muon:** Independent track reconstruction is performed in the MS and ID. The combined muon track is formed by combining a stand-alone muon track with one from the ID.
- **Segment-tagged muons:** An ID muon track is identified if a track in the ID is extrapolated to the MS and is associated with at least one track segment in the MDT or CSC when a stand-alone MS track does not match the segment.

- **Calorimeter-tagged muons:** An ID muon track is identified if the track can be associated to a minimum ionizing energy deposit in the calorimeter. Calorimeter-tagged muons have the lowest purity of the four muon types but can be used to enhance the muon acceptance in deficient regions of the MS.

4.5 Missing E_T reconstruction

Neutrinos are neutral and massless particles that only interact via the weak force and generally pass through matter unimpeded and unaltered, therefore they leave no tracks nor energy deposits in the ATLAS detector. The presence of neutrinos is deduced by measuring the transverse momentum imbalance in a given event. The missing transverse momentum, E_T^{miss} , is reconstructed by vectorially summing all energy deposits in the calorimeter in addition to all muon transverse momentum. Since the initial momentum in the transverse plane must be zero by energy conservation, any imbalance between zero and the vectorial sum is equivalent to E_T^{miss} .

4.6 Summary

The principal procedures used to reconstruct and identify photons, electrons, muons, hadronic jets, and missing transverse energy have been introduced and discussed. In addition to reconstruction and identification, physics objects generally require calibration. Calibration refers to the porting of a measured energy in the detector to the actual energy of the particles emitted from the interaction, accounting for effects like pileup, noise, non-

uniform response, leakage, and dead material in the detector. Each physics object has a dedicated calibration strategy. A detailed account of the calibration of hadronic jets will be given in the next chapter as this is an area in which the author played a significant role. The calibration of all other physics objects will not be discussed but should be assumed in the remainder of this work.

Table 4.1: Calorimeter shower profile variables used for Loose, Loose', and Tight photon identification selection.

Category	Description	Name	Tight	Loose	Loose'
Hadronic leakage		\mathcal{R}_{had}			
	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (applied to the range $ \eta < 0.8$ and $ \eta > 1.37$)	\mathcal{R}_{had_1}	✓	✓	✓
	Ratio of E_T in the full hadronic calorimeter to E_T of the EM cluster (applied to the range $ \eta < 0.8$ and $ \eta > 1.37$)	\mathcal{R}_{had_2}	✓	✓	✓
EM middle layer					
	Ratio in η of cell energies in 3×7 versus 7×7 cells	\mathcal{R}_η	✓	✓	✓
	Total lateral shower width	ω_2	✓	✓	✓
	Ratio in Φ of cell energies in 3×3 versus 3×7 cells	\mathcal{R}_Φ	✓		✓
EM first layer ('strips')					
	Shower width for three strips around maximum strip	ω_{s3}	✓		
	Total lateral shower width	$\omega_{s\,tot}$	✓		✓
	Fraction of energy outside core of three central strips but within seven strips	F_{side}	✓		
	Difference between the energy associated with the second maximum in the strip layer, and the energy measured in the strip with minimum value found between the first and second maxima	ΔE	✓		
	Ratio of the energy difference associated with the largest and second largest energy deposits over the sum of these energies	E_{ratio}	✓		

Chapter 5

Hadronic jet calibration in ATLAS

Hadronic jets are reconstructed from the EM-scale or LCW-scale topo-clusters using the anti- k_t algorithm, as described in Section 4.1. The calibration must account for numerous effects:

1. Differences in the energy scales measured for hadronic and EM showers (calorimeter non-compensation).
2. Energy losses due to detector dead material.
3. Energy leakage out of the calorimeter.
4. Energy losses due to low energy particles that do not deposit enough energy to pass the cell noise thresholds.
5. Additional energy deposits due to pileup.

The final jet energy scale calibration is obtained using a combination of simulation based and data-driven methods. First, a simulation based primary jet energy calibration as

a function of jet pseudorapidity and transverse momentum is applied, then residual data-driven corrections are applied in order to minimize any effects due to mis-modelling in the simulation.

The primary jet calibration is implemented through the following four-step procedure. First, the direction of the jet is corrected so that it points back to the primary interaction vertex (PV_0) instead of the detector centre [37]. Second, energy deposits originating from additional proton-proton interactions (pileup) are subtracted from the jet using a jet-area based procedure [37]. The jet-area technique utilizes the pileup energy density (ρ) in the $\eta \times \Phi$ -plane multiplied by the area of a jet (A_j) in the same plane to subtract the pileup energy contribution to the jet ($E_{\text{jet}}^{\text{corrected}} = E_{\text{jet}} - \rho \times A_j$). Third, simulation based energy-dependent and pseudorapidity-dependent calibration factors are applied to correct the jet energy to the particle level [34]; these calibration factors are determined using simulated events by matching calorimeter jets to particle-level jets and taking the ratio of the calorimeter and particle-level jet energies. Finally, a Global Sequential energy scale calibration (GSC) is applied. The GSC calibration is parameterized in terms of the jet transverse momentum (p_T), the detector pseudorapidity (η_{det}), and a set of global jet observables such as the longitudinal structure of the calorimeter energy depositions, information from tracks associated to the jet, and information related to the amount of activity in the muon chamber behind a jet [38]. The detector pseudorapidity, η_{det} , is defined as the pseudorapidity of the object with respect to the detector centre, with all stages of the jet calibration procedure performed in terms of η_{det} . After the full calibration, the jet energy scale (JES) of jets built from EM- or LCW-scale topo-clusters are referred to as EM+JES or LCW+JES, respectively.

The jet calibration procedure outlined above is dependent on Monte Carlo (MC) simulations to correct the jet response. After calibrating jets using the MC-based procedure just described, *in-situ* calibration factors are applied in order to correct for features not captured by the initial calibration or due to mis-modelling in the simulation. The term "in-situ" refers to the procedure of using the data to determine the calibration scale factors. All the *in-situ* techniques utilized exploit the momentum balance of physics objects in the transverse plane. The transverse momentum (p_T) of a well defined reference object (photon, Z boson, or other jets) is compared to the p_T of the jet being calibrated in both data and MC, with the data to MC ratio defining the residual scale corrections:

$$\frac{\mathcal{R}_{\text{MC}}}{\mathcal{R}_{\text{data}}} = \frac{\langle p_T^{\text{jet}} / p_T^{\text{ref}} \rangle_{\text{MC}}}{\langle p_T^{\text{jet}} / p_T^{\text{ref}} \rangle_{\text{data}}} . \quad (5.1)$$

The first stage of the *in-situ* calibration chain is to apply an η -dependent forward jet energy calibration (η -intercalibration) derived using dijet events. The η -intercalibration is a relative calibration that corrects forward jets ($|\eta_{\text{det}}| > 0.8$) to the scale of central jets ($|\eta_{\text{det}}| \leq 0.8$). Following the η -intercalibration, residual corrections to the central jet energy scale are derived by exploiting the p_T -balance of Z bosons or photons recoiling against jets in Z +jet and γ +jet events. The Z +jet and γ +jet measurements are obtained for jets with $20 \text{ GeV} \leq p_T \leq 200 \text{ GeV}$ and $30 \text{ GeV} \leq p_T \leq 800 \text{ GeV}$, respectively. Finally, the central jet calibration is extended to high- p_T ($p_T > 800 \text{ GeV}$) using a multi-jet balance (MJB) technique. MJB exploits the transverse momentum balance in events where the highest- p_T jet is produced back-to-back with a recoil system composed of multiple lower- p_T jets. The low- p_T jets used in the construction of the recoil system are fully calibrated, including all

other *in-situ* calibrations.

The following sections describe in detail the *in-situ* method used for deriving the η -intercalibration, as this is the area in which the author made a significant contribution. All results presented correspond to measurements obtained using the 8 TeV dataset.

5.1 Forward jet energy scale calibration using dijet events: η -intercalibration

5.1.1 Monte Carlo simulation

Inclusive QCD jet production is simulated using the POWHEG BOX [39, 40, 41], SHERPA 1.4.5 [42], PYTHIA 8.160 [43] and HERWIG++ 2.5.2 [44, 45] event generators.

The POWHEG BOX event generator produces inclusive dijet events at next-to-leading-order (NLO) accuracy in perturbative QCD [46], using the CT10 parton distribution functions (PDFs) [47]. The fully hadronic final state is produced by interfacing the POWHEG BOX to PYTHIA 8.175, which provides parton showering, hadronisation and multiple parton interactions (MPI). These predictions are produced using the AU2 tune [48] for underlying event activity and are referred to as POWHEGPYTHIA8 in the remainder of this chapter. A “tune” refers to adjusting (*i.e.* tuning) the parameters of PDF set to better conform to experimental observations. An additional sample of POWHEG events is generated using HERWIG 6.520.2 for parton showering and hadronisation, and JIMMY for modelling MPI. These predictions are referred to as POWHEG-HERWIG in the remainder of this chapter.

SHERPA is a matrix-element plus parton-shower generator that is used to produce n -jet

predictions ($n = 2, 3, 4 \dots$) at leading-order (LO) accuracy in perturbative QCD. The CKKW method [49] is used to combine these various final-state topologies to produce an inclusive dijet sample. Parton-shower, hadronisation and MPI algorithms create the fully hadronic final state. The SHERPA predictions are produced using the CT10 PDFs and SHERPA’s default generator tune for underlying event activity.

5.1.2 Event selection

Events are required to have at least one reconstructed interaction vertex with at least two associated tracks. Reconstructed jets are required to pass data quality criteria that remove fake jets originating from noise bursts in the calorimeters, from non-collision backgrounds, and from cosmic rays [50]. Jets are required to originate from the primary interaction vertex using the jet vertex fraction (JVF) [51]. Tracks are first associated to jets, then the JVF variable is defined as the summed transverse momentum of associated tracks that originate from the primary vertex, divided by the summed transverse momentum of associated tracks from all vertices. Any jet with $|\eta_{\text{det}}| < 2.4$ and $p_{\text{T}} < 50$ GeV is required to have $\text{JVF} > 0.25$. This selection has been shown to efficiently reject jets originating from “pileup”. η_{det} is defined as the pseudorapidity of the jet with respect to the detector centre.

Dijet events were retained using a combination of central ($|\eta_{\text{det}}| < 3.1$) and forward ($|\eta_{\text{det}}| > 2.8$) jet triggers. The event selection is designed such that the trigger efficiency, for a specific region of $p_{\text{T}}^{\text{avg}}$, is greater than 99 % and approximately flat as a function of the pseudorapidity of the probe jet. Due to the different prescales for the central and forward jet triggers, the data collected by the different triggers correspond to different integrated luminosities. The data are assigned luminosity-dependent and prescale-dependent weights

according to the ‘exclusion’ method described in [52]. Selected events are required to have at least two jets with $p_T > 25$ GeV and $|\eta_{\text{det}}| < 4.5$. To select events with an exclusive dijet topology, the azimuthal angle between the two leading jets is required to be $\Delta\phi(\text{jet}_1, \text{jet}_2) > 2.5$, additionally events are rejected if they contain a third jet with $p_T^{\text{jet3}} > 0.4 \cdot p_T^{\text{avg}}$.

5.1.3 Dijet balance

The relative response of the ATLAS calorimeter to jets is investigated as a function of pseudorapidity using QCD events with dijet topologies. At leading-order in QCD, the two jets are expected to have equal transverse momentum, with any imbalance in the transverse momentum arising from different responses to jets in different calorimeter regions. The standard observable used to quantify the transverse momentum balance is the dijet asymmetry, defined as:

$$\mathcal{A} = \frac{p_T^{\text{probe}} - p_T^{\text{ref}}}{p_T^{\text{avg}}}, \quad (5.2)$$

where p_T^{ref} is the transverse momentum of a jet in a well-calibrated reference region (the central barrel, $|\eta_{\text{det}}| \leq 0.8$), p_T^{probe} is the transverse momentum of the jet in the calorimeter region being probed, and $p_T^{\text{avg}} = (p_T^{\text{probe}} + p_T^{\text{ref}})/2$.

The calorimeter response to a jet relative to the reference region, $1/c$, is then defined as

$$\frac{1}{c} = \frac{p_T^{\text{probe}}}{p_T^{\text{ref}}} = \frac{2 + \langle \mathcal{A} \rangle}{2 - \langle \mathcal{A} \rangle}, \quad (5.3)$$

where c is the correction factor and $\langle \mathcal{A} \rangle$ is the mean of the asymmetry distribution in a given bin of p_T^{avg} and η_{det} . The mean of the asymmetry distribution is obtained from the fit of a

Gaussian function. Presented in Figure 5.1 are two representative asymmetry distributions corresponding to the central reference method (see below).

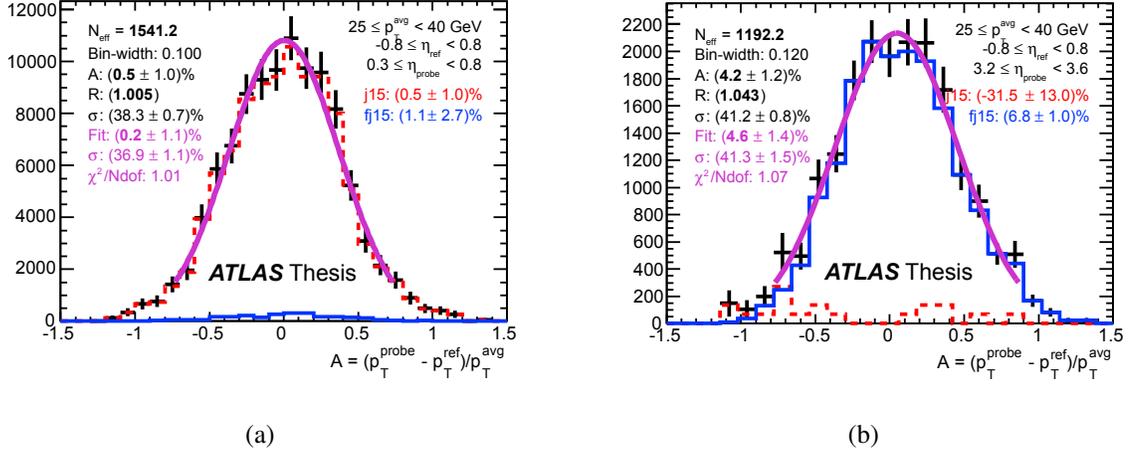


Figure 5.1: Two representative asymmetry distributions corresponding to the central reference method, obtained for jets with $25 \text{ GeV} \leq p_T^{\text{avg}} < 40 \text{ GeV}$ and a probe jet pseudo rapidity in the range of (a) $0.3 \leq \eta_{\text{probe}} < 0.8$ and (b) $3.2 \leq \eta_{\text{probe}} < 3.6$. The red and blue histograms correspond to data collected by central and forward triggers, respectively, with the black points being the sum of the two. The thick magenta line is the Gaussian function used to obtain the mean and width of the asymmetry distribution. N_{eff} is the number of events in the distribution, while A , R , and σ (left side, black text) are the mean asymmetry, the response, and the root mean square (RMS) of the asymmetry distribution, respectively. The results of the Gaussian fit are given on the left (magenta text).

The standard approach for η -intercalibration with dijet events is to define a single central calorimeter region ($|\eta_{\text{det}}| < 0.8$) as the reference region, referred to as the *central reference method*. If both jets fall into the reference region, each jet is used, in turn, to probe the other. As a consequence, the average asymmetry in the reference region will be zero by design, and the average response will therefore be unity. A notable disadvantage with the central reference method is that all events are required to have a jet in the central reference region. This results in a significant loss of statistics, particularly in the forward calorimeter

region due to the dijet cross section falling rapidly as the rapidity interval between the jets increases.

An alternative approach is to choose multiple reference regions and measure the calorimeter response of a given region relative to all other reference regions. The calorimeter response relative to the central region as a function of p_T^{avg} and η_{det} is obtained by solving a matrix of linear equations. This approach is therefore referred to as the *matrix method*. The advantage of the matrix method is that the full dataset is used thereby reducing the statistical uncertainty on the measurement.

Two versions of the analysis are performed, the main results are obtained using the matrix method and the central reference method is only used for validation. Further details of the matrix method are given in the following section.

5.1.4 The matrix method

The matrix method extends the central reference method by replacing the “probe” and “reference” jets by “left” and “right” jets defined from $\eta^{\text{left}} < \eta^{\text{right}}$, exploiting the full statistics of the available dataset. Equations 5.2 and 5.3 then become:

$$\mathcal{A} = \frac{p_T^{\text{left}} - p_T^{\text{right}}}{p_T^{\text{avg}}}, \quad \text{and} \quad \mathcal{R} = \frac{p_T^{\text{left}}}{p_T^{\text{right}}} = \frac{c^{\text{right}}}{c^{\text{left}}} = \frac{2 + \mathcal{A}}{2 - \mathcal{A}}, \quad (5.4)$$

where the term \mathcal{R} denotes the ratio of the responses, and c^{left} and c^{right} are the η -intercalibration factors for the left and right jets, respectively.

In this method there is a response ratio distribution, \mathcal{R}_{ijk} , whose average value $\langle R_{ijk} \rangle$ is evaluated for each η^{left} -bin i , η^{right} -bin j and p_T^{avg} -bin k . For a specified p_T^{avg} -bin k , the

relative correction factor c_{ik} for given jet η -bin i is obtained by minimizing a matrix of linear equations:

$$S(c_{1k}, \dots, c_{Nk}) = \sum_{j=1}^N \sum_{i=1}^{j-1} \left(\frac{1}{\Delta \langle \mathcal{R}_{ijk} \rangle} (c_{ik} \langle \mathcal{R}_{ijk} \rangle - c_{jk}) \right)^2 + X(c_{1k}, \dots, c_{Nk}), \quad (5.5)$$

where N are the number of η -bins, $\Delta \langle \mathcal{R} \rangle$ is the statistical uncertainty of $\langle \mathcal{R} \rangle$ and the function $X(c_{ik})$ is used to quadratically suppress large deviations from unity of the average corrections¹. The minimization of Equation 5.5 is done separately for each p_T -bin k , and the resulting calibration factors c_i (for each jet η -bin i) are scaled such that the average calibration factor in the reference region $|\eta_{\text{det}}| < 0.8$ equals unity.

The statistical uncertainties assigned to the relative responses obtained from the matrix method are estimated using pseudo-experiments. In each pseudo-experiment a random number is drawn from a standard normal, $N(0,1)$, distribution and the mean of the response ratio distribution, $\langle R_{ijk} \rangle$, is shifted by adding to it the value obtained from multiplying the normally distributed random number by $\Delta \langle \mathcal{R} \rangle$. The minimization is then redone. This procedure is repeated a hundred times and the RMS of the resulting response distributions is taken as the statistical uncertainty.

¹ $X(c_{1k}, \dots, c_{Nk}) = K \left(N_{\text{bins}}^{-1} \sum_{i=1}^{N_{\text{bins}}} c_{ik} - 1 \right)^2$, with K being a constant and N_{bins} being the number of η -bins (number of indices i). This term prevents the minimization from choosing the trivial solution: all c_{ik} equal to zero. The value of the constant K does not influence the solution as long as it is sufficiently large ($K \approx N_{\text{bins}}$ or larger).

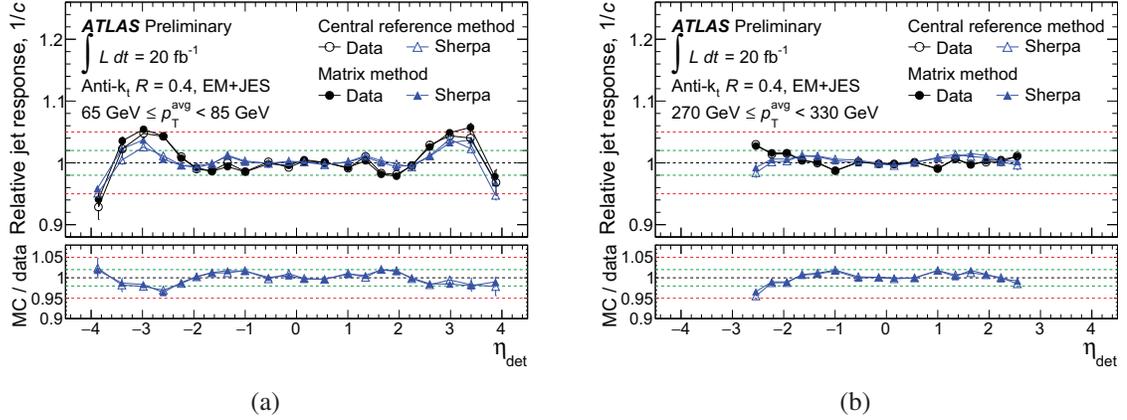


Figure 5.2: Relative response measured using the matrix and the central reference methods for anti- k_t jets with $R = 0.4$ calibrated with the EM+JES scheme as a function of the probe jet η . Results are presented for $65 \text{ GeV} \leq p_T^{\text{avg}} < 85 \text{ GeV}$ and $270 \text{ GeV} \leq p_T^{\text{avg}} < 330 \text{ GeV}$. The green and red dashed lines indicate $1 \pm 2\%$ and $1 \pm 5\%$, respectively.

5.1.5 Comparison of matrix and central reference methods

Figure 5.2 compares the relative jet response obtained using the matrix and central reference methods for $65 \text{ GeV} \leq p_T^{\text{avg}} < 85 \text{ GeV}$ and $270 \text{ GeV} \leq p_T^{\text{avg}} < 330 \text{ GeV}$. The matrix method exhibits a slight shift in the relative response compared to the central reference method, most notably in the forward regions. However, the same shift appears in both data and the MC simulation, resulting in consistent data-to-MC ratios. Since it is the ratio that determines the residual calibration the slight shift between the two methods is of no concern. For the most statistically limited p_T -bin ($25 \text{ GeV} \leq p_T^{\text{avg}} < 40 \text{ GeV}$) the matrix method generally results in 40% increase in precision compared to the central reference method. The gained precision decreases with increasing p_T^{avg} and is typically less than 10% for $p_T^{\text{avg}} > 400 \text{ GeV}$. In the remainder of this work all results presented correspond to the matrix method.

5.1.6 Comparison of data with simulation

The relative response as a function of the jet η for data and the MC simulations are shown in Figure 5.3. Four different p_T^{avg} regions are presented: $25 \text{ GeV} \leq p_T^{\text{avg}} < 40 \text{ GeV}$, $85 \text{ GeV} \leq p_T^{\text{avg}} < 115 \text{ GeV}$, $220 \text{ GeV} \leq p_T^{\text{avg}} < 270 \text{ GeV}$ and $760 \text{ GeV} \leq p_T^{\text{avg}} < 1200 \text{ GeV}$. Figure 5.4 shows the relative response as a function of p_T^{avg} for two representative η_{det} -bins, specifically $-1.5 \leq \eta_{\text{det}} < -1.2$ and $2.1 \leq \eta_{\text{det}} < 2.4$. The SHERPA and POWHEGPYTHIA8 predictions reproduce reasonably well the general features observed in data. Furthermore, the theoretical predictions are in good agreement with each other.

5.1.7 Derivation of a residual Jet Energy Scale correction

The ratio of η -intercalibration factors obtained from data and SHERPA, *i.e.* $C_i = c_i^{\text{data}} / c_i^{\text{SHERPA}}$, are used to derive the residual calibration factors. A smooth calibration function is obtained by combining the calibration factors from many bins of p_T^{avg} and η_{det} using a two dimensional Gaussian kernel,

$$F_{\text{rel}}(p_T, \eta_{\text{det}}) = \frac{\sum_{i=1}^{N_{\text{bins}}} C_i w_i}{\sum_{i=1}^{N_{\text{bins}}} w_i}, \text{ with} \quad (5.6)$$

$$w_i = \frac{1}{\Delta C_i^2} \times \text{Gaus} \left(\frac{\log p_T - \log \langle p_T^{\text{probe}} \rangle_i}{\sigma_{p_T}} \oplus \frac{\eta_{\text{det}} - \langle \eta_{\text{det}} \rangle_i}{\sigma_\eta} \right),$$

where i is the index of a $(p_T^{\text{avg}}, \eta_{\text{det}})$ -bin, ΔC_i is the statistical uncertainty of C_i , $\langle p_T^{\text{probe}} \rangle_i$ and $\langle \eta_{\text{det}} \rangle_i$ are the average p_T and η_{det} of the probe jets in bin- i , $\text{Gaus}(x)$ is the amplitude

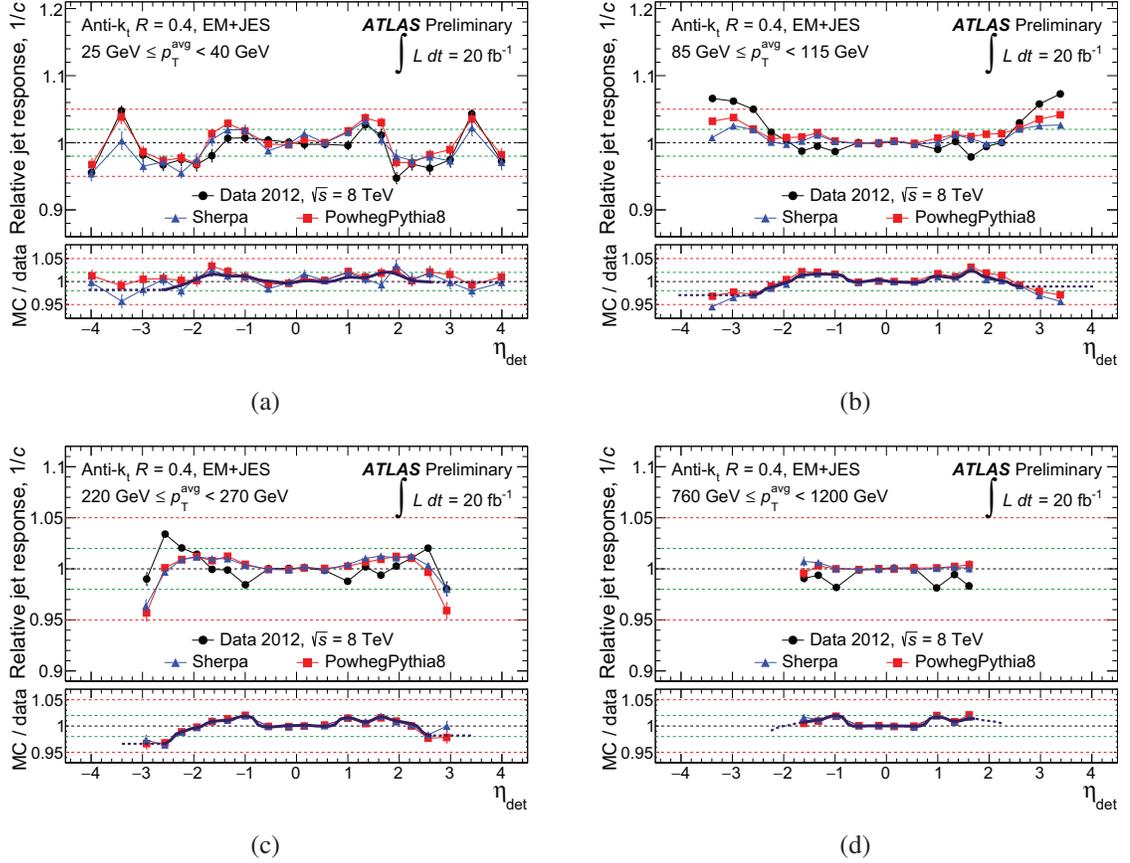


Figure 5.3: Relative jet response, $1/c$, as a function of the jet η for anti- k_t jets with $R = 0.4$ calibrated with the EM+JES scheme, separately for $25 \text{ GeV} < p_T^{\text{avg}} < 40 \text{ GeV}$, $85 \text{ GeV} < p_T^{\text{avg}} < 115 \text{ GeV}$, $220 \text{ GeV} < p_T^{\text{avg}} < 270 \text{ GeV}$ and $760 \text{ GeV} < p_T^{\text{avg}} < 1200 \text{ GeV}$. The lower parts of the figures show the ratios between the data and MC relative response. These measurements are performed using the matrix method. The green and red dashed lines indicate $1 \pm 2\%$ and $1 \pm 5\%$, respectively.

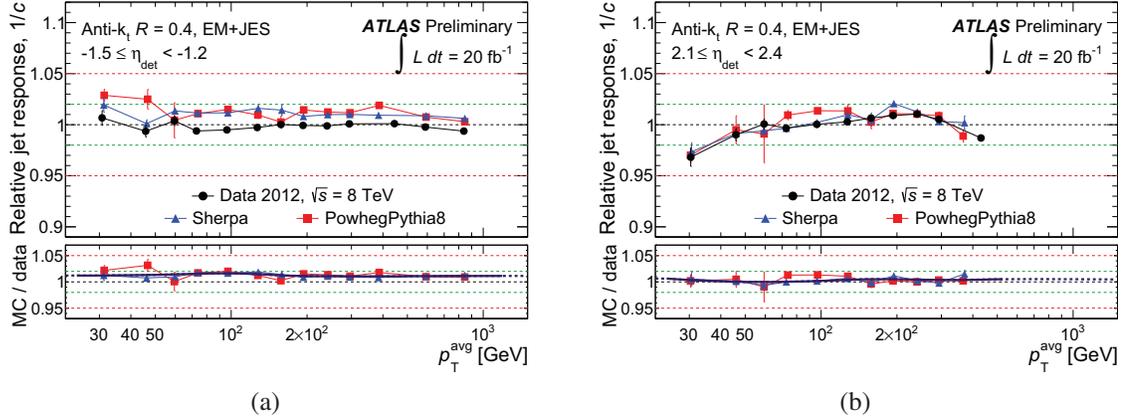


Figure 5.4: Relative jet response, $1/c$, as a function of the jet p_T for anti- k_t jets with $R = 0.4$ calibrated with the EM+JES scheme, separately for $-1.5 \leq \eta_{\text{det}} < -1.2$ and $2.1 \leq \eta_{\text{det}} < 2.4$. The lower parts of the figures show the ratios between the data and MC relative response. The green and red dashed lines indicate $1 \pm 2\%$ and $1 \pm 5\%$, respectively.

of a Gaussian function with $\mu = 0$ and $\sigma = 1$, σ_{p_T} and σ_η are width-parameters² of the Gaussian kernel and \oplus denotes addition in quadrature. The kernel-width parameters are chosen to capture the shape of the data-MC ratio across p_T and η_{det} , while also providing stability against statistical fluctuations, particularly in the low- p_T region. The derivation of the correction function excludes measurements in the region $|\eta_{\text{det}}| > 2.7$, with the correction function for further forward jets fixed to the value obtained at $|\eta_{\text{det}}| = 2.7$. Therefore the same residual correction factor is applied to all jets with a given p_T and $|\eta_{\text{det}}| \geq 2.7$.

