



THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ SAVOIE MONT BLANC

Spécialité : **Physique Subatomique et Astroparticules**

Arrêté ministériel : 25 Mai 2016

Présentée par

Mohamed BELFKIR

Thèse dirigée par **Stéphane JEZEQUEL**

préparée au sein du **Laboratoire d'Annecy de Physique des Particules**
dans l'**École Doctorale de Physique de Grenoble**

Recherche de production de paires de bosons de Higgs au collisionneur LHC (CERN): Premier test du potentiel de Higgs et recherche de nouvelle physique

Search for Higgs pair production at LHC collider (CERN): The first measurement for Higgs potential and search for new physics

Thèse soutenue publiquement le **15 Septembre 2021**,
devant le jury composé de :

Madame Lucia DI CIACCIO

Professeure, LAPP, Président

Monsieur Maarten BOONEKAMP

Cadre scientifique des EPIC, CEA, Rapporteur

Monsieur Roberto SALERNO

Chargeé de recherche, LLR, Rapporteur

Monsieur Arnaud DUPERRIN

Directeur de recherche, CPPM, Membre

Madame Élisabeth PETIT

Chargeée de recherche, CPPM, Membre

Madame Lydia ROOS

Directrice de recherche, LPNHE, Membre



³ Abstract

⁴ From the discovery of the Higgs boson in 2012, most of its properties such as mass, spin, production
⁵ cross-section and its coupling to fermions and bosons have been measured. However, The trilinear self-
⁶ coupling λ_{HHH} of the Higgs boson has not been measured yet. This parameter controls the shape of the
⁷ Higgs potential, explaining the importance of its measurement. Deviation from its Standard Model (SM)
⁸ predicted value would indicate new physics beyond the SM (BSM). Deviations are quantified through
⁹ the κ_λ modifier. At the LHC, it is measured through the rate of the rare Higgs boson pair production
¹⁰ (HH) process, which is the only direct way to access it. This thesis presents the search for Higgs boson
¹¹ pair production in its $H(\rightarrow \gamma\gamma)H(\rightarrow b\bar{b})$ decay channel. The search is performed with the full Run-2
¹² proton-proton collision data collected with the ATLAS detector at the LHC between 2015 and 2018 at a
¹³ centre-of-mass energy of $\sqrt{s} = 13$ TeV. This amount of data corresponding to an integrated luminosity
¹⁴ of 139 fb^{-1} . Despite low statistics, the $b\bar{b}\gamma\gamma$ channel profits from a very clean experimental signature,
¹⁵ an excellent $m_{\gamma\gamma}$ mass resolution, a smooth background and a large $H \rightarrow b\bar{b}$ branching ratio. Therefore,
¹⁶ It is an excellent channel to perform κ_λ measurement.
¹⁷ Photon objects are so crucial for this channel that work concerning the improvement of their identifi-
¹⁸ cation efficiency is presented in this thesis. A new algorithm based on Convolutional Neural Network
¹⁹ (CNN) is implemented to improve identification efficiency. The CNN algorithm uses images from the
²⁰ electromagnetic calorimeter layers to learn the shower development of prompt photons and separates
²¹ them from photons originating from jet fragmentation or fake photons. In addition, the mis-modelling
²² of the shower shape in the simulation is corrected by redistributing the energy between the cluster cells
²³ to become consistent with the real data.
²⁴ The sensitivity to the $HH \rightarrow b\bar{b}\gamma\gamma$ signal is improved using a specific b -jet calibration technique. The
²⁵ signal is separated from the single Higgs boson and the continuum $\gamma\gamma+\text{jets}$ backgrounds through a
²⁶ Multivariate Analysis (MVA). Additional tools are developed to improve the analysis sensitivity. They
²⁷ are not included in the publication of ATLAS collaboration presented in this thesis, but they would be
²⁸ used for the next analysis. No significant excess of $HH \rightarrow b\bar{b}\gamma\gamma$ events is observed in the data with
²⁹ respect to the expected background. Thus, an upper limit on the HH cross-section is set. The observed
³⁰ (expected) 95% confidence level (CL) limit corresponds to 4.1 (5.5) times the cross-section predicted by
³¹ the SM. The observed (expected) 95% CL limit on the κ_λ modifier is also derived to be [-1.5, 6.7] ([-2.4,
³² 7.7]).
³³ An extrapolation of this result to Run-3 and to High-Luminosity LHC (HL-LHC) is also presented in
³⁴ this thesis.

³⁵ Résumé

³⁶ Depuis la découverte du boson de Higgs en 2012, la plupart de ses propriétés telles que sa masse, son spin,
³⁷ sa section efficace de production et son couplage aux fermions et bosons ont été mesurées. Cependant,
³⁸ la composante trilinéaire de l'auto-couplage du boson de Higgs λ_{HHH} n'a pas encore été mesuré. Ce
³⁹ paramètre contrôle la forme du potentiel de Higgs, donnant de l'importance à sa mesure. Une déviation
⁴⁰ par rapport à la valeur prédite par le Modèle Standard (MS) indiquerait la présence d'une nouvelle
⁴¹ physique au-delà du MS. Les déviations sont quantifiées via le modificateur κ_λ . Au LHC, il est mesuré
⁴² à travers le taux de production de paires de boson de Higgs (HH) qui est un processus rare et aussi le
⁴³ seul moyen direct d'y accéder. Cette thèse présente la recherche de la production de paires de boson
⁴⁴ de Higgs dans le canal de désintégration $H(\rightarrow \gamma\gamma)H(\rightarrow b\bar{b})$. Le travail est effectué avec les données de
⁴⁵ collisions de protons du Run-2 enregistrées à une énergie de $\sqrt{s} = 13$ TeV par le détecteur ATLAS au
⁴⁶ LHC entre 2015 et 2018. Cette quantité de données correspond à une luminosité intégrée de 139 fb^{-1} .
⁴⁷ Malgré le faible nombre d'événements signaux, le canal $b\bar{b}\gamma\gamma$ bénéficie d'une signature expérimentale
⁴⁸ très propre grâce d'une excellente résolution de masse $m_{\gamma\gamma}$, d'un fond lisse et d'un grand rapport de
⁴⁹ branchement $H \rightarrow b\bar{b}$. C'est un canal excellent pour effectuer la mesure de κ_λ .

⁵⁰ Les photons sont tellement cruciaux pour ce canal qu'un travail d'amélioration de leur identification est
⁵¹ présenté dans cette thèse. Un nouvel algorithme basé sur les réseaux de neurones à convolution (CNN)
⁵² est implémenté pour améliorer l'efficacité de l'identification. L'algorithme CNN utilise des images des
⁵³ cellules du calorimètre électromagnétique pour apprendre à reconnaître les gerbes électromagnétiques
⁵⁴ des photons et les séparer des photons provenant de la fragmentation des jets ou des faux photons.
⁵⁵ De plus, la modélisation du développement des gerbes électromagnétiques en simulation est corrigée en
⁵⁶ répartissant l'énergie entre les cellules du calorimètre électromagnétique pour se rapprocher des données
⁵⁷ réelles.

⁵⁸ La sensibilité au signal $HH \rightarrow b\bar{b}\gamma\gamma$ est améliorée en appliquant une méthode d'étalonnage spécifique
⁵⁹ aux jets de b . Le signal est ensuite séparé des bruits de fonds (boson de Higgs et continus $\gamma\gamma$ +jets)
⁶⁰ grâce à une technique d'analyse multivariée (MVA). Des outils supplémentaires sont développés pour
⁶¹ améliorer la sensibilité de l'analyse. Ils ne sont pas inclus dans la publication de la collaboration ATLAS,
⁶² mais sont disponible pour la prochaine analyse. Aucun excès significatif d'événements $HH \rightarrow b\bar{b}\gamma\gamma$ n'est
⁶³ observé dans les données par rapport au bruit de fond attendu. Ainsi, une limite supérieure sur la section
⁶⁴ efficace HH est fixée. La limite observée (attendue) à un niveau de confiance (CL) de 95 % correspond
⁶⁵ à 4,1 (5,5) fois la section efficace prévue par le MS. La limite de 95 % CL observée (attendue) sur le
⁶⁶ modificateur κ_λ est égale à $[-1,5 ; 6,7]$ ($[-2,4 ; 7,7]$).

⁶⁷ Une extrapolation de ce résultat au Run-3 et au LHC Haute-Luminosité (HL-LHC) est également
⁶⁸ présentée dans cette thèse.

69 Contents

70	Abstract	i
71	Résumé	ii
72	1 Introduction	1
73	2 Theoretical Framework	5
74	2.1 Introduction	5
75	2.2 The Standard Model (SM) of particle physics	5
76	2.2.1 Elementary Particles	6
77	2.2.2 Elementary Interactions	7
78	2.3 Electroweak Symmetry Breaking (EWSB)	9
79	2.4 Higgs boson : Production and observation	11
80	2.4.1 Previous searches for Higgs boson	11
81	2.4.2 Proton-proton collisions	12
82	2.4.3 Higgs boson production	13
83	2.4.4 Higgs boson decay channels	14
84	2.4.5 Higgs boson discovery	16
85	2.5 Looking for di-Higgs boson events	19
86	2.5.1 Di-Higgs boson production and decays	19
87	2.5.2 Di-Higgs boson as a probe of BSM physics	22
88	2.5.3 Current measurements	22
89	2.6 Conclusion	25
90	3 The Large Hadron Collider and the ATLAS detector	26
91	3.1 The Large Hadron Collider	26
92	3.1.1 Acceleration chain	27
93	3.1.2 Luminosity	28
94	3.1.3 Pile-up events	30
95	3.2 ATLAS : A Toroidal LHC ApparatuS detector	30
96	3.2.1 System of coordinates	31
97	3.2.2 Inner Tracker	31
98	3.2.3 Calorimeter system	36
99	3.2.4 Muon Spectrometer	40
100	3.2.5 Trigger	41
101	3.3 Upgrade plans towards Run-3 and the HL-LHC	42
102	3.4 Physics objects reconstruction	43

103	3.4.1	Track and vertex reconstruction	43
104	3.4.2	Electron and photon reconstruction	43
105	3.4.3	Muon reconstruction and identification	52
106	3.4.4	Jet reconstruction	54
107	3.4.5	Missing transverse energy	55
108	4	Selection of photons in ATLAS	56
109	4.1	Photon Isolation	56
110	4.2	Photon Identification	56
111	4.3	Shower shape mis-modelling	62
112	4.3.1	Cell-based reweighting correction	62
113	4.3.2	Electron results	62
114	4.3.3	Cell-based reweighting for photons	62
115	4.4	Convolutional Neural Network for Photon Identification	70
116	4.4.1	CNN algorithm	70
117	4.4.2	Photon efficiency estimation : Out-of-sample validation	74
118	4.5	Conclusion	80
119	5	Jet reconstruction and Tagging in ATLAS	82
120	5.1	Reconstruction algorithms	82
121	5.2	Jet Energy Calibration	83
122	5.2.1	Jet Calibration Chain	83
123	5.3	Jet Tagging : Pile-up and flavour taggings	85
124	5.3.1	Jet Vertex Tagging (JVT) algorithms	86
125	5.3.2	b -jet tagging	89
126	5.4	b -jet energy calibration	91
127	5.4.1	μ -in-jet correction	92
128	5.4.2	p_T -Reco correction	92
129	5.4.3	b -jet energy correction improvement	94
130	5.5	Conclusion	95
131	6	Measurement of Higgs boson self-coupling	96
132	6.1	Data and simulation	96
133	6.2	Object and event selection	98
134	6.2.1	Object selection	98
135	6.2.2	Event selection	99
136	6.2.3	Event selection using Deep Neural Network	103
137	6.3	Signal and background modelling	106
138	6.3.1	Signal and single Higgs boson parameterization	106
139	6.3.2	Continuum background parameterization	107
140	6.4	Systematic uncertainties	109
141	6.5	Statistical model	112
142	6.6	Results	112
143	6.6.1	Observed events and $m_{\gamma\gamma}$ fit	112
144	6.6.2	Cross-section limits and κ_λ constraint	114
145	6.7	Comparison to 36 fb^{-1} results	117
146	6.8	Comparison to CMS $HH \rightarrow b\bar{b}\gamma\gamma$ results	118

147	6.9	Comparison to other HH decay channels	120
148	6.9.1	$b\bar{b}b\bar{b}$ channel	120
149	6.9.2	$b\bar{b}\tau^+\tau^-$ channel	121
150	6.9.3	Combination	122
151	6.10	Conclusion	124
152	7	Prospects of Higgs boson pair production at Run-3 and HL-LHC	125
153	7.1	Potential $HH \rightarrow b\bar{b}\gamma\gamma$ improvements from analysis tools	125
154	7.1.1	Photon identification with CNN in $HH \rightarrow b\bar{b}\gamma\gamma$	125
155	7.2	Expected results at the end of Run-3	126
156	7.3	Prospects at HL-LHC	130
157	7.3.1	Estimation for European strategy 2019 of $b\bar{b}\gamma\gamma$	130
158	7.3.2	Combination with $b\bar{b}b\bar{b}$ and $b\bar{b}\tau^+\tau^-$ channels	131
159	7.4	New extrapolation of full Run-2 $HH \rightarrow b\bar{b}\gamma\gamma$ analysis	133
160	7.5	Conclusion	135
161	8	Conclusion	136
162	Bibliography		148
163	A Cell-by-cell reweighting		149
164	A.1	Electron reweighting applied to photons	149
165	A.2	Specific photon reweighting	149
166	B Additional details on Convolutional Neural Network ID		158
167	B.1	Background rejection for CNN Tight and Loose WPs	158
168	B.2	CNN efficiency as a function of $ \eta $	158
169	B.3	CNN efficiency in additional $ \eta $ bins	158
170	B.4	Purity evaluation using the template fit method	158
171	C Future improvement for Convolutional Neural Network ID		166
172	C.1	Low efficiency at high η	166
173	C.2	Multi-Task Cascade Convolutional Network (MTCNN)	166
174	C.3	Pile-Up dependency	168
175	D Kinematic Fit for $HH \rightarrow b\bar{b}\gamma\gamma$		170
176	D.1	Kinematic Fit (KF)	170
177	D.2	HH system balance	171
178	D.3	Likelihood	171
179	D.4	$\sigma_{b\bar{b}\gamma\gamma}$ determination	172
180	D.5	Kinematic Fit results	173
181	D.5.1	HH system balance	173
182	D.5.2	$m_{b\bar{b}}$ resolution	173
183	D.6	Conclusion	173
184	List of Figures		185
185	List of Tables		187

186	Acknowledgements	188
-----	------------------	-----

¹⁸⁷

Chapter 1

¹⁸⁸

Introduction

¹⁸⁹ The Standard Model (SM), developed in the early 1970s, encapsulates our best understanding of the fundamental particles and their interactions. It has demonstrated huge successes in providing experimental predictions which have been confirmed by the experimental observations over-time. The discovery of the Higgs boson by the ATLAS and CMS experiments in 2012 at the Large Hadron Collider (LHC) was a breakthrough for the experimental tests of the SM. Currently, precise measurements of the Higgs boson couplings to fermions and bosons, mass and cross-section are performed with the Run-1 and Run-2 datasets collected with both ATLAS and CMS detectors. The triple Higgs boson self-coupling, λ_{HHH} , still resists physicists. The λ_{HHH} is present in the SM. This parameter controls the shape of the Higgs potential. The only direct way to measure this coupling is through the Higgs boson pair production process (HH). This process is mainly produced at the LHC via gluon-gluon fusion (ggF) through destructive interference of two Feynman diagrams involving quark loops and the triple Higgs boson self-interaction. At the LHC centre-of-mass energy of 13 TeV, the cross-section of the Higgs boson pair production is $31.05^{+2.2\%}_{-5.0\%}$ fb which is 1000 smaller than the single Higgs boson production cross-section. This small cross-section requires large integrated luminosity to make the first observation which will be delivered at HL-LHC. Since the Higgs boson is not a stable particle, the di-Higgs boson event is reconstructed through Higgs boson decay products. The HH decay channels are the combination of single Higgs boson decay channels. At LHC, the most sensitive decay channels in order of expected significance are $HH \rightarrow b\bar{b}\tau^+\tau^-$, $HH \rightarrow b\bar{b}\gamma\gamma$, $HH \rightarrow b\bar{b}bb$, and $HH \rightarrow b\bar{b}VV$ ($V = W^\pm, Z^0$).
²⁰⁷ If SM expectation hold, the observation of Higgs boson pair production is not possible with the currently available data, while many Beyond Standard Model (BSM) hypotheses predict enhancements of its cross-section either through new resonances (resonant production) or non-resonant production. The non-resonant production assumes new physics anomalies that predict a different value for the Higgs boson self-coupling. This deviation from the SM predicted value is quantified with $\kappa_\lambda = \frac{\lambda^{\text{new physics}}}{\lambda^{\text{SM}}}$ modifier, where $\lambda^{\text{SM}} = 0.13$ is the predicted Standard Model λ_{HHH} value. Searches for non-resonant and resonant di-Higgs boson production were performed by the ATLAS and CMS collaborations in several decay channels using about 36 fb^{-1} of pp collisions data from 2015-2016 at a centre-of-mass energy of 13 TeV. The ATLAS statistical combination of the $b\bar{b}bb$, $b\bar{b}\tau^+\tau^-$ and $b\bar{b}\gamma\gamma$ channels set an observed (expected) upper limit on the non-resonant HH production cross-section of around 6.9 (10) times the SM expectation and a constraint of the κ_λ parameters of $[-5, 12]$ ($[-5.8, 12]$) at 95% C.L. The CMS statistical combination of the same channels set an observed (expected) upper limit on the non-resonant HH production cross-section of around 22.2 (12.8) times the SM expectation and a constraint on κ_λ of $[-11.8, 18.8]$ ($[-7.1, 13.6]$) at 95% C.L. These results show that the sensitivity to this process is still limited with statistics used, making it a flagship analysis for the HL-LHC. It is nevertheless important to continue exploring HH production with the increase of the luminosity in data 2015-2018 (139 fb^{-1})

and to improve analysis techniques for this process.

This thesis focuses on the search for the non-resonant Higgs boson pair production in the $b\bar{b}\gamma\gamma$ final state, exploring the decay where one Higgs boson decays to two b -quark and the other to two photons. This search is translated into the measurement of the self-coupling quantifier κ_λ using the full Run-2 data collected by ATLAS detector. The branching ratio of this decay channel is only 0.3%, but still considered as a "golden" channel thanks to the well-measured photon energy. This channel was already explored in early Run-2 using 36 fb^{-1} of pp collisions. The full Run-2 analysis aims to improve the 36 fb^{-1} results on top of the luminosity increase by implementing new analysis techniques.

This thesis is structured as follows: Chapter 2 presents an overview of the SM and its particles, the Brout-Englert-Higgs mechanism, Higgs boson production, its decay at LHC and summary of the Higgs boson discovery, as well as a summary of the Higgs boson pair production and decay channels and the 36 fb^{-1} results. A description of the LHC and ATLAS detector, as well as the methods used to reconstruct the physics object, are presented in Chapter 3. To reconstruct the $HH \rightarrow b\bar{b}\gamma\gamma$ event an excellent knowledge of the photon, and therefore of the detector, is crucial. Chapter 4 focus on the different methods used to reconstruct photon and a new photon identification algorithm is presented in this chapter. Similarly to photons, jets and b -jets are mandatory for $HH \rightarrow b\bar{b}\gamma\gamma$ analysis. Jets reconstruction and tagging as well as a specific calibration of the b -jet energy is presented in Chapter 5. The complete description of the $HH \rightarrow b\bar{b}\gamma\gamma$ analysis, including signal and background Monte Carlo simulations, objects and events selection, event categorization, background modelling, systematic uncertainties, statistical interpretation and results as well as comparison to the $HH \rightarrow b\bar{b}\gamma\gamma 36 \text{ fb}^{-1}$ results, $bb\bar{b}\bar{b}$, $b\bar{b}\tau^+\tau^-$ and the 139 fb^{-1} CMS $HH \rightarrow b\bar{b}\gamma\gamma$ results, is given in Chapter 6. Finally, a description of the prospective for Run-3 analysis and the HL-LHC is given in Chapter 7.

²⁴⁵ Personal contributions

²⁴⁶ The work presented in this thesis was performed within the ATLAS collaboration. Even though this
²⁴⁷ thesis will focus on my work, all the work described would not have been possible without major
²⁴⁸ contributions from many other people and common tools and frameworks. In the following section, I
²⁴⁹ list my contributions to the analysis and ATLAS developments.

²⁵⁰ Selection of photons in ATLAS

²⁵¹ Chapter 4 details the study of photon objects in ATLAS starting from reconstruction, calibration,
²⁵² isolation and identification. My contributions are the following:

²⁵³ 1. EM objects shower shape mis-modelling: (ATLAS qualification task)

- ²⁵⁴ • I implemented the developed (for electrons) shower shape reweighting technique to photon
²⁵⁵ objects, and showed that the function developed for electron, was not enough to correct the
²⁵⁶ mis-modelling for photons.
- ²⁵⁷ • Therefore, I developed a specific reweighting function for photons and demonstrated that the
²⁵⁸ reweighting procedure is not enough to correct the mis-modelling.
- ²⁵⁹ • I provided a new idea for a three-dimension reweighting function and demonstrated its power
²⁶⁰ using pseudo data.
- ²⁶¹ • This work is documented in an internal note.

²⁶² 2. Photon identification:

- ²⁶³ • I developed a new photon identification algorithm based on Convolutional Neural Network
²⁶⁴ (CNN).
- ²⁶⁵ • I quantified the improvement of the new algorithm in different event topologies.
- ²⁶⁶ • I checked the performance of the CNN photon ID algorithm on 2017 data and evaluated the
²⁶⁷ efficiency using the Radiative Z method.
- ²⁶⁸ • I computed the scale factors using the 2017 data.
- ²⁶⁹ • I presented ideas for possible improvements of the CNN algorithm.

²⁷⁰ Jet reconstruction and Tagging in ATLAS

²⁷¹ Chapter 5 details the jet reconstruction and calibration procedures as well as the different tagging
²⁷² algorithms. In this chapter, my contributions were mainly focusing on the *b*-jet calibration method in
²⁷³ which:

- 274 • I optimized the muon selection for the μ -in-jet correction.
- 275 • I derived the p_T -Reco correction function using $t\bar{t}$ sample for both reconstruction algorithms (Topo
276 and PFlow) and both b -tagging algorithms (MV2c10 and DL1r).
- 277 • I developed an analysis tool, that applies the μ -in-jet and p_T -Reco corrections to the b -jet object.

278 Measurement of Higgs boson self-coupling

279 Chapter 6 focuses on the measurement of Higgs boson self-coupling. My contribution to this analysis
280 are the following:

- 281 • I studied the improvement of the b -jet calibration on the analysis sensitivity.
- 282 • I supervised an M1 student to work on the estimation of the systematic uncertainties related to
283 the b -jet calibration.
- 284 • I developed the likelihood fit (Kinematic Fit) and studied its impact on the $m_{b\bar{b}}$ resolution and
285 analysis sensitivity.
- 286 • I developed a tool interfacing the kinematic fit to the $b\bar{b}\gamma\gamma$ event.
- 287 • I integrated both the b -jet and the kinematic fit tools in the ATLAS software used for the analysis
288 (HGamCore).
- 289 • I participated in the ntuples (MxAOD) production for the analysis group.
- 290 • I developed the analysis strategy using the Deep Neural Network (DNN).
- 291 • I estimated the impact of the CNN photon identification on $HH \rightarrow b\bar{b}\gamma\gamma$ analysis.
- 292 • I contribute to the internal note documentation.

293 Chapter 7 presents an extrapolation of the measurement of the Higgs boson self-coupling performed
294 in Chapter 6. Extrapolations to Run-3 and to HL-LHC are performed by me.

295 Plots and tables with caption in the box,

296

297 are my personal contribution and produced by my self.

²⁹⁸ **Chapter 2**

²⁹⁹ **Theoretical Framework**

³⁰⁰ **2.1 Introduction**

³⁰¹ How the world around was made? How does it work? These are fundamental questions asked a long-time
³⁰² ago to understand the Universe. The first efforts to elucidate those questions referred to ancient Greek
³⁰³ philosophers. They gave much to physics by developing the fundamental basis of modern principles as
³⁰⁴ the conservation of matter and the atomic theory. Democritus model introduced the notion of Atom as
³⁰⁵ small indivisible building blocks (particles) of matter. At that moment, atoms described a variety of
³⁰⁶ phenomena.

³⁰⁷ Rutherford came in 1909 with his experience demonstrating that atoms consist of mostly space with
³⁰⁸ electrons surrounding a dense central nucleus (made up of protons and neutrons). At that time, Newtonian
³⁰⁹ laws of motion and atoms provided a solid framework. A continuing collaboration between theorists
³¹⁰ and experimentalists brought us to a simple theory upon which all modern physics is based, the Standard
³¹¹ Model (SM) of particle physics. Currently, SM is the most accurate model describing the universe
³¹² composition. There are two types of elementary particles: the fundamental constituents of matter called
³¹³ "Fermions", and the quanta of fields called "Bosons" exchanged when an interaction occurs between
³¹⁴ fermions and bosons. The SM has successfully predicted the results from the measurements performed
³¹⁵ in the past 50 years. The following sections provide more details about the Standard Model and its
³¹⁶ particles.

³¹⁷ **2.2 The Standard Model (SM) of particle physics**

³¹⁸ One of the physicist goals is to describe all natural phenomena with a minimal set of fundamental laws
³¹⁹ and theories which, at least in principle, quantitatively explain and predict experimental results. At
³²⁰ the microscopic scale, all matter behaviour and phenomenology, including molecular, atomic, nuclear,
³²¹ and subnuclear physics, are explained under three fundamental interactions: electromagnetic, weak and
³²² strong forces. At the macroscopic scale, the fourth force, the gravitational interaction, has an essential
³²³ role but is negligible at the microscopic scale. All the three interactions are described within a local relativistic
³²⁴ Quantum Field Theory (QFT) based on the principal gauge invariance of $SU(3) \times SU(2) \times U(1)$.
³²⁵ This is "The Standard Model" of particle physics. Within the SM, matter consists of fermions and
³²⁶ interactions are mediated by bosons.

327 2.2.1 Elementary Particles

328 In the SM, particles are classified either as fermions or bosons depending on the associated statistics.
 329 Fermions obey to Fermi-Dirac statistics and respect the Pauli exclusion principle, i.e., two fermions in
 330 the same quantum state can not exist in the same place and time. Such particles have an intrinsic
 331 angular momentum, called spin J , equal to a half-integer. Bosons obey to Bose-Einstein statistics, due
 332 to the spin-statistics theorem, corresponding to integer spin value. Through an interaction, a boson
 333 emitted by a matter particle and then absorbed by another particle. Fermions are divided into two
 334 categories: Leptons and Quarks.

335 2.2.1.1 Leptons

336 Leptons whose name comes from the Greek word meaning "light", are grouped into three families or
 337 generations formed by three charged leptons: electron e , muon μ and tau τ with an electric charge $-e$,
 338 where e is equal to the elementary electric charge of $\sim 1.6 \cdot 10^{-19} \text{ C}$ [1], and their neutral complemented
 339 partners neutrinos : ν_e , ν_μ and ν_τ . Only the electron and neutrinos are stable. A quantum number called
 340 Leptonic number (L) is associated with each lepton. Electron, muon, and tau have identical properties
 341 (e.g., charge, spin), however tau is 3477 times heavier than the electron while the muon is 17 times the
 342 mass of an electron. The rest mass of an electron is $9.10938356 \cdot 10^{-31} \text{ kg}$ with a relative uncertainty of
 343 $3 \cdot 10^{-10}$. In high-energy physics, the particle mass is expressed in terms of its energy with the equivalence
 344 $E = mc^2$, where c is the light velocity in the vacuum, thus the electron mass is 511 keV.

345 2.2.1.2 Quarks

346 Quarks are electrically charged particles, with a charge of $+\frac{2}{3}e$ for the so-called up-type quarks and $-\frac{1}{3}e$
 347 for the down-type quarks. There are six known quark flavours, similarly to leptons, quarks are paired
 348 into three generations. The first generation consists of up (u) and down (d) quarks, the second has the
 349 charm (c) and strange (s) quarks, and top (t) and bottom (b) quarks for the third generation. Quarks
 350 do not exist in a free state due to the QCD colour confinement.

351 Table 2.1 summarizes leptons and quarks. Each of the higher generations has particles with higher
 352 mass and tends to decay to the lower one, which explains why the ordinary matter made off the first-
 353 generation particles. Each fermion is associated with a corresponding anti-particle. The Standard
 354 Model of elementary particles assumes neutrinos to be mass-less, while some experiments demonstrate
 355 that neutrinos are changing flavours which can be explained by having a non-zero mass, which makes
 356 the SM an incomplete theory.

357 2.2.1.3 Bosons

358 As mentioned before, bosons are particles of integer angular momentum and obeying the Bose-Einstein
 359 statistics. They are the carriers of the gauge interactions between fermions. The photon (γ) is a boson
 360 known as a quantum of the electromagnetic field including electromagnetic radiation such as light.
 361 Photons are electrically neutral and mass-less particles. W^\pm, Z^0 are bosonic particles that carries the
 362 weak interaction. Z^0 boson is neutral while W^\pm have a charge of $\pm e$. Contrary to photons, weak bosons
 363 are massive. W^\pm and Z^0 have been predicted at the end of 1972 as the charged (CC) and neutral
 364 currents (CN) of the weak interaction. Their masses are predicted to be so large that it took many years
 365 to build powerful accelerators to produce them. They are directly observed at CERN in 1983 by the
 366 UA1 and UA2 collaborations [2, 3] and their masses were measured to be about 80 GeV and 91 GeV

	1^{st} Generation	2^{nd} Generation	3^{rd} Generation
Quarks	u 2.16 MeV $+\frac{2}{3}$	c 1.27 GeV $+\frac{2}{3}$	t 172.4 GeV $+\frac{2}{3}$
	d 4.67 MeV $-\frac{1}{3}$	s 93 MeV $-\frac{1}{3}$	b 4.18 GeV $-\frac{1}{3}$
	e 0.511 MeV -1	μ 105.7 MeV -1	τ 1.8 GeV -1
	ν_e 0 0	ν_μ 0 0	ν_τ 0 0

Table 2.1: Generations of quarks and leptons with their masses and electrical charges

367 respectively. Gluons are the neutral quantum of the strong force known as the "glue" that links quarks
 368 to form hadrons. The mass of the gluons is known to be strictly zero.

369 2.2.2 Elementary Interactions

370 At the microscopic level, there are three fundamental forces: electromagnetic, weak and strong. The
 371 interactions with matter (fermions) is transmitted by boson. The SM content and interactions are
 372 expressed more formally through the concepts of symmetries and gauge invariance. Each interaction
 373 is described through a gauge group. The group generators correspond to the gauge bosons that are
 374 mediators of the fundamental force and responsible for the interactions. The electromagnetic interaction
 375 is mediated by photons, while the weak force have W^\pm and Z^0 as mediators. The gluons are the mediators
 376 for the strong interaction.

377 2.2.2.1 Electromagnetic interaction

Quantum Electrodynamics (QED) describes the dynamics of electromagnetic interaction for electrically charged fermions and bosons. Each quantum field theory is represented by a Lagrangian density. The QED Lagrangian representing the behaviour of a freely propagating fermion field $\psi(x, t)$ is written as:

$$\mathcal{L} = \bar{\psi} i\gamma^\mu \partial_\mu \psi - m\psi\bar{\psi}, \quad (2.1)$$

where m is the particle mass. The Einstein convention is used here, with the indices $\mu = 0, 1, 2, 3$ representing the space-time components x and the time t .

To be a valid gauge theory, the QED Lagrangian should be invariant under a U(1) (electromagnetic force group) local gauge transformation of the field: $\psi \rightarrow e^{i\alpha(x)}\psi$. This condition leads to additional terms to be added to the Lagrangian, a new gauge field A_μ that represents the photon:

$$\mathcal{L}_{QED} = \bar{\psi} i\gamma^\mu \partial_\mu \psi - m\psi\bar{\psi} + q\psi\gamma^\mu\psi A_\mu - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \quad (2.2)$$

where $F^{\mu\nu} = -F^{\mu\nu} = \partial^\nu A^\mu - \partial^\mu A^\nu$ is the field-strength tensor for the electromagnetic force which describes the kinetic propagation of the field. Note that U(1) is an Abelian group inducing that photons can not self-interact and the EM tensor does not have the photon self-interaction term included. The term $q\bar{\psi}\gamma^\mu\psi A_\mu$ reflects the interaction between a fermion and the electromagnetic force (photon). The strength q (q=-e) is the electromagnetic interaction charge (electric charge). The mass term of the photon is not added to the Lagrangian since it spoils the gauge invariance. This is in agreement with the observation that the photon is mass-less. U(1) has one generator which corresponds to the photon mediator.

2.2.2.2 Electro-Weak interaction

The electroweak interaction consists of a unification of the electromagnetic and weak interactions. It is described by the combined gauge symmetry $SU(2)_T \times U(1)_Y$, where the $U(1)_Y$ symmetry mimics the QED one with the weak hypercharge Y. The $SU(2)$ represents the weak interaction with its generator vector T called weak isospin. The Lagrangian of the theory can be written as:

$$\mathcal{L}_{EW} = \bar{\psi}i\gamma^\mu\partial_\mu\psi - eY\bar{\psi}\gamma^\mu B_\mu\psi - g_W\bar{\psi}\gamma^\mu\mathbf{T}\cdot\mathbf{W}_\mu\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{4}W^{i\mu\nu}W_{\mu\nu}^i, \quad (2.3)$$

where g_W is the weak coupling to fermionic fields. The first two terms are similar to \mathcal{L}_{QED} . However, the term representing the coupling of fermions to photons is replaced by more general terms: the W_μ field and the hyper-photon B_μ field. The two charged vector bosons W^\pm appear as a linear combination of the W_μ field components, $W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp W_\mu^2)$. The photon is now created through the mixing of the W_μ and B_μ fields, as:

$$A_\mu = B_\mu \cos\theta_W + W_\mu^3 \sin\theta_W, \quad (2.4)$$

where θ_W is the weak mixing angle, and the Z_μ field, corresponding to the Z boson, is generated similarly as:

$$Z_\mu = -B_\mu \sin\theta_W + W_\mu^3 \cos\theta_W. \quad (2.5)$$

The field strength tensor for the weak gauge fields W^i is defined as:

$$W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g_W\epsilon_{ijk}W_\mu^jW_\nu^k, \quad (2.6)$$

The non-Abelian nature of SU(2) group generates the third term in Eq. 2.6, which gives rise to the weak boson self-interactions.

The electric charge q is related to Y and T through the Gell-Mann-Nishijima relation [4]:

$$q = T_3 + \frac{Y}{2}, \quad (2.7)$$

where T_3 is the third component of the isospin T . Fermions are decomposed into left-handed and right-handed chirality types. The chirality is defined for mass-less particles as the same as helicity and refers to the relation between the spin and the momentum direction. For massive particles, chirality is trickier to define. Note that the electroweak force interacts only with left-handed particles because interaction with right-handed particles would violate the parity. The chirality separates left-handed doublets and right-handed singlets:

$$(\nu_i \ i)_L^T, i_R \text{ with } i = e, \mu \text{ or } \tau, \quad (2.8)$$

and,

$$(u \ d)_L^T, (c \ s)_L^T, (t \ b)_L^T, u_R, d_R \dots . \quad (2.9)$$

387 **2.2.2.3 Strong interaction**

Additionally to QED, the strong interaction is described by Quantum Chromodynamics (QCD), which is a local gauge symmetry under $SU(3)_C$. $SU(3)$ has eight generators corresponding to the eight gluons mediators of the strong force. The expression for the locally gauge invariant Lagrangian of the QCD, which describes interactions of quarks of mass m and quark-field spinors ψ via mass-less gluons is:

$$\mathcal{L}_{QCD} = \sum_{flavours} \bar{\psi}_a (i\gamma^\mu \partial_\mu \delta_{ab} - g_s \gamma^\mu t_{ab}^C G_\mu^C - m \delta_{ab}) \psi_b - \frac{1}{4} F_{\alpha\beta}^A F^{A\alpha\beta}, \quad (2.10)$$

where $a = 1, 2, 3$ is the colour index. The G_μ^C are the gluon fields with $C = 1, 2, \dots, 8$ referring to the eight gluons. The t_{ab}^C are the generators of the $SU(3)$ group. The quantity g_s corresponds to the strong coupling constant that determines the strength of interactions between coloured particles (quarks and gluons). The strong field tensor is defined as:

$$F_{\alpha\beta}^A = \partial_\mu G_\nu^A - \partial_\nu G_\mu^A - g_s f^{ABC} G_\mu^B G_\nu^C, \quad (2.11)$$

where f^{ABC} are the structure constants of the $SU(3)$ group. In contrast to QED, the field tensor of the QCD includes the gluon triplet and quartic self-interaction. Only particles carrying colour charge (i.e. quarks) couple to these gluons. Colour charge is conserved in each interaction. Since the $SU(3)$ group is non-Abelian, the gluons themselves must carry (anti-)colour charge and thus interact with each other. Finally, the Standard Model Lagrangian density could be summarized as:

$$\mathcal{L}_{SM} \sim \mathcal{L}_{EW} + \mathcal{L}_{QCD}. \quad (2.12)$$

388 Note that, in the \mathcal{L}_{EW} no mass term is introduced for the two W/Z appearing bosons, while experimental
389 observations indicate large masses for them. Additionally, the fermion mass term that appeared in \mathcal{L}_{QED}
390 is now absent in \mathcal{L}_{EW} , requiring the introduction of an additional mechanism to explain the origin of
391 their mass. The introduction of terms related to mass is handled through the process of Electroweak
392 Symmetry Breaking (EWSB), and the introduction of the Higgs boson mechanism.

393 **2.3 Electroweak Symmetry Breaking (EWSB)**

The gauge bosons mass terms cannot be simply introduced as they would break gauge invariance. In 1964, the Brout-Englert-Higgs (BEH) mechanism introduced a spontaneous symmetry breaking ad-hoc in the SM to preserve the local gauge invariance [5] called Electroweak Symmetry Breaking (EWSB). In this minimal model, the mass of the particle is generated by introducing the Higgs field represented by a weak isospin doublet of one charged and one neutral complex scalar ($J=0$) field:

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_+ \\ \phi_0 \end{pmatrix}. \quad (2.13)$$

The Higgs Lagrangian is added to the SM Lagrangian as:

$$\mathcal{L}_{Higgs} = (D^\mu \phi)^\dagger (D_\mu \phi) - V(\phi), \quad (2.14)$$

where D^μ is the covariant derivative $D^\mu = (\partial_\mu + igT^i W_\mu^i + i\frac{1}{2}g'B_\mu)$. The kinetic and the interaction of the Higgs field with weak gauge bosons is described by the term $D^\mu \phi$. $V(\phi)$ is the Higgs potential:

$$V(\phi) = \mu^2 \phi^+ \phi + \lambda (\phi^+ \phi)^2, \quad (2.15)$$

the first term is associated with the mass of the field and the second represents the quadratic self-interaction of the scalar field. The parameter λ has to be positive to obtain a potential with finite minima, while μ can be chosen freely. For $\mu^2 > 0$, the potential has a single ground state at zero with all fields being zero ($\phi_0 = 0$). Hence, the vacuum expectation value of the Higgs field would be zero and the symmetry is preserved. However, for $\mu^2 < 0$ the potential has an infinite set of minima v given by:

$$\phi_0 = \langle \phi^+ \phi \rangle = \frac{v^2}{2} = -\frac{\mu^2}{2\lambda}, \quad (2.16)$$

$v \neq 0$ being the vacuum expectation value (vev) which is illustrated in Figure 2.1.

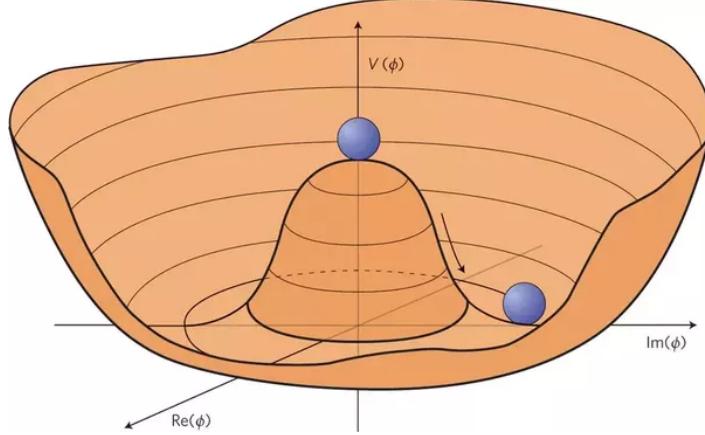


Figure 2.1: Shape of the Higgs potential for $\mu^2 < 0$. This potential has an infinite amount of minima. [6]

394

The choice of the physical vacuum state spontaneously breaks the Lagrangian symmetry. An expansion of ϕ_0 around its vacuum state v introduces a massive scalar and three mass-less Goldstone bosons. However, the Goldstone bosons appear to be not physical and can be eliminated using the gauge Unitary, enforcing the Higgs doublet to be real:

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \quad (2.17)$$

where $h(x)$ is the physical field linked to a new massive particle: the Higgs boson.

The form of Higgs potential becomes:

$$V(\phi) = -\frac{1}{2}m_H^2 h^2(x) + \lambda_{HHH} h^3(x) + \lambda_{HHHH} h^4(x), \quad (2.18)$$

where $m_H^2 = 2\lambda^2 = 2\lambda v^2$ is the square of Higgs boson mass, λ_{HHH} is the coupling in a vertex with tree Higgs bosons (trilinear coupling) and λ_{HHHH} for the case of four Higgs bosons. Accordingly, the terms giving the mass to the gauge bosons can be identified:

$$m_W = \frac{1}{2}gv \quad (2.19)$$

$$m_A = 0 \quad (2.20)$$

$$m_Z = \frac{m_W}{\cos \theta_W}. \quad (2.21)$$

395 The Higgs mechanism associates the degrees of freedom of the hypothetical scalar (Goldstone) bosons
 396 with the longitudinal components of gauge bosons and consequently become massive. The coupling
 397 of the Higgs boson to the gauge bosons appears to be proportional to the gauge boson masses. The
 398 mass of the fermions cannot be explained in the same way as the fermions acquire masses through the
 399 corresponding Yukawa couplings. Weak bosons masses are predicted in the SM, while Higgs boson mass
 400 is unknown and can not be constrained from other SM measurement. The Higgs boson discovery is
 401 necessary to confirm the EWSB [7], and the Higgs boson properties needs to be measured.

402 2.4 Higgs boson : Production and observation

403 2.4.1 Previous searches for Higgs boson

404 The Large Electron-Positron (LEP) collider, approved by the CERN Council in 1981 and commissioned
 405 eight years later, provided a detailed study of the electroweak interaction until 2000. It accelerated
 406 and collided electron and positron beams at a centre-of-mass energy \sqrt{s} from 91 to 200 GeV. Its results
 407 have been used to perform stringent tests of the SM by comparing the precise measurements with theory
 408 predictions. This checked the correctness of the SM theory. Observing the Higgs boson was of particular
 409 interest because this fundamental ingredient of SM has not been observed and needed to complete the
 410 SM picture.

411 At the beginning of the LEP program, no solid prediction existed on the mass of the Higgs boson.
 412 Searches for the SM Higgs boson carried out by the four LEP experiments extended the sensitive range
 413 well beyond that anticipated. Similarly, the Tevatron experiments CDF and D0 have excluded a range
 414 of mass $162 < m_H < 166$ GeV. Tevatron was a proton-antiproton ($p\bar{p}$) collider at a centre-of-mass energy
 415 of 3 TeV. Figure 2.2 shows the LEP and Tevatron exclusion limits [8, 9, 10]. Combination of LEP and
 Tevatron results yields to the best estimated Higgs boson mass at that time of $m_H = 116.4^{+15.6}_{-1.3}$ GeV.

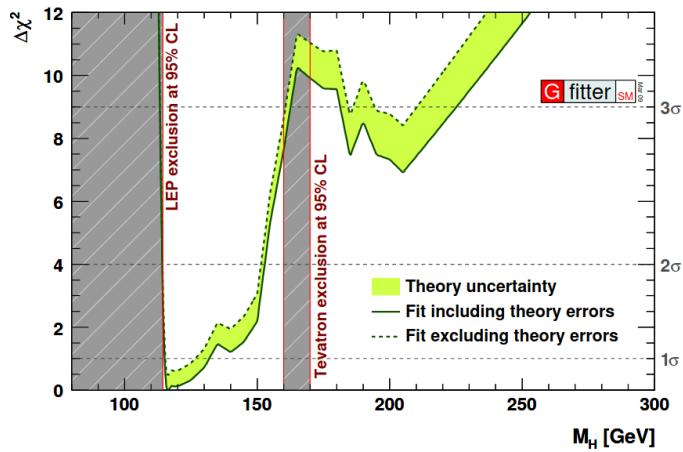


Figure 2.2: The χ^2 function for the SM as a function of the Higgs boson mass, LEP and Tevatron exclusions limits are shown.

416
 417 This is due to the higher energy achieved and to more sophisticated detectors and analysis techniques.
 418 However, circular accelerators with electrons are limited by synchrotron radiation due to the small
 419 electron mass. Since muon acceleration is not technically possible at the current stage due to the short
 420 lifetime of the muon, proton collisions with accelerator technology represented a good way to achieve

high-energy collisions needed for Higgs boson discovery.

LHC provides a nominal proton-proton collision at a centre-of-mass energy \sqrt{s} up to 14 TeV. But at the beginning of LHC the energy was limited to 7-8 TeV for safety reason on the accelerator. During the last four years, the achieved energy is 13 TeV. LHC is detailed in the dedicated Chapter 3.

2.4.2 Proton-proton collisions

Protons are fermions made of two up and one down quarks (uud) called "valence" quarks. They interact with each other through the strong force via gluon exchange. Due to the nature of QCD (non-abelian), gluons can fluctuate into quark anti-quark pairs forming the so-called "sea" quarks. Effectively, a proton is therefore a bound state of quarks and gluons called partons each carries a fraction x of the total proton momentum. At the LHC, process are produced by colliding mainly protons beams. The Higgs boson is produced in a proton-proton collision as:

$$p_1 + p_2 \rightarrow H. \quad (2.22)$$

The probability for a process to occur is expressed in terms of its cross-section, and the dynamic of its products is determined by the dynamics of partons involved in the collision. The hard scattering cross-section for such a process is given by:

$$d\sigma^{p_1 p_2 \rightarrow H} = \int_0^1 dx_1 \int_0^1 dx_2 \sum_{a,b} f_{a/p_1}(x_1, \mu_F^2) f_{b/p_2}(x_2, \mu_F^2) d\hat{\sigma}^{ab \rightarrow H}(x_1, x_2, \mu_F^2), \quad (2.23)$$

where a, b are the partons involved in the process and $f_{n/p_i}(x_i)$ is the parton distribution function (PDF), which is the probability to find a parton of type n inside the proton p_i with a longitudinal momentum fraction x_i at the energy scale Q . $\sigma^{ab \rightarrow H}$ is the parton cross-section of the process.

In general, the x dependence at a given Q^2 cannot be calculated analytically but rather are extracted from global fits to data from many experiments. Different parton distribution function sets use different fitting methods and experimental data [11]. Figure 2.3 shows two examples of CT14 parton distribution functions at different Q^2 computed using data from LHC experiments, and the new DØ charged lepton rapidity asymmetry data [12]. The production of the Higgs boson at the LHC occurs in different modes.

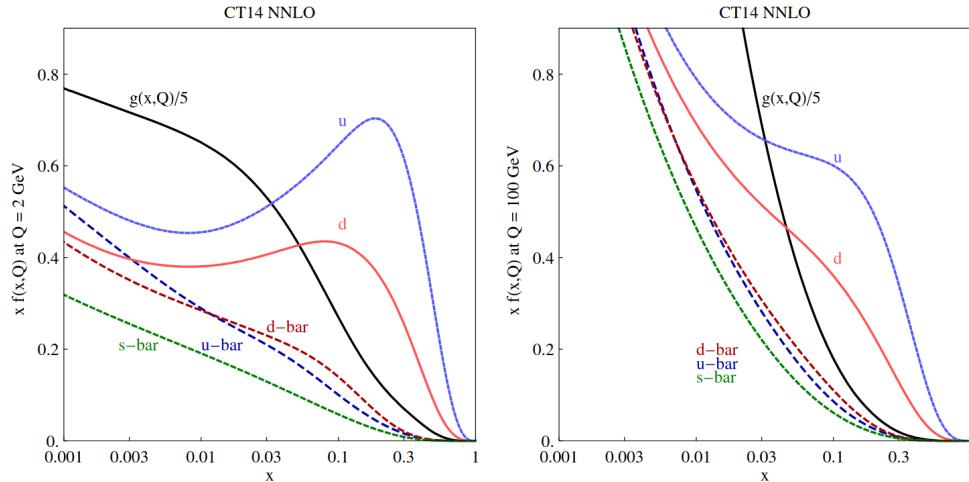


Figure 2.3: The CT14 PDFs at $Q=2 \text{ GeV}$ (left) and $Q = 100 \text{ GeV}$ (right) for different partons. These PDFs are computed at the next-to-next-to-leading order.

433
434
435

Their cross-sections depend on the coupling of the Higgs boson to specific particles, but also essentially on the PDFs described above.

436

2.4.3 Higgs boson production

437

The main processes contributing to the Higgs boson production at LHC are represented by their leading Feynman diagrams displayed in Figure 2.4. In the 7-13 TeV centre-of-mass energy range, the most

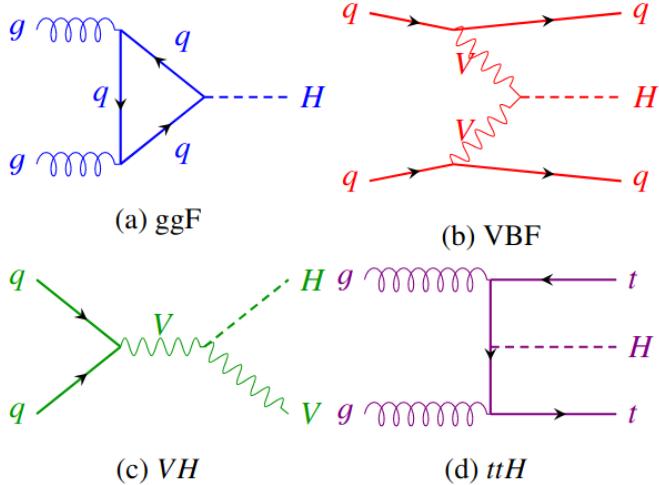


Figure 2.4: Feynman diagrams for the main production modes of Higgs boson.

438

439

important Higgs boson production is the fusion of gluons (ggF).

440

In order of decreasing cross-section as shown in Figure 2.5 (a), the Higgs boson production modes are:

441
442
443
444

- Gluon-gluon fusion (ggF): the main Higgs boson production at LHC. The process involves the fusion of two incoming gluons that produce the Higgs boson through a heavy quark loop, whose main contribution comes from the top quark. The corresponding Feynman diagram is shown in Figure 2.4 (a).

445
446
447
448
449
450
451

- Vector Boson Fusion (VBF): each of the two interacting quarks emit a W^\pm or Z^0 boson which, in turn, interact to produce the Higgs boson, as shown in figure 2.4 (b). Quarks deriving from the incoming partons after the emission of vector bosons proceed in the forward direction and represent the peculiar signature of this production mode, two high-energy forward jets separated by a large pseudo-rapidity gap, a $\Delta\eta$ region with reduced particle density. This process has a cross-section which is one order of magnitude lower than ggF, and it would have become comparable to ggF for Higgs boson mass of the order of 1 TeV as shown in Figure 2.5 (b).

452
453
454
455
456

- Vector boson associated production (VH): also known as Higgs-strahlung, this process is characterized by the emission of a Higgs boson from a W^\pm or Z^0 boson produced by two incoming quarks, as shown in figure 2.4 (c). The VH cross-section is several orders of magnitude lower than the ggF and VBF cross-sections for m_H larger than about 300 GeV, while the VH and VBF cross-sections are comparable around $m_H = 125$ GeV as shown in Figure 2.5 (b).

- 457 • Top quark associated production ($t\bar{t}H$): a pair of top quarks, originated from the splitting of two
 458 incoming gluons, interacts to give rise to a Higgs boson, as shown in figure 2.4 (d). The production
 459 in association with a pair of top quarks allows a direct measurement of the Higgs boson coupling
 460 to the top quark. Another production mechanism analogous to the $t\bar{t}H$ process and with a similar
 461 cross-section is the b-quark associated production.
- 462 The SM Higgs boson production cross-section for the various production modes depends on the Higgs
 boson mass and the centre-of-mass energy, as shown in Figure 2.5. The Higgs boson is an unstable

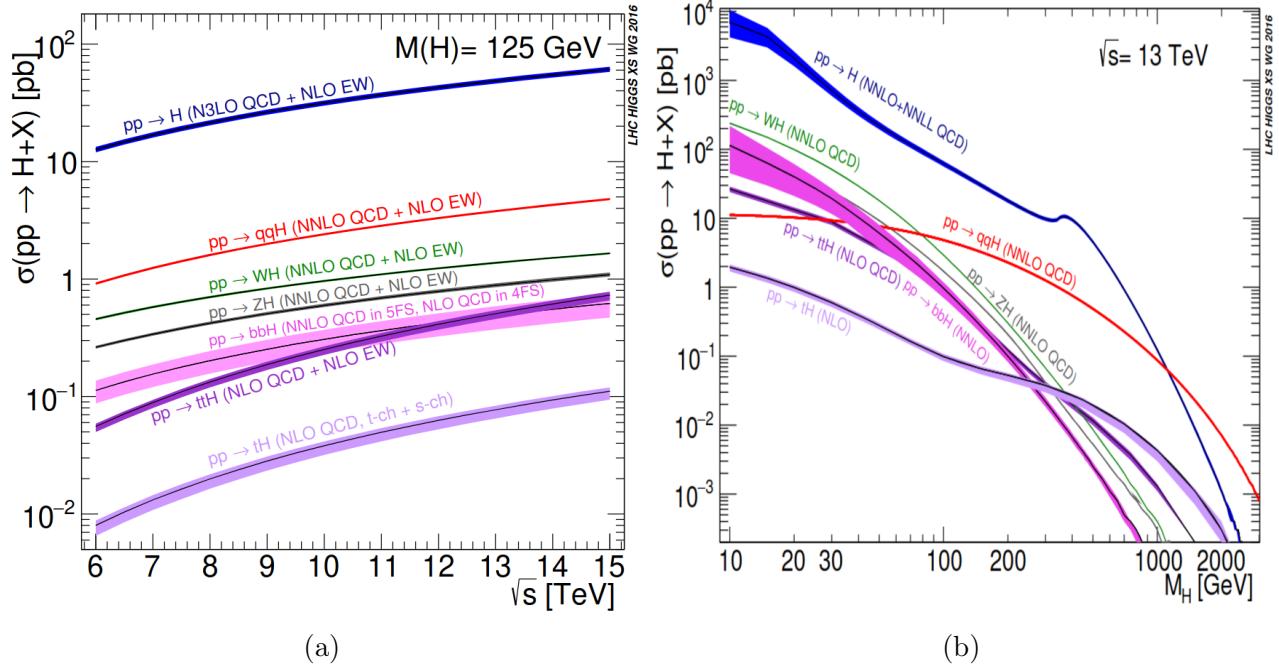


Figure 2.5: Cross-sections for different Higgs boson production modes at a proton-proton collider with a centre-of-mass energy of 6-15 TeV (a). A Higgs boson mass of 125 GeV is assumed in this plot. Additionally, cross-sections as a function of Higgs boson mass are shown (b) [13].

- 463 particle as shown in Figure 2.6. Note that the probability to decay within a given time "decay width"
 464 is linked to the lifetime τ by $\Gamma\tau = \hbar$ where \hbar is the reduced Planck constant. The Higgs boson decay
 465 width for a Higgs boson mass $m_H = 125$ GeV is around 4 MeV [14].
 466 In order to identify the Higgs boson and its production modes, it is reconstructed from its products for
 467 a chosen decay channel.

469 2.4.4 Higgs boson decay channels

- 470 The possible SM Higgs boson decay modes are very dependent on the Higgs boson mass as shown in
 471 Figure 2.7. If the Higgs boson was heavy enough to decay into two real vector bosons, the modes
 472 $H \rightarrow WW^*$ and $H \rightarrow ZZ^*$ would have dominated the decay with small contribution from Higgs boson
 473 to di-top quarks. At very low masses of the Higgs boson, decays into the vector boson or $t\bar{t}$ would have
 474 played almost no role and the dominant decay mode would have been the experimentally challenging
 475 decay mode $H \rightarrow b\bar{b}$. Since photons are mass-less particles, the direct coupling of the Higgs boson to
 476 photons is zero in the SM. However, the Higgs boson can decay into a pair of photons via loop processes.
 477 The main Feynman diagrams of this decay are shown in Figure 2.8. At a Higgs boson mass of 125 GeV,

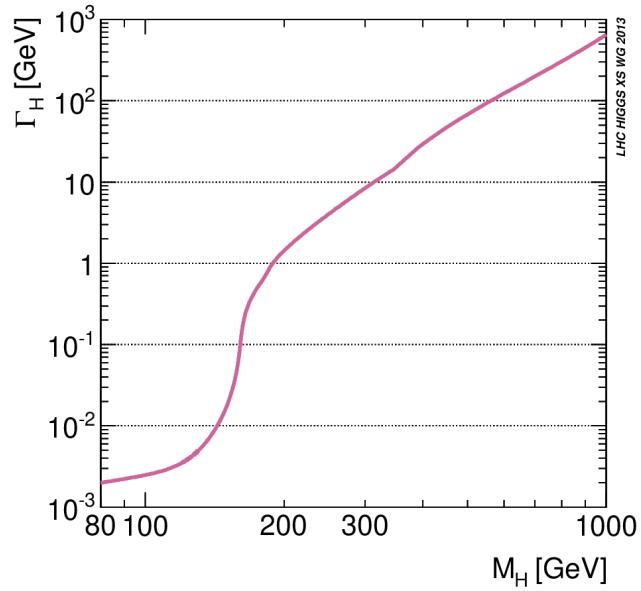


Figure 2.6: Higgs boson total decay width as a function of Higgs boson mass [15].