The resulting residual correction is shown as a black line in the lower panels of Figures 5.3 and 5.4. The line is solid over the range where the measurements are used to constrain the calibration, and dashed in the range where extrapolation is performed. The

²The kernel width parameters varied with p_T , with stronger smoothing applied in low p_T ranges to reduce effects from statistical fluctuations: $\sigma_{p_T} = 0.55$ and $\sigma_\eta = 0.16$ for $p_T = 25$ GeV transitioning smoothly to $\sigma_{p_T} = 0.28$ and $\sigma_\eta = 0.12$ for $p_T = 50$ GeV and beyond.

calibration factors range from 0.96 to 1.02 depending on the η of the jet and can essentially be considered constant as a function of p_T .

5.1.8 Systematic uncertainties

Observed differences between the relative response obtained in data and simulation could be caused by simulation mis-modelling of physics and/or detector effects. All systematic uncertainties are derived as a function of p_T and $|\eta_{\text{det}}|$, with no uncertainty assigned in the reference region ($|\eta_{\text{det}}| < 0.8$). The parameterization of the systematic uncertainty in terms of $|\eta_{\text{det}}|$ is justified since no statistically significant difference is observed for positive and negative η_{det} for any of the uncertainties.

The SHERPA and POWHEGPYTHIA8 are used to assess the physics modelling uncertainty. Both of these generators are accurate to leading-order in QCD for variables sensitive to the third jet modelling (such as the dijet balance). Since there is no *a priori* reason to trust one generator over the other, the difference between the two predictions is used to assess the physics modelling uncertainty. For $0.8 \leq |\eta_{\text{det}}| < 2.7$, where the data are corrected to the SHERPA predictions, the full difference between POWHEGPYTHIA8 and SHERPA is taken as the uncertainty. For $|\eta_{\text{det}}| \geq 2.7$, where the calibration is frozen, the uncertainty is taken as the maximum difference between the calibration and either POWHEGPYTHIA8 or SHERPA. The modelling uncertainty is the dominant systematic uncertainty on the measurement.

The physics modelling uncertainty on the relative response can also be cross-checked at particle-level by varying the POWHEGPYTHIA8 predictions. First, the renormalisation and factorisation scales in the POWHEG Box are each varied by a factor of 0.5 and 2.0. The change in the relative response is observed to be much smaller than the difference between

the POWHEGPYTHIA8 and SHERPA relative responses. Secondly, the change in the particle-level relative response of the POWHEG-HERWIG sample with respect to the POWHEGPYTHIA8 sample is found to be similar to the difference in the particle-level relative response between the POWHEGPYTHIA8 and SHERPA samples. It is concluded that the difference between SHERPA and POWHEGPYTHIA8 is a good reflection of the underlying physics modelling uncertainty.

The event topology selection requires that the two leading jets have a $\Delta\phi$ separation greater than 2.5 radians. In order to assess the influence of this selection on the transverse momentum balance, the residual calibration is re-derived after shifting the cut up and down by ± 0.3 radian. The maximum difference between the shifted and nominal calibrations is taken as the uncertainty.

To assess the impact of pile-up on the calibration the difference between low and high pile-up subsets is investigated. The data is divided into high and low $\langle\mu\rangle$ subsets ($\langle\mu\rangle < 14$ and $\langle\mu\rangle \geq 17$), and high and low N_{PV} subsets ($N_{PV} < 9$ and $N_{PV} \geq 11$), where $\langle\mu\rangle$ is the average number of interactions per bunch crossing and N_{PV} is the number of reconstructed primary vertices in the inner detector. The calibration is re-evaluated for each subset and the uncertainty due to pile-up effects is taken to be the maximum fractional difference between the varied and nominal calibrations.

To assess the uncertainty on the calibration due to the JVF requirement on jets, the JVF cut is shifted up and down. The up/down variations account for the extent to which JVF is mis-modelled for jets originating from the primary interaction vertex. The calibration is re-evaluated for both the JVF-up and JVF-down cases and the uncertainty is taken as the maximum fractional difference between the JVF-varied calibration and the nominal

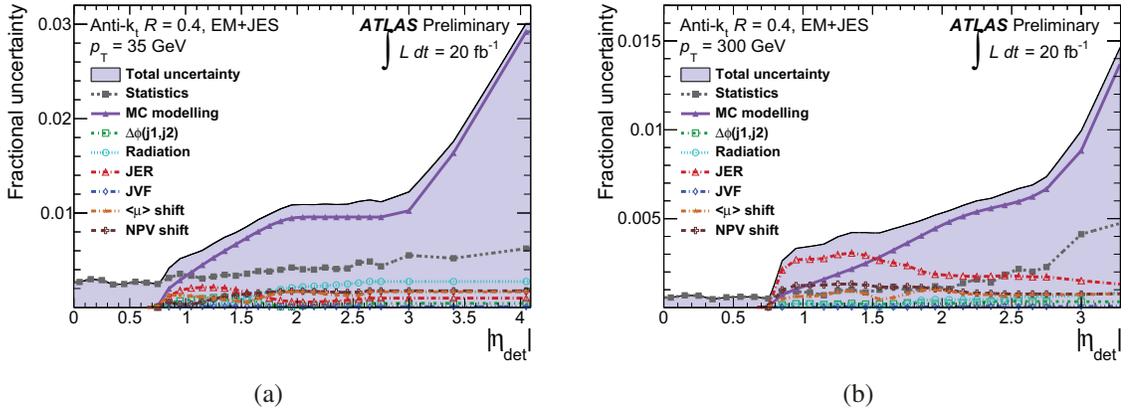


Figure 5.5: Summary of uncertainties on the intercalibration as a function of the jet η_{det} for anti- k_t jets with $R = 0.4$ calibrated with the EM+JES scheme, separately for (a) $p_T = 35$ GeV and (b) $p_T = 300$ GeV. The individual components are added in quadrature to obtain the total uncertainty. The modelling uncertainty is the dominant component.

calibration.

The jet energy resolution (JER) [53, 37] in the simulation is comparable to the resolution observed in data. To assess the impact of the JER on the transverse momentum balance, the jet energies in simulated events are smeared using a Gaussian function with width set to the JER uncertainty in the data. The difference between the calibrations obtained with nominal and smeared simulation is taken as the uncertainty due to JER effects.

The total systematic uncertainty is obtained as the quadratic sum of the various components mentioned. Figure 5.5 presents a summary of the uncertainties as a function of η_{det} for two representative values of jet transverse momentum, namely $p_T = 35$ GeV and $p_T = 300$ GeV. The uncertainties have a strong η dependence, increasing with η_{det} . The uncertainties have a weaker p_T dependence, decreasing with increasing jet transverse momentum.

Figure 5.6 shows the final η -intercalibration and its associated uncertainties as a function

of jet pseudorapidity for two representative intervals of p_T^{avg} .

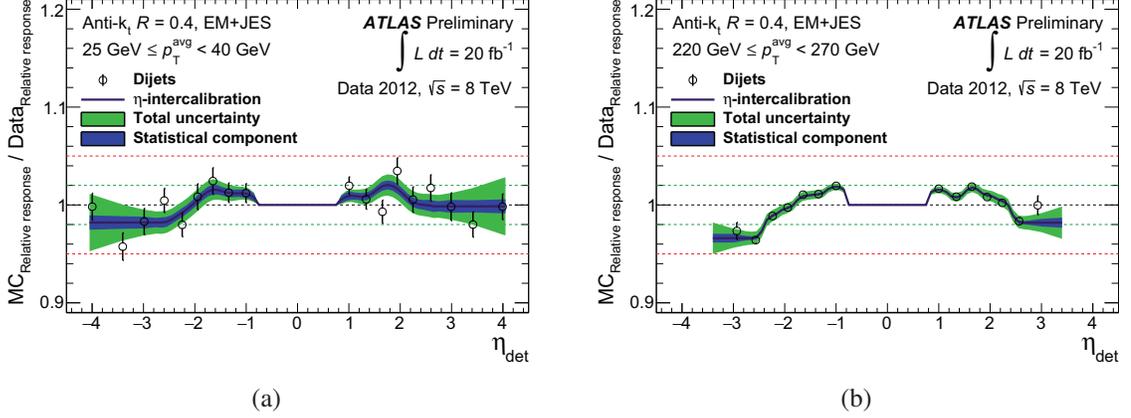


Figure 5.6: Final $\sqrt{s} = 8$ TeV η -intercalibration and associated uncertainties as a function of the jet pseudorapidity for anti- k_t jets with $R = 0.4$ calibrated with the EM+JES scheme, separately for (a) $25 \leq p_T^{\text{avg}} < 40$ GeV and (b) $220 \leq p_T^{\text{avg}} < 270$ GeV. The green and red dashed lines indicate $1 \pm 2\%$ and $1 \pm 5\%$, respectively.

5.2 Summary

The response of the ATLAS calorimeter to jets has been studied using proton-proton collisions at $\sqrt{s} = 8$ TeV recorded by the ATLAS experiment in 2012. An overview of the full hadronic jet calibration procedure has been provided, with a particular focus on the forward jet energy scale calibration. The forward jet energy scale calibration and its associated uncertainties have been derived as a function of transverse momentum and η through studying the η dependence of the forward calorimeter's response to jets relative to a central calorimeter region by evaluating the transverse momentum balance in dijet events. A residual transverse momentum and η -dependent jet calibration is derived for jets in data, in order to correct for effects not captured by the default calibration derived using MC

simulation. These calibration factors range from 0.96 to 1.02 depending on the η of the jet. The uncertainties on these calibration factors are typically below 1% for central jets rising to 3.5% for low transverse momentum jets at high absolute η . The measurements presented in this chapter were reported in two ATLAS public conference notes [54, 55]. While not presented here, the author was also responsible for providing the forward jet energy measurement and uncertainty using data recorded at $\sqrt{s} = 7$ TeV. The 7 TeV measurements were reported in an ATLAS public conference note [56] and also in the ATLAS publication on jet energy measurement at 7 TeV [34].

Chapter 6

Event and physics object selection for the $H \rightarrow \gamma\gamma$ analysis

The ATLAS physics object reconstruction and definitions described in Chapter 3 are used in this analysis. This chapter provides the details of the data and simulated samples and the event and object selections used in the $H \rightarrow \gamma\gamma$ analysis. Furthermore, it provides the definition of the fiducial measurement regions. A summary of the event and object selections used in this analysis is given in the following sections.

Two classifications of physics object will be discussed: reconstruction-level objects, and particle-level objects. Reconstruction-level refers to physics objects that are reconstructed at the detector-level. By definition, all objects in data are reconstruction-level objects. Simulated Monte Carlo (MC) events that have been pulled through the full ATLAS detector simulation and digitized are also used to build reconstruction-level physics objects. Particle-level physics objects are defined as physics objects built from stable particles from the MC

simulation with no consideration of the ATLAS detector. Particle-level is colloquially referred to as MC ‘truth’.

6.1 Dataset and simulated signal samples

This analysis exploits the full 7 TeV and 8 TeV diphoton datasets, corresponding to total recorded integrated luminosities of 4.5 fb^{-1} and 20.3 fb^{-1} , respectively. The datasets were collected using a diphoton trigger that selects events containing two photons that are also identified as Loose (see Section 4.3.1). The exact central values of the 7 TeV and 8 TeV integrated luminosities used in the measurements are 4519.24 pb^{-1} and 20276.9 pb^{-1} , with a luminosity uncertainty of 1.8 % and 2.8 % [57], respectively.

The production and decay of a Higgs boson are simulated for each of the dominant LHC production modes: gluon fusion (GF), vector boson fusion (VBF), associated production with a weak vector boson (WH , ZH), and associated production with a top anti-top pair ($t\bar{t}H$). Individual simulated MC samples are generated for Higgs boson masses spanning the range 100 GeV to 160 GeV in steps of 5 GeV. The samples corresponding to a Higgs boson mass of 125 GeV are used to obtain the detector correction (unfolding) factors and were generated with high statistics. Further details of the simulated MC samples can be found in Appendix A.

6.2 Trigger, data quality, and primary vertex selection applied to data and reconstructed MC

Physics events are collected using diphoton triggers specific to the 7 TeV and 8 TeV data taking conditions. The 7 TeV trigger requires two reconstructed photon candidates with E_T of at least 20 GeV, with the candidate photons required to fulfil the Loose photon identification criteria. The 7 TeV trigger efficiency for the events that pass the remaining event selection is found to be 98.9 %. The 8 TeV trigger requires two reconstructed, Loose identified, photon candidates with E_T of at least 35 GeV and 25 GeV for the leading p_T and sub-leading p_T photon candidates, respectively. The 8 TeV trigger efficiency for the events that pass the remaining event selection is found to be 99.6 %.

All ATLAS detector components relevant to the analysis are required to be fully operational during data-taking. Standard event cleaning is applied based on the detector quality flags. Events with data integrity errors in the calorimeters are rejected, as are incomplete events with missing detector information. Additionally, events which are corrupted due to power supply trips in the Tile calorimeter are rejected.

Each event is required to contain at least one reconstructed collision vertex, required to have at least three inner detector tracks with sufficient transverse momentum ($p_T > 400$ MeV).

6.3 Reconstruction-level physics object selection

6.3.1 Photons

In the $H \rightarrow \gamma\gamma$ analysis, photons take the highest priority in the object selection. Topological energy clusters formed in the electromagnetic calorimeter are used to seed photon reconstruction as described in Section 4.3. Cluster energies are calibrated separately for unconverted and converted candidates [58], with the calibration accounting for energy losses upstream of the calorimeter and also for energy leakage outside of the calorimeter cluster. Additionally, the energy measurement of converted photons is improved by applying corrections based on dedicated simulation studies that utilize the radial distance of the conversion [58].

The event selection described above requires events to have at least two Loose identified photons, with the two highest p_T photons defining the “Higgs boson candidate”.

The two Loose identified photon candidates are required to have a transverse momentum satisfying $p_T > 25$ GeV, a pseudorapidity in the interval $|\eta| < 2.37$, excluding the transition region between the barrel and endcap calorimeters ($1.37 \leq |\eta| < 1.56$). The Higgs boson candidate photons must further satisfy track and calorimeter isolation requirements. The track isolation transverse momentum is defined as the scalar sum of the p_T of tracks ($p_T^{\text{track}} > 1$ GeV) originating from the diphoton vertex that reside in a cone of $\Delta R < 0.2$ around the photon. The track isolation transverse momentum must be less than 2.2 GeV (2.6 GeV) for the 7 TeV (8 TeV) analysis. The calorimeter isolation energy is defined as the energy contained in all calorimeter topoclusters residing in a cone of $\Delta R < 0.4$ surrounding the photon, excluding the region directly surrounding the barycentre of the photon cluster.

The size of the excluded region is 0.125×0.175 in $\eta \times \Phi$. This calorimeter isolation energy is corrected for photon energy leakage out of excluded region, as well as contamination from pileup interactions [59, 60]. The calorimeter isolation energy is required to be less than 5.5 GeV (6.0 GeV) for the 7 TeV (8 TeV) analysis. Both Higgs candidate photons must also satisfy the Tight identification requirements on the calorimeter shower profiles as described in Section 4.3.1. Note that the final kinematic selection is applied as part of the fiducial definition specified in Section 6.6.

6.3.2 Electrons and muons

Electrons and muons are reconstructed as described in Section 4.2 and Section 4.4, respectively. Both electrons and muons are required to have a transverse momentum satisfying $p_T > 15$ GeV and a pseudorapidity in the interval $|\eta| < 2.47$. Electrons and muons must also pass isolation and identification criteria similar to photons. In addition, muons are required to have a transverse impact parameter of less than 1 mm and a longitudinal impact parameter of less than 10 mm. The vast majority of events entering this analysis contain no electrons or muons.

6.3.3 Jets and missing transverse energy (E_T^{miss})

Jets are built from calorimeter energy clusters at the EM-scale using the anti- k_t algorithm with clustering parameter $R = 0.4$, as described in Section 4.1. Each jet is required to have a transverse momentum satisfying $p_T > 30$ GeV and a rapidity in the interval $|y| < 4.4$. Jets originating from pile-up interactions are suppressed by applying additional selection

criteria on the JVF: each jet within the tracking acceptance ($|\eta_{\text{det}}| < 2.4$) and with a transverse momentum satisfying $p_T < 50$ GeV must also satisfy $|\text{JVF}| > 0.5$ (0.25) for 7 TeV (8 TeV). For a description of JVF refer to the introduction of Chapter 5. Jets overlapping with photons ($\Delta R(\text{jet}, \gamma) < 0.4$) or electrons ($\Delta R(\text{jet}, e) < 0.2$) are discarded.

Missing transverse energy reconstruction is described in Chapter 4. No additional requirements are made on E_T^{miss} .

6.4 Particle-level physics object selection

A robust particle-level definition must be adopted for the calculation of the unfolding factors (c_i) introduced in Equation 1.1. A suitable definition ensures the signal yields extracted from data can be translated to a set of model-independent, detector corrected cross section measurements. Stable particles from the MC simulation are used to reconstruct the particle-level physics objects. Particles produced during hadronization are excluded from particle-level physics objects. Particles originating from hard-scatter tau decays (such as W decay to leptons via an intermediate tau) are kept, while particles from tau decays originating from hadronization are discarded.

The kinematic acceptance selection criteria applied to particle-level objects is identical to that applied to reconstruction-level objects, with the only exception being for photons where the transition region between the barrel and endcap calorimeters is excluded at reconstruction-level but not at particle-level. The exclusion of this phase-space region is treated as a detector-level inefficiency and corrected for during the unfolding procedure. Section 6.6 provides the definition of fiducial measurement regions.

6.4.1 Particle-level photons

Particle-level photons are required to have a transverse momentum satisfying $p_T > 25$ GeV and a pseudorapidity in the interval $|\eta| < 2.37$, which is the same as for reconstruction-level, however the barrel-endcap transition region ($1.37 < |\eta| < 1.56$) is not removed. A particle-level isolation energy requirement that is similar to the calorimeter isolation energy requirement at the reconstruction-level is applied to particle-level photons. Each particle-level photon must have $E_T^{\text{iso}} < 14$ GeV, where E_T^{iso} is E_T of the 4-momentum sum of stable particles within a cone of $\Delta R < 0.4$ around the photon, excluding muons and neutrinos. This definition was adopted to minimize the topological model-dependence of the unfolding factors that are used to correct for experimental effects and obtain Higgs boson cross sections at the particle-level. Section 9.2.1 provides the full motivation and details of this particle-level isolation requirement. The two highest p_T photons from the collection of isolated particle-level photons defines the Higgs candidate diphoton system. Note that the final kinematic selection is applied as part of the fiducial definition specified in Section 6.6.

6.4.2 Particle-level electrons and muons

Particle-level electrons and muons are ‘dressed’ with nearby photons. Dressing refers to combining all photons within a cone of $\Delta R < 0.1$ around the electron or muon through 4-vector addition. Before dressing, the electrons and muons correspond to final state leptons after QED final-state radiation (FSR), referred to as bare leptons. By recombining the photons within a cone of $\Delta R < 0.1$ around the bare lepton, the dressing procedure

corrects for FSR effects. The dressed lepton definition is considered to be closest to what is experimentally measured. Dressed electrons and muons are required to have a transverse momentum satisfying $p_T > 15$ GeV and a pseudorapidity in the interval $|\eta| < 2.47$, which is the same as for reconstruction-level.

6.4.3 Particle-level jets and missing transverse energy (E_T^{miss})

Particle-level jets are reconstructed from all stable particles except neutrinos using the anti- k_t algorithm with a radius parameter of $R = 0.4$, as described in Section 4.1. Each jet is required to have a transverse momentum satisfying $p_T > 30$ GeV and a rapidity in the interval $|y| < 4.4$, which is the same as for reconstruction-level.

The particle-level E_T^{miss} is defined as the transverse-vector sum of all neutrinos in an event not originating from hadron decays. There are no additional requirements made on E_T^{miss} .

6.4.4 Particle-level object overlap removal

Isolated photons take the highest priority in the analysis, therefore, any electron, muon or jet overlapping with a photon is removed. Jets built from isolated electrons are also removed. Muons are not included as inputs to the jet algorithm when building particle-level jets and therefore it is not possible for a muon to be reconstructed as a jet. This definition is consistent with jets at the reconstruction-level, as muons deposit very minimal energy in the calorimeter. For the collection of particle-level objects that pass the particle-level selection criteria the overlap removal proceeds in the following order:

- Electrons overlapping with photons are removed ($\Delta R(e, \gamma) < 0.4$).
- Jets overlapping with photons ($\Delta R(\text{jet}, \gamma) < 0.4$) or electrons ($\Delta R(\text{jet}, e) < 0.2$) are removed.
- Muons overlapping with photons $\Delta R(\mu, \gamma) < 0.4$ are removed.

6.5 Corrections to the simulation

To account for observed differences between data and the simulation, the following corrections are applied to simulated signal events:

- Simulated samples are generated with a prior assumption of the expected pileup conditions (see Section 3.4). To match the pileup conditions of the collected dataset, events in the simulation are weighted based on the distributions of the average number of interactions per bunch crossing, $\langle \mu \rangle$, a quantity that is directly proportional to the instantaneous luminosity. To account for a known offset between $\langle \mu \rangle$ in data and simulation a scale factor is applied to the simulated $\langle \mu \rangle$ distributions prior to the re-weighting.
- The width of the beam spot along the beam axis, σ_z , is slightly smaller in the data compared to the MC samples and a re-weighting is applied to match the spread (and mean) of the beam spot observed in data.
- Correction factors are applied to the simulated photon shower shape variables prior to photon identification. This correction accounts for observed data-simulation differences in the shower-shape distributions of inclusive photon samples.

- The photon energy is smeared to account for differences in resolution between data and simulation.

6.6 Definition of fiducial regions

The inclusive fiducial signal region is defined from the sample of events containing two isolated Higgs candidate photons after the application of the full event selection described in the previous sections. For the enlarged 8 TeV dataset, additional fiducial measurement regions are defined corresponding to different phase-space regions. The fiducial regions are:

Baseline inclusive fiducial region:

The diphoton invariant mass is defined as:

$$m_{\gamma\gamma} = \sqrt{2E_{\gamma_1}E_{\gamma_2} [1 - \cos(\alpha)]}, \quad (6.1)$$

where $E_{\gamma_{1,2}}$ are the energies of the two photons, and α is the angle between the two photons. The diphoton invariant mass ($m_{\gamma\gamma}$) is required to be in the range $105 \text{ GeV} \leq m_{\gamma\gamma} < 160 \text{ GeV}$. Additionally, the leading (sub-leading) photon must satisfy $\frac{p_T}{m_{\gamma\gamma}} > 0.35$ (0.25).

Fiducial regions characterized by the level of inclusive hadronic jet activity:

Events are classified as having at least one jet ($N_{\text{jet}} \geq 1$), or at least two jets ($N_{\text{jet}} \geq 2$), or at least three jets, ($N_{\text{jet}} \geq 3$) with $p_T > 30 \text{ GeV}$. It should be noted here that measurements in the three fiducial regions just described are obtained from the

N_{jet} differential spectrum. The N_{jet} differential spectrum is defined in terms of three exclusive N_{jet} bins: $N_{\text{jet}} = 0$, $N_{\text{jet}} = 1$, $N_{\text{jet}} = 2$, and the inclusive $N_{\text{jets}} \geq 3$ bin. Further details on measurements in these N_{jet} fiducial regions will be given in Chapter 8.

VBF-enhanced fiducial region:

A VBF enriched phase-space region is defined by requiring events to have at least two hadronic jets with an invariant mass $m_{jj} > 400$ GeV, a rapidity separation of $\Delta y > 2.8$, and an azimuthal angular separation between the diphoton and dijet system of $\Delta\phi_{\gamma\gamma,jj} > 2.6$.

$N_{\text{leptons}} \geq 1$ fiducial region:

Events must contain at least one reconstructed muon or electron. This region of phase-space has enhanced sensitivity to Higgs boson production in association with a weak vector boson (WH or ZH). The vast majority of events in this analysis do not contain any electrons or muons.

$E_{\text{T}}^{\text{miss}} \geq 80$ GeV fiducial region:

Events are required to have a large amount of missing transverse energy ($E_{\text{T}}^{\text{miss}} > 80$ GeV). This phase-space region has enhanced sensitivity to Higgs boson production in association with a weak vector boson (WH or ZH) and also to potential contributions from Higgs boson production in association with dark matter particles. The vast majority of events in this analysis do not contain large amounts of $E_{\text{T}}^{\text{miss}}$.

6.7 Dalitz decays

By default PYTHIA8 includes Dalitz decays ($H \rightarrow \gamma\gamma^* \rightarrow \gamma f\bar{f}$) when it decays the Higgs boson to two photons. For a Higgs mass of 125 GeV, approximately 6 % of the events decayed by PYTHIA8 are Dalitz decays, while the other 94 % of the events produce two stable photons. A small portion of the Dalitz decays make it through the full analysis selection. The $H \rightarrow \gamma\gamma$ branching ratio used in this analysis does not include Dalitz decays and therefore any such contributions to the Higgs boson signal need to be accounted for. Dalitz contributions to the signal are removed through the unfolding procedure (see Section 9.2).

Chapter 7

Data-driven background rate estimates in the $H \rightarrow \gamma\gamma$ channel

The dominant non-resonant SM background processes contributing to the $m_{\gamma\gamma}$ spectrum are generic diphoton production ($\gamma\gamma$), photon plus jet production where either the photon or the jet is the leading p_T object (γ +jet and jet+ γ) and the jet is mis-identified as a photon, and dijet production (jet+jet) where both jets are mis-identified as photons. The analytic function used to parameterize the non-resonant background contributions to the $m_{\gamma\gamma}$ spectrum is chosen based on ‘spurious signal’ studies on high statistics simulated background samples as described in detail in Section 8.1.1.1. Spurious signal refers to a bias in the extracted signal due to the chosen background parameterization. Simulated background samples are generated for each of the dominant background processes individually and then combined to form the total background model of the $m_{\gamma\gamma}$ spectrum. The individual processes are combined by scaling each contribution using data-driven rate estimates for each process.

The data-driven rate estimates are obtained using a $2 \times 2D$ Sideband Method. The remaining sections of this chapter describe the details of the method and present the results for the background rate estimates used in the $H \rightarrow \gamma\gamma$ cross section analysis.

7.1 The $2 \times 2D$ Sideband Method

The $2 \times 2D$ Sideband Method categorizes an input sample of photon candidates into 16 regions: 1 signal region and 15 background control regions. The background rates in the signal region are obtained by extrapolating the rates extracted in the background control regions into the signal region. The $2 \times 2D$ Sideband Method was first applied in the isolated diphoton cross section measurements [61], with an improved version implemented for the $H \rightarrow \gamma\gamma$ analysis. The input to the method is a sample of events consisting of Loose' identified diphoton candidates (see Section 4.3.1) that pass the nominal $H \rightarrow \gamma\gamma$ selection (see Section 6.2). Recall that the Loose' criterion is defined by relaxing four (W_{S3} , F_{side} , ΔE , E_{ratio}) out of five of the shower profile variables that define the Tight photon identification. The sample of Loose' diphoton candidates is composed of contributions of $\gamma\gamma$, γ +jet, jet+ γ , and jet+jet events. The Higgs boson signal region is included in the input data sample, however, this does not pose any problems since the Higgs boson signal events represent an essentially negligible contribution to the overall sample.

Application of the isolation criteria¹ (I) and the Tight photon identification (T) define the signal region where both photons are Tight and isolated ($AA \equiv TI-TI$), with the background control regions defined by all possible combinations of requiring Loose' but not Tight

¹Calorimeter energy isolation, $E_{\text{T}}^{\text{iso}} > 5.5$ GeV (6.0 GeV), and track transverse momentum isolation, $p_{\text{T}}^{\text{iso}} > 2.2$ GeV (2.6 GeV), for 7 TeV (8 TeV).

identification (L') and/or inverting the isolation cut. The notation used to describe the various photon classifications is summarized in Table 7.1. In each event the two photon

Table 7.1: Notations and descriptions used for the photon candidate categorization.

notation	Description
T	Loose' photon candidate is identified as Tight
L'	Loose' photon candidate is not identified as Tight
I	photon candidate passes the isolation criterion
NI	photon candidate fails the isolation criterion
$A \equiv TI$	Tight identified photon candidate passes the isolation criteria
$B \equiv TNI$	Tight identified photon candidate fails the isolation criterion
$C \equiv L'I$	Loose' photon candidate is not identified as Tight and passes the isolation criterion
$D \equiv L'NI$	Loose' photon candidate is not identified as Tight and fails the isolation criterion

candidates are simultaneously classified, always distinguishing between the leading and sub-leading candidate. This defines $4 \times 4 = 16$ possible combinations, hence there are 16 equations for the diphoton background composition, one for each control region. This categorization is illustrated in the isolation plane of the diphoton candidates in Figure 7.1, sorted first by the leading photon candidate (x-axis) and then by the sub-leading photon candidate (y-axis). The event yields in each control region are extracted from data and denoted $N_{i,j}$, with $i, j = A, B, C,$ or D .

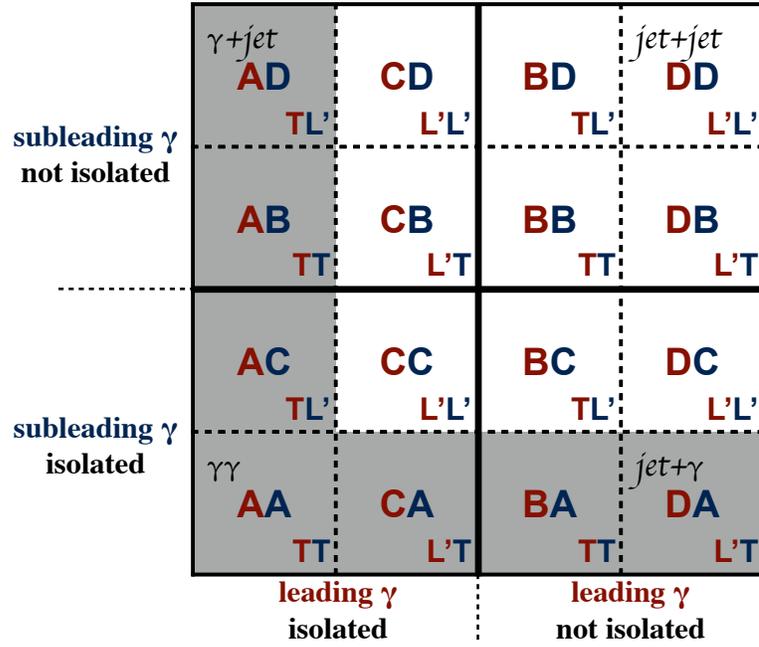


Figure 7.1: Illustration of the 1 signal and 15 background control regions of the 2x2D sideband method, with the leading (sub-leading) photon candidate in red (blue). All regions are enriched with various compositions of signal and background contributions ($\gamma\gamma$, γ +jet, jet+ γ , and jet+jet). Regions enriched with a particular background type are indicated by the text in the upper left corner of the ‘cells’ representing each region. Text in lower right of each cell gives each photon candidates identification. See the text for the meaning of the grey-shaded regions.

The solution of the system of equations requires knowledge of the photon efficiencies ($\epsilon_{1,2}$) and jet fake rates ($f_{1,2}$), both of which rely on the signal photon leakage fractions (s_{Ai} and s_{iA}). The leakage fractions are the fractional number of leading and sub-leading signal photons that ‘leak’ out of the signal region into the background control regions. The signal leakages are obtained from studies on simulated $\gamma\gamma$ background samples ($s_{Ai} = N_{Ai}/N_{AA}$ and $s_{iA} = N_{iA}/N_{A,A}$). The photon efficiencies are defined as the efficiency for a leading or sub-leading photon to pass the Tight identification and the isolation criteria. The photon efficiencies will later be obtained solely from the signal leakage fractions. The jet fake rates

are defined as the efficiency for a not Tight identified jet to pass the isolation selection and are parameterized in terms of the number of extracted data events in the control regions and also by the signal leakage fractions. In order to simplify the system of equations the following assumptions are made: (i) jet fake rates are the same for γ +jet, jet+ γ , and jet+jet events, (ii) there are no correlations between jets (photons) in jet+jet ($\gamma\gamma$) events. By using these assumptions the full information is truncated into information contained in the 7 shaded regions in Figure 7.1, which represent the regions where either one or both photon candidates are identified as Tight. The 7 regions representing the truncated information are folded into two 2D sidebands, one for the leading p_T photon candidate and one for the sub-leading p_T candidate. An illustration of the two 2D sidebands is given in Figure 7.2. The background composition in the signal region (AA) is obtained by simultaneously solving the system of equations contained in the two 2D sidebands, the detailed solution is now presented.

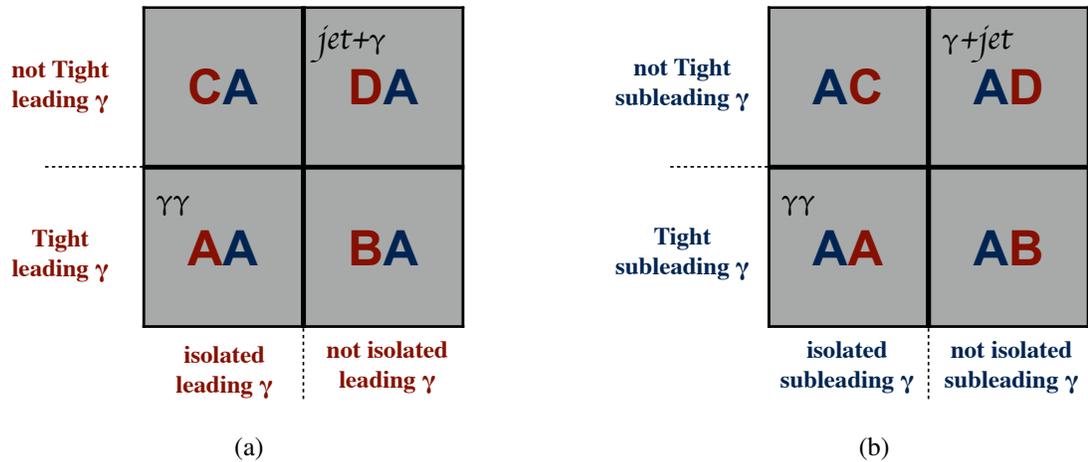


Figure 7.2: (a) and (b) illustrate the 2D sidebands corresponding to the leading and sub-leading photon candidates, respectively.

The number of true diphoton events in the signal region (AA) can now be extracted using the information contained in the two 2D sidebands corresponding to the leading and sub-leading p_T photons, respectively. First, consider the case where signal leakages are neglected ($s_{Ai} = s_{iA} = 0$), then the number of events where the leading (sub-leading) candidate is a prompt photon can be written as:

$$\begin{aligned} N_{AA}^{\gamma_1} &= N_{AA} - N_{AA}^{\text{bkg}} \\ N_{AA}^{\gamma_2} &= N_{AA} - N_{AA}^{\text{bkg}}, \end{aligned} \quad (7.1)$$

where $N_{AA}^{\gamma_1}$ ($N_{AA}^{\gamma_2}$) is the number of true leading (sub-leading) signal photon events in the signal region, and N_{AA}^{bkg} is the number of background events in the signal region. At this stage one further assumption is considered: in both sidebands the ratio of isolated to non-isolated background candidates in the not Tight bin is equal to the same ratio in the Tight bin, *i.e.*

$$\frac{N_{BA}^{\text{bkg}}}{N_{AA}^{\text{bkg}}} = \frac{N_{DA}^{\text{bkg}}}{N_{CA}^{\text{bkg}}} \quad \text{and} \quad \frac{N_{AB}^{\text{bkg}}}{N_{AA}^{\text{bkg}}} = \frac{N_{AD}^{\text{bkg}}}{N_{AC}^{\text{bkg}}}. \quad (7.2)$$

Then Equations 7.1 for the signal yield in the AA region become:

$$\begin{aligned} N_{AA}^{\gamma_1} &= N_{AA} - N_{CA} \frac{N_{BA}}{N_{DA}} \\ N_{AA}^{\gamma_2} &= N_{AA} - N_{AC} \frac{N_{AB}}{N_{AD}}. \end{aligned} \quad (7.3)$$

The introduction of signal leakage fractions s_{Ai} and s_{iA} transforms the number of events in each background control region according to:

$$N_{iA} \rightarrow N_{iA} - s_{iA} N_{AA}^{\gamma_1} \quad \text{and} \quad N_{Ai} \rightarrow N_{Ai} - s_{Ai} N_{AA}^{\gamma_1}, \quad \text{where } i \in B, C, D. \quad (7.4)$$

Including the signal leakages in Equations 7.1 for the signal yields in the AA region gives:

$$\begin{aligned} N_{AA}^{\gamma_1} &= N_{AA} - (N_{CA} - s_{CA}N_{AA}^{\gamma_1}) \frac{N_{BA} - s_{BA}N_{AA}^{\gamma_1}}{N_{DA} - s_{DA}N_{AA}^{\gamma_1}} \\ N_{AA}^{\gamma_2} &= N_{AA} - (N_{AC} - s_{AC}N_{AA}^{\gamma_2}) \frac{N_{AB} - s_{AB}N_{AA}^{\gamma_2}}{N_{AD} - s_{AD}N_{AA}^{\gamma_2}}. \end{aligned} \quad (7.5)$$

Equation 7.5 provides the final definition for the number of events where the leading (sub-leading) candidate is a prompt photon, $N_{AA}^{\gamma_1}$ ($N_{AA}^{\gamma_2}$). Furthermore, the jet fake rates (f_1 , f_2) can be defined in terms of the number of events in the not Tight regions of the sidebands, the signal leakage fractions (s_{Ai} , s_{iA}), and the signal yields ($N_{AA}^{\gamma_1}$, $N_{AA}^{\gamma_2}$):

$$\begin{aligned} f_1 &= \frac{N_{CA}}{N_{CA} + N_{DA}} \rightarrow \text{signal leakage} \rightarrow \frac{f_1}{1 - f_1} = \frac{N_{CA} - s_{CA}N_{AA}^{\gamma_1}}{N_{DA} - s_{DA}N_{AA}^{\gamma_1}} \\ f_2 &= \frac{N_{AC}}{N_{AC} + N_{AD}} \rightarrow \text{signal leakage} \rightarrow \frac{f_2}{1 - f_2} = \frac{N_{AC} - s_{AC}N_{AA}^{\gamma_2}}{N_{AD} - s_{AD}N_{AA}^{\gamma_2}}. \end{aligned} \quad (7.6)$$

The efficiency, ϵ_1 (ϵ_2), for a Tight identified leading (sub-leading) signal photon to pass the isolation criteria is then expressed in terms of the signal leakage fractions:

$$\begin{aligned} \epsilon_1 &= \frac{1}{1 + s_{BA}} \\ \epsilon_2 &= \frac{1}{1 + s_{AB}}. \end{aligned} \quad (7.7)$$

Now that the fake rates and efficiencies have been defined in terms of the number of events in the sideband regions and the leakage fractions, the Tight-Tight (TT) region of the two 2D sidebands can be used to extract the background composition in the diphoton signal region ($AA \equiv TI-TI$). In the TT sample region each photon can pass or fail the isolation

cut, resulting in four subsamples: Isolated Isolated (I-I), Isolated Not-Isolated (I-NI), Not-Isolated Isolated (NI-I), Not-Isolated Not-Isolated (NI-NI). In the TT sample, the equations describing the sample composition can be written as the following matrix equation: .

$$\begin{bmatrix} N^{\text{TI-TI}} \\ N^{\text{TI-TNI}} \\ N^{\text{TNI-TI}} \\ N^{\text{TNI-TNI}} \end{bmatrix} = \begin{bmatrix} \epsilon_1\epsilon_2 & \epsilon_1f_2 & f_1\epsilon_2 & f_1f_2 \\ \epsilon_1(1-\epsilon_2) & \epsilon_1(1-f_2) & f_1(1-\epsilon_2) & f_1(1-f_2) \\ (1-\epsilon_1)\epsilon_2 & (1-\epsilon_1)f_2 & (1-f_1)\epsilon_2 & (1-f_1)f_2 \\ (1-\epsilon_1)(1-\epsilon_2) & (1-\epsilon_1)(1-f_2) & (1-f_1)(1-\epsilon_2) & (1-f_1)(1-f_2) \end{bmatrix} \begin{bmatrix} N_{\gamma\gamma}^{\text{TT}} \\ N_{\gamma+\text{jet}}^{\text{TT}} \\ N_{\text{jet}+\gamma}^{\text{TT}} \\ N_{\text{jet}+\text{jet}}^{\text{TT}} \end{bmatrix},$$

where $N^{\text{TI-TI}}$, $N^{\text{TNI-TI}}$, $N^{\text{TI-TNI}}$, and $N^{\text{TNI-TNI}}$, are the event yields in each of the four sub-regions defined by the isolation selection, $f_{1,2}$ are the jet fake rates, $\epsilon_{1,2}$ are the photon efficiencies, and $N_{\gamma\gamma}^{\text{TT}}$, $N_{\gamma+\text{jet}}^{\text{TT}}$, $N_{\text{jet}+\gamma}^{\text{TT}}$, and $N_{\text{jet}+\text{jet}}^{\text{TT}}$, are the yields for each event type in the TT sample. The above system of equations is non-linear and the solution is obtained through minimization. The minimization returns the fake rates and event yields. The solution in the signal region ($\text{AA} \equiv \text{TI-TI}$) is given by:

$$\begin{aligned} N^{\text{TI-TI}} &= N_{\gamma\gamma}^{\text{TI-TI}} + N_{\gamma+\text{jet}}^{\text{TI-TI}} + N_{\text{jet}+\gamma}^{\text{TI-TI}} + N_{\text{jet}+\text{jet}}^{\text{TI-TI}} \\ &= \epsilon_1\epsilon_2 N_{\gamma\gamma}^{\text{TT}} + \epsilon_1f_2 N_{\gamma+\text{jet}}^{\text{TT}} + f_1\epsilon_2 N_{\text{jet}+\gamma}^{\text{TT}} + f_1f_2 N_{\text{jet}+\text{jet}}^{\text{TT}}. \end{aligned} \quad (7.8)$$

The above procedure is used to obtain the background decomposition in the inclusive fiducial region for the 7 TeV and 8 TeV datasets. Additionally, for the 8 TeV dataset, the background decomposition is obtained in four fiducial regions characterized by the hadronic activity in an event, parameterized in terms of the number of reconstructed jets ($N_{\text{jet}} = 0$, $N_{\text{jet}} = 1$, $N_{\text{jet}} = 2$, and $N_{\text{jet}} \geq 3$). The two 2D sidebands are defined corresponding to

the object selection that defines the fiducial region and the system of equations is solved to obtain the fake rates and yields for the given fiducial region.

7.2 Systematic uncertainties

The following systematic uncertainties are considered:

- **Loose' definition:** Affects the definition of the background control regions and the background photon ID-isolation correlation. The default Loose' identification definition corresponds to relaxing four of the calorimeter shower profile variables with respect to the Tight identification (see Section 4.3.1). The impact of this definition on the method is investigated by relaxing 2, 3 or 5 of the shower profile variables instead of four. The variation introducing the maximal deviation from the nominal definition is taken as the systematic uncertainty.
- **Modelling of the isolation efficiencies:** The dependence of the background decomposition on the simulation modelling of the isolation efficiencies is evaluated by using PYTHIA8 instead of SHERPA to compute the leakage fractions. The difference between the yields obtained corresponding to the two efficiencies obtained from each event generator is assigned as the systematic uncertainty.
- **Isolation cut threshold:** The dependence on the thresholds used for the calorimeter and track based photon isolation selection is assessed by shifting the leading and sub-leading photon isolation energies up and down by 0.2 GeV. The difference between the obtained yield and the nominal yield is taken as a systematic uncertainty.

- Tight identification efficiencies: The uncertainty due to the photon Tight identification efficiency is also evaluated. Indeed, the effect of the correction factors that are applied to the simulated photon shower shape variables are assessed by comparing the yields obtained without the application of the correction factors to the nominal yields.
- Correlations between the isolation and identification variables: Correlations between the variables defining the isolation and Tight identification can result in different jet rejections between the Loose' and Tight regions. The effect of such correlations is evaluated by introducing pseudo-correlation factors for the leading and sub-leading photons, namely θ_1 and θ_2 , respectively. The introduction of the pseudo-correlation factors multiplies the right side of Equations 7.5 by θ_1 and θ_2 . These correlations are neglected in the nominal results ($\theta_1 = \theta_2 = 1$), and are varied between 0.8 and 1.2 to assess the impact on the method. The yields obtained corresponding to the up and down variations are compared to the nominal result and the difference is assigned as an uncertainty.

The method relies on the procedure of extrapolating the background composition obtained in the control regions to the signal region. If the photon efficiency estimates from simulation are overestimated (underestimated) the resulting $\gamma\gamma$ purity will be lower (higher) than the actual $\gamma\gamma$ purity. To insure that no biases were introduced due to the isolation efficiency estimates from MC, a data-driven study of the isolation efficiencies neglecting signal leakages was performed. The results of the study exhibit no signs of an introduced bias due to the simulation efficiency estimates. Furthermore, the generator portion of the isolation uncertainty is designed to cover such biases.

7.3 Results

This section presents the results of the background decomposition obtained using the $2 \times 2D$ Sideband Method. All plots correspond to the results obtained using the 8 TeV dataset as this was the focus of the author's work. The 7 TeV result is included in Table 7.4 for completeness.

The photon identification and isolation efficiencies obtained from the SHERPA simulation are shown as a function of $m_{\gamma\gamma}$ in Figures 7.3(a), (b), (c), (d), and (e) for the inclusive selection and in jet multiplicity bins, $N_{\text{jet}} = 0$, $N_{\text{jet}} = 1$, $N_{\text{jet}} = 2$, and $N_{\text{jet}} \geq 3$, respectively. The Tight identification efficiency is essentially constant versus $m_{\gamma\gamma}$ in all cases. The isolation efficiency is essentially constant across $m_{\gamma\gamma}$ in all cases, but tends to decrease as the number of jets increases.

Figures 7.4(a), (b), (c), (d), and (e) show the extracted $\gamma\gamma$, γ +jet (γ +jet \equiv γ +jet + jet+ γ), and jet+jet yields in the signal region as well as the fractional composition (purity) of each type of event as a function of $m_{\gamma\gamma}$, for the inclusive selection and in jet multiplicity bins, $N_{\text{jet}} = 0$, $N_{\text{jet}} = 1$, $N_{\text{jet}} = 2$, and $N_{\text{jet}} \geq 3$, respectively. The left y-axis is the number of events per bin and corresponds to the stacked histogram. The dashed lines represent the purity of each type of event and corresponds to the right y-axis.

Figure 7.5(a) shows the average photon identification and isolation efficiencies as a function of the number of hadronic jets. The Tight identification efficiency is essentially constant versus N_{jet} . The isolation efficiency tends to decrease as the number of jets increases, more notably for the leading photon candidate. Figure 7.5(b) gives the yields and fractional decomposition as a function of the number of hadronic jets. The y-axis labelings

plot style are the same as in Figure 7.4.

Table 7.2 summarizes the extracted background yields in the signal region obtained at 8 TeV and Table 7.3 gives the extracted fake rates. Table 7.4 gives the fractional background decomposition for each event type in the signal region, the 7 TeV results are included for completeness.

Table 7.2: Extracted background yields for each event type in the signal region, including statistical and systematic uncertainties.

8 TeV				
jet selection	fitted yield \pm stat. \pm syst.			
	$N_{\gamma\gamma}$	$N_{\gamma\text{-jet}}$	$N_{\text{jet-}\gamma}$	$N_{\text{jet-jet}}$
Inclusive	$73345 \pm 372^{+1852}_{-2000}$	$14059 \pm 157^{+1592}_{-1578}$	$4859 \pm 112^{+555}_{-270}$	$2332 \pm 51^{+249}_{-432}$
$N_{\text{jet}} = 0$	$45070 \pm 302^{+1002}_{-1251}$	$10398 \pm 131^{+996}_{-851}$	$2982 \pm 88^{+324}_{-115}$	$1820 \pm 35^{+202}_{-381}$
$N_{\text{jet}} = 1$	$19129 \pm 181^{+568}_{-442}$	$2637 \pm 68^{+344}_{-436}$	$1498 \pm 56^{+54}_{-133}$	$385 \pm 21^{+100}_{-86}$
$N_{\text{jet}} = 2$	$6636 \pm 103^{+152}_{-154}$	$794 \pm 37^{+108}_{-134}$	$362 \pm 27^{+75}_{-41}$	$74 \pm 9^{+3}_{-21}$
$N_{\text{jet}} \geq 3$	$2409 \pm 62^{+44}_{-63}$	$278 \pm 22^{+37}_{-35}$	$127 \pm 16^{+30}_{-20}$	$23 \pm 6^{+14}_{-3}$

Table 7.3: 8 TeV jet fake rates (f_1 , f_2) obtained from the minimization. The fake rates (Equation 7.6) are defined as the efficiency for a not Tight identified jet to pass the isolation selection.

8 TeV	f_1	f_2
inclusive	0.23	0.18
$N_{\text{jet}} = 0$	0.27	0.17
$N_{\text{jet}} = 1$	0.16	0.22
$N_{\text{jet}} = 2$	0.13	0.21
$N_{\text{jet}} \geq 3$	0.10	0.20

Table 7.4: Fractional decomposition of dominant backgrounds in the signal region, in %. Here $\gamma\text{-jet} \equiv \gamma\text{-jet} + \text{jet-}\gamma$.

	$\gamma\gamma$ [%]	$\gamma\text{-jet}$ [%]	jet-jet [%]
7 TeV [62]			
inclusive	84 ± 8	15 ± 8	1 ± 1
8 TeV			
inclusive	77 ± 2	20 ± 2	3 ± 1
$N_{\text{jet}} = 0$	75 ± 2	22 ± 2	3 ± 1
$N_{\text{jet}} = 1$	81 ± 2	18 ± 2	2 ± 1
$N_{\text{jet}} = 2$	85 ± 2	14 ± 2	1 ± 1
$N_{\text{jet}} \geq 3$	85 ± 2	14 ± 2	1 ± 1

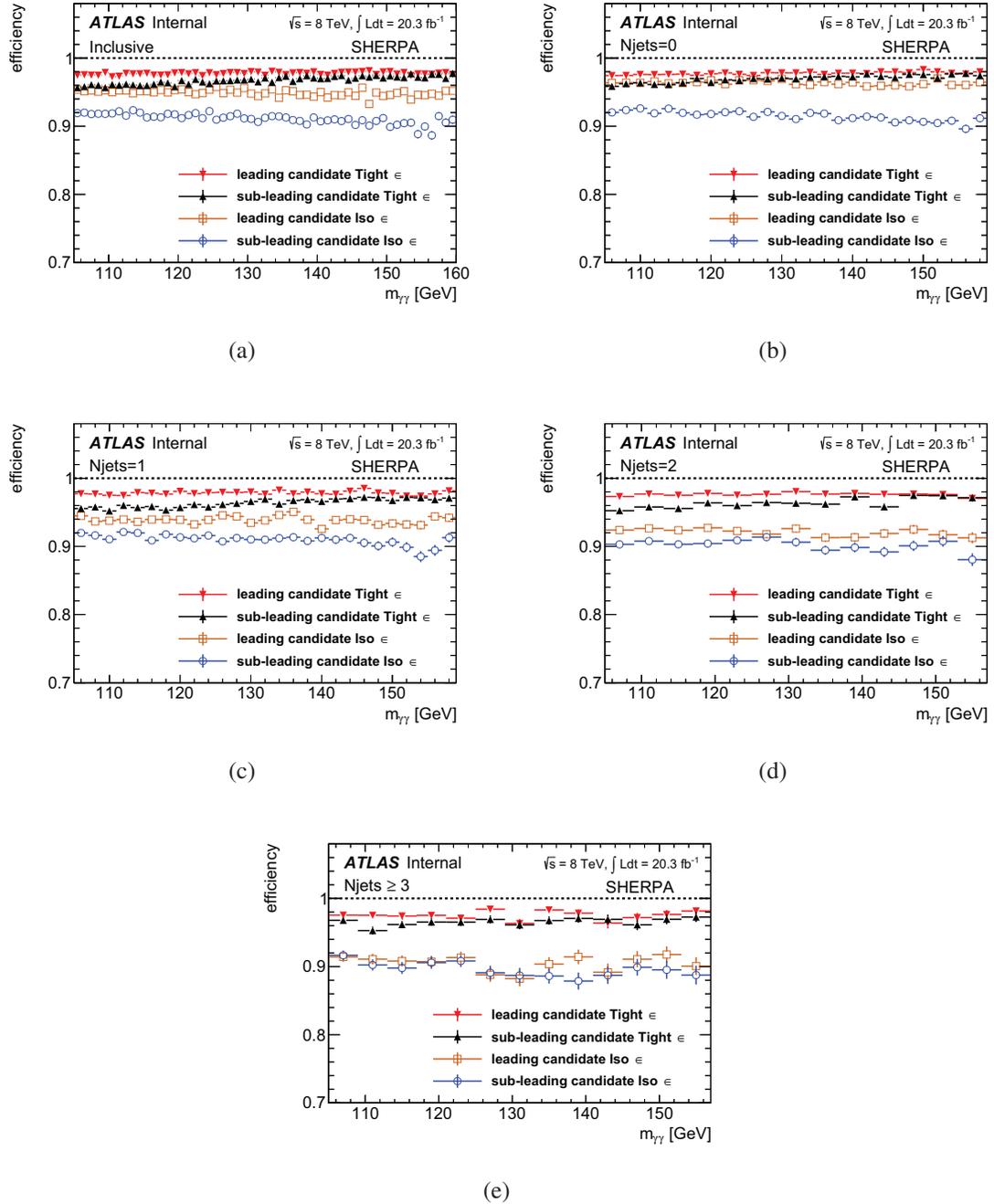


Figure 7.3: Photon efficiencies from the SHERPA simulation (Jet $p_T > 30$ GeV & $|y_{\text{jet}}| < 4.4$) as a function of $m_{\gamma\gamma}$ for (a) the inclusive selection, (b) $N_{\text{jet}} = 0$, (c) $N_{\text{jet}} = 1$, (d) $N_{\text{jet}} = 2$, and (e) $N_{\text{jet}} \geq 3$.

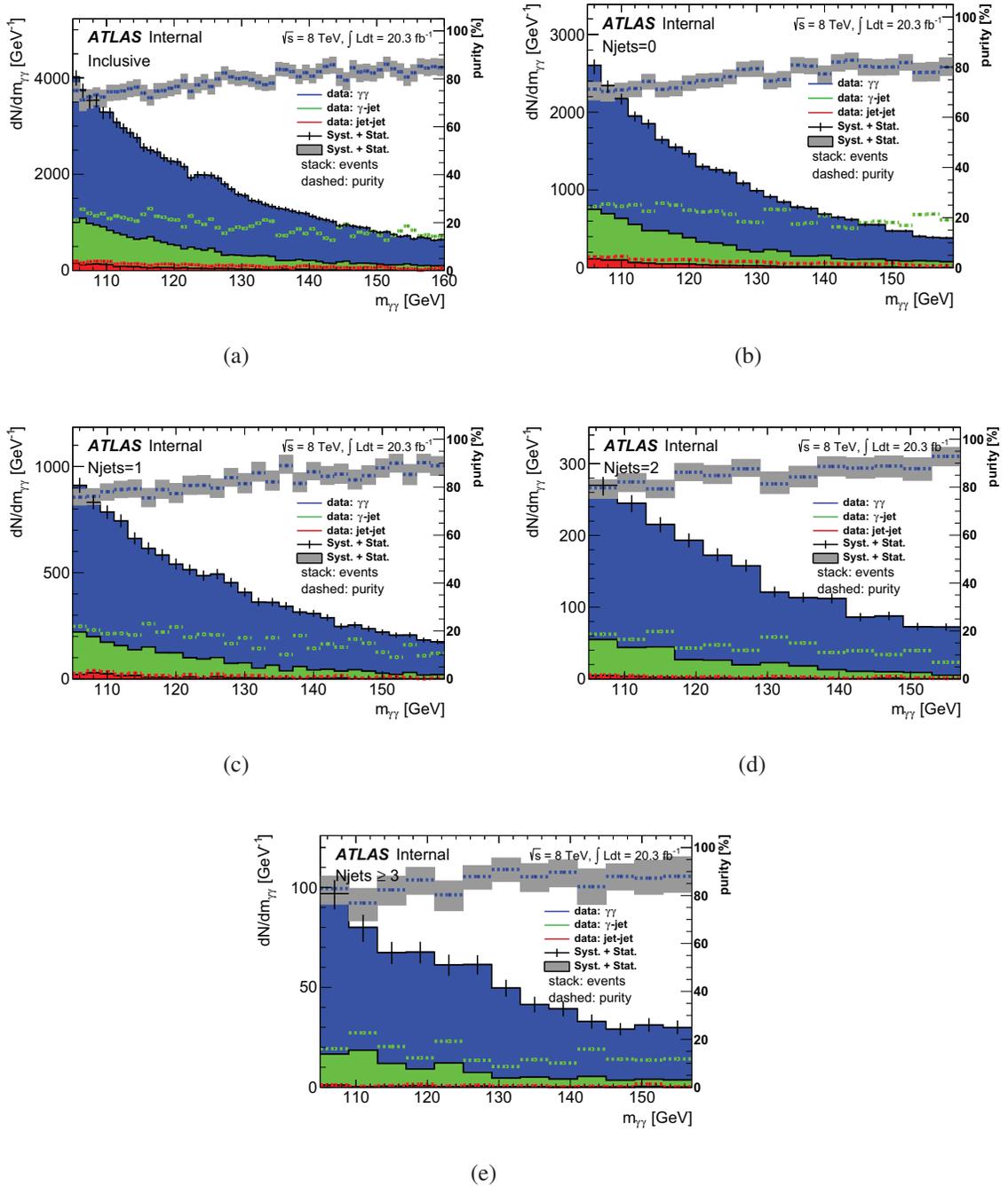
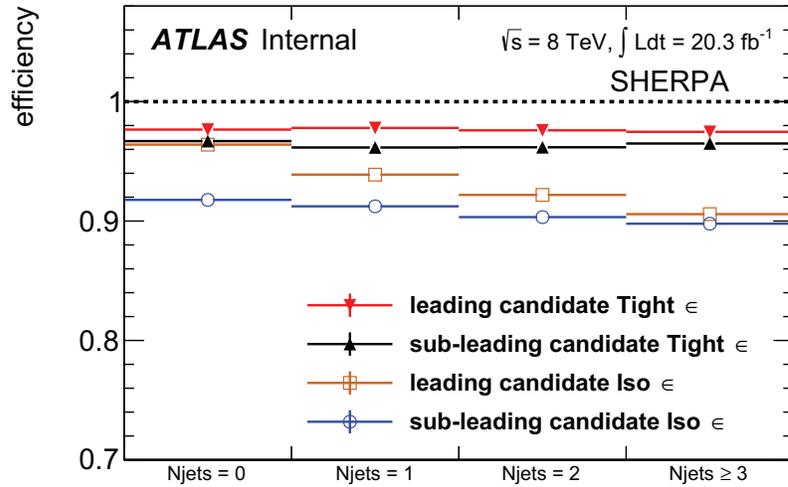
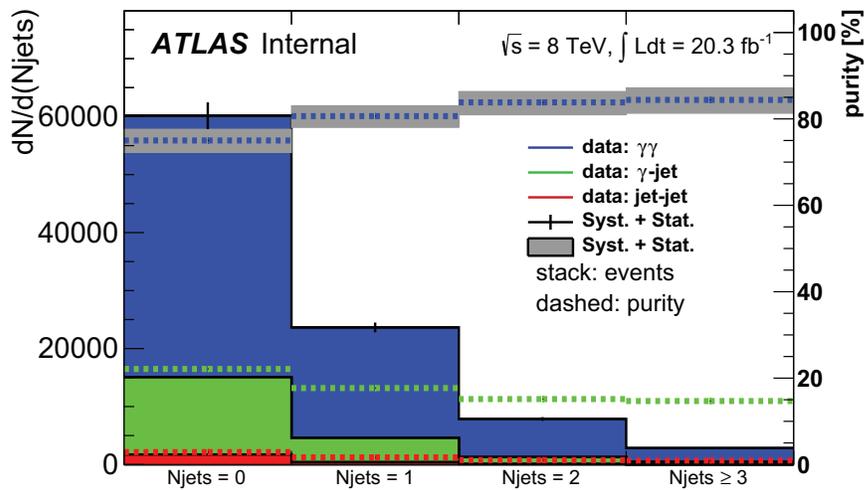


Figure 7.4: Data-driven fractional background decomposition (Jet $p_T > 30$ GeV & $|y_{\text{jet}}| < 4.4$) as a function of $m_{\gamma\gamma}$ for (a) the inclusive selection, (b) $N_{\text{jet}} = 0$, (c) $N_{\text{jet}} = 1$, (d) $N_{\text{jet}} = 2$, and (e) $N_{\text{jet}} \geq 3$. The left y-axis is the number of events per bin and corresponds to the stacked histogram. The dashed lines represent the purity of each type of event and corresponds to the right y-axis.



(a)



(b)

Figure 7.5: (a) gives the average photon efficiencies as a function of N_{jet} . (b) gives the background decomposition as a function of N_{jet} . The left vertical axis is the number of events per bin and the right vertical axis is the purity of each background process per bin.