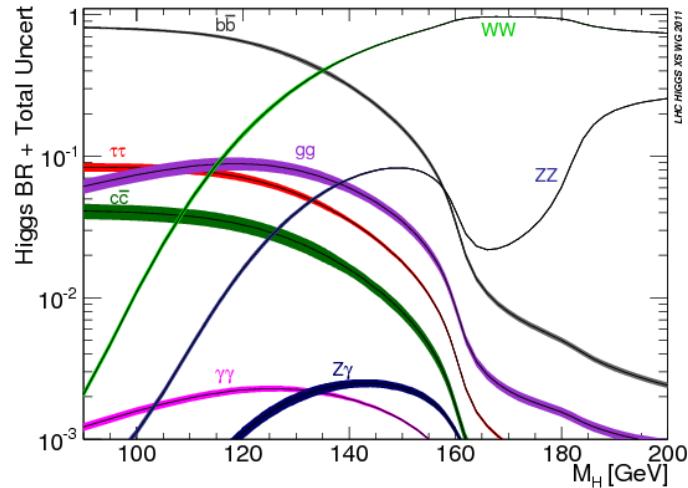


Figure 2.7: Higgs boson branching ratio of the possible Higgs boson decay channels as a function of the Higgs boson mass [16]

478 the dominant decay of the Higgs boson is $H \rightarrow b\bar{b}$ with a branching ratio of roughly 58%. At the same mass $H \rightarrow \gamma\gamma$ branching ratio is around 0.23% [17].

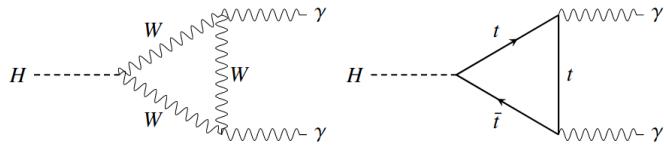


Figure 2.8: Feynman diagrams of the $H \rightarrow \gamma\gamma$ decay. Other particles can also contribute to the loop.

479

480 2.4.5 Higgs boson discovery

481 The ATLAS and CMS collaborations, located at the LHC collider, announced on 4 July 2012 the
 482 identification with a confidence of 5σ of a new boson within a mass range of 125-127 GeV [18, 19]. The
 483 new boson was consistent with the Standard Model Higgs boson decay modes and signal rate. Later,
 484 it was confirmed that the new particle corresponds to the Higgs boson. The Higgs boson was predicted
 485 in 1964, but the world had to wait until 2012 for its discovery. The evidence of the new boson was
 486 performed in the three bosonic decay channels $H \rightarrow ZZ^* \rightarrow 4l$, $H \rightarrow \gamma\gamma$ and $H \rightarrow WW^* \rightarrow l\nu l\nu$.
 487 The $H \rightarrow \gamma\gamma$ has a small branching ratio, while the reconstruction of the two photons has a clean
 488 signature in the detector which allows for a good photon energy resolution. This signal is therefore well
 489 separated from the dominating continuum background from the QCD $\gamma\gamma + \text{jets}$, which has a smooth and
 490 well parametrisable shape. This makes the $H \rightarrow \gamma\gamma$ decay mode one of the most sensitive channel.

491 Events collected in 2010 and 2011 with centre-of-mass energy $\sqrt{s} = 7$ TeV and 8 TeV respectively,
 492 containing two photons were combined. Figure 2.9 shows the distribution of the invariant mass of pho-
 493 ton pairs measured by ATLAS experiments, and demonstrates a statistically significant excess of events
 near 125 GeV.

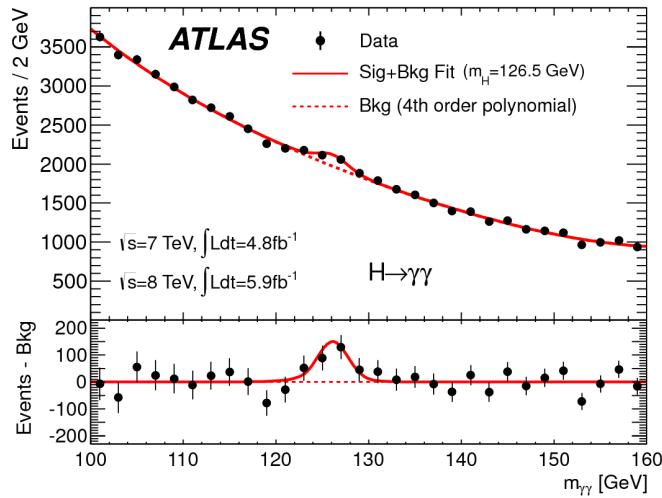


Figure 2.9: Distributions of the invariant mass of di-photon system [18].

494

495 Measurements of the three bosonic channels were combined to confirm the observation of the new par-
 496 ticle [18]. The observed local significance reaches 6σ (the observed signal is $\sim 10^{-9}$ to be a background
 497 fluctuation) around 125 GeV making the first observation of the Higgs boson, as shown in Figure 2.10.

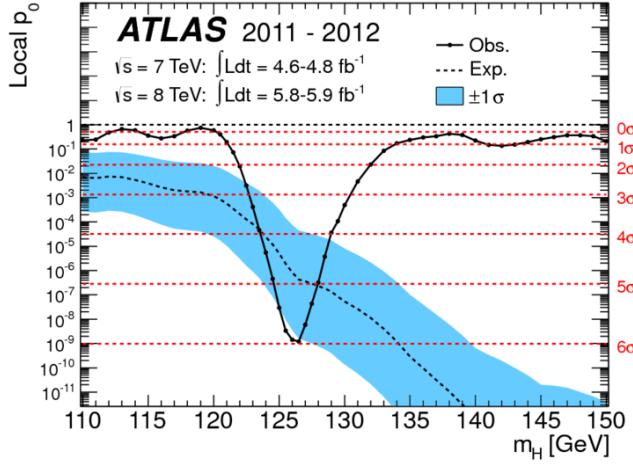


Figure 2.10: The observed local p-value as a function of m_H (solid line) and the expectation with its $\pm 1\sigma$ band assuming the presence of a Standard Model Higgs boson at that mass (dashed line) from the combination of the ZZ^* and $\gamma\gamma$ channels by the ATLAS experiment. The horizontal dashed lines indicate the p-values corresponding to significance of 1 to 6σ [18].

The latest measurement of the Higgs boson using the data collected during 2015–2018 corresponding to 139 fb^{-1} yields to a mass of $m_H = 125.09 \pm 0.24 \text{ GeV}$ [20]. Figure 2.11 shows the diphoton invariant mass distribution of event collected during 2015–2018. Figure 2.12 shows the best-fit values of the production cross-sections times branching fraction in the different channel to their SM values.

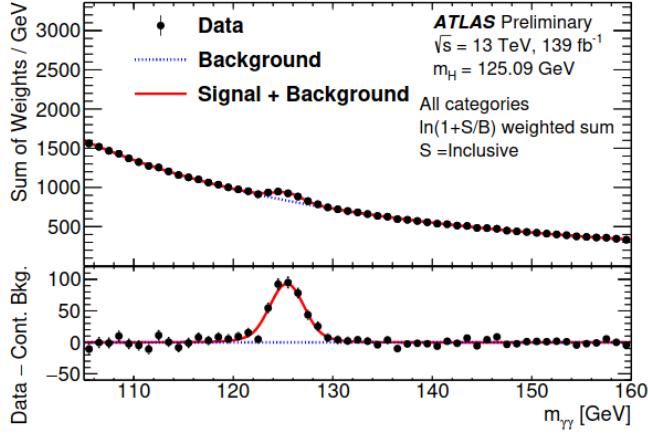


Figure 2.11: The inclusive diphoton invariant mass distribution [21].

Finally, by the observation of the Higgs boson, the last piece of the standard model of elementary particles was found. SM particles are summarized in Figure 2.13. Measuring the properties and couplings of the last piece of SM is a priority for both ATLAS and CMS. Higgs boson self-couplings λ_{HHH} is vital, providing a direct probe on the EWSB and is a precision test of the electroweak theory. A direct probe of the trilinear coupling is possible through studying Higgs boson pair production where two Higgs bosons are produced in the same event, making di-Higgs boson analyses particularly interesting and the main subject of this thesis.

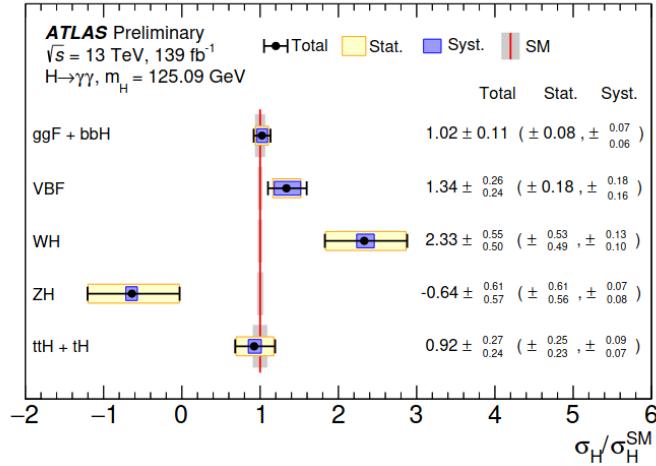


Figure 2.12: Cross-section times branching fraction for $ggF + b\bar{b}H$, VBF , VH and $t\bar{t}H + tH$ production, normalized to their SM predictions [21].

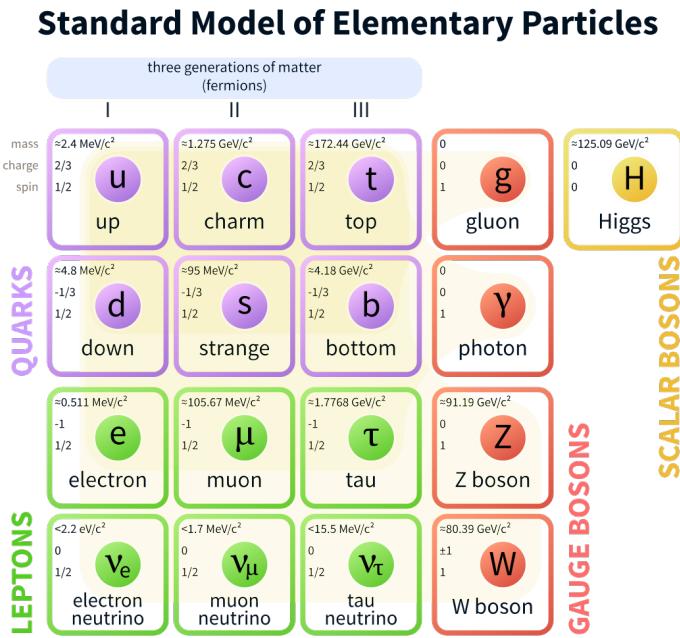


Figure 2.13: The particle content of the standard model of particle physics.

510 2.5 Looking for di-Higgs boson events

511 Double Higgs boson production (HH) presents the only direct probe of the Higgs boson self-coupling.
 512 Confirming the existence and amplitude of the Higgs boson self-coupling is one of the main key pre-
 513 diction of the SM not observed yet. In particular, the measurement of the Higgs boson self-coupling λ
 514 (also referred to as the Higgs boson trilinear coupling λ_{HHH}) is of great importance to yield a deeper
 515 understanding of particle physics and cosmology. This measurement makes it possible to experimen-
 516 tally reconstruct the Higgs potential and check whether the Higgs boson discovered in 2012 at CERN
 517 is the one predicted by the Brout-Englert-Higgs mechanism. The di-Higgs boson production rate gives
 518 a handle to measure more accurately the Higgs potential. The main goal of this thesis is the search of
 519 Higgs boson pair production.

520 2.5.1 Di-Higgs boson production and decays

As the SM provides the trilinear coupling of the Higgs boson, all SM Higgs boson production modes are known for di-Higgs boson production. Similarly to the single Higgs boson, HH production is dominated by gluon-gluon fusion (ggF) through the destructive interference of two LO Feynman diagrams, shown in Figure 2.14, involving top-quark loops and the triple Higgs boson self-coupling. In the box diagram, the top-quark Yukawa coupling λ_t is present in two vertices so the contribution of this diagram to the amplitude is proportional to λ_t^2 . In the triangle diagram, there is λ_t in one vertex and the triple Higgs boson self-coupling λ_{HHH} in the other vertex, thus the contribution of this diagram is proportional to $\lambda_t \cdot \lambda_{HHH}$. The amplitude of the process can be written as:

$$A(\lambda_t, \lambda_{HHH}) \equiv \lambda_t^2 \cdot \square + \lambda_t \cdot \lambda_{HHH} \Delta, \quad (2.24)$$

where \square represents the contribution of the box diagram and Δ the one of the triangle diagram. The

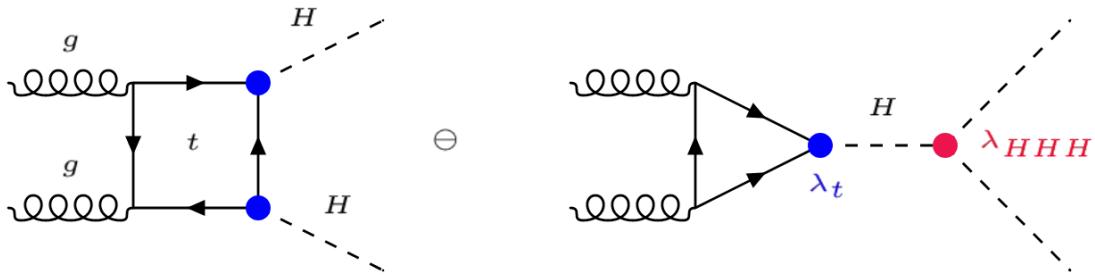


Figure 2.14: Leading Order (LO) Feynman diagrams contributing to ggF Higgs boson pair production through a top-quark loop (box) and through the triple self-coupling of the Higgs boson (triangle).

SM cross-section for Higgs boson pair production via ggF at $\sqrt{s} = 13$ TeV, calculated at Next-to-Next Leading Order (NNLO) [22, 23], is:

$$\sigma_{\text{ggF HH}}^{\text{NNLO}} = 31.05^{+2.2\%}_{-5.0\%} \text{ fb}, \quad (2.25)$$

521 three orders of magnitude smaller than the single Higgs boson production cross-section. This accounts
 522 for more than 90% of the total Higgs boson pair production cross-section. VBF mode also contributes
 523 to the di-Higgs boson production with a lower cross-section of $\sigma_{\text{VBF HH}}^{\text{NNLO}} = 1.723^{+0.03\%}_{-0.04\%}$ fb.

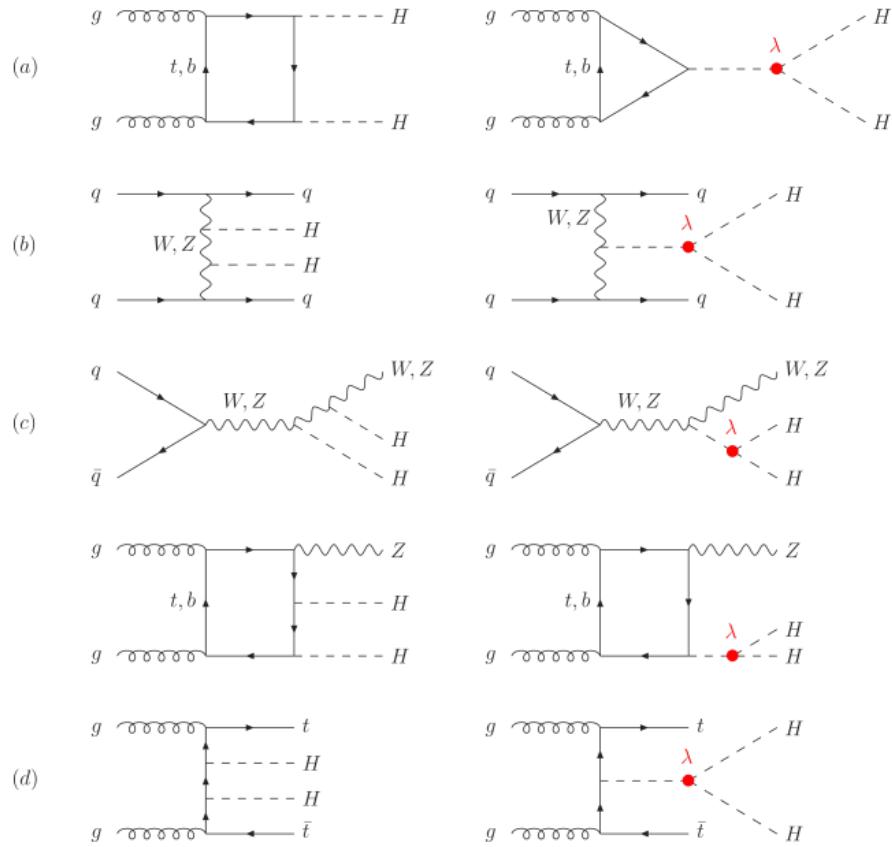


Figure 2.15: HH production modes, in decreasing order of cross-section. λ refers to Higgs boson trilinear coupling: (a) ggF mode, (b) VBF mode, (c) Vector boson association (VH) and (d) top associated production.

524 Figure 2.15 shows the other production modes contributing to Higgs boson pair production. The
 525 corresponding cross-section as a function of centre-of-mass energy is shown in Figure 2.16.
 526 Figure 2.17 represents the matrix of the decay channels and their branching ratio of a pair of Higgs bosons
 527 resulting from all possible combination of decays of the two Higgs bosons. The dominant channels are
 $HH \rightarrow b\bar{b}b\bar{b}$, $HH \rightarrow b\bar{b}\tau\tau$ and $HH \rightarrow b\bar{b}WW^*$.

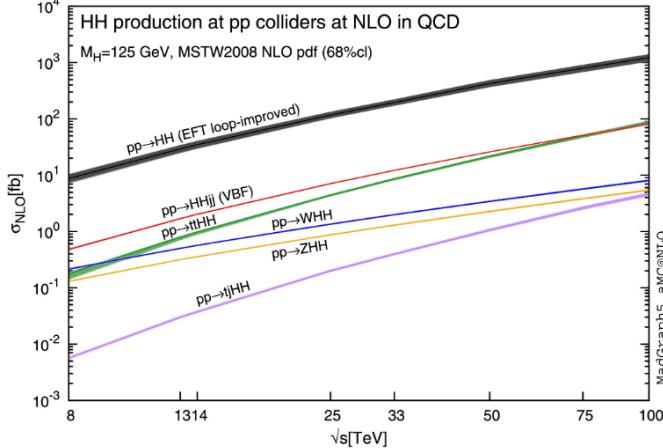


Figure 2.16: Total cross-sections at the NLO for HH production as a function of centre-of-mass energy. Assuming κ_λ and κ_t equals to their SM values.

	bb	WW	$\tau\tau$	ZZ	$\gamma\gamma$
bb	33%				
WW	25%	4.6%			
$\tau\tau$	7.4%	2.5%	0.39%		
ZZ	3.1%	1.2%	0.34%	0.076%	
$\gamma\gamma$	0.26%	0.10%	0.029%	0.013%	0.0005%

Figure 2.17: Di-Higgs boson system decay branching ratios assuming SM Higgs bosons with $m_H = 125.09$ GeV.

528 Even if there are other channels with branching ratios higher by one or two orders of magnitude, $b\bar{b}\gamma\gamma$
 529 final state is one of the most promising decay channels for the search for di-Higgs boson production as it
 530 has a good compromise between the large branching ratio from $H \rightarrow b\bar{b}$ decay and the small Higgs boson
 531 to diphoton branching ratio. The $b\bar{b}\gamma\gamma$ channel is appealing thanks to an excellent diphoton invariant
 532 mass resolution leading to a clean diphoton signature through a narrow peak at Higgs boson mass in
 533 the $m_{\gamma\gamma}$ invariant mass spectrum on top of a smoothly falling background. This help to separate the
 534 signal from the background processes contrary to other channels such as $b\bar{b}b\bar{b}$, $b\bar{b}WW^*$ and $b\bar{b}\tau^+\tau^-$. This
 535 thesis is performed on $b\bar{b}\gamma\gamma$ final state.

537 2.5.2 Di-Higgs boson as a probe of BSM physics

The small cross-section (Eq. 2.25) is making the di-Higgs boson observation challenging. If SM expectations hold, the production of a Higgs boson pair in a single pp interaction should not be observed with the Run-2 data unless its cross-section is enhanced by an anomalous component (new physics). This makes HH a promising process to probe new physics beyond the standard model (BSM). Many of those BSM theories predict the existence of heavy particles that can decay into a pair of Higgs bosons. These could be identified as a resonance in the di-Higgs boson invariant mass spectrum. In addition to the resonant production, there can also be non-resonant scenarios which can bring substantial enhancement of the cross-section by modifying the relative sign of Δ and \square , and by increasing the Δ which is proportional to λ_{HHH} . Only the non-resonant search is considered in this thesis. These can either originate from loop corrections involving new particles, such as a light-coloured scalar, or through non-SM couplings. Anomalous couplings of Higgs boson with the top-quark or triple Higgs boson self-coupling can either be extensions to the SM, such as contact interactions between two top quarks and two Higgs bosons, or deviation from the SM values of the trilinear Higgs boson self-coupling. Considering possible modifications of them, the deviation is quantified by $\kappa_\lambda = \frac{\lambda_{HHH}}{\lambda_{HHH}^{SM}}$ and $\kappa_t = \frac{\lambda_t}{\lambda_t^{SM}}$, where λ_i is the coupling i of the new physics and λ_i^{SM} its SM value.

Given the κ_t and κ_λ modifiers, the Higgs boson pair production cross-section can be parameterized as:

$$\sigma \approx k_t^4 \left[|\square|^2 + \frac{k_\lambda}{k_t} (\square \Delta + \Delta \square) + \left(\frac{k_\lambda}{k_t} \right)^2 |\Delta|^2 \right], \quad (2.26)$$

538 this shows that the production cross-section depends on both parameters κ_t and κ_λ , while the kinematics
 539 only depends on their ratio, that modifies the relative contribution of the two diagrams and thus the
 540 shape of the kinematic distributions. The maximum interference between the two diagrams corresponds
 541 to the cross-section minimum located at $\kappa_\lambda = 2.4\kappa_t$. Figure 2.18 shows an illustration of diagram
 542 contribution to the invariant mass distribution of di-Higgs boson system m_{HH} . The box diagram has
 543 an invariant mass spectrum peaking around $2m_t$. When including the triangle diagram with the triple
 544 Higgs boson self-coupling, the invariant mass spectrum becomes generally softer with the increase of
 545 its contribution. This effect causes a large change in m_{HH} distribution as shown in Figure 2.19. The
 546 interference of the two diagrams also generates local minima in the differential cross-section around
 547 $m_{HH} = 2m_t$ for the case of $\kappa_\lambda = 2$. Figure 2.20 displays the total LO and NLO cross-sections for the
 548 six dominant HH production channels at the LHC, as a function of the self-interaction coupling κ_λ .
 549 As mentioned above, Higgs boson pair production is a promising process to probe new BSM physics. In
 550 this thesis, in addition to the search for the SM Higgs boson pair production process, constrain on κ_λ is
 551 also extracted. In this thesis, κ_t is assumed to be one.

552 2.5.3 Current measurements

553 At the beginning of this thesis (2018), the latest measurement of Higgs boson pair cross-section in
 554 $HH \rightarrow b\bar{b}\gamma\gamma$ channel was performed based on dataset of 36 fb^{-1} representing the data collected by
 555 ATLAS detector between 2015 and 2016 (Chapter 3 is dedicated to ATLAS detector). Results of this
 556 analysis were consistent with the SM expectations [24]. The analysis set an observed (expected) upper
 557 limit at 95% Confidence Level (CL) on HH production cross-section of 0.73 (0.93) pb, corresponding
 558 to 22 (28) times the predicted SM value. The Higgs boson self-coupling was constrained to be between
 559 $-8.2 < \kappa_\lambda < 13.2$ at 95% CL ($-8.3 < \kappa_\lambda < 13.2$ expected), other SM parameters were fixed to their SM
 560 value when extracting limits. The limit scan as a function of κ_λ is shown in Figure 2.21.

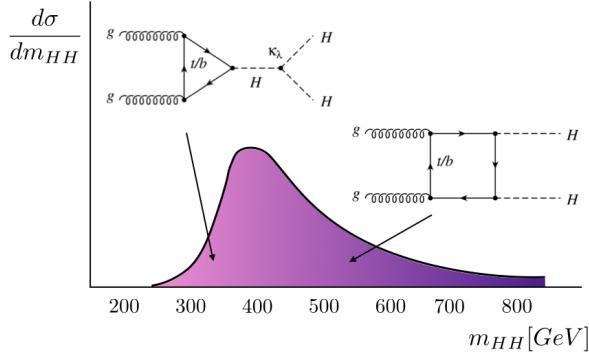


Figure 2.18: Illustration of both box and triangle diagrams contribution to di-Higgs boson invariant mass spectrum.

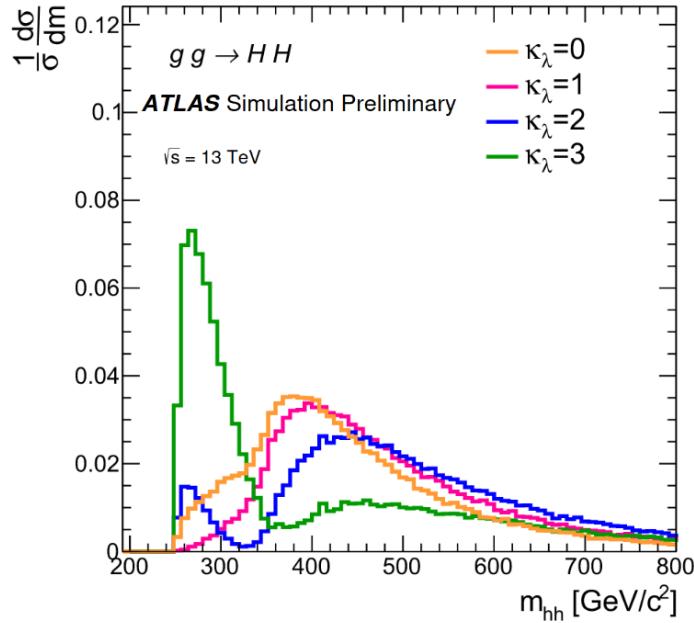


Figure 2.19: Higgs boson pair invariant-mass distribution in ggF for various values of κ_λ assuming $\kappa_t = 1$.

	Observed	Expected	-1σ	$+1\sigma$
$\sigma_{gg \rightarrow HH}$ [pb]	0.73	0.93	0.66	1.3
As a multiple of σ_{SM}	22	28	20	40

Table 2.2: The 95% CL observed and expected limits on the Higgs boson pair cross-section in pb and as a multiple of the SM production cross-section. The $\pm 1\sigma$ band around each 95% CL limit is also indicated.

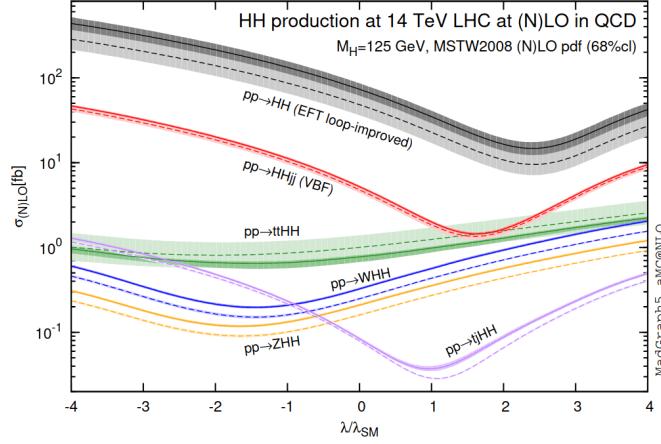


Figure 2.20: The total LO and NLO cross-sections for HH production at $\sqrt{s} = 14$ TeV, as a function of the self-interaction coupling κ_λ .

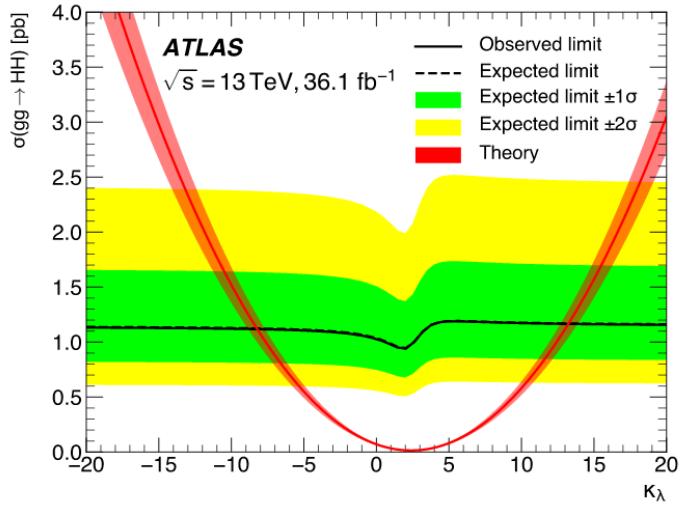


Figure 2.21: The expected and observed 95% CL limits on the non-resonant production cross-section $\sigma_{gg \rightarrow HH}$ as a function of κ_λ . The red line indicates the predicted HH cross-section as a function of κ_λ with all other couplings fixed at their SM values. The red band indicates the theoretical uncertainty of this prediction.

562 Improving this measurement by contributing to the $HH \rightarrow b\bar{b}\gamma\gamma$ non-resonant analysis with the
563 full Run-2 (2015-2018) data, and improve the current limits on the cross-section and constrain on the
564 self-coupling is the aim of this thesis.

565 2.6 Conclusion

566 The Higgs boson mechanism is a key component of the SM to explain the origin of the mass of the
567 particles. A particle with similar properties to the SM Higgs boson was discovered at collisions of
568 the LHC in 2012. After its discovery, a priority of the ATLAS and CMS collaborations has been
569 to measure more precisely its properties and couplings. Understanding Higgs boson self-coupling is
570 providing a direct probe on electroweak symmetry breaking (EWSB) and experimentally reconstruction
571 of the Higgs potential. The theoretical backgrounds needed to understand the main subject of the
572 thesis is introduced. The current HH search results with 2015-2016 data are also presented. The next
573 chapter focuses on the experimental setup to produce and collect data, the Large Hadron Collider and
574 the ATLAS detector.

575 **Chapter 3**

576 **The Large Hadron Collider and the ATLAS
577 detector**

578 The physics described in this thesis uses data collected by the ATLAS (A Toroidal LHC Apparatus)
579 detector from high-energy proton-proton collisions, accelerated and being collided by the Large Hadron
580 Collider (LHC). In terms of achievable centre-of-mass energy, LHC is currently the most powerful particle
581 accelerator. The structure, parameters, and principles of LHC and ATLAS, besides the complex sub-
582 detecting system, is introduced in this chapter.

583 **3.1 The Large Hadron Collider**

584 The LHC is a circular collider of hadrons with a circumference of 27 kilometres. This machine is
585 installed in a tunnel located underground between 50 m and 175 m deep, and was built between 1984
586 and 1989 for the LEP e^+e^- machine. In 2001, LEP was dismounted to give way to the LHC. It was
587 designed to accelerate and collide proton beams into a centre-of-mass energy up to 14 TeV. It also
588 accelerates and collides heavy ions, in particular lead nuclei (Pb), at 2.3 TeV per nucleon. Inside this
589 circular accelerator, a set of protons (or ions) so-called bunch racing clockwise close to the speed of light
590 (99.9999991%) collides with another bunch speeding anticlockwise. The energy involved in the collisions
591 is so high that in a sub-microscopic region at the heart of the collisions, it briefly generates conditions
592 similar to those that occurred shortly after the birth of the universe.

593 Within the tunnel are two adjacent parallel beam pipes and surrounded by superconductive magnets. In
594 total, 1232 dipole magnets bend the beams into their circular orbit and 392 quadrupole magnets to focus
595 the beams. The strength of the focusing magnets is required to be high to squeeze the transverse beam
596 sizes and, thus, to increase the probability of collisions. The adopted design of the LHC is approximately
597 80% of the arcs filled with dipole magnets. Dipoles are also equipped with sextupoles, octupoles and
598 decapoles to correct non-linear dynamics of the beams. Keeping 7 TeV proton energy beam on the
599 designed orbit implies the use of magnetic bending fields of 8.4 T. Generation of such field required
600 using superconducting magnets at the limit of the existing technologies. Approximately 96 tonnes of
601 liquid helium is needed to maintain the superconductivity of the magnets at an operational temperature
602 of 1.9 K (-271.3°C), making the LHC the largest cryogenic facility in the world. A detailed description
603 of the LHC and the CERN accelerator complex is given in Ref. [25].

3.1.1 Acceleration chain

A succession of small to large accelerators is used to accelerate the protons extracted from hydrogen gas to the energy needed for injection into the LHC. Figure 3.1 displays the CERN accelerator complex including all pre-acceleration steps before the LHC.

CERN's Accelerator Complex

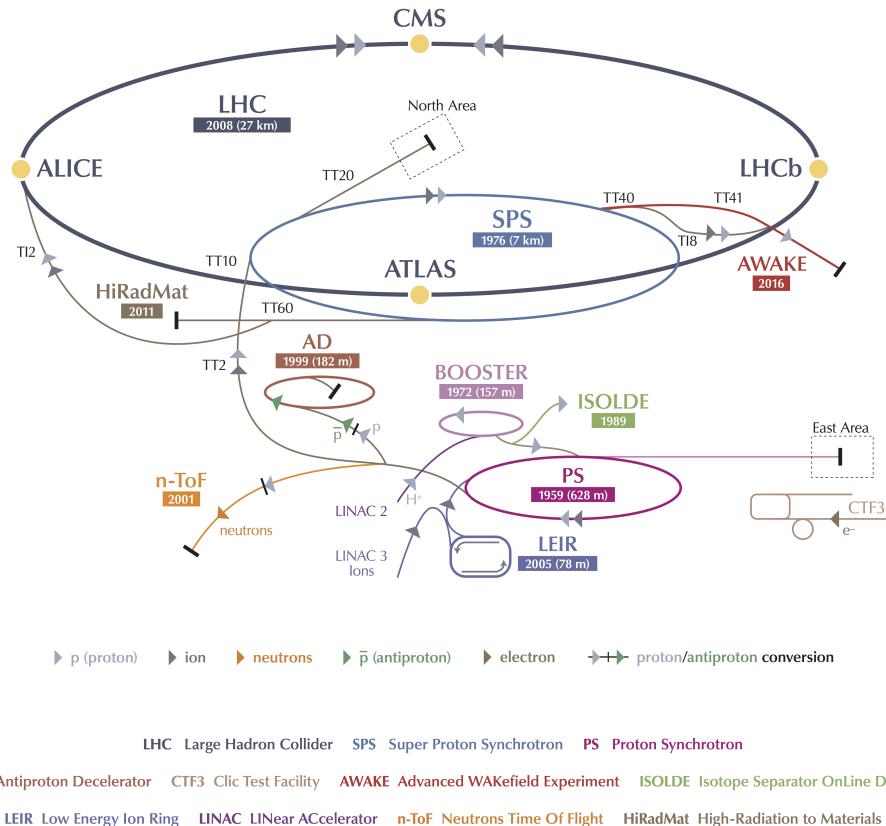


Figure 3.1: Overview of the CERN accelerator complex, including the LHC and its pre-accelerators [26]. The four main LHC experiments are depicted, too.

The process of acceleration starts from the linear accelerator Linac 2, which accelerates protons up to 50 MeV. The beam is then injected into the Booster. The Booster accelerates the beam to 1.4 GeV and feeds the Proton Synchrotron (PS), where protons are further accelerated to 25 GeV. The next chain is the Super Proton Synchrotron (SPS), which is 6.9 km long. Here protons reach the energy of 450 GeV before they are transferred to two beam-pipes of the LHC main ring. It takes several minutes to fill the LHC ring and about 15 minutes to accelerate beams to their maximum energy of 6.5 TeV using eight radio frequency (RF) cavities at $f_{RF} = 400$ MHz. Each time a beam passes the electric field in the RF cavity, some energy from the radio waves is transferred to the particles, nudging them forwards. The beams are injected in bunches spaced by 25 ns. Each bunch contains approximately 10^{11} protons. Each beam contains 2808 bunches.

Protons are colliding in four interaction points (IP). At each point, the created particles are distributed over the full solid angle around the IP where four detectors record those particles and identify

them are installed. They are ATLAS, CMS, ALICE and LHCb. LHCb was built to study flavour physics looking at the properties of b -hadrons, and the ALICE detector is specialized on measurements on heavy-ion collisions. CMS and ATLAS are general-purpose detectors. They allow making precision measurements of SM processes, including the properties of the Higgs boson, and to search for BSM physics. Section 3.2 is dedicated to ATLAS detector description.

3.1.2 Luminosity

The number of produced events is proportional to the integrated luminosity \mathcal{L}_{int} multiplied by the total cross-section:

$$N_{events} = \int \mathcal{L} dt \times \sigma_{process}. \quad (3.1)$$

The instantaneous luminosity \mathcal{L} is the quality factor for colliders, measuring the intensity of the beam, and is defined as:

$$\mathcal{L} = \frac{N_b^2 n_b f_r \gamma_r}{4\pi \epsilon_n \beta^*} F, \quad (3.2)$$

where for the design luminosity (nominal parameters for the LHC are given in parentheses):

- N_b is the number of particle per bunch ($\sim 10^{11}$).
- n_b is the number of bunch per beam (2808).
- f_r is the revolution frequency (11245 Hz).
- γ_r is the relativistic γ factor (~ 7000).
- ϵ_n is the normalized traverse beam emittance which characterizes its spread in coordinate and momentum phase space (3.75 μm).
- β^* is the beta function at the collision point determined by the magnet configuration (for ATLAS 0.55 m).
- F is the geometric luminosity reduction factor due to the crossing angle at the interaction point.

For standardize the analysis procedure, ATLAS has defined a basic time unit called Luminosity Block (LB) where the instantaneous luminosity is assumed to be stable. The typical LB duration is one to two minutes. Data are analysed under the assumption that each LB contains data taken under uniform conditions (data quality). To define a data sample for physics, quality criteria are applied to select LBs where the conditions are acceptable. The average luminosity in the LB is multiplied by the LB duration to provide the integrated luminosity delivered in the given LB.

The design luminosity of LHC is $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. The integrated good quality data of Run-1 is approximately 25 fb^{-1} as shown in Figure 3.2. During the first long shutdown (LS1), the LHC beam energy was increased from 3.5 TeV to 6.5 TeV. Figure 3.3 shows the delivered and recorded integrated luminosity during the Run-2 data taking [27]. The integrated luminosity of Run-2 associated to good data period is $\mathcal{L}_{int} = 139 \text{ fb}^{-1}$. The analysis described in this thesis is performed with Run-2 data.

Knowing the cross-section of the $HH \rightarrow b\bar{b}\gamma\gamma$ production, one can evaluate the number of events available for the analysis as $N_{HH \rightarrow \gamma\gamma b\bar{b}} = \mathcal{L}_{int} \cdot \sigma_{pp \rightarrow HH} \cdot Br(HH \rightarrow \gamma\gamma b\bar{b})$ that corresponds to about 12 events.

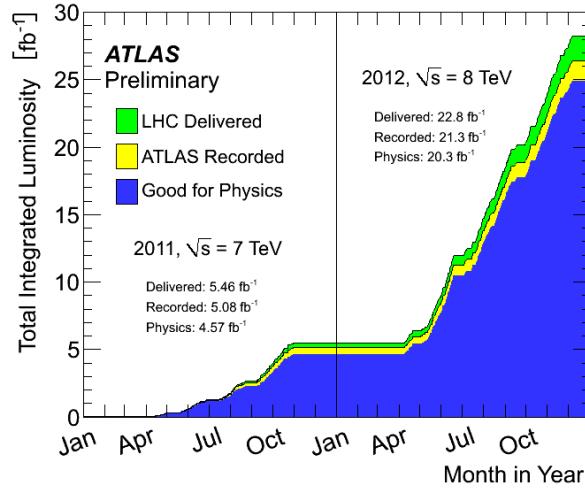


Figure 3.2: Cumulative luminosity versus time delivered to (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams and for pp collisions at 7 and 8 TeV centre-of-mass energy in 2011 and 2012 (Run-1).

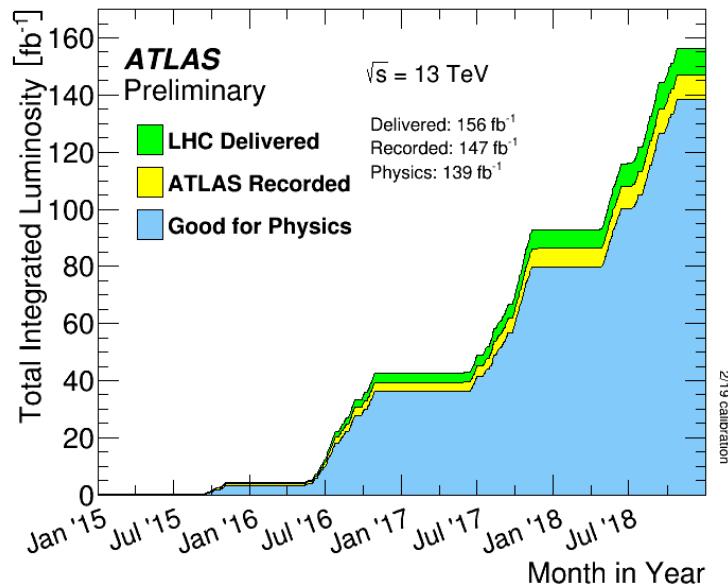


Figure 3.3: Luminosity delivered by the LHC during the Run-2 data taking. ATLAS recorded this data with an efficiency above 90%.

651 3.1.3 Pile-up events

652 Because of the very high proton density at the collision points, more than one proton interact when
 653 two LHC bunches cross each other at the centre of the experiment. This is commonly referred to as
 654 "pile-up". On top of the usual *in-time* pile-up, defined as the collision events occurring during the
 655 same bunch-crossing as the event of interest, one also has to consider *out-of-time* pile-up, coming from
 656 remnants of information found in some of the detector subsystems that end up being attributed to the
 657 wrong bunch-crossing, and therefore to the wrong event typically from previous collisions. Figure 3.4
 658 shows the average number of simultaneous interactions per bunch crossing for Run-2.

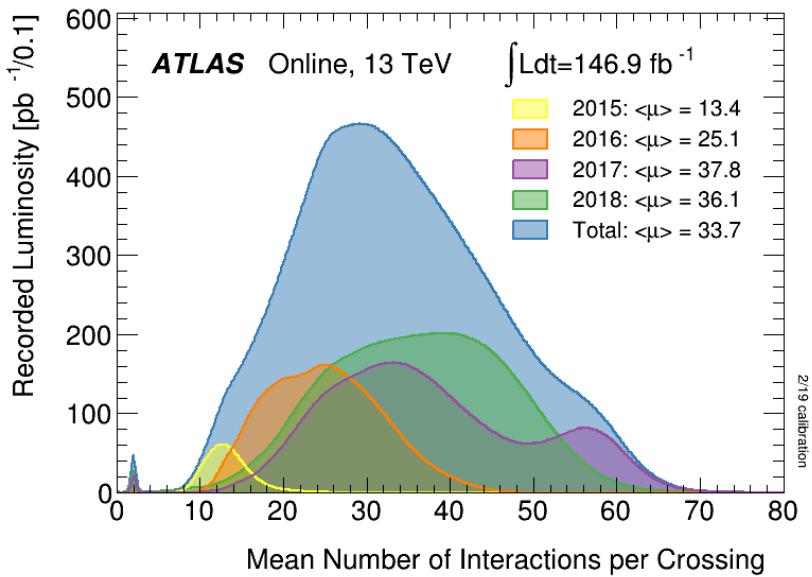


Figure 3.4: Recorded integrated luminosity as function of the mean number of interaction per bunch crossing in pp collisions recorded by the ATLAS detector during Run-2 [27].

659
 660 The presence of pile-up interactions significantly increase the activity inside the detector which degrades
 661 the performances of different sub-detector components and algorithms efficiency.

662 3.2 ATLAS : A Toroidal LHC ApparatuS detector

663 The ATLAS detector is one of the four experiments placed on the interaction points of the LHC beams.
 664 It is currently the largest experiment of particle physics with a length of 46 m along the beam pipe
 665 and a transverse diameter of 25 m. Its weight is larger than 7000 tons [28]. It is a superposition of
 666 four sub-detectors, each optimized for the identification and the measurement of a specific category
 667 of particles: Inner Tracker, Electromagnetic Calorimeter (ECAL), Hadronic Calorimeter (HCAL) and
 668 Muon Spectrometer. It is composed of a central component called Barrel and two End-Cap to cover
 669 the 4π solid angle. Its geometry is optimized to measure particle properties produced orthogonally to
 670 the beam pipe and allow for detection of forward particles to estimate the energy of invisible particles.
 671 The detector has been operating since 2008, taking alignment data with cosmic rays before the LHC
 672 collisions start, and its data is exploited by a collaboration of about 3000 scientific authors from 181
 673 institutions in 38 countries. An overview sketch of the ATLAS detector is shown in Figure 3.5.

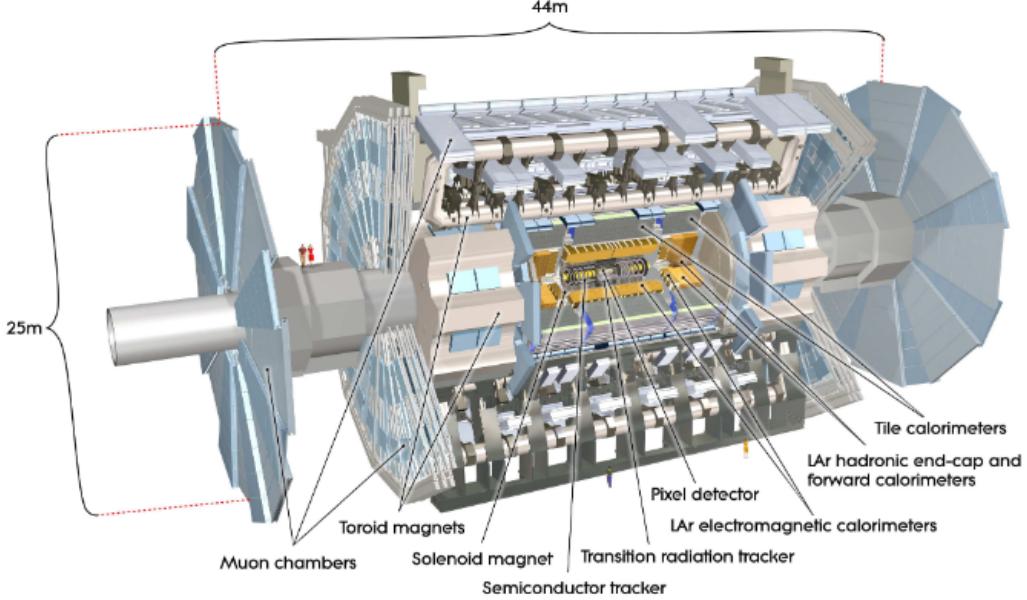


Figure 3.5: Sketch of the ATLAS detector with its different sub-detectors.

⁶⁷⁴ 3.2.1 System of coordinates

The coordinate system used by the ATLAS experiment is cylindrical where the z-axis is along the LHC beam pipe, the x-axis pointing toward the centre of the LHC ring and the y-axis pointing upwards. A physic object (particle) is characterized by its transverse component of the three-momentum $p_T = \sqrt{p_X^2 + p_Y^2}$, its azimuthal angle $\phi \in [-\pi, \pi]$ formed by the three-momentum and the x-axis and its polar angle $\theta \in [0, \pi]$, i.e., the angle between the three-momentum and the z-axis. Figure 3.6 shows the coordinate system common to ATLAS and CMS experiments.

The polar angle is expressed in terms of the pseudo-rapidity η , defined as:

$$\eta = -\log[\tan(\theta/2)]. \quad (3.3)$$

In collisions involving protons, the adoption of η instead of θ ensures the detector balance over particles and particles distribution recorded is approximately flat in η :

$$\frac{\partial \sigma_{QCD}}{\partial \eta} = cte. \quad (3.4)$$

For relativistic particles, the pseudo-rapidity η corresponds to the rapidity y , defined as:

$$y = \frac{1}{2} \left(\frac{E + p_Z}{E - p_Z} \right), \quad (3.5)$$

⁶⁷⁵ where E is the particle energy.

⁶⁷⁶ 3.2.2 Inner Tracker

⁶⁷⁷ The Inner Detector (ID) has been designed to detect and reconstruct the path of the electrically charged
⁶⁷⁸ particle bent by a 2 T solenoid magnetic field. It also provides a good momentum resolution by recon-
⁶⁷⁹ structing the curvature and the direction, and both primary and secondary vertex measurements for

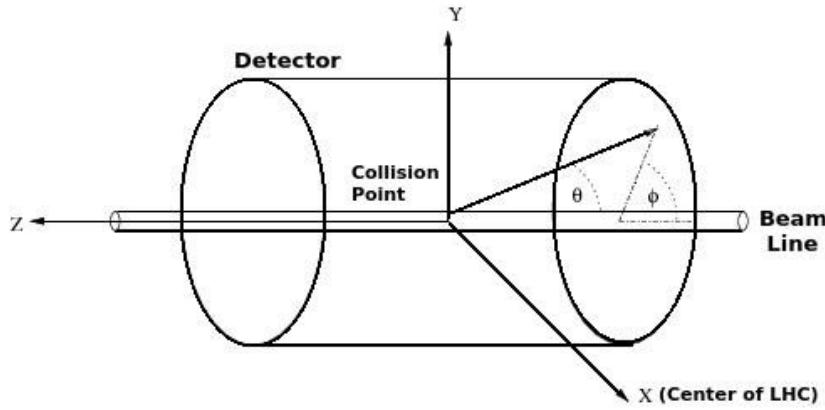


Figure 3.6: Coordinate system used by the ATLAS and CMS experiments at the LHC.

680 tracks above approximately 0.5 GeV [29, 30]. In terms of acceptance, the ID covers the region up to
 681 $|\eta| \leq 2.5$. To achieve the momentum and vertex resolution requirements imposed by the physics goals
 682 at the LHC and the very large track density environment, the ID high-precision measurements must
 683 have fine detector granularity. The ID is composed of four complementary sub-detectors: IBL, the Pixel
 684 Detector, the Semi-Conductor Tracker (SCT) and the Transition Radiation (TRT). A magnetic field of
 685 2 T is provided by a solenoid inserted between the ID and the EM calorimeter. The layout of the Inner
 Detector (ID) is illustrated in Figure 3.7.

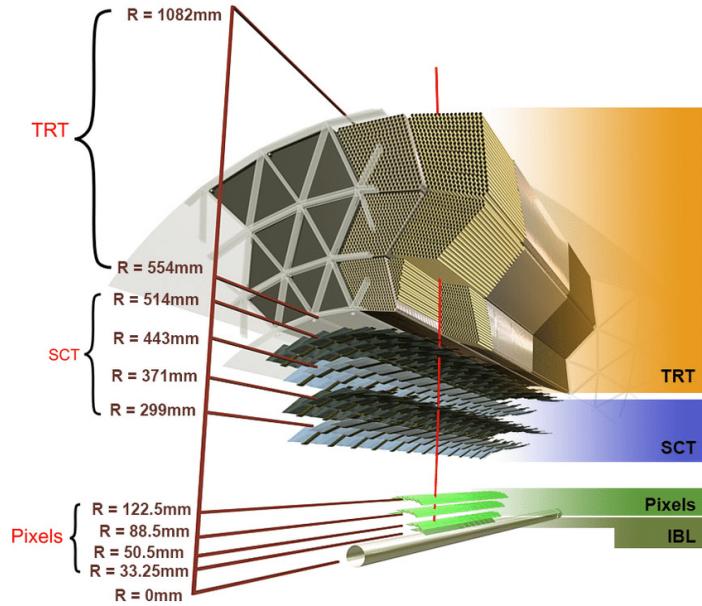


Figure 3.7: Layout of the Inner Detector (ID) [31].

687 3.2.2.1 IBL

688 In 2014, during the first LHC long shutdown (LS1), the ATLAS pixel detector was upgraded with
 689 an additional pixel layer installed close to the beam pipe called Insertable B-Layer (IBL) [32]. Its
 690 motivations are:

- 691 • Increase the number of measurement point of tracks to improve their reconstruction efficiency.

- 692 • Improve the identification of the primary vertex which plays an important role in b -jet identification
 693 (b -tagging), which in turn significantly improves the sensitivity of many analyses. Inefficiencies in
 694 the other layers can be partially compensated during the reconstruction at the cost of an increased
 695 fake rate, the IBL restore the full b -tagging efficiency even in case of a complete B-layer failure.
- 696 • Luminosity effects: The pre-Run-2 pixel detector was designed for a peak luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$.
 697 With high luminosity the number of tracks from pile-up increases, leading to high occupancy that
 698 can induce readout inefficiencies, would thereby limits the b -tagging efficiency. The addition of the
 699 IBL layer helps to preserve tracking performance in face of luminosity increase.
- 700 Strong constraints and project specifications have a substantial impact on the technologies required for
 701 the IBL. IBL covers the region up to $|\eta| < 2.58$ and is located at a mean radius of 33.2 mm around
 702 the beam pipe. The size of IBL pixels is $50 \times 250 \mu\text{m}$. The IBL consists of 14 staves each containing 20
 703 modules made of planar sensors. The temperature of the IBL is controlled using a bi-phase CO_2 cooling
 704 system. Figure 3.8 shows the IBL within the Pixel Detector volume and around the beam pipe.

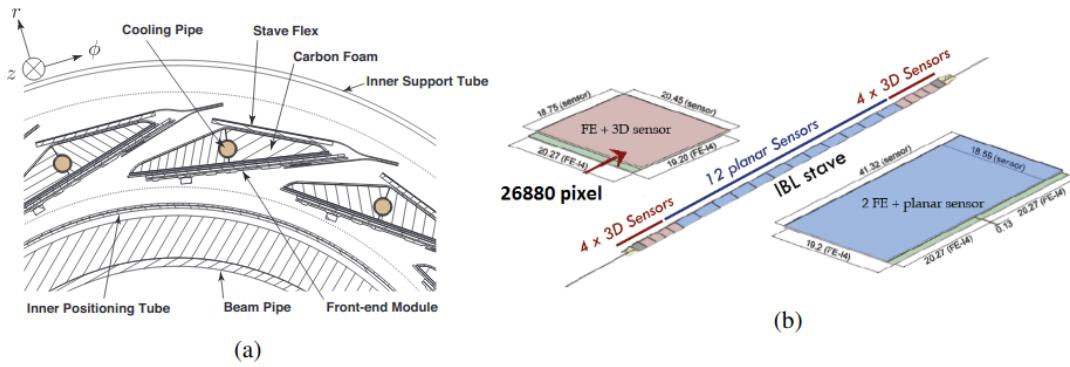


Figure 3.8: (a) Transverse view of 3 of the Insertable-B-Layer (IBL) staves, located directly on the beam pipe. (b) The layout of one of the 14 IBL staves [31].

704 The additional measurement point provided by IBL improves significantly the reconstructed parameters
 705 by the tracker. Figure 3.9 shows the improvement in impact parameter resolution thanks to the IBL
 706 as measured from early Run-2 data with respect to Run-1. In addition to a factor of 4 reduction in
 707 reconstruction speed [34], Figure 3.10 shows the improvement achieved with IBL for the performance of
 708 tracking in dense environments (TIDE). Improvements in tracking have a direct impact on vertex iden-
 709 tification and b -tagging. Figure 3.11 shows the improvement in the b -tagging efficiency with IP3D+SV1
 710 b -tagging algorithms compared to Run-1 layout and using new algorithm [36, 37]. The IP3D and SV1
 711 algorithms will be explained later in the thesis.

713 3.2.2.2 Pixel Detector

714 With the existing technology in the 2000s, the Pixel Detector (PD) installed at the start of ATLAS was
 715 designed to provide high-granularity, high-precision measurements as close as possible to the interaction
 716 point. It consists of three barrel layers placed at the radius of 50.5 mm, 88.5 mm and 122.5 mm centred
 717 around the beam axis and two end-cap with three disc layers each positioned at $|z| = 495.580$ and 650
 718 mm. It provides three measurement points per track. The system is designed to be highly modular,
 719 containing approximately 1500 identical barrel modules and 1000 identical disk modules, each module is
 720 composed of 61440 pixel elements of silicon semi-conductor. In total there are about 80 million readout

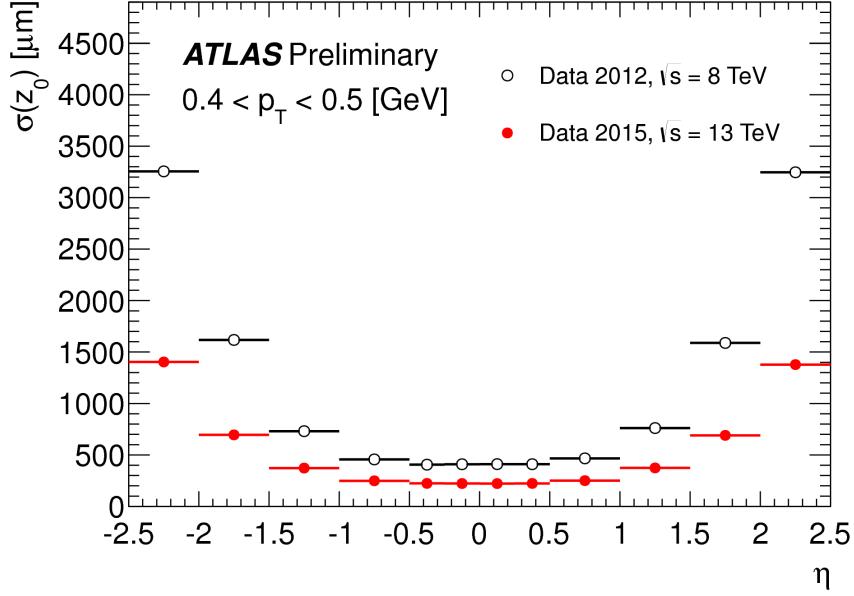


Figure 3.9: Unfolded longitudinal impact parameter resolution measured from data in 2015, $\sqrt{s} = 13 \text{ TeV}$, with the Inner Detector including the IBL, as a function of η for values of $0.4 < p_T < 0.5 \text{ GeV}$ compared to that measured from data in 2012 [33].