7.4 Summary

The data-driven procedure used for extracting the background rates in the diphoton signal region for the $H \rightarrow \gamma\gamma$ analysis has been presented. The fractional decomposition of the $m_{\gamma\gamma}$ spectrum into the three dominant background contributions has been determined. For the inclusive fiducial region the percent contributions of the three background processes to the $m_{\gamma\gamma}$ spectrum are: $84 \pm 8\%$ ($77 \pm 2\%$), $15 \pm 8\%$ ($20 \pm 2\%$) γ +jet, and $1 \pm 1\%$ ($3 \pm 1\%$) jet+jet, for 7 TeV (8 TeV) data. For 8 TeV data, the percent contributions of the three background processes to the $m_{\gamma\gamma}$ spectrum in each of the four regions characterized by the number of hadronic jets are: $75 \pm 2\%$ $\gamma\gamma$, $22 \pm 2\%$ γ +jet, $3 \pm 1\%$ jet+jet, for $N_{\text{jet}} = 0$; $81 \pm 2\%$ $\gamma\gamma$, $18 \pm 2\%$ γ +jet, $2 \pm 1\%$ jet+jet, for $N_{\text{jet}} = 1$; $85 \pm 2\%$ $\gamma\gamma$, $14 \pm 2\%$ γ +jet, $1 \pm 1\%$ jet+jet, for $N_{\text{jet}} = 2$; and $85 \pm 2\%$ $\gamma\gamma$, $14 \pm 2\%$ γ +jet, and $1 \pm 1\%$ jet+jet, for $N_{\text{jet}} \geq 3$.

The 8 TeV background rates presented in this chapter were used in the published ATLAS result reporting the first ever measurements of Higgs boson cross sections in the diphoton decay channel [63]. All figures are included in the publication's internal supporting documentation [64].

Chapter 8

Higgs boson signal extraction in the

$H \rightarrow \gamma\gamma$ channel

This thesis focuses on cross section measurement of fiducial of the Higgs boson in the $H \rightarrow \gamma\gamma$ decay channel using data collected at 7 TeV and 8 TeV by the ATLAS detector. Differential cross sections are also studied using the enlarged 8 TeV dataset.

This chapter provides a detailed description of the analysis procedure used for extracting the Higgs boson signal yields in the $H \rightarrow \gamma\gamma$ decay channel. The following sections will focus primarily on the analysis details of the signal extraction for the 7 TeV inclusive fiducial cross section measurement, as this was the author's principal contribution. Additionally, details and results will be presented corresponding to the 8 TeV inclusive fiducial measurement and also in regions characterized by the level of hadronic jet activity (N_{jet}). The N_{jet} results directly reflect the author's contributions in the areas of non-resonant diphoton background decomposition (see Chapter 7) and hadronic jet calibration (see Chapter 5). Where

appropriate, additional details relevant to the measurement of differential cross sections will also be provided for completeness. The final cross section measurements will be presented in Chapter 10.

Fiducial cross sections are defined as: the measured cross section in a given fiducial region. Differential cross sections are defined as: the measured cross section in a bin of a differential distribution. The definition of the different fiducial regions are given in Section 6.6. In order to calculate the cross sections in each fiducial region or bin of a differential distribution, the number of Higgs boson signal events in each region or bin must be measured. The Higgs boson signal is extracted, and the background contribution subtracted via a simultaneous signal plus background (s+b) fit to the diphoton invariant mass ($m_{\gamma\gamma}$) spectrum in each fiducial region (or bin of a differential distribution).

This chapter is organized as follows. First, a description of the signal and non-resonant background modelling of the $m_{\gamma\gamma}$ spectrum is provided in Section 8.1. Then, the details of the signal extraction procedure are given in Section 8.2, and a discussion of the various sources of uncertainty on the extraction procedure is given in Section 8.3. A brief discussion on the choice of the Higgs boson mass is presented in Section 8.4. Finally, a summary of the extracted signal yields obtained from the fitting procedure is given in Section 8.5.

8.1 Background and signal modelling of $m_{\gamma\gamma}$

The Higgs boson signal is extracted by performing a simultaneous s+b fit to the $m_{\gamma\gamma}$ spectrum. The fitting function is parameterized in terms of signal and background probability density functions (PDFs) that describe the $m_{\gamma\gamma}$ spectrum. Parameters that describe

the signal and background PDFs that are not being measured or constrained by the analysis are referred to as “nuisance parameters”. Systematic uncertainties that affect the $m_{\gamma\gamma}$ signal shape are included as additional nuisance parameters. These additional nuisance parameters are constrained to their measured values using either normal or log-normal functions.

8.1.1 Non-resonant background modelling of $m_{\gamma\gamma}$

The non-resonant background contributions to the $m_{\gamma\gamma}$ spectrum are parameterized using an analytic PDF. Generally, the functional form of the background parameterization is chosen to be either an exponential or the exponential of a second order polynomial.

Fitting the background can potentially overestimate or underestimate the actual background contribution in the signal region, resulting in a bias on the extracted signal yield. It is therefore highly important to choose a parameterization that minimizes any potential bias due to the background modelling. Biases in the measured signal yield due to the choice of background parameterization are referred to as ‘spurious signal’. A spurious signal is the extracted signal resulting from fitting a pure background prediction, where the expected signal yield is zero. Therefore, spurious signal provides a direct measure of the bias on the signal extraction due to the background modelling. The non-resonant background parameterizations were chosen based on spurious signal studies using high statistics simulated background samples, as described in the next section.

8.1.1.1 Choice of non-resonant background $m_{\gamma\gamma}$ parametrization

The choice of background function used in the signal extraction is a trade-off between the statistical uncertainty (additional degrees of freedom result in a larger statistical uncertainty),

and the systematic uncertainty (additional degrees of freedom result in a better described background shape and smaller spurious signal). The spurious signal is studied using high-statistics diphoton background simulated samples, generated using SHERPA and PYTHIA. The total background prediction is obtained by combining the predictions for the dominant background processes ($\gamma\gamma$, γ -jet and jet-jet). The combination proceeds by scaling each background process based on the relative background fractions obtained from data-driven estimates using the 2x2D Sideband Method as detailed in Chapter 7. The spurious signal corresponding to each potential background model is determined by performing a signal plus background fit to the predicted pure background samples. The same signal model is used in both the spurious signal studies and in extracting the signal yield in the data. When fitting a pure background distribution, the signal yield from a s+b fit should in principle always be profiled to zero. Therefore fitting a spurious signal directly measures the signal bias due to the background modelling. Starting from the functional forms discussed in the previous section, the background parameterization is chosen as the one that results in the smallest spurious signal.

To illustrate the use of the data-driven background fractions derived in Chapter 7 in constructing the simulated background distributions used in the spurious signal studies Figure 8.1 shows the 8 TeV simulated background samples for the inclusive selection and also in bins of N_{jet} . The contributions from the dominant backgrounds, after scaling to the data-driven background fractions, are indicated by the different colours. The $H \rightarrow \gamma\gamma$ data is also included in the plots for comparison with the simulated background modelling. Note that the Higgs boson signal is included in the data spectrum. The plots exhibit a good agreement between data and the simulated samples in the regions that do not contain

the Higgs boson signal. The functional form of the background PDF that resulted in the smallest spurious signal for the 7 TeV and 8 TeV inclusive selection and also for the 8 TeV selection in bins of N_{jet} was the exponential of a second order polynomial.

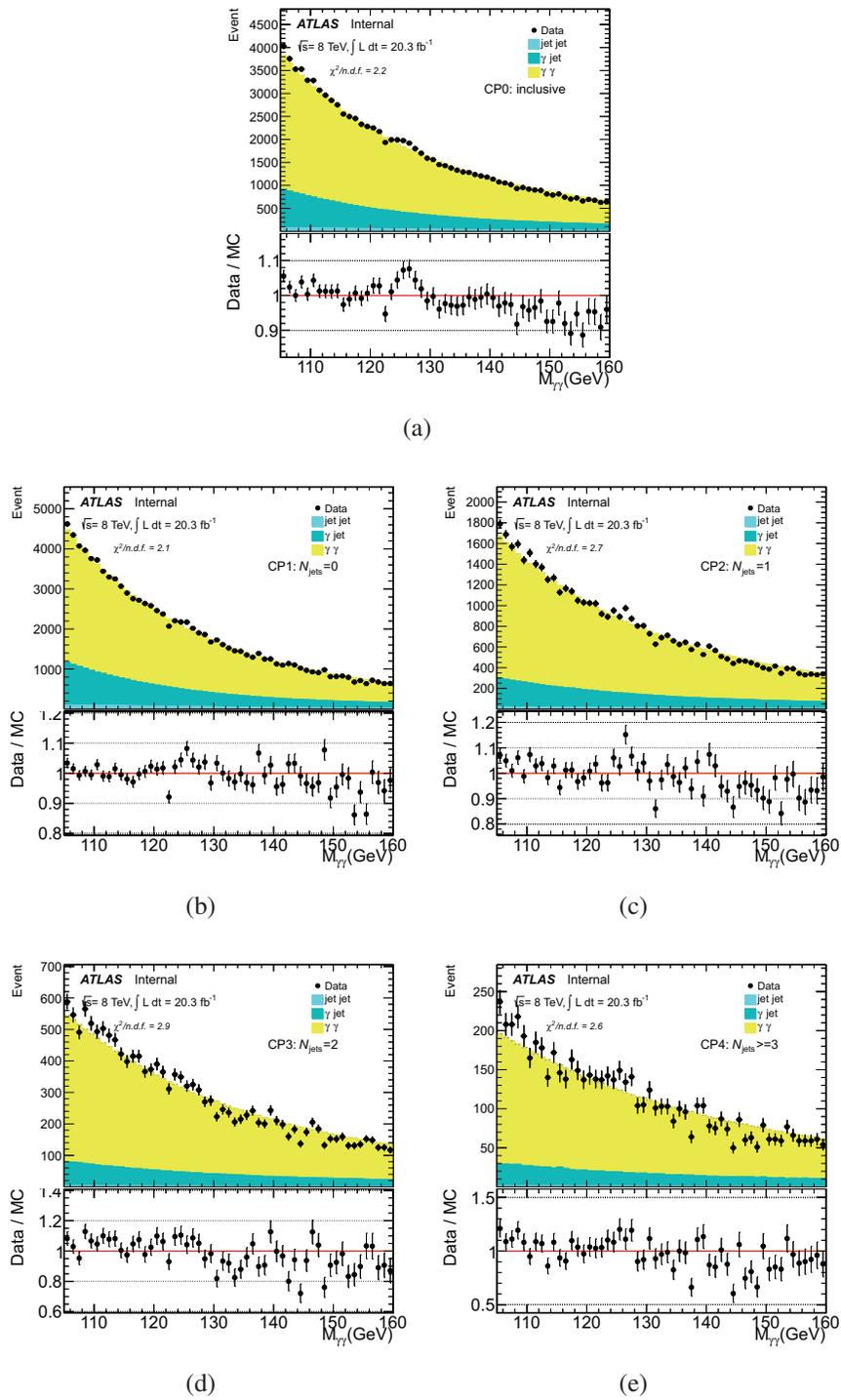


Figure 8.1: Comparison of the diphoton mass spectra between data and simulated samples for the inclusive selection (CP0) and in the different N_{jet} bins: $N_{\text{jet}} = 0$ (label: CP1), $N_{\text{jet}} = 1$ (label: CP2), $N_{\text{jet}} = 2$ (label: CP3), and $N_{\text{jet}} \geq 3$ (label: CP4) [64].

8.1.2 Signal modelling of $m_{\gamma\gamma}$

The signal function used to model a resonance is usually a Breit-Wigner convolved with a Gaussian distribution. For the Higgs boson resonance, a similar method can still be used. However, the Higgs boson width is predicted to be around 4 MeV to 5 MeV, which is very small compared to resolution of the ATLAS detector. In this case, the width of the Breit-Wigner distribution would be a very small number and the Breit-Wigner distribution would not contribute much to the signal determination. Instead, the Higgs boson signal portion of the reconstructed $m_{\gamma\gamma}$ distribution is modelled as the sum of a Crystal Ball function, $\mathcal{C}(m_{\gamma\gamma}; \mu_{CB}, \sigma_{CB}, \alpha_{CB}, n_{CB})$, and a wide Gaussian function, $\mathcal{G}(m_{\gamma\gamma}; \mu_G, \sigma_G)$.

The Crystal Ball function consists of a Gaussian core portion and a power-law low-end tail, below a certain threshold defined by the parameter α_{CB} . It is defined as:

$$\mathcal{C}(m_{\gamma\gamma}; \mu_{CB}, \sigma_{CB}, \alpha_{CB}, n_{CB}) = N \cdot \begin{cases} \exp(-t^2/2) & \text{if } t > -\alpha_{CB} \\ A \cdot (B - t)^{-n_{CB}} & \text{otherwise} \end{cases},$$

with :

$$\begin{aligned} t &= \frac{(m_{\gamma\gamma} - \mu_{CB})}{\sigma_{CB}}, \\ A &= \left(\frac{n_{CB}}{\alpha_{CB}}\right)^{n_{CB}} \cdot e^{-\alpha_{CB}^2/2}, \\ B &= \frac{n_{CB}}{\alpha_{CB}} - \alpha_{CB}, \end{aligned} \tag{8.1}$$

where N is a normalization parameter, μ_{CB} is the peak of the Crystal Ball's Gaussian core, σ_{CB} represents the resolution of the Crystal Ball's Gaussian core, and n_{CB} and α_{CB} parameterize the non-Gaussian tail of the Crystal Ball function. The value of n_{CB} is set to

10, and N is allowed to float in the s+b fit. The majority of signal events are captured by the narrow Gaussian core and asymmetric low side of the Crystal Ball function (\mathcal{C}). The wide Gaussian distribution (\mathcal{G}) accounts for outliers not captured by the Crystal Ball function. The full signal probability density function (PDF) is parameterized as:

$$\begin{aligned} \mathcal{S}(x = m_{\gamma\gamma}; \boldsymbol{\theta}^{\text{sig}}, \mu_{CB}) = \\ f_{CB} \mathcal{C}(m_{\gamma\gamma}; \mu_{CB}, \sigma_{CB}, \alpha_{CB}, n_{CB}) + (1 - f_{CB}) \mathcal{G}(m_{\gamma\gamma}; \mu_G, \sigma_G), \end{aligned} \quad (8.2)$$

where f_{CB} parameterizes the fraction of the total PDF that is represented by the Crystal Ball function, and $\boldsymbol{\theta}^{\text{sig}}$ is a vector of nuisance parameters. The width of the wide Gaussian function is defined through its proportional relation to the width of the Crystal Ball function:

$$\sigma_G = \kappa_G \cdot \sigma_{CB}. \quad (8.3)$$

The full set of nuisance parameters that define the signal shape are:

$$\boldsymbol{\theta}^{\text{sig}} = (f_{CB}, \mu_{CB}, \sigma_{CB}, \alpha_{CB}, n_{CB}, \mu_G, \kappa_G). \quad (8.4)$$

The values of the shape parameters are determined by fitting the $m_{\gamma\gamma}$ distribution from simulated $H \rightarrow \gamma\gamma$ decays at a specified Higgs boson mass. The Crystal Ball normalization parameter is left as a floating parameter in the s+b fit. The signal shapes are fitted over the range -20 GeV to +15 GeV centred on the Higgs boson mass corresponding to the distribution of simulated $H \rightarrow \gamma\gamma$ decays. The restriction of the fit range surrounding each Higgs boson mass removes a tiny portion of events in the high and low mass tails of the

distributions which distorted the fitted signal shapes, particularly in regions with one or more jets.

8.1.2.1 Global resolution model

The Higgs boson mass that is used in the signal extraction is not necessarily the same as the Higgs boson mass corresponding to a simulated sample, therefore, the mass dependence of the nuisance parameters defining the signal shape must be quantified. A linear global resolution model is adopted in order to model the mass dependence of the signal shape nuisance parameters given in Equation 8.4:

$$\begin{aligned}
 \mu_{CB} &= m_H + \mu_{CB \text{ offset}} + (m_H - 125) \mu_{CB \text{ slope}} , \\
 \mu_G &= m_H + \mu_{G \text{ offset}} + (m_H - 125) \mu_{G \text{ slope}} , \\
 \sigma_{CB} &= \sigma_{\text{offset}} + (m_H - 125) \sigma_{\text{slope}} .
 \end{aligned} \tag{8.5}$$

The values of $\mu_{CB \text{ offset}}$, $\mu_{CB \text{ slope}}$, σ_{offset} , σ_{slope} , $\mu_{G \text{ offset}}$, $\mu_{G \text{ slope}}$, κ_G , and f_{CB} are determined by simultaneously fitting multiple $m_{\gamma\gamma}$ distributions from simulated $H \rightarrow \gamma\gamma$ decays, with each distribution generated for a specific Higgs boson mass. The available simulated $H \rightarrow \gamma\gamma$ decay samples for the various Higgs boson production processes cover a mass range of $m_H \in [100 \text{ GeV}, 160 \text{ GeV}]$ in steps of 5 GeV. The production processes corresponding to each mass point are combined assuming the Standard Model composition. The Gaussian signal shape parameters are allowed to vary independently of the Crystal Ball signal shape parameters. Individual signal parameterizations are obtained in each fiducial region and in each bin of a differential distribution.

Figure 8.2 shows the results of the global resolution model fit to the ten invariant mass distributions from simulated $H \rightarrow \gamma\gamma$ decays for the (a) 7 TeV and (b) 8 TeV inclusive selection. Figure 8.3 shows the results of the global resolution model fit in each of the N_{jet} measurement regions. In both figures the functional form is given in Equation 8.2, where the parameters are constrained as a function of $m_{\gamma\gamma}$ according to Equation 8.5. Table 8.1 presents the values for the nuisance parameters that define the 7 TeV signal shape obtained from the global resolution model fit. The mean value of the nuisance parameters defining the signal shape (Equation. 8.4) are fixed to the values obtained from the global resolution fits. Uncertainties on the nuisance parameters due to the modelling of the simulated $m_{\gamma\gamma}$ mass resolution are discussed in the next section.

Table 8.1: Signal PDF nuisance parameters obtained from the global resolution model corresponding to the 7 TeV inclusive fiducial region, for a Higgs boson mass of 125.36 GeV.

inclusive	$\mu_{\text{CB offset}}$	$\mu_{\text{CB slope}}$	σ_{offset}	σ_{slope}	$\mu_{\text{G offset}}$	$\mu_{\text{G slope}}$	κ_{G}	f_{CB}	α_{CB}
7 TeV	-0.0168	-0.0003	1.3707	0.0057	1.0560	-0.0011	1.7216	0.9192	1.3615

8.1.2.2 Signal parameter uncertainties

Systematic uncertainties on the modelling of the $m_{\gamma\gamma}$ distribution in the Monte Carlo simulation are accounted for in the fit through the introduction of additional nuisance parameters. Systematic uncertainties associated with the photon energy scale and resolution, and also the bias on the diphoton mass scale due to the background modelling, modify the

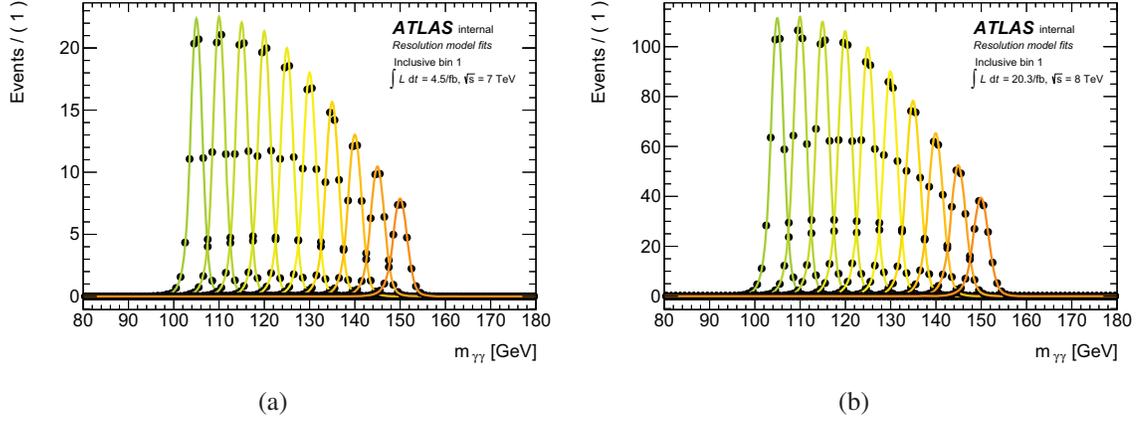


Figure 8.2: Results from simultaneous fits to the $m_{\gamma\gamma}$ spectrum over the ten “mass points” spanning Higgs boson masses from 105 GeV to 150 GeV at (a) 7 TeV and (b) 8 TeV. The functional form is given in Equation 8.2 where the parameters are constrained as a function of $m_{\gamma\gamma}$ according to Equation 8.5.

means and widths of the signal parameterization in the following way:

$$\begin{aligned}
 \mu_{CB} &\rightarrow \mu_{CB} \cdot \Theta_{\text{EnScale}} \cdot \Theta_{\text{BkgBias}} , \\
 \mu_{GA} &\rightarrow \mu_{GA} \cdot \Theta_{\text{EnScale}} \cdot \Theta_{\text{BkgBias}} , \\
 \sigma_{CB} &\rightarrow \sigma_{CB} \cdot \Theta_{\text{EnRes}} .
 \end{aligned}
 \tag{8.6}$$

Θ_{EnScale} , Θ_{EnRes} , and Θ_{BkgBias} are the correction terms accounting for each uncertainty source.

Each correction term is constructed from the product of the individual correction terms corresponding to each systematic uncertainty component in the three classifications (EnScale,

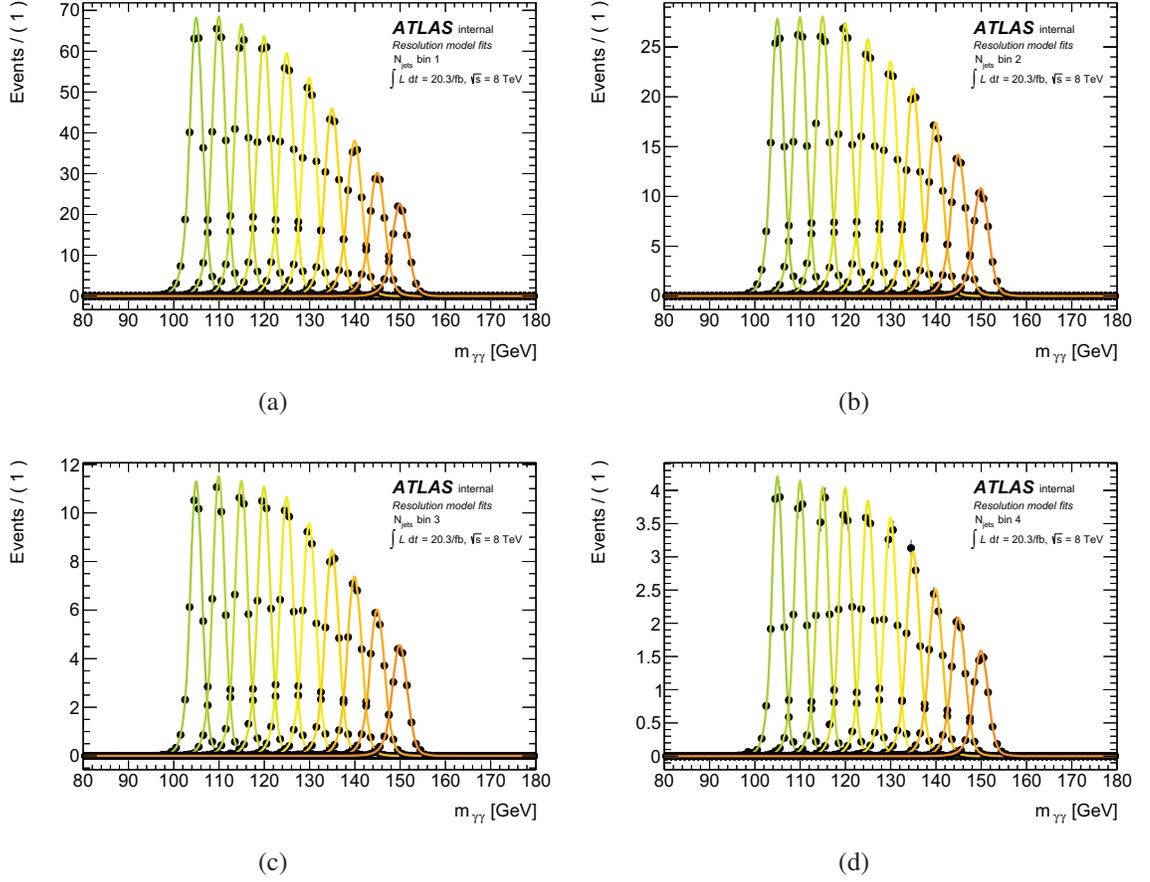


Figure 8.3: Results from simultaneous fits to the $m_{\gamma\gamma}$ spectrum over the ten “mass points” spanning Higgs boson masses from 105 GeV to 150 GeV in at 8 TeV. The functional form is given in Equation 8.2 where the parameters are constrained as a function of $m_{\gamma\gamma}$ according to Equation 8.5. The fit is performed separately for each N_{jet} bin.

EnRes, BkgBias):

$$\begin{aligned}
 \Theta_{\text{EnScale}} &= \prod_i^{N_{\text{EnScale}}} \Theta_i^{\text{EnScale}}, \\
 \Theta_{\text{BkgBias}} &= \prod_i^{N_{\text{BkgBias}}} \Theta_i^{\text{BkgBias}}, \\
 \Theta_{\text{EnRes}} &= \prod_i^{N_{\text{EnRes}}} \Theta_i^{\text{EnRes}}.
 \end{aligned} \tag{8.7}$$

A simple Gaussian form is used for the correction terms corresponding to the uncertainties from the photon energy scale and the background modelling. A log-normal form is used instead of the simple Gaussian form for the correction term corresponding to the photon energy resolution uncertainties. The choice of a log-normal form prevents negative corrections to the width parameter. More explicitly:

$$\begin{aligned}\Theta_i^{\text{EnScale}} &= (1 + \delta_i^{\text{EnScale}} \theta_i^{\text{EnScale}}) , \\ \Theta_i^{\text{BkgBias}} &= (1 + \delta_i^{\text{BkgBias}} \theta_i^{\text{BkgBias}}) , \\ \Theta_i^{\text{EnRes}} &= \exp \left(\theta_i^{\text{EnRes}} \sqrt{\log(1 + (\delta_i^{\text{EnRes}})^2)} \right) ,\end{aligned}\tag{8.8}$$

where δ_i denotes the value of the various systematic uncertainties in each classification and the θ_i are free nuisance parameters corresponding to each uncertainty source. The value of each nuisance parameter, θ_i , is nominally 0 ± 1 . The nuisance parameters are constrained by Gaussian functions (G_k) of width 1.

There are 29 uncorrelated systematic uncertainties from the photon energy scale calibration and 4 systematic uncertainties from the photon energy resolution. The four energy resolution uncertainties are added in quadrature and introduced in the form of one nuisance parameter. The non-dominant energy scale uncertainties are added in quadrature to form a single nuisance parameter, while the dominant energy scale uncertainties are introduced as individual nuisance parameters. This results in eight (five) energy scale nuisance parameters introduced to the fit for the signal extraction at 7 TeV (8 TeV). Further details of the energy scale and resolution systematics are documented in Section 8.3. There is a single systematic uncertainty on the mass scale due to the choice of background PDF.

8.2 Higgs boson signal extraction via a simultaneous s+b fit to the $m_{\gamma\gamma}$ spectrum

A simultaneous un-binned extended maximum likelihood fit to the inclusive $m_{\gamma\gamma}$ distribution is used to extract the signal yields. The likelihood function, \mathcal{L} , of the data sample $\mathbf{x}_n = \mathbf{m}_{\gamma\gamma}^n$ is given by:

$$\begin{aligned} \mathcal{L}(m_{\gamma\gamma}, \nu^{\text{sig}}, \nu^{\text{bkg}}, m_H) = & \prod_i \left\{ \frac{e^{-\nu_i}}{n_i!} \prod_j^{n_i} [\nu_i^{\text{sig}} \mathcal{S}_i(m_{\gamma\gamma}^j; \boldsymbol{\theta}^{\text{sig}}, m_H) + \nu_i^{\text{bkg}} \mathcal{B}_i(m_{\gamma\gamma}^j; \boldsymbol{\theta}^{\text{bkg}})] \right\} \\ & \times \prod_k G_k, \end{aligned} \quad (8.9)$$

where i labels the fiducial regions (or bins of a differential distribution) being simultaneously fitted, ν_i^{sig} (ν_i^{bkg}) are the extracted number of signal (background) events in data, $\nu_i = \nu_i^{\text{sig}} + \nu_i^{\text{bkg}}$ is the mean value of the underlying Poisson distribution of the n_i events (treated as an independent parameter of the fit), $m_{\gamma\gamma}^j$ is the diphoton invariant mass of event j , and $\mathcal{S}_i(m_{\gamma\gamma}^j; \boldsymbol{\theta}^{\text{sig}}, m_H)$ and $\mathcal{B}_i(m_{\gamma\gamma}^j; \boldsymbol{\theta}^{\text{bkg}})$ are the signal and background PDFs, respectively. The signal and background PDFs are both functions of $m_{\gamma\gamma}$ and a vector of nuisance parameters ($\boldsymbol{\theta}^{\text{sig}}, \boldsymbol{\theta}^{\text{bkg}}$) which are shared between all regions (or bins) being simultaneously fitted. The G_k are the Gaussian functions used to constrain the nuisance parameters of the normal and log-normal constraint terms described in the previous section, where the index k runs over the different uncertainty components that affect the signal shape and fitted peak position.

In the simultaneous fit of the $m_{\gamma\gamma}$ spectrum to extract the Higgs boson signal, three separate uncertainty sources on the signal extraction were considered: (i) the uncertainty

on the diphoton mass resolution due to the photon energy resolution; and two uncertainties on the peak position, (ii) the photon energy scale, and (iii) the bias due to the chosen background parameterization. The effect of including a spurious signal nuisance parameter can be completely factorized out of the likelihood function; therefore no nuisance parameter is assigned for the spurious signal. Instead, the uncertainty due to spurious signal is applied to the extracted yield, post-fit. Additional uncertainty sources that do not impact the signal shape or peak position are not included in the fit. Instead they are treated as uncertainties on the unfolding factor as will be described in Section 9.2. The uncertainty on the luminosity measurement is treated as an uncertainty on the measured cross sections.

All events contained in the inclusive fiducial region are included in the signal extraction for all other fiducial regions and observables. Events that are un-categorized are put into an additional bin (overflow) which is simultaneously fit along with the bin containing the categorized events belonging to a given fiducial region. For example, in the VBF-enhanced fiducial region all events that do not pass the VBF-enhanced selection but are present in the inclusive fiducial region are placed in the overflow bin which is simultaneously fit with the VBF-enhanced bin. The inclusion of the un-categorized events makes use of all the available information to help stabilize the fit and constrain the nuisance parameters of the signal PDF. The fitted signal yield is not constrained to be positive.

8.3 Systematic uncertainties on the signal extraction

The following systematic uncertainties impact the shape of the signal PDF and are included in the s+b fit to extract the signal in the form of additional nuisance parameters,

as described in Section 8.1.2.2.

8.3.1 Mass resolution uncertainties

Systematic uncertainties on the photon energy resolution affect the resolution of the diphoton invariant mass ($m_{\gamma\gamma}$). The uncertainties are categorized into the following four sources:

- Constant term: the energy of photons in the simulation is smeared to match the energy resolution measured in data using a sample containing $Z \rightarrow ee$ events. The uncertainty on extraction of the measured resolution is propagated to an uncertainty on the photon's energy resolution.
- Intrinsic calorimeter resolution: a 10 % uncertainty in the intrinsic calorimeter resolution is assigned as ± 10 % relative uncertainty on the single photon resolution.
- Material mis-modeling in the simulation.
- Electronic and pileup noise mis-modeling in the simulation.

The impact of these uncertainty sources on $m_{\gamma\gamma}$ is obtained by propagating the single photon energy resolution uncertainties to the diphoton mass.

A simplified correlation procedure is adopted in this analysis: the four invariant mass resolution systematic uncertainties are combined into a single uncertainty source by adding them in quadrature. This combined systematic uncertainty is introduced as a single nuisance parameter in the fit to extract the signal.

8.3.2 Mass energy scale uncertainties

There are 29 different sources of systematic uncertainty due to the photon energy scale; each one is propagated to $m_{\gamma\gamma}$ using the mass bias method [65]. In this method, the nominal photon energies are used to build the $m_{\gamma\gamma}$ distribution in a window of 120 GeV to 130 GeV. The mean (μ) and the RMS (R) of the $m_{\gamma\gamma}$ distribution are extracted. The mean and RMS are used to define a new mass window around the peak ($\mu - 1.5R \leq m_{\gamma\gamma} \leq \mu + 2R$). The mean found in the new window is considered as the nominal $m_{\gamma\gamma}$ peak position. The energies of the photon are shifted according to the photon energy scale uncertainty for each different source of systematic error and the same procedure is repeated. Each uncertainty source results in a shifted peak position. The percent difference between the nominal peak position and shifted peak position for each uncertainty source is taken as the systematic uncertainty. As for the resolution uncertainties, a simplified model was adopted where the non-dominant uncertainty sources are added in quadrature to form a single photon energy scale uncertainty source while the dominant uncertainties are kept separate. For the 7 TeV analysis this results in seven uncertainty components being kept separate, with the rest merged into a eighth. For the 8 TeV analysis the procedure resulted in four uncertainty components being kept separate, with the rest merged into a fifth. Each of the eight (five) uncertainties is treated as a single nuisance parameter in the signal extraction as described in Section 8.2.

8.3.3 Background modelling

To account for any potential uncertainty on the diphoton invariant mass scale due to the background parameterization, an uncertainty is assigned to the mean of the Crystal Ball and Gaussian PDFs used to model the signal. This uncertainty is calculated using pseudo-data generated by injecting the signal PDF into the high statistics simulation samples. The signal+background model is then fit to this combined distribution, and the difference between the injected mass and the fitted mass is taken as the bias on the mass scale. A scan is performed in the invariant mass range 119-137 GeV. The signal is injected in steps of 0.2 GeV over the $m_{\gamma\gamma}$ range of 119 GeV to 137 GeV and the bias evaluated for each step. The maximum absolute bias obtained over the mass range is assigned as the background mass bias uncertainty. This uncertainty was evaluated to be 0.2 % for the 7 TeV and 8 TeV inclusive fiducial region. The background mass bias was found to be 0.3 % in the $N_{\text{jet}} = 0$ bin, increasing to 0.5 % in the $N_{\text{jet}} \geq 3$ bin.

8.3.4 Summary of systematic uncertainties affecting the signal PDF

The naming and definition of each of the various uncertainty components that affect the shape of the signal PDF and are included in the fit is given in Table 8.2.

Table 8.2: Naming convention and definitions for the systematic uncertainty sources affecting the signal PDF and included in the fit.

uncertainty label	definition
BackgroundUncert_bin	The relative systematic uncertainty from the mass bias due to the choice of background model. The <code>_bin</code> denotes the uncertainty corresponding to a particular bin of a distribution. In the case of fitting the inclusive fiducial region <code>_bin0</code> refers to the bin representing the inclusive distribution.
MergedRes	The relative systematic uncertainty corresponding to the quadratic sum of the four mass resolution uncertainties.
MergedPES	The relative systematic uncertainty corresponding to the quadratic sum of the non-dominant mass energy scale uncertainties.
MatD1	Relative photon energy scale systematic uncertainty from material mis-modelling in the simulation.
Zee	Relative photon energy scale systematic uncertainty from the <i>in-situ</i> calibration using a data sample containing $Z \rightarrow ee$ events.
LeakageConv	Relative photon energy scale systematic uncertainty due to leakage of converted photons.
LeakageUnconv	Relative photon energy scale systematic uncertainty due to leakage of unconverted photons.
L2Gain	Relative photon energy scale systematic uncertainty due to the level-two LAr calorimeter gain.
LArCalib_bar	Relative photon energy scale systematic uncertainty due to LAr calibration in the barrel calorimeter.
LArElecUnconv_bar	Relative photon energy scale systematic uncertainty due to unconverted electron in the LAr barrel calorimeter.

8.3.5 Separation of statistical and systematic uncertainties on the yield

To separate the statistical and systematic uncertainties on the fitted signal yield, a pseudo-dataset is generated using the maximum likelihood estimation from the fit to data. This pseudo-dataset is then fit twice: once allowing the systematic uncertainties to float, and once where they are kept fixed to their nominal values. The statistical contribution to the total uncertainty is obtained from the quadratic difference of the uncertainties from the two fits.

8.4 Choice of the Higgs boson mass used in the signal extraction

In the above s+b fit used to extract the signal from data, the Higgs boson mass was fixed to the value measured using the full Run-I ATLAS dataset, combining the $H \rightarrow ZZ^* \rightarrow 4\ell$ and $H \rightarrow \gamma\gamma$ decay channels: $m_H = 125.36 \pm 0.37$ (stat.) ± 0.18 (syst.) GeV [66]. A more recent and precise Higgs boson mass was obtained using the combined ATLAS+CMS Run-I data [12]: $m_H = 125.09 \pm 0.21$ (stat.) ± 0.11 (syst.) GeV. This combined ATLAS+CMS Higgs boson mass will be used in future results. Additionally it should be noted that the Higgs boson mass obtained exclusively in the $H \rightarrow \gamma\gamma$ decay channel has been measured to be $124.30_{-0.60}^{+0.59}$ GeV at 7 TeV and $126.21_{-0.54}^{+0.54}$ GeV at 8 TeV [67]. Additional studies and cross-checks were performed to assess the effect of the choice of the mass on the extracted signal yields:

- Using a fixed Higgs boson mass, a scan over a range of diphoton masses (120 GeV

$< m_{\gamma\gamma} < 135$ GeV) was performed in steps of 0.2 GeV. The extracted signal yield as a function of the chosen mass was profiled from the scan.

- The Higgs boson signal yields are extracted using the ATLAS+CMS combined Higgs boson mass and compared to those obtained using the ATLAS $H \rightarrow ZZ^* \rightarrow 4\ell + H \rightarrow \gamma\gamma$ combined Higgs boson mass.

The results of these studies are presented in the next section.

8.5 Fit results

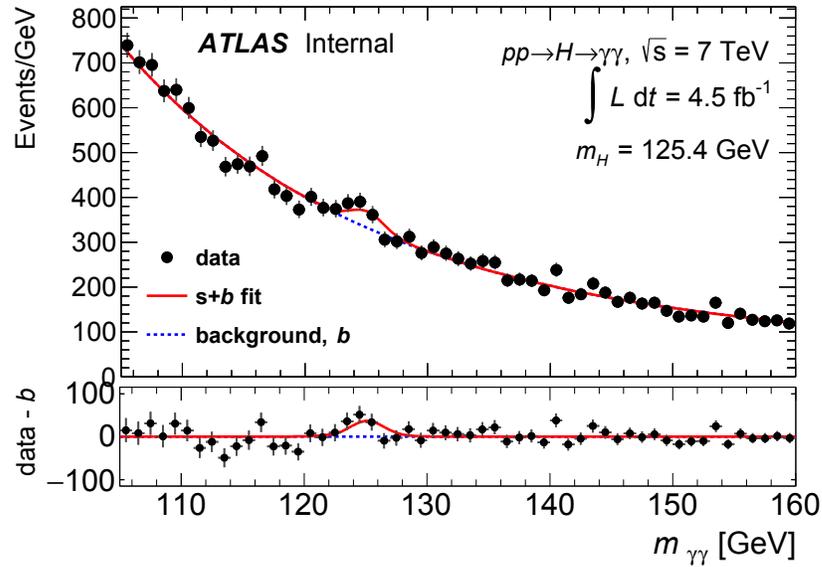
The s+b fit results presented in this section are obtained using a fixed Higgs boson mass corresponding to the final ATLAS Run-I $H \rightarrow ZZ^* \rightarrow 4\ell + H \rightarrow \gamma\gamma$ combined Higgs boson mass. The s+b fit to the inclusive $m_{\gamma\gamma}$ spectrum for 7 TeV and 8 TeV are shown in Figures 8.4(a) and (b), respectively. Figures 8.5(a), (b), (c), and (d) show the s+b fit $m_{\gamma\gamma}$ spectrum in bins of $N_{\text{jet}} = 0$, $N_{\text{jet}} = 1$, $N_{\text{jet}} = 2$, and $N_{\text{jet}} \geq 3$, respectively. All N_{jet} bins are fit simultaneously. Figures 8.6(a), (b), and (c) compare the fit yields in data to the expected number of reconstructed signal events from simulation for the inclusive selection corresponding to the 7 TeV and 8 TeV datasets, and as a function of N_{jet} for the 8 TeV dataset, respectively. Table 8.3 provides a summary of the total number of events in data, the extracted Higgs boson signal yield, and the expected number of events from simulation. Note that the number of extracted signal events in the fiducial region characterized by the inclusive hadronic jet activity ($N_{\text{jet}} \geq 1$, $N_{\text{jet}} \geq 2$) are obtained by combining the extracted signal yields obtained in the exclusive N_{jet} bins.

The $N_{\text{lepton}} \geq 1$ and $E_{\text{T}}^{\text{miss}} \geq 80$ GeV fiducial regions result in yields that are consistent with zero. In these cases the data are used to obtain limits on the fiducial cross section. Further details on the limit setting procedure can be found in Reference [63].

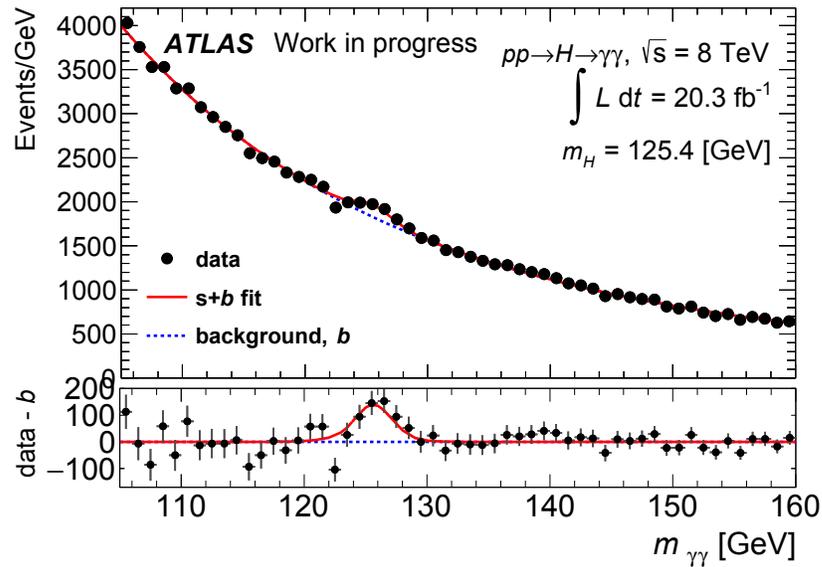
The nuisance parameters corresponding to the different detector uncertainty components were defined such that they should be standard normal variables. Figure 8.7 shows the distribution of the resulting nuisance parameters and their errors compared to the prior of a standard normal distribution, $N(0,1)$. This provides a measure of the impact (‘pull’) the nuisance parameter has on the signal shape. Additional details on the uncertainty sources are provided in Section 8.3. Deviations of the nuisance parameter central value from zero in the positive (negative) direction indicate that the fit would have preferred a higher (lower) Higgs boson mass, to accommodate the non-preferred mass choice the nuisance parameters pulled from their measured values to slightly alter the signal shape. As previously quoted the Higgs boson mass measured at 7 TeV (8 TeV) was lower (higher) than the ATLAS Run-I $H \rightarrow ZZ^* \rightarrow 4\ell + H \rightarrow \gamma\gamma$ combined. Therefore, the pulls on the nuisance parameters are negative (positive) for the fits to data recorded at 7 TeV (8 TeV).

Table 8.3: Presented in the table are the event yields in each fiducial region: the total number of events selected in data (N_{data}), the signal yield obtained from the s+b fit to data (ν_i^{sig}), and the expected signal yield from the MC simulation ($N_{\text{sim}}^{\text{sig}}$). The extracted signal yields (ν_i^{sig}) are always larger than the expected number of signal events ($N_{\text{sim}}^{\text{sig}}$), reflecting the fact that the measured signal strength is larger than the expected signal strength in the $H \rightarrow \gamma\gamma$ decay channel, though consistent with expectations for a Standard Model Higgs boson within the statistical precision of the measurements. The quoted uncertainty on the fitted yield represents both the statistical and systematic uncertainties on the signal extraction. The uncertainty on the expected yields takes into account both theoretical and experimental uncertainties.

fiducial region	N_{data}	ν_i^{sig}	$N_{\text{sim}}^{\text{sig}}$
7 TeV			
inclusive	17225	138 ± 52	74 ± 50
8 TeV [63]			
inclusive	94627	570 ± 130	403 ± 45
$N_{\text{jet}} \geq 1$	34293	308 ± 79	178^{+31}_{-26}
$N_{\text{jet}} \geq 2$	10699	141 ± 43	63 ± 11
$N_{\text{jet}} \geq 3$	2840	64 ± 22	17 ± 4
VBF-enhanced	334	24 ± 9	13 ± 2



(a)



(b)

Figure 8.4: Signal plus background fit to the inclusive $m_{\gamma\gamma}$ spectrum for data recorded at (a) $\sqrt{s} = 7 \text{ TeV}$ and (b) $\sqrt{s} = 8 \text{ TeV}$.

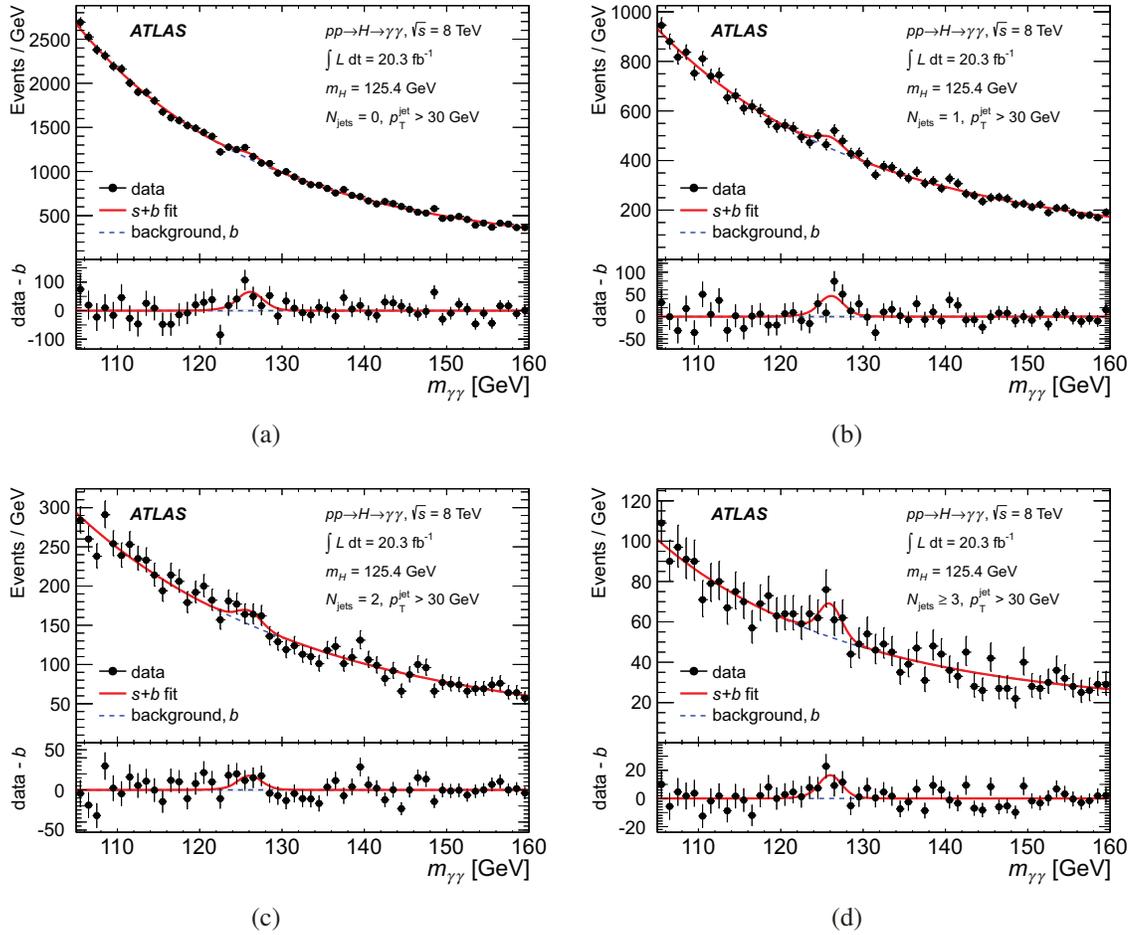


Figure 8.5: Signal plus background fit to the $m_{\gamma\gamma}$ spectrum corresponding to (a) $N_{\text{jet}} = 0$, (b) $N_{\text{jet}} = 1$, (c) $N_{\text{jet}} = 2$, and (d) $N_{\text{jet}} \geq 3$, for data recorded at $\sqrt{s} = 8$ TeV [63].

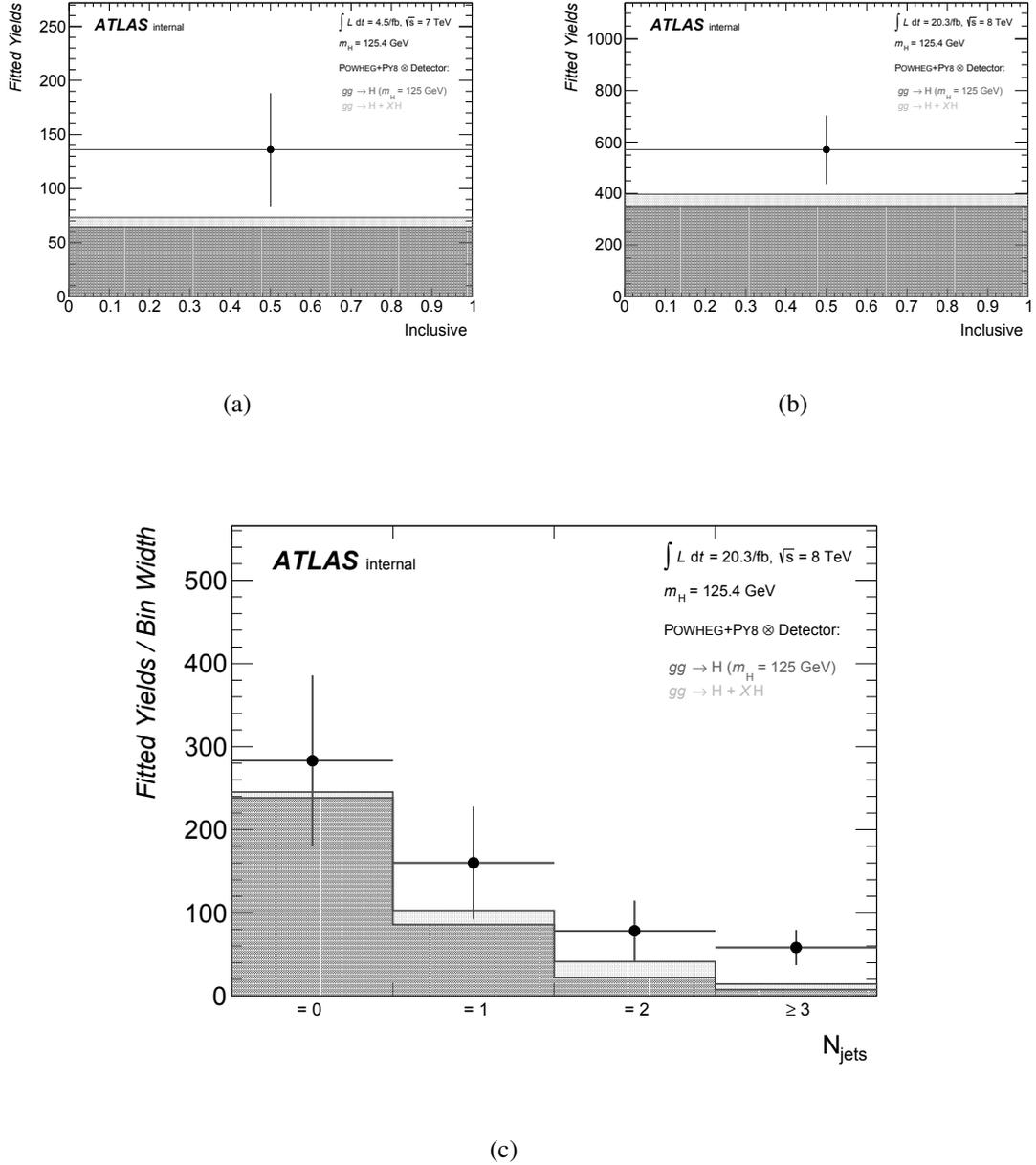


Figure 8.6: The fitted signal yields in data compared to the expected number of reconstructed signal events from simulation corresponding to the (a) 7 TeV and (b) 8 TeV datasets for the inclusive selection, and (c) in bins of N_{jet} for the 8 TeV dataset. The signal contribution due to non-GF processes is denoted as XH ($XH \equiv VBF + WH + ZH + ttH$).

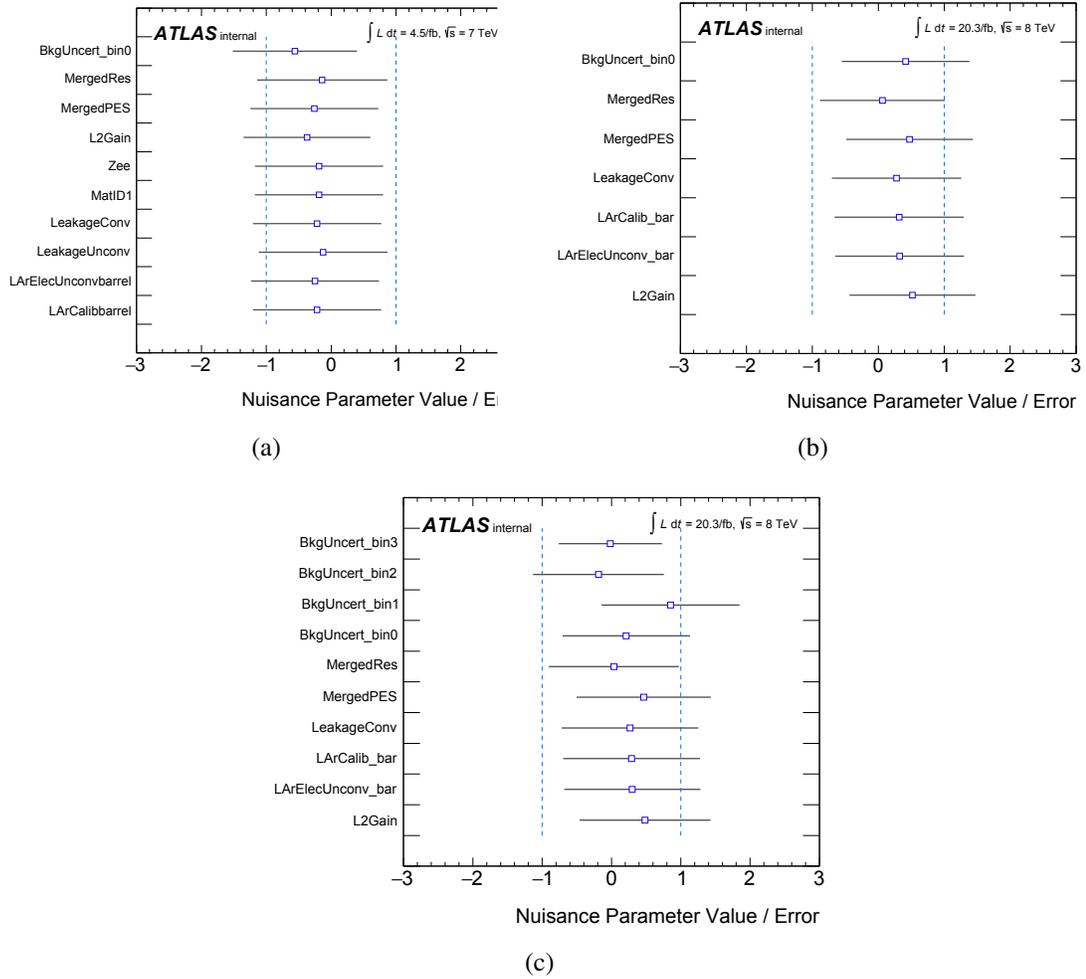


Figure 8.7: Pulls on the nuisance parameters for (a) the 7 TeV inclusive fit, (b) the 8 TeV inclusive fit, and (c) the simultaneous fit of the N_{jet} spectrum at 8 TeV. See Table 8.2 for the definitions of the various nuisance parameters in the figures.

8.5.1 Fit results: impact of the chosen Higgs boson mass on the extracted signal yield

8.5.1.1 Scan across fixed Higgs boson masses

To check the sensitivity of the measurements on the choice of Higgs boson mass the extracted signal yield was profiled as a function of fixed Higgs boson mass. The results of performing these mass scans are presented in Figure 8.8 for the inclusive results at 7 TeV and 8 TeV. The black solid line is the central value of the signal yield and the red dashed lines are the 1σ error bands. The mass scan peaks around 124.5 GeV (126.0 GeV) for the 7 TeV (8 TeV) data, falling fairly steeply on either side of the peak value. These results illustrate a good stability in the extracted signal yield over the Higgs boson mass range of 124 GeV to 125.5 GeV, with any deviations well within the uncertainty on the signal extraction.

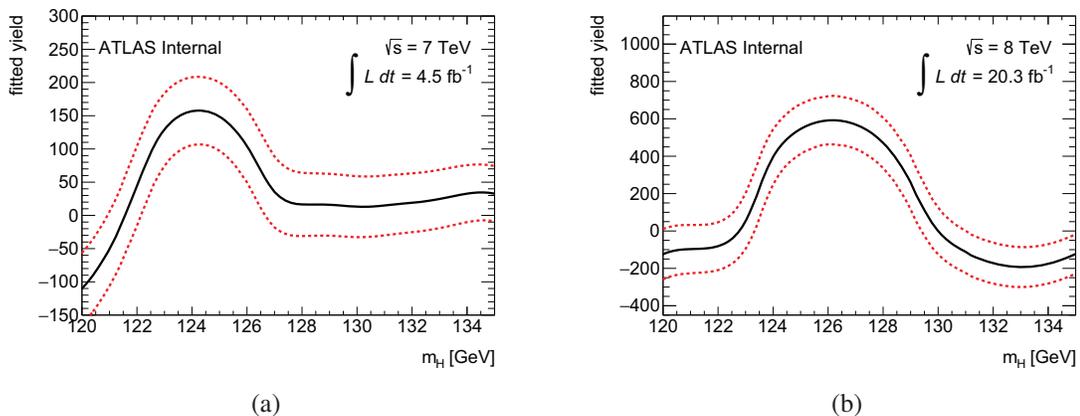


Figure 8.8: Scan across fixed Higgs boson masses for the inclusive selection at (a) 7 TeV and (b) 8 TeV. The black solid line is the central value of the signal yield and the red dashed lines are the 1σ error bands.

8.5.1.2 Signal yields corresponding to the Run-I ATLAS+CMS combined Higgs boson mass

Table 8.4 gives the percent differences between the inclusive and N_{jet} signal yields obtained using the Run-I ATLAS+CMS combined Higgs boson mass versus the Run-I ATLAS $H \rightarrow ZZ^* \rightarrow 4\ell + H \rightarrow \gamma\gamma$ combined Higgs boson mass value used in the current measurements. The reported differences are smaller much than the statistical uncertainties on the signal extraction.

Table 8.4: Percent difference in the extracted signal yields due to using the Run-I ATLAS+CMS combined Higgs boson mass instead of the Run-I ATLAS $H \rightarrow ZZ^* \rightarrow 4\ell + H \rightarrow \gamma\gamma$ combined Higgs boson mass value used in the current measurements.

inclusive fiducial region	ν_i^{sig} $m_H = 125.09 \text{ GeV}$	ν_i^{sig} $m_H = 125.36 \text{ GeV}$	% difference w.r.t. $m_H = 125.36 \text{ GeV}$	stat. uncertainty on yield for $m_H = 125.36 \text{ GeV}$
7 TeV	148	138	+7 %	34 %
8 TeV	554	570	-3 %	22 %

8.6 Summary of the extracted signal yields and uncertainties

Table 8.5 summarizes the signal yield and associated uncertainties for the 7 TeV inclusive fiducial measurement. Table 8.6 summarizes the signal yields and associated uncertainties for the 8 TeV analysis.

Table 8.5: Signal yield associated uncertainties measured using the 7 TeV dataset, for $m_H = 125.36$ GeV.

Signal yield from s+b fit	138
Total uncertainty on s+b fit	$\pm 38\%$
Statistical uncertainty on s+b fit	$\pm 37\%$
Systematic uncertainty on s+b fit	$\pm 7.5\%$
Spurious signal	$\pm 1\%$

Table 8.6: Signal yields measured in data and their associated uncertainties corresponding to the 8 TeV analysis [63].

	inclusive	$N_{\text{jet}} \geq 1$	$N_{\text{jet}} \geq 2$	$N_{\text{jet}} \geq 3$	VBF enhanced
Signal yield from fit	570	308	141	64	334
Total uncertainty on s+b fit	$\pm 23\%$	$\pm 26\%$	$\pm 31\%$	$\pm 34\%$	$\pm 35\%$
Statistical uncertainty on s+b fit	$\pm 22\%$	$\pm 25\%$	$\pm 30\%$	$\pm 33\%$	$\pm 34\%$
Systematic uncertainty on s+b fit	$\pm 6.5\%$	$\pm 7.4\%$	$\pm 7.1\%$	$\pm 6.5\%$	$\pm 9.0\%$
Spurious signal	$\pm 2\%$	$\pm 6\%$	$\pm 5\%$	$\pm 4\%$	$\pm 8\%$

Chapter 9

Correction factors

This chapter provides a detailed description of the model-independent unfolding procedure that is used to correct for experimental effects and translate the extracted detector-level signal yields to particle-level signal yields, producing cross section measurements at the particle-level. Model-independent particle-level cross sections allow the measurements to be compared to a wide variety of theoretical predictions. Reconstruction-level object and particle-level object selections are described in Section 6.3 and Section 6.4, respectively.