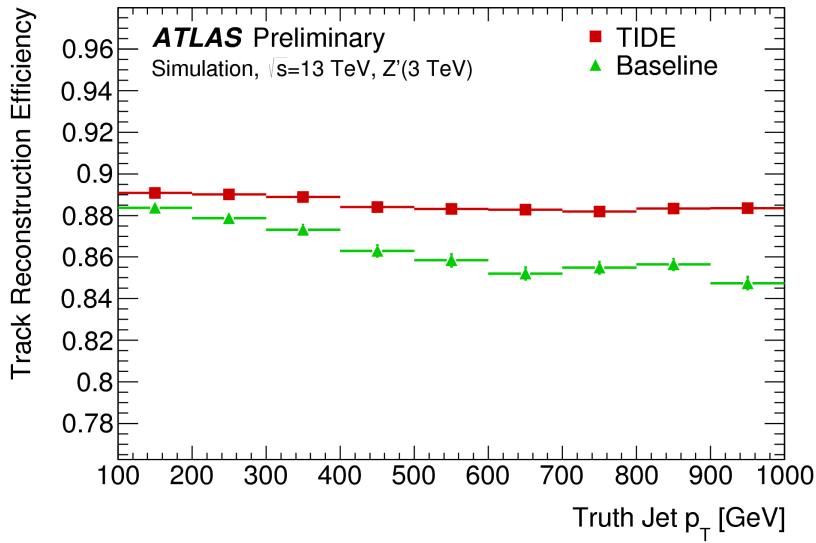


Figure 3.10: The average efficiency to reconstruct primary tracks with a production vertex before the first layer in jets as a function of jet p_T . The same sample generation, with limited statistics, is used for both reconstruction algorithms, resulting in correlated features [35].

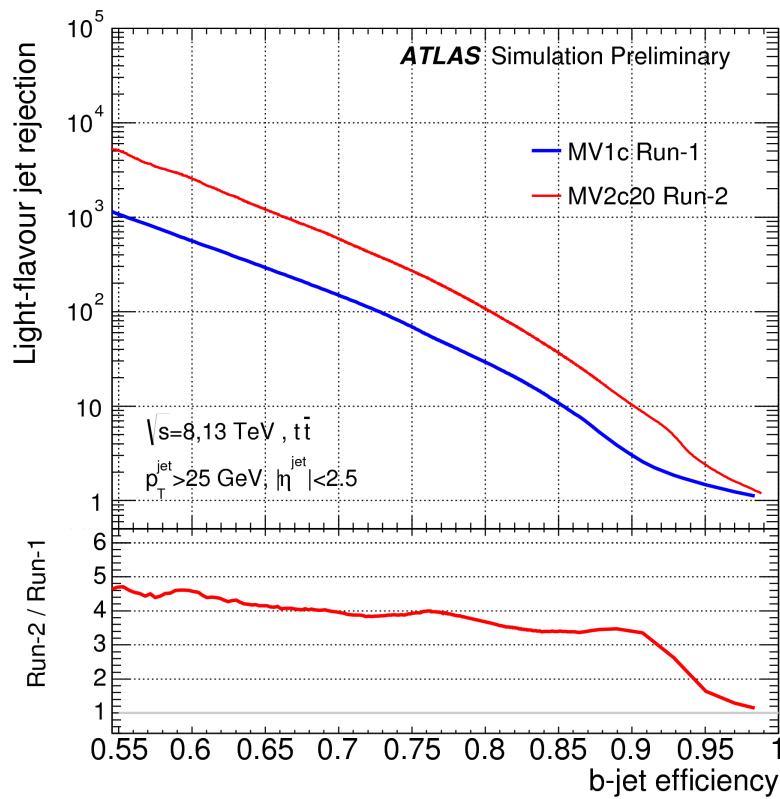


Figure 3.11: Rejection factor against light jets as a function of b -jet efficiency for the combined IP3D+SV1 tagger. Run-1 and Run-2 performances are compared.

⁷²¹ channels in the whole PD. The spatial resolution for the barrel modules is $10 \mu\text{m}$ in $r\phi$ and $66 \mu\text{m}$ in
⁷²² z , for the end-caps the spatial resolution in $r\phi$ is the same as the barrel and $115 \mu\text{m}$ in z . The main
⁷²³ limitation of the pixel detector is the radiation hardness, as the expected flounce is at the tolerable limit.

⁷²⁴ 3.2.2.3 Semi-Conductor Tracker

⁷²⁵ The SCT system is designed to provide four precision points per track in the intermediate radial range,
⁷²⁶ contributing to the measurement of momentum, impact parameter and vertex position. The barrel and
⁷²⁷ end-caps SCT are four layers of silicon microstrip for barrel and nine disks for end-caps. The spatial
⁷²⁸ resolution is $16 \mu\text{m}$ in $r\phi$ for both the barrel and the end-caps. The four complete barrels are positioned
⁷²⁹ in a radius of 300, 373, 447 and 520 mm. Tracks can be reconstructed if separated by more than ~ 200
⁷³⁰ μm . There are 6.3 millions readout channels for the SCT.

⁷³¹ 3.2.2.4 Transition Radiation Tracker

⁷³² The TRT is positioned at the outer part of the ID. It consists of 370000 drift tubes called straws.
⁷³³ Each straw has a diameter of 4 mm and a length of 1.44 m. The straws are filled with a gas mix-
⁷³⁴ ture of 70% Xe , 27% CO_2 and 3% O_2 . Its wall acts as a cathode kept at high voltage. The anode
⁷³⁵ is a $30 \mu\text{m}$ diameter plated tungsten wire placed in the centre of the straw. When a charged particle
⁷³⁶ crosses a straw, it ionizes the gas and the produced electrons travel through the anode generating an
⁷³⁷ electric signal. To keep the TRT performance constant, the close-loop gas system is used, maintain-
⁷³⁸ ing the correct gas fractions. The straws are arranged to be parallel to the beam-pipe in the barrel
⁷³⁹ and perpendicular in the end-cap region. There are about 50k straws in the barrel and 320k straws in
⁷⁴⁰ the end-cap providing high-precision measurement for each track. The radial resolution is about $130 \mu\text{m}$.

⁷⁴¹ In the Perigee representation, tracks are described using the parameters of a helicoidal trajectory at
the point of the closest approach to the z -axis: the transverse impact parameter d_0 , the z coordinate z_0 ,
the angles θ and ϕ and the inverse of the particle momentum multiplied by the charge q/p , as illustrated
in Figure 3.12.

The expected momentum resolution of the inner detector at Run-1 is given by:

$$\sigma(1/p_T) \cdot p_T = 0.036\% \cdot p_T[\text{GeV}] \oplus 1.3\%, \quad (3.6)$$

⁷⁴² \oplus denote the quadrature addition.

⁷⁴³ 3.2.3 Calorimeter system

⁷⁴⁴ The ATLAS calorimeter system is designed to provide a precise energy measurement and position
⁷⁴⁵ reconstruction for electromagnetic (EM) particles (electrons, photons) and jets (hadrons). The good
⁷⁴⁶ hermiticity of the calorimeter ($|\eta|$ up to 5) also allows to measure the missing transverse energy and
⁷⁴⁷ provides the separation of electrons and photons from hadrons and jets. The calorimeter system is
⁷⁴⁸ composed of two calorimeters: the electromagnetic calorimeter (ECal) and the hadronic calorimeter
⁷⁴⁹ (HCal). Both are sampling calorimeters, with alternating layers of a heavy absorber material and an
⁷⁵⁰ active material in which an ionization signal is produced. Figure 3.13 shows a three-dimensional view
⁷⁵¹ of the ATLAS calorimeter system.

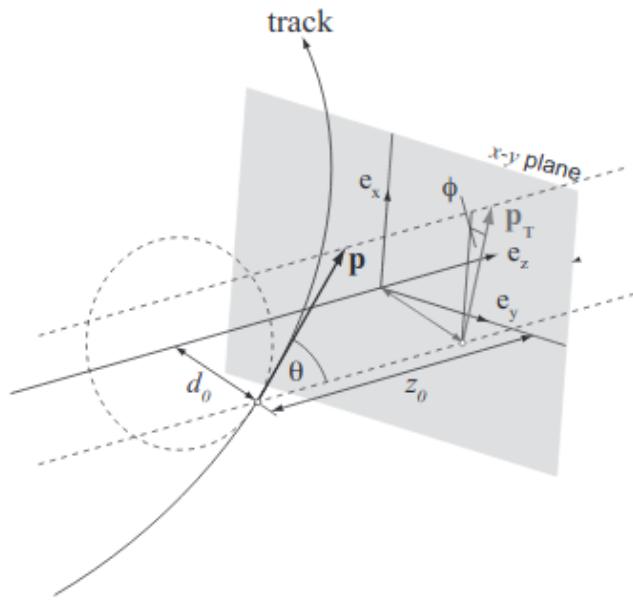


Figure 3.12: The Perigee representation of the track [38].

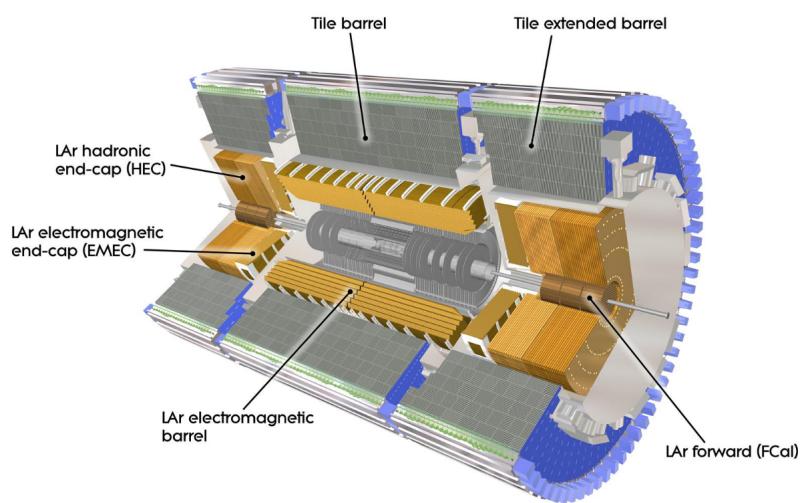


Figure 3.13: ATLAS Calorimeter system.

⁷⁵² **3.2.3.1 Electromagnetic calorimeter**

The ECal is the first sub-detector after the ID. It is optimized for the energy reconstruction of electrons and photons by triggering EM shower and measure its properties [39]. It covers the region of $|\eta| < 3.2$ excluding the region $1.375 < |\eta| < 1.52$ which corresponds to the transition region between the barrel and end-caps. The barrel part covers $|\eta| < 1.475$, while the two end-caps cover $1.375 < |\eta| < 3.2$. The barrel and end-caps are composed of alternated layers of absorbing material lead plates (~ 1 mm thick), to enforce the development of the whole EM showers within EM longitudinal envelop and separated by active medium Liquid Argon (LAr) of 2 mm thick. The advantages of LAr are its radiation hardness, intrinsic linear behaviour, cheapness compared to other noble gases. It has been considered as outweighing the drawbacks associated with the need of cryostats and signal feed-throughs. The total thickness of the calorimeter is at least 22 radiation lengths in the barrel, and more than 24 radiation lengths in the end-caps. The radiation lengths X_0 is defined as the scale after which high-energy electrons lose all but $1/e$ of their initial energy.

Electrons and photons entering the calorimeters initiate electromagnetic cascades, in which e^+e^- pair production and bremsstrahlung processes occur. High-energy electrons predominantly lose energy in matter through bremsstrahlung, while high-energy photons create e^+e^- pair. EM particle develops a shower until its energy falls below the critical energy E_c . E_c is defined as the energy for which the energy loss per X_0 due to ionization of the material is equal to the particle energy. In lead, $E_c = 7.4$ MeV for electrons. The e^+/e^- of EM shower ionizes atoms of the LAr generating an electric signal proportional to the energy deposited by the particle. The ionization is drifted to the electrode under electric field generated by a high voltage of 2000 V. To provide a large signal response and hermiticity, the ATLAS Collaboration adopted a particular geometry for the ECal: *accordion* geometry. In the barrel, the accordion waves are axial and run in ϕ ; the folding angles of the waves vary with radius to keep the liquid-argon gap constant and reduce the non-instrumented zones. In the end-caps, the waves are parallel to the radial direction and run axially. The size of the drift gap on each side of the electrode is 2.1 mm. Figure 3.14 shows the accordion shape of the EM calorimeter.

The ECal is further segmented in three longitudinal layers, to measure the longitudinal shower development. They are called respectively strip, middle, and back. It has different EM cells granularity $\Delta\eta \times \Delta\phi$ per layer. Cells of the middle layer (Lr2) in the barrel region have a size of 0.025×0.025 , while for the strip (Lr1) cells are 8 times finer in the $|\eta|$ direction providing a precise η measurement of incident particles. The back layer (Lr3) cells has a twice coarser granularity in η and the same ϕ segmentation as in Lr2. In front of the EM barrel calorimeter there is a Presampler (PrS) LAr detector (Lr0), covering range $|\eta| < 1.8$ and placed to start the shower before the calorimeter. The PrS has the finest granularity with a cell size of $\Delta\eta \times \Delta\phi = 0.003 \times 0.1$ used for $\pi^0 \rightarrow \gamma\gamma$ background separation. The number of samplings and the granularity in each of the samplings are summarized in Table 3.1.

Sampling	$ \eta < 1.5$	$1.5 < \eta < 1.8$	$1.8 < \eta < 2.0$	$2.0 < \eta < 2.5$	$2.5 < \eta < 3.2$
Presampling	0.025×0.1	0.025×0.1			
Strip	$0.025/8 \times 0.1$	$0.025/8 \times 0.1$	$0.025/6 \times 0.1$	$0.025/4 \times 0.1$	0.1×0.1
Middle	0.025×0.025	0.025×0.025	0.025×0.025	0.025×0.025	0.1×0.1
Back	0.050×0.025	0.05×0.025	0.05×0.025	0.05×0.025	

Table 3.1: Granularity of the EM calorimeter ($\Delta\eta \times \Delta\phi$) [39].

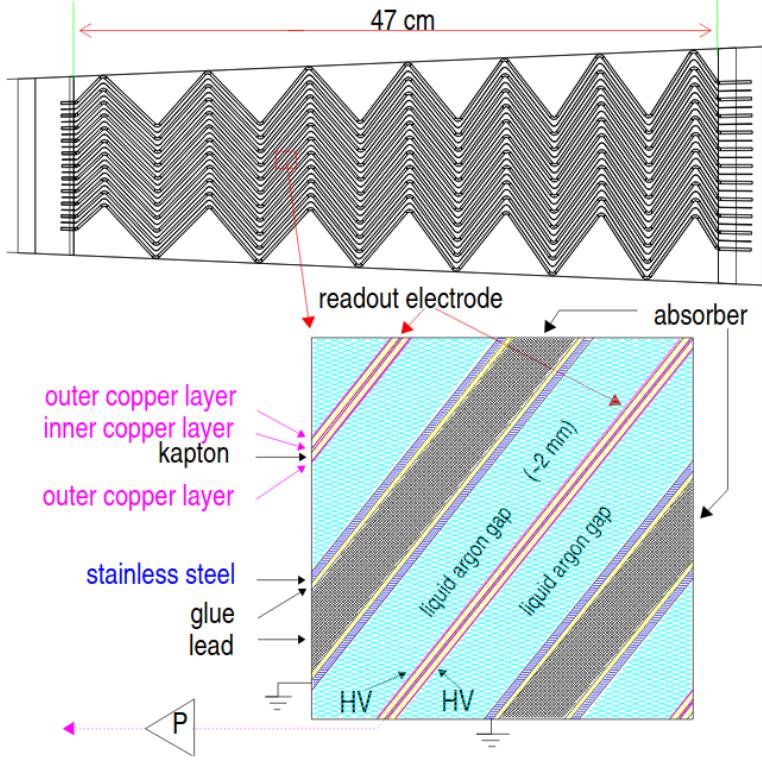


Figure 3.14: Accordion shape of the EM calorimeter.

The target EM calorimeter energy resolution is given by:

$$\sigma_E/E = \frac{10\%}{E} \oplus 0.7\%. \quad (3.7)$$

3.2.3.2 Hadronic calorimeter

At high-energy colliders, quarks and gluons fragment into a shower of particles called jets. The Hadronic Calorimeter (HCal) completes the measurement of the jet energy initiated by the ECal [40]. Hadronic showers are larger than electromagnetic ones, thus HCal needs to be large enough to contain the hadronic shower and reduce the punch-through hadrons reaching the muon system. The total thickness was chosen to be about $11 X_0$ to obtain good performances on the resolution for high-energy jets. The hadronic calorimeter is divided into the Tile calorimeter and the LAr hadronic end-caps calorimeter. The barrel calorimeter is made of steel as absorbing material and scintillating plastic tiles as the active medium and covers the range $|\eta| < 1$ with two extensions in the range $0.8 < |\eta| < 1.7$. The end-caps cover the range $1.5 < |\eta| < 3.2$, and they are composed of two wheels made of parallel copper plates with LAr as active material in between. For the determination of the missing energy, a good hermetic coverage is essential. Therefore, the ATLAS calorimeter is also equipped with a calorimeter covering the very forward region of $3.1 < |\eta| < 4.9$, the LAr Forward Calorimeter (FCal).

Tile: provides signal by the tiles scintillation. Tiles are 3 mm thick and perpendicular to the beam pipe. It consists of a barrel ($|\eta| < 1.0$) and two extended barrels. The Tile calorimeter is segmented into three layers. The dimension of the cells corresponds to $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the first two layers and $\Delta\eta \times \Delta\phi = 0.2 \times 0.1$ in the last layer. A vertical gap of 68 cm wide between the barrel and extended

771 barrel regions is used for the transition from the ID and the EM calorimeter.

772

LAr: covers end-cap and forward regions with $1.5 < |\eta| < 4.9$. Each end-cap calorimeter (HEC) consists of two independent wheels of equal diameter with copper absorber plates. The end-cap calorimeter is divided into front, middle and back longitudinal layers. The FCal is placed at a distance of about 5 meters from the interaction point. It is a high-density detector facing a very high particle flux and consisting of three wheels on each side employing liquid argon as an active material. The innermost wheel is optimized for electromagnetic showers and employs copper as the absorber. While the other two wheels measure hadronic showers using tungsten as absorbing material. The granularity of the hadronic LAr calorimeter is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for $1.5 < |\eta| < 2.5$ and $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ for $2.5 < |\eta| < 3.2$ while the forward calorimeter has $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$.

The HCal is designed to measure the energy with a target resolution of [41]:

$$\sigma_E/E = \frac{50\%}{E} \oplus 3\%. \quad (3.8)$$

773 3.2.4 Muon Spectrometer

774 Muons are minimally ionizing particles due to their relatively high mass, therefore are not stopped by
 775 the calorimeter system. The Muon Spectrometer (MS) is the outermost part of the ATLAS detector
 776 dedicated to detect muons exiting the calorimeter and measure their momentum in the range of $|\eta| <$
 777 2.7 [42]. A large superconducting air-cored toroid magnet is used to bend muons trajectories. Over
 778 the range of $|\eta| < 1.4$, magnetic bending is provided by the large barrel toroid. For $1.6 < |\eta| < 2.7$
 779 region, muon tracks are bent by two smaller end-cap magnets inserted into both ends of the barrel
 780 toroid. Over $1.4 < |\eta| < 1.6$, usually referred to as the transition region, magnetic deflection is provided
 781 by a combination of barrel and end-cap fields. Two different functions are accomplished by the MS:
 782 triggering and high-precision tracking. The tracking is performed by the Monitored Drift Tube (MDT)
 783 and by Cathode Strips Chamber (CSC) at large pseudo-rapidity. However, the trigger system covers
 784 the region up to $|\eta| < 2.4$, and it is composed of Resistive Plate Chamber (RPC) in the barrel and Thin
 785 Gap Chamber (TGC) in the end-caps. The conceptual layout of the spectrometer is shown in Figures
 786 3.15.

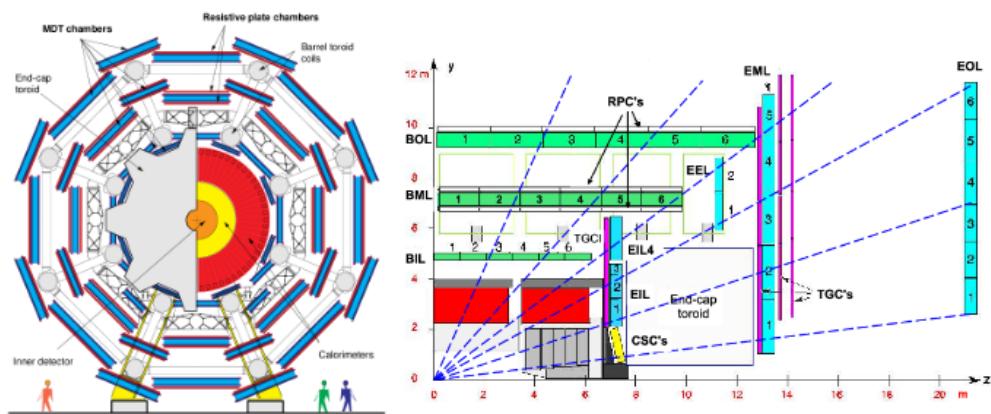


Figure 3.15: Side view of one quadrant (*right*) and transverse view (*left*) of the muon spectrometer.

787

The MDTs are aluminium drift tubes with a diameter of 30 mm as Figure 3.16 shows, operating with a gas mixture of 93% of Ar and 7% of CO_2 at 3 bar pressure. Each chamber consists of two sections with three (inner station) or four (middle and outer station) layers of the drift tubes. Tungsten-rhenium wire of 50 μm collects the electrons resulting from ionization at a potential of 3 kV. The maximum drift time can reach about 700 ns. The MDTs are designed for precise tracking of muons and cover most of the MS pseudorapidity total coverage, with a single-hit resolution of about 35 μm per chamber.

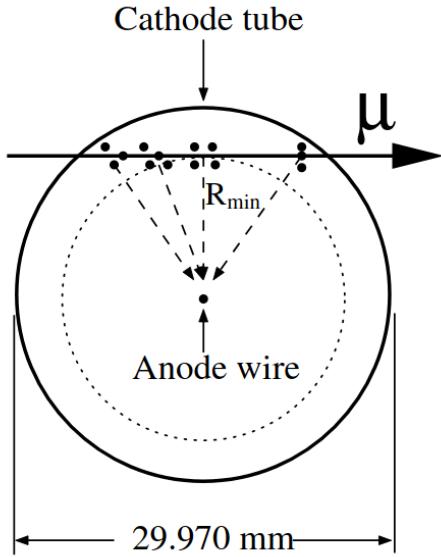


Figure 3.16: Transverse cross-section of the MDT tube.

CSCs are multi-wire proportional chambers filled with a mixture of 80% of Ar and 20% of CO_2 with cathode planes segmented into strips in an orthogonal direction to the beam axis. The CSCs cover the range of $2.0 < |\eta| < 2.7$ which is partially covered by the ID and has higher particle flux, because of their better time resolution and rate capability than MDTs. The CSC chambers have a slightly lower resolution than the MDTs, with 40 μm in the tracks bending plane, and 5 mm in the transverse plane. For trigger purpose, the RPCs placed in the barrel, covering the range $|\eta| < 1.05$, are arranged into three concentric layers around the beam axis and placed before or after the MDT layers as Figure 3.15 shows. The RPC unit is composed of a gap of 2 mm formed by two parallel electrodes. The gap is filled with a mixture gas of 94.7% of $C_2H_2F_4$, 5% of $Iso - C_2H_2F_4$ and 0.3% of SF_6 . The electric field between electrodes is about 4.9 kV/mm. The end-caps are equipped with the TGCs up to $|\eta| = 2.4$. The TGCs are multi-wire proportional chambers filled with a mixture 55% of CO_2 and 45% of $n - C_5H_{12}$. The TGCs are arranged in seven layers on each side. The overall momentum resolution $\frac{\sigma_{p_T}}{p_T}$ provided by the muon system is 4% (1%) at 5 GeV (1 TeV) [43].

3.2.5 Trigger

The collision rate at the centre of ATLAS is 40 MHz due to the bunch crossing in every 25 ns provided by LHC. At such a rate, it is technically not possible to fully reconstruct, process and store all data recorded by the ATLAS detector. The ATLAS trigger system was designed to handle this problem. Its purpose is to reduce the input of 40 MHz bunch crossing rate to an output rate of about 200 Hz to record them for offline analysis. This is done thanks to two separated trigger systems. The first level,

so-called Level-1 (L1) [44], is a hardware-based trigger that uses reduced granularity signals from the calorimeter and the muon spectrometer to identify Regions-Of-Interest (ROI) with high-energy objects or high multiplicity at a latency of $2.5 \mu\text{s}$. The L1 trigger system is responsible for reducing the rate to at least 75 kHz. Events passing the L1 trigger are then sent to the software-based High-Level Trigger (HLT) [45]. At the HLT, a fast analysis of the ROIs followed by an offline-like reconstruction of the events takes place with a processing time of around 0.2 s. The HLT creates an output with a frequency of around 1 kHz [46]. For final states with large production rates, it can be necessary to keep only a fraction to match the offline storage budget.

3.3 Upgrade plans towards Run-3 and the HL-LHC

The long-term plan for the instantaneous luminosity delivered by the LHC foresees a step-wise increase due to continuous upgrades of the accelerator. To cope with the increased event rate and higher pile-up, the LHC experiments perform upgrades to various detector components, too. The main changes in the detector setup are performed during the long shutdown (LS) phases that are accrued in 2019-20 (LS2 with the so-called Phase-1 upgrades) and planned in 2025-27 (LS3 and Phase-2 upgrades). During LS2 and LS3, different detector and accelerator upgrades will be installed in ATLAS and LHC. After LS2, LHC could provide around 300 fb^{-1} of data at the end of Run-3, and possibly run with an increased centre-of-mass energy ($\sqrt{s} = 14 \text{ TeV}$). The Run-3 data taking was planned to start at the beginning of 2021, while given the coronavirus (Covid-19) epidemic the plan was delayed to 2022. A subsequent upgrade of the LHC (LS3), the so-called High-Luminosity (HL-LHC), will increase instantaneous luminosity and deliver 3000 fb^{-1} in 10 years. An overview of the LHC operation plan is shown in Figure 3.17.



Figure 3.17: Operation plan for the LHC scientific program.

The mean pile-up is expected to be up to $<\mu> = 200$ during the HL-LHC operation. To cope with the more difficult conditions, the planned upgrades of ATLAS include replacement of the ID during the LS3 by the so-called ITk, a silicon detector made of pixels and strips. ITk will maintain at least similar tracking and vertex performances as in the HL-LHC conditions. In addition to ITk, the planned upgrades also include the extension of the tracking system with the so-called High-Granularity Timing Detector (HGTD) [47], based on low gain avalanche detector technology. HGTD will improve the identification

840 and subtraction of pile-up in the forward direction. High-precision timing greatly improves the track-
 841 to-vertex association, leading to performance similar to that in the central region for the reconstruction
 842 of both jets and leptons, as well as for the tagging of heavy-flavour jets. These improvements in object
 843 reconstruction performance translate into important sensitivity gains and enhance the reach of the HL-
 844 LHC physics program, that double Higgs boson search will profit from it. The full ATLAS upgrade
 845 program can be found in the Ref. [48, 49, 50, 51, 52, 53, 54].

846 3.4 Physics objects reconstruction

847 ATLAS provides energy deposits collected by elements of sub-detectors. To interpret this raw output
 848 in terms of the event particles, an advanced particle reconstruction chains was developed. These are
 849 described in the following sections.

850 3.4.1 Track and vertex reconstruction

851 Tracks are reconstructed in the ID (Section 3.2.2) using a sequence of algorithms [55, 56]. The inside-out
 852 algorithm starts from three points seeds in the SCT. A combinatorial Kalman filter (Iterative algorithm
 853 providing the best estimate of the state based on the projection of earlier measurements and current
 854 measurement [57]) is used to build track candidates from the chosen seeds by incorporating additional
 855 space-points from the remaining layers of the PD and SCT which are compatible with the preliminary
 856 trajectory. Ambiguities in the track candidates are resolved, and tracks are extended into the TRT.
 857 The inside-out is the baseline algorithm designed for the efficient reconstruction of primary charged
 858 particles. In the second stage, a back-tracking algorithm is used in the track search starting from
 859 segments reconstructed in the TRT and extending them inwards by adding silicon hits. Back-tracking
 860 is designed to reconstruct secondary particles.

861 Finally, tracks with a TRT segment but no extension into the silicon detectors are referred to as TRT-
 862 standalone tracks. To minimize the rate of the fake track in a high pile-up environment, two-track
 863 quality requirements are defined. The robust requirement requires at least 9 hits in the silicon detectors
 864 and zero holes (non-existing but expected hit) in the PD. In contrast to the robust, the default requires
 865 at least 7 silicon hits and allow at most two-pixel holes.

866 The track reconstruction efficiency is defined as the fraction of primary particles with $p_T > 400$ MeV
 867 and $|\eta| < 2.5$ matched to a reconstructed track. Figure 3.18 shows the track reconstruction efficiency as
 868 a function of p_T and η .

869 Primary vertices are reconstructed using an iterative vertex finding algorithm [58]. Vertex seeds are
 870 obtained from the z-position at the beam axis of the reconstructed tracks. An iterative χ^2 fit constrained
 871 with the beam spot position is made using the seed and nearby tracks. Tracks are weighted depending
 872 on the χ^2 to measure the compatibility with the fitted vertex [59]. Tracks displaced by more than 7σ
 873 from the vertex are used to seed a new vertex and the procedure is repeated until no additional vertex
 874 is found. During the reconstruction, vertices are required to contain at least two tracks. The efficiency
 875 to reconstruct a vertex from a minimum bias interaction is shown in Figure 3.19. Vertices are matched
 876 to interactions by calculating the sum of the weights of the tracks associated to the interaction vertex.

877 3.4.2 Electron and photon reconstruction

878 Photons and electrons are reconstructed similarly using the ECal (Section 3.2.3.1). When these particles
 879 pass through the calorimeter dense medium, a showering process starts through cascading bremsstrahlung

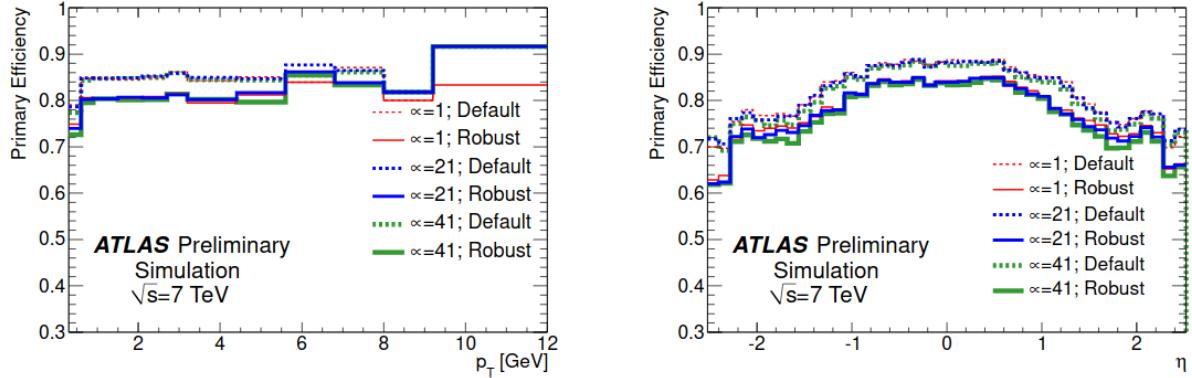


Figure 3.18: The primary track reconstruction efficiency in minimum bias simulated samples containing exactly one and on average 21 or 41 interactions. The distributions are shown for tracks passing the default(dashed) and robust (solid) requirements.

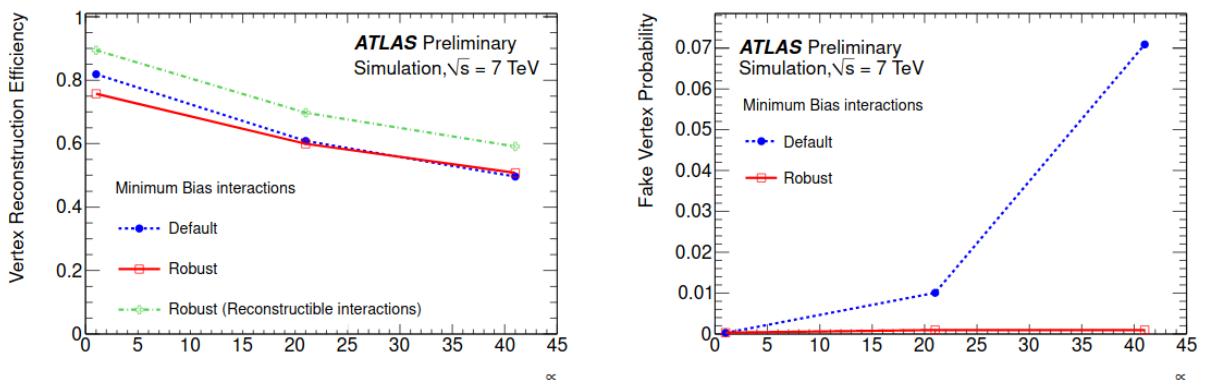


Figure 3.19: The vertex reconstruction efficiency (*left*) and fake probability (*right*) as a function of the average number of interactions.

and electron pair production. The electrical signal induced by electrons from ionized active material (LAr) is proportional to the deposited energy in the active volume of the calorimeter and used to compute the cell energy. Note that the cell energy is affected by fluctuations from both the electronic noise which is about 10 MeV per cell in the strip and 30 MeV in the middle and back layers and the pile-up noise. To reconstruct the EM clusters, the EM calorimeter is divided into a grid of $N_\eta \times N_\phi$ towers of the same granularity as the middle layer (0.025×0.025). Inside each of these elements, the energy of all cells in all longitudinal layers is summed into tower energy. A window of fixed size 3×5 is moved across each element of the tower. If the window transverse energy E_T (defined as the sum of the transverse energy of the towers contained in the window) is a local maximum and is above a threshold (2.5 GeV), a cluster seed is formed. Then, if the reconstructed cluster is associated with at least one reconstructed track, the candidate is classified as an electron. This reconstruction algorithm is called *fixed-size* [60]. A second reconstruction algorithm has been developed to benefit from dynamic, variable-size, clusters called *super-clusters* [61]. This allows the recovery of low-energy photons radiated due to bremsstrahlung interactions in the ID or electrons from photon conversions. In this scenario, the electron is defined as an object consisting of a supercluster and matched track. A converted photon is a cluster matched to a conversion vertex, and an unconverted photon is a cluster matched to neither an electron, track nor conversion vertex. In contrast to the sliding window (fixed-size), the super-cluster selects clusters based on topologically connected calorimeter cells [62]. Figure 3.20 shows an illustration of super-clusters. The topological cluster reconstruction starts from a cell with $|E_{cell}| > 4\sigma$, where σ is the expected cell

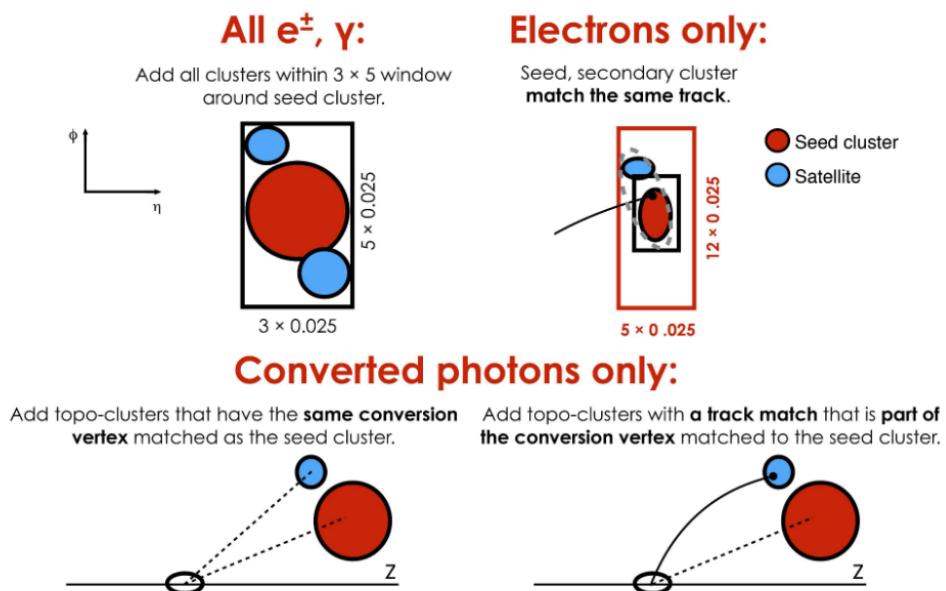


Figure 3.20: Diagram of the super clustering algorithm for electrons and photons. Seed clusters are shown in red, satellite clusters in blue.

noise based on the known electronic noise and an estimation of the pile-up noise. Then successively, all neighbouring cells with $|E_{cell}| > 2\sigma$ are added. The absolute cell energy is used to avoid biasing the cluster energy upwards due to negative energy induced by the calorimeter noise. The list of reconstructed topo-clusters is sorted according to descending total energy. The topo-clusters are tested one by one for use as a super-cluster. For an electron, the topo-cluster is required to have a minimum E_T of 1 GeV and matched to a track with at least four hits in the silicon detector [63], while for a photon the threshold is

set to 1.5 GeV with no matched track or conversion vertex. For both cases, to recover radiative losses, satellite clusters around that reconstructed cluster are matched to the super-cluster. Tracks are matched to electron super clusters and conversion vertices to converted photon super clusters. The matching is performed in the same way as the matching to EM topo-clusters was performed, but using the super clusters instead.

The reconstruction efficiency using super clusters for an electron is shown in Figure 3.21.

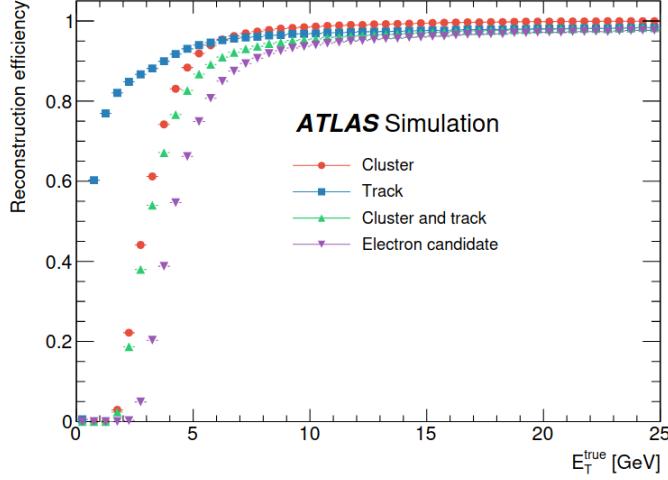


Figure 3.21: The cluster, track, cluster and track, and electron reconstruction efficiencies as a function of the generated electron E_T .

Figure 3.22 shows the reconstruction efficiency for converted photons as a function of the true E_T of the simulated photon for the previous version of the reconstruction software (fixed-size) and the current version (dynamic-size). An important reason to use super clusters is the improved energy resolution by collecting more of the deposited energy.

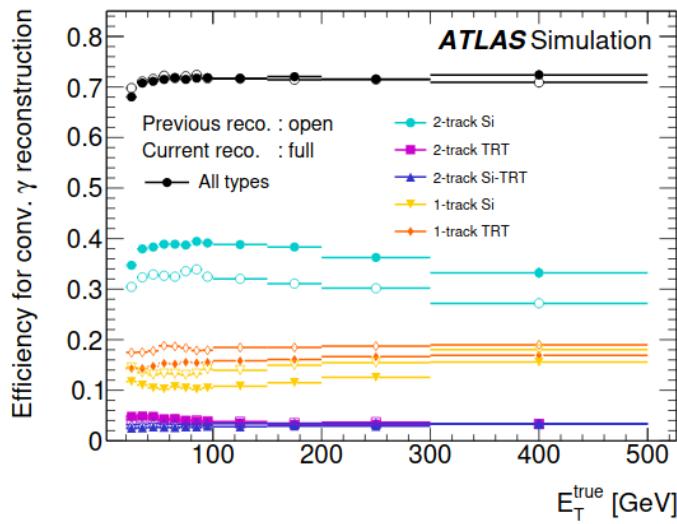


Figure 3.22: The converted photon reconstruction efficiency and contributions of the different conversion types as a function of E_T^{true} .

915 Additionally, an ambiguity resolution is performed to remove a part of the overlap in case one object
 916 is reconstructed at the same time as electron and photon since electron and photon super clusters are
 917 built independently. However, in order to maintain a high reconstruction efficiency, a residual overlap is
 918 allowed to match physics analysis needs. Figure 3.23 shows the procedure used for ambiguity resolution.

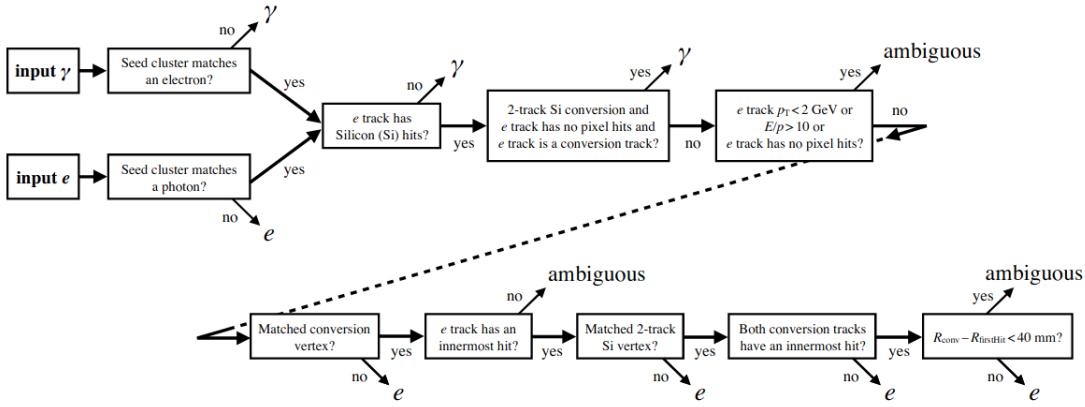


Figure 3.23: Flowchart showing the logic of the ambiguity resolution for particles initially reconstructed both as electrons and photons.

919
 920 Most LHC physics require the identification of prompt electrons and/or photons. Prompt particles are
 921 those not produced by a hadron or tau decay. Non-fake particles are those whose type was properly
 922 reconstructed (i.e., a non-fake reconstructed electron is a true electron, not another misidentified parti-
 923 cle). Their identification and separation are obtained with several selection criteria and algorithms.
 924 The identification for the reconstructed electrons and photons is finally based on the reconstructed
 925 super-cluster objects. A set of discriminating variables are calculated for the identification using cells
 926 in the fixed-size window. A list is given in Table 3.2 with an indication if they are used for electron
 927 or photon identification. The prompt and non-fake leptons/photons are usually isolated, without much
 928 activity around them. It is therefore important to define proper "isolation criteria" to reduce the con-
 929 tamination from non-prompt and fake objects. Electron isolation and identification are described in
 930 Sections 3.4.2.2 and 3.4.2.3 respectively, while Chapter 4 is dedicated to photon properties.

931 3.4.2.1 Electron and Photon calibration

932 After the electron and photon super clusters are built, initial energy calibration and position correction
 933 are applied to them [64]. A complex chain with several successive calibration steps is used to calibrate
 934 the electron and photon energies based on simulation and well-known reference processes. A schematic
 935 overview of the whole calibration chain is shown in Figure 3.24.
 936 The calibration steps are:

- 937 1. A multivariate regression algorithm is trained on the properties of the shower development in the
 938 EM calorimeter to optimize the energy resolution and to minimize the impact of material in front
 939 of the calorimeter. The algorithm used in this step is the Boosted Decision Tree (BDT) tuned
 940 in intervals of $|\eta|$ and E_T on samples of simulated single particles without pile-up, separately for
 941 electrons, converted and unconverted photons.

Table 3.2: Discriminating variables used for electron and photon identification. The usage column indicates if the variables are used for the identification of electrons, photons, or both. [61].

Category	Description	Name	Usage
Hadronic leakage	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (used over the ranges $ \eta < 0.8$ and $ \eta > 1.37$)	R_{had_1}	e/γ
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $0.8 < \eta < 1.37$)	R_{had}	e/γ
EM third layer	Ratio of the energy in the third layer to the total energy in the EM calorimeter	f_3	e
EM second layer	Ratio of the sum of the energies of the cells contained in a 3×7 ($\eta \times \phi$) rectangle (measured in cell units) to the sum of the cell energies in a 7×7 rectangle, both centred around the most energetic cell	R_η	e/γ
	Lateral shower width, $\sqrt{(\sum E_i \eta_i^2)/(\sum E_i) - ((\sum E_i \eta_i)/(\sum E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells	w_{η_2}	e/γ
	Ratio of the sum of the energies of the cells contained in a 3×3 ($\eta \times \phi$) rectangle (measured in cell units) to the sum of the cell energies in a 3×7 rectangle, both centred around the most energetic cell	R_ϕ	e/γ
EM first layer	Total lateral shower width, $\sqrt{(\sum E_i (i - i_{\max})^2)/(\sum E_i)}$, where i runs over all cells in a window of $\Delta\eta \approx 0.0625$ and i_{\max} is the index of the highest-energy cell	$w_{s\,\text{tot}}$	e/γ
	Lateral shower width, $\sqrt{(\sum E_i (i - i_{\max})^2)/(\sum E_i)}$, where i runs over all cells in a window of 3 cells around the highest-energy cell	$w_{s\,3}$	γ
	Energy fraction outside core of three central cells, within seven cells	f_{side}	γ
	Difference between the energy of the cell associated with the second maximum, and the energy reconstructed in the cell with the smallest value found between the first and second maxima	ΔE_s	γ
	Ratio of the energy difference between the maximum energy deposit and the energy deposit in a secondary maximum in the cluster to the sum of these energies	E_{ratio}	e/γ
	Ratio of the energy measured in the first layer of the electromagnetic calorimeter to the total energy of the EM cluster	f_1	e/γ
Track conditions	Number of hits in the innermost pixel layer	$n_{\text{innermost}}$	e
	Number of hits in the pixel detector	n_{Pixel}	e
	Total number of hits in the pixel and SCT detectors	n_{Si}	e
	Transverse impact parameter relative to the beam-line	d_0	e
	Significance of transverse impact parameter defined as the ratio of d_0 to its uncertainty	$ d_0/\sigma(d_0) $	e
	Momentum lost by the track between the perigee and the last measurement point divided by the momentum at perigee	$\Delta p/p$	e
	Likelihood probability based on transition radiation in the TRT	$e\text{ProbabilityHT}$	e
Track–cluster matching	$\Delta\eta$ between the cluster position in the first layer of the EM calorimeter and the extrapolated track	$\Delta\eta_1$	e
	$\Delta\phi$ between the cluster position in the second layer of the EM calorimeter and the momentum-rescaled track, extrapolated from the perigee, times the charge q	$\Delta\phi_{res}$	e
	Ratio of the cluster energy to the measured track momentum	E/p	e

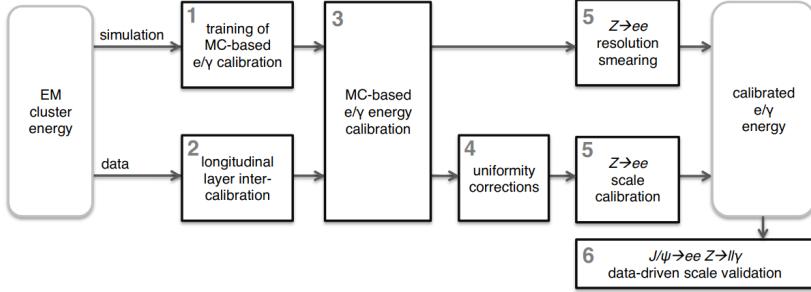


Figure 3.24: Schematic overview of the calibration chain for the electron and photon energies [65].

- 942 2. An inter-layer correction is applied to the relative energy response in data to match the relative
 943 response in simulation in different calorimeter layers. The inter-layer calibration is independent of
 944 the material upstream of the calorimeter.
- 945 3. The BDT correction is applied to both data and simulation. This step performs the main correction
 946 of the absolute energy scale and improves significantly the energy resolution.
- 947 4. The missing detector non-uniformities are therefore corrected in data. Correcting the effects of
 948 energy loss between the barrel calorimeter modules and the high-voltage inhomogeneities.
- 949 5. Scale factors are applied to the energy in data to correct for the residual miscalibration between
 950 data and simulation using $Z \rightarrow e^+e^-$ events.
- 951 6. The validity of the calibration is cross-checked from data using different processes. For the low-
 952 energy range, $J/\Psi \rightarrow e^+e^-$ events are used. The calibration for photons is cross-checked using
 953 radiative $Z \rightarrow l^+l^-\gamma$ ($l = e, \mu$) decays.

954 3.4.2.2 Electron Isolation

The activity near leptons and photons is quantified from the tracks of nearby charged particles or energy deposits in the calorimeters, leading to two classes of isolation variables, calorimeter-based isolation E_T^{coneXX} and track-based p_T^{coneXX} [66]. The corrected E_T^{coneXX} is computed as the sum of the transverse energies of topo-clusters inside a cone of $\Delta R = \frac{XX}{100}$ around the reconstructed electrons. This is illustrated in Figure 3.25, after subtraction of the energy deposited by the electrons, pile-up and underlying event [67]. p_T^{coneXX} is the sum of the transverse momentum of selected tracks of $p_T > 1$ GeV and $|\eta| < 2.5$, within a cone centred around the electron track. Tracks matched to the electron are excluded. In a high-momentum heavy particles decay, the electron is produced very close to other decay products for this reason, an isolation $p_T^{varconeXX}$ with a variable cone size ΔR^{XX} is defined as:

$$\Delta R^{XX} = \min\left(\frac{10}{p_T[\text{GeV}]}, \frac{XX}{100}\right). \quad (3.9)$$

955 Based on these variables, three different selection criteria, called working point (WP), are implemented.
 956 The WPs were defined in two different ways, either targeting a fixed value of efficiency or with fixed
 957 cuts on the isolation variables. Table 3.3 lists the different electron-isolation WPs used in ATLAS. An
 958 additional WP named **Gradient** is designed to give an efficiency of 90% at $p_T = 25$ GeV and 99% at 60
 959 GeV, uniform in η .

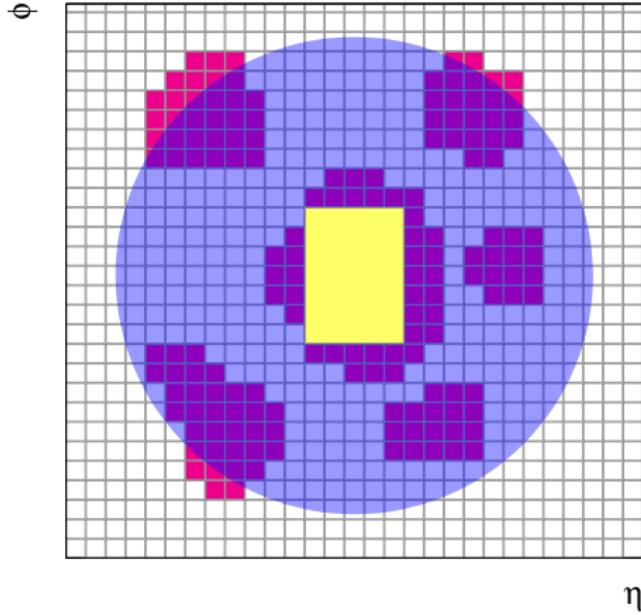


Figure 3.25: Schema of the calorimeter isolation method: the grid represents the second-layer calorimeter cells in the η and ϕ directions. The candidate electron is located in the centre of the purple circle representing the isolation cone. All topological clusters, represented in red, for which the barycentres fall within the isolation cone are included in the computation of the isolation variable. The 5×7 cells represented by the yellow rectangle correspond to the subtracted cells in the core subtraction method.

WP	Calorimeter-based isolation	Track-based isolation
HighPtCaloOnly	$E_T^{cone20} < \max(0.015 \times p_T, 3.5\text{GeV})$	-
Loose	$E_T^{cone20}/p_T < 0.2$	$p_T^{varcone20}/p_T < 0.15$
Tight	$E_T^{cone20}/p_T < 0.06$	$p_T^{varcone20}/p_T < 0.06$

Table 3.3: Definition of the electron isolation WPs.

Figure 3.26 shows the electron isolation efficiency in data recorded in 2017 as a function of electron E_T [68]. The method used to compute the electron isolation efficiency and the associated uncertainties are described in Ref. [66]. The jump observed in Gradient efficiency at 15 GeV is due to the isolation efficiency which is process dependent: the Gradient cut maps is optimized with $J/\Psi \rightarrow e^+e^-$ events below 15 GeV, while the efficiency measurement is performed with $Z \rightarrow e^+e^-$.

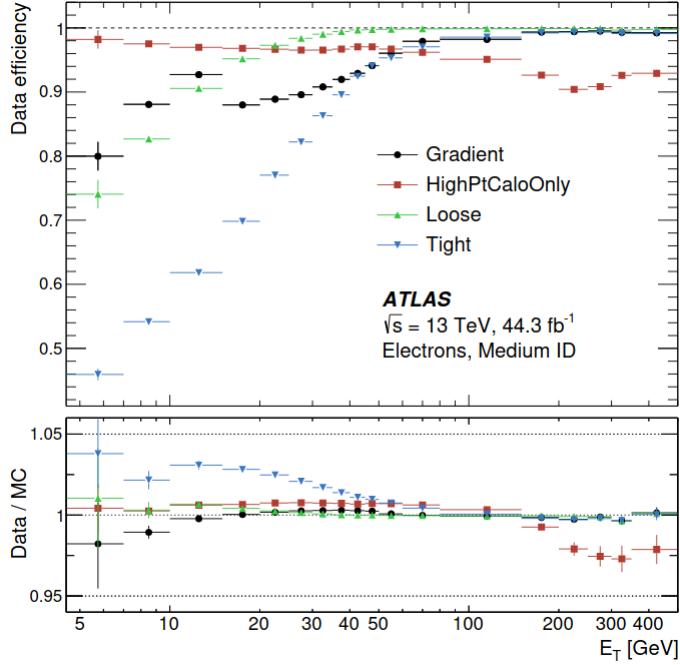


Figure 3.26: Efficiency of the different isolation WPs for electrons from inclusive $Z \rightarrow e^+e^-$ events as a function of the electron E_T . The lower panel shows the ratio of the efficiencies measured in data and in simulations [68].

965

3.4.2.3 Electron Identification

The identification of prompt electrons relies on a likelihood (LH) discriminant constructed from quantities measured in the different sub-detectors and listed in Table 3.2. The electron LH is based on the products of probability density functions (pdf) for signal L_S , and for background L_B . The pdfs are created by smoothing histograms of the n discriminating variables with an adaptive kernel density estimator (KDE) [69] as implemented in the TMVA framework [70] in 9 bins of $|\eta|$ and 7 bins of E_T :

$$L_{S(B)}(\mathbf{x}) = \prod_{i=1}^n P_{S(B),i}(x_i), \quad (3.10)$$

where \mathbf{x} is the vector of the various variables. For each electron candidate, a discriminant d_L is formed using an inverse sigmoid function transformation:

$$d_L = -\frac{1}{\tau} \ln\left(\frac{L_S + L_B}{L_S} - 1\right), \quad (3.11)$$

where τ is fixed to 15 [70]. Figure 3.27 shows an example of the distribution of the transformed discriminant for prompt electrons and non-prompt ones. This distribution illustrates the effective separation

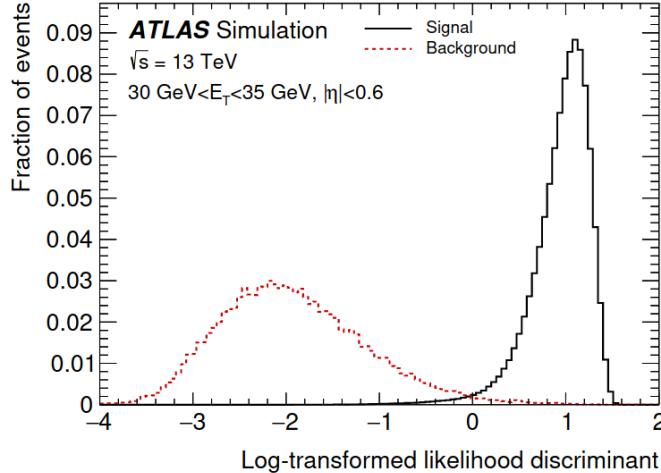


Figure 3.27: The transformed LH-based identification discriminant d_L for reconstructed electron candidates with good quality tracks with $30 < E_T < 35 \text{ GeV}$ and $|\eta| < 0.6$ [71].

969 between signal and background encapsulated in this single quantity.

970 For physics purposes, three WPs are defined, and are referred to as **Loose**, **Medium**, and **Tight**. Their
971 efficiency for identifying a prompt electron with $E_T = 40 \text{ GeV}$ are 93%, 88% and 80% respectively.

972 Figure 3.28 shows the resulting efficiencies in data [61].

973 3.4.3 Muon reconstruction and identification

974 3.4.3.1 Muon reconstruction

975 Muon reconstruction is performed in two sub-detectors independently. Firstly, the muon tracks are
976 reconstructed in the ID like any track as described in Section 3.4.1. Then, the ID reconstruction is
977 combined with the muon reconstruction in MS sub-detector to perform the muon object used in physics
978 analysis. In the MS, the muons are triggered in RPC/TGC if at least one hit exists, defining the region
979 of activity (ROA). All the muon chambers intersecting with the ROA are then selected as muon track
980 candidates. The MDT segments are reconstructed by performing a straight-line fit (the bending of
981 muons larger than few GeV is sufficiently small) to the hits found in each layer [72]. The fitted segments
982 are required to point loosely towards the IP, to reject background events and random hit combinations.
983 Muon track candidates are then built by extrapolating each of these segments to the other. At least two
984 matching segments using their relative positions and angles are required to build a track, except in the
985 barrel-endcap transition region where a single high-quality segment can be used.

986 The combined reconstruction is performed according to various algorithms based on the provided information
987 by sub-detectors [73, 74]:

- 988 • Combined (CB) muon: the combined muon is formed with a global refit that used the hits from
989 both the ID and MS sub-detectors. The reconstruction is done following two complementary
990 approaches, the outside-in in which the reconstructed track in the MS are extrapolated inward
991 and match to an ID track, and the inside-out reconstruction, in which ID tracks are extrapolated
992 outward and matched to MS tracks.
- 993 • Segment-tagged (ST) muon: used when muons cross only one layer of MS chamber, either because
994 of their low p_T or because of MS acceptance.

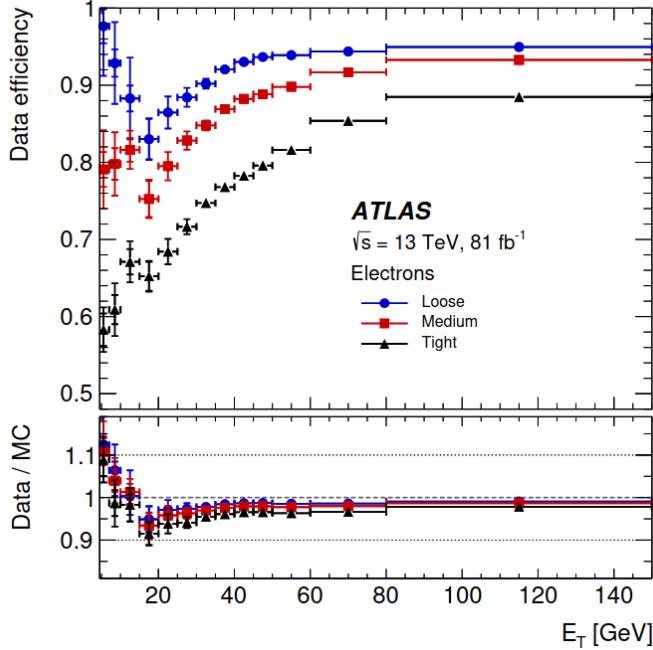


Figure 3.28: The electron identification efficiency in $Z \rightarrow e^+e^-$ events in data as a function of E_T for the Loose, Medium, and Tight WPs.

- Calorimeter-tagged (CT) muon: reconstructed track with an energy deposit in the calorimeter compatible with a minimum-ionizing particle is identified as CT muon.
- Extrapolated (ME) muon: reconstructed based only on the MS track and a loose requirement on compatibility with originating from the IP.

Overlaps between different muon types are resolved with preferences to CB, ST and CT muons respectively before producing the collection of muons used in physics analyses.

3.4.3.2 Muon identification

In order to suppress non-prompt muons, mainly coming from pion and kaon decay, a muon identification is performed. Muon identification uses several variables, for CB tracks:

- q/p significance defined as:

$$q/p \text{ significance} = \frac{|q/p_{\text{ID}} - q/p_{\text{MS}}|}{\sqrt{\sigma^2(q/p_{\text{ID}}) + \sigma^2(q/p_{\text{MS}})}}, \quad (3.12)$$

where q/p_{ID} and q/p_{MS} are the measurements in the ID and MS of the ratio of the charge q to the momentum p of the muon, expressed at the IP and σ is the corresponding uncertainties.

- ρ' defined as the absolute value of the difference between the transverse momentum measurements in the ID and MS divided by the p_T of the combined track.
- Normalized χ^2 of the combined track fit.

Four muon identification WPs are provided to address specific needs of different physics analysis:

- 1010 ● **Loose**: designed to maximize the reconstruction efficiency while providing good-quality muon
 1011 tracks. They are specifically optimized for reconstructing Higgs boson candidates in the four-
 1012 lepton final state [75].
- 1013 ● **Medium**: provides the default selection for muons in ATLAS. This selection minimizes the system-
 1014 atic uncertainties associated with muon reconstruction and calibration. Only CB and ME tracks
 1015 are used. About 0.5% of Medium muon originates from the inside-out reconstruction strategy, in
 1016 the central region.
- 1017 ● **Tight**: maximizes the purity of muons at the cost of efficiency. Only CB muons with hits in at
 1018 least two stations of the MS and satisfying the **Medium** selection criteria are considered.
- 1019 ● **High- p_T** : aims to maximize the momentum resolution for tracks with transverse momentum above
 1020 100 GeV. The selection is optimized for searches for high-mass Z' and W' resonances [76, 77]. CB
 1021 muons passing the **Medium** selection and having at least three hits in three MS stations are selected.

1022 Figure 3.29 shows the muon identification efficiency for **Loose**, **Medium** and **Tight** muons as measured
 in $J/\Psi \rightarrow \mu\mu$.

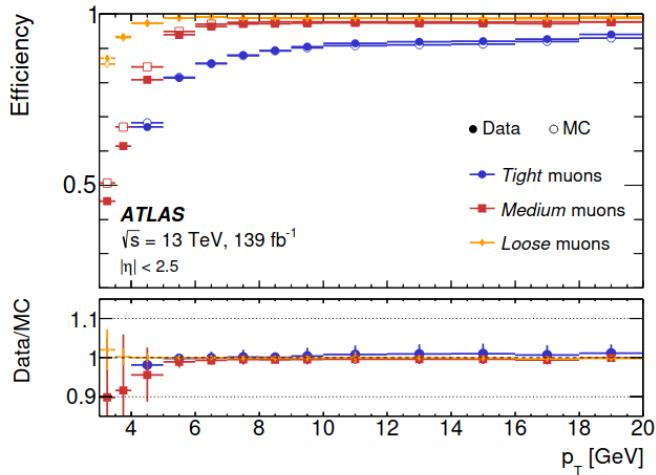


Figure 3.29: Muon identification efficiency for the **Loose**, **Medium** and **Tight** WPs as function of p_T . The panel at the bottom shows the ratio of the measured to predicted efficiencies, with statistical and systematic uncertainties.

1023

1024 3.4.4 Jet reconstruction

1025 Jets are made of many partons coming from the initial quark or gluon hadronization, appearing in the
 1026 detector as a collimated shower. Many jets reconstruction algorithms exist in ATLAS, but only two
 1027 approaches are of interest for the work presented in this thesis. The first approach is based exclusively
 1028 on electromagnetic and hadronic topological clusters, so-called EMTopo jets. The second approach uses
 1029 both tracking and calorimetric information through a Particle Flow algorithm to build PFlow jets [78].
 1030 Jets are reconstructed using the anti- k_t algorithm [79], with a distance parameter of $R = 0.4$. Jets
 1031 reconstruction, calibration and identification of b -jets will be discussed in more details in Chapter 5.

₁₀₃₂ **3.4.5 Missing transverse energy**

Neutrinos and other BSM particles interact extremely weakly with matter, making them hard to detect. Their energy is reconstructed as missing transverse energy (MET). Thanks to momentum conservation, the transverse momentum of all particles generated in a collision should sum up to zero, since the original transverse momentum of the partons is negligible. MET is defined as the sum of the transverse energy momenta of all reconstructed objects i :

$$\vec{E}_T^{miss} = - \sum_i \vec{p}_T^i \quad (3.13)$$

₁₀₃₃ The MET resolution is typically 10-20 GeV [80].

1034 Chapter 4

1035 Selection of photons in ATLAS

1036 The photon object is a key component to study the $HH \rightarrow b\bar{b}\gamma\gamma$ properties. They are reconstructed
1037 similarly to electrons as introduced in Section 3.4.2, and their energy is calibrated in the same way. This
1038 chapter details photon isolation and identification. I implemented a new photon identification algorithm
1039 based on Machine Learning tools to improve photon identification (ID).

1040 4.1 Photon Isolation

1041 Photon isolation is almost similar to the one from electrons (Section 3.4.2.2). However for photons,
different requirements are defined for three WPs. Table 4.1 summarizes these WPs [68].

WP	Calorimeter-based isolation	Track-based isolation
Loose	$E_T^{cone20} < 0.065 \times E_T$	$p_T^{varcone20}/E_T < 0.05$
Tight	$E_T^{cone40} < 0.022 \times E_T + 2.45 \text{ GeV}$	$p_T^{varcone20}/E_T < 0.05$
TightCaloOnly	$E_T^{cone40} < 0.022 \times E_T + 2.45 \text{ GeV}$	-

Table 4.1: Definition of the photon isolation WPs.

1042 A discrepancy between the peak positions of the simulated and real distributions of the calorimeter-
1043 based variables is observed since Run-1 [81], pointing to a mismodelling of the lateral profile development
1044 of the electromagnetic showers and resulting different isolation efficiency in data and simulations, thus
1045 large scale factors. A data-driven shift is applied to handle the mismodelling in simulation. The shift
1046 values are computed by taking the difference of the fitted peak values between data and simulation.
1047 The fit is performed using the Crystal Ball pdf [82] on the E_T^{cone40} and E_T^{cone20} variables in bin of η ,
1048 E_T and conversion type (converted or unconverted) separately. The resulting shift is then added to the
1049 simulation distribution. Figure 4.1 shows the distribution of E_T^{cone40} after the data-driven shifts applied.
1050 Figure 4.2 shows the efficiency of the isolation WPs defined in Table 4.1, using $Z \rightarrow ll\gamma$ ($l = e, \mu$). The
1051 radiative Z decays signature is used to estimate the photon efficiencies as it provides a clean environment
1052 of prompt photons, especially in the low- E_T range.
1053

1054 4.2 Photon Identification

1055 In contrary to electron identification, photon identification still relies on rectangular selection criteria,
1056 *cut-based algorithm*, using a set of global variables which characterize the lateral and longitudinal elec-

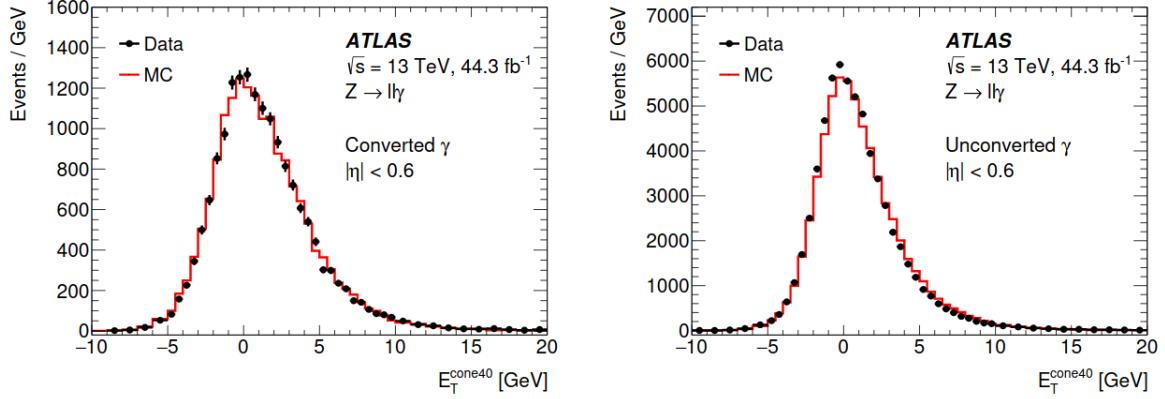


Figure 4.1: Distribution of E_T^{cone40} in data and simulation, in the central region of the detector ($|\eta| < 0.6$), separately for converted (left) and unconverted (right) photons after the data-driven shifts are applied.

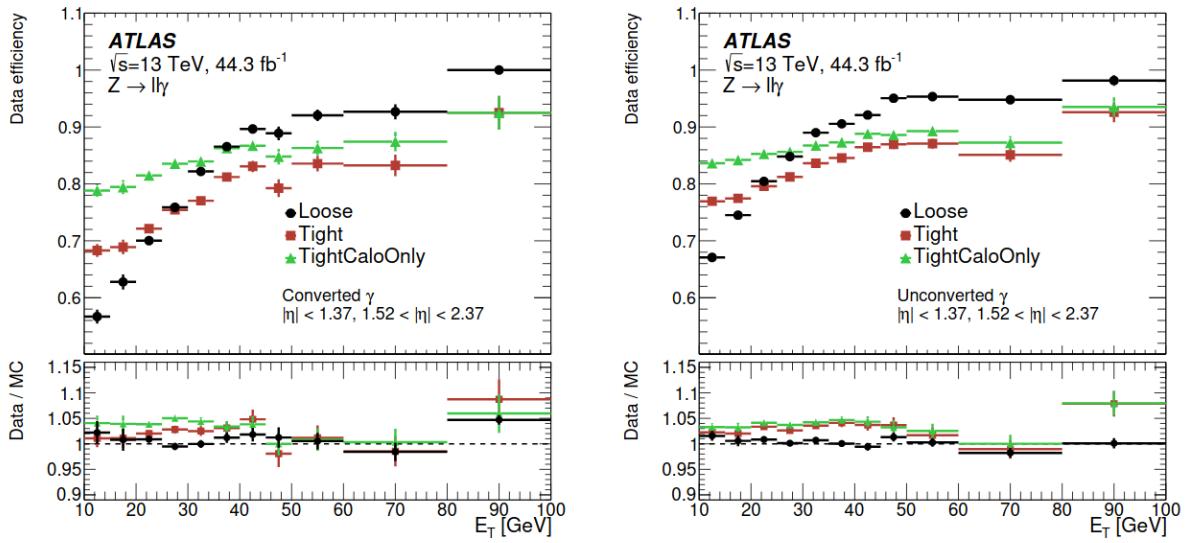


Figure 4.2: Efficiency of the isolation working points for converted(left) and unconverted (right) photons as a function of photon E_T . The lower panel shows the ratio of the efficiencies measured in data and in simulation (scale factors). The error bars account for the systematic and statistical errors.

tromagnetic shower development in the calorimeter and encode separation between prompt-photons and background photons: non-prompt originating from the decay of QCD jets, and objects misidentification as photons. Such variables, listed in Table 3.2 and depicted in Figure 4.3 with their respective definitions, are called "shower shapes". Given the fine granularity of the first EM layer, the associated shower shapes play an important role to reject the collimated photons from $\pi^0 \rightarrow \gamma\gamma$ decay.

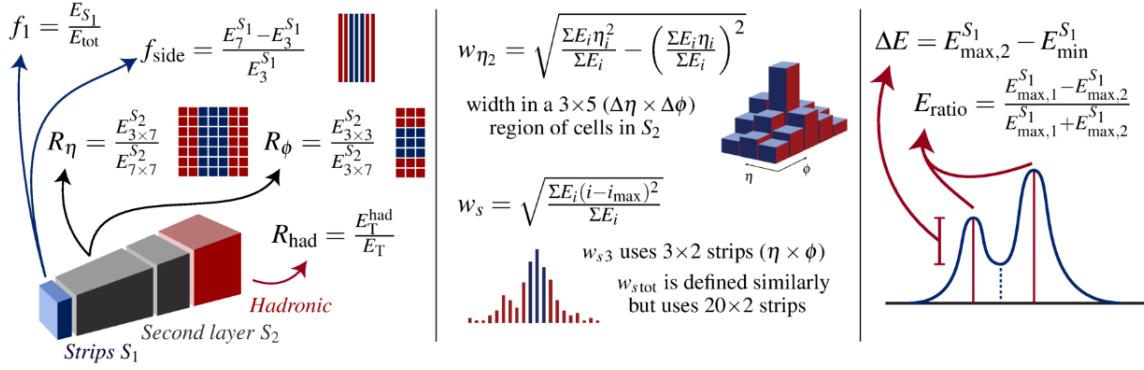


Figure 4.3: Schematic representation of the photon identification discriminating variable [83].

The cut-based algorithm provides three WPs: **Tight**, **Medium** and **Loose**, with less restrictive selections respectively. **Medium** and **Loose** WPs are mainly used for trigger purposes. Since the trigger system is transparent between converted and unconverted photons, they are similar for both converted and unconverted. Table 4.2 lists the variables set used in each WP. The optimization of the WPs is done using TMVA [70] in bins of $|\eta|$ since the shower shapes vary due to calorimeter granularity. Since the shower shapes are different between the converted and unconverted photons due to the opening angle of the e^+e^- conversion pair and the material interaction, the optimization of tight WP is performed separately for converted and unconverted photons. The tight WP comes with two versions: E_T -independent selection and the E_T -dependent which is tuned in separate bins of E_T . The tuning is done in two energy regimes with two different series of the simulated dataset. For photons with $10 < E_T < 25$ GeV, the $Z \rightarrow ll\gamma$ simulated sample is used to define the signal, while the corresponding background is derived from the $Z \rightarrow ll + jets$ simulated sample. For $E_T > 25$ GeV, the inclusive-photon production simulated sample is used as a signal, and the background is the dijet sample. The simulated samples are described in Section 3.1 and 3.2 of Ref. [68].

Working Point	Variables set
Loose	$R_{\text{had}}, R_{\text{had}_1}, R_\eta$ and w_{η_2}
Medium	Loose + E_{ratio}
Tight	Medium + $R_\phi, w_{s3}, w_{\text{stot}}, f_{\text{side}}, \Delta E_s$ and f_1

Table 4.2: Discriminating variables used for **Loose**, **Medium** and **Tight** photon identification.

Similarly to the calorimeter based variables for isolation, the shower shapes variables suffer from the mis-modelling. The mis-modelling is addressed by setting a data-driven correction similar to the one applied for isolation variables. Corrections are applied as a simple shift to simulated events to align with the real data distributions. Shifts are derived using a χ^2 minimization between the data and simulation in bins of $|\eta|$ and E_T as described in Ref [84]. Figure 4.4 shows examples of shower shape distributions

1082 before and after correction compared to the real distributions using radiative Z photons. Section 4.3 is
dedicated to the shower shape mis-modelling and its correction using an alternative method.

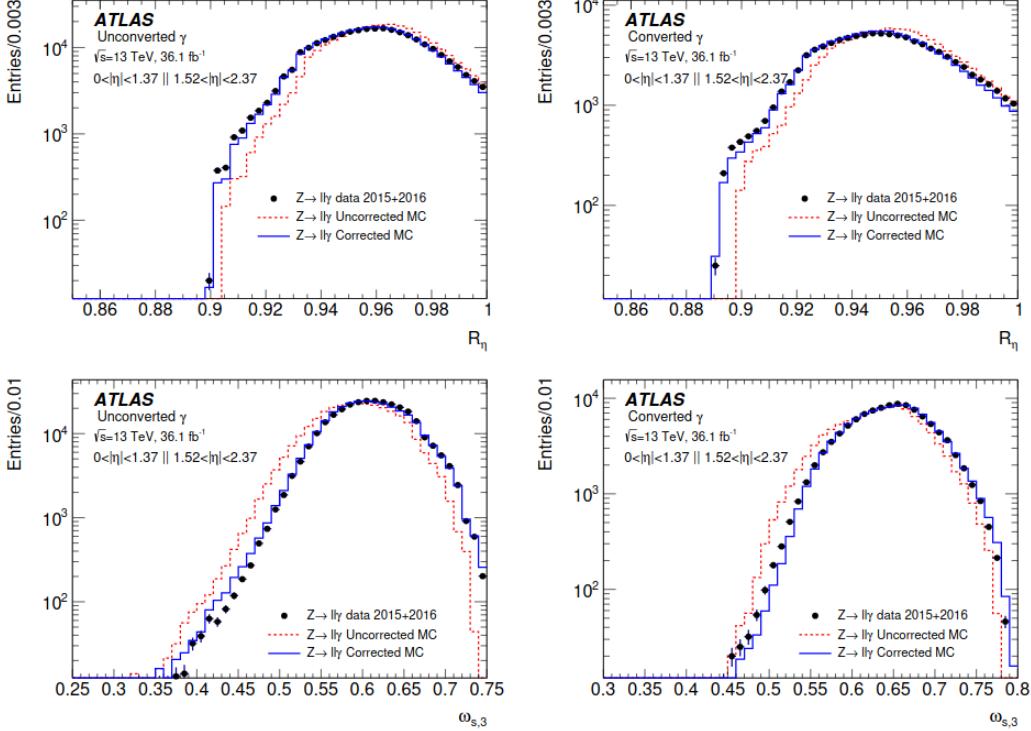


Figure 4.4: Distributions of the R_η and $\omega_{s,3}$ for converted and unconverted photon candidates with E_T in [10, 50] GeV and $|\eta| < 2.37$ selected from radiative Z events for data (black), uncorrected simulation (dashed red line) and corrected simulation (solid blue line).

1083
1084 Photon identification efficiency is measured using three different methods over distinct energy regimes:

- 1085 • The Matrix Method (MM): based on the inclusive-photon production sample collected using single-
1086 photon triggers and performed over a wide kinematic range from 25 GeV to 1.5 TeV in E_T .
- 1087 • The Radiative Z method: uses the clean signature of radiative Z at low-energy and allows a
1088 measurement of identification efficiency ϵ_{ID} from $E_T = 10$ GeV to 100 GeV. This method handles
1089 small background contamination from Z+jets where the jet is misidentified as a photon using a
1090 maximum-likelihood fit [85].
- 1091 • Electron Extrapolation (EE): uses a transformed electron shower shapes sample form $Z \rightarrow e^+e^-$
1092 and measures ϵ_{ID} in the region $25 < E_T < 150$ GeV.

1093 These methods are fully described in Section 5 of Ref. [84]. The tight identification efficiencies from
1094 each method are then combined to provide a ϵ_{ID} measurement for the E_T -dependent selection as shown
1095 in Figures 4.5 and 4.6. A comparison between the two versions of the tight WP is shown in Figure 4.7.
1096

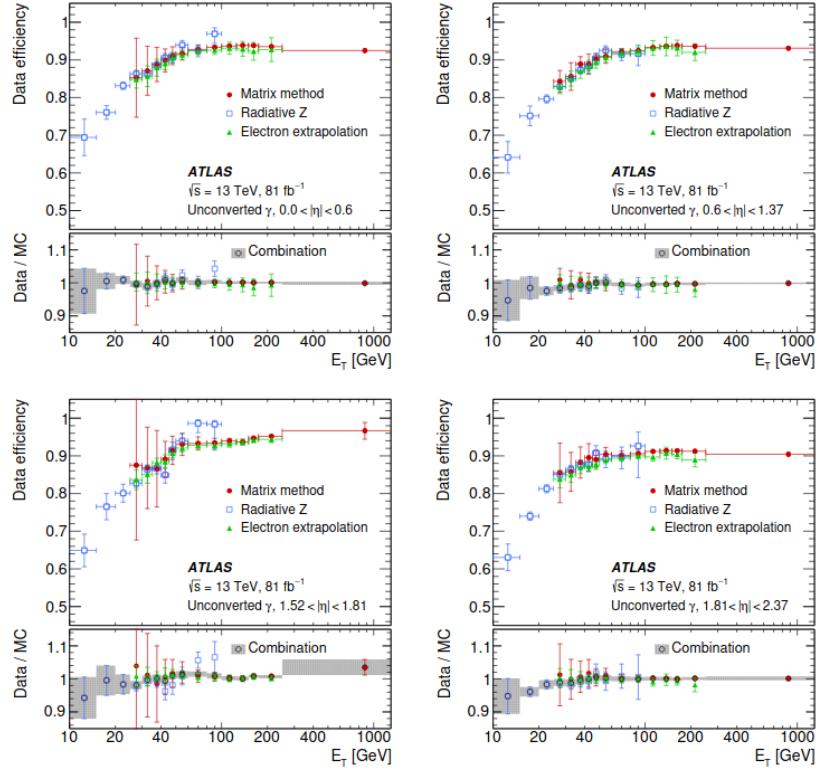


Figure 4.5: The photon identification efficiency with E_T -dependent selection, and the ratio of data to simulation efficiencies, for unconverted photons with a Loose isolation requirement applied as pre-selection, as a function of E_T in four different $|\eta|$ regions. The combined scale factor, obtained using a weighted average of scale factors from the individual measurements, is also presented; the band represents the total uncertainty.

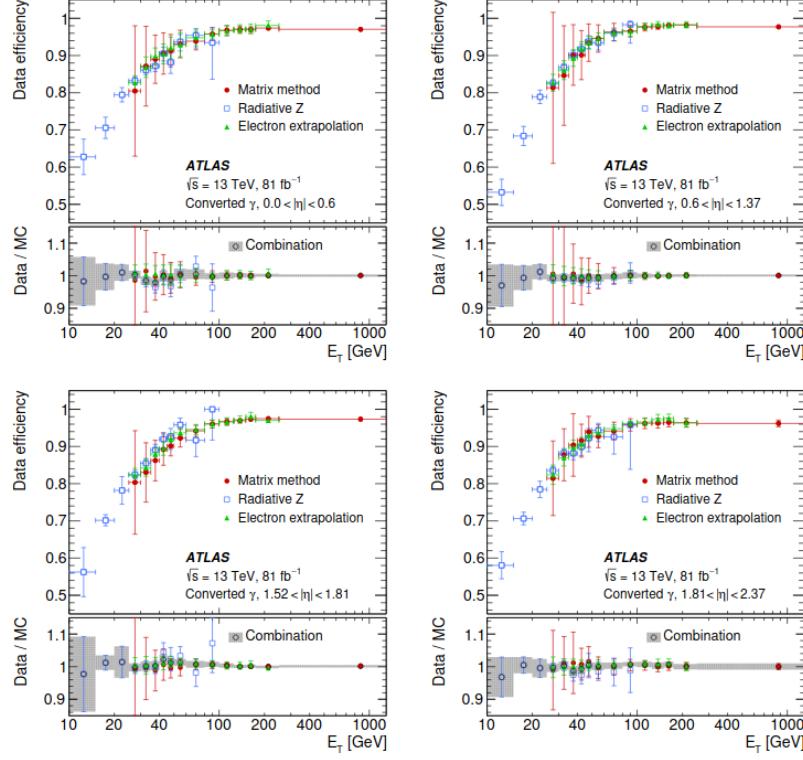


Figure 4.6: The photon identification efficiency with E_T -dependent selection, and the ratio of data to simulation efficiencies, for converted photons with a Loose isolation requirement applied as pre-selection, as a function of E_T in four different $|\eta|$ regions. The combined scale factor, obtained using a weighted average of scale factors from the individual measurements, is also presented; the band represents the total uncertainty.

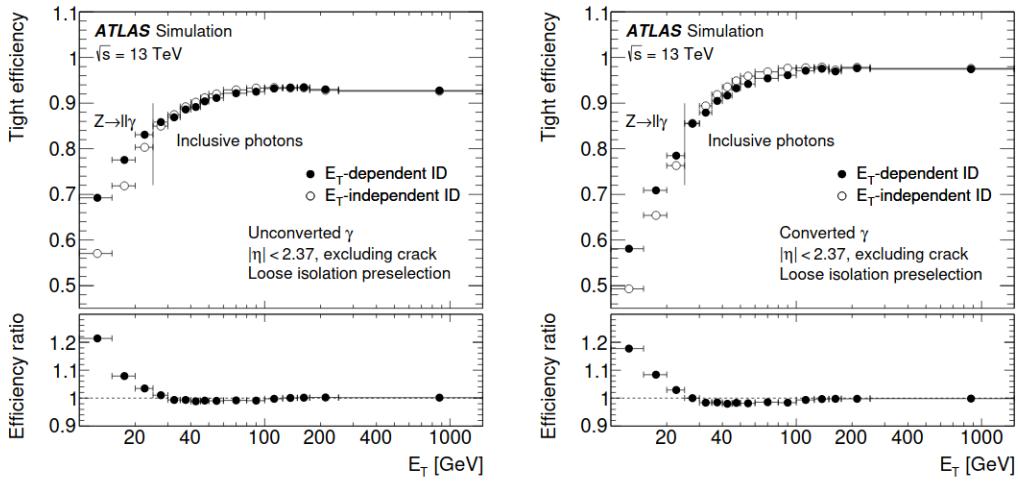


Figure 4.7: Efficiencies of the tight photon identification for unconverted (left) and converted (right) signal photons, plotted as a function of photon E_T . The loose isolation is applied as a pre-selection. For both plots, the bottom panel shows the ratios between the E_T -dependent and the E_T -independent identification efficiencies.

1097 4.3 Shower shape mis-modelling

1098 The source of the disagreement in shower shapes between the data and the Monte Carlo simulation
 1099 (MC) described above is still unclear. Many sources can contribute to this mis-modelling, mainly from
 1100 the detector geometry, material distribution and EM shower modelling. It is difficult to handle those
 1101 sources and fix the mis-modelling. To fix the issue for electron, a cell-based reweighting method has been
 1102 developed to apply a global correction to reduce systematic uncertainties related to the mis-modelling [86,
 1103 87]. The approach shows a promising electron result, which motivated its application to photons. This
 1104 work was done in the context of my ATLAS qualification task to allow me to sign ATLAS publications.

1105 4.3.1 Cell-based reweighting correction

1106 The aim is to redistribute the energy between the cluster cells in simulation to become consistent with
 1107 the real data. The cell-based reweighting is derived for the second layer of the EM calorimeter. A
 1108 cluster of 7×11 (7 cells in η and 11 cells in ϕ) centred around the cell with the highest energy (hottest)
 1109 is considered. The correction is derived as a matrix (7×11) in bins of η in two steps:

- 1105 1. Compute single cell corrections as:

$$M_i^{\text{Correction}} = \frac{E_i^{\text{data}}}{\sum E_i^{\text{data}}} - \frac{E_i^{\text{simulated}}}{\sum E_i^{\text{simulated}}}, \quad (4.1)$$

1110 where $E_{i=1..77}$ denotes the cell energy in the 7×11 cluster of layer 2.

- 1110 2. Compute the reweighted energy for each cluster cell as:

$$E_i^{\text{reweighted}} = E_i^{\text{non-reweighted}}(1 + M_i^{\text{Correction}}). \quad (4.2)$$

The main idea behind the reweighting is to factorize out the material effects and not the physics behaviours, especially the bremsstrahlung tails in the energy profile for electrons and positrons. For electron only, the bremsstrahlung tails are extracted from the e^+ and e^- energy profiles as:

$$E_i = \begin{cases} E_i^{\text{electron}} & \text{if } E_i^{\text{electron}} < E_i^{\text{positron}} \\ E_i^{\text{positron}} & \text{otherwise} \end{cases} \quad (4.3)$$

1111 so it is derived independently of bremsstrahlung and similarly for both electrons and positrons.

1112 4.3.2 Electron results

1113 Figure 4.8 shows a reduction in data-simulation discrepancy after the reweighting and leads to a similar
 1114 distribution for shower shapes computed from layer 2 (R_η and R_ϕ) between data and simulation. The cell-
 1115 based reweighting seems to be promising to reduce data-simulation mis-modelling on electrons shower
 1116 shapes. Applying the same strategy to photons is the aim of this study.

1117 4.3.3 Cell-based reweighting for photons

1118 4.3.3.1 Event selection

1119 For this study, inclusive photons from $Z \rightarrow ll\gamma$ events are considered. The **SHERPA** generator is used for
 1120 simulated samples while the data correspond to the 2017 proton-proton collisions. The selection criteria
 1121 applied are the following:

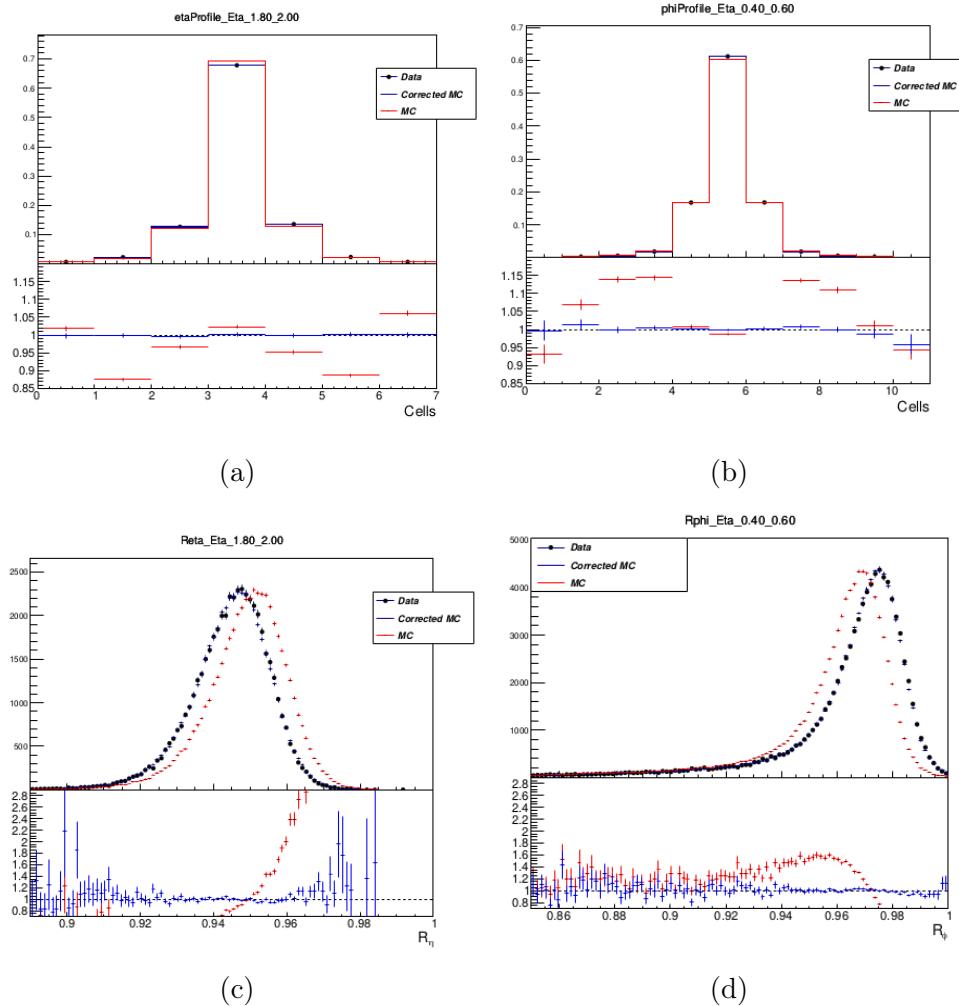


Figure 4.8: (a) Energy profile in η direction, (b) Energy profile in ϕ direction, (c) R_η variable and (d) R_ϕ variable before (red) and after reweighting for electrons from $Z \rightarrow e^+e^-$ for $1.80 < |\eta| < 2.00$ first column and $0.40 < |\eta| < 0.60$ in the second column (the bottom plot is the data/simulation ratio) [87].

1122 1. Photon selection:

- 1123 • Transverse energy $E_T > 10$ GeV.
- 1124 • Pseudorapidity $|\eta|$ within the calorimeter acceptance: $|\eta| < 2.37$, excluding the transition
1125 region between the barrel and end-cap (crack region): $1.37 < |\eta| < 1.52$
- 1126 • Isolation: tight.
- 1127 • Identification: None.

1128 2. Electron selection:

- 1129 • Transverse energy $E_T > 10$ GeV.
- 1130 • Pseudorapidity $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$.
- 1131 • Longitudinal impact parameter $z_0 < 10$ mm, transverse impact parameter significance $|d_0|/\sigma_{d_0}$
1132 less than 10.
- 1133 • Isolation: loose.
- 1134 • Identification: medium.

1135 3. Combined muon selection:

- 1136 • Transverse energy $E_T > 10$ GeV.
- 1137 • Pseudorapidity $|\eta| < 2.5$.
- 1138 • Longitudinal impact parameter $z_0 < 10$ mm, transverse impact parameter significance $|d_0|/\sigma_{d_0}$
1139 less than 10.
- 1140 • Isolation: loose.
- 1141 • Identification: medium.

1142 4. $Z \rightarrow ll\gamma$ event selection:

- 1143 • $\Delta R_{min}(e, \gamma) > 0.4$ and $\Delta R_{min}(\mu, \gamma) > 0.2$: the selection is tighter for electron channel to
1144 reduce photon from electron bremsstrahlung.
- 1145 • $80 < m_{ll\gamma} < 100$ GeV to select Z peak, and $40 < m_{ll} < 83$ GeV for final state radiation (FSR)
1146 event selection.

1147 No identification criteria is applied to photons to avoid any bias from the identification.

1148 4.3.3.2 Electrons reweighting applied to photons

1149 The first strategy has been to apply the electron reweighting values to photons. The developed reweight-
1150 ing matrix for electrons is directly applied to the selected photons from radiative Z events. Figure 4.9
1151 shows the energy profiles and layer-2 shower shapes variables before and after the reweighting. The
1152 reweighting goes in the correct direction but not enough to correct photons shower shape mis-modelling;
1153 this motivates to implement a specific reweighting for photons. Additional control plots for other η
1154 regions and converted photons are provided in Appendix A.1.

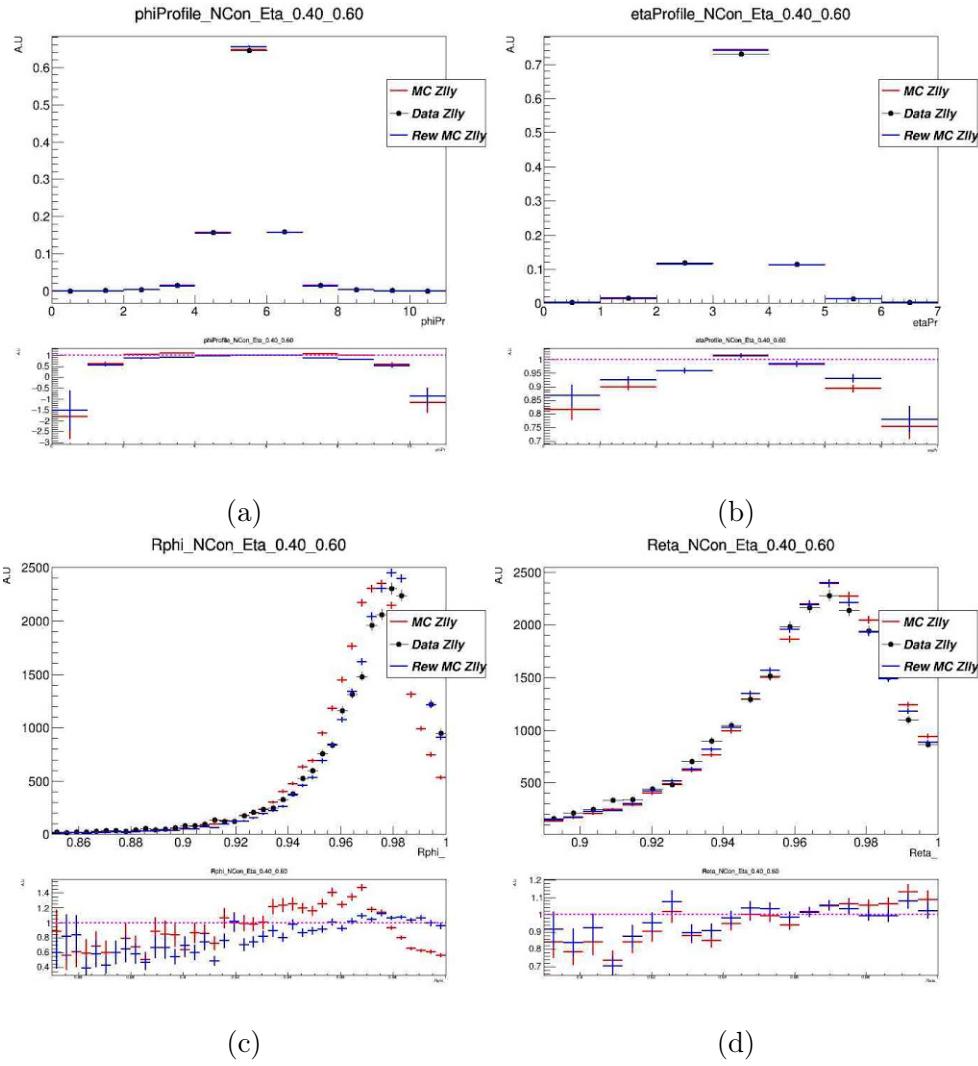


Figure 4.9: The energy profile in ϕ and η directions (a, b) and the corresponding R_ϕ and R_η variables, for unconverted photons with $0.40 < |\eta| < 0.60$. The black points correspond to Data 2017, red points to non-reweighted simulation and blue points to the reweighted simulation from $Z \rightarrow ll\gamma$.

4.3.3.3 Photon reweighting

The second step was to derive specific reweighting values for photons. Similarly to electrons, the reweighting is recomputed for photons using radiative Z events inclusively in conversion to reduce statistical uncertainties. Figure 4.10 shows the profiles and shower shapes after applying the derived reweighting on photons. A good agreement is observed in energy profiles for both direction η and ϕ . However, for shower shape variables, no significant improvement is observed and the reweighting method seems not to be working for photons. From the energy profiles in Figures 4.9 and 4.10 photons reweighting from radiative Z are larger than the one for electrons. Additional control plots for other η regions and converted photons are provided in Appendix A.2. To validate the reweighting procedure a closure test is performed on pseudo data (PS).

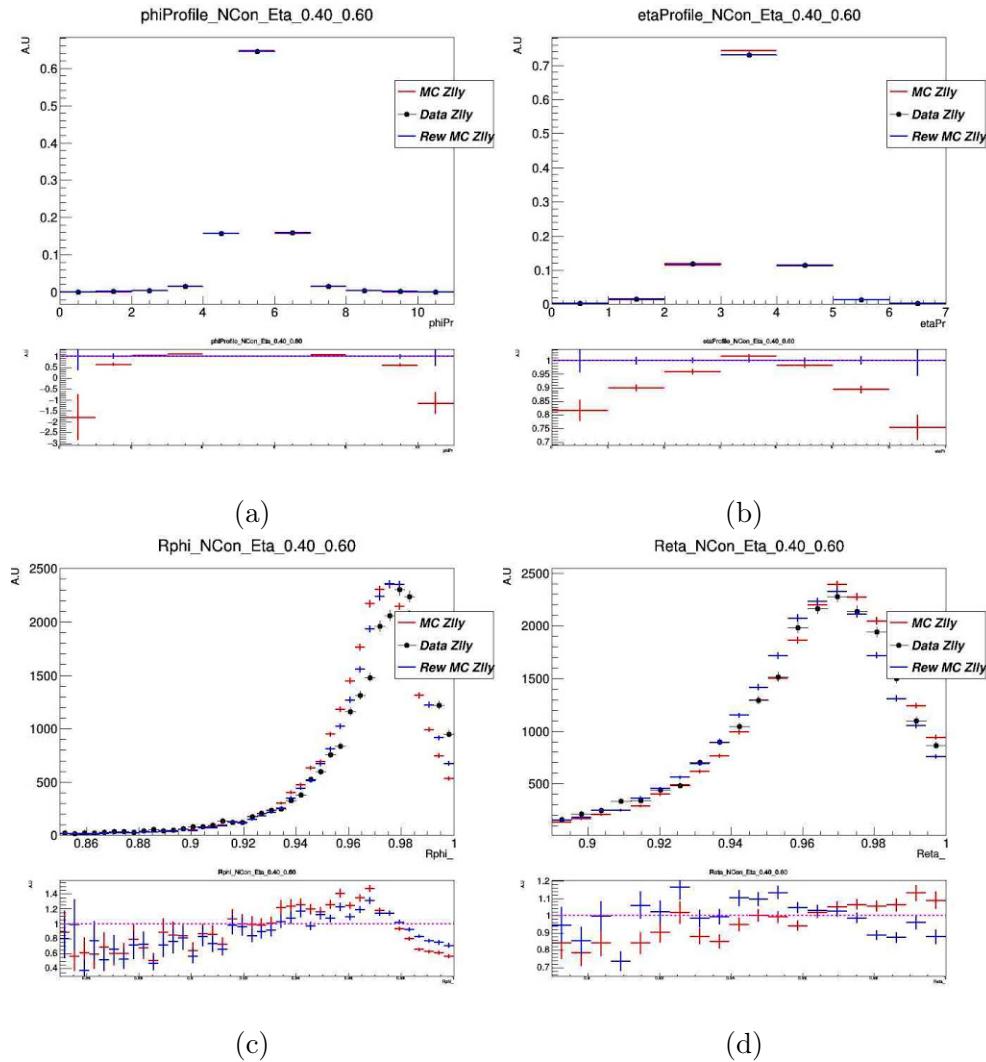


Figure 4.10: The energy profile in ϕ and η directions (a, b) and the corresponding R_ϕ and R_η variables, for unconverted photons with $0.40 < |\eta| < 0.60$. The black points correspond to Data 2017, red points to non-reweighted simulation and blue points to the reweighted simulation from $Z \rightarrow ll\gamma$.

1165 4.3.3.4 Closure test

1166 To validate the implemented code, a closure test was performed using PS from simulation, for that:

- 1167 1. PS is produced by adding 50 MeV (arbitrary value) to each simulated cell.
- 1168 2. Re-compute the reweighting function with the same procedure as above using unconverted photons
1169 (for simplicity).
- 1170 3. Correct simulated cell to match the PS.

Results of the closure test are shown in Figure 4.11.

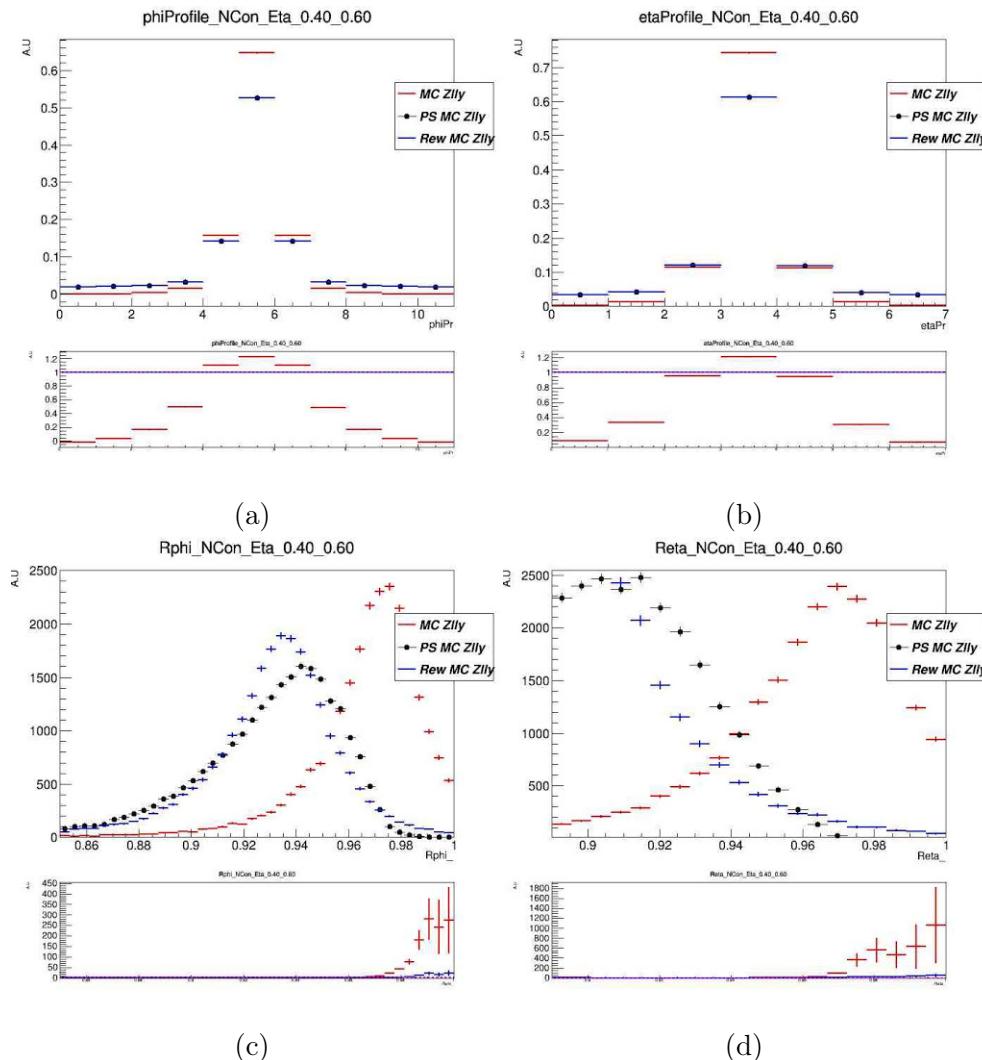


Figure 4.11: The energy profile in ϕ and η directions (a, b) and the corresponding R_ϕ and R_η variables, for unconverted photons with $0.40 < |\eta| < 0.60$. The black points correspond to the pseudo data, red points to non-reweighted simulation and blue points to the reweighted simulation from $Z \rightarrow ll\gamma$.

1171

1172 Perfect agreement in energy profiles for both directions (η and ϕ) is observed. The closure test reproduces

the PS, while for shower shapes the agreement is improved in the correct direction but not enough to correct the "fake" mis-modelling. The closure test demonstrates that the reweighting procedure was not able to catch the discrepancy between data and simulation, and seems to be working only on average and not event-by-event.

Figure 4.12 shows the averages of R_η and R_ϕ in real data, simulation and after reweighting in bin of photon $|\eta|$. After reweighting, the average of shower shapes in simulation perfectly matches the data distribution, which demonstrates that the reweighting is working on average and not event-by-event.

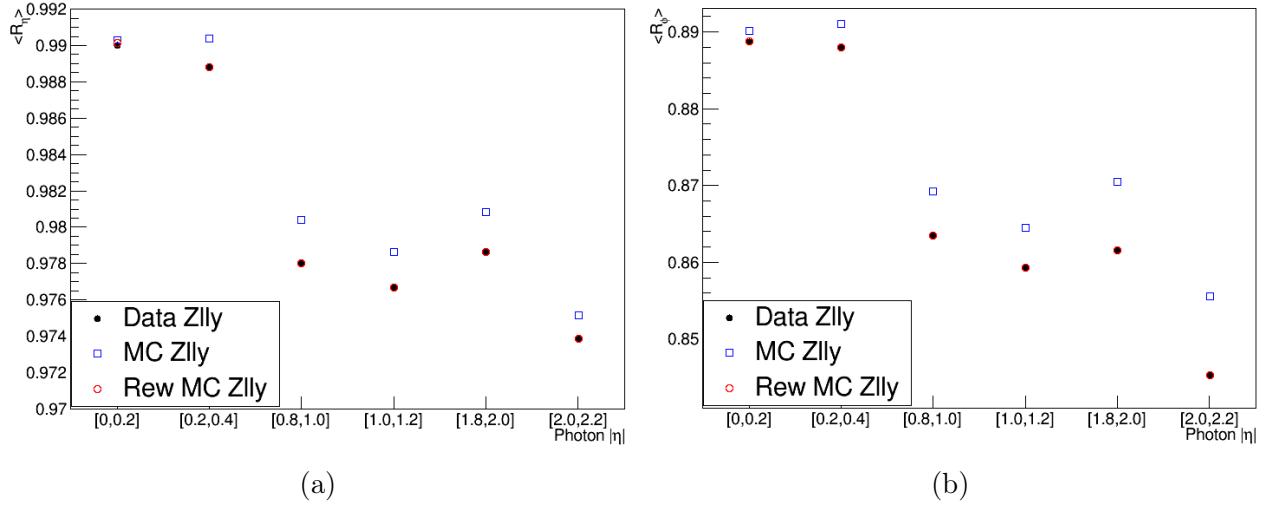


Figure 4.12: Average R_η (a) and R_ϕ (b) in real data (black), simulation (blue) and after reweighting (red) using photon reweighting values.

4.3.3.5 Three-dimensional reweighting

Since the reweighting is only correcting on average and not event-by-event, a new reweighting technique was developed which is presented here. The idea is to add cell energy as an additional dimension to the previous reweighting procedure. The reweighting factor for each cell will be a function of cell energy. Instead of 2D reweighting the reweighting will be in three dimensions, defined as:

$$E_{k,n}^{\text{reweighted}} = E_{k,n}^{\text{non-reweighted}} \times \alpha ; \text{with } \alpha = f(k, n, E_{k,n}^{\text{simulated}}) = E_{k,n}^{\text{data}} / E_{k,n}^{\text{simulated}}, \quad (4.4)$$

where $k = 1..77$ denotes the cell number, n the corresponding photon $|\eta|$ bin and $E_{k,n}^{\text{simulated}}$ the energy of the cell k in η bin n .

The new method requires matching between photons from data and simulation to compute the α factors, which is technically complicated. The procedure is tested with PS from simulation. The PS is derived from simulation by scaling the energy of the cell k by 0.26 (the factor 0.26 is arbitrary). For simplicity, the reweighting is evaluated using unconverted photons only. The result of the new reweighting procedure is presented in Figure 4.13. A good agreement is observed after the three-dimensional reweighting for both energy profiles and their corresponding shower shapes variables for layer 2. The improvement in data-simulation agreement observed makes the new method promising to be applied on real data.

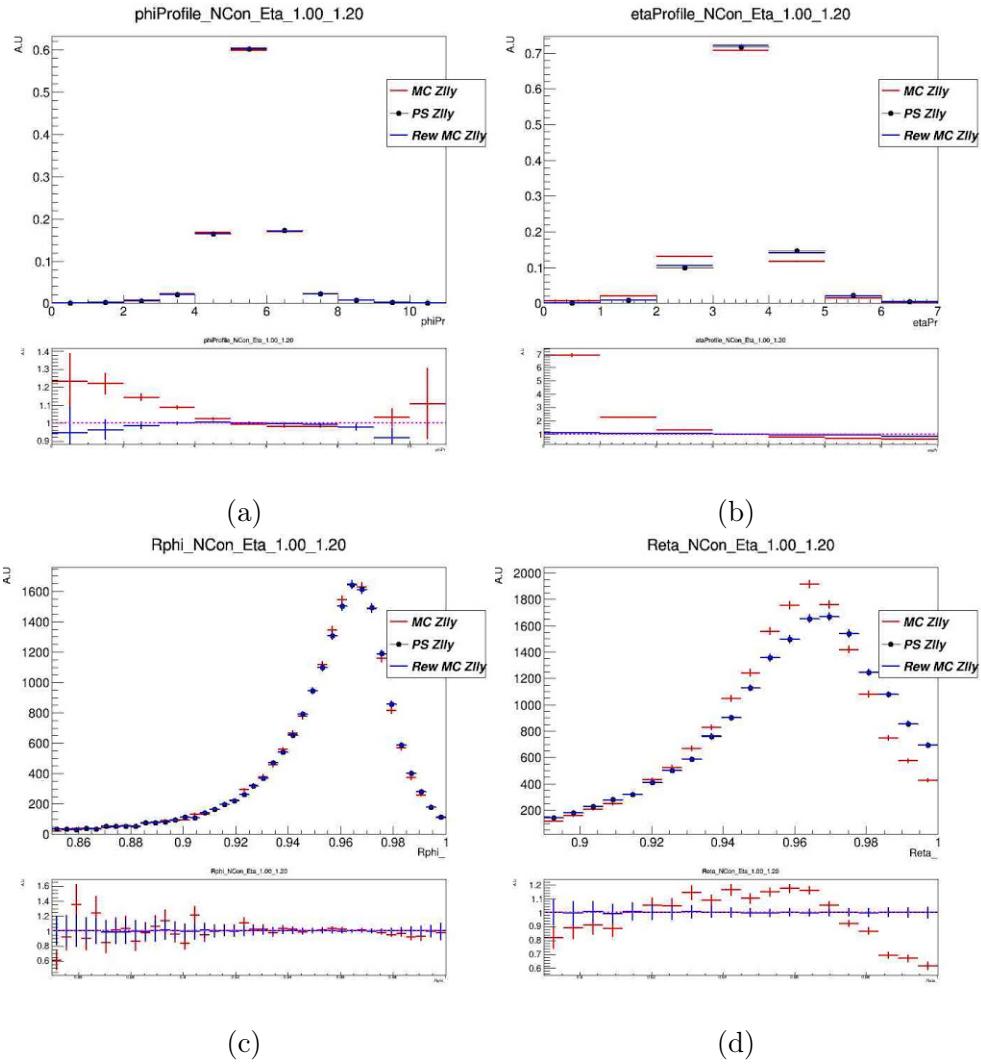


Figure 4.13: The energy profile in ϕ and η directions (a, b) and the corresponding R_ϕ and R_η variables, for unconverted photons with $1.00 < |\eta| < 1.20$. The black points correspond to the pseudo data, red points to non-reweighted simulation and blue points to the reweighted simulation from $Z \rightarrow ll\gamma$.

1190 4.4 Convolutional Neural Network for Photon Identification

1191 Applying cuts on shower shape variables for the identification algorithm limits the potential to improve
 1192 the separation between prompt and background photons. Correcting the mis-modelling seems to be
 1193 a critical problem without any solution for the moment. For these reasons, a further improvement
 1194 is possible by developing a new identification algorithm based on the EM cells and using advanced
 1195 Machine Learning (ML) techniques. The EM cells of the photon cluster are commonly represented as
 1196 an image, where each pixel contains the cell energy, leading to the use of images of the deposited energy
 1197 in the calorimeter to learn the shower properties. The most adapted ML technique for image processing
 1198 is Convolutional Neural Networks (CNN). In the following, its implementation using EM calorimeter
 1199 images is described. A complete description of NNs goes beyond the scope of this thesis. The interested
 1200 reader can find an introduction to NNs in Ref. [88].