Additional fiducial acceptance factors and non-perturbative correction factors are also derived. The fiducial acceptance factor is split into two separate components: kinematic acceptance factors and isolation correction factors. The kinematic acceptance factors translate inclusive theoretical predictions at the parton-level to parton-level predictions in a fiducial region. The isolation correction factor accounts for the particle-level photon isolation requirement. The isolation factor is multiplied by the kinematic acceptance factor to obtain the full translation from inclusive parton-level predictions to fiducial parton-level

predictions. The non-perturbative (NP) correction factor translates fiducial parton-level predictions to fiducial particle-level predictions. Figure 9.1 provides an illustration of how the unfolding, fiducial acceptance, and NP correction factors are applied to data and inclusive parton-level predictions.

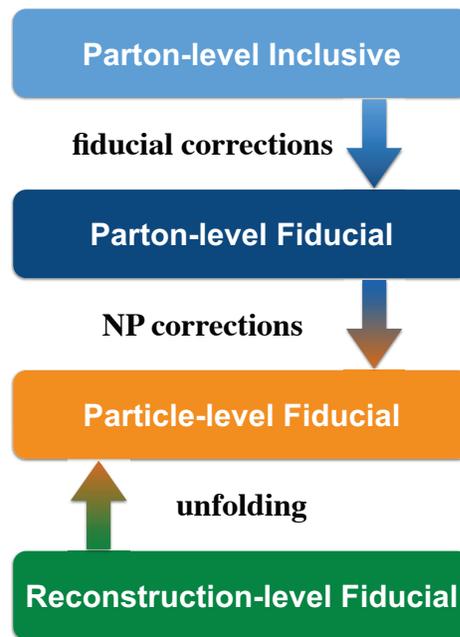


Figure 9.1: An illustration of the application of the different correction factors. Unfolding (correcting for detector effects) takes the data from the reconstruction-level to the particle-level within a fiducial region or bin of a differential distribution. The kinematic acceptance and isolation correction factors translate inclusive theoretical predictions at the parton-level to fiducial theoretical predictions at the parton-level. The non-perturbative (NP) correction factor translates fiducial parton-level theoretical predictions to fiducial particle-level theoretical predictions.

This chapter is organized as follows. First, a reminder of the Higgs boson cross section definition is provided in the next section. This is followed by a description of the unfolding procedure in Section 9.2. Section 9.3 provides details of the various systematic

uncertainties on the unfolding procedure. Section 9.4 summarizes the unfolding factors and their associated uncertainties. Finally, a description of the evaluation of the additional fiducial correction factors is provided in Section 9.5.

The following sections will focus primarily on the details pertaining to the 7 TeV correction factors as this was the author’s main contribution. Additional details and discussion will be presented on 8 TeV correction factors where appropriate.

9.1 Higgs boson cross section definition

The cross-section is defined as:

$$\sigma_i = \frac{\nu_i^{\text{data}}}{c_i \cdot \mathcal{L}_{\text{int}}}, \quad (9.1)$$

where i denotes the particular fiducial region or bin of a differential distribution being considered, ν_i^{data} being the number of measured signal events in region i , c_i is a correction factor accounting for differences in the event yield between detector-level and particle-level due to experimental effects, and \mathcal{L}_{int} the integrated luminosity of the recorded dataset.

9.2 Correcting for detector effects (“unfolding factors”)

The unfolding factors, c_i , correct for detector effects including the trigger inefficiency, photon identification efficiency, and detector acceptance losses. The unfolding factors are derived using simulated samples for the different Higgs boson production processes: gluon fusion (GF), weak vector boson fusion (VBF), associated production with a vector boson

($VH \equiv WH + ZH$), and associated production with a top anti-top pair (ttH). The total fiducial unfolding factor is obtained from the combined sample, in which each production process is normalized to the SM production cross sections for a 125.4 GeV Higgs boson.

The unfolding factors are obtained by dividing the number of selected events at reconstruction-level by the number of events selected at particle-level in each fiducial region (or bin of a differential distribution), defined as:

$$c_i = \frac{n_i^{\text{det}}}{n_i^{\text{part}}}, \quad (9.2)$$

where n_i^{part} and n_i^{det} are the particle-level and reconstruction-level event yields corresponding to the particular fiducial region (or bin of a differential distribution).

Unfolding factors are applied to the extracted event yields in data (ν_i^{data}) in order to translate them to particle-level event yields:

$$n_i^{\text{data,part}} = \frac{\nu_i^{\text{data}}}{c_i}, \quad (9.3)$$

where $n_i^{\text{data,part}}$ is the particle-level event yield in data.

A point worth re-emphasizing is that the barrel-endcap calorimeter transition region ($1.37 < |\eta| < 1.56$) is included in the particle-level fiducial region definitions but excluded in the reconstruction-level selection. Due to the difference in the reconstruction- and particle-level fiducial definitions n_i^{part} contains contributions from the transition region while n_i^{det} does not, hence, the unfolding procedure extrapolates across the transition region correcting for the event losses at reconstruction-level.

A second point worth re-emphasizing is that Dalitz contributions to the signal are removed through the unfolding procedure. At reconstruction-level no attempt is made to remove the Dalitz decays and therefore n_i^{det} contains the Dalitz contributions. At particle-level Dalitz decays are identified and removed, therefore n_i^{part} does not contain contributions from the decays and the unfolding procedure removes all contributions due to Dalitz decays from the $H \rightarrow \gamma\gamma$ cross section measurement.

9.2.1 Particle-level photon isolation

To minimize the topology dependence of the unfolding factors needed to translate the extracted detector-level data yields to particle-level, a selection based on the particle-level photon isolation energy is applied. Reconstructed photons are required to be isolated (see Section 6.3), and part of this requirement is that the transverse energy due to calorimeter energy clusters in a cone of $\Delta R < 0.4$ around the photon is less than 5.5 GeV (6.0 GeV) for the 7 TeV (8 TeV) data sample. A corresponding particle-level isolation energy is obtained by mapping the value of the calorimeter isolation energy to a corresponding particle-level isolation energy. The particle-level isolation energy is defined as the 4-momentum sum of stable particles within a cone of $\Delta R < 0.4$ around the photon, excluding muons and neutrinos. This value is chosen by considering each production mode individually and also the combination of the production modes, weighted by the respective cross sections. Figure 9.2 shows the particle-level isolation versus the calorimeter isolation for the combination of all production modes. The calorimeter isolation energy is profiled against the particle-level isolation energy and a linear fit is performed to the profiled points; the mapped particle-level isolation energy is obtained from this fit function. Table 9.1 gives the mapped values for

each individual production mode and for the combination. For 7 TeV (8 TeV) data a particle-level isolation energy that best described all cases was chosen to be 13.0 GeV (14.0 GeV). As it is desirable to have a consistent fiducial region definition for measurements made at 7 TeV and 8 TeV the impact of choosing a 14 GeV isolation energy for the 7 TeV data was also investigated.

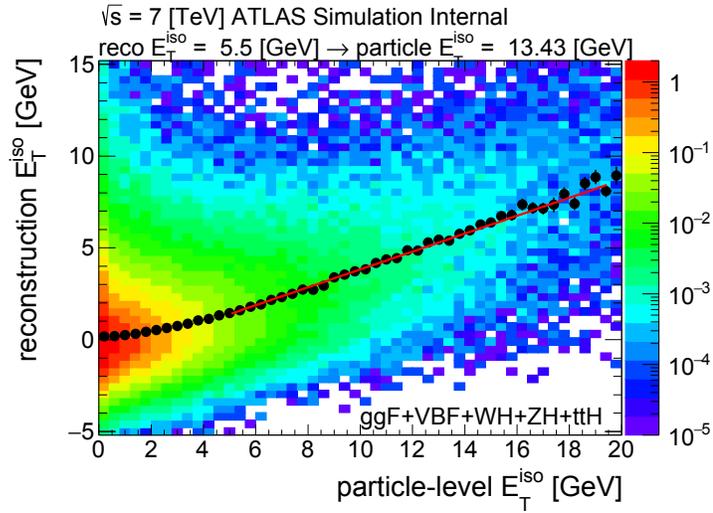
Figures 9.3(a) and (b) show the inclusive fiducial unfolding factors as a function of the production mode with and without the application of the particle-level isolation for 7 TeV and 8 TeV, respectively. The plots clearly illustrate how the application of the particle-level isolation drastically reduces the correction factor dependence on the Higgs boson production mode, thereby mitigating the model dependence of the unfolding procedure. Figure 9.3(a) also shows results corresponding to using a 14 GeV isolation energy for the 7 TeV data. Table 9.2 presents the unfolding factors for each Higgs boson production mode considered in this analysis. The difference between the 13 GeV and 14 GeV isolation cuts is on the order of 0.2% for all Higgs boson production modes except for ttH where it has a 2% impact. Given the generally small impact of applying the 14 GeV isolation selection to 7 TeV data, and the desire to have a consistent fiducial region definition, the 14 GeV selection is adopted for both 7 TeV and 8 TeV measurements.

Table 9.1: Production mode dependence of the particle-level photon isolation.

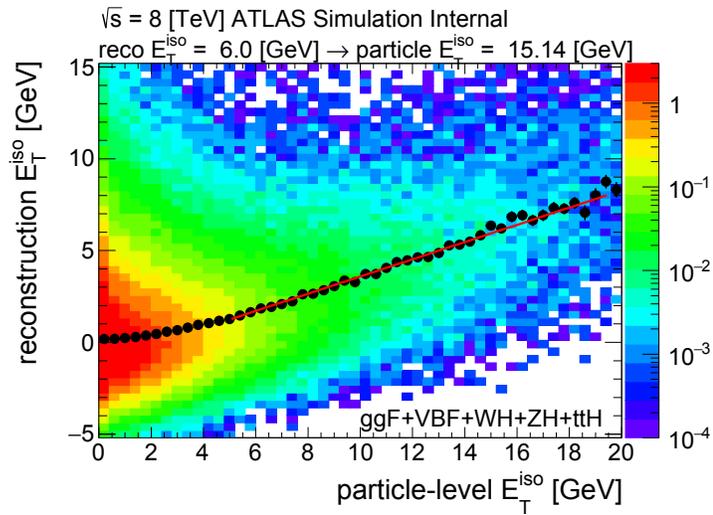
	GF	VBF	WH	ZH	ttH
7 TeV particle-level photon isolation in GeV	13.7	12.9	12.5	12.5	12.8
8 TeV particle-level photon isolation in GeV	15.4	14.6	13.8	14.0	14.3

Table 9.2: Unfolding factors (c) for the inclusive fiducial region, for each SM $H \rightarrow \gamma\gamma$ production process and the total. Unfolding factors are given for each particle-level isolation definition. The 8 TeV unfolding factors are also included for comparison.

	GF	VBF	WH	ZH	ttH	total
7 TeV						
no isolation cut	0.660	0.678	0.618	0.620	0.480	0.659
$E_T^{\text{iso}} < 13$ GeV	0.674	0.695	0.666	0.670	0.683	0.676
$E_T^{\text{iso}} < 14$ GeV	0.673	0.694	0.664	0.667	0.675	0.674
8 TeV						
no isolation cut	0.638	0.643	0.593	0.593	0.453	0.635
$E_T^{\text{iso}} < 14$ GeV	0.650	0.656	0.635	0.637	0.640	0.650

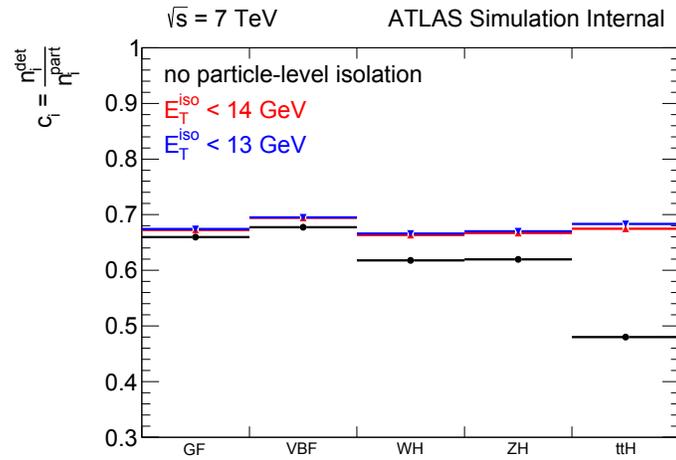


(a)

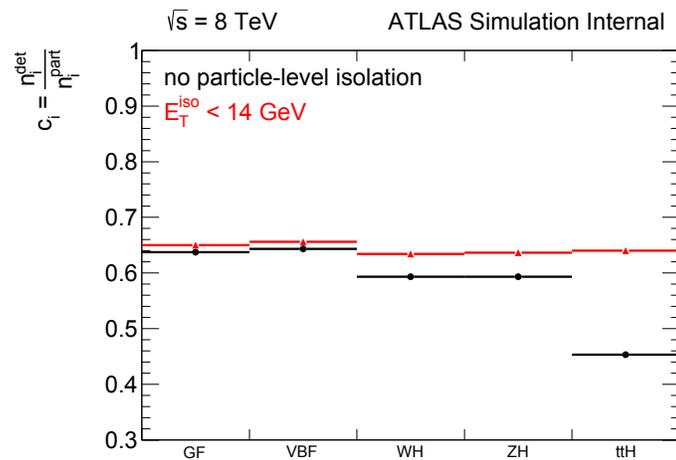


(b)

Figure 9.2: Calorimeter isolation energy against particle-level isolation energy for simulated events at (a) 7 TeV and (b) 8 TeV. The calorimeter isolation energy was profiled against truth isolation energy (black points) and a linear fit (red line) was performed over the range $5 \text{ GeV} \leq \text{particle-level } E_T^{\text{iso}} \leq 19 \text{ GeV}$. These results are independent of the track isolation transverse momentum selection.



(a)



(b)

Figure 9.3: Fiducial unfolding factors as a function of the Higgs boson production mode with and without the particle-level photon isolation: (a) 7 TeV and (b) 8 TeV.

9.3 Systematic uncertainties on the unfolding factors

The following sections discuss the systematic uncertainties that are associated with the experimental and theoretical modelling of the simulated Higgs boson samples that are used to calculate the unfolding factors.

9.3.1 Photon selection

Uncertainties associated with the selection of reconstructed photons impact the unfolding factors by altering the number of reconstruction-level events. The uncertainties on the photon selection due to the trigger efficiency, the photon identification efficiency and the photon isolation are included as uncertainties on the unfolding factor.

- **Trigger:** The trigger efficiency corresponding to diphoton events entering the signal region was found to be $98.9^{+0.2}_{-0.3}$ % for the majority of the 7 TeV dataset and $98.8^{+0.3}_{-0.4}$ % for the remaining 7 TeV dataset. The trigger efficiency corresponding to diphoton events entering the signal region collected at 8 TeV was found to be 99.61 ± 0.5 %. A 0.5 % uncertainty is assigned to the measurements obtained at both 7 TeV and 8 TeV. The trigger efficiency and the associated uncertainty were obtained from dedicated studies by the $H \rightarrow \gamma\gamma$ analysis group [62].
- **Photon identification efficiency:** The photon identification efficiency is measured using three procedures.
 - Photons ‘tagged’ in $Z \rightarrow \ell\ell\gamma$ decays provide a very pure sample at low p_T .

- Simulated MC events are used to create a mapping from electron showers to photon showers, and applied to $Z \rightarrow ee$ events in data.
- From a data sample of isolated photon candidates from prompt γ +jet production after subtracting the measured fraction of jet+jet background events.

These three methods agree within their uncertainties and are combined to obtain the final uncertainty on the photon identification efficiency. An uncertainty of 8.4 % (1.4 %) on the diphoton identification is assigned to the 7 TeV (8 TeV) measurements.

- **Photon isolation efficiency:** The photon isolation selection efficiency uncertainty is estimated from a data sample containing $Z \rightarrow e^+e^-$ events. The same cuts are applied on the photon isolation variables' electron counterparts, and the selection efficiency is compared between data and simulation. An isolation selection efficiency uncertainty of 1.4 % and 1 % is assigned to the 7 TeV and 8 TeV inclusive analyses, respectively. The uncertainty increases as a function of the number of hadronic jets, reaching 4 % for events containing three or more jets.

9.3.2 Detector modelling

The unfolding factors are derived from simulated samples as described in Section 9.2. Therefore, differences between data and the simulation due to modelling of the detector can potentially affect the correction factors, this is accounted for by assigning an uncertainty. The uncertainty due to detector modelling is estimated by re-evaluating the correction factors after systematically shifting detector-level distributions in the simulation and then comparing them to the nominal unfolding factors. For the inclusive 7 TeV fiducial analysis

the only detector-modelling considerations come from the photon energy scale and resolution uncertainties. For the 8 TeV analysis additional uncertainties due to the jet energy scale and resolution, the JVF and pileup, and E_T^{miss} must also be accounted for.

9.3.2.1 Photon and electron energy scale and resolution

The unfolding factors can potentially be affected by the photon and electron energy calibrations. Photon migrations in and out of the fiducial regions or bins of a differential distribution at reconstruction-level can occur due to the photon energy scale calibration. Since jets overlapping with electrons are removed the electron scale calibrations can also cause bin migrations. To assess the impact of photon and electron energy calibration on the unfolding factors each of the systematic uncertainty sources on the scale and resolution of reconstructed photons was varied and propagated through the unfolding procedure independently. The uncertainty on the unfolding factor was taken as the percent difference between the nominal unfolding factors and those obtained from each systematic shift, with the total uncertainty obtained by summing all the uncertainties in quadrature. Both the photon energy scale and resolution uncertainty components were found to have a very small effect on the unfolding factors, $\leq 0.1\%$ for the photon energy calibration and $\leq 0.01\%$ for the electron energy calibration, therefore these uncertainty sources were neglected.

9.3.2.2 Jet energy scale and resolution uncertainties

The uncertainty on the jet energy resolution (JER) measurement is propagated to the unfolding factor by smearing the jet energy in simulation to match that measured in data. The JER uncertainty was measured using a combination of *in situ* techniques [55].

The uncertainty on the unfolding factors due to uncertainties on the jet energy scale (JES) are split into 16 uncorrelated categories [55]. To assess the impact of the JES uncertainties on the unfolding factors each of the 16 JES uncertainty sources was varied and propagated through the unfolding procedure independently. The total uncertainty is assigned as the quadratic sum. The JES uncertainties are categorized according to:

- Baseline JES uncertainty: six uncertainty sources originating from the various *in situ* JES measurements for central jets.
- η -intercalibration: two uncertainty sources from the forward jet calibration procedure presented in Chapter 5. The uncertainty has two components, one uncertainty source from simulation modelling of forward jets, and one from the statistical precision of the calibration.
- Pileup: four sources of uncertainty due to the pileup dependence of the jet calibrations.
- Flavour composition: two sources of uncertainty from the different calorimeter response expected for jets initiated by quarks or gluons. Quark jets are narrower than gluon jets and have a response that is several percent higher [68].
- Detector dead material: a single uncertainty component to account for mis-modelling of dead material in the simulation.
- High p_T : a single uncertainty component only applied to high p_T jets ($p_T^{\text{jet}} > 800$ GeV). Essentially only low p_T jets enter the analysis and therefore this source of uncertainty has no impact on the analysis.

9.3.2.3 Jet vertex fraction (JVF)

The impact of the JVF cut due to the mis-modelling of the JVF distribution in the simulation is considered. The unfolding factors are re-derived after shifting the JVF selection thresholds up or down

9.3.2.4 Missing transverse energy uncertainties

All physics objects used in calculating E_T^{miss} are shifted according to their uncertainties and E_T^{miss} is re-calculated and propagated through the analysis.

9.3.3 Theoretical modelling

The unfolding factor used to correct the data for detector effects and translate the measurements to particle-level can be sensitive to the underlying physics modelling of the simulation used to derive it. The goal is to obtain unfolding factors that have minimal dependence on the assumed theoretical inputs to the simulation, therefore it is necessary to assess the impact on the unfolding factor (modelling uncertainty) from using simulations with different theoretical inputs. To obtain an estimate of the modelling uncertainty a number of approaches are investigated: varying the signal composition, data-based re-weighting of the Higgs boson's p_T and rapidity spectrums, using different generators to simulate the $H \rightarrow \gamma\gamma$ decays, and turning off multiple parton interactions in the simulation.

9.3.3.1 Signal composition

The unfolding factors are derived using a simulated sample composed of the sum of different Higgs boson production processes: gluon fusion (GF), weak vector boson fusion (VBF), associated production with a vector boson ($VH \equiv WH + ZH$), and associated production with a top anti-top pair (ttH). Each production process is normalized to the SM production cross sections for a 125.4 GeV Higgs. To estimate the uncertainty due to the assumption of a SM composition, the cross sections of the vector boson contribution (VBF + VH) and the ttH component are individually varied up and down with respect to the GF contribution. The estimated variations are obtained from limits set by data in order to cover the lack of knowledge of the couplings. The up and down vector boson component variations are $0.5 \times \sigma_{\text{VBF}+VH}$ and $2 \times \sigma_{\text{VBF}+VH}$, respectively, while the ttH component is varied from $0 \times \sigma_{ttH}$ to $5 \times \sigma_{ttH}$. The percent difference between the nominal unfolding factors and those obtained after varying the signal composition are given in Table 9.3 for the 7 TeV analysis corresponding to the $E_T^{\text{iso}} < 14$ GeV particle-level photon isolation criteria. The 7 TeV uncertainties are consistent with those obtained at 8 TeV and are essentially negligible compared to the other theoretical modelling uncertainties discussed in the following sections.

Table 9.3: Signal composition uncertainties on the 7 TeV unfolding factor.

7 TeV inclusive	$E_T^{\text{iso}} < 14$ GeV
$2 \times \sigma_{\text{VBF}+VH}$	0.14 %
$0.5 \times \sigma_{\text{VBF}+VH}$	-0.08 %
$5 \times \sigma_{ttH}$	0.002 %
$0 \times \sigma_{ttH}$	-0.0005 %

9.3.3.2 Generator dependence: matrix element (ME) , parton-shower (PS), and underlying event (UE)

The nominal unfolding factors are derived using POWHEG+PYTHIA8 for the GF and VBF Higgs production processes, and PYTHIA8 for the WH , ZH , and ttH production processes. Since there is no a priori reason to select one generator over another and different generators can potentially result in different unfolding factors, an uncertainty is assigned to cover any observed difference in the unfolding factor. The impact of changing GF modelling of the generated samples was investigated by generating GF samples using a variety of different generators. To assess the impact of different parton-showering models, the POWHEG GF and VBF samples were showered using a variety of different showering implementations and tunes. The impact of the underlying event on the unfolding factors was estimated by turning off multiple parton interactions (MPI) in the GF and VBF samples. The uncertainty on the unfolding factors due to generator dependence is taken as the envelope of the unfolding factors derived from the aforementioned variations.

Table 9.4 summarizes the generator modelling uncertainties obtained for the 7 TeV analysis corresponding to the $E_T^{\text{iso}} < 14$ GeV particle-level photon isolation criteria. Turning off MPI is the dominant uncertainty for both the 7 TeV and 8 TeV analyses, it is evaluated to be: 2.4 % (1.1 %) for the 7 TeV (8 TeV) inclusive fiducial region, and 1.0 %, 1.3 %, 3.1 %, 3.9 % for the $N_{\text{jet}} = 0$, $N_{\text{jet}} = 1$, $N_{\text{jet}} = 2$, and $N_{\text{jet}} \geq 3$ regions, respectively.

Table 9.4: Uncertainties on the 7 TeV unfolding factor due to generator modelling, in percent.

7 TeV inclusive	$E_T^{\text{iso}} < 14 \text{ GeV}$
MC@NLO	-0.21 %
HERWIG	0.04 %
POWHEG+JIMMY	-0.64 %
POWHEG+HERWIG	-0.20 %
POWHEG+HERWIG+PERUGIA	0.09 %
MPI off	2.4 %

9.3.3.3 Data-driven Higgs boson p_T and rapidity weighting

To account for differences between the spectra of diphoton kinematic variables observed between data and simulation, a re-weighting was applied to the simulation. The Higgs boson's p_T and rapidity were the two kinematic distributions considered for the re-weighting, with these two distributions being approximately uncorrelated. In 8 TeV, the Higgs boson p_T (rapidity) spectrum in data tends to be harder (more forward) than predicted by the simulation. Continuous weighting functions were obtained by comparing the unfolded p_T and rapidity spectra in data to simulated particle-level predictions. The simulation was then simultaneously weighted to the p_T and rapidity spectra in data. This procedure was developed using the 8 TeV data. For the 7 TeV analysis it was not possible to obtain stable diphoton p_T and rapidity spectra from the data due to insufficient statistics, therefore the weighting functions derived using 8 TeV data were applied to the 7 TeV simulation samples. The uncertainty due to the Higgs boson p_T and rapidity weighting is the dominant negative signed uncertainty on the unfolding factors, it is evaluated to be: -2.4 % (-3.2 %) for the 7 TeV (8 TeV) inclusive fiducial region, and -3.5 %, -6.6 %, -3.5 %, -1.1 % for the $N_{\text{jet}} = 0$,

$N_{\text{jet}} = 1$, $N_{\text{jet}} = 2$, and $N_{\text{jet}} \geq 3$ regions, respectively.

9.3.3.4 Combination of generator modelling uncertainties

The assigned simulation modelling uncertainty is taken as the envelope of the data-driven weighting uncertainty and the quadratic sum of the signal composition and generator modelling uncertainties. The signs of the uncertainties are preserved, resulting in an asymmetric final uncertainty. The envelope of the data weighting with the other uncertainties is taken in order to avoid any potential double-counting. The positive signed uncertainty is dominated by the data-driven weighting of the Higgs boson's p_T and rapidity spectrums, while the negative signed uncertainty is dominated by the generator modelling component, which is generally dominated by the effect of turning off MPI.

9.4 Unfolding factors and uncertainties

Table 9.5 summarizes the unfolding factors and their associated uncertainties for each measurement region, corresponding to a particle-level photon isolation cut of 14 GeV.

Table 9.5: Unfolding factors and uncertainties for the 7 TeV and 8 TeV analyses, corresponding to a particle-level photon isolation cut of $E_T^{\text{iso}} < 14$ GeV.

	7 TeV inclusive	8 TeV inclusive [63]
unfolding factor	0.674	0.650
stat.	$\pm 0.09\%$	$\pm 0.05\%$
photon ID	$\pm 8.4\%$	$\pm 1.4\%$
photon isolation	$\pm 1.4\%$	$\pm 1.0\%$
detector modelling	$< 0.01\%$	$< 0.01\%$
theory modelling	$^{+2.4}_{-2.4}\%$	$^{+1.0}_{-3.3}\%$
trigger efficiency	0.5%	0.5%

9.5 Fiducial acceptance factors and non-perturbative correction factors

9.5.1 Fiducial acceptance factors

Fiducial acceptance factors are derived in order to compare the measurements to theoretical predictions that do not decay the Higgs boson. These acceptance factors are calculated at parton-level to allow translation from the full parton-level phase-space to the particle-level fiducial region of the measurements as described in Section 6.6.

The fiducial acceptance factors have two components: a factor to account for the kinematic acceptance of the two photons from the Higgs boson decay, and a factor to account for the particle-level isolation efficiency. The product of the kinematic and isolation factors gives the total fiducial acceptance factor.

The error on the kinematic fiducial factors is calculated for the one sigma (*i.e.* the 68 % confidence level) uncertainty of the parton distribution functions considered. Three different parton distribution function sets are used to calculate this uncertainty. The envelope of the acceptance factors obtained from all parton distribution function variations is taken as the error on the fiducial correction factor. The final fiducial correction factors were then taken as the central value of this envelope.

The isolation efficiency correction factor is calculated for different parton-showering procedures and tunes, with the central value taken as the acceptance factor and the envelope of these acceptance factors assigned as the uncertainty. Table 9.6 gives the resulting kinematic and isolation fiducial acceptance factors and their associated uncertainties.

Table 9.6: Diphoton fiducial acceptance factors in percent. The kinematic acceptance factor is defined as the probability to fulfil the diphoton kinematic acceptance criteria: $|y| < 2.37$ and $p_T/m_{\gamma\gamma} < 0.35(0.25)$ for the leading (sub-leading) photon. The isolation efficiency factor is defined as the probability for both photons to pass the particle-level isolation requirement ($E_T^{\text{iso}} < 14$ GeV) for events that fulfil the diphoton kinematic criteria.

kinematic acceptance factor [%]	
7 TeV inclusive fiducial region	63.3 ± 0.9
8 TeV inclusive fiducial region [64]	62.6 ± 0.9
isolation efficiency factor [%]	
7 TeV inclusive fiducial region	97.8 ± 1.1
8 TeV inclusive fiducial region [64]	98.1 ± 0.5

9.5.2 Non-Perturbative correction factor

Perturbative QCD calculations provide predictions at the parton level, therefore non-perturbative (NP) effects must be accounted for in order to compare between theoretical predictions and the unfolded measurements. Two different data samples are required to derive the NP corrections; a sample in which the parton-level final state is kept, and a sample in which the hadronic final state is kept. Partonic and hadronization effects cannot be factorized since multiple parton interactions (MPI) can feed into the hadronization process. The NP correction factor is evaluated as:

$$C_{\text{NP}} = \frac{N^{\text{MPI+Had}}}{N^{\text{noMPI+noHad}}} , \quad (9.4)$$

where C_{NP} is the NP correction factor, $N^{\text{MPI+Had}}$ is the number of events obtained at the hadron level and $N^{\text{noMPI+noHad}}$ is the number of events obtained at the parton-level.

The impact on the NP correction factors due to NP contributions which are incalculable (MPI and hadronization), are evaluated by implementing dedicated phenomenological models. In this analysis, PYTHIA8 and HERWIG++, are utilized in order to extract the NP cor-

Table 9.7: Non-perturbative correction factors that account for the impact of hadronization and underlying event, in percent.

NP correction factor [%]	
7 TeV inclusive fiducial region	99.9 ± 1.1
8 TeV inclusive fiducial region [64]	99.9 ± 0.6

rection to the NLO calculation. The PYTHIA and HERWIG Monte Carlo (MC) generators implement different parton-shower evolution prescriptions, p_T -ordered and angular-ordered, respectively. The two generators also utilize different hadronization and fragmentation models, Lund String Model versus Cluster Model, respectively. Additionally, the two MC generators utilize different implementations of underlying event (UE); PYTHIA models UE using MPI which affect both the parton and hadron level, while HERWIG models UE through interactions with the beam remnants which only impacts the hadron level. Furthermore, the possible interdependence of effects between MPI and parton-showering are also considered.

The baseline MC sample is generated with POWHEG, using the CT10 PDF set and PYTHIA8 to perform the parton-showering, hadronization, and UE modelling. Comparisons are made using the POWHEG generator interfaced with PYTHIA8, HERWIG++, and PYTHIA6, and implementing various showering tunes. The central value of all modelling and tune variations was taken as the NP correction factor and the envelope of the variations was taken as the uncertainty. The NP correction factors and their associated uncertainties are summarized in Table 9.7.