1201 4.4.1 CNN algorithm

1202 4.4.1.1 Event selection

1203 Simulated samples of inclusive photon production are used to train the CNN. The inclusive photon
 1204 sample includes $\gamma + \text{jet}$ events from the hard subprocesses $qg \rightarrow q\gamma$ and $q\bar{q} \rightarrow g\gamma$ enriched by prompt
 1205 photon and used as a signal for the training. Background photons are taken from quark fragmentation
 1206 in QCD di-jet events. Photon candidates are requested to pass the same photon selection defined for
 1207 radiative Z in Section 4.3.3.1. Additionally, photons used as a signal in the training (prompt) are
 1208 required to be stable particle, not originate from a hadron and to match to a truth photon. Photons
 1209 failing one of these requirements are flagged as a background. Photons from radiative Z decay are used
 1210 as a control sample to evaluate the performance of the trained model (out-of-sample validation).

1211 4.4.1.2 Images pre-processing

1212 To compromise between collecting more energy and minimizing the impact of the activity around the
 1213 cluster, the 7×11 cluster around the hottest cell is considered similarly to the shower shapes computation.
 1214 As the EM cell granularity changes with η (Table 3.1), it leads to a different number of cells in each η
 region. The corresponding number of cells in 7×11 windows are summarized in Table 4.3.

$ \eta $ range	0 to 1.4	1.4 to 1.8	1.8 to 2.0	2.0 to 2.5
Sampling 1	112	112	84	56
Sampling 2	77	77	77	77
Sampling 3	44	44	44	44

Table 4.3: Number of cells in 7×11 EM calorimeter windows.

1215
 1216 The difference in shape and size of the EM calorimeter windows complicates the network training
 1217 procedure. To avoid this issue, zeros are appended to the calorimeter cluster image to end up with the
 1218 same image size in all η region for each sampling (e.g., sampling 1 with 84 cells is completed by zeros,
 1219 to have 112 cells). Table 4.4 shows the final image shape used for each sampling.
 1220 I decided to use the cell energy normalized to the total cluster energy as an image pixel to build an
 1221 energy independent algorithm which performs in the same way for all energy regime.

Sampling	Shape
Sampling 1	(56, 2)
Sampling 2	(7, 11)
Sampling 3	(4, 11)

Table 4.4: Image shape of 7×11 windows in each sampling in (η, ϕ) .

4.4.1.3 CNN Architecture

A CNN classifier is built to separate between prompt and background photons using images of the 7×11 cluster windows of sampling 1, 2, and 3. Figure 4.14 shows an illustration of the classifier.

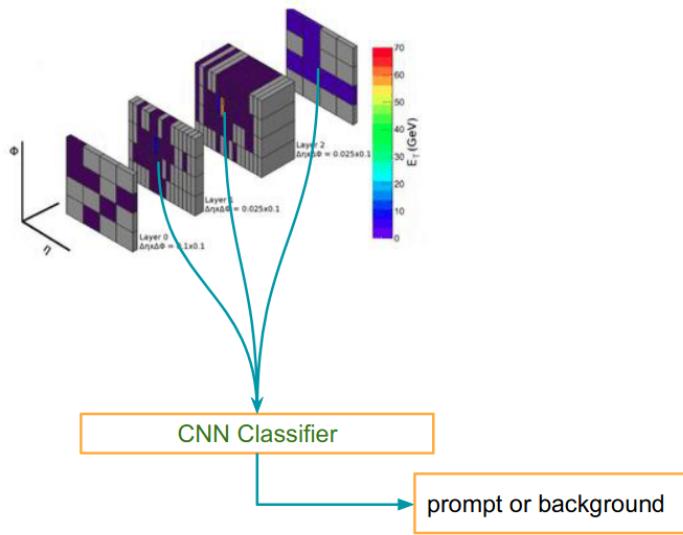


Figure 4.14: CNN classifier schema.

The Global CNN classifier is constructed using the KERAS library, with TENSORFLOW as a backend [89, 90]. Information from EM layers images is extracted using three baby CNNs. Each one is connected to an EM layer and built from two sets of two-dimensional convolution (2DCov) and two-dimensional max-pooling (2DMaxPool) [91] layers followed by a flatten layer that prepares a vector for the fully connected network. Output vectors of the three flatten layers are combined using a concatenate layer, then pass through a Dense Neural Network (DNN) for the classification task. The DNN is made of two Dense (fully connected) layers, a dropout, another Dense layer then a final output layer. An illustration of the architecture is given in Figure 4.15.

All layers in the network have their weights randomly initialized by sampling from a truncated normal distribution centred on zero with the width given by $\sqrt{1/N_{input}}$, where N_{input} is the number of input units in the weight tensor. The activation function of each layer is Scaled Exponential Linear Unit (SELU) [92] to preserve the mean and variance of the inputs between two consecutive layers and handle the normalization issues, except for the final output layer which has "sigmoid" as an activation function. Layers in the babies CNN are 256 nodes wide, while in the DNN 128 nodes except for the output layer is one node wide.

Using an output layer with a sigmoid activation function allows interpreting the output as the probability

for each class (prompt or background photon), given the input images.

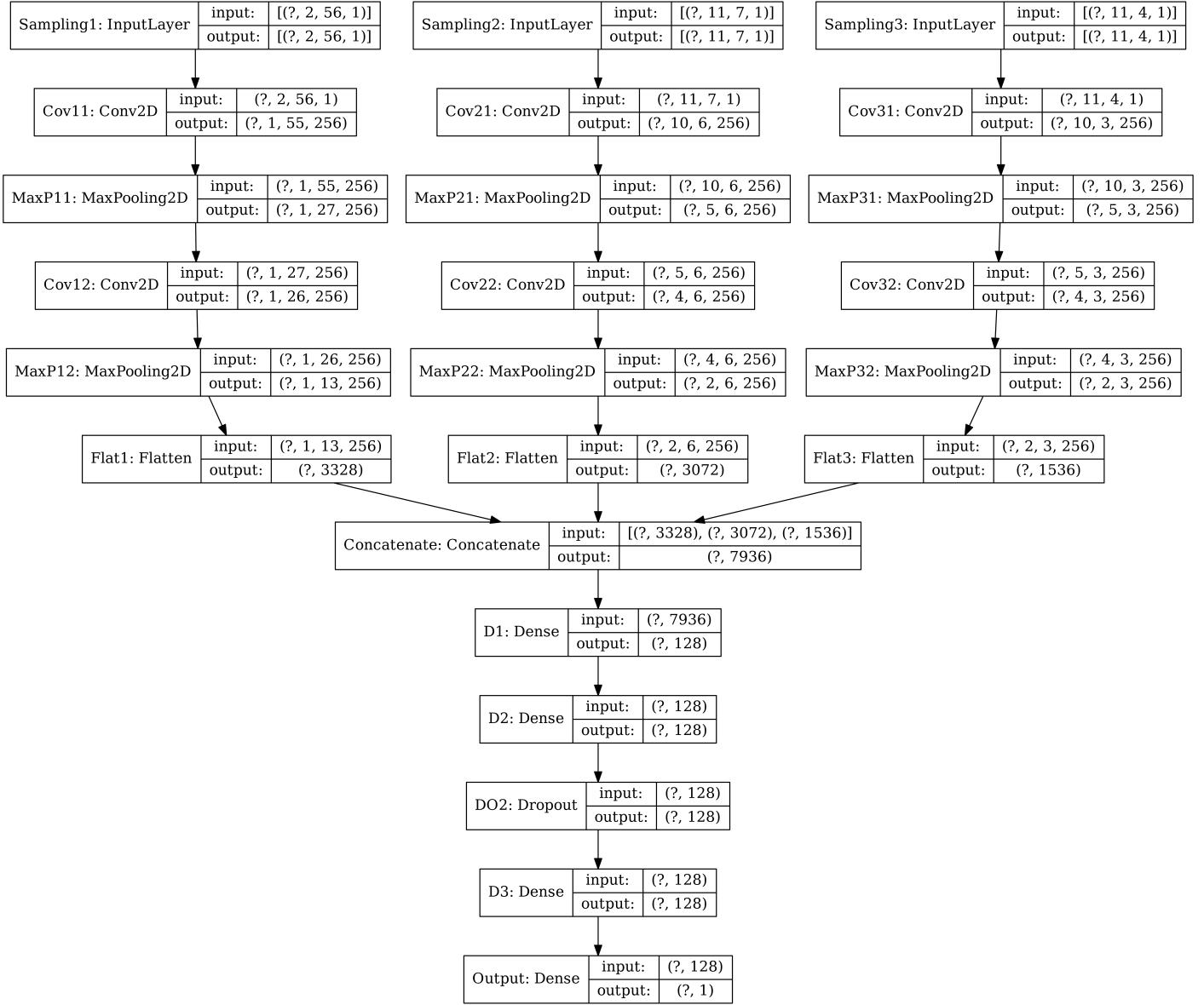


Figure 4.15: Illustration of the global CNN classifier graph.

In the network presented, the dropout layer has a dropout rate of 8%. During the training, the binary cross-entropy is used as a loss metric, the optimization is done with Adam optimizer.

4.4.1.4 CNN Training

The CNN of ~ 1.4 million parameters is trained with a learning rate of $1e^{-4}$ and a batch size fixed to 96 images (32 photons) due to memory limitations, with a maximum number of epochs set to 20. CNN network is trained using two NVIDIA Tesla K80 GPUs with a memory of 12 GB each run on CC-In2p3 cluster (France) [93]. Each training epoch lasts ~ 15 minutes. An early stopping metric [94] is imposed

1250 during the training phase to reduce the over-training/fitting, such that if the network learning reaches
 1251 the plateau it will stop. The metric used to determine the early stopping is the network loss as evaluated
 1252 on the validation data sample throughout the training. The network aims to minimize the weighted loss
 1253 function between the truth and predicted labels.

1254 The network is trained inclusively in E_T , η and photon conversion type. Only cluster images are
 1255 included during the training. Figure 4.16 shows the evaluation of loss function and the accuracy during
 1256 the training time for the training and validation dataset. After the 4th epoch the network achieves the
 1257 plateau and the validation loss does not decrease. To reduce the overtraining, the early stopping metric
 stops the training at epoch 4.

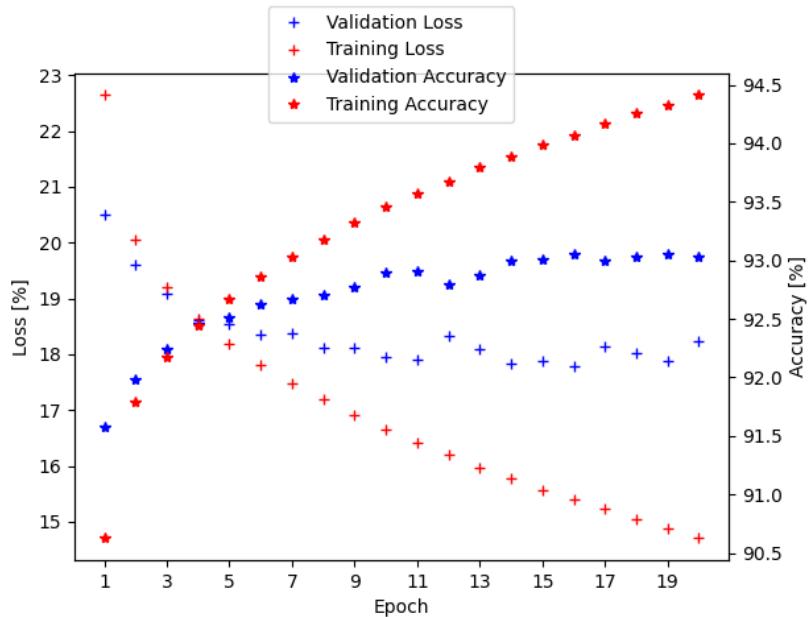


Figure 4.16: Neural network loss and accuracy as function of epoch number.

1258

1259 4.4.1.5 CNN Validation

1260 The CNN output is scanned to compute the improvement in two different approaches in bins of photon
 1261 p_T and $|\eta|$:

- 1262 • Background rejection for the same cut-based signal efficiency.
- 1263 • Signal efficiency for the same cut-based background rejection.

The signal efficiency ϵ_{eff} and background rejection ϵ_{rej} are evaluated as the following:

$$\epsilon_{eff} = \frac{TP}{TP + FN}; \quad \epsilon_{rej} = \frac{TN}{TN + FP}, \quad (4.5)$$

1264 where TP is the number of prompt photons correctly classified as a signal by the CNN, TN is the
 1265 number of background photons correctly classified as background, FP is the number of prompt photons
 1266 misclassified as background and FN is the number of background photons misclassified as a signal. The

following binning is used $p_T = [10, 20, 30, 40, 60, 80, \infty]$ GeV and $|\eta| = [0., 0.6, 1.37, 1.52, 1.8, 2.4]$. Figure 4.17 shows the receiver-operating characteristic (ROC) which illustrates the diagnostic ability of the CNN as its discrimination threshold is varied. The blue dot shows the signal and background rejection of the current cut-based tight WP. A significant improvement is obtained with the CNN compared to tight WP for both converted and unconverted photons.

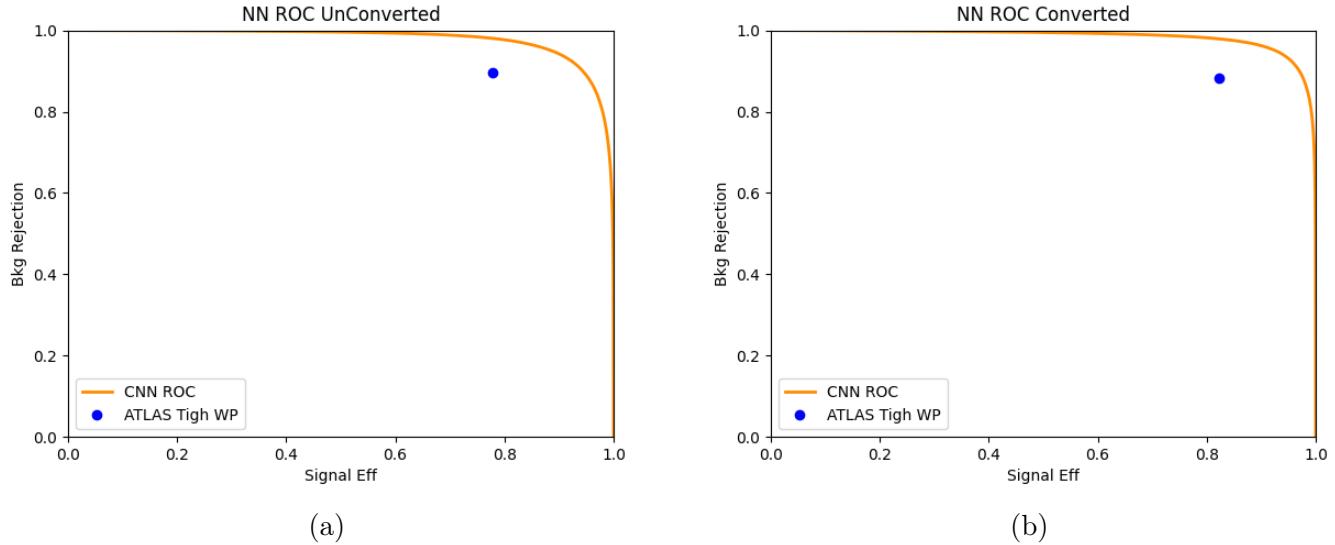


Figure 4.17: Network ROC curve for inclusive (a) unconverted photons (b) converted photons, the orange line show the ROC curve and the blue dot show the tight WP, x-axis is the ϵ_{eff} and y-axis show the ϵ_{rej} as defined in Eq. 4.5.

Figure 4.18 shows the relative improvement over the tight WP in terms of signal efficiency for the same background rejection. An improvement of up to 93% is achieved.

4.4.2 Photon efficiency estimation : Out-of-sample validation

Radiated photons from Z boson decay are used to evaluate the final improvement achieved by CNN after an appropriate optimization to define the CNN WPs. The optimization is done by maximizing the $S/\sqrt{(S+B)}$ with at least the same or higher background rejection as cut-based WPs in order to maximize the signal efficiency. CNN output is scanned in bins of $|\eta|$ for both converted and unconverted photons separately to define two WPs (Loose and Tight). To be more consistent with cut-based tight WP the following binning is chosen: $[0, 0.6, 0.8, 1.15, 1.37, 1.81, 2.01, 2.37]$.

The identification efficiency is evaluated using the Radiative Z method already described in Section 4.2. Due to the contamination from the $Z \rightarrow ll + \text{jet}$, where the jet is misidentified as photons. The efficiency should be corrected by subtracting this background component from the counted numbers (N). Therefore, a template fit method was performed to estimate the signal purity (P) both before and after applying the CNN tight WP. To perform the template fit, the $m_{ll\gamma}$ probability density function for signal ($Z \rightarrow ll\gamma$) is extracted from simulated events while the background ($Z \rightarrow ll + \text{jet}$) is fitted using a second-order polynomial function. The likelihood sum of the signal and background pdfs, with floating normalization, is fitted to the data $m_{ll\gamma}$ distribution in the range [60, 120] GeV. Then, the signal purities

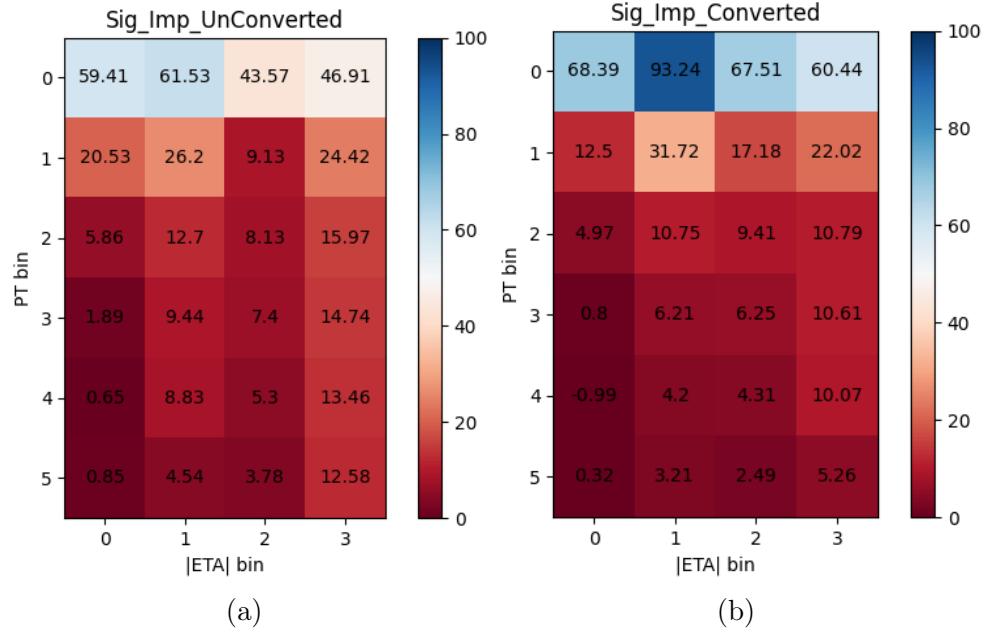


Figure 4.18: Relative signal efficiency improvement for similar tight WP background rejection (a) for unconverted photons and (b) converted photons.

(P) are calculated within the signal region [80, 100] GeV, making the efficiency corrected as:

$$\epsilon_{ID} = \frac{N_{\text{after ID}} \times P_{\text{after ID}}}{N_{\text{before ID}} \times P_{\text{before ID}}}, \quad (4.6)$$

¹²⁷⁵ The purities are evaluated in 5 E_T bins separately for converted and unconverted photons as well as for
¹²⁷⁶ $e\gamma\gamma$, $\mu\mu\gamma$ and $ll\gamma$ channels. The purity before the CNN selection is reported in Table 4.5.

	Unconverted			Converted		
	$ee\gamma$	$\mu\mu\gamma$	$ll\gamma$	$ee\gamma$	$\mu\mu\gamma$	$ll\gamma$
$10 < E_T < 15$	98.1 ± 4.7	96.2 ± 1.9	96.6 ± 1.8	97.8 ± 7.7	95.7 ± 3.4	96.2 ± 3.1
$15 < E_T < 20$	99.4 ± 12.7	99.1 ± 5.1	90.7 ± 1.9	99.3 ± 20.2	98.7 ± 7.5	98.8 ± 6.9
$20 < E_T < 25$	99.5 ± 24.2	99.6 ± 10.3	99.6 ± 9.5	99.2 ± 29.4	99.4 ± 13.3	99.2 ± 11.2
$25 < E_T < 30$	99.4 ± 33.8	99.7 ± 15.2	99.7 ± 14.5	99.3 ± 51.4	99.5 ± 19.1	99.5 ± 18.2
$30 < E_T < 100$	99.1 ± 33	99.7 ± 17.8	99.7 ± 16.5	99.1 ± 50.8	99.6 ± 23.4	99.5 ± 21.4

Table 4.5: Fitted photon purity in % of all probes, before tight CNN WP. The uncertainties are only statistical.

1278 Table 4.6 reports the purity after the CNN selection. Template fit results in the $ll\gamma$ channel is shown
in Appendix B.4.

	Unconverted			Converted		
	$ee\gamma$	$\mu\mu\gamma$	$ll\gamma$	$ee\gamma$	$\mu\mu\gamma$	$ll\gamma$
$10 < E_T < 15$	99.6 ± 11.2	98.9 ± 4.1	99.1 ± 3.8	99.3 ± 16.1	98.5 ± 6.4	98.7 ± 6.02
$15 < E_T < 20$	99.8 ± 23.1	99.7 ± 8.9	99.7 ± 8.1	99.7 ± 34.1	99.6 ± 13.2	99.6 ± 12.0
$20 < E_T < 25$	99.8 ± 38.1	99.8 ± 15.2	99.8 ± 14.1	99.7 ± 46.6	99.6 ± 17.9	99.7 ± 18.1
$25 < E_T < 30$	99.7 ± 49.1	99.8 ± 18.8	99.8 ± 18.5	99.6 ± 66.1	99.7 ± 25.6	$99.7 \pm 25.$
$30 < E_T < 100$	99.5 ± 45.6	99.8 ± 21.2	99.8 ± 20.7	98.9 ± 49.4	99.7 ± 28.5	99.7 ± 25.8

Table 4.6: Fitted photon purity in % of all probes, after tight CNN WP. The uncertainties are only statistical.

1279
1280 The small amount of data at high E_T leads to large uncertainties. It indicates that the probes with
1281 $E_T > 30$ GeV are very pure photon samples both before and after applying the CNN. Therefore, the
1282 template fit method is only applied for the first E_T bins, and the efficiencies for the rest of bins are
1283 obtained by counting. The combined $ll\gamma$ channel is used to estimate the efficiency and evaluate the scale
1284 factors.

1285 Figures 4.19 and 4.20 show photon identification efficiency for tight CNN WP compared to cut-based
1286 tight as a function of photon E_T for unconverted and converted photons respectively. CNN over performs
1287 the cut-based algorithm. Ratio plots show the scale factor.

1288 The CNN scale factors are reported in Table 4.7 and Table 4.8, and found to be closer to unity except
1289 for the first E_T bin where the scale factor is slightly large but still better than cut-based as shown in
the ratio plots.

	$0 < \eta < 0.6$	$0.6 < \eta < 1.37$	$1.81 < \eta < 2.01$	$2.01 < \eta < 2.37$
$10 < E_T < 15$	0.965 ± 0.0046	0.984 ± 0.0051	1.045 ± 0.0155	1.04 ± 0.011
$15 < E_T < 20$	1.01 ± 0.0032	1.03 ± 0.0036	1.01 ± 0.0114	0.982 ± 0.008
$20 < E_T < 25$	1.002 ± 0.0026	1.003 ± 0.003	0.980 ± 0.0099	0.992 ± 0.0093
$25 < E_T < 30$	0.999 ± 0.0025	0.998 ± 0.003	0.997 ± 0.0114	1.001 ± 0.091
$30 < E_T < 35$	0.999 ± 0.0032	0.999 ± 0.0032	1 ± 0.0746	
$35 < E_T < 40$	0.996 ± 0.0051	1.002 ± 0.0043		
$40 < E_T < 45$	0.999 ± 0.0067	0.995 ± 0.0077		
$45 < E_T < 50$	1 ± 0.0086	1.001 ± 0.0093		
$50 < E_T < 60$	1.004 ± 0.0114	1 ± 0.0097		
$60 < E_T < 80$	1.006 ± 0.0179	1.012 ± 0.0326		
$80 < E_T < 100$	1 ± 0.07			

Table 4.7: Scale factors for tight CNN WP efficiency measured with unconverted photons from
 $Z \rightarrow ll\gamma$ decays, in various bins of pseudorapidity and transverse energy. The uncertainty includes
only statistical components.

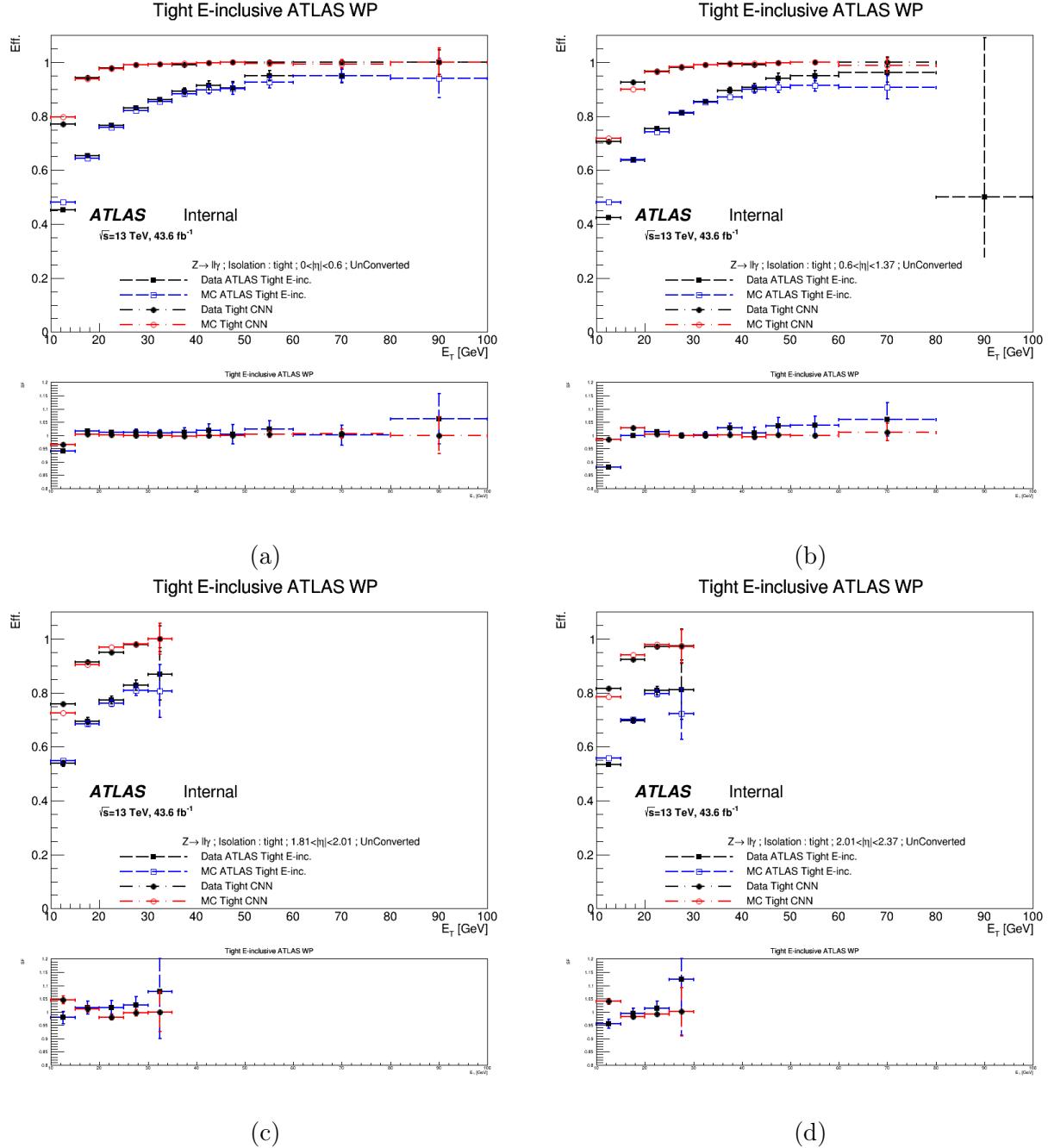


Figure 4.19: The unconverted photon identification efficiency as a function of photon E_T for CNN (red) and cut-based (blue) indicated with ATLAS, for data (full marks) and simulation (open marks) for 4 $|\eta|$ regions. Bottom ratios show the scale factors.

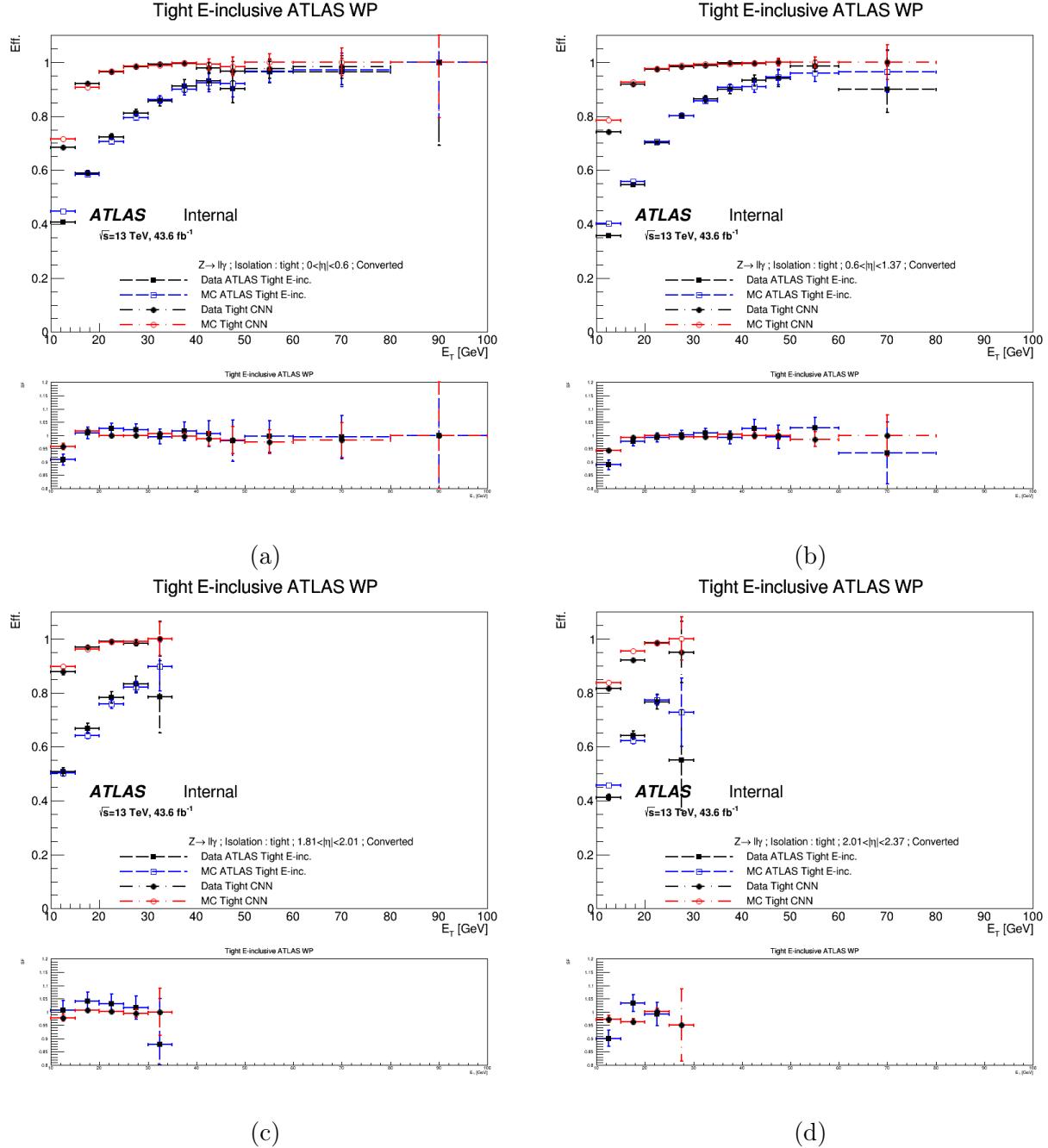


Figure 4.20: The converted photon identification efficiency as a function of photon E_T for CNN (red) and cut-based (blue) indicated with ATLAS, for data (full marks) and simulation (open marks) for 4 $|\eta|$ regions. Bottom ratios show the scale factors.

	$0 < \eta < 0.6$	$0.6 < \eta < 1.37$	$1.81 < \eta < 2.01$	$2.01 < \eta < 2.37$
$10 < E_T < 15$	0.96 ± 0.012	0.944 ± 0.0085	0.978 ± 0.013	0.973 ± 0.013
$15 < E_T < 20$	1.02 ± 0.008	0.99 ± 0.006	1.008 ± 0.01	0.963 ± 0.011
$20 < E_T < 25$	0.999 ± 0.0067	0.99 ± 0.0044	1.001 ± 0.008	1.003 ± 0.012
$25 < E_T < 30$	0.999 ± 0.0063	0.995 ± 0.0044	0.994 ± 0.014	0.95 ± 0.137
$30 < E_T < 35$	1.007 ± 0.0079	0.995 ± 0.0054		1 ± 0.089
$35 < E_T < 40$	0.997 ± 0.0114	1.004 ± 0.0066		
$40 < E_T < 45$	0.987 ± 0.0277	0.998 ± 0.012		
$45 < E_T < 50$	0.983 ± 0.0508	$1. \pm 0.018$		
$50 < E_T < 60$	0.976 ± 0.0455	0.99 ± 0.0279		
$60 < E_T < 80$	0.983 ± 0.0654	$1. \pm 0.078$		
$80 < E_T < 100$		1 ± 0.37		

Table 4.8: Scale factors for tight CNN WP efficiency measured with converted photons from $Z \rightarrow ll\gamma$ decays, in various bins of pseudorapidity and transverse energy. The uncertainty includes only statistical components.

1291 The CNN is not sensitive to shower shapes mis-modelling as it uses the cell fraction energy where the
 1292 mis-modelling is infinitesimal and not caught by the CNN algorithm. This is illustrated in Figure 4.21
 1293 where no significant difference in the real and simulated data shapes of the CNN distribution is observed.

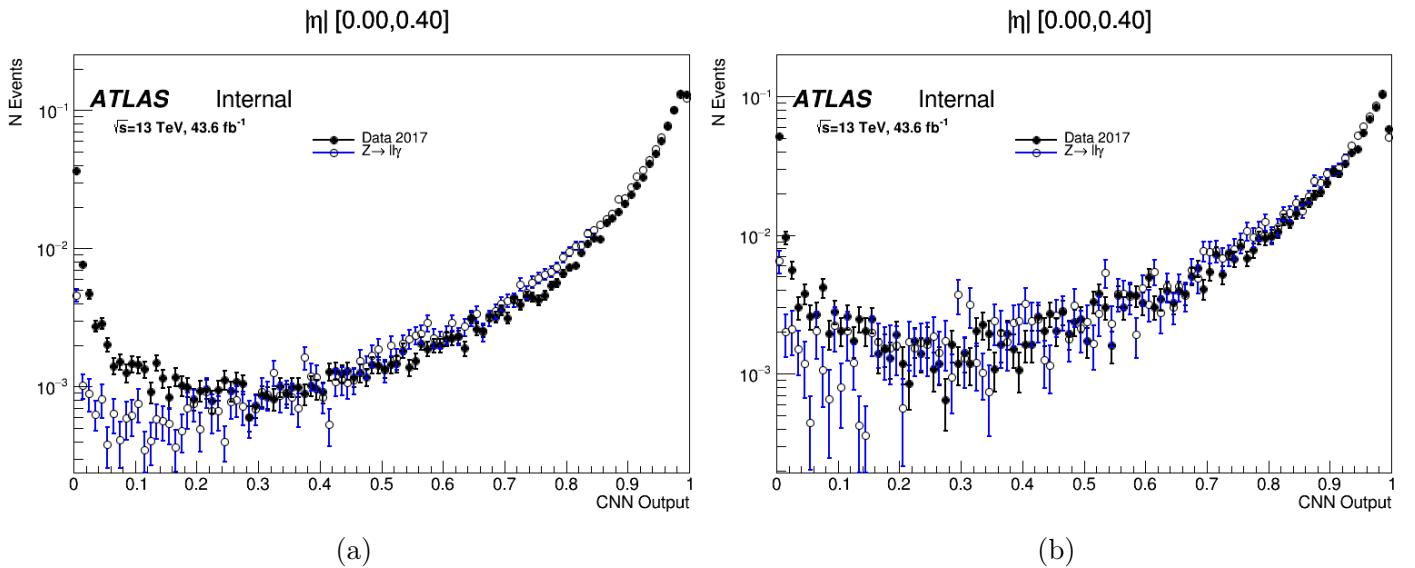


Figure 4.21: CNN output distribution as computed for data (black) and simulation (blue) for (a) for unconverted photons and (b) converted photons with $|\eta| < 0.4$.

1294
 1295 At low E_T and especially for high $|\eta|$, the photon efficiency dropped because photons with low E_T
 1296 deposit more energy in the first sampling compared to photons with high E_T , and the zeros added to
 1297 the first sampling whitens more than half of the image which reduce the CNN learning ability for this

1298 category of photons.

1299 The presence of additional pp interactions in the event reduces the photon identification and isolation
1300 efficiencies. Figure 4.22 shows the evolution of photon identification efficiency as a function of the average
1301 number of interaction per bunch crossing ($\langle \mu \rangle$) and the number of primary vertices (nPV). A clear
1302 drop by $\sim 9\text{-}16\%$ when going from $\mu \sim 5$ to $\mu \sim 70$, depending on the photon candidate pseudorapidity
1303 and conversion status. The drop is explained by the additional activities around the hottest cell affecting
1304 the CNN performance of extracting the shower.

1305 The proposed identification algorithm is getting integrated into the ATLAS Athena framework and is
1306 planned to be used as a possible baseline for Run-3.

1307 Additional control plots for Tight CNN and Loose CNN WPs are shown in Appendix B.

1308 4.5 Conclusion

1309 As mentioned at the beginning of this chapter, photon identification is an essential ingredient to select
1310 and study the $HH \rightarrow b\bar{b}\gamma\gamma$ decay. ATLAS provides a cut-based algorithm that uses global shower shapes
1311 variables to identify prompt photons. However, the shower shapes are affected by the mis-modelling
1312 discrepancy, leading to a scale factor far from unity. Besides, the cut-based algorithm suffers from the
1313 features space dimensionality which limits its performance. The CNN algorithm provides a solution to
1314 these problems by using low-level detector information. Using the neural network with low-level detector
1315 information scale up the features space. The CNN over-performs the cut-based method by generating its
1316 complicated variables and handling their correlations for better separation. Even with the performance
1317 achieved, CNN still needs to be improved to handle some issues: the low efficiency at low E_T and the
1318 pile-up effect which reduces its performance for high pile events. Appendix C is dedicated to a discussion
1319 on future improvements for the CNN identification algorithm.

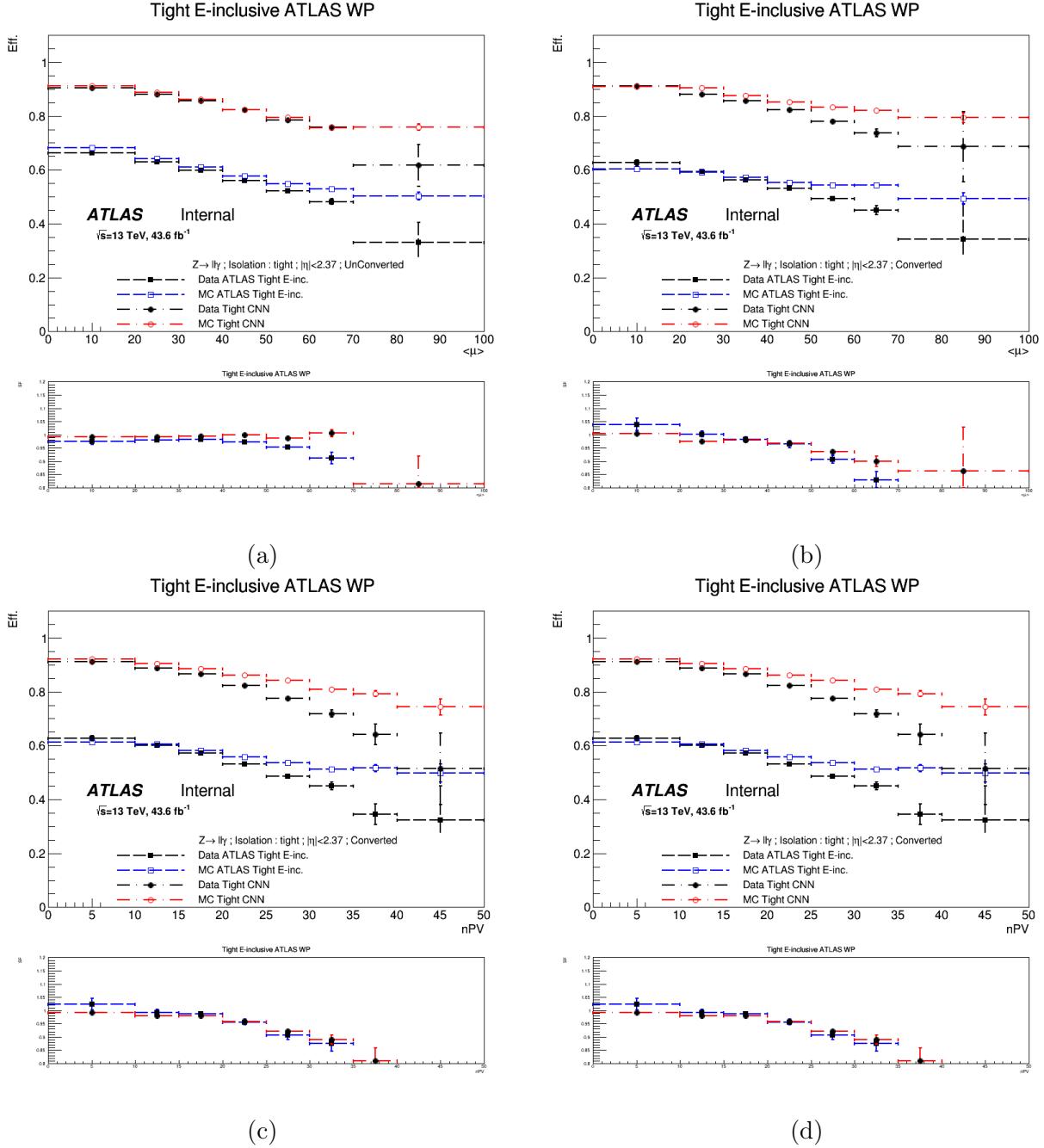


Figure 4.22: Photon identification efficiency for CNN (red) and cut-based (blue) indicated with ATLAS, for data (full marks) and simulation (open marks). Unconverted photons (right) and converted photons (left). As a function of pile up $\langle \mu \rangle$ (top) and number of primary vertices (nPV) (bottom). Bottom ratios show the scale factors.

1320 Chapter 5

1321 Jet reconstruction and Tagging in ATLAS

1322 Jet objects and specifically b -quark jets are crucial to maximize the precision on the measurement of
1323 the $HH \rightarrow b\bar{b}\gamma\gamma$ properties. Jets in ATLAS are currently reconstructed in two ways (EMTopo and
1324 PFlow) as introduced in Section 3.4.4. This chapter details jet reconstruction, b -jet identification and
1325 calibration implemented in $HH \rightarrow b\bar{b}\gamma\gamma$ analysis. I implemented a calibration method specific to b -jets
1326 which provides better reconstruction performances, thus better $m_{b\bar{b}}$ invariant mass resolution compared
1327 to generic jet corrections.

1328 5.1 Reconstruction algorithms

For both jet algorithm types at Run-2, jet reconstruction starts with a clustering algorithm based on calorimeter objects to build topological clusters [95]. The algorithm identifies calorimeter cells with energy larger than $4 \sigma_{noise}$ (expected noise standard deviation) as cluster seeds. Then, neighbouring cells with $|E_{cell}| > 2\sigma_{noise}$ are grouped to form the cluster as well as all adjacent cells. Clusters with two or more local energy maxima are split in two with cells being associated with each maximum according to their relative distance.

The EMTopo jets are built later by using the Anti- k_t algorithm. This algorithm selects the most energetic cluster and associates it to all neighbouring clusters in descending transverse momentum order satisfying:

$$\min(k_{T,i}^{-2}, k_{T,j}^{-2}) \frac{\Delta R_{ij}}{R^2} < k_{T,i}^{-2}, \quad (5.1)$$

1329 where $k_{T,i}^{-2}$ is the transverse momentum of cluster i, ΔR_{ij} is the relative distance between clusters, and
1330 R is the jet radius fixed to 0.4 for all jets used in this thesis.

1331 Alternatively, the PFlow jets are built by combining the topo-clusters to the ID tracks. Firstly, well-
1332 measured tracks are selected by requesting a tight quality selection. Tracks are required to be within
1333 $|\eta| < 2.5$ and have $p_T > 0.5$ GeV. They must have at least nine hits in the silicon detector and no missing
1334 pixel hit. Tracks with $p_T > 40$ GeV are excluded as well as tracks matched to electrons or muons with
1335 medium quality criteria and their deposited energy in the calorimeter is taken into account when building
1336 the jet [96]. No requirement on the track-vertex is applied to the track in this stage. Then, the algorithm
1337 assigns each track to a single topological cluster to form a track/topo-cluster system. The extrapolated
1338 (to the second sampling of the EM) distance between the cluster barycentre and the track is used in
1339 addition to the $E^{cluster}/p^{track}$ in the matching. The expected energy in the calorimeter, deposited by
1340 particles, is then computed. Additional topo-clusters are added to the track/topo-cluster system if the
1341 particle deposits its energy in more than one topo-cluster. At this stage, the expected energy deposited

1342 by the particle in the calorimeter is subtracted cell by cell from the set of the matched topo-clusters.
 1343 Finally, if the remaining energy in the system is consistent with the expected shower fluctuations, the
 1344 topo-cluster remnants are removed. Finally, the PFlow algorithm combines the calorimeter and soft
 1345 tracks which do not reach the calorimeter to improve jet resolution at low p_T compare to EMTopo jets
 (As Figure 5.3 demonstrates). This algorithm is sketched in Figure 5.1.

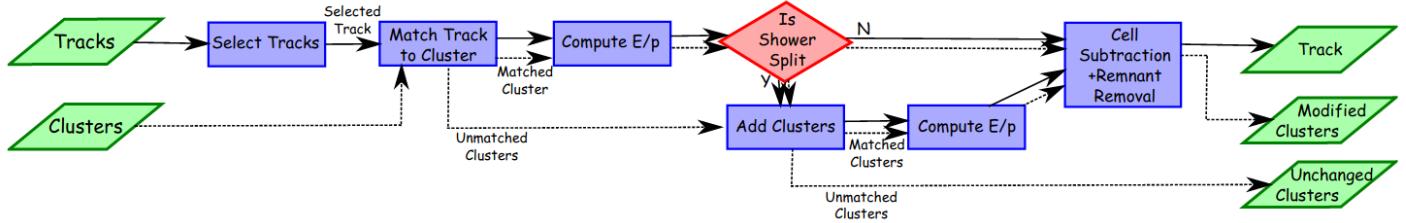


Figure 5.1: A flow chart of how the particle flow algorithm proceeds, starting with track selection and continuing until the energy associated with the selected tracks have been removed from the calorimeter [96].

1346
 1347 Similarly to EMTopo jets, the anti- k_t algorithm with $R=0.4$ is used to reconstruct the PFlow jets.
 1348 The algorithm takes as inputs the survived topo-clusters and the associated tracks that are matched
 1349 to the hard scatter primary vertex. These criteria are applied to reject tracks originating from pile-up
 1350 interaction [97].

1351 The PFlow is the jet type used in this thesis, while in the rest of this chapter the word "jet" refers to
 1352 both EMTopo and PFlow.

1353 5.2 Jet Energy Calibration

1354 The resulting jets from the simulation have to be calibrated such that, on average, the truth jet energy
 1355 corresponds to the reconstructed one in the ATLAS detector. A jet calibration procedure is developed
 1356 to calibrate jets. This calibration proceeds in several sequential stages of calibration derived from
 1357 a combination of simulation-based methods, which correct the jet 4-momentum to the truth one in
 1358 simulation, and *in-situ* techniques that measure the difference in jet response between real and simulated
 1359 data with residual correction applied to jets in the real data only.

1360 5.2.1 Jet Calibration Chain

1361 The first step in the calibration chain aims to recalculate the jet direction to point to the hard scatter
 1362 primary vertex rather than the centre of the detector. This correction only affects the jet η and resulting
 1363 in a better η resolution without changing the jet energy called origin correction. Figure 5.2 presents an
 1364 overview of the ATLAS calibration scheme for jets.

1365 5.2.1.1 Area-based and residual pile-up correction

This correction aims at subtracting the per-event energy excess due to in-time and out-of-time pile-up
 contaminants (Section 3.1.3). The contribution from pile-up is evaluated from the median transverse
 energy density ρ . The correction is linked to the jet area A and derived for jet within $|\eta| < 2$ given
 the higher calorimeter occupancy in the forward region. An additional residual correction is derived
 to complete correction for the pile-up rate dependence which remains after the area correction [99].

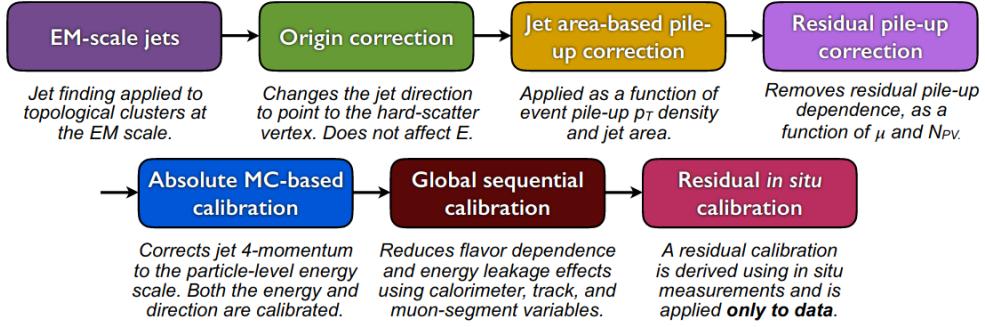


Figure 5.2: Calibration stages for jets. Other than the origin correction, each stage of the calibration is applied to the four-momentum of the jet [98].

The dependence is seen on the number of reconstructed primary vertices N_{PV} , sensitives to the in-time pile-up, and the number of pp interactions in the event μ , sensitives to the out-of-time pile-up. The residual correction is derived as the difference between the reconstructed jet p_T and the truth jet p_T using a geometric reconstructed-truth matching with $\Delta R = 0.3$. The area-based and pile-up correction is applied as:

$$p_T^{\text{corr}} = p_T^{\text{reco}} - \rho A - \alpha (N_{PV} - 1) - \beta \mu, \quad (5.2)$$

where α and β are the slopes of the linear dependency of jet response in respect to N_{PV} and μ .

5.2.1.2 Absolute MC-based calibration

This correction aims to correct, through a simulation-based procedure, the reconstructed jet energy scale (JES) to the truth-level scale as well as the direction biases caused by the different calorimeter technologies and granularity. The average JES response is defined as the mean of a Gaussian fit to the core of the $R = E^{\text{reco}}/E^{\text{truth}}$ distribution in (E^{truth}, η) bins [100, 101]. A numerical inversion is used to parameterize the correction as a function of E^{reco} , the correction is applied as:

$$E_{\text{JES}}^{\text{jet}} = \frac{E^{\text{jet}}}{F(E^{\text{jet}}, \eta)}, \quad (5.3)$$

where $E_{\text{JES}}^{\text{jet}}$ and E^{jet} are jet energy after and before the calibration, $F(E^{\text{jet}}, \eta)$ is the inverted jet response for a given (E, η) bin. A second correction is derived to account for a bias seen in the reconstructed jet η . The η correction is computed as the difference between the reconstructed η^{reco} and the truth η^{truth} as a function of E^{truth} and η . Similarly to JES correction, a numerical inversion is used to derive the correction in E^{reco} . This correction is evaluated on a multi-jet events sample. At this step, EMTopo (PFlow) jet is denoted EM+JES (PFlow+JES). Comparisons of the jet energy resolution measurement for PFlow+JES and EM+JES jets as a function of jet p_T is shown in Figure 5.3.

5.2.1.3 Global Sequential Calibration

The Global Sequential Calibration (GSC) corrects effects due to intrinsic properties of the jets not properly taken into account in previous steps and the hadron composition yielded by the fragmentation process. The average particle composition and shower shape of a jet vary between initiating particles, most notably between quark- and gluon-initiated jets. The GSC uses five observables to improve the resolution of the JES jets. For each observable, an independent jet 4-momentum correction is derived as a function of p_T^{truth} and $|\eta|$ by inverting the reconstructed jet response in simulated events. The effect

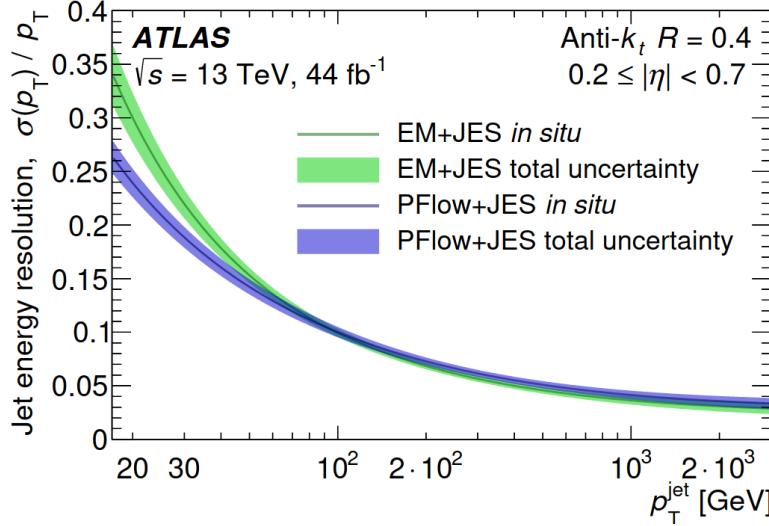


Figure 5.3: The relative jet energy resolution in simulation for fully calibrated PFlow+JES jets (green curve) and EM+JES jets (blue curve) as a function of p_T jet.

1382 of each correction is therefore to remove the dependence of the jet response on each observable while
 1383 conserving the overall energy scale at the JES step. No correlation between observables is taken into
 1384 account. The five observables are:

- 1385 1. $f_{\text{Tile}0}$, the fraction of jet energy measured in the first layer of the hadronic Tile calorimeter.
- 1386 2. $f_{\text{LAr}3}$, the fraction of jet energy measured in the last layer of the electromagnetic LAr calorimeter.
- 1387 3. n_{trk} , the number of tracks with $p_T > 1 \text{ GeV}$ associated to the jet.
- 1388 4. \mathcal{W}_{trk} , the average p_T -weighted transverse distance between the jet axis and all tracks of $p_T > 1$
 1389 GeV associated to the jet.
- 1390 5. n_{segments} , the number of muon track segments associated with the jet.

1391 5.2.1.4 Residual $in - situ$ calibration

1392 The last step of the jet calibration chain recovers the disagreement between real and simulated data
 1393 in the jet response coming from the imperfect detector simulation and mismodelling. The correction is
 1394 extracted by balancing the jet p_T against that of well-measured reference objects from $Z + \text{jets}$, $\gamma + \text{jets}$
 1395 and di-jet events. Figure 5.4 shows the final $in - situ$ combination as a function of jet p_T . To complete
 1396 the calibration, the inverse of the curve ($R_{\text{MC}}/R_{\text{data}}$) is taken as the scaling factor and applied to data.
 1397 Additional details on each calibration steps can be found in the Ref. [98].

1398 Figure 5.3 demonstrates the performance of PFlow at low jet p_T coming from combining calorimeter
 1399 and tracking information as well as the low threshold on track p_T , for high jet p_T the performances of
 1400 both EMTopo and PFlow are comparable.

1401 5.3 Jet Tagging : Pile-up and flavour taggings

1402 Jets coming from pile-up vertices can have a disastrous impact on physics analysis, by leading to fake jets
 1403 and therefore biasing the event selection. PU jets are coming either from hard-scatter interaction from

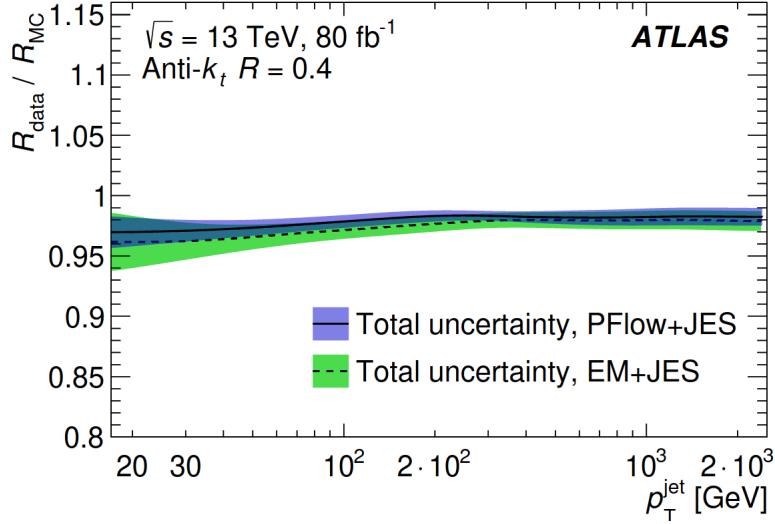


Figure 5.4: Ratio of the PFlow+JES (blue curve) and EM+JES (green curve) jet response in data to that in the nominal event generators as a function of p_T jet.

1404 a different primary vertex, defining the most frequent type of jet QCD-PU, or from interactions that are
1405 not related to the initial one defining the stochastic PU. Removing these PU jets has a critical importance
1406 [102, 103]. *Jet Vertex Tagger* (JVT) and *forward Jet Vertex Tagger* (fJVT) are two algorithms used
1407 to separate hard-scatter (HS) and pile-up jets [104, 105]. In addition to pile-up jets, identifying the
1408 initial nature of the jets is needed for specific analyses such as the $HH \rightarrow b\bar{b}\gamma\gamma$ which require their
1409 jets to be originating from the b -quark fragmentation instead of being from gluons or other quarks. For
1410 that, specific b -tagging algorithms are designed to discriminate between b -jets and other jets [106]. This
1411 section includes details on each tagging algorithm.

5.3.1 Jet Vertex Tagging (JVT) algorithms

5.3.1.1 JVT

Jet can be matched to its vertex origin only within $|\eta| < 2.5$ where the tracking information is available. Information from this matching is the main ingredient of the JVT algorithm [107]. A global JVT discriminant is constructed from two track-based variables using k -nearest-neighbour algorithm [108]. The two variables are defined as:

$$\text{corrJVF} = \frac{\sum_k p_T^{\text{track}_k} (\text{PV}_0)}{\sum_l p_T^{\text{track}_l} (\text{PV}_0) + \frac{p_T^{\text{PU}}}{k.n_{\text{tracks}}^{\text{PV}}}}, \quad (5.4)$$

$$R_{p_T} = \frac{\sum_k p_T^{\text{track}_k} (\text{PV}_0)}{p_T^{\text{jet}}}, \quad (5.5)$$

1414 The $\sum_k p_T^{\text{track}_k} (\text{PV}_0)$ term, appearing in both formulas, represents the scalar sum of tracks p_T originating
1415 from the hard scattered vertex and associated to a jet, corrJVF also uses a similar term for pile-up
1416 vertices $p_T^{\text{PU}} = \sum_{n \geq 1} \sum_l p_T^{\text{track}_l} (\text{PV}_n)$ where pile-up vertices are noted as PV_n , $\frac{1}{k.n_{\text{tracks}}^{\text{PU}}}$ where $k = 0.01$
1417 is a correction factor used to correct the behaviour from linear increase with the number of tracks. For
1418 both EMTopo and PFlow the JVT is built similarly, while WPs are defined differently for the two types

1419 of jets given the fact that PFlow has a built-in pile-up jet suppression in its reconstruction algorithm.
 1420 For EMTopo three WPs are defined (**Loose**, **Medium** and **Tight**) and only two WPs (**Medium** and **Tight**)
 1421 are defined for PFlow jets. Figure 5.5 shows the JVT reconstructed variable as well as the fake rate
 efficiency for JVF, R_{p_T} , corrected JVF and the combined JVT discriminate variables.

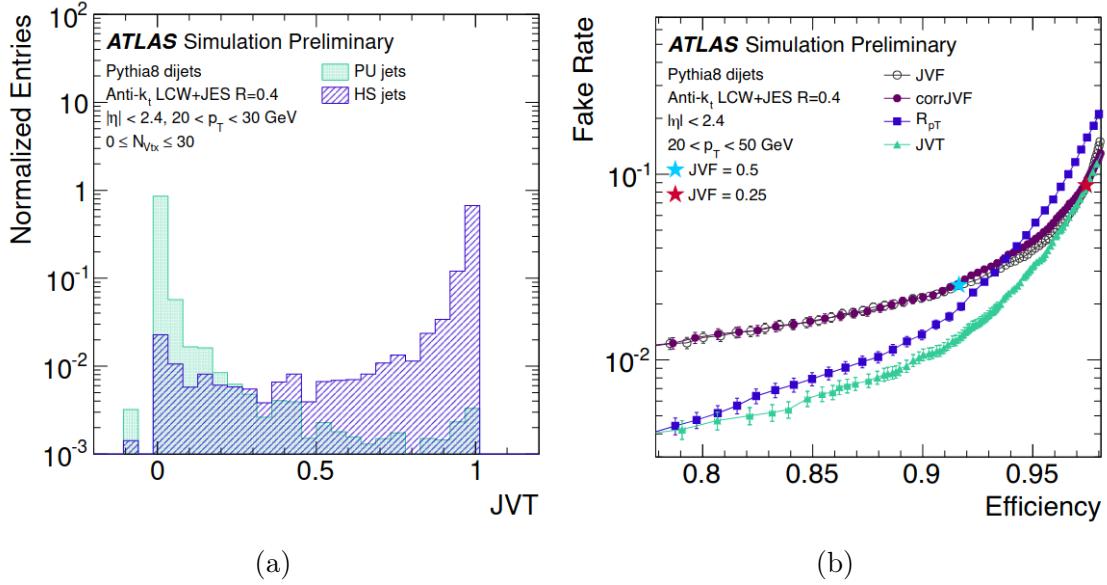


Figure 5.5: (a) Comparative distribution, scaled to unity, of JVT for truth-matched HS and pile-up.
 (b) pile-up jet fake rate as a function of the HS-jet efficiency for cuts on the JVT discriminant and JVT
 input variables. The performance is compared to that of the JVT inputs, corrJVF and R_{p_T} , as well as
 that of the no corrected JVF.

1422

1423 5.3.1.2 fJVT

For jets outside the ID acceptance ($|\eta| > 2.5$), the JVT can not be used due to the absence of tracking information. The fJVT algorithm is defined for this forward region. It starts by selecting central QCD pile-up jets and associate them with the corresponding vertex. Then for each vertex i , the missing transverse momentum $-\vec{p}_{T,i}$ is computed, averaging the jet and track component, as:

$$-\vec{p}_{T,i} = \frac{1}{2} \left(k \cdot \sum_{tracks \in PV_i} \vec{p}_T^{track} + \sum_{jets \in PV_i} \vec{p}_T^{jet} \right) \quad (5.6)$$

where $k = 2.5$ is a factor to account for intrinsic differences between the jet and tracks terms. The fJVT discriminant is therefore for a given jet is defined as:

$$fJVT = \max(fJVT_i) = \max\left(\frac{-\vec{p}_{T,i} \cdot \vec{p}_T^{jet}}{\vec{p}_T^{jet} \cdot \vec{p}_T^{jet}}\right) \quad (5.7)$$

1424 where $fJVT_i$ is the jet fJVT to the vertex i . Figure 5.6 shows the fJVT discriminant which tends to
 1425 have large values for QCD pile-up jets, while the distribution for hard-scatter jets falls steeply.
 1426 It is important to note that the fJVT discriminant is only clearly defined for QCD pile-up jets, as opposed
 1427 to stochastic pile-up jets. To be able to select stochastic jets, fJVT is accompanied by an additional
 1428 requirement on jet timing. Two fJVT WPs are defined: **Loose** which corresponds to $fJVT < 0.5$ and

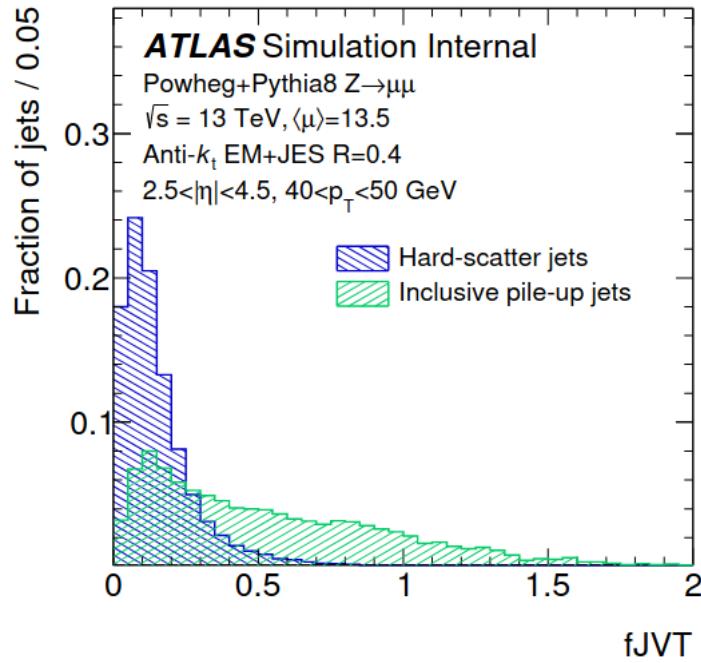


Figure 5.6: The fJVT distribution for hard-scatter (blue) and pile-up (green) forward jets.

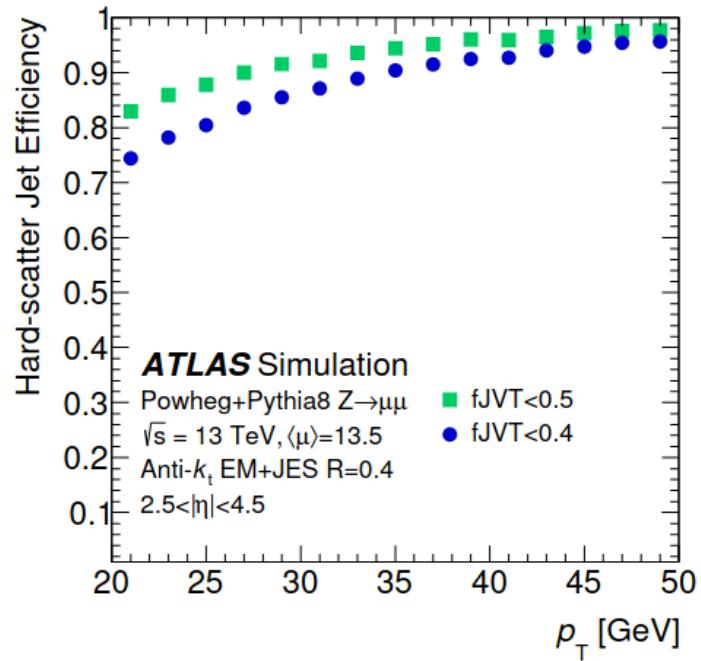


Figure 5.7: Efficiency for hard-scatter jets as a function of the forward jet p_T .

¹⁴²⁹ Tight with $f_{\text{JVT}} < 0.4$. Their efficiencies are shown in Figure 5.7. An additional improvement in the
¹⁴³⁰ f_{JVT} performances is achieved by combining f_{JVT} , timing, and jets shapes and kinematics information
¹⁴³¹ using machine learning techniques [109].

¹⁴³² 5.3.2 b -jet tagging

¹⁴³³ The identification of jets originating from b -quark decay, called b -tagging, is important for all analyses
¹⁴³⁴ involving b -quark in their final state, such as the $HH \rightarrow b\bar{b}\gamma\gamma$ analysis presented in this thesis. The
¹⁴³⁵ b -jet identification relies on identifying a secondary displaced vertex given the large lifetime of b -hadrons
¹⁴³⁶ $\sim 491.1 \mu\text{m}$ compared to c -jet or gluon- (light) jets. b -tagging algorithms are designed to discriminate
¹⁴³⁷ between b -jets, c -jet and other types. They are based on several low-level algorithms that exploit in-
¹⁴³⁸ formation from reconstructed tracks in the ID and MS which are combined by a high-level b -tagging
¹⁴³⁹ algorithm into a single discriminant variable used at the analysis level. Two high-level b -tagging algo-
¹⁴⁴⁰ rithms are developed MV2 which is based on the BDT, and the DL1 based on NNs. Both algorithms
¹⁴⁴¹ use almost the same input from low-level algorithms except for the DL1 that includes additional inputs
¹⁴⁴² related to the semi-leptonic decay of heavy-flavoured hadrons topology. Each low-level flavour tagging
¹⁴⁴³ algorithm focuses on different flavour-sensitive information, deriving different sets of variables based on
¹⁴⁴⁴ the impact parameter algorithms of tracks in jets, the presence and properties of an inclusive secondary
¹⁴⁴⁵ vertex and the reconstruction of the b - and c -hadron multi-vertices chain topology:

- ¹⁴⁴⁶ • Impact Parameter (IP) based algorithms: use a log-likelihood ratio (LLR) to separate tracks
¹⁴⁴⁷ associated with jets according to whether or not they are compatible with the primary vertex
¹⁴⁴⁸ hypothesis. The LLR is computed as the sum of per-track contributions $\sum_{i=1}^N \log(p_b/p_u)$ where
¹⁴⁴⁹ N is the number of tracks and p_b , p_u are the pdfs for the b - and light-jet hypotheses respectively.
¹⁴⁵⁰ The algorithms do not take into account tracks correlation [110, 111].
- ¹⁴⁵¹ • Secondary Vertex (SV) finding: explicitly reconstructs an inclusive displaced secondary vertex
¹⁴⁵² within the jet. All track pairs within a jet are tested for a two-track vertex hypothesis using a
¹⁴⁵³ χ^2 . Two-track vertices are discarded if they are likely to originate from the decay of a long-lived
¹⁴⁵⁴ particle, photon conversions or hadronic interactions with the detector material [112].
- ¹⁴⁵⁵ • Decay Chain Multi-Vertex algorithm (JetFitter): exploits the topological structure of weak b - and
¹⁴⁵⁶ c -hadron decays inside the jet and tries to reconstruct the full b -hadron decay chain. A Kalman
¹⁴⁵⁷ filter is used to find a common line on which the primary vertex and the bottom and charm vertices
¹⁴⁵⁸ lie, approximating the b -hadron flight path, as well as their positions [113].

Outputs from these algorithms construct input to MV2 and DL1. Many variations of DL1 exist with different NN configurations and additional inputs. The DL1r adopted in the work of this thesis as well as ATLAS publications also includes input from a Recurrent Neural Network (RNN) based IP algorithm to consider the correlation among tracks [114]. The final discriminant for b -tagging efficiency is calculated as:

$$\text{DL1} = \ln \left(\frac{p_b}{f_{c-\text{jets}} \cdot p_c + (1 - f_{c-\text{jets}}) \cdot p_{\text{light-flavour}}} \right), \quad (5.8)$$

¹⁴⁵⁹ where p_b , p_c and $p_{\text{light-flavour}}$ are the NN outputs probabilities that the jet is a b -jet, c -jet or light-jet
¹⁴⁶⁰ respectively. $f_{c-\text{jets}}$ is the c -jet fraction taken into account, which can be tuned separately according to
¹⁴⁶¹ physics analysis needs.

¹⁴⁶² For both MV2 and DL1, four WPs are computed for different b -tagging efficiencies. The WPs are defined
¹⁴⁶³ as a fixed cut or p_T dependent cut on the final discriminant variable (Eq. 5.8) and labelled with the

¹⁴⁶⁴ corresponding integrated efficiency on b -jets. The working points are 85%, 77%, 70% and 60%. Figure
¹⁴⁶⁵ 5.8 shows a comparison of performances of MV2 and DL1 built in 2018, and DL1r optimized in 2019.
¹⁴⁶⁶ The 2019 DL1r shows a significant improvement over other algorithms. Figure 5.9 shows the efficiency
¹⁴⁶⁷ of identifying b -jets as a function of jet p_T for 77% WP requirement, for 2018 recommended versions of
MV2 and DL1, and 2019 DL1r optimization [115].

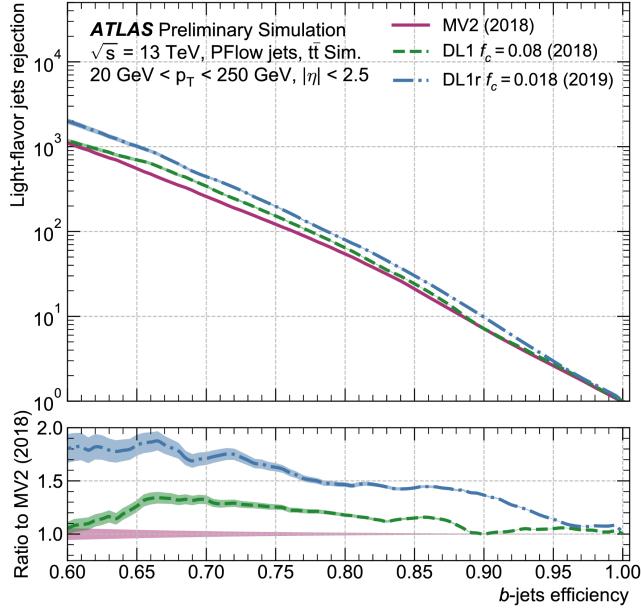


Figure 5.8: ROC curves for 2018 recommended versions of MV2 and DL1, and 2019 DL1r optimization. The x-axis corresponds to the b -jet efficiency, while the y-axis corresponds to light-flavor jets rejection. The shaded bands represent the statistical uncertainty.

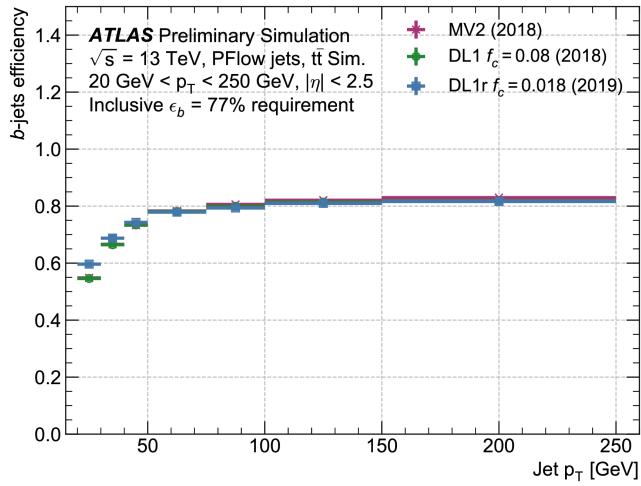


Figure 5.9: Efficiency of identifying b -jets as a function of jet p_T for an inclusive 77% efficiency requirement, for 2018 recommended versions of MV2 and DL1, and 2019 DL1r optimization.

5.4 b -jet energy calibration

The default JES calibration described in Section 5.2.1.2 and applied to all jets is a correction optimized on multi-jet events which are mainly light-quark jets. This correction is not optimal for b -jets due to:

- large b -quark mass which reduces the boost of the jet and, as a consequence, a larger fraction of the jet particles lay outside the usual jet cone leading to out-of-cone radiation.
- larger semi-leptonic branching ratio of b -hadrons which produce elusive neutrino.

Figure 5.10 illustrates these statements through the schematic of a b -jet.

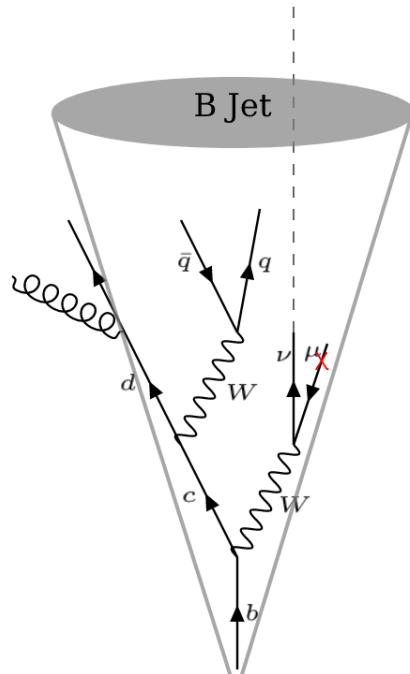


Figure 5.10: Typical schema of a b -jet.

To overcome this issue, a step-by-step approach, firstly developed in $VH(\rightarrow b\bar{b})$ analysis [116], is used:

- The first correction called " μ -in-jet" adds soft muon in the jet to address the semi-leptonic decay of b -hadrons inside b -jet.
- The second correction called " p_T -Reco" corrects b -jets energy for the neutrino presence in the semi-leptonic decay and out-of-cone radiation effect by introducing a global correction factor to correct the jet transverse momenta.

The developed approach in $VH(b\bar{b})$ was computed only for EMTopo b -jets from the $VH(b\bar{b})$ simulated events. To generalize this correction to any channel with b -jets its amplitude is computed from b -jets produced in $t\bar{t}$ simulated events for both EMTopo and PFlow jets. This correction is added as an additional step to the calibration chain already introduced (see Figure 5.11).

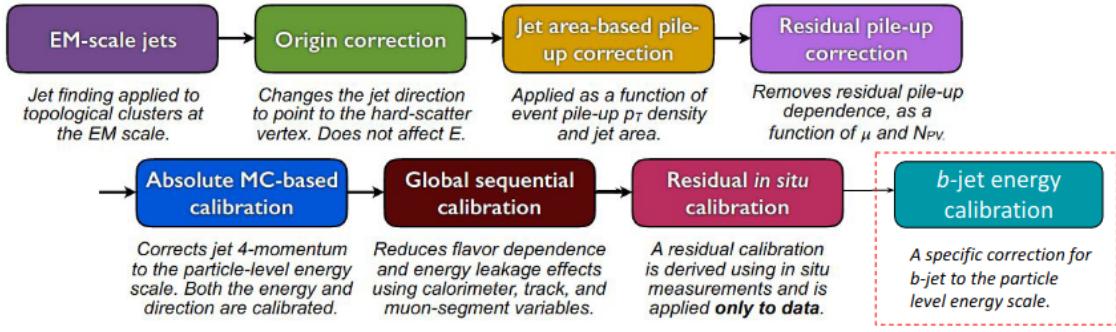


Figure 5.11: Calibration stages for jets including energy correction for tagged b -jet.

1487 5.4.1 μ -in-jet correction

The decays of B and D hadrons produce each type of the three leptons with a branching ratio of $\sim 11\%$. These leptons are usually mixed with other particles within jets. For electrons and taus, their identification efficiency is low and, as a consequence, will not be flagged as semi-leptonic and their energies are fully included in the jet reconstruction as mentioned in Section 5.1. The energy coming from their associated neutrinos will be mixed with p_T -Reco. Instead, the muon is a minimum ionizing particle, depositing very little energy in the calorimeter. To correct for the semi-leptonic decay, the four-vector momenta of a muon found inside the jet is added to the four-vector momenta of the jet. To select the associate muon to the jet, selection criteria listed in Table 5.1, were defined.

Criteria	Selection
p_T	$> 4 \text{ GeV}$
ID	Medium
ΔR	$< \text{VR}$

Table 5.1: Muon selection requested for μ -in-jet correction

ΔR represents the distance between jet and muon axis. VR is computed using the formula:

$$\text{VR} = \min(0.4, 0.04 + 10/p_T) \quad (5.9)$$

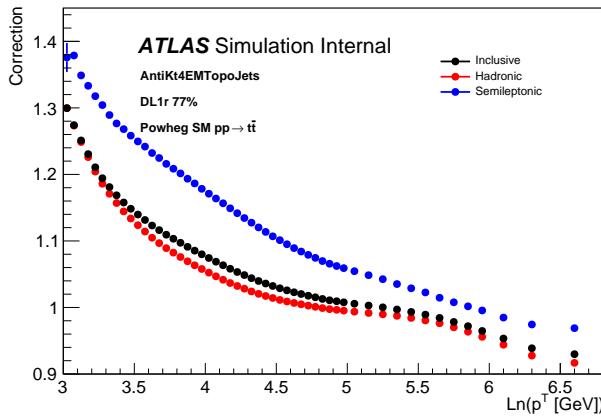
1488 with muon p_T in GeV. Instead of a 0.4 fixed cut on ΔR , a variable cut is applied which takes into
 1489 account the fact that jet boost reduces the angular distance of the muon. If more than one muon is
 1490 identified, only the highest transverse momentum (closest) is used. Such muons are found for 12% of
 1491 the b-tagged jets. The 4-vector momenta associated with the muon passing the selection is added back
 1492 to jet one after having subtracted the estimated energy deposited in the calorimeter. The jets including
 1493 a muon are labelled as "semi-leptonic" while others are labelled "hadronic".

1494 5.4.2 p_T -Reco correction

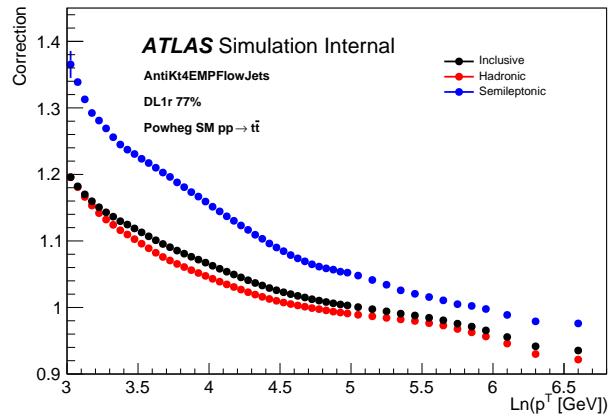
1495 On top of μ -in-jet correction, a global p_T dependent scale factor is applied to the jet 4-vector to correct
 1496 for the presence of neutrino in semi-leptonic decays and out-of-cone radiation. The relevant target is
 1497 the truth-level TruthWZ jets container, which contains reconstructed jets with the same reconstruction

1498 algorithm (Antik_t), using all stable hadrons and non-isolated muons and neutrinos. In order to be used by
 1499 any analysis with b -jet in the final state, the correction factor is derived from a $t\bar{t}$ sample and benefit from
 1500 large p_T spectrum. It is binned in $\ln(p_T)$ of reconstructed jet up to 400 GeV. The log scale is used because
 1501 of the exponential shape of the p_T spectrum and to provide a smooth distribution for the interpolation
 1502 step. The scale factor is computed separately for semi-leptonic and hadronic jets and evaluated as the
 1503 mean of the distribution of the ratio between the reconstructed b -jet p_T and the corresponding **TruthWZ**
 1504 jet p_T , matched within $\Delta R < 0.3$ for the given p_T bin. For each b -tagged jet, the 4-vector is scaled
 1505 by the scale factor computed using a linear interpolation by `TH1::Interpolate(ln(pT))` function.
 1506 The correction is derived for each jet reconstruction algorithms EMTopo or PFlow and each b -tagging
 1507 algorithms MV2 or DL1r.

1508 Figure 5.12 shows the correction factor distribution for different jet labels and reconstruction algorithms.



(a) topological cluster jets



(b) particle flow jets

Figure 5.12: p_T -Reco correction distribution for (a) topological cluster jets (b) particle flow jets,
 b -tagging 77% WP is applied with DL1r algorithm.

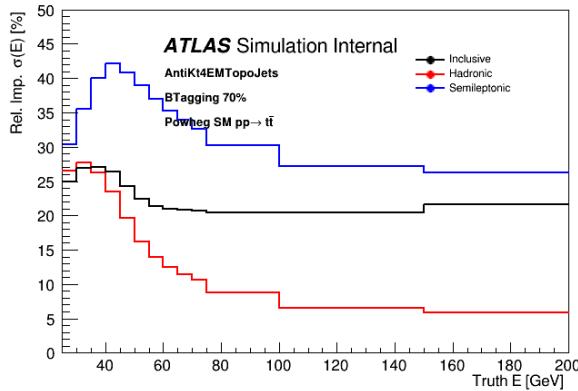
1509

1510 Figure 5.12 shows that :

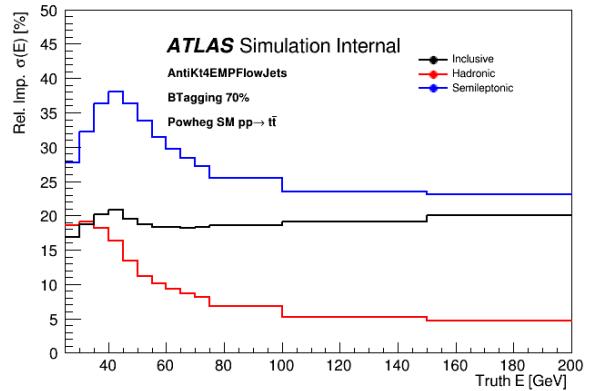
- 1511 • EMTopo hadronic correction factor is $\sim 10\%$ higher than the one for PFlow at low jet p_T . This is
 1512 explained by the fact that PFlow combines tracking, soft activity and calorimetric information to
 1513 reconstruct the jet, so it includes low p_T ($p_T > 0.5$ GeV) tracks that do not reach the calorimeter,
 1514 which improves the jet resolution at low p_T , as mentioned in Section 5.1.
- 1515 • The presence of neutrino in semi-leptonic jets explains the higher correction compared to hadronic
 1516 jets (up to $\sim 10\%$ higher for low p_T jet).
- 1517 • At high jet p_T , the distribution of the ratio between the reconstructed p_T and the truth one has a
 1518 tail below unity indicating that the reconstructed jet p_T is higher than the associated truth. This
 1519 can be explained by the fact that at truth-level, some boosted b -quark gives two closed jets where
 1520 only one is identified as b -jet. At the reconstructed-level, the b -quark shower is reconstructed as a
 1521 single b -jet.

5.4.3 b -jet energy correction improvement

Figure 5.13, shows the maximum possible gain in b -jet energy resolution in $t\bar{t}$ sample (used to compute the p_T -Reco correction). The energy resolution is improved by 40% (20%) for low (high) energy for semi-leptonic jet, and about 20% (5%) for low (high) energy for hadronic jet. Similar improvement is observed for both jet types (EMTopo and PFlow).

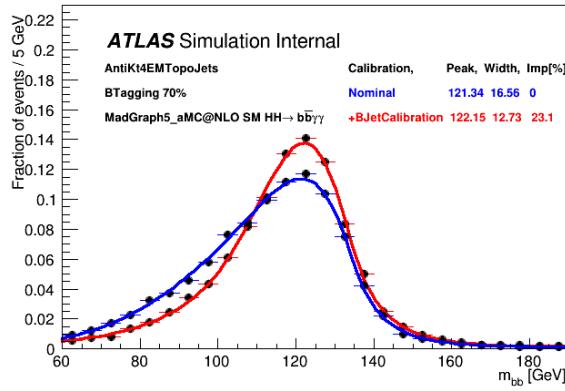


(a) topological cluster jets

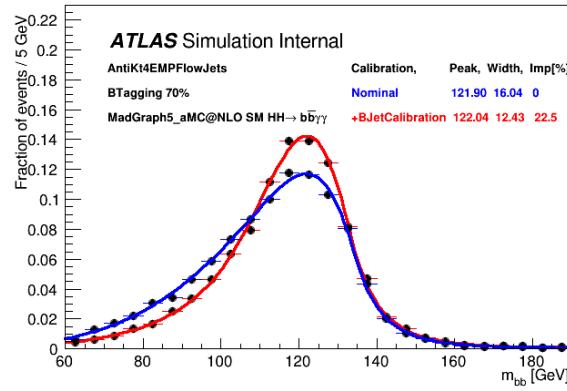


(b) particle flow jets

Figure 5.13: Relative energy resolution improvement after b -jet energy calibration in bin of truth energy for $t\bar{t}$ using (a) topological cluster jets (b) particle flow jets.



(a) topological cluster jets



(b) particle flow jets

Figure 5.14: Distribution of m_{bb} before calibration (blue) and after calibration (red) in $HH \rightarrow bb\gamma\gamma$ for (a) topological cluster jets (b) particle flow jets.

5.4.3.1 Improvement on $H \rightarrow \bar{b}b$

The $HH \rightarrow b\bar{b}\gamma\gamma$ channel benefits from the $H \rightarrow \gamma\gamma$ part which has a very narrow peak for $m_{\gamma\gamma}$ helping in rejection of a huge amount of the continuum background by defining a narrow window of $m_{\gamma\gamma}$.

spectrum to select the Higgs boson signal. The $H \rightarrow b\bar{b}$ is suffering from the b -jet miscalibration and has large $m_{b\bar{b}}$ resolution compared to $m_{\gamma\gamma}$. Any gain on $m_{b\bar{b}}$ should be beneficial for background rejection. The gain on $m_{b\bar{b}}$ resolution from the decoupled method is quantified on $HH \rightarrow b\bar{b}\gamma\gamma$. The gain in b -jet energy resolution reported before is translated into $\sim 22\%$ improvement in the $m_{b\bar{b}}$ invariant mass resolution, as Figure 5.14 shows. In addition, the $m_{b\bar{b}}$ peak position get closer to the Higgs boson mass. Similar improvements are obtained in both topological and particle flow jets.
 The two steps μ -in-jet and p_T -Reco are combined in a decoupled method denoted b -jet energy calibration under a simple **BJetCalibration** tool. The tool is documented and public to be used by analysis with b -jet in their final state. It has been adopted as the standard b -jet calibration across all di-Higgs boson channels including b -jets ($HH \rightarrow b\bar{b}\gamma\gamma$, $HH \rightarrow b\bar{b}b\bar{b}$, $HH \rightarrow b\bar{b}\tau\tau$, $HH \rightarrow b\bar{b}ll + MET$ and $HH \rightarrow b\bar{b}VV$).

5.5 Conclusion

The related features to calibrate and identify jets and b -jets have direct impact on the $HH \rightarrow b\bar{b}\gamma\gamma$ selection. Reconstruction algorithms, energy calibration, b -tagging and the specific b -jet calibration were introduced in this chapter. A significant improvement is achieved on the $m_{b\bar{b}}$ resolution which leads to a better signal versus background separation.
 The developed method shows similar performance to an alternative approach using a simplified NNs regression to calibrate the b -jet energy trained on $t\bar{t}$ events. The simplified NNs regression uses similar requested as the decoupled method, while the full NNs regression includes additional information about the tracks and secondary vertices (~ 90 variables) and achieves a resolution of ~ 10 GeV in the $m_{b\bar{b}}$. The huge number of inputs needed by the full NNs regression makes its implementation at the analysis level not trivial. As a consequence, its usage was postponed to a future Run-3 analysis.

1552 Chapter 6

1553 Measurement of Higgs boson self-coupling

1554 The central subject of this thesis is to search for di-Higgs boson events decaying into $HH \rightarrow b\bar{b}\gamma\gamma$ using
1555 the full Run-2 data sample. The absence of a significant number of signal events is translated into
1556 a measurement of the Higgs boson self-coupling. All the ingredients needed to select and reconstruct
1557 the $HH \rightarrow b\bar{b}\gamma\gamma$ events were presented in the previous chapters. This chapter describes the analysis
1558 strategy to measure the Higgs boson self-coupling, the results and their comparison to the previous
1559 ATLAS publication with 36 fb^{-1} of data and the CMS full Run-2 results.

1560 6.1 Data and simulation

1561 As mentioned before, the analysis uses the full Run-2 dataset recorded by the ATLAS experiment be-
1562 tween 2015 and 2018 corresponding to a total integrated luminosity of $139 \pm 2.4 \text{ fb}^{-1}$ (Figure 3.3). The
1563 values of the integrated luminosity recorded in each year are shown in Table 6.1 [117].

Year	int. luminosity [fb^{-1}]
2015-2016	36.21
2017	44.39
2018	58.45

Table 6.1: Integrated luminosity recorded by ATLAS in Run 2 per year of data taking.

1564 Similarly to the $H \rightarrow \gamma\gamma$ analysis, the $HH \rightarrow b\bar{b}\gamma\gamma$ analysis relies on different sets of di-photon triggers
1565 for each year. The trigger used for 2015-2016 requires at least two photons with E_T greater than 35
1566 GeV for the leading and 25 for the sub-leading photon, both passing the loose identification criterion.
1567 The di-photon trigger used in 2017-2018 requires at least two photons with the same requirement on
1568 their transverse energies but passing the medium identification criterion to reduce fake rates. Figure 6.1
1569 shows the di-photon trigger efficiency for each year as a function of photon E_T and η . Slightly lower
1570 efficiencies are observed in 2017-2018 due to the tightening in photon identification. In addition, no
1571 significant dependence on η is observed in the trigger efficiency, which remains close to 100%.

1572 For boosted Higgs bosons ($p_T^{Higgs} > 650 \text{ GeV}$), two high momentum single-photon triggers available in
1573 2018 were added in a logical OR with the di-photon trigger, requiring a loose photon with an energy
1574 of 120 GeV and 140 GeV respectively. The high momentum single-photon triggers have no impact on
1575 the HH analysis, since only 0.1% of signal events pass this trigger. The efficiency of the trigger selec-

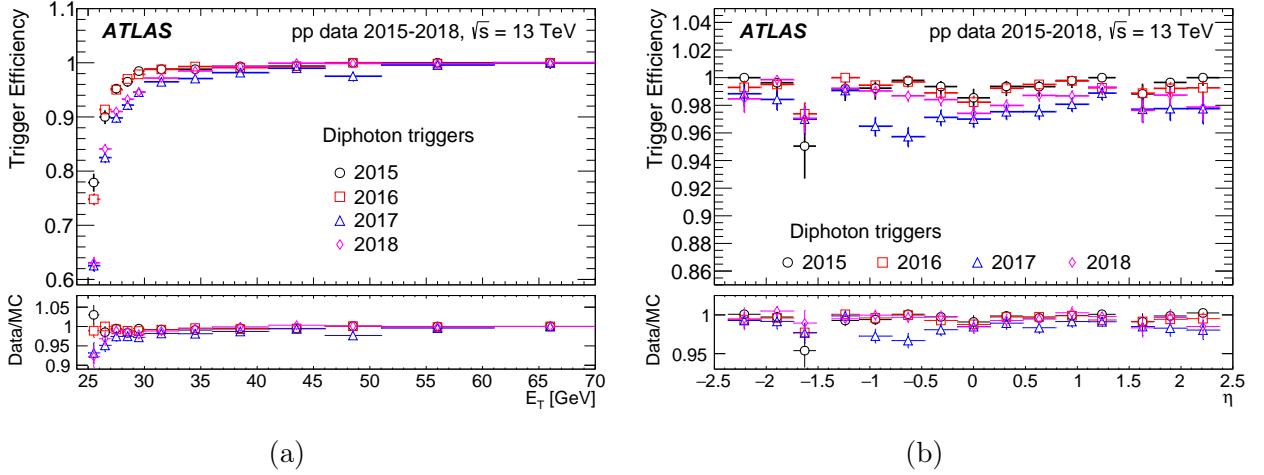


Figure 6.1: Di-photon trigger efficiency in data 2015-2018 as a function of (a) E_T and (b) η . The ratios of data to MC simulation efficiencies are also shown.

1577 tition is $\sim 83\%$ for the HH signal, and the selection criteria remains looser than the final analysis selection.

1578

1579 The simulated $HH \rightarrow b\bar{b}\gamma\gamma$ sample (MC) are generated for two production modes: the ggF HH
 1580 and VBF HH productions described in Section 2.5.1. Events from ggF HH production were gener-
 1581 ated at NLO using the POWHEG-BOX v2 generator in the finite top-quark mass approximation with
 1582 PDFLHC15 PDF set [118, 119, 120]. The PYTHIA 8.2 code is used for parton showering. Two samples
 1583 were generated for $\kappa_\lambda = 1$ and $\kappa_\lambda = 10$. Events from the VBF HH production were generated at LO
 1584 using MADGRAPH5_AMC@NLO [121] with the NNPDF3.0nnlo PDF set [122] and PYTHIA 8.2 for
 1585 showering.

1586

A reweighting procedure was developed for both ggF and VBF to avoid generating simulation samples for each κ_λ . Truth level HH samples with 10 million events were produced for $\kappa_\lambda = 0, 1, 10$ and 20 with specific HH decay modes. A linear combination of samples with $\kappa_\lambda = 0, 1$ and 20 derived from Eq. 2.26 is implemented to derive the sample at a given κ_λ value. The method was validated for $\kappa_\lambda = 10$. The total cross-section as a function of κ_λ and κ_t is parameterized as:

$$|A(\kappa_t, \kappa_\lambda)|^2 = \kappa_t^2 \left[\left(\kappa_t^2 + \frac{\kappa_\lambda^2}{20} - \frac{399}{380} \kappa_t \kappa_\lambda \right) |A(1, 0)|^2 + \left(\frac{40}{38} \kappa_t \kappa_\lambda - \frac{2}{38} \kappa_\lambda^2 \right) |A(1, 1)|^2 + \frac{\kappa_\lambda^2 - \kappa_t \kappa_\lambda}{380} |A(1, 20)|^2 \right], \quad (6.1)$$

1587 with weights derived by dividing the binned m_{HH} distribution of the target κ_λ by the binned m_{HH}
 1588 distribution of the SM sample. A bin width of 10 GeV is adopted and $\kappa_t = 1$. For each κ_λ value, the
 1589 inclusive cross-section is normalized to its theoretical prediction [123]. The reweighting procedure is
 1590 common across all HH analyses and has been validated for $HH \rightarrow b\bar{b}\gamma\gamma$ by comparing the $m_{\gamma\gamma}$ and
 1591 event yields in the generated $\kappa_\lambda = 10$ and the reweighted $\kappa_\lambda = 1$ to $\kappa_\lambda = 10$. A good agreement is found
 1592 with a maximum discrepancy of 3-4% taken as systematic uncertainty related to the method.

1593

1594 For background processes, different generators were used depending on the process. Only the $H \rightarrow \gamma\gamma$
 1595 decay channel is considered for single Higgs boson backgrounds. An alternative showering with HERWIG7
 1596 for the single Higgs boson backgrounds is used to estimate the systematic uncertainties of the parton
 1597 shower. Table 6.2 lists the generators, parton showering and PDF used for each component.

1598 The dominant background is the reducible continuum $\gamma\gamma + \text{jets}$ process which is not evaluated from the

Process	Generator	PDF set	Showering
ggF	NNLOPS	PDFLHC	PYTHIA 8.2
VBF	POWHEG-BOX v2	PDFLHC	PYTHIA 8.2
WH	POWHEG-BOX v2	PDFLHC	PYTHIA 8.2
$q\bar{q} \rightarrow ZH$	POWHEG-BOX v2	PDFLHC	PYTHIA 8.2
$gg \rightarrow ZH$	POWHEG-BOX v2	PDFLHC	PYTHIA 8.2
$t\bar{t}H$	POWHEG-BOX v2	NNPDF2.3lo	PYTHIA 8.2
bbH	POWHEG-BOX v2	NNPDF3.0nnlo	PYTHIA 8.2
$tHqj$	MADGRAPH5_AMC@NLO	NNPDF3.0nnlo	PYTHIA 8.2
tHW	MADGRAPH5_AMC@NLO	NNPDF3.0nnlo	PYTHIA 8.2
$\gamma\gamma + \text{jets}$	SHERPA v2.2.4	NNPDF3.0nnlo	SHERPA v2.2.4
$t\bar{t}\gamma\gamma$	MADGRAPH5_AMC@NLO	NNPDF2.3lo	PYTHIA 8.2

Table 6.2: Summary of single Higgs boson background samples, split by production modes, and continuum background samples. The generators used, and the PDF sets are also listed.

1599 generated SHERPA sample. Instead a data-driven method is implemented to estimate its contribution
1600 by parametrizing its shape with an analytical function fit, which its form is fixed using the simulated
1601 SHERPA sample.

6.2 Object and event selection

6.2.1 Object selection

6.2.1.1 Di-photons

1605 The selected photons are reconstructed using the dynamical topological clusters method, as defined in
1606 Section 3.4.2, and must fulfil a pre-selection of $p_T > 25$ GeV and $|\eta| < 2.37$ (excluding the transition
1607 region). They are calibrated using the latest Run-2 calibration correction for the energy scale and res-
1608 olution as detailed in Section 3.4.2.1. Photons are required to pass the tight identification as defined
1609 in Section 4.2. Events must contain at least two photons ($N_{\text{photons}} \geq 2$). The $H \rightarrow \gamma\gamma$ candidate is
1610 reconstructed from the two highest p_T photons in the event.

1611 The two photons are used to re-determine the primary vertex using an algorithm based on NNs that
1612 makes use of the pointing information from the electromagnetic calorimeter along with tracking infor-
1613 mation [124]. Track-based isolation quantities are re-computed based on the defined primary vertex.

1614 A loose isolation criterion is applied to select isolated photons as described in Section 4.1.

1615 The selected Higgs bosons candidates must pass an additional selection which reflects the online trigger.
1616 It requires that the leading (subleading) photon p_T is at least 35% (25%) of the invariant mass of the
1617 two photons. A mass window cut in the range [105, 160] GeV is also applied, thus the cut on the leading
1618 (subleading) photon momentum fraction translates in a cut on its energy to be larger than at least 36.75
1619 (26.25) GeV.

6.2.1.2 Leptons as veto

1621 Electron candidates are required to have $p_T > 10$ GeV and $|\eta| < 2.47$ (excluding the transition region).
1622 Additionally, they should pass the medium identification and tight isolation as defined in Sections 3.4.2.3

1623 and 3.4.2.2.

1624 Muon candidates are required to have $p_T > 10$ GeV, $|\eta| < 2.7$ and required to satisfy a medium
 1625 identification and loose isolation. Both the electrons and muons are matched to the primary vertex via
 1626 requirements on the track longitudinal and transverse impact parameters, $|z_0|$ and $|d_0|$.

1627 6.2.1.3 Jets

1628 As described in Chapter 5, particle flow jets are adopted for this analysis. They are required to have
 1629 $|\eta| < 2.5$ and $p_T > 25$ GeV. A tight JVT is applied to separate jets arising from the hard-scatter and
 1630 those from pile-up, as defined in Section 5.3. The selected jets are calibrated using the calibration chain
 1631 described in Section 5.2.

1632 The jet flavour is determined using the 77% efficiency WP of the DL1r tagger as described in Section
 1633 5.3.2. Events are required to have exactly two b -jets ($N_{b\text{-jet}}^{77\%} = 2$) to preserve the orthogonality with
 1634 the $HH \rightarrow b\bar{b}b\bar{b}$ channel which requires more than two b -jets. The two b -jets are used to reconstruct the
 1635 $H \rightarrow b\bar{b}$ candidate.

1636 The energy of b -tagged jets is corrected with the calibration method introduced in Section 5.4. The
 1637 improvement on the $m_{b\bar{b}}$ resolution translates to a $7.2\% \pm 2\%$ improvement in the expected signal sig-
 1638 nificance (the error is estimated using the bootstrap method [125]).

1639 An additional 5-10% improvement on the $m_{b\bar{b}}$ resolution could have been achieved through the use of
 1640 a likelihood-based kinematic fit denoted "kinematic fit". The kinematic fit uses the $H \rightarrow \gamma\gamma$ component
 1641 reconstructed with a percent-level precision to improve the $H \rightarrow b\bar{b}$ resolution through the constraint
 1642 of the good overall balance in the transverse plane. I developed the kinematic fit for this $HH \rightarrow b\bar{b}\gamma\gamma$
 1643 analysis. It is not used in the final analysis as a bug was found in its implementation after sample
 1644 production. Fixing the bug would have delayed the publication. A detailed description of the kinematic
 1645 fit can be found in Appendix D.

1647

1648 6.2.2 Event selection

1649 Events are reconstructed from $H \rightarrow b\bar{b}$ and $H \rightarrow \gamma\gamma$ defined above. In addition to the requirement on
 1650 the number of photons ($N_{\text{photons}} \geq 2$) and the b -jet veto ($N_{b\text{-jet}}^{77\%} = 2$), the following criteria are applied
 1651 to reduce the $t\bar{t}H$ background contamination:

- 1652 • less than six jets in the central detector region ($|\eta| < 2.5$) to reject $t\bar{t}H$ with the hadronic decays.
- 1653 • veto any event with at least one lepton to reject $t\bar{t}H$ with leptonic decays.

1654 Selected events are then divided in two regions using the $m_{b\bar{b}\gamma\gamma}^*$ variable, which is defined as $m_{b\bar{b}\gamma\gamma}^* =$
 1655 $m_{b\bar{b}\gamma\gamma} - m_{b\bar{b}} - m_{\gamma\gamma} + 250$ GeV. The $m_{b\bar{b}\gamma\gamma}^*$ was implemented to improve the m_{HH} resolution for the resonant
 1656 analysis. A high-mass region, with $m_{b\bar{b}\gamma\gamma}^* > 350$ GeV, targets the SM signal, while a low-mass region,
 1657 with $m_{b\bar{b}\gamma\gamma}^* < 350$ GeV, is used to retain sensitivity for BSM signals and provide better constraints on
 1658 κ_λ are defined. The distribution of $m_{b\bar{b}\gamma\gamma}^*$ is shown in Figure 6.2.

1659 In each mass region, a separate BDT is trained using XGBoost [126] to categorize the ggF HH signal
 1660 against a combination of the dominant backgrounds (continuum $\gamma\gamma + \text{jets}$, $t\bar{t}H$, ggF and ZH). In the high-
 1661 mass region, the SM $\kappa_\lambda = 1$ process is considered as the ggF HH signal while in the region targeting
 1662 BSM scenarios (low-mass) the $\kappa_\lambda = 10$ process ggF HH is used. Table 6.3 lists the BDT input variables
 1663 which are used for both the low and high-mass regions.

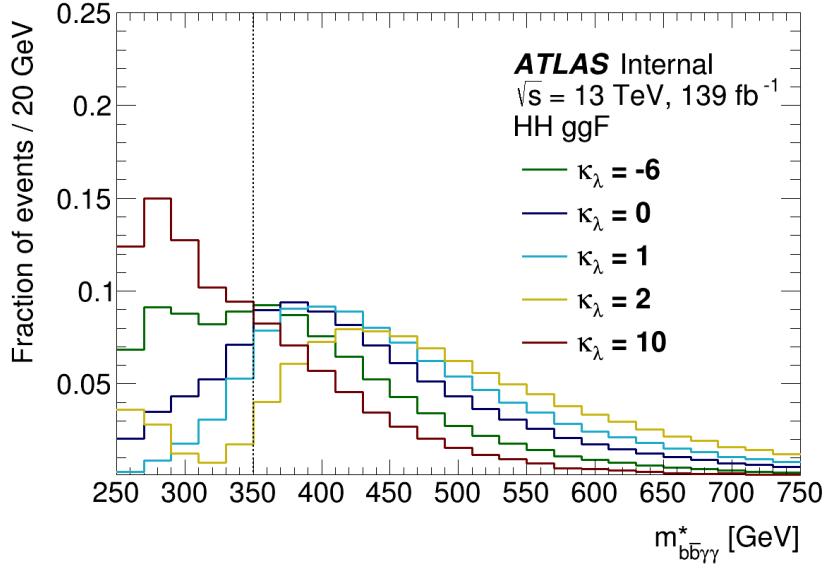


Figure 6.2: $m_{b\bar{b}\gamma\gamma}^*$ distributions for the ggF HH signal with several κ_λ values. $m_{b\bar{b}\gamma\gamma}^* = 350$ GeV is chosen as the separating boundary between categories targeting the SM and BSM κ_λ signals.

Variable	Definition
$p_T/m_{\gamma\gamma}$	p_T of the two photon system scaled by its invariant mass $m_{\gamma\gamma}$.
η and ϕ	Pseudo-rapidity and azimuthal angle of the two photon system.
b -tagging score	b -tagging score of the two jets.
η , ϕ and p_T	p_T , pseudo-rapidity and azimuthal angle of the two jets.
$m_{b\bar{b}}$	$H \rightarrow b\bar{b}$ invariant mass.
H_T	Scalar sum of the p_T of the jets in the event.
χ_{Wt}	Single topness defined in Equation 6.2.
E_T^{miss} and ϕ	Missing transverse momentum and its azimuthal angle.

Table 6.3: Input variables used in the BDT.

The "single-topness" variable is defined as:

$$\chi_{Wt} = \min \sqrt{\left(\frac{m_{j_1 j_2} - m_W}{m_W}\right)^2 + \left(\frac{m_{j_1 j_2 j_3} - m_t}{m_t}\right)^2}, \quad (6.2)$$

where $m_W = 80$ GeV, $m_t = 173$ GeV, and the minimum is taken over all possible combinations of 3 jets in the event. No requirement on the b -tagging is applied on the third jet.

The BDT discriminant distribution for the low mass and the high mass categories are shown in Figure 6.3. In each mass region, two categories are defined based on the BDT score resulting in total of four categories as listed in Table 6.4. Defining additional categories does not bring any significant improvement to the analysis sensitivity.

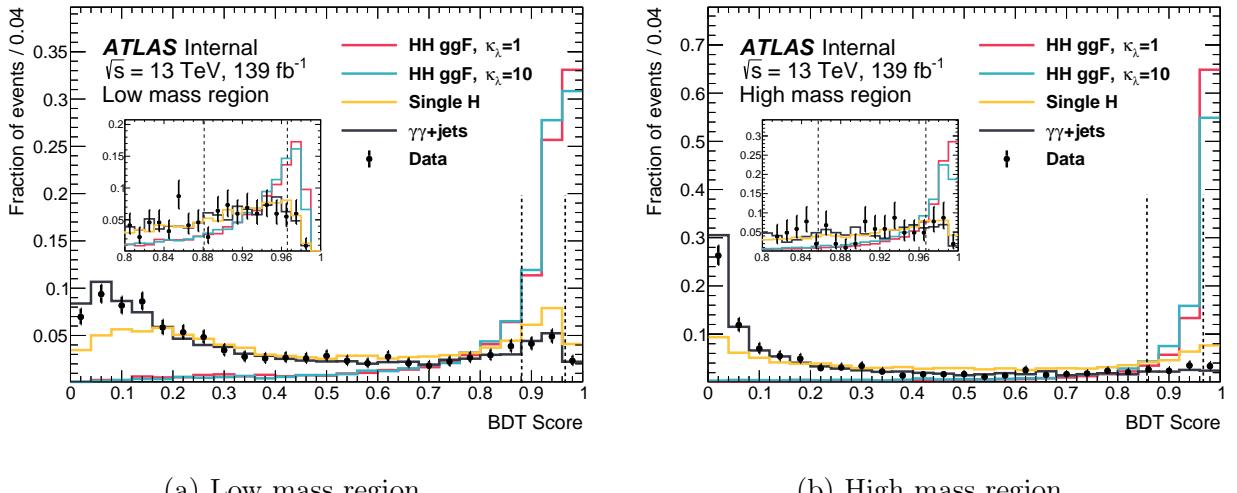


Figure 6.3: The BDT score for the benchmark signals and the main backgrounds in the low- (a) and high- (b) mass categories. Distributions are normalized to unit area. The dotted lines denote the category boundaries. Events with a BDT score below 0.881 in the low mass region or below 0.857 in the high mass region are rejected.

Category	Selection criteria
High mass BDT tight	$m_{b\bar{b}\gamma\gamma}^* \geq 350$ GeV, BDT score $\in [0.967, 1]$
High mass BDT loose	$m_{b\bar{b}\gamma\gamma}^* \geq 350$ GeV, BDT score $\in [0.857, 0.967]$
Low mass BDT tight	$m_{b\bar{b}\gamma\gamma}^* < 350$ GeV, BDT score $\in [0.966, 1]$
Low mass BDT loose	$m_{b\bar{b}\gamma\gamma}^* < 350$ GeV, BDT score $\in [0.881, 0.966]$

Table 6.4: Definition of the analysis categories.

The BDT approach improves the analysis significance by approximately 20% compared to the previous cut-based selection used in the 36 fb^{-1} analysis [24]. The $m_{\gamma\gamma}$ distribution in each of the BDT categories are shown in Figure 6.4.

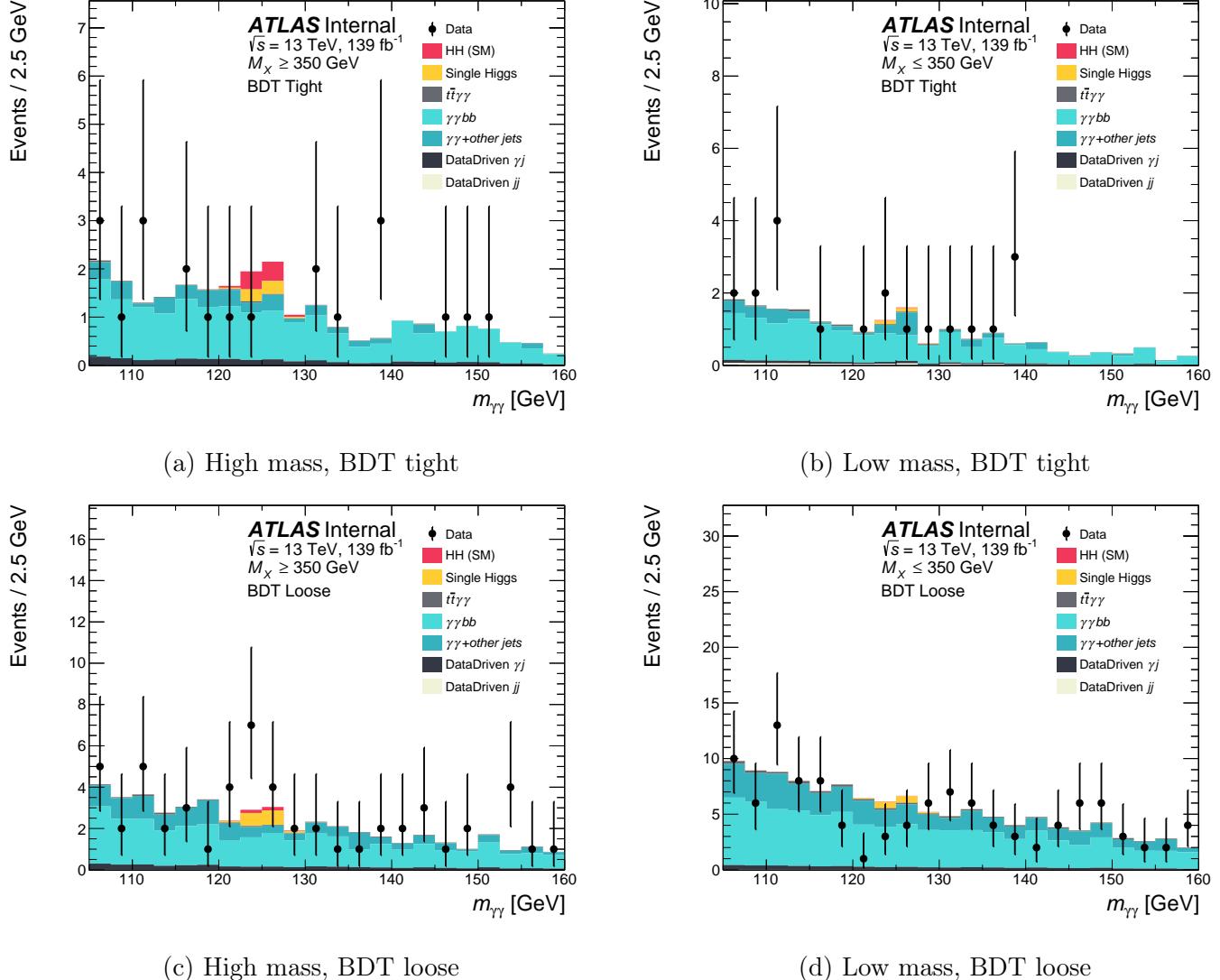


Figure 6.4: Distributions of $m_{\gamma\gamma}$ in all signal categories. The simulated continuum background is scaled by the $\gamma\gamma$, γ -jet/jet- γ , and jet-jet fractions and normalized to the data side-band ($m_{\gamma\gamma} \in [105, 120] \cup [130, 160]$ GeV). The background decomposition is described in Section 6.3.2.