Chapter 10

Results

10.1 Theoretical predictions

Cross section predictions in the inclusive fiducial region utilize the highest order theoretical calculations available for Higgs boson production via gluon fusion (GF) at the parton-level. Following the recommendations of the LHC Higgs cross section working group, the inclusive fiducial measurements are compared to theoretical predictions that incorporate calculations of the GF production cross section that are accurate to next-to-next-to-leading order (NNLO) [69, 70, 71, 72, 73, 74] plus next-to-next-to-leading-logarithmic order (NNLL) [75, 76, 77] in QCD. Electroweak (EW) corrections calculated at next-to-leading order (NLO) are also included [78, 79]. These predictions are referred to as LHC-XS. For the 8 TeV analysis two additional cross section predictions of Higgs boson production via GF are included (STZW [80] and HRes [81]), both calculations are accurate to NNLO+NNLL.

Theoretical predictions corresponding to fiducial regions characterized by the hadronic jet activity are obtained from simulated samples of GF Higgs boson production in association with one and two jets. Higgs boson plus one jet (HJ) and Higgs boson plus two jets (HJJ) [82] events are generated using the POWHEG processes with MINLO (Multi-scale improved NLO) [83, 84]. HJ events are accurate at NLO for Higgs boson plus one jet events, while HJJ events are accurate at NLO for Higgs boson plus two jets events. Additionally, a parton-level GF cross section has been calculated using soft-collinear effective theory for events containing a Higgs boson and one or more hadronic jets. The prediction combines NNLO+NNLL zero-jet and NLO+NLL one-jet cross sections and is referred to as BLPTW [85]. Furthermore, a NNLO+NNLL parton-level prediction is provided by JetVHeto [86].

Particle-level predictions from the POWHEG+PYTHIA8 and PYTHIA8 simulations are used to obtain the SM predictions for the VBF , WH , ZH , and ttH production processes in all measurement regions. Each prediction is normalized to state-of-the-art theoretical calculations, for more details on the normalizations see Appendix A.

The fiducial correction factors presented in Section 9.5 are used to translate all theoretical predictions at the parton-level to the particle-level in order to facilitate comparison with data. The $H \rightarrow \gamma\gamma$ branching ratio is taken to be $(0.228 \pm 0.011) \%$ [87].

10.2 Higgs boson cross section measurements

As detailed in Section 6.1, the 7 TeV and 8 TeV datasets used in this analysis correspond to integrated luminosities of 4.5 fb^{-1} and 20.3 fb^{-1} with associated uncertainties of 1.8 %

and 2.8 %, respectively. The uncertainty on the luminosity is treated as a relative uncertainty on the measured cross sections and is therefore quoted separately.

Figure 10.1 compares the measured 7 TeV and 8 TeV fiducial cross sections to the theoretical predictions constructed using the LHC-XS predictions for all production processes. Figure 10.2 presents the fiducial cross sections and cross section limits measured using the 8 TeV ATLAS dataset. The measurements are compared to a selection of state-of-the-art theoretical predictions for the production of a SM Higgs boson with a mass of 125.4 GeV. The theoretical predictions correspond to the sum of the cross section predictions for the five dominant Higgs boson production modes, GF, VBF, WH , ZH , and ttH .

The inclusive fiducial cross sections measured at 7 TeV and 8 TeV are:

$$\begin{aligned}\sigma_{\text{fid}}(pp \rightarrow H \rightarrow \gamma\gamma) &= 45.3 \pm 16.9 \text{ (stat.) } \begin{matrix} +5.3 \\ -5.3 \end{matrix} \text{ (syst.) } \pm 0.8 \text{ (lumi.) fb,} \\ \sigma_{\text{fid}}(pp \rightarrow H \rightarrow \gamma\gamma) &= 43.2 \pm 9.4 \text{ (stat.) } \begin{matrix} +3.2 \\ -2.9 \end{matrix} \text{ (syst.) } \pm 1.2 \text{ (lumi.) fb,}\end{aligned}$$

respectively. These can be compared with the SM predictions for inclusive Higgs boson production of 24.4 ± 2.3 fb and 30.5 ± 3.3 fb at centre-of-mass energies of 7 TeV and 8 TeV, respectively. The predictions are constructed using the LHC-XS prediction. The ratio of the data to the LHC-XS prediction is 1.83 ± 0.73 and 1.41 ± 0.36 for 7 TeV and 8 TeV, respectively. These ratios are consistent with dedicated signal strength measurements of the Higgs boson in the $H \rightarrow \gamma\gamma$ channel [62]. The selection of theoretical predictions considered all fall below the data, however, given the statistical limitation of the data the excess is not significant. The measured and predicted fiducial cross sections obtained at centre-of-mass energies of 7 TeV and 8 TeV are summarized in Table 10.1 and Table 10.2, respectively. Table 10.3 and Table 10.4 summarize the various uncertainties on the fiducial

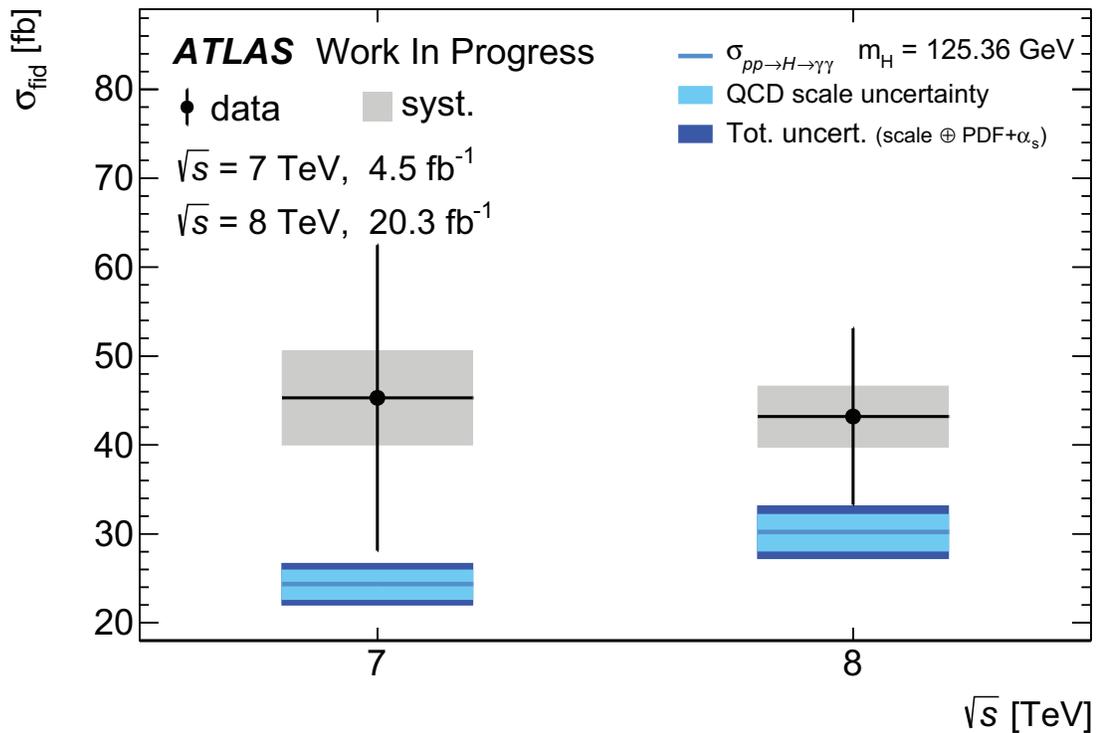


Figure 10.1: Fiducial cross sections for $pp \rightarrow H \rightarrow \gamma\gamma$ as a function of \sqrt{s} . Black circles represent the data and the error bar on each data point is the total measurement uncertainty. The grey rectangles on each data point represent the systematic uncertainty. The data are compared to theoretical predictions constructed entirely from the LHC-XS predictions [87].

cross sections measured at 7 TeV and 8 TeV, respectively.

Three theoretical predictions are compared to the measured cross section in the fiducial regions characterized by requiring one or more hadronic jets. The data are generally in agreement with the theoretical predictions using the BLPTW and JetVHeto calculations for the GF component of the cross section. For the fiducial region characterized by requiring at least two hadronic jets, the data are in good agreement with the prediction based on the BLPTW calculation. In both the inclusive one and two jet fiducial regions the data

Table 10.1: Measured cross sections and 95 % cross-section limits in the various fiducial regions for the 7 TeV and 8 TeV analyses. The fiducial regions are defined in Section 6.6. The uncertainty associated with the luminosity (lumi.) is quoted separately along with the statistical (stat.) and systematic (syst.) uncertainties.

7 TeV	
Fiducial region	measured cross section [fb]
inclusive (diphoton baseline)	45.3 ± 16.9 (stat.) $^{+5.3}_{-5.3}$ (syst.) ± 0.8 (lumi.)
8 TeV [63]	
Fiducial region	measured cross section [fb]
inclusive (diphoton baseline)	43.2 ± 9.4 (stat.) $^{+3.2}_{-2.9}$ (syst.) ± 1.2 (lumi.)
$N_{\text{jet}} \geq 1$	21.5 ± 5.3 (stat.) $^{+2.4}_{-2.2}$ (syst.) ± 0.6 (lumi.)
$N_{\text{jet}} \geq 2$	9.2 ± 2.8 (stat.) $^{+1.3}_{-1.2}$ (syst.) ± 0.3 (lumi.)
$N_{\text{jet}} \geq 3$	4.0 ± 1.3 (stat.) ± 0.7 (syst.) ± 0.1 (lumi.)
VBF-enhanced	1.68 ± 0.58 (stat.) $^{+0.24}_{-0.25}$ (syst.) ± 0.05 (lumi.)
$N_{\text{leptons}} \geq 1$	< 0.80
$E_{\text{T}}^{\text{miss}} \geq 80$ GeV	< 0.74

Table 10.2: Predicted cross sections and 95 % cross-section limits in the various fiducial regions for the 7 TeV and 8 TeV analyses. The fiducial regions are defined in Section 6.6.

7 TeV		
Fiducial region	Theory prediction [fb]	Source
inclusive (diphoton baseline)	24.4 ± 2.3	LHC-XS + XH
8 TeV [63]		
Fiducial region	Theory prediction [fb]	Source
inclusive (diphoton baseline)	30.5 ± 3.3 $34.1^{+3.6}_{-3.5}$ $27.2^{+3.6}_{-3.2}$	LHC-XS + XH STWZ + XH HRes + XH
$N_{\text{jet}} \geq 1$	13.8 ± 1.7 $11.7^{+2.0}_{-2.4}$ $11.7^{+1.8}_{-1.2}$	BLPTW + XH JetVHeto + XH MINLO HJ + XH
$N_{\text{jet}} \geq 2$	5.65 ± 0.87 $3.99^{+0.56}_{-0.59}$	BLPTW + XH MINLO HJJ + XH
$N_{\text{jet}} \geq 3$	0.94 ± 0.15	MINLO HJJ + XH
VBF-enhanced	0.87 ± 0.08	MINLO HJJ + XH
$N_{\text{leptons}} \geq 1$	0.27 ± 0.02	XH
$E_{\text{T}}^{\text{miss}} \geq 80$ GeV	0.14 ± 0.01	XH

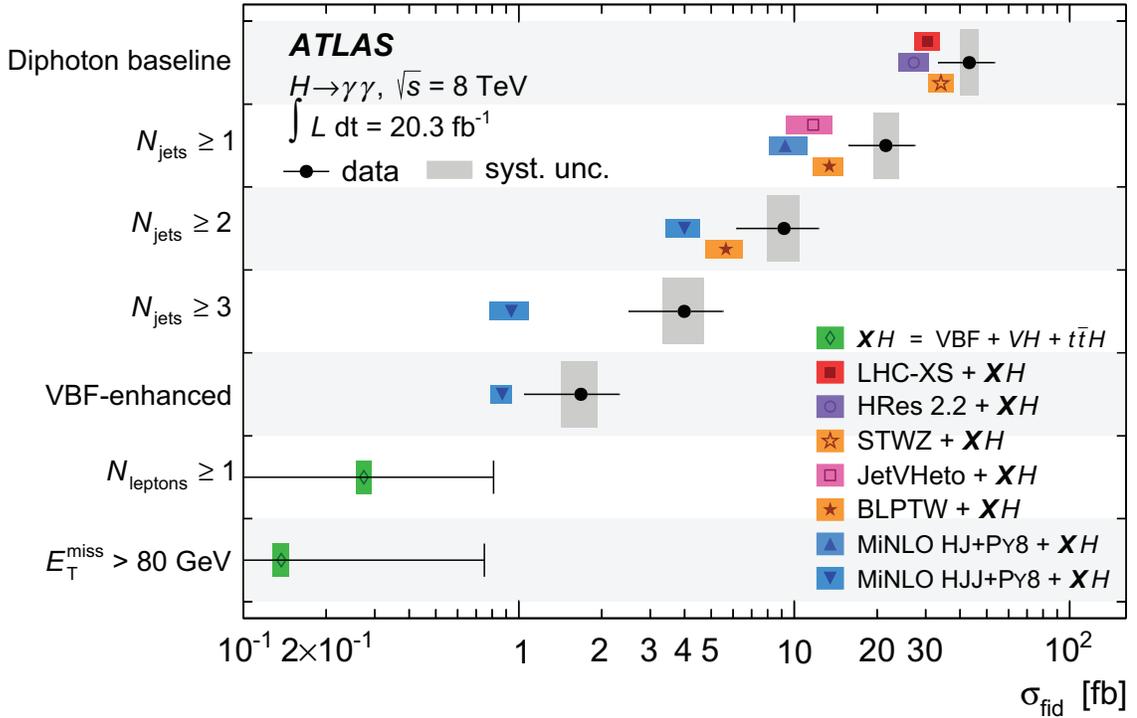


Figure 10.2: Fiducial cross sections (and limits) for $pp \rightarrow H \rightarrow \gamma\gamma$ measured at $\sqrt{s} = 8$ TeV compared to theoretical predictions [63]. The vertical axis defines the fiducial regions defined in Section 6.6 and the horizontal axis is the cross section in fb. Black circles represent the data and the error bar on each data point is the total measurement uncertainty. The grey rectangles on each data point represent the systematic uncertainty. The two cross section limits for $N_{\text{leptons}} \geq 1$ and $E_T^{\text{miss}} \geq 80$ GeV are shown at the 95 % confidence level. The data are compared to theoretical predictions as described in the previous section, where $VH \equiv WH + ZH$ and $XH \equiv \text{VBF} + VH + ttH$. The total uncertainty on each prediction is represented by the width of the marker.

are less well described by the predictions based on MINLO HJ or MINLO HJJ, indicating that the higher-order corrections included in the BLPTW and JetVHeto calculations play an important role. In the VBF-enhanced fiducial region the theoretical predictions based on MINLO HJJ for GF production and POWHEG for VBF production are in reasonable agreement with the data. Given that the VBF process is the dominant contribution to

Table 10.3: Uncertainties on the measured cross sections, expressed as percentages, corresponding to the inclusive fiducial region, for the 7 TeV analysis.

Uncertainty source	Uncertainty on the fiducial cross section [%]
Signal extraction stat.	± 37
Signal extraction syst.	± 7.5
Spurious signal	± 1
Photon efficiency	± 8.4
Theoretical modelling	$+2.4$ -2.4
Luminosity uncertainty	± 1.8

Table 10.4: Uncertainties on the measured cross sections, expressed as percentages, corresponding to the inclusive, $N_{\text{jet}} \geq 1$, $N_{\text{jet}} \geq 2$, $N_{\text{jet}} \geq 3$ and VBF-enhanced fiducial regions, for the 8 TeV analysis [63].

Uncertainty source	Uncertainty on the fiducial cross section [%]				
	inclusive	$N_{\text{jet}} \geq 1$	$N_{\text{jet}} \geq 2$	$N_{\text{jet}} \geq 3$	VBF enhanced
Signal extraction stat.	± 22	± 25	± 30	± 33	± 34
Signal extraction syst.	± 6.5	± 7.4	± 7.1	± 6.5	± 9.0
Spurious signal	± 2	± 6	± 5	± 4	± 8
Photon efficiency	± 1.5	± 2.1	± 3.1	± 4.2	± 2.3
Jet energy scale and resolution	-	$+6.2$ -5.8	$+11$ -10	$+15$ -13	$+12$ -11
JVF and pileup-jet	-	± 1.3	± 2.2	± 3.3	± 0.5
Theoretical modelling	$+3.3$ -1.0	$+5.0$ -2.6	± 4.1	$+6.3$ -4.9	$+2.2$ -3.2
Luminosity	± 2.8	± 2.8	± 2.8	± 2.8	± 2.8

the cross section in this fiducial region, the comparison of the data with the predictions is sensitive to the theoretical modelling of the VBF process. For the $N_{\text{leptons}} \geq 1$ and $E_{\text{T}}^{\text{miss}} \geq 80$ GeV fiducial regions cross section limits of 0.80 fb and 0.74 fb were obtained, respectively. Given that the obtained limits are a factor of three to five larger than the predictions, they cannot be used to constrain the theoretical predictions for these processes, however, the limits can potentially be used to constrain theories involving Higgs boson production in association with dark matter.

The differential cross section as a function of the hadronic jet multiplicity for jets with

$p_T > 30$ GeV is shown in Figure 10.3. The data are compared to the theoretical predictions based on MINLO HJ for GF production and the nominal simulation samples for the other production processes. The MINLO HJ prediction is normalized to the LHC-XS prediction using a scaling factor (K_{ggF}) of 1.54. The agreement between the data and theoretical predictions is reasonably good, with the data exhibiting a small, statistically insignificant, excess in the highest jet multiplicity bin.

Figure 10.4 and Figure 10.5 show the differential cross sections of the transverse momentum and the rapidity of the diphoton system, respectively. Both differential distributions were measured for the inclusive fiducial region. The HRes prediction for GF production is combined with the predictions obtained from the nominal simulation samples for the other production mechanisms. The HRes predictions are normalized to the total LHC-XS cross section using a scaling factor (K_{ggF}) of 1.15. The SM prediction reasonably describes the shape of the measured differential distributions, with the data exhibiting an overall excess that is consistent with the total inclusive fiducial cross section measurement presented above.

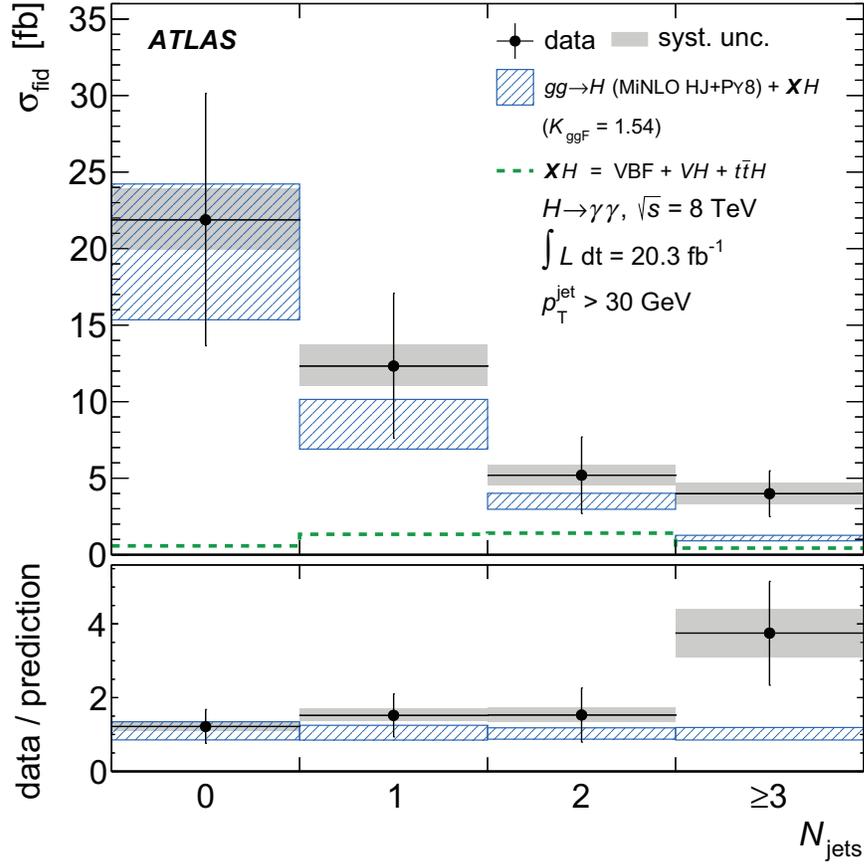


Figure 10.3: The differential cross section for $pp \rightarrow H \rightarrow \gamma\gamma$ as a function of the jet multiplicity for jets satisfying $p_{\text{T}} > 30 \text{ GeV}$ [63]. Black circles represent the data and the error bar on each data point is the total measurement uncertainty. The grey rectangles on each data point represent the systematic uncertainty. The blue hatched band represents the SM prediction, with the size of the band reflecting the total theoretical uncertainty. The HRes prediction for GF production is combined with the predictions obtained from the nominal simulation samples for the other production mechanisms. The MINLO HJ prediction for GF production is combined with the predictions obtained from the default simulation samples for the other production mechanisms. The MINLO HJ prediction is normalized to the LHC-XS prediction [87] using a scaling factor (K_{ggF}) of 1.54. XH denotes the small contribution from VBF, $VH \equiv WH + ZH$ and $t\bar{t}H$ production and is presented separately as a dashed green line.

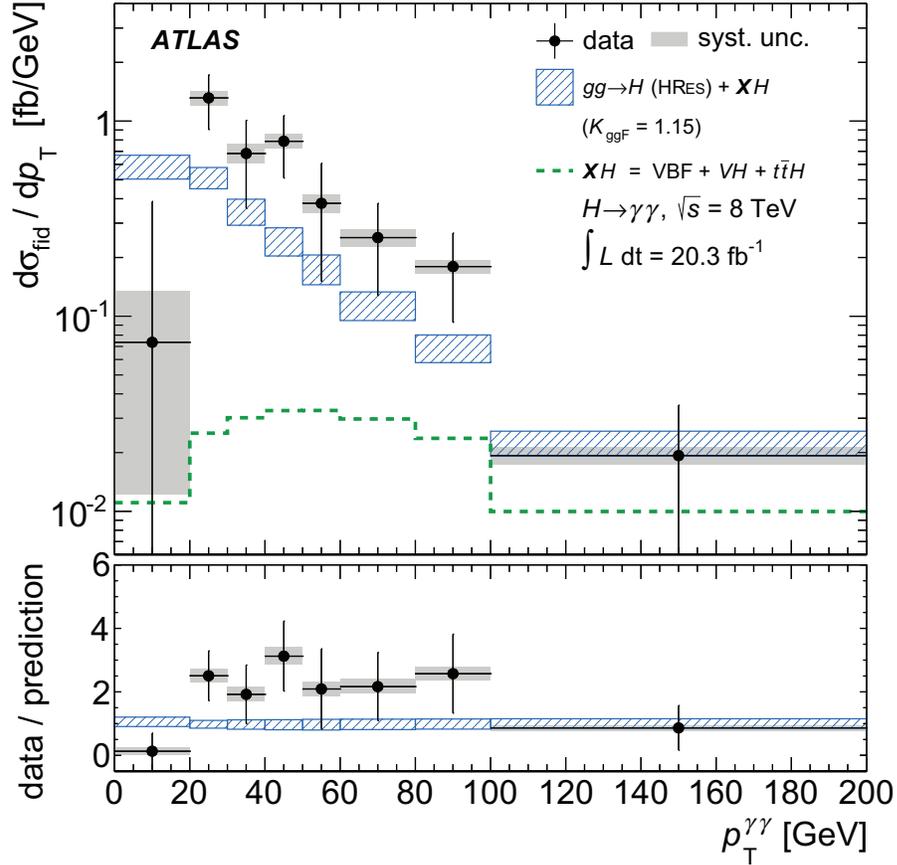


Figure 10.4: The differential cross section for $pp \rightarrow H \rightarrow \gamma\gamma$ as a function of the diphoton transverse momentum, $p_T^{\gamma\gamma}$ [63]. Black circles represent the data and the error bar on each data point is the total measurement uncertainty. The HRes prediction for GF production is combined with the predictions obtained from the nominal simulation samples for the other production mechanisms. The HRes predictions are normalized to the total LHC-XS cross section [87] using a scaling factor (K_{ggF}) of 1.15. XH denotes the small contribution from VBF, $VH \equiv WH + ZH$ and $t\bar{t}H$ production and is presented separately as a dashed green line.

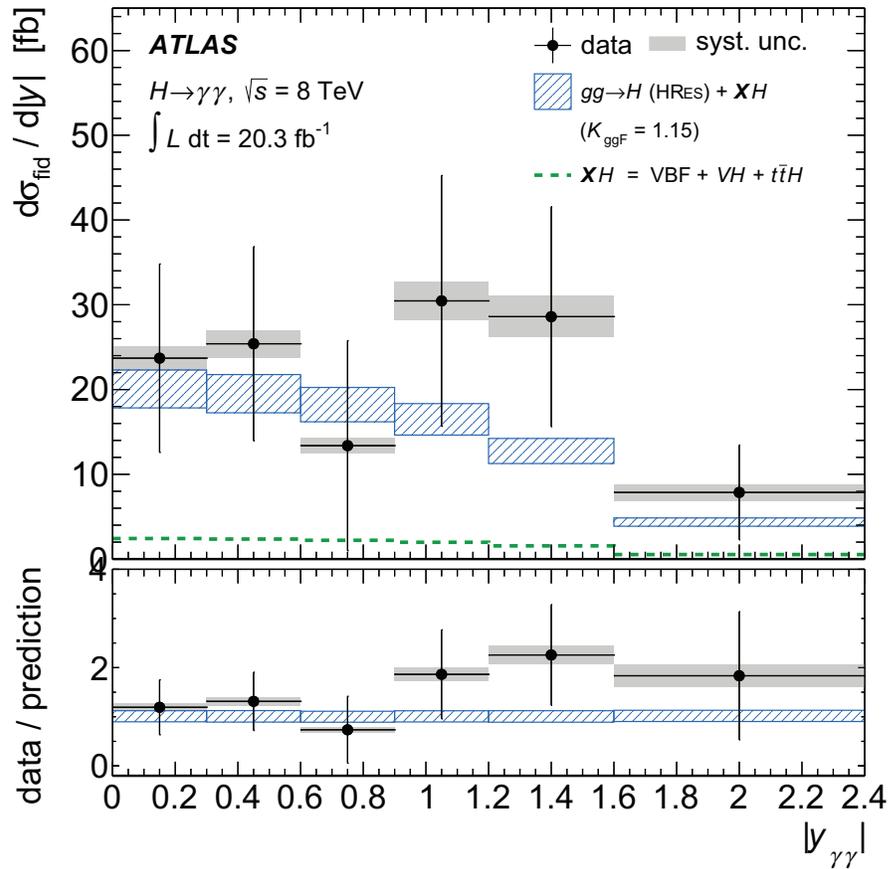


Figure 10.5: The differential cross section for $pp \rightarrow H \rightarrow \gamma\gamma$ as a function of the absolute rapidity of the diphoton system, $|y_{\gamma\gamma}|$ [63]. Black circles represent the data and the error bar on each data point is the total measurement uncertainty. The data and theoretical predictions are presented the same way as in Figure 10.3. The theoretical cross section is also constructed in the same way as described in the caption of Figure 10.4.

Chapter 11

Summary

In 2012, both ATLAS and CMS announced the discovery of a new particle consistent with the expectations for a Standard Model (SM) Higgs boson. In order to fully understand the exact nature of the newly discovered Higgs boson and the role it plays in generating the masses of the elementary particles, detailed investigation of all its production and decay modes, as well as precision measurements of all of its properties must be performed. A large assortment of first measurements based on the Run-I ATLAS and CMS datasets have been published in high peer-reviewed journals, with no significant deviations from the Standard Model expectations observed.

The diphoton decay channel served as a highly important discovery channel due to its very clean decay signature and narrow mass resolution, providing efficient event selection and sufficient ability to discriminate against backgrounds. These same features also make the $H \rightarrow \gamma\gamma$ channel ideal for studying the cross section of the Higgs boson. Cross section measurements provide a powerful test of the consistency of the data with SM predictions and

are designed to be as model independent as possible facilitating the comparison to a variety of theoretical predictions. Cross section measurements also provide an alternative method to study the Higgs boson. This thesis has focused on exploiting the ATLAS Run-I dataset to measure fiducial and differential cross sections of the Higgs boson in the diphoton decay channel, with the 8 TeV results representing the first published measurements of Higgs boson cross sections.

The measurement of fiducial cross sections of the Higgs boson in phase-space regions characterized by the amount of hadronic jet activity or enriched with VBF-like events requires a precise understanding of the jet energy scale (JES). The forward JES is particularly important in the study of VBF-like events which are characterized by the presence of two forward tagging jets. To obtain a precise measurement of the JES and its associated uncertainty dedicated simulation based and *in-situ* techniques were employed. The forward JES was investigated using the η -intercalibration procedure to probe the calorimeters' response to jets as a function of pseudorapidity. A residual calibration that is applied to jets in data was derived along with its associated uncertainty. All ATLAS analyses involving jets apply these calibrations.

In the $H \rightarrow \gamma\gamma$ analysis the Higgs boson signal was extracted by performing a s+b fit to the diphoton invariant mass spectrum ($m_{\gamma\gamma}$). The $m_{\gamma\gamma}$ spectrum is dominated by non-resonant SM backgrounds. The dominant non-resonant SM background processes that contribute to the $m_{\gamma\gamma}$ spectrum are generic diphoton production ($\gamma\gamma$), photon plus jet production (γ +jet) where one jet is mis-identified as a photon, and dijet production (jet+jet) where both jets are mis-identified as photons. The Higgs boson signal appears as an excess of events on top of the non-resonant SM backgrounds, in the region of the

Higgs boson mass. Obtaining an accurate background parameterization that introduces the minimum possible bias on the extracted signal is therefore highly important. In each phase-space region, or bin of a differential distribution, the form of the background function was chosen to minimize the bias observed in the extracted yield when fitting a background-only distribution constructed from $\gamma\gamma$, γ -jet, and jet-jet simulated samples. The simulated samples were normalized using data-driven scale factors determined using the $2 \times 2D$ Sideband Method. A $2 \times 2D$ Sideband Method was implemented to extract the background rates in the Higgs boson signal region by extrapolating the background rates measured in designated control regions. The control regions were defined by inverting the isolation and tight identification criteria for each photon. The relative composition of each background process was determined in the inclusive fiducial region and also as a function of the hadronic jet multiplicity. Each fiducial region or bin of a differential distribution received a dedicated background parameterization.