1675 6.2.3 Event selection using Deep Neural Network

1676 The BDT categorization is used as baseline for the publication. In parallel, I developed a multi-label
 1677 classification based on a Deep Neural Network (DNN) to classify events, and shows similar performance
 1678 to the BDT.

1679 6.2.3.1 Multi-label classification DNN

1680 Following the same strategy as the BDT, a potential gain can be achieved using a multi-label classifica-
 1681 tion DNN. DNNs usually show a better handling of correlations between variables to enhance classifier
 1682 performance. The multi-label classification attempts to classify each event in four process categories
 1683 (HH signal, ZH, $t\bar{t}H$ or continuum $\gamma\gamma+jets$). This helps in learning event topology for a better separation
 1684 between signal and each background component separately. In addition, the multi-label classification al-
 1685 lows for a control region where ZH is dominant, which could be used as a validation for the $HH \rightarrow b\bar{b}\gamma\gamma$
 1686 analysis.

1687 A separate DNN classifier is trained in each mass region using the same signal and background
 definitions as the BDT. The architecture of the two classifiers is identical. The classifiers are built using
 the KERAS library with TENSORFLOW as a backend [89, 90]. It contains an input layer constructed
 from a batch normalization and a fully connected layer. Then five fully connected hidden layers with
 each one followed with a dropout layer, and finally an output layer. Each of the fully connected layers
 has 128 nodes using a Rectified Linear Unit (ReLU) activation function, and their weights are randomly
 initialized by sampling from a truncated distribution centred on zero with width given by $\sqrt{\frac{2}{N_{inputs}}}$ where
 N_{inputs} is the number of input features. The dropout layers are used to improve the robustness of the
 training and reduce overfitting effects. For the high mass region, the dropout rate is 11%, while for the
 low mass region, the dropout rate is 22.5%. The batch normalization layer is used to standardize the
 inputs to the first layer. The output layer is 4-node wide and activated using a softmax function which
 allows one to interpret the output as the probabilities for each associated class (ggF HH, ZH, $t\bar{t}H$ or
 $\gamma\gamma+jets$) for the input event. During the training, a weighted categorical cross-entropy is used as loss
 function and the Adam optimizer for network weight optimization.

Weights of the loss function correspond to the event class weight computed as the weight of an event j
 associated to class i as:

$$\text{weight}_i^j = \frac{N}{n \times N_{\text{class}_i}}, \quad (6.3)$$

1688 where $n = 4$ is the number of classes, N is the total number of events and N_{class_i} is the total number
 1689 of events in the given class i . The event class weight enables the models to perform similarly between
 1690 classes and reduces the issue of unbalanced data which affects the model classification ability.

1691 The two DNN models are trained with a learning rate of 10^{-4} and a batch size fixed to 1000 events, with
 1692 a maximum number of epochs set to 100. Similar to the CNN implemented for photon identification,
 1693 an early stopping metric evaluated on the validation data sample throughout the training is imposed
 1694 during the training phase. The early-stopping stops the training around epoch 40 to avoid overfitting.
 1695 DNNs use almost the same input variables as the BDT, as listed in Table 6.3, except for the single
 1696 topness, H_T , the missing transverse energy and its azimuthal angle variables which are not included.
 1697 The class-invariant symmetries (detector symmetry) complexity is removed by rotating all events around
 1698 the beam axis in such a way that the leading photon has $\phi = 0$.

1699 Figure 6.5 shows the evaluation of the loss function during the training time for the training and
 1700 validation dataset. The low mass model shows a high loss error compared to the model trained in the

high mass region, which translates the signal like background distribution, reducing the model separation in the low mass region.

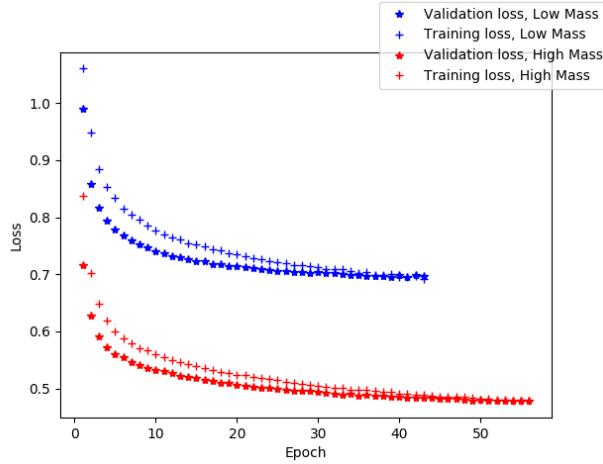


Figure 6.5: Evaluation of the loss function for high mass (red) and low mass (blue) DNNs as a function of epoch number.

One of the performance measurement of a classifier is the confusion matrix: the matrix compares the actual target values with those predicted by the model. This gives a holistic view of the performance of the classification model as well as the associated errors. Figure 6.6 shows the confusion matrix for high mass and low mass models. Note that the threshold value for the prediction of the model is set to 0.5 on the softmax outputs. Both models show a high ($\sim 19\%$) confusion between the ggF HH signal and the ZH background, coming from the similarity between the event topology of the two processes. In the low mass region, the performance clearly decreases due to the similarity in the kinematics and event topology between ggF HH and backgrounds which degrades the sensitivity at high κ_λ values.

6.2.3.2 Discriminant variable

The four outputs of the softmax layer are combined in a single discriminant denoted d_{HH} for each region, the probabilities are normalized to the corresponding process cross-section.

$$d_{HH}^{HM} = \log \left(\frac{\sigma_{SM} \cdot p_{HH}}{\sigma_{ZH} \cdot p_{ZH} + \sigma_{t\bar{t}H} \cdot p_{t\bar{t}H} + \sigma_{\gamma\gamma+jets} \cdot p_{\gamma\gamma+jets}} \right) \quad (6.4)$$

$$d_{HH}^{LM} = \log \left(\frac{\sigma_{\kappa_\lambda 10} \cdot p_{HH}}{\sigma_{ZH} \cdot p_{ZH} + \sigma_{t\bar{t}H} \cdot p_{t\bar{t}H} + \sigma_{\gamma\gamma+jets} \cdot p_{\gamma\gamma+jets}} \right) \quad (6.5)$$

d_{HH}^{HM} in Eq. 6.4 represents the d_{HH} discriminant for the high mass region, while Eq. 6.5 is the d_{HH} discriminant in the low mass region.

Figure 6.7 shows the distribution of d_{HH} discriminant for each process (ggF HH, ZH, $t\bar{t}H$ and $\gamma\gamma+jets$) in each region. At high mass, d_{HH} shows a clear separation between the HH and other processes, while for low mass the discrepancy is low as expected. The side-band data follows the continuum background closely, indicating a good modelling of the d_{HH} distribution.

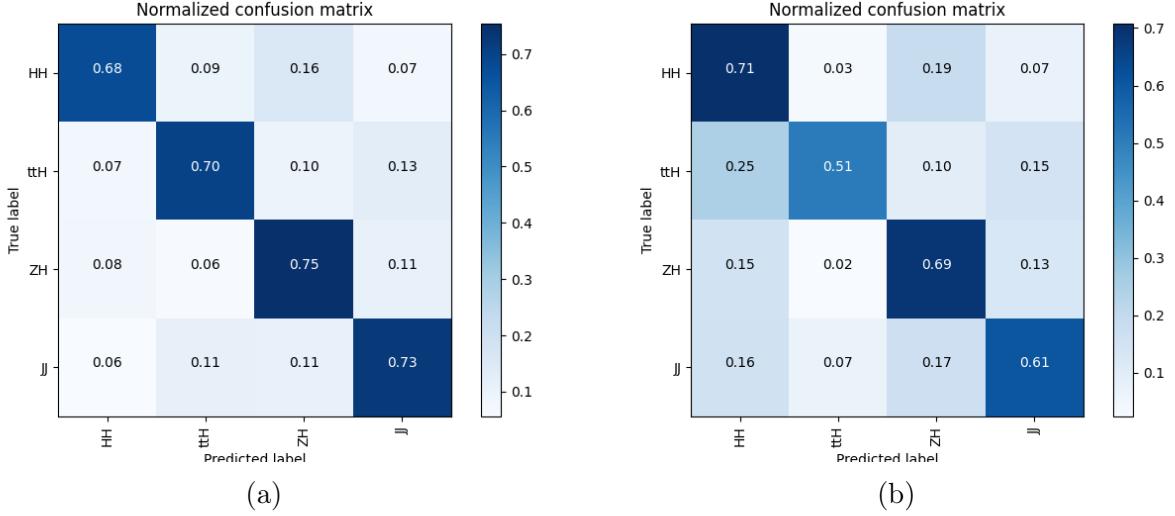


Figure 6.6: Normalized (to unity) confusion matrix for (a) high mass and (b) low mass models. JJ here denote the continuum $\gamma\gamma+\text{jets}$.

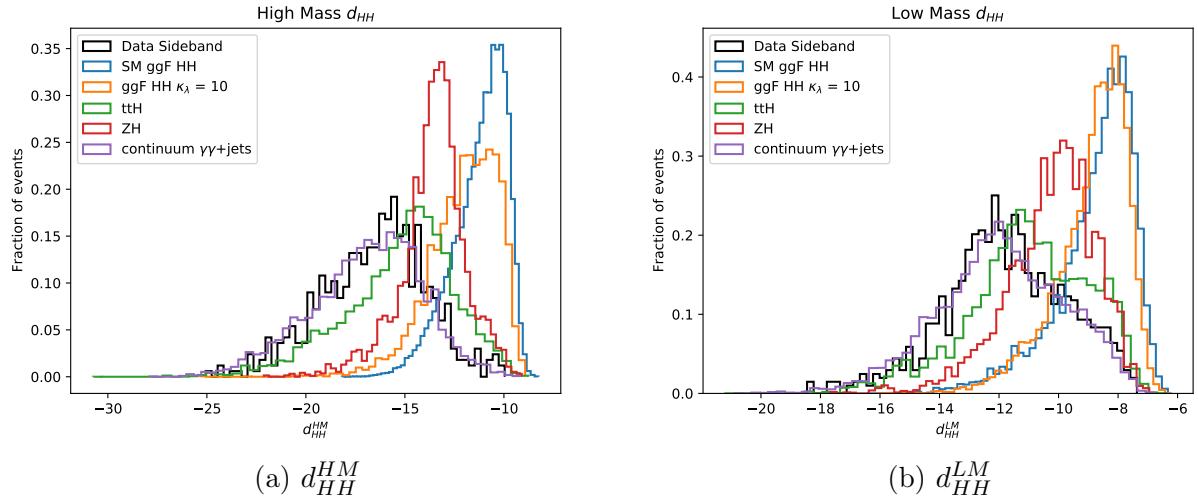


Figure 6.7: d_{HH} discriminant distribution in (a) high mass and (b) low mass. The black line shows the real data in the sideband region ($m_{\gamma\gamma} \in [105, 120] \cup [130, 160]$).

¹⁷¹⁹ **6.2.3.3 Significance scan**

In each mass region, the corresponding discriminant is scanned after applying $90 < m_{b\bar{b}} < 140$ GeV requirement to build two orthogonal categories high d_{HH} and low d_{HH} with a maximum significance. The category with the highest significance must have at least 0.8 continuum $\gamma\gamma + \text{jets}$ events in the $123 < m_{\gamma\gamma} < 127$ GeV mass range sufficient to perform reliable $m_{\gamma\gamma}$ fit. This is done similarly to the BDT. The significance is computed using the Asimov formula [127] defined as:

$$Z = \sqrt{2[(s+b) \times \log(1+s/b) - s]}, \quad (6.6)$$

¹⁷²⁰ where, s is the ggF HH signal yield and b is the total background yield.

¹⁷²¹ In total, four categories are defined, two for the low mass region and two for the high mass region.

¹⁷²² The significances of the four categories are then combined for the final significance. Table 6.5 lists the significance in each category.

Categories	SM ggF HH	BSM $\kappa_\lambda=10$ ggF HH
High mass, High d_{HH}	0.53	2.45
High mass, Low d_{HH}	0.11	0.94
Low mass, High d_{HH}	0.03	2.21
Low mass, Low d_{HH}	0.01	0.54
Combined	0.54σ	3.47σ

Table 6.5: Significance in each category for SM ggF HH and BSM $\kappa_\lambda = 10$ ggF signals.

¹⁷²³ The proposed selection based on a DNN shows similar performances with respect to the BDT with
¹⁷²⁴ less inputs variables. The missing variables in the DNN are very discriminant in the BDT case, thus
¹⁷²⁵ including those variables in the DNN could have enhanced the DNN performance. Given the timescale
¹⁷²⁶ of the analysis, the DNN work was stopped here to focus on work with higher priority. Future DNN
¹⁷²⁷ optimization are kept for the Run-3 analysis. The DNN based selection is available and documented for
¹⁷²⁸ the next publication with Run-3 data.

¹⁷³⁰

¹⁷³¹ **6.3 Signal and background modelling**

¹⁷³² To extract the $HH \rightarrow b\bar{b}\gamma\gamma$ signal yield, a fit to the $m_{\gamma\gamma}$ distribution in the range [105, 160] GeV is
¹⁷³³ performed in the data. Analytical functions are used to describe the signal, single Higgs boson and
¹⁷³⁴ the continuum backgrounds. Parameters of the signal and the peaking single Higgs boson background
¹⁷³⁵ shapes are fixed on simulated distributions.

¹⁷³⁶ **6.3.1 Signal and single Higgs boson parameterization**

¹⁷³⁷ Both signal and peaking single Higgs boson backgrounds are described analytically by a Double-Sided
¹⁷³⁸ Crystal Ball (DSCB) function which is characterized by a Gaussian core and asymmetric power law tails

1739 [128], defined as:

$$f_{\text{DSCB}}(m_{\gamma\gamma}) = N \times \begin{cases} e^{-t^2/2} & \text{if } -\alpha_{low} \leq t \leq \alpha_{high} \\ \frac{e^{-\frac{1}{2}\alpha_{low}^2}}{\left[\frac{1}{R_{low}}(R_{low}-\alpha_{low}-t)\right]^{n_{low}}} & \text{if } t < -\alpha_{low} \\ \frac{e^{-\frac{1}{2}\alpha_{high}^2}}{\left[\frac{1}{R_{high}}(R_{high}-\alpha_{high}-t)\right]^{n_{high}}} & \text{if } t > \alpha_{high} \end{cases} \quad (6.7)$$

1740 with $t = \frac{m_{\gamma\gamma} - \mu_{CB}}{\sigma_{CB}}$. μ_{CB} and σ_{CB} describe the mean and the width of the Gaussian core, $\alpha_{low/high}$ 1741 the transition from the core to the tails of the distribution and $n_{low/high}$ the tails. N normalizes the 1742 distribution. The explicit separation of the Gaussian core from the tails in this function allows an easy 1743 treatment of the systematic uncertainties, as the energy scale uncertainty impacts the mean μ_{CB} and 1744 the energy resolution systematic the width σ_{CB} .

1745 The normalizations are obtained from the expected MC yields. For each BDT category, the parameters 1746 of the DSCB are obtained from a fit of the $m_{\gamma\gamma}$ distribution obtained from simulation of the ggF and 1747 VBF HH processes with $\kappa_\lambda = 1$, since no significant dependence of the functional form with κ_λ was 1748 found.

1749 Fitted parameters are further allowed to vary only within the systematic uncertainties on the energy 1750 scale and resolution. Injection tests are performed to quantify potential biases from the signal HH 1751 only fit and signal + single Higgs boson fit. No statistically significant bias ($< 10\%$) was observed in 1752 both tests, thus the same parameterized functions are used to fit the signal and single Higgs boson 1753 backgrounds.

1754 6.3.2 Continuum background parameterization

1755 The continuum background estimation is done in a data-driven way. The simulated $\gamma\gamma$ +jets sample 1756 is used only to select the functional form and its parameters are then left completely free in the fit to 1757 the data, reducing uncertainties due to background mismodelling. The background is decomposed into 1758 several components to take into account contribution from the continuous $\gamma\gamma$, γ -jet (γj) and jet-jet 1759 (jj) fake photons components. The decomposition is done using a 2x2 ABCD side-band method [129], 1760 in which an equation system is built and then solved to determine the relative contribution of each 1761 component (purity). The purities are measured inclusively and in each BDT category. Inclusively, the 1762 $\gamma\gamma$ event purity is found to be $85 \pm 2.9\%$ and the remaining $15 \pm 2.8\%$ consists of γj events and a 1763 negligible amount of jj events. The $\gamma\gamma$ contribution increases after the BDT selection since the BDT 1764 tends to select events with higher photon p_T , thus more real photons. The decomposition is shown in 1765 Figure 6.8.

1766 The functional form is chosen by fitting the simulated background template. Given the high $\gamma\gamma$ purity, 1767 the template is constructed from the SHERPA $\gamma\gamma$ +jets normalized to the data side-band. The bias on the 1768 specific choice of functional form (to absorb a potential signal or create an artificial signal) is determined 1769 by the spurious signal test.

1770 6.3.2.1 Spurious signal test

1771 The spurious signal (SS) refers to any potential fake signal N_{SS} obtained from fitting a smooth back- 1772 ground only distribution to a function that describes the signal plus background shape. The N_{SS} is taken 1773 to be the largest number of fitted signal events computed for Higgs boson masses varying in intervals of 1774 1 GeV from 121 to 129 GeV. The SS requires the tested function to fulfil the following criteria:

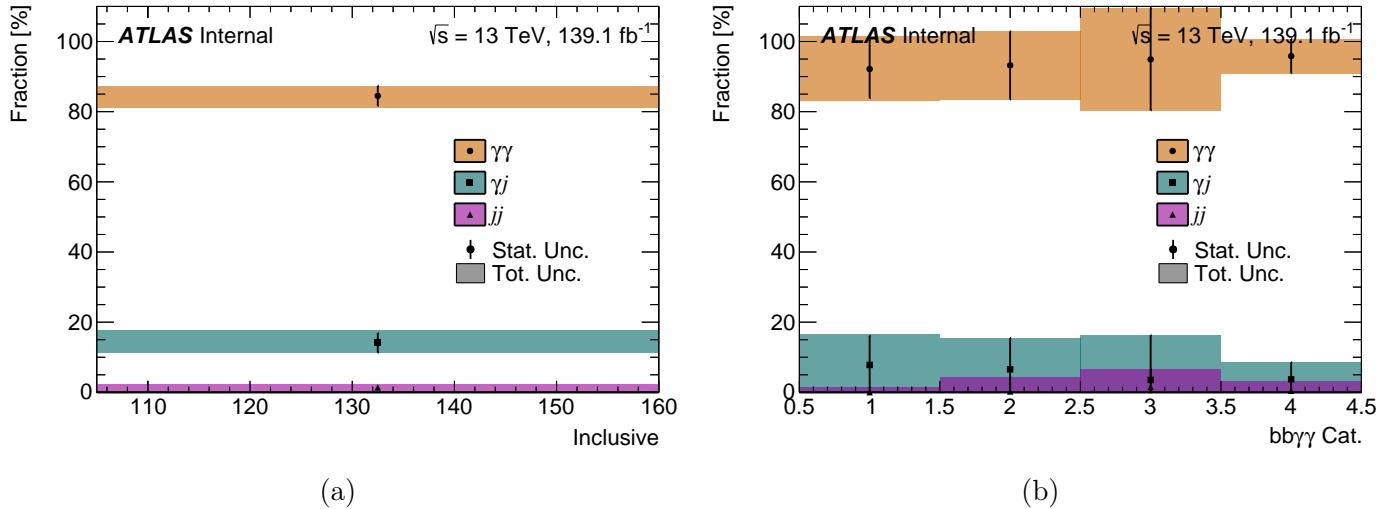


Figure 6.8: $\gamma\gamma$, γj , and jj fractions in the inclusive (a) and the BDT categories (b). In the BDT figure bin 1 corresponds to the High mass BDT tight category, bin 2 to the High mass BDT loose category, bin 3 to the Low mass BDT tight category, and finally bin 4 to the Low mass BDT loose category.

Category	N_{SS}	Z_{sp}	$p(\chi^2)$ [%]
High mass BDT tight	0.688	0.394	68.8
High mass BDT loose	0.990	0.384	30.5
Low mass BDT tight	0.594	0.378	29.8
Low mass BDT loose	1.088	0.272	26.9

Table 6.6: Spurious signal result for the exponential functional form for the various BDT categories. In each category, the spurious signal value, its ratio to the expected statistical error from data, and χ^2 the probability of the background-only fit.

- $N_{SS} < 10\% N_s^{exp}$, where N_s^{exp} is the expected number of signal events within the BDT category.
- $N_{SS} < 20\% \sigma_{bkg}$, where σ_{bkg} is the statistical uncertainty on the fitted signal events when performing a signal+background fit to the template.

Due to the limited background statistics, a new variable ξ_{SS} is defined to relax the criterion defined before. The ξ_{SS} is defined in a such way as to allow for a 2σ local statistical fluctuation in the background template. ξ_{SS} is defined as:

$$\xi_{SS} = \begin{cases} N_{SS} + 2\Delta_{MC}, & N_{SS} + 2\Delta_{MC} < 0 \\ N_{SS} - 2\Delta_{MC}, & N_{SS} - 2\Delta_{MC} > 0 \\ 0, & otherwise \end{cases} \quad (6.8)$$

where Δ_{MC} is the statistical uncertainty on N_{SS} due to limited statistics. ξ_{SS} should then pass the same criteria as N_{SS} .

The function that gives the smallest N_{SS} is selected as the final background function and the N_{SS} is added as a systematic on the signal yield. The exponential function " $\exp(a.m_{\gamma\gamma})$ " is found to be the best choice for all BDT categories and the corresponding parameters for SS are shown in Table 6.6.

1783 6.4 Systematic uncertainties

1784 Due to the small number of expected signal events (Section 2.5.1), and the restrictive event selection, the
 1785 dominant source of uncertainty in this analysis is still the statistical one. Nevertheless, the amplitude
 1786 of systematic sources have to be evaluated. They can be split into two categories: the experimental
 1787 and theoretical uncertainties. The impact of each systematic on the shape and/or the expected event
 1788 yield (Normalization) is evaluated in each BDT category. Since the continuum background is the main
 1789 source of background, and is estimated with data-driven method, no associated systematic are assigned.
 1790 Systematic are estimated for the single Higgs boson backgrounds and HH signal using simulation. Tech-
 1791 nically, these systematic uncertainties are implemented as nuisance parameters in the statistical model
 1792 and constrained by a Gaussian or log-normal function in most cases.

1793 6.4.0.1 Experimental systematic uncertainties

1794 Experimental systematic uncertainties are mostly related to the reconstruction, the calibration, the tag-
 1795 ging and the identification of the physics objects used in the analysis. In addition to, a 1.7% uncertainty
 1796 on the integrated luminosity used in this analysis (full Run-2). The following sources of systematic are
 1797 also considered:

- 1798 • Spurious signal: as described in Section 6.3.2.1, the potential bias arising from the background
 1799 modelling is assessed as an additional source of uncertainty in the total number of expected HH
 1800 signal events in each category according to Table 6.6.
- 1801 • Photon energy scale and resolution: related to the measurement of photon energy and its calibra-
 1802 tion arising from the different component such as the amount of material in front of the calorimeter
 1803 and cell energy non-linearity. It affects both the shape and normalization of the signal model and
 1804 its amplitude is taken from [130].
- 1805 • Photon identification and isolation efficiencies: resulting mainly from the mis-modelling of the
 1806 shower shape described in Section 4.3 and bias of the ID measurement methods listed in Section
 1807 4.2.
- 1808 • Trigger and vertex efficiencies: these are considered to account for the di-photon trigger efficiency
 1809 uncertainty which affects the acceptance by 1% for each category, and the photon-pointing vertex
 1810 efficiency.
- 1811 • Jet energy scale and resolution: similarly to photon energy related-uncertainties, they are related
 1812 to the jet energy measurement and its calibration chain described in Section 5.2.1. It affects the
 1813 $m_{b\bar{b}}$ distribution and are propagated to the E_T^{miss} calculation.
- 1814 • Flavour-tagging: this account for the b -tagging uncertainties resulting from the impact of parton
 1815 shower and hadronisation on the b -tagging efficiency [111]. This affects the acceptance.

1816 In addition to the systematic listed above, an additional uncertainty on the yield from the p_T -
 1817 Reco correction is examined. Three variations of the correction function were generated using different
 1818 b -tagging WPs (60%, 70%, 85%). Figure 6.9 shows the p_T -Reco correction distribution for the four
 1819 b -tagging WPs (60%, 70%, 77%, 85%). All WPs show almost the same correction distribution for high
 1820 p_T jet while for low p_T the variation is larger. At low p_T , a large fraction of light-jet mis-identified as
 1821 b -jet contributes to compute the p_T -Reco correction, which leads to a low correction factor, specially for

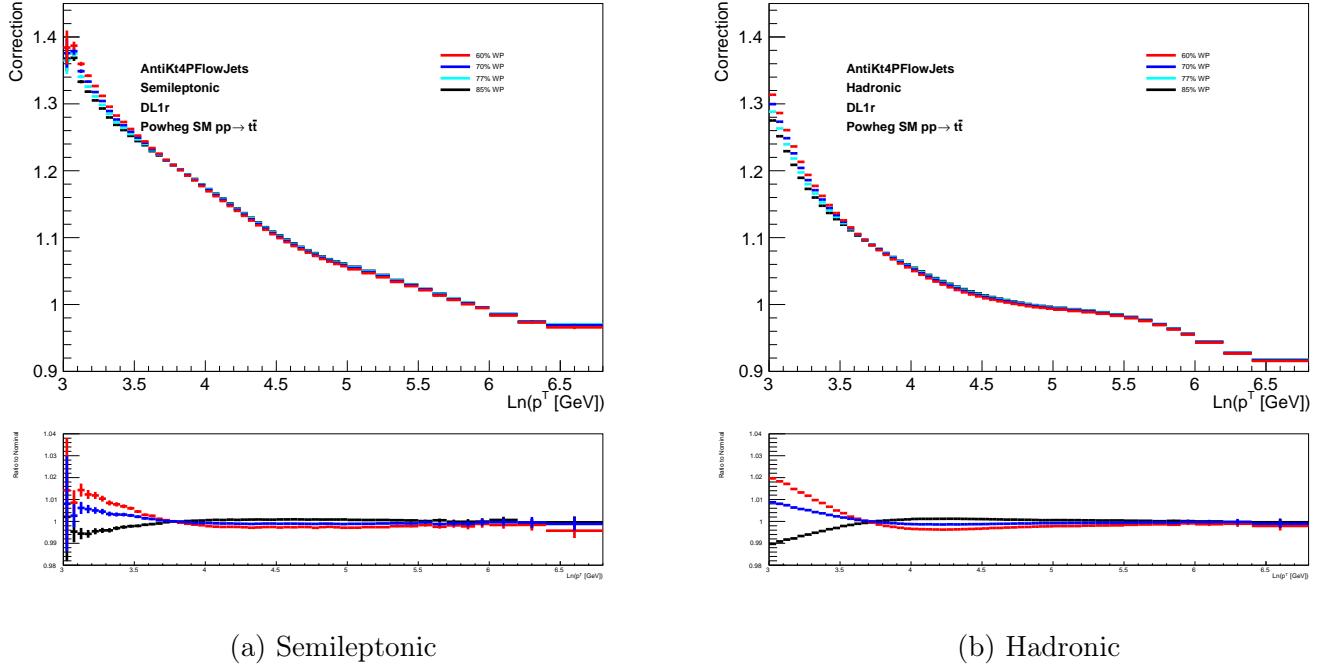


Figure 6.9: p_T -Reco correction distribution using different DL1r b -tagging WPs for both semileptonic (a) and hadronic (b) jets. Bottom panel shows the ratio of the p_T -Reco to the nominal (77% WP).

1822 85% WPs which is the looser WP, and more the WP is tighter more the correction factor is higher.
 1823 The relative impact on the yield for a given correction function is used as a systematic uncertainty.
 1824 Table 6.7 summaries the relative effect of the nominal p_T -Reco correction on the yield. The impact on
 the shape is negligible for both the position and the spread.

p_T -Reco variation	% relative effect to nominal
60% WP	± 0.094
70% WP	± 0.065
85% WP	± 0.12

Table 6.7: Relative p_T -Reco systematic uncertainty on the yield for given variation.

1825
 1826 Even for 85% WP (with a large acceptance) which is related to the high amplitude of systematic
 1827 uncertainties, it is of the same order as the flavour tagging uncertainty and therefore negligible compared
 1828 to the jet energy scale and resolution uncertainties. Thus this systematic is neglected.

1829 6.4.0.2 Theoretical systematic uncertainties

1830 Sources of systematic uncertainties related to the theory are considered:

- 1831 • QCD scale uncertainties: these are defined to account for the potential effect of missing higher
 1832 order in the matrix element. These are estimated by varying the re-normalisation μ_R and the
 1833 factorization μ_F scales coherently by a factor of 2 or 0.5 from their nominal values and recalculating
 1834 the cross-section in each case. The largest deviation from the nominal cross-section across the κ_λ
 1835 range [-10, 10] is taken as the uncertainty.
- 1836 • PDF+ α_s : account for the impact on the cross-sections of the choice of the PDF set as well as the
 1837 α_s variation from its central value. These are taken from Ref. [131] to be 3% and kept constant
 1838 over the full κ_λ range.
- 1839 • Higgs boson branching ratio: accounts for the $H \rightarrow \gamma\gamma$ and $H \rightarrow b\bar{b}$ branching fractions uncer-
 1840 tainties of the order of 2.9% and 1.7% respectively, taken also from Ref. [131].
- 1841 • Higgs boson mass: accounts for the Higgs boson mass measurement uncertainty of the order of
 1842 0.2%, taken from Ref. [20].
- 1843 • Heavy flavour (HF): accounts for the uncertainty on the amount of select single ggF H, VBF H
 1844 and WH events in which the two b -jets are probing a special phase space. This is considered to be
 1845 100%.
- 1846 • non-closure of κ_λ reweighting: account for the discrepancy in the κ_λ reweighting method as de-
 1847 scribed in Section 6.1.

1848 **6.5 Statistical model**

The observed data is interpreted using the Modified Frequentist method (CL_s) [132]. A hypothesis test is made to evaluate the compatibility between the observed data and the background-only hypothesis for limit computation, which assumes no signal in the data. The test statistic used here is the profile likelihood ratio [127] built from signal+background model. The signal is extracted by performing a simultaneous maximum likelihood fit of the $m_{\gamma\gamma}$ distribution in the range [105, 160] GeV over the four categories. The likelihood is defined as:

$$\mathcal{L} = \prod_c \left(Pois(n_c | N_c(\theta)) \cdot \prod_{i=1}^{n_c} f_c(m_{\gamma\gamma}^i, \theta) \cdot G(\theta) \right), \quad (6.9)$$

where for each event i in a category c , n_c is the observed number of events in real data, while N_c is the expected number of event from simulation, f_c the value of the pdf function, θ the nuisance parameters (systematic uncertainties) with G is their constraint pdfs. The expected number of events combines the signal HH from both ggF and VBF production modes, single Higgs boson backgrounds, the spurious signal and the continuum background, defined as:

$$N_c(\theta) = \mu \cdot N_{\text{HH},c}(\theta_{yield}) + N_{\text{H},c}(\theta_{yield}) + N_{\text{SS},c} \cdot \theta_{\text{SS},c} + N_{\text{continuum},c}, \quad (6.10)$$

where θ_{yield} and θ_{SS} are the systematics on the yield from experimental, theoretical uncertainties and spurious signal uncertainties. μ is the parameter of interest (POI) of the fit, (the signal strength) which measures the HH cross-section normalized to its SM predicted value.

The pdf f_c is the sum of the DSCBs modelling the HH, single Higgs boson backgrounds and the spurious signal and the exponential function modelling the continuum background described in Section 6.3. Only the signal strength, the exponential function parameter a , its normalization $N_{\text{continuum}}$ and the systematics uncertainties are allowed to float in the fit.

The profile likelihood ratio test is defined as:

$$\tilde{q}_\mu = \begin{cases} -2 \log \frac{L(\mu, \hat{\theta}(\mu))}{L(0, \hat{\theta}(0))} & \hat{\mu} < 0 \\ -2 \log \frac{L(\mu, \hat{\theta}(\mu))}{L(\hat{\mu}, \hat{\theta})} & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \quad (6.11)$$

1849 where $\hat{\mu}$ is the fit favoured signal strength, $\hat{\theta}$ is the fit favoured nuisance parameter values when μ is 1850 fixed in the fit and $\hat{\theta}$ is the fit favoured nuisance parameter when μ is also floating.

1851 In the absence of signal, an exclusion limit at 95% Confidence Level (CL) is set on the $HH \rightarrow b\bar{b}\gamma\gamma$ cross- 1852 section for each κ_λ value in the range [-10, 10] with step of 0.2. The asymptotic approximation [127] is 1853 adopted here. A pseudo-data is used for the computation of the expected limits, called Asimov dataset. 1854 An Asimov dataset provides an idealized description of the predicted data with a given integrated 1855 luminosity without statistical fluctuations.

1856 **6.6 Results**

1857 **6.6.1 Observed events and $m_{\gamma\gamma}$ fit**

1858 The results of the background-only fit of the $m_{\gamma\gamma}$ distribution for each category are shown in Figure 1859 6.10. The number of the observed and expected events in each category is summarized in Table 6.8.

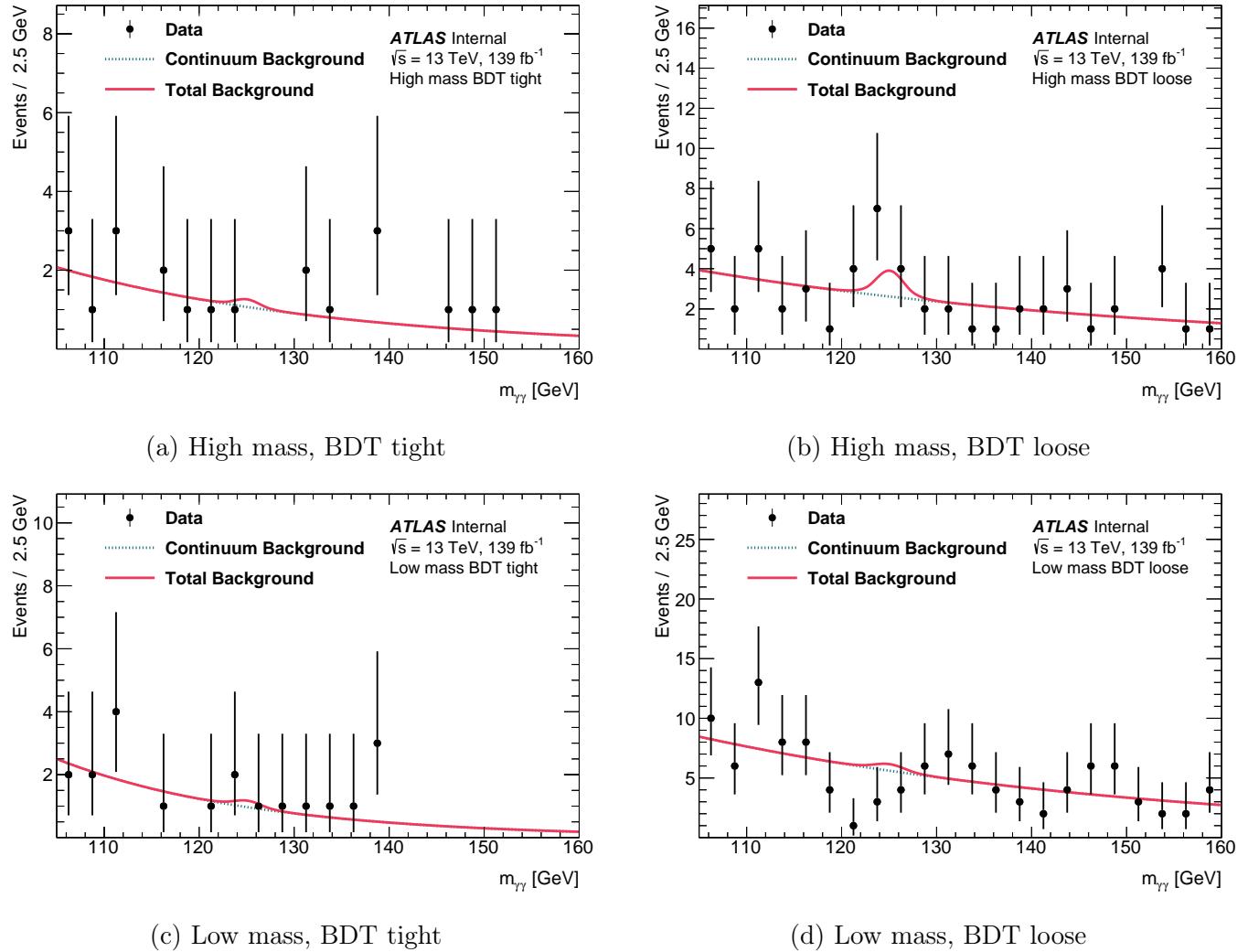


Figure 6.10: The data points are compared to the background-only fit for the four BDT categories. Both the continuum background and the background from single Higgs boson production are considered.

	High mass BDT tight	High mass BDT loose	Low mass BDT tight	Low mass BDT loose
Continuum	4.9 ± 1.1	9.5 ± 1.5	3.7 ± 1.0	24.9 ± 2.5
Single Higgs boson	0.670 ± 0.032	1.57 ± 0.04	0.220 ± 0.016	1.39 ± 0.04
ggF	0.261 ± 0.028	0.44 ± 0.04	0.063 ± 0.014	0.274 ± 0.030
$t\bar{t}H$	0.1929 ± 0.0045	0.491 ± 0.007	0.1074 ± 0.0033	0.742 ± 0.009
ZH	0.142 ± 0.005	0.486 ± 0.010	0.04019 ± 0.0027	0.269 ± 0.007
Rest	0.074 ± 0.012	0.155 ± 0.020	0.008 ± 0.006	0.109 ± 0.016
SM HH ($\kappa_\lambda = 1$)	0.8753 ± 0.0032	0.3680 ± 0.0020	$(49.4 \pm 0.7) \cdot 10^{-3}$	$(78.7 \pm 0.9) \cdot 10^{-3}$
ggF	0.8626 ± 0.0032	0.3518 ± 0.0020	$(46.1 \pm 0.7) \cdot 10^{-3}$	$(71.8 \pm 0.9) \cdot 10^{-3}$
VBF	0.01266 ± 0.00016	0.01618 ± 0.00018	$(3.22 \pm 0.08) \cdot 10^{-3}$	$(6.923 \pm 0.011) \cdot 10^{-3}$
BSM HH ($\kappa_\lambda = 10$)	6.36 ± 0.05	3.691 ± 0.038	4.65 ± 0.04	8.64 ± 0.06
Data	2	17	5	14

Table 6.8: Expected and observed numbers of events in the signal region ($[120, 130]$ GeV) for the four BDT categories. The uncertainties on the continuum background are those arising from the fitting procedure. The uncertainties on the single Higgs boson and Higgs boson pair production are from MC statistical error.

1860 6.6.2 Cross-section limits and κ_λ constraint

1861 Since no significant excess over the background prediction is found, exclusion limits at 95% CL are
 1862 computed on the Higgs boson pair production cross-section for each κ_λ as described in Section 6.5.
 1863 The upper limit obtained for the double Higgs boson production cross-section σ_{HH} in the ggF + VBF
 1864 production modes is 140 fb assuming $\kappa_\lambda = 1$. A median value of 180 fb is estimated. Expressed in terms
 1865 of multiple of the SM production cross-section, the observed limit is 4.1 times the SM prediction while
 1866 the expected limit is 5.5 times the SM prediction. The expected measured signal strength μ is 1 ± 2.37 .
 1867 Constraint on the Higgs boson self-coupling is derived from the cross-section limit scan versus κ_λ by
 1868 taking the intersection between the theoretical prediction and the measured cross-section, as shown in
 1869 Figure 6.11. The observed constraint at 95% CL is $-1.5 < \kappa_\lambda < 6.7$, while the expected constraint is
 1870 $-2.4 < \kappa_\lambda < 7.7$. BSM effects on the single Higgs boson cross-section and Higgs boson branching ratios
 1871 are neglected. The couplings of Higgs boson to other particles are set to their SM values [133]. Table
 6.9 summarizes the results.

	Expected	Observed
$\sigma_{HH}/\sigma_{HH}^{SM}$ limit	5.5	4.1
κ_λ interval	$[-2.4, 7.7]$	$[-1.5, 6.7]$

Table 6.9: Summary of $HH \rightarrow b\bar{b}\gamma\gamma$ limits at 95% CL from the 139 fb^{-1} analysis.

1872 In addition to the constraint, a measurement of κ_λ is performed by floating κ_λ in the fit with all other
 1873 couplings set to their SM values. The best-fit κ_λ is $2.7^{+2.1}_{-2.2}$ which is close to the κ_λ value giving the
 1874 minimum HH cross-section. The likelihood scan as a function of κ_λ is shown in Figure 6.12.
 1875 Table 6.10 shows the expected limit at 95% CL on the HH cross-section for each category. The signifi-
 1876 1877 cance is mainly driven by the High mass high BDT category.

Figure 6.13 shows the expected limit at 95% CL on the cross-section of the HH pair production as
 1878 a function of the κ_λ modifier for individual contributions of the statistical and statistical + systematic

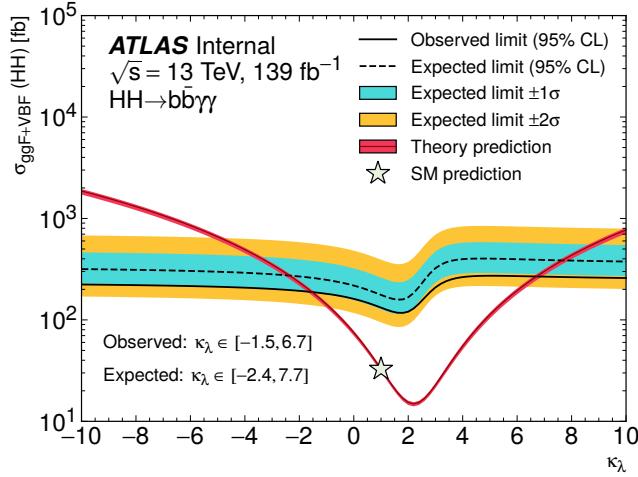


Figure 6.11: Observed and expected limits at 95% CL on the cross-section of Higgs boson pair production as a function of the Higgs boson self-coupling modifier $\kappa_\lambda = \lambda_{HHH}/\lambda_{HHH}^{\text{SM}}$. The $\pm 1\sigma$ and $\pm 2\sigma$ bands show the expected variations on the expected limit due to statistical and systematic uncertainties. The theory prediction curve represents the scenario where all parameters and couplings are set to their SM values except for κ_λ . The uncertainty band of the theory prediction curve shows the cross-section uncertainty.

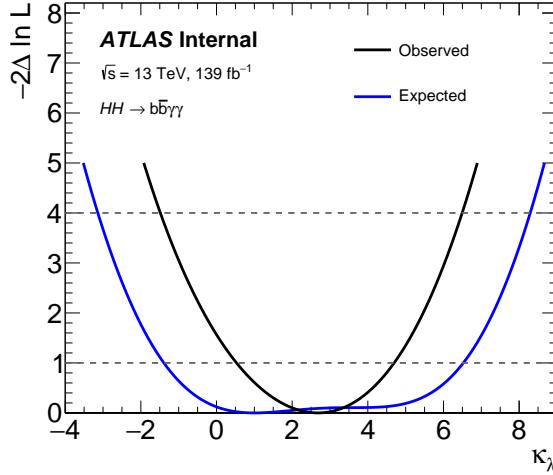


Figure 6.12: The likelihood function as a function of κ_λ , with $\mu = 1$. The observed 68% and 95% CL ranges are [0.5, 4.7] and [-1.4, 6.5] respectively, while the expected are [-1.4, 6.4] and [-3.1, 8.2].

Category	$\sigma_{HH}/\sigma_{HH}^{SM}$ limit
High mass, High BDT	5.8
High mass, Low BDT	21.0
Low mass, High BDT	102.9
Low mass, Low BDT	125.6
Combined	5.5

Table 6.10: The expected limit on the HH cross-section at 95% CL for each analysis category for SM case. The combined limit is also shown.

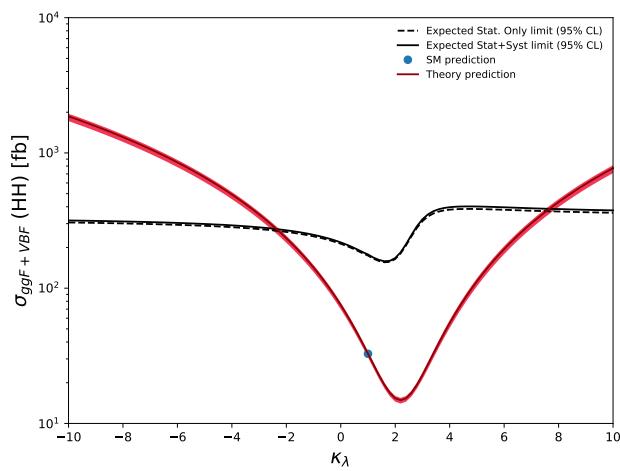


Figure 6.13: Expected limits at 95% CL on the cross-section of Higgs boson pair production as a function of the Higgs boson self-coupling modifier κ_λ for statistical only and statistical+systematic configurations. The luminosity is equal to 139 fb^{-1} .

1880 uncertainties. The statistical only limits are computed by setting all the systematic uncertainties to zero.
 1881 Table 6.11 summarizes the upper limit for $\kappa_\lambda = 1$ case and the constraint on κ_λ for each systematic
 1882 configuration. It shows that this search is statistically limited, and the total impact of systematic
 uncertainties on the results is about 4%.

1883
 1884 Table 6.12 shows the impact of the systematic uncertainties on the upper limit on the signal strength
 1885 μ . The dominant systematics are related to the spurious signal and the photon energy scale (PES)
 1886 uncertainties. As mentioned in Section 6.4.0.1, the PES systematic arises from the measurement of
 1887 photon energy and its calibration which is mainly induced by the cell non-linearity and the under-
 1888 estimation of passive material in front of the calorimeter in simulation. The bias due to the background
 1889 function choice leads to the large spurious signal uncertainties and mainly dominated by the statistical
 1890 fluctuations of the continuum background statistics used in the spurious signal test.
 1891 Since simulating more events was however not possible in the timescale of the analysis, several attempts
 1892 are done to improve the continuum statistics and reduce statistical fluctuation:

	Stat.	Only	Stat.+Syst.
$\sigma_{HH}/\sigma_{HH}^{SM}$ limit	5.3	5.5	
κ_λ interval	[−2.3, 7.6]	[−2.4, 7.7]	

Table 6.11: Expected upper limit at 95% CL on the cross-section of the HH pair production and allowed κ_λ interval for statistical only and statistical + systematic uncertainties

- The Gaussian Process Regression (GPR) [134] is used to smooth and reduce fluctuations in the $m_{\gamma\gamma}$ distribution, this method reduces the N_{SS} by 84%.
 - Increasing statistics by inverting the b -tagging requirements reduces the systematic on μ by $\sim 3\%$.
- Since the analysis is statistically dominated, the significance does not improve significantly, thus these improvements are not included in these results.

Source	Type	Relative impact in %
Experimental		
Photon energy scale	Norm. + Shape	5.2
Photon energy resolution	Norm. + Shape	1.8
Flavour tagging	Normalization	0.5
Theoretical		
Heavy flavour content	Normalization	1.5
Higgs boson mass	Norm. + Shape	1.8
PDF+ α_s	Normalization	0.7
Spurious signal	Normalization	5.5

Table 6.12: Breakdown of the dominant systematic uncertainties. Only systematic uncertainties with an impact of at least 0.5% are shown. Uncertainties of Norm. + Shape type have effects on both the yield and the shape of the $m_{\gamma\gamma}$ distributions, the rest of uncertainties affect only the yields.

6.7 Comparison to 36 fb^{−1} results

- Table 6.13 compares the observed and expected limits as well as the constraint on κ_λ derived with 36 fb^{−1} and 139 fb^{−1} datasets [24, 135].

	Expected 36 fb ^{−1}	Observed 36 fb ^{−1}	Expected 139 fb ^{−1}	Observed 139 fb ^{−1}
$\sigma_{HH}/\sigma_{HH}^{SM}$ limit	28	22	5.5	4.1
κ_λ interval	[−8.3, 13.2]	[−8.2, 13.2]	[−2.4, 7.7]	[−1.5, 6.7]

Table 6.13: Summary of $HH \rightarrow b\bar{b}\gamma\gamma$ results from 36 fb^{−1} and 139 fb^{−1} analyses.

- The limit in the 139 fb^{−1} is improved compared to 36 fb^{−1} analysis by a factor 5. The improvement is coming from both the increase by factor a 3.9 of the integrated luminosity and the new analysis

strategy. By extrapolating the 36 fb^{-1} results to 139 fb^{-1} , the new analysis strategy improves the limit by a factor 2.4. Such improvement is mainly coming from the implementation of machine learning in event categorization (BDT) and the split in the two $m_{b\bar{b}\gamma\gamma}^*$ bins (High/Low mass).

In addition to the analysis strategy, the 139 fb^{-1} analysis benefits from several improvements such as the usage of the PFlow algorithm for jet reconstruction which improves jet resolution as well as the specific b -jet energy calibration which enhances the analysis sensitivity by $\sim 10\%$. The b -tagging efficiency is also improved by moving from the MV2c10 algorithm to the DL1r which has a higher efficiency as shown in Figure 5.9.

Considering the VBF production mode for the first time as a part of the signal when setting limits on κ_λ brings an improvement of 6-8% in κ_λ interval.

6.8 Comparison to CMS $HH \rightarrow b\bar{b}\gamma\gamma$ results

The CMS collaboration published its results of the search of $HH \rightarrow b\bar{b}\gamma\gamma$ based on the data collected by the CMS detector between 2015 and 2018 which corresponds to 137 fb^{-1} [136]. Similarly to ATLAS, no significant deviation from the background-only hypothesis was found. The analysis strategy is very different from the ATLAS one presented in this thesis. It leads to different limits and κ_λ measurement. Due to the differences in the analysis strategies, a direct comparison is not trivial. The results obtained by CMS on the limit are presented in Figure 6.14 and compared to ATLAS in Table 6.14. The κ_λ likelihood scan is shown in Figure 6.15. CMS results are extracted from a simultaneous fit of ($m_{\gamma\gamma}$, $m_{b\bar{b}}$) divided in fourteen analysis categories (on MVA output and $m_{b\bar{b}\gamma\gamma}^*$) defined using several MVA techniques and containing twelve ggF HH categories and two VBF HH dedicated categories. In contrary to ATLAS, CMS defines dedicated categories to target the VBF HH. In addition, CMS provides a specific orthogonal category to target $t\bar{t}H$ to set constraints on the κ_t modifier. Both HH and single Higgs boson are combined to provide an improved constraint of the κ_λ and κ_t parameters. The inclusion of the $t\bar{t}H$ significantly improves the constraint on κ_t as demonstrated in the 2D likelihood scan of κ_λ and κ_t in Figure 6.16. Besides this, CMS uses two machine learning based techniques to improve the b -jet energy:

- a DNN approach to improve the energy of resolution similarly to the method presented in Section 5.4.
- additional BDT approach which uses the fact that there is no missing transverse momentum from the hard-scattering process in the $HH \rightarrow b\bar{b}\gamma\gamma$ final state to correct the $m_{b\bar{b}}$ invariant mass which is similar to the kinematic fit described in Appendix D.

CMS uses a 2D fit of $m_{\gamma\gamma}$ and $m_{b\bar{b}}$, while ATLAS only $m_{\gamma\gamma}$ as the final discriminant variable. The 2D fit was considered in the ATLAS analysis but not implemented, since similar improvements were obtained between the 2D fit and including the $m_{b\bar{b}}$ in the MVA training.

CMS shows a slightly better expected cross-section limit than ATLAS, which could be explained by the dedicated VBF category. ATLAS shows a more precise κ_λ constraint which is related to the under-optimization done by CMS when defining the analysis ggF HH regions. In the CMS analysis, categories on $m_{b\bar{b}\gamma\gamma}^*$ are defined after the MVA is trained, while in ATLAS the training is done separately in each $m_{b\bar{b}\gamma\gamma}^*$ category with specific signal as described in Section 6.2.2. The signal used by CMS as training sample is a mixture of SM HH and 12 BSM benchmarks proposed in Ref. [137], but the mixing procedure is not explained in their published article. The training strategy used by ATLAS seems to improve the MVA performance in each $m_{b\bar{b}\gamma\gamma}^*$ region, giving better sensitivity to κ_λ which is mainly derived by the low mass category. As consequence, even with lower di-photon trigger threshold and better photon

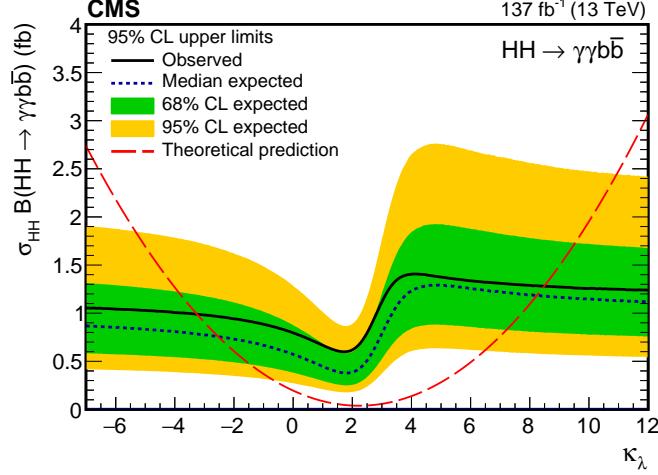


Figure 6.14: Expected and observed 95% CL upper limits on the $HH \rightarrow b\bar{b}\gamma\gamma$ cross-section obtained for different values of κ_λ assuming $\kappa_t = 1$. The green and yellow bands represent, respectively, the one and two standard deviation. The long-dashed red line shows the theoretical prediction.

	Expected	Observed
CMS $\sigma_{HH}/\sigma_{HH}^{SM}$ limit	5.2	7.7
CMS κ_λ interval	[-2.5, 8.2]	[-3.3, 8.5]
ATLAS $\sigma_{HH}/\sigma_{HH}^{SM}$ limit	5.5	4.1
ATLAS κ_λ interval	[-2.4, 7.7]	[-1.5, 6.7]

Table 6.14: Summary of ATLAS and CMS $HH \rightarrow b\bar{b}\gamma\gamma$ results. The limit is presented as a multiple of the SM expected value.

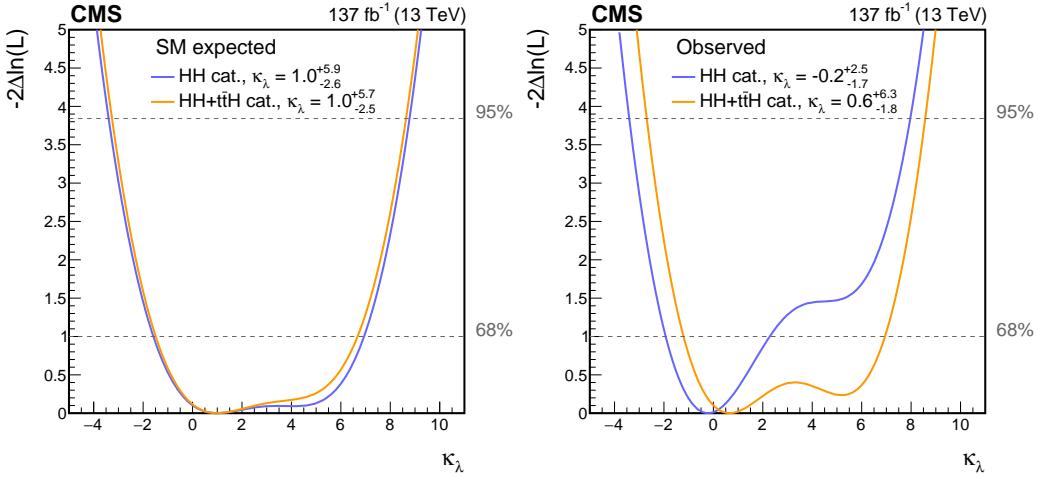


Figure 6.15: Negative log-likelihood, as a function of κ_λ , evaluated with an Asimov data set assuming the SM hypothesis. The 68 and 95% CL intervals are obtained from the intersection with the dashed grey lines. The two curves are shown for the HH (blue) and HH+ $t\bar{t}H$ (orange) analysis categories. All other couplings are set to their SM values.