The Higgs boson signal was extracted and the backgrounds were subtracted by performing an un-binned maximum likelihood fit to the $m_{\gamma\gamma}$ spectrum. The yields were then corrected for experimental effects such as detector inefficiencies and resolution through the application of a model independent unfolding procedure. The s+b fit incorporated uncertainties on the photon energy scale and resolution, as well as the uncertainty on the peak position due to the chosen background parameterization. Additional uncertainties that had no impact on the shape of the signal PDF were not included in the fit and were treated as an uncertainty on the unfolding factors. The unfolded yields were then used to obtain model-independent cross section measurements at the particle-level, therefore, the measurements do not rely on the description of the detector and can be compared to a variety

of theoretical predictions. Additionally, fiducial acceptance factors and non-perturbative correction factors were derived in order to translate inclusive parton-level predictions to predictions at the particle-level and in the fiducial regions. This facilitates comparisons to theoretical predictions that do not decay the Higgs boson.

Higgs boson production cross section measurements were presented in the diphoton decay channel for proton-proton collisions recorded at centre-of-mass energies of $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV. The data recorded at 7 TeV and 8 TeV correspond to integrated luminosities of 4.5 fb^{-1} and 20.3 fb^{-1} , respectively. The inclusive fiducial cross sections measured at 7 TeV and 8 TeV for a Higgs boson of mass 125.36 GeV are:

$$\sigma_{\text{fid}}(pp \rightarrow H \rightarrow \gamma\gamma) = 45.3 \pm 16.9 \text{ (stat.) } {}_{-5.3}^{+5.3} \text{ (syst.) } \pm 0.8 \text{ (lumi.) fb,}$$

$$\sigma_{\text{fid}}(pp \rightarrow H \rightarrow \gamma\gamma) = 43.2 \pm 9.4 \text{ (stat.) } {}_{-2.9}^{+3.2} \text{ (syst.) } \pm 1.2 \text{ (lumi.) fb,}$$

respectively. Four additional fiducial cross sections and two cross-section limits were also presented for the 8 TeV dataset. Additionally, two differential cross sections probing the Higgs boson kinematics were reported within the inclusive fiducial volume. Higgs boson production in association with jets was probed by measuring the cross section in the fiducial regions characterized by the inclusive amount of hadronic activity and also through the differential N_{jet} distribution. Many additional differential distributions were measured using the 8 TeV data but were not reported here. The interested reader is directed to the publication that reported the full ensemble of results [63]. In all cases, the data are in agreement with the expectations for a Standard Model Higgs boson within the precision of the measurements.

The LHC began its second operational run (Run-II) in March 2015 after a scheduled two

year long shutdown in order to perform maintenance and upgrades. In June 2015 the LHC began delivering physics data at a record breaking centre-of-mass energy of $\sqrt{s} = 13$ TeV. By the end of 2015, the LHC delivered more than 4 fb^{-1} of 13 TeV data to the ATLAS experiment. This dataset is not large enough to improve on the precision of the cross section measurements performed with the 7 TeV and 8 TeV data samples, however it does provide the opportunity to make a preliminary measurement of the Higgs boson production cross section at $\sqrt{s} = 13$ TeV. The total Higgs boson cross section measured at 7, 8, and 13 TeV obtained from the $H \rightarrow \gamma\gamma$ analysis channel were combined with complementary measurements made in the $H \rightarrow ZZ^* \rightarrow 4\ell$ channel (see References [88, 89]). The total Higgs boson cross section is obtained by dividing the fiducial cross sections by the acceptance factors presented in this thesis, and by the branching ratio for $H \rightarrow \gamma\gamma$. It should be noted that the 7 TeV and 8 TeV fiducial cross sections presented in this thesis were updated to correspond to the combined ATLAS+CMS Higgs boson mass measurement $m_H = 125.09$ GeV. The preliminary total Higgs boson production cross sections are found to be 34_{-10}^{+11} pb, $33.3_{-5.5}^{+5.8}$ pb, and 24_{-18}^{+21} pb for centre-of-mass energies of 7, 8 and 13 TeV, respectively. Figure 11.1 compares the preliminary cross section measurements at 7, 8, and 13 TeV with the SM predictions. The preliminary results presented here are a good example of the legacy of the research presented in this doctoral dissertation.

The aim of Run-II of the LHC is to collect approximately 100 fb^{-1} of data by the end of 2018. Using this enlarged Run-II dataset, the properties of the Higgs boson will be measured with a much greater precision than was possible in Run-I, which will allow testing the consistency of the data with theoretical predictions at a much higher precision than currently available. The Run-II dataset will also enable the measurement of rare

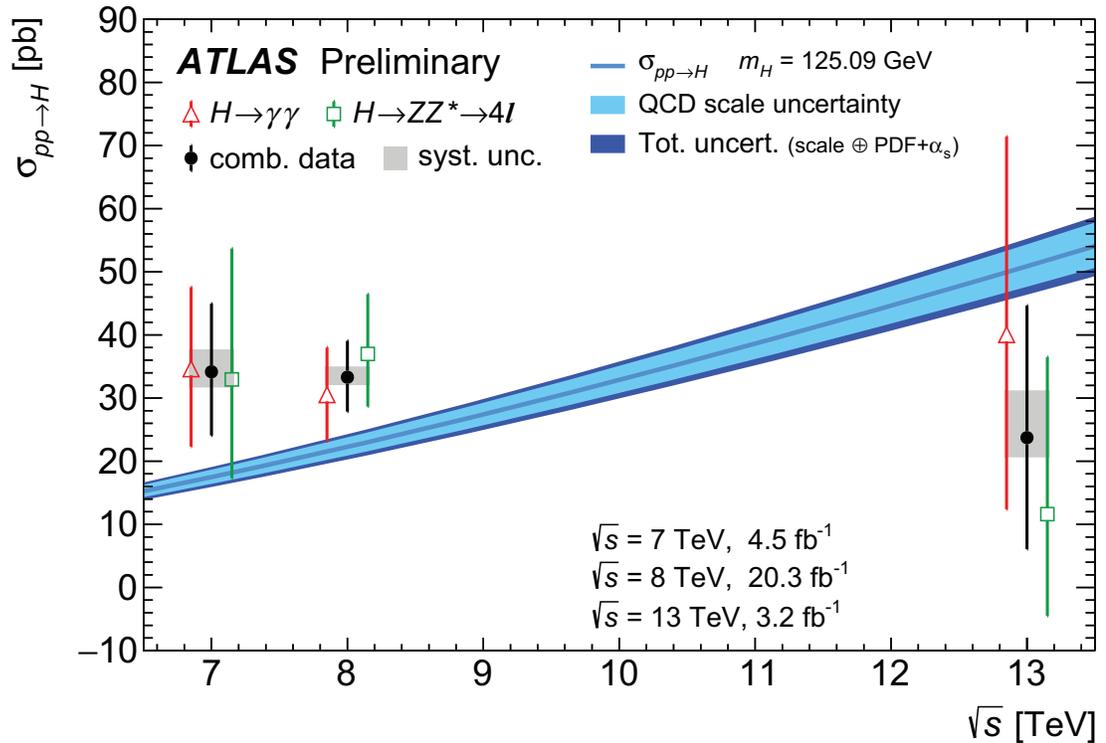


Figure 11.1: Total $pp \rightarrow H$ cross sections measured at different center-of-mass energies compared to LHC-XS theory predictions at up to NNLO in QCD [89]. The red markers are the measurements from the $H \rightarrow \gamma\gamma$ channel, the green markers are the measurements from the $H \rightarrow ZZ^* \rightarrow 4\ell$ channel, and the black markers are the combinations of these two channels. The grey bands on the combined measurements represent the systematic uncertainty, while the black lines are the total uncertainty. Theory predictions refer to a Higgs boson mass of 125.09 GeV, with the light (dark) blue band representing the QCD scale (total) uncertainty. The smooth curve of the SM prediction is obtained by applying a third-order polynomial fit to the values available in Reference [9]. The theoretical uncertainties, described by the blue band, are partially correlated between different values of the centre-of-mass energy.

production and decay modes of the Higgs boson that have yet to be observed. All these measurements are required in order to confirm whether the newly discovered particle is indeed the Higgs boson of the Standard Model or if it is better described by theories beyond the Standard Model.

Appendix A

Simulated samples used in the $H \rightarrow \gamma\gamma$ analysis

A.1 Nominal simulated signal samples

The production and decay of a Higgs boson are simulated for each of the dominant LHC production modes: gluon fusion (GF), vector boson fusion (VBF), associated production with a weak vector boson (WH , ZH), and associated production with a top anti-top pair ($t\bar{t}H$). Individual samples are generated for Higgs boson masses spanning the range 100 GeV to 160 GeV in steps of 5 GeV. The samples corresponding to a Higgs boson mass of 125 GeV are generated with high statistics and are used to obtain the detector correction (unfolding) factors.

GF events are simulated with the POWHEG [90] event generator interfaced with PYTHIA8 [91] (denoted POWHEG+PYTHIA8) for the implementation of underlying event, parton shower-

ing, and hadronization. POWHEG is tuned to match predictions including finite mass effects and soft-gluon re-summations, calculated at next-to-next-to-leading logarithmic order (NNLL) [75, 76, 77]. Calculations at next-to-next-to-leading order (NNLO) [69, 70, 71, 72, 73, 74] in QCD for the GF process are used to obtain the overall normalization of the POWHEG+PYTHIA8 samples. Electroweak corrections calculated at next-to-leading order (NLO) are also included [78, 79].

VBF events are also generated using POWHEG+PYTHIA8 [92]. A cross section calculated with full NLO QCD and EW corrections [93, 94, 95] with approximate NNLO QCD corrections [96] is used to normalize the VBF cross section.

PYTHIA8 (including parton showering, modelling of multiple parton interactions, and hadronization) is used to generate samples containing Higgs bosons produced in association with a W or Z boson, or $t\bar{t}$ pair. The W/ZH predictions are normalized to cross sections calculated at NLO [97] and at NNLO [98] including NLO electroweak radiative corrections [99]. The $t\bar{t}H$ samples are normalized to calculations including full NLO QCD corrections [100, 101, 102, 103].

Samples generated using POWHEG utilize the CT10 [104] parton distribution function (PDF) set, while samples generated using PYTHIA8 utilize the CTEQ6L1 [105] PDF set. A summary of the Higgs boson cross sections and branching ratios was given in Section 2.3 and Section 2.4 with further details provided in References [106, 107, 108], with the uncertainties detailed in References [109, 77].

A.2 Additional simulated signal samples

Parton-level diphoton p_T and rapidity distributions generated at NNLO, with NNLL re-summation terms, using HRes2.2 [110, 111] are used to compare unfolded Higgs boson p_T and rapidity distributions for use in evaluating the generator modelling uncertainty as described in Section 9.3.3.3. The HRes2.2 simulation includes the dependence on the finite quark mass effects up to NLO. Parton-showering was not implemented for these samples.

Unfolded data distributions in fiducial regions defined by the hadronic jet activity are compared to events generated for GF Higgs boson production in association with one and two jets. Higgs boson plus one jet (HJ) and Higgs boson plus two jets (HJJ) [82] events are generated using the processes of POWHEG with MINLO (Multi-scale improved NLO) [83, 84]. HJ events are accurate at NLO while HJJ events are accurate at NLO. Evaluation of the Born cross section term is implemented in MINLO using a standard procedure for merging parton showers with QCD matrix elements (referred to as the CKKW method, named after the authors who first proposed the procedure) [49, 112, 113]. Events produced with MINLO use the CT10 parton distribution function [104] and the events are showered using PYTHIA8 as described above.

Additional GF samples were generated using SHERPA v1.4.3 [114] for assessing the systematic uncertainty associated with the kinematic distributions in GF production. The SHERPA samples implement the CT10 parton density functions [104] and the authors' default tune for multiple parton interactions. SHERPA simulates Higgs boson production with up to four partons at leading order in the matrix element. The fully hadronic final state is produced using SHERPA's models for hadronization and multiple parton interaction, implementing the

CKKW method for matrix element / parton shower matching.

A.3 Simulated background samples

High statistics prompt diphoton background samples are generated using the SHERPA and PYTHIA8 generators previously described. Events are generated from SM processes that result in two photons in the final state but do not involve the production and decay of a Higgs boson. These samples are required for use in the data-driven method for estimating the background rates (see Chapter 7).

Bibliography

- [1] ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the {ATLAS} detector at the {LHC}. *Physics Letters B*, 716(1):1 – 29, 2012.
- [2] CMS Collaboration. Observation of a new boson at a mass of 125 gev with the {CMS} experiment at the {LHC}. *Physics Letters B*, 716(1):30 – 61, 2012.
- [3] Cliff Burgess and Guy Moore. *The Standard Model, A Primer*. Cambridge University Press, Shaftesbury Road, Cambridge, CB2 8BS, United Kingdom, 2007.
- [4] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Westview Press, a member of the Perseus Books Group, Boston, MA, US, 1995.
- [5] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.
- [6] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

-
- [7] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.*, 13:585–587, Nov 1964.
- [8] K. A. Olive et al. Review of Particle Physics. *Chin. Phys.*, C38:090001, 2014.
- [9] LHC Higgs Cross Section Working Group, S. Heinemeyer, C. Mariotti, G. Passarino, and R. Tanaka (Eds.). Handbook of LHC Higgs Cross Sections: 3. Higgs Properties. *CERN-2013-004*, CERN, Geneva, 2013. <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG>.
- [10] ATLAS Collaboration. Evidence for the higgs-boson yukawa coupling to tau leptons with the atlas detector. *Journal of High Energy Physics*, 2015(4), 2015.
- [11] The CMS Collaboration. Evidence for the direct decay of the 125 gev higgs boson to fermions. *Nat Phys*, 10:557–560, 2014.
- [12] ATLAS Collaboration and CMS Collaboration. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett.*, 114:191803, May 2015.
- [13] ATLAS Collaboration. Evidence for the spin-0 nature of the higgs boson using atlas data. *Physics Letters B*, 726(1-3):120–144, 2013.
- [14] ATLAS Collaboration and 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. Technical Report ATLAS-CONF-2015-044, CERN, Geneva, Sep 2015.

- [15] Lyndon R Evans and Philip Bryant. LHC Machine. *J. Instrum.*, 3:S08001. 164 p, 2008. This report is an abridged version of the LHC Design Report (CERN-2004-003).
- [16] The ATLAS Collaboration. The atlas experiment at the cern large hadron collider. *Journal of Instrumentation*, 3(08):S08003, 2008.
- [17] The CMS Collaboration. The cms experiment at the cern lhc. *Journal of Instrumentation*, 3(08):S08004, 2008.
- [18] The LHCb Collaboration. The lhcb detector at the lhc. *Journal of Instrumentation*, 3(08):S08005, 2008.
- [19] The ALICE Collaboration. The alice experiment at the cern lhc. *Journal of Instrumentation*, 3(08):S08002, 2008.
- [20] J.J. Goodson. *Search for Supersymmetry in States with Large Missing Transverse Momentum and Three Leptons including a Z-Boson*. PhD thesis, Stony Brook University, May 2012. Presented 17 Apr 2012.
- [21] M Bajko et al. Report of the Task Force on the Incident of 19th September 2008 at the LHC. Technical Report LHC-PROJECT-Report-1168. CERN-LHC-PROJECT-Report-1168, CERN, Geneva, Mar 2009.
- [22] ATLAS Collaboration. ATLAS luminosity public results. Technical report. <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults>.

- [23] The ATLAS collaboration. Operation and performance of the atlas semiconductor tracker. *Journal of Instrumentation*, 9(08):P08009, 2014.
- [24] Performance of the atlas trigger system in 2010. *The European Physical Journal C*, 72(1), 2012.
- [25] S. Agostinelli et al. GEANT4: a simulation toolkit. *Nucl. Instrum. Methods*, 506:250–303, 2003.
- [26] J. Allison et al. GEANT4 Developments and Applications. *IEEE Trans. Nucl. Sci.*, 53:270–278, 2006.
- [27] D. Costanzo, A. Dell’Acqua, A Di Simone, M. Gallas, A. Nairz, A. Rimoldi, J. Boudreau, and V. TSulaia. ATLAS detector simulation. (ATL-SOFT-PUB-2005-004, CERN-ATL-SOFT-PUB-2005-004, ATL-COM-SOFT-2005-008), 2005.
- [28] A Rimoldi and the ATLAS Collaboration. Simulation strategies for the atlas experiment at lhc. *Journal of Physics: Conference Series*, 331(3):032026, 2011.
- [29] The Optimization of ATLAS Track Reconstruction in Dense Environments. Technical Report ATL-PHYS-PUB-2015-006, CERN, Geneva, Mar 2015.
- [30] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti- k_t jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063, 2008.
- [31] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *Eur. Phys. J.*, C72:1896, 2012.

-
- [32] Matteo Cacciari and Gavin P. Salam. Dispelling the N^3 myth for the k_t jet-finder. *Phys. Lett.*, B641:57–61, 2006.
- [33] T. Barillari et al. Local Hadronic Calibration. Technical Report ATL-LARG-PUB-2009-001-2, CERN, Geneva, Jun 2008.
- [34] ATLAS Collaboration. Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Eur. Phys. J.*, C75:17, 2015.
- [35] Commissioning of the atlas muon spectrometer with cosmic rays. *The European Physical Journal C*, 70(3):875–916, 2010.
- [36] Measurement of the muon reconstruction performance of the atlas detector using 2011 and 2012 lh proton-proton collision data. *The European Physical Journal C*, 74(11), 2014.
- [37] Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS. Technical Report ATLAS-CONF-2015-037, CERN, Geneva, Aug 2015.
- [38] ATLAS Collaboration. Jet global sequential corrections with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV. Technical Report ATLAS-CONF-2015-002, CERN, Geneva, Mar 2015.
- [39] Paolo Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 0411:040, 2004.

-
- [40] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.
- [41] Simone Alioli et al. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010.
- [42] T. Gleisberg, Stefan. Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al. Event generation with SHERPA 1.1. *JHEP*, 0902:007, 2009.
- [43] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput. Phys. Commun.*, 178:852–867, 2008.
- [44] Stefan Gieseke et al. Herwig++ 1.0: An Event generator for e+ e- annihilation. *JHEP*, 02:005, 2004.
- [45] J. Bellm et al. Herwig++ 2.7 Release Note. (IPPP-13-88, MCNET-13-15, DCPT-13-176, DESY-13-186, KA-TP-31-2013, –ZU-TH-23-13), 2013.
- [46] Simone Alioli et al. Jet pair production in POWHEG. *JHEP*, 1104:081, 2011.
- [47] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, et al. New parton distributions for collider physics. *Phys. Rev.*, D82:074024, 2010.
- [48] ATLAS Collaboration. New ATLAS event generator tunes to 2010 data. Technical Report ATL-PHYS-PUB-2011-008, CERN, Geneva, Apr 2011.
- [49] S. Catani, F. Krauss, R. Kuhn, and B.R. Webber. QCD matrix elements + parton showers. *JHEP*, 0111:063, 2001.

-
- [50] Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of $\sqrt{s} = 7$ TeV. Technical Report ATLAS-CONF-2010-038, CERN, Geneva, Jul 2010.
- [51] ATLAS Collaboration. Pile-up subtraction and suppression for jets in ATLAS. Technical Report ATLAS-CONF-2013-083, CERN, Geneva, Aug 2013.
- [52] Victor Lendermann et al. Combining Triggers in HEP Data Analysis. *Nucl. Instrum. Meth., A* 604:707–718, 2009.
- [53] Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector. *Eur. Phys. J., C* 73:2306, 2013.
- [54] Data-driven determination of the energy scale and resolution of jets reconstructed in the ATLAS calorimeters using dijet and multijet events at $\sqrt{s} = 8$ TeV. Technical Report ATLAS-CONF-2015-017, CERN, Geneva, Apr 2015.
- [55] ATLAS Collaboration. Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS. Technical Report ATLAS-CONF-2015-037, CERN, Geneva, Aug 2015.
- [56] ATLAS Collaboration. In situ jet pseudorapidity intercalibration of the ATLAS detector using dijet events in $\sqrt{s} = 7$ TeV proton-proton 2011 data. Technical Report ATLAS-CONF-2012-124, CERN, Geneva, Aug 2012.
- [57] ATLAS Collaboration. Improved luminosity determination in pp collisions at $\sqrt{s} = 7$

- TeV using the ATLAS detector at the LHC. *The European Physical Journal C*, 73(8), 2013.
- [58] The ATLAS Collaboration. Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data. *Eur.Phys.J.*, C74, 2014.
- [59] 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. D*, 83:052005, Mar 2011.
- [60] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Phys. Lett. B*, 659:119 – 126, 2008.
- [61] ATLAS Collaboration. Measurement of the isolated diphoton cross-section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Phys. Rev. D* 85, 012003, 2012.
- [62] The 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. D*, 90:112015, Dec 2014.
- [63] The 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. *Journal of High Energy Physics*, 2014(9), 2014.
- [64] The HSG1 Group. Measurement of fiducial and differential cross sections of the Higgs boson in the diphoton decay channel using centre-of-mass energy of 8 TeV. Technical Report ATL-COM-PHYS-2013-1664, CERN, Geneva, Dec 2013.

- [65] N Lorenzo Martinez and R Turra. Mass measurement in the $H \rightarrow \gamma\gamma$ channel: Supporting documentation for the Mass Paper. Technical Report ATL-COM-PHYS-2014-018, CERN, Geneva, Jan 2014.
- [66] ATLAS Collaboration. Measurement of the Higgs boson mass from the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ channels in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector. *Phys. Rev. D*, 90:052004, Sep 2014.
- [67] The HSG1 Group. Coupling measurement with the $H \rightarrow \gamma\gamma$ channel : Supporting document. Technical Report ATL-COM-PHYS-2014-025, CERN, Geneva, Jan 2014.
- [68] ATLAS Collaboration. Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with ATLAS 2011 data. *ATLAS-CONF-2013-004*, 2013.
- [69] A. Djouadi, M. Spira, and P.M. Zerwas. Production of Higgs bosons in proton colliders: QCD corrections. *Phys. Lett.*, B264:440–446, 1991.
- [70] S. Dawson. Radiative corrections to Higgs boson production. *Nucl. Phys.*, B359:283–300, 1991.
- [71] M. Spira, A. Djouadi, D. Graudenz, and P.M. Zerwas. Higgs boson production at the LHC. *Nucl. Phys.*, B453:17–82, 1995.
- [72] Robert V. Harlander and William B. Kilgore. Next-to-next-to-leading order Higgs production at hadron colliders. *Phys. Rev. Lett.*, 88:201801, 2002.

-
- [73] Charalampos Anastasiou and Kirill Melnikov. Higgs boson production at hadron colliders in NNLO QCD. *Nucl. Phys.*, B646:220–256, 2002.
- [74] 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.*, B665:325–366, 2003.
- [75] Stefano Catani, Daniel de Florian, Massimiliano Grazzini, and Paolo Nason. Soft gluon resummation for Higgs boson production at hadron colliders. *JHEP*, 0307:028, 2003.
- [76] Daniel de Florian, Giancarlo Ferrera, Massimiliano Grazzini, and Damiano Tommasini. Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC. *JHEP*, 1111:064, 2011.
- [77] LHC Cross Section Working Group. Handbook of LHC Higgs Cross Sections: 2. Differential Distributions. Technical Report arXiv:1201.3084. CERN-2012-002, Geneva, 2012.
- [78] U. Aglietti, R. Bonciani, G. Degrossi, and A. Vicini. Two loop light fermion contribution to Higgs production and decays. *Phys. Lett.*, B595:432–441, 2004.
- [79] Stefano Actis, Giampiero Passarino, Christian Sturm, and Sandro Uccirati. NLO electroweak corrections to Higgs boson production at hadron colliders. *Phys. Lett.*, B670:12–17, 2008.
- [80] Iain W. Stewart, Frank J. Tackmann, Jonathan R. Walsh, and Saba Zuberi. Jet p_T

- resummation in Higgs production at NNLL' + NNLO. *Phys. Rev. D*, 89:054001, Mar 2014.
- [81] Daniel de Florian, Giancarlo Ferrera, Massimiliano Grazzini, and Damiano Tommasini. Higgs boson production at the LHC: transverse momentum resummation effects in the $H \rightarrow \gamma\gamma$, $H \rightarrow WW \rightarrow l\nu l\nu$ and $H \rightarrow ZZ \rightarrow 4l$ decay modes. *Journal of High Energy Physics*, 2012(6), 2012.
- [82] John M. Campbell, R. Keith Ellis, Rikkert Frederix, Paolo Nason, Carlo Oleari, et al. NLO Higgs Boson Production Plus One and Two Jets Using the POWHEG BOX, MadGraph4 and MCFM. *JHEP*, 1207:092, 2012.
- [83] Keith Hamilton, Paolo Nason, Carlo Oleari, Giulia Zanderighi. Merging H/W/Z + 0 and 1 jet at NLO with no merging scale: a path to parton shower + NNLO matching. 2012.
- [84] Keith Hamilton, Paolo Nason, and Giulia Zanderighi. MINLO: Multi-Scale Improved NLO. *JHEP*, 1210:155, 2012.
- [85] Radja Boughezal, Xiaohui Liu, Frank Petriello, Frank J. Tackmann, and Jonathan R. Walsh. Combining resummed Higgs predictions across jet bins. *Phys. Rev. D*, 89:074044, Apr 2014.
- [86] Andrea Banfi, Pier Francesco Monni, Gavin P. Salam, and Giulia Zanderighi. Higgs- and Z-boson Production with a Jet Veto. *Phys. Rev. Lett.*, 109:202001, Nov 2012.
- [87] S. Heinemeyer et al. Handbook of LHC Higgs Cross Sections: 3. Higgs Proper-

- ties: Report of the LHC Higgs Cross Section Working Group. Technical Report arXiv:1307.1347. CERN-2013-004, Geneva, 2013.
- [88] The ATLAS Collaboration. Measurement of the Higgs boson production cross section at 7, 8 and 13 TeV center-of-mass energies in the $H \rightarrow \gamma\gamma$ channel with the ATLAS detector. Technical Report ATLAS-CONF-2015-060, CERN, Geneva, Dec 2015.
- [89] The ATLAS Collaboration. Measurements of the total cross sections for Higgs boson production combining the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ decay channels at 7, 8 and 13 TeV center-of-mass energies with the ATLAS detector. Technical Report ATLAS-CONF-2015-069, CERN, Geneva, Dec 2015.
- [90] S. Alioli, P. Nason, C. Oleari and E. Re. NLO Higgs boson production via gluon fusion matched with shower in POWHEG. *JHEP*, 04:002, 2009.
- [91] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. A Brief Introduction to PYTHIA 8.1. *Comput.Phys.Commun.*, 178:852–867, 2008.
- [92] P. Nason and C. Oleari. NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG. *JHEP*, 02:037, 2010.
- [93] M. Ciccolini, Ansgar Denner, and S. Dittmaier. Strong and electroweak corrections to the production of Higgs + 2-jets via weak interactions at the LHC. *Phys. Rev. Lett.*, 99:161803, 2007.
- [94] Mariano Ciccolini, Ansgar Denner, and Stefan Dittmaier. Electroweak and QCD

- corrections to Higgs production via vector-boson fusion at the LHC. *Phys. Rev.*, D77:013002, 2008.
- [95] K. Arnold, M. Bahr, Giuseppe Bozzi, F. Campanario, C. Englert, et al. VBFNLO: A Parton level Monte Carlo for processes with electroweak bosons. *Comput. Phys. Commun.*, 180:1661–1670, 2009.
- [96] Paolo Bolzoni, Fabio Maltoni, Sven-Olaf Moch, and Marco Zaro. Higgs production via vector-boson fusion at NNLO in QCD. *Phys. Rev. Lett.*, 105:011801, 2010.
- [97] Tao Han and S. Willenbrock. QCD correction to the $pp \rightarrow WH$ and ZH total cross-sections. *Phys. Lett.*, B273:167–172, 1991.
- [98] Oliver Brein, Abdelhak Djouadi, and Robert Harlander. NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders. *Phys. Lett.*, B579:149–156, 2004.
- [99] M.L. Ciccolini, S. Dittmaier, and M. Krämer. Electroweak radiative corrections to associated WH and ZH production at hadron colliders. *Phys. Rev.*, D68:073003, 2003.
- [100] W. Beenakker, S. Dittmaier, M. Kramer, B. Plumper, M. Spira, et al. Higgs radiation off top quarks at the Tevatron and the LHC. *Phys. Rev. Lett.*, 87:201805, 2001.
- [101] W. Beenakker, S. Dittmaier, M. Kramer, B. Plumper, M. Spira, et al. NLO QCD corrections to $t\bar{t}H$ production in hadron collisions. *Nucl. Phys.*, B653:151–203, 2003.

-
- [102] S. Dawson, L.H. Orr, L. Reina, and D. Wackerth. Next-to-leading order QCD corrections to $pp \rightarrow t\bar{t}h$ at the CERN Large Hadron Collider. *Phys. Rev.*, D67:071503, 2003.
- [103] S. Dawson, C. Jackson, L.H. Orr, L. Reina, and D. Wackerth. Associated Higgs production with top quarks at the Large Hadron Collider: NLO QCD corrections. *Phys. Rev.*, D68:034022, 2003.
- [104] H.-L. Lai, et al. New parton distributions for collider physics. *Phys. Rev. D*, 82:074024, 2010.
- [105] P.M. Nadolsky, et al. Implications of CTEQ global analysis for collider observables. *Phys. Rev. D*, 78:013004, 2008.
- [106] A. Djouadi, J. Kalinowski, and M. Spira. HDECAY: A Program for Higgs boson decays in the Standard Model and its supersymmetric extension. *Comput. Phys. Commun.*, 108:56–74, 1998.
- [107] A. Bredenstein, Ansgar Denner, S. Dittmaier, and M.M. Weber. Precise predictions for the Higgs-boson decay $H \rightarrow WW/ZZ \rightarrow 4$ leptons. *Phys. Rev.*, D74:013004, 2006.
- [108] Stefano Actis, Giampiero Passarino, Christian Sturm, and Sandro Uccirati. NNLO computational techniques: The cases $H \rightarrow \gamma\gamma$ and $H \rightarrow gg$. *Nucl. Phys.*, B811:182–273, 2009.
- [109] LHC Cross Section Working Group. Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. Technical Report arXiv:1101.0593v3, Geneva, 2012.

-
- [110] D. de Florian, G. Ferrera, M. Grazzini, and D. Tommasini. Higgs boson production at the LHC: transverse momentum resummation effects in the $H \rightarrow \gamma\gamma$, $H \rightarrow WW \rightarrow l\nu l\nu$ and $H \rightarrow ZZ \rightarrow 4l$ decay modes. *JHEP*, 1206:132, 2012.
- [111] Massimiliano Grazzini and Hayk Sargsyan. Heavy-quark mass effects in Higgs boson production at the LHC. *JHEP*, 1309:129, 2013.
- [112] F. Krauss. Matrix elements and parton showers in hadronic interactions. *JHEP*, 0208:015, 2002.
- [113] Stephen Mrenna and Peter Richardson. Matching matrix elements and parton showers with HERWIG and PYTHIA. *JHEP*, 0405:040, 2004.
- [114] T. Gleisberg et al. SHERPA 1. α , a proof-of-concept version. *JHEP*, 0402(056), 2004.