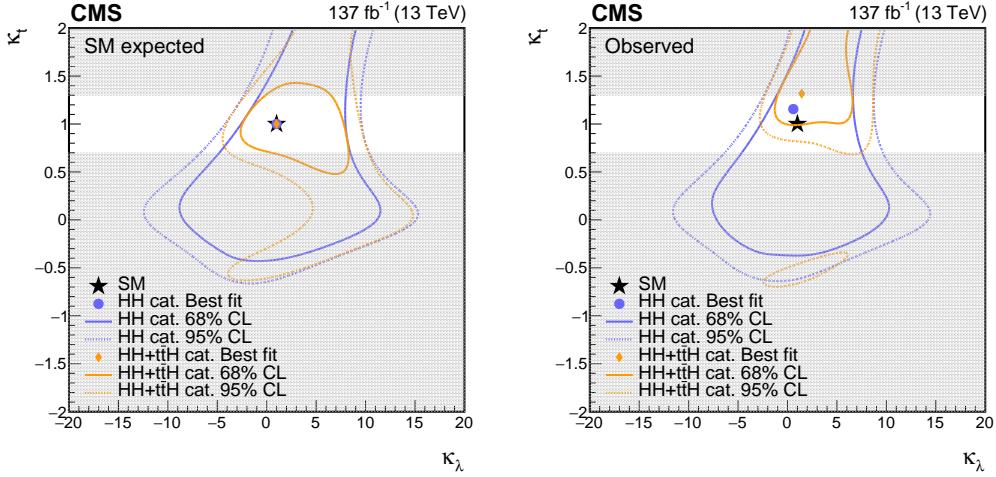


Figure 6.16: Negative log-likelihood contours at 68 and 95% CL in the $(\kappa_\lambda, \kappa_t)$ plane evaluated with an Asimov data set assuming the SM hypothesis (left) and the observed data (right). The contours obtained using the HH analysis categories only are shown in blue, and in orange when combined with the $t\bar{t}H$ categories. The best fit value for the HH categories only ($\kappa_\lambda = 0.6, \kappa_t = 1.2$) is indicated by a blue circle, for the HH + $t\bar{t}H$ categories ($\kappa_\lambda = 1.4, \kappa_t = 1.3$) by an orange diamond, and the SM prediction ($\kappa_\lambda = 1.0, \kappa_t = 1.0$) by a black star. The grey band shows the region where LO parametrization of κ_t used in CMS analysis is not reliable.

1945 energy resolution, CMS measurement is slightly less precise.

1946
1947 Since ATLAS and CMS results could be considered similar, the ATLAS+CMS $HH \rightarrow b\bar{b}\gamma\gamma$ combination
1948 is equivalent to twice the ATLAS only. Given this assumption, the ATLAS+CMS combination
1949 is approximated by scaling the ATLAS $HH \rightarrow b\bar{b}\gamma\gamma$ luminosity with a factor 2.

1950 The statistical-only expected limit on the HH cross-section for both ATLAS and ATLAS+CMS approx-
1951 imation is shown in Figure 6.17. The expected 95% CL upper limits on the HH cross-section for the SM
1952 case from the ATLAS+CMS approximation is 3.5 times the SM expectation and the expected allowed
1953 κ_λ interval is $[-1.2, 6.6]$ at 95% CL.

1954 6.9 Comparison to other HH decay channels

1955 As discussed in Section 2.5.1, other production modes are sensitive to the Higgs boson self-coupling. In
1956 the following, the final states $b\bar{b}b\bar{b}$, $b\bar{b}\tau^+\tau^-$ channels are discussed. When this thesis is written, the full
1957 Run-2 analyses of these two channels are not yet published. As a consequence, only the 2015-2016 data
1958 results will be considered.

1959 6.9.1 $b\bar{b}b\bar{b}$ channel

1960 In the $b\bar{b}b\bar{b}$ analysis, the two Higgs boson of the HH system are reconstructed from 4 b -tagged jets with
1961 an invariant mass of each pair of b -tagged jets close to m_H . Only 27.5 fb^{-1} of the data collected between
1962 2015 and 2016 was used due to an inefficiency in the online vertex reconstruction affecting b -jet triggers.
1963 The invariant mass m_{HH} of the two Higgs boson is used a discriminant variable. A large background
1964 remains after selection, mainly coming from QCD multi-jets (95%) and $t\bar{t}$ events (5%). A fully data-

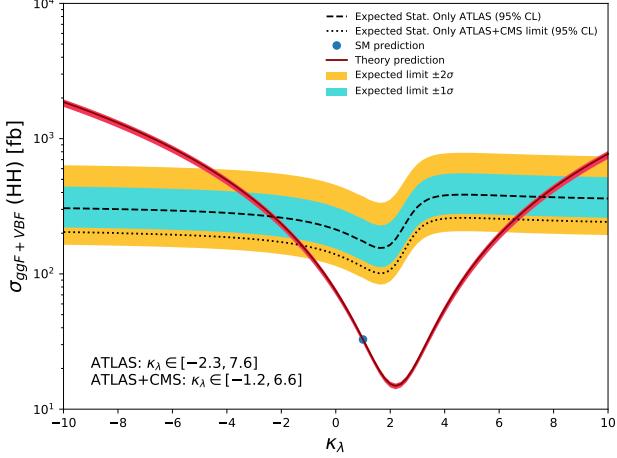


Figure 6.17: The stat. only expected 95% CL upper limits on the HH cross-section obtained for different values of κ_λ . Approximate ATLAS+CMS expected limits are obtained by scaling the ATLAS luminosity by a factor 2. The $\pm 1\sigma$ and $\pm 2\sigma$ bands are shown for the ATLAS expected limits.

¹⁹⁶⁵ driven method is used to estimate the QCD multi-jets and $t\bar{t}$ backgrounds. Even with a large branching
¹⁹⁶⁶ ratio ($\sim 33\%$), the $b\bar{b}b\bar{b}$ channel suffers from very large backgrounds. Table 6.15 summarizes the limit
on the HH cross-section and the constraint on the κ_λ .

	Expected	Observed
$\sigma_{HH}/\sigma_{HH}^{SM}$	21	12.9
κ_λ interval	[-11.6, 18.8]	[-10.9, 20.1]

Table 6.15: The 95% confidence limits of the HH production cross-section in $b\bar{b}b\bar{b}$ channel and κ_λ interval.

1967

6.9.2 $b\bar{b}\tau^+\tau^-$ channel

¹⁹⁶⁹ For the $b\bar{b}\tau^+\tau^-$ analysis, HH events are reconstructed from two b -tagged jets and two τ -leptons. The
¹⁹⁷⁰ analysis consists in splitting events into two categories according to the $\tau^+\tau^-$ decay channels:

- ¹⁹⁷¹ • the $\tau_{lep}\tau_{had}$ category, in which one of the two τ s is decaying leptonically and the other hadronically.
¹⁹⁷² A single lepton trigger (SLT) is used.
- ¹⁹⁷³ • the $\tau_{had}\tau_{had}$ in which both τ s are decaying hadronically. A SLT or di-tau triggers (DTT) are used.

¹⁹⁷⁴ The two categories are combined. A BDT is used to improve the separation between signal and back-
¹⁹⁷⁵ ground. Similarly to the $b\bar{b}b\bar{b}$ channel, the dominant background processes are the $t\bar{t}$, QCD multi-jets
¹⁹⁷⁶ and Z bosons produced in association with jets originating from heavy-flavour quarks as well as the
¹⁹⁷⁷ irreducible ZH background. Data-driven methods are used to estimate the background contamination.
¹⁹⁷⁸ The $b\bar{b}\tau^+\tau^-$ uses the BDT score as a discriminating variable to extract the results, which is less sensitive

1979 to κ_λ variations as shown in Figure 6.18. Table 6.16 summarizes the limit on the HH cross-section and
 1980 the constraint on the κ_λ parameter.

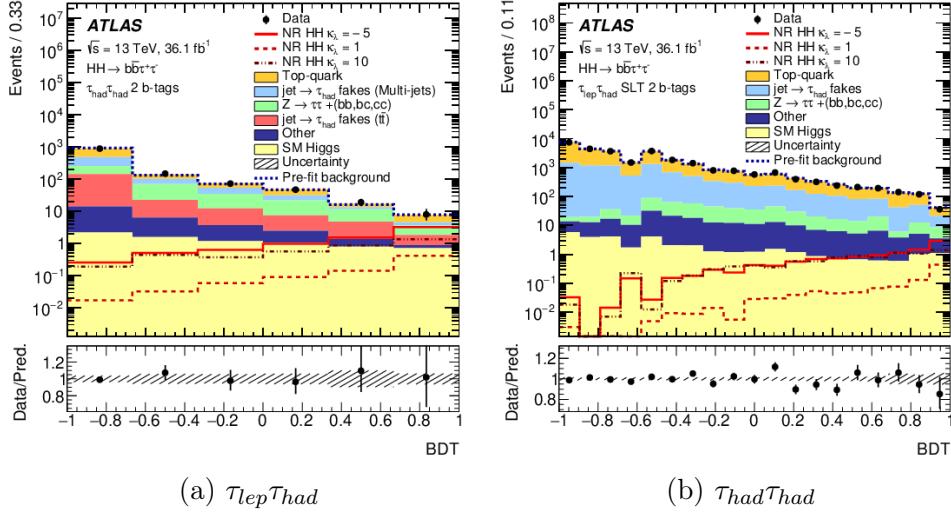


Figure 6.18: The BDT distributions in the $b\bar{b}\tau^+\tau^-$ analysis for the $\tau_{lep}\tau_{had}$ and $\tau_{had}\tau_{had}$ categories.

	Expected	Observed
$\tau_{lep}\tau_{had} \sigma_{HH}/\sigma_{HH}^{SM}$	28.4	23.5
$\tau_{had}\tau_{had} \sigma_{HH}/\sigma_{HH}^{SM}$	17.4	16.4
Combined $\sigma_{HH}/\sigma_{HH}^{SM}$	14.8	12.7
Combined κ_λ interval	[-8.9, 16.8]	[-7.4, 15.7]

Table 6.16: The 95% confidence limits on the HH production cross-section and κ_λ for the $b\bar{b}\tau^+\tau^-$ analysis.

1981 6.9.3 Combination

1982 A combination of the three channels for 36 fb^{-1} was published to set a better constraint on κ_λ and the
 1983 HH production cross-section [138]. In this combination, the $b\bar{b}b\bar{b}$ and $b\bar{b}\tau^+\tau^-$ channels were combined
 1984 with the 36 fb^{-1} $b\bar{b}\gamma\gamma$ to set the constraint on κ_λ parameter. Other channels are also considered for
 1985 the cross-section limit combination such as $b\bar{b}W^+W^-$, $b\bar{b}\gamma\gamma$ and $W^+W^-W^+W^-$ but with limited gain.
 1986 Figure 6.19 shows the upper limits for individual final states and their combination. The combined
 1987 observed (expected) upper limit on the SM HH production is 6.9 (10) times the SM prediction. The
 1988 sensitivity of the SM HH is driven by the $b\bar{b}b\bar{b}$, $b\bar{b}\tau^+\tau^-$ and $b\bar{b}\gamma\gamma$ modes which explains the motivation
 1989 to consider only these three channels for the κ_λ constraint.

1990 The upper limit as a function of κ_λ is shown in Figure 6.20. In the $b\bar{b}\tau^+\tau^-$ final state the observed
 1991 limit is more stringent than the expected limits over the whole range of κ_λ , due to a deficit in data
 1992 relative to the background predictions at high values of the BDT score. The $b\bar{b}\gamma\gamma$ limit shows a weaker
 1993 dependence on κ_λ than the $b\bar{b}b\bar{b}$ and $b\bar{b}\tau^+\tau^-$ limits because its acceptance is less dependent on κ_λ . The
 1994 combination results are summarized in Table 6.17.

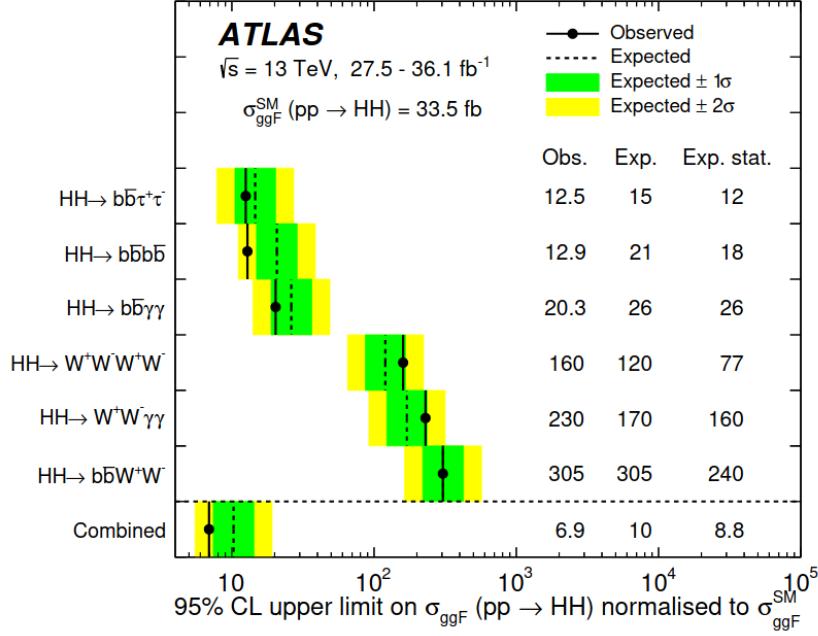


Figure 6.19: Upper limits at 95% CL on the ggF SM HH production normalized to its SM expectation from the $b\bar{b}\tau^+\tau^-$, $b\bar{b}b\bar{b}$, $b\bar{b}\gamma\gamma$, $\bar{b}bW^+W^-$, $\bar{b}b\gamma\gamma$ and $W^+W^-W^+W^-$ decay channels, and their combination.

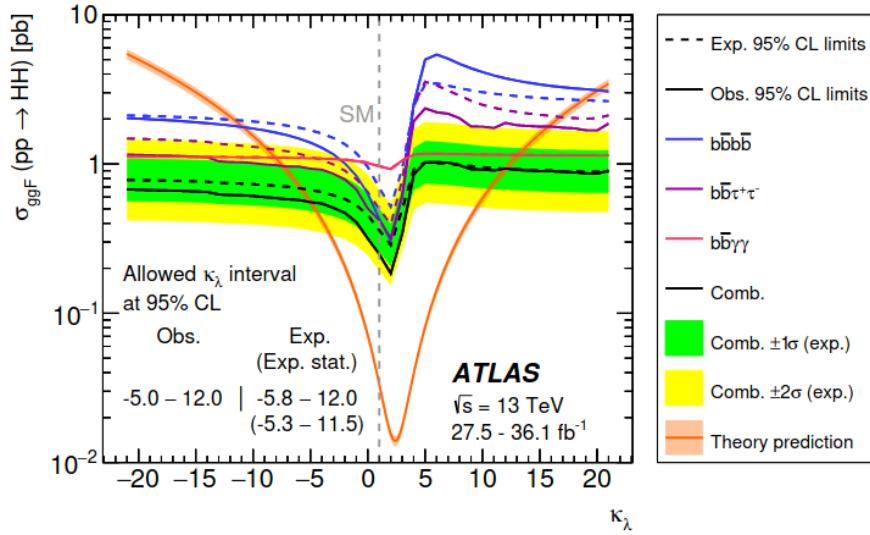


Figure 6.20: Upper limits at 95% CL on the cross-section of the ggF SM HH production as a function of κ_λ .

	Observed	Expected
$\sigma_{gg \rightarrow HH}/\sigma_{gg \rightarrow HH}^{\text{SM}}$	6.9	10
Allowed κ_λ interval	[-5.0, 12.0]	[-5.8, 12.0]

Table 6.17: The 95% CL observed and expected limits on the Higgs boson pair cross-section normalized to the SM production cross-section and the allowed κ_λ interval.

1995 6.10 Conclusion

1996 The full Run-2 $HH \rightarrow b\bar{b}\gamma\gamma$ analysis sets the best constraint to date on the Higgs boson self-coupling.
1997 The previous results are improved by a factor of 2.4 on top of the improvement due to the integrated
1998 luminosity increase. The improvement is mainly coming from the analysis strategy by dividing events
1999 in sub-categories of m_{HH} , implementing machine learning to increase separation between signal and
2000 background and improving the b -jet energy resolution. The developed b -jet energy calibration leads to
2001 an improvement of 10% on the analysis significance. The presented full Run-2 $HH \rightarrow b\bar{b}\gamma\gamma$ analysis sets
2002 a slightly better constraint than the full Run-2 CMS analysis and the combined 36 fb^{-1} results. No HH
2003 event is observed in the full Run-2 data, neither by ATLAS nor CMS. The observed limit at 95% CL
2004 on HH cross-section is set to 4.2 times the SM prediction and an allowed κ_λ interval to [-1.5, 6.7]. The
2005 results are better than the 36 fb^{-1} combining all decay channels. New data taking period is foreseen at
2006 the LHC Run-3 (and beyond) in which an additional integrated luminosity of 300 fb^{-1} is expected with
2007 the same pile up conditions. The HL-LHC program with 20 times more data than Run-2 as described
2008 in Section 3.3. The next chapter is dedicated to discuss the possible improvements and prospective for
2009 the HH search at Run-3 and HL-LHC.

2010 Chapter 7

2011 Prospects of Higgs boson pair production at 2012 Run-3 and HL-LHC

2013 With the full Run-2 data, the 95% CL interval for the κ_λ modifier is $-1.5 < \kappa_\lambda < 6.7$ using the $HH \rightarrow b\bar{b}\gamma\gamma$
2014 decay channel. At the end of Run-3, the integrated luminosity is expected to reach 300 fb^{-1} . At the
2015 end of HL-LHC, the integrated luminosity should reach 3000 fb^{-1} at a centre-of-mass energy of 14 TeV.
2016 As the measurement is statistically limited and will remain so even at HL-LHC, the Higgs boson pair
2017 production search will benefit from the integrated luminosity increase and results presented at the last
2018 European strategy predict to achieve a precision of 50% on κ_λ at 68% CL at the end of HL-LHC [139] by
2019 combining both ATLAS and CMS extrapolations. New prospects for the measurement of the trilinear
2020 Higgs boson self-coupling at Run-3 and HL-LHC will be presented in this chapter.

2021 7.1 Potential $HH \rightarrow b\bar{b}\gamma\gamma$ improvements from analysis tools

2022 Several improvements to the $HH \rightarrow b\bar{b}\gamma\gamma$ analysis strategy are possible in the next round. Significant
2023 enhancement would be achieved if we consider the different items presented in this thesis (not applied
2024 in the Run-2 publication):

- 2025 Even with fewer variables, the DNN categorization described in Section 6.2.3 shows promising
2026 results. Considering this DNN at next round with the addition of the topness, HT and missing
2027 transverse energy variables, around 10% improvement in the sensitivity of the analysis should be
2028 achievable.
- 2029 Additional improvement to the $m_{b\bar{b}}$ resolution can be achieved with the usage of the Kinematic
2030 Fit (KF), described in Appendix D. The KF brings an additional 10% improvement to the $m_{b\bar{b}}$
2031 resolution which can be translated to 2-5% improvement on the analysis significance.

2032 7.1.1 Photon identification with CNN in $HH \rightarrow b\bar{b}\gamma\gamma$

2033 Improving photon identification (PID) can have important impact for $HH \rightarrow b\bar{b}\gamma\gamma$. The CNN-based
2034 algorithm introduced in Section 4.4 provides an improved PID efficiency. In the $HH \rightarrow b\bar{b}\gamma\gamma$ channel,
2035 this gain leads to an enhancement of the signal acceptance. The developed CNN is applied to photons
2036 from the $HH \rightarrow b\bar{b}\gamma\gamma$ simulated events. Figure 7.1 shows the efficiencies of the Tight WP for both
2037 the cut-based and the CNN algorithms as a function of photon energy and pseudorapidity. The CNN
2038 over-performs the cut-based and shows an efficiency close to 100% over the full energy spectrum.

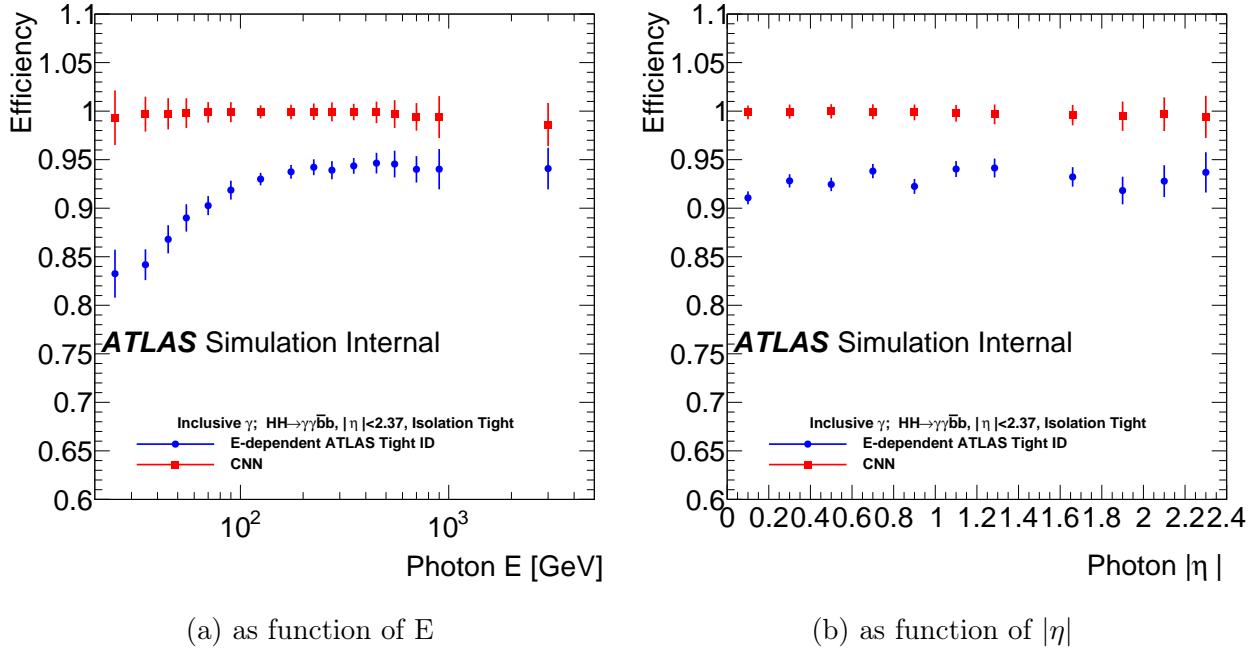


Figure 7.1: Photon identification efficiency of cut-based Run-2 ID algorithm (blue) and the CNN ID algorithm (red) as function of photon energy and $|\eta|$ from $HH \rightarrow b\bar{b}\gamma\gamma$ simulated events. Tight isolation is applied as baseline.

2039 The increase in the PID efficiency for each photon can be translated to $\sim 15\%$ improvement in the
 2040 signal acceptance as shown in the $m_{\gamma\gamma}$ distribution in Figure 7.2. This is also true for single Higgs boson
 2041 backgrounds.

2042
 2043 For the continuum $\gamma\gamma + \text{jets}$ background, the PID efficiency is shown in Figure 7.3. The same im-
 2044 provement can be achieved for the continuum background. Figure 7.4 shows the $m_{\gamma\gamma}$ spectrum with
 2045 an improvement of 15% in the statistics. With the CNN algorithm, the continuum statistics can be
 2046 enhanced by 15% which will slightly improve the background modelling, thus reduce the spurious signal
 2047 systematics and improve the analysis significance with approximately 7.3%. The b -jet requirement in
 2048 $HH \rightarrow b\bar{b}\gamma\gamma$ leads to a high $\gamma\gamma$ purity as described in Section 6.3.2. For the $H \rightarrow \gamma\gamma$ analysis, an
 2049 improvement in the $\gamma\gamma$ purity is expected with the CNN algorithm.

2050 7.2 Expected results at the end of Run-3

2051 Some detector upgrades are planned for the Run-3, but most of them do not have a significant impact
 2052 on the $HH \rightarrow b\bar{b}\gamma\gamma$ election efficiency. As mentioned before, the expected integrated luminosity col-
 2053 lected at Run-3 is around 300 fb^{-1} . Given the expected similarity in the LHC beams and data taking
 2054 conditions, a simple extrapolation from the Run-2 to Run-3 is reasonable. The extrapolation is set by
 2055 scaling the full Run-2 workspace to $300/139 \sim 2.16$. To allow for an easy extrapolation, impacts from
 2056 reconstruction algorithms are neglected. Given the assumption that $HH \rightarrow b\bar{b}\gamma\gamma$ analysis is statistically
 2057 dominated, no systematic uncertainties are considered.

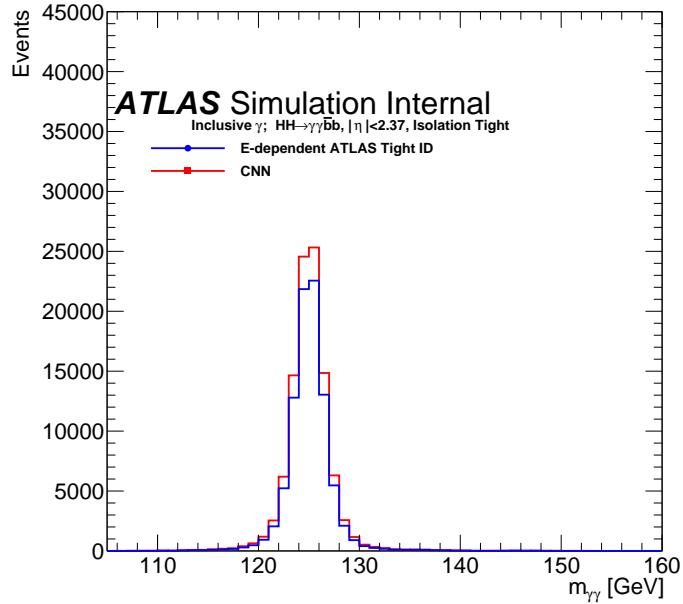


Figure 7.2: $m_{\gamma\gamma}$ invariant mass distribution for two selected photons with cut-based algorithm (blue) and CNN algorithm (red).

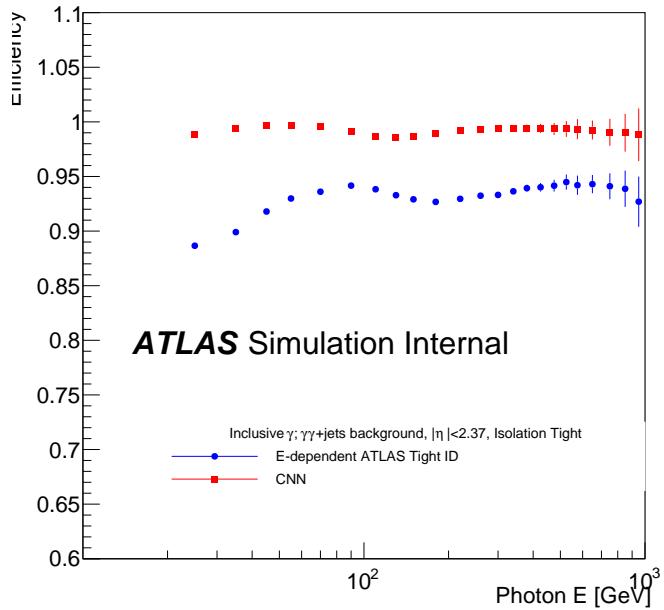


Figure 7.3: Photon identification efficiency of cut-based Run-2 ID algorithm (blue) and the CNN ID algorithm (red) as function of photon energy from $\gamma\gamma + \text{jets}$ simulated events. Tight isolation is applied as baseline.

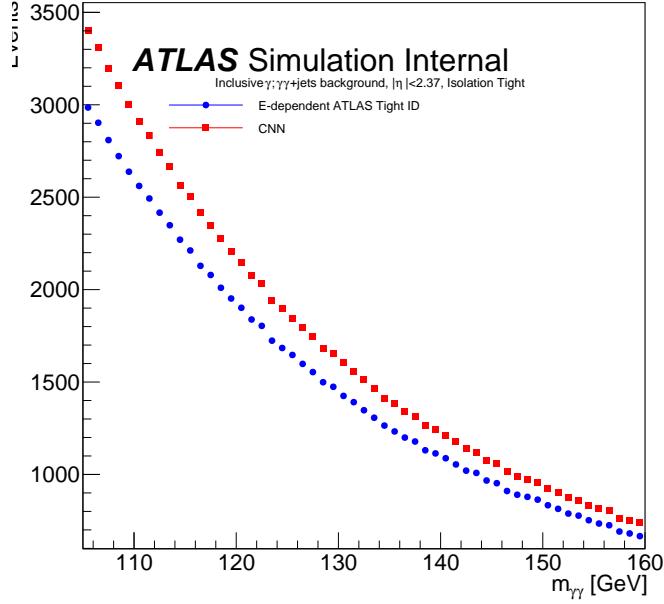


Figure 7.4: $m_{\gamma\gamma}$ invariant mass distribution of the continuum $\gamma\gamma+\text{jets}$ for two selected photons with cut-based algorithm (blue) and CNN algorithm (red).

2059 The expected upper 95% CL limits on the Higgs boson pair production cross-section as a function of
 2060 the Higgs boson self-coupling κ_λ is shown in Figure 7.5 for both full Run-2 and extrapolation to Run-3.
 2061 The expected limit for the SM at $\kappa_\lambda = 1$ is found to be 3.8 times SM expectations which is 1.4 lower.
 2062 The expected κ_λ interval at 95% CL from the limit scan is [-1.1, 6.5].

2063 An Asimov dataset with the background-only hypothesis was created, and maximum likelihood fits
 2064 of this dataset were performed with different κ_λ hypotheses. The negative logarithm of the ratio of
 2065 the maximum likelihood for κ_λ to the fit with $\kappa_\lambda = 1$ is shown in Figure 7.6 for both Run-2 and Run-
 2066 2+Run-3. The 1σ and the 2σ confidence interval (CI) constraints on κ_λ are shown in Table 7.1. It can
 2067 be seen in Figure 7.6 that there are typically two minima. The first minimum is located at $\kappa_\lambda = 1$ as
 2068 the signal hypothesis used to create the Asimov dataset. The second minimum is observed at a κ_λ value
 2069 that corresponds to a identical cross-section with respect to the κ_λ point at the first minimum. The
 2070 measured signal strength is 1.0 ± 1.52 considering only statistical uncertainties. The expected median
 2071 significance of the SM HH signal relative to the background-only hypothesis is 0.71σ .

Scenario	1σ CI	2σ CI
Run-2 Stat. Only	[-1.3, 6.4]	[-2.9, 8]
Run-2+Run-3 Stat. Only	[-0.7, 5.6]	[-1.9, 7.0]

Table 7.1: Constraints on κ_λ from the likelihood ratio test performed on the Asimov dataset created from the background and the $\kappa_\lambda = 1$ signal, as shown in Figure 7.6.

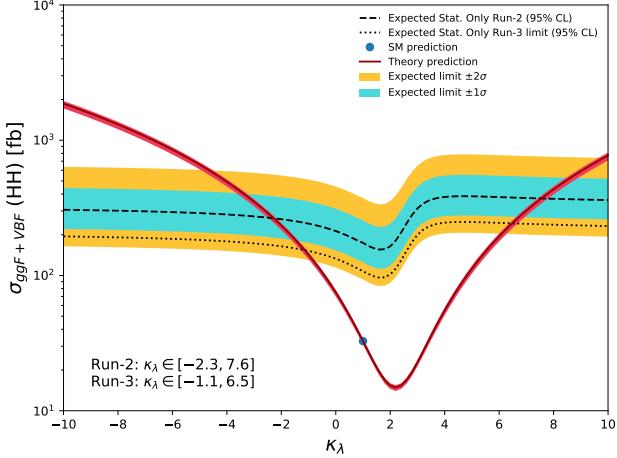


Figure 7.5: The 95% CL expected upper limits on the HH production cross-section as function of κ_λ . The ± 1 and ± 2 bands are shown for the Run-2 results. The expected allowed κ_λ intervals are also reported.

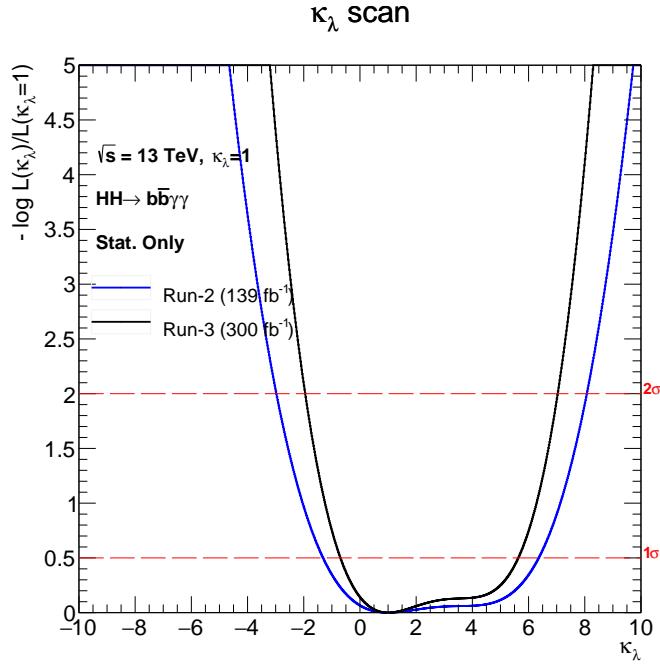


Figure 7.6: Negative natural logarithm of the ratio of the maximum likelihood to the maximum likelihood for $\kappa_\lambda = 1$. The dashed lines at 0.5 and 2.0 indicate the values corresponding to a 1σ and 2σ confidence intervals, respectively (assuming an asymptotic χ^2 distribution of the test statistic).

7.3 Prospects at HL-LHC

As described in Section 3.3, the HL-LHC is a planned upgrade of the LHC accelerator to deliver an integrated luminosity of 3000 fb^{-1} in ten years. It should be operated at $\sqrt{s} = 14 \text{ TeV}$ and will start after the long shut-down 3 (LS3) finishing in 2027. An upgrade of the ATLAS detector and trigger system are planned are assumed to compensate the higher pileup rate which will be around 200 collisions per bunch crossing. Due to the higher pileup, photon and jet energy resolutions are expected to slightly degrade, but this should be compensated by better reconstruction techniques. With the ITk, the b -tagging efficiency will be improved by 8% per jet, while remaining the same rejection efficiency. A prospects study of Higgs boson pair production at HL-LHC was performed in the context of the European strategy in 2019. In this study, three HH decay channels were considered : $b\bar{b}b\bar{b}$, $b\bar{b}\tau^+\tau^-$ and $b\bar{b}\gamma\gamma$. For $b\bar{b}b\bar{b}$ and $b\bar{b}\tau^+\tau^-$ the 36 fb^{-1} analysis were kept unchanged and extrapolated to 3000 fb^{-1} at 14 TeV , while a completely new truth simulation-based analysis was performed for $b\bar{b}\gamma\gamma$. In the following the European strategy results for $HH \rightarrow b\bar{b}\gamma\gamma$, its combination with $b\bar{b}b\bar{b}$ and $b\bar{b}\tau^+\tau^-$ and new extrapolation of the full Run-2 $HH \rightarrow b\bar{b}\gamma\gamma$ analysis to HL-LHC integrated luminosity are presented.

7.3.1 Estimation for European strategy 2019 of $b\bar{b}\gamma\gamma$

As mentioned above, the prospect of $HH \rightarrow b\bar{b}\gamma\gamma$ at HL-LHC was performed through the use of truth-level Monte Carlo samples generated at $\sqrt{s} = 14 \text{ TeV}$. To emulate the response of the upgraded ATLAS detector at the HL-LHC, the final state particles at truth level are smeared assuming a pile-up density of 200 at HL-LHC. In contrary to the 36 fb^{-1} analysis, the projection improves the sensitivity by including a BDT categorization. The BDT is trained to separate the SM HH signal from the backgrounds. The BDT response for signal and background is shown in Figure 7.7.

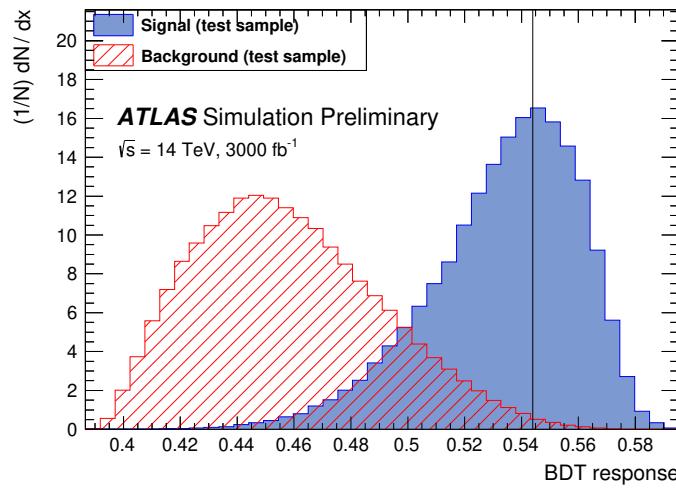


Figure 7.7: BDT response for signal and background. The vertical line denotes the optimal cut on the BDT response that maximizes the statistical-only significance,

A BDT cut is set at 0.54 to define a single category that maximizes the computed significance accounting for statistical uncertainties only. The BDT selects around 40% of the SM HH signal, while it rejects almost 99% of backgrounds. The analysis uses the $m_{\gamma\gamma\bar{b}b}$ shape fit in the $m_{\gamma\gamma} \in [123, 127] \text{ GeV}$ window to extract the results. The sensitivity to κ_λ is enhanced by defining eight bins in the $m_{\gamma\gamma\bar{b}b}$ distribution. The negative log-likelihood ratio performed on an Asimov dataset generated with $\kappa_\lambda = 1$

is shown in Figure 7.8. Table 7.2 summarizes constraints on κ_λ from the likelihood ratio test. It shows that even with 3000 fb^{-1} the analysis is still statistically dominated.

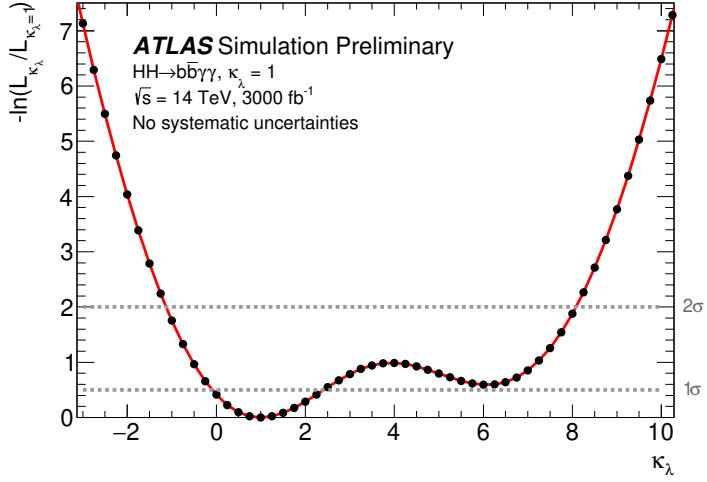


Figure 7.8: Negative natural logarithm of the ratio of the maximum likelihood for κ_λ to the maximum likelihood for $\kappa_\lambda = 1$ for the fit with only statistical uncertainties.

2099

Scenario	1σ CI	2σ CI
No systematic uncertainties	$-0.1 < \kappa_\lambda < 2.4$	$-1.1 < \kappa_\lambda < 8.1$
Systematic uncertainties included	$-0.2 < \kappa_\lambda < 2.5$	$-1.4 < \kappa_\lambda < 8.2$

Table 7.2: Constraints on κ_λ from the likelihood ratio test performed on the Asimov dataset created from the backgrounds and the SM HH signal.

2100 7.3.2 Combination with $b\bar{b}b\bar{b}$ and $b\bar{b}\tau^+\tau^-$ channels

An extrapolation of the 36 fb^{-1} analysis of $b\bar{b}b\bar{b}$ and $b\bar{b}\tau^+\tau^-$ channels to HL-LHC was also performed (More details in Ref. [139]). Their statistical combination is performed by constructing a combined likelihood function that takes into account signal and background models and correlated systematic uncertainties from all channels. Figure 7.9 shows the combined likelihood scan on an Asimov dataset. The combined significance for SM HH is evaluated to be 3.5σ without systematic uncertainties and 3σ including systematic and its mainly driven by $b\bar{b}\tau^+\tau^-$ and $b\bar{b}\gamma\gamma$ channels as shown in Figure 7.10. The significance depends on the expected signal yield, and therefore it is lower for those κ_λ values for which the cross-section and the acceptance times efficiency is low. The measured combined signal strength is 1.0 ± 0.4 including systematics and 1.0 ± 0.3 without systematics. Tables 7.3 and 7.4 show the significance and the signal strength measured in the individual HH channels, respectively. Table 7.5 summarizes the constraints on κ_λ from the likelihood test performed on the Asimov dataset created from the backgrounds and HH signal with $\kappa_\lambda = 1$.

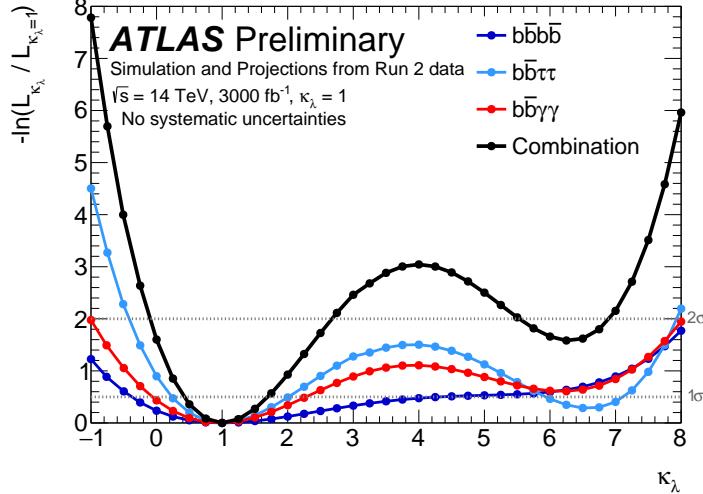


Figure 7.9: Negative natural logarithm of the ratio of the maximum likelihood for κ_λ to the maximum likelihood for $\kappa_\lambda = 1$ for the fit with only statistical uncertainties. The black circles show the results for the combination, while the coloured markers show the values coming from the individual channels.

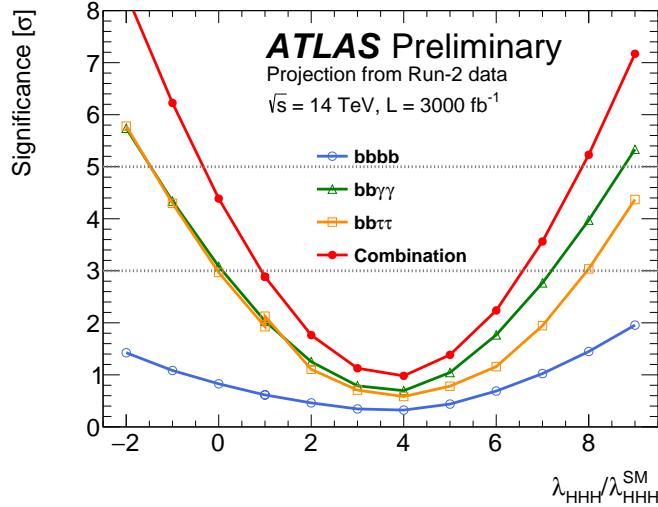


Figure 7.10: Expected significance of observing Higgs boson pair production as a function of κ_λ without any systematic uncertainties. The two horizontal dashed lines show the 3σ and 5σ thresholds.

Channel	Statistical-only	Statistical + Systematic
$HH \rightarrow bbbb$	1.4	0.61
$HH \rightarrow b\bar{b}\tau^+\tau^-$	2.5	2.1
$HH \rightarrow b\bar{b}\gamma\gamma$	2.1	2.0
Combined	3.5	3.0

Table 7.3: Significance of the individual $b\bar{b}b\bar{b}$, $b\bar{b}\tau^+\tau^-$ and $b\bar{b}\gamma\gamma$ channels as well as their combination.

Channel	Measured μ (Statistical-only)	Measured μ (Statistical + Systematic)
$HH \rightarrow bbbb$	1.0 ± 0.6	1.0 ± 1.6
$HH \rightarrow b\bar{b}\tau^+\tau^-$	1.0 ± 0.4	1.0 ± 0.5
$HH \rightarrow b\bar{b}\gamma\gamma$	1.0 ± 0.6	1.0 ± 0.6
Combined	1.0 ± 0.3	1.0 ± 0.4

Table 7.4: Signal strength measured in the individual channels and their combination using an Asimov dataset with SM HH signal injected.

Scenario	1σ CI		2σ CI		
	Statistical uncertainties only	$0.4 \leq \kappa_\lambda \leq 1.7$	$-0.10 \leq \kappa_\lambda \leq 2.7 \cup 5.5 \leq \kappa_\lambda \leq 6.9$	Systematic uncertainties	$0.25 \leq \kappa_\lambda \leq 1.9$

Table 7.5: Constraints on κ_λ from the likelihood ratio test performed on the Asimov dataset created from the backgrounds and the SM HH signal. Results are presented as 1σ and 2σ CI on κ_λ .

2113 7.4 New extrapolation of full Run-2 $HH \rightarrow b\bar{b}\gamma\gamma$ analysis

2114 An increase of the analysis sensitivity is expected from further analysis strategy optimization, while the
 2115 global developments of the analysis are hard to predict over many years of R&D. The extrapolation
 2116 aims at giving a realistic estimation of the expected sensitivity gain using the current state-of-the-art
 2117 analysis methods.

2118 To evaluate the performance of the new Run-2 $HH \rightarrow b\bar{b}\gamma\gamma$ analysis at HL-LHC, an extrapolation
 2119 to 3000 fb^{-1} is performed. The assumption that the planned upgrades to the ATLAS detector and
 2120 the improvements to reconstruction algorithms will mitigate effects of higher pile-up, resulting in the
 2121 same performance as the Run-2 detector, is assumed. For simplicity, the 139 fb^{-1} analysis strategy is
 2122 kept unchanged and a simple integrated luminosity scaling is performed. No systematic uncertainty is
 2123 considered since the European strategy study showed that $b\bar{b}\gamma\gamma$ analysis still be statistically dominated
 2124 even at HL-LHC. The effect of the planned energy increase to 14 TeV on the cross-section is considered.
 2125 Scale factors are applied to account for the increase from 13 TeV to 14 TeV [131]. The values of these
 2126 scale factors are listed in Table 7.6. An average scaling of $+18\%$ is used for the continuum background
 2127 normalization.

Process	ggF HH	VBF HH	ggF H	VBF H	$t\bar{t}H$	WH	ZH	tHjb	tWH
$\sigma_{14\text{TeV}}/\sigma_{13\text{TeV}}$	1.19	1.2	1.13	1.13	1.21	1.1	1.11	1.21	1.22

Table 7.6: Scale factors between the production cross-sections at 13 TeV (LHC Run 2) and 14 TeV (HL-LHC).

2128 An Asimov dataset with the SM signal+background hypothesis was created, and maximum likelihood
 2129 fits of this dataset were performed with different κ_λ hypotheses. The negative logarithm of the ratio
 2130 of the maximum likelihood for κ_λ to the fit with $\kappa_\lambda = 1$ is shown in Figure 7.11. The 1σ and the 2σ
 2131 confidence interval (CI) constraints on κ_λ from the full Run-2 analysis, the European strategy and the
 2132 extrapolation of the full Run-2 analysis to HL-LHC are shown in Table 7.7.

2133 The expected median significance of the SM HH signal relative to the background-only hypothesis
 2134 for the individual category and their combination is shown in Table 7.8. The precision on signal strength
 2135 is 44%.

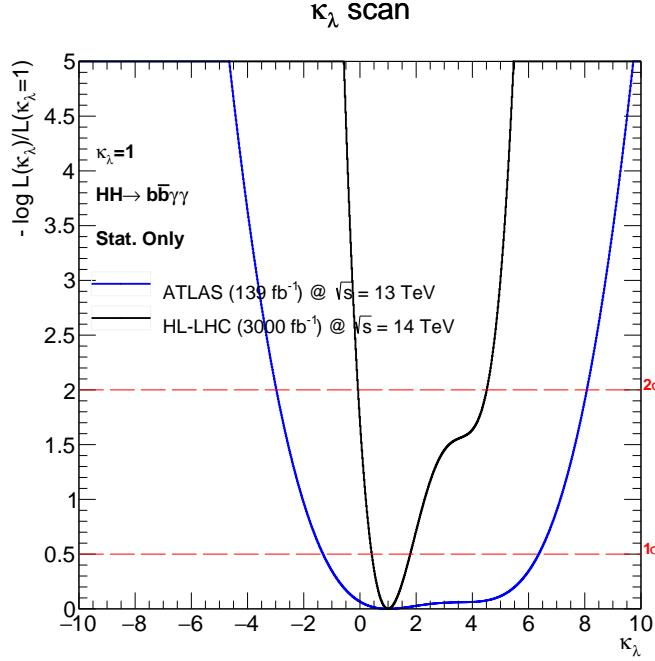


Figure 7.11: Negative natural logarithm of the ratio of the maximum likelihood for κ_λ to the maximum likelihood for $\kappa_\lambda = 1$. The dashed lines at 0.5 and 2.0 indicate the values corresponding to a 1σ and 2σ confidence intervals, respectively (assuming an asymptotic χ^2 distribution of the test statistic).

Scenario	1σ CI	2σ CI
Run-2 Stat. Only	[-1.3, 6.4]	[-2.9, 8]
European Strategy Stat. Only	[-0.1, 2.4]	[-1.1, 8.1]
Extrapolation from to HL-LHC Stat. Only	[0.4, 1.8]	[-0.1, 4.4]

Table 7.7: Constraints on κ_λ from the likelihood ratio test performed on the Asimov dataset created from the background and the $\kappa_\lambda = 1$ signal, as shown in Figure 7.11.

Scenario	Run-2 Stat. Only	HL-LHC Stat. Only
High mass, High BDT	0.47	2.38
High mass, Low BDT	0.13	0.64
Low mass, High BDT	0.03	0.15
Low mass, Low BDT	0.02	0.10
Combined	0.48σ	2.47σ

Table 7.8: The Statistical-only significance of the individual categories as well as their combination.

The HL-LHC extrapolation of the Run-2 analysis shows similar performance for the SM HH to the European strategy results while it shows a better κ_λ constraint. This is mainly driven by the BDT strategy. As mentioned before, the European strategy study used a single BDT trained on SM HH and define a single signal region. By construction, this signal region is equivalent to the High mass, High BDT category defined for full Run-2 analysis from which the significance is mostly driven. This explains the similarity in analysis performance. For κ_λ sensitivity, having a dedicated BDT targeting BSM scenarios trained on $\kappa_\lambda = 10$ HH signal, mostly improves the full Run-2 sensitivity for κ_λ variations. This explains the improved κ_λ constraint. Figure 7.12 shows the significance as a function of κ_λ from the extrapolation of full Run-2 analysis and the European strategy results.

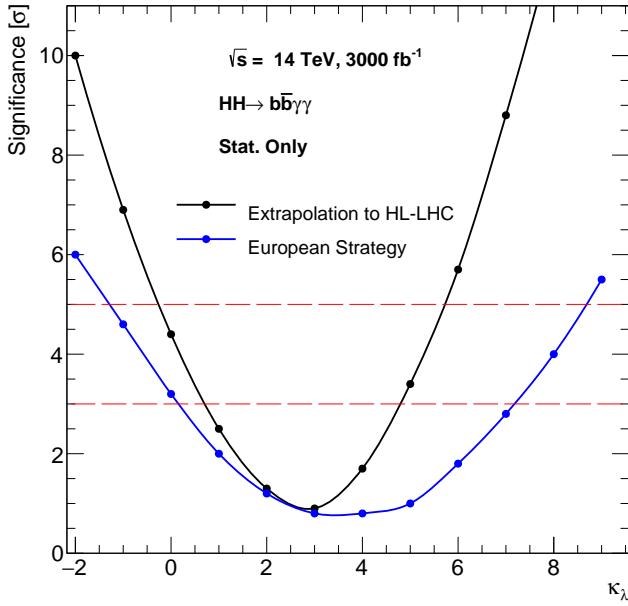


Figure 7.12: Expected statistic-only significance of observing Higgs boson pair production as a function of κ_λ from the European Strategy $b\bar{b}\gamma\gamma$ (blue) and the extrapolation of the full Run-2 $b\bar{b}\gamma\gamma$ to HL-LHC (black). The European Strategy numbers are taken from Figure 7.10.

2144

2145 7.5 Conclusion

A prospect study for the search for non-resonant Higgs boson pair production at the HL-LHC has been performed, using the $b\bar{b}b\bar{b}$, $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau^+\tau^-$ channels and their combination. From the European strategy combination, the signal strength is expected to be measured with an accuracy of 30% and the κ_λ is constrained at 95% to $[-0.1, 2.7] \cup [5.5, 6.9]$. The full Run-2 $HH \rightarrow b\bar{b}\gamma\gamma$ analysis presented in this thesis is statistically extrapolated to HL-LHC. From 139 fb^{-1} $HH \rightarrow b\bar{b}\gamma\gamma$ extrapolation, the signal strength is expected to be measured with an accuracy of 44% and the Higgs boson self-coupling is constrained at 95% CL to $[-0.1, 4.4]$ and measured with a precision of order of 60% at 68% CI. Improving the sensitivity of the $HH \rightarrow b\bar{b}\gamma\gamma$ in the future through the enhancement of $m_{b\bar{b}}$ resolution, the analysis strategy and photon identification would be beneficial.

²¹⁵⁵ Chapter 8

²¹⁵⁶ Conclusion

²¹⁵⁷ In this thesis, a search for the Standard Model (SM) Higgs boson pair production and new physics
²¹⁵⁸ phenomena in its decay to two photons and two b -quarks channel is presented. The search is based
²¹⁵⁹ on the full Run-2 dataset recorded by the ATLAS detector between 2015 and 2018 at a centre-of-mass
²¹⁶⁰ energy of $\sqrt{s} = 13$ TeV provided by the Large Hadron Collider (LHC) leading to a total of the integrated
²¹⁶¹ luminosity of 139 fb^{-1} .

²¹⁶²
²¹⁶³ The search is restricted to the non-resonant mode as a way to test the sensitivity to the SM process
²¹⁶⁴ and new physics that could manifest as modifications of the Higgs boson self-coupling. Modifications
²¹⁶⁵ of the Higgs boson self-coupling are quantified through the trilinear κ_λ parameter. Despite the small
²¹⁶⁶ branching ratio (0.3%) of the $b\bar{b}\gamma\gamma$ decay channel, it is a "golden" channel since it compromises between
²¹⁶⁷ the $H \rightarrow b\bar{b}$ large branching ratio and a clean di-photon signature in the detector and its smoothly falling,
²¹⁶⁸ continuous backgrounds.

²¹⁶⁹
²¹⁷⁰ To perform such analysis, an excellent identification of photon is needed. Such identification is cur-
²¹⁷¹ rently performed on high-level shower shape variables. A new photon identification algorithm based on
²¹⁷² advanced Machine Learning (ML) technique is introduced. This algorithm is based on the Convolutional
²¹⁷³ Neural Network (CNN) technique to learn prompt photon shower shape from electromagnetic calorime-
²¹⁷⁴ ter images. With this algorithm, a photon identification efficiency of 85-95% is obtained. The algorithm
²¹⁷⁵ is validated in different event topology and scale factors are extracted using radiative $Z \rightarrow \gamma l^- l^+$ method.
²¹⁷⁶ In addition, a cell-by-cell reweighting technique is presented to correct the discrepancy between real and
²¹⁷⁷ simulated events. The sources of this discrepancy is still remaining unclear, with many sources can
²¹⁷⁸ contribute to this mis-modelling, mainly from the detector geometry, material distribution and electro-
²¹⁷⁹ magnetic shower modelling.

²¹⁸⁰
²¹⁸¹ As the b -jet energy is under-estimated which degrades $m_{b\bar{b}}$ resolution, excellent calibration of the b -jet
²¹⁸² energy scale and resolution is also needed for this search. A jet-by-jet calibration technique specific for
²¹⁸³ b -jet is presented. The method attempts to correct the b -jet from the presence of muons and neutrinos
²¹⁸⁴ in its semi-leptonic decay and the out-of-cone radiation effects. With this method, an improvement of
²¹⁸⁵ 23% in $m_{b\bar{b}}$ resolution is achieved and an enhancement of 10% of the $HH \rightarrow b\bar{b}\gamma\gamma$ analysis sensitivity is
²¹⁸⁶ obtained. The related systematic to this calibration method is also studied and found to be negligible.
²¹⁸⁷ The b -jet calibration is complemented with a kinematic fit to constrain the di-Higgs boson system bal-
²¹⁸⁸ ance using the well-reconstructed $H \rightarrow \gamma\gamma$, which bring an additional 10% improvement in $m_{b\bar{b}}$ resolution.

²¹⁸⁹
²¹⁹⁰ The results of this search are presented as an upper limit on the cross-section as a function of the κ_λ

2191 modifier. The observed (expected) upper limit on the Higgs boson pair production cross-section of 4.1
 2192 (5.5) times the SM expectation which corresponds to an observed (expected) cross-section upper limit
 2193 140 (180) fb, assuming the SM couplings. The allowed observed (expected) interval at 95% CL obtained
 2194 with this search is $-1.5 < \kappa_\lambda < 6.7$ ($-2.4 < \kappa_\lambda < 7.7$). They are the best upper limit on the Higgs boson
 2195 pair production and κ_λ constraint up to date. This search is improved with a factor of 3 on top of the
 2196 luminosity increase with respect to the previous search performed on subsets of the Run-2 dataset of 36
 2197 fb^{-1} . The improvement of the sensitivity of the search presented in this thesis is the result of the use
 2198 of Multi-Variate Analysis (MVA) techniques to target the HH signal and the improvement of the b -jet
 2199 energy resolution as well as the categorization of the event over the m_{HH} invariant distribution which
 2200 enhances the sensitivity to κ_λ measurement. The results are compared to the one obtained by CMS
 2201 collaboration in the same decay channel with the same amount of data. CMS reported an observed
 2202 (expected) upper limit on the SM HH production of 7.7 (5.2) and an observed (expected) 95% CL
 2203 constrain on κ_λ parameter of $-3.3 < \kappa_\lambda < 8.5$ ($-2.5 < \kappa_\lambda < 8.2$). One should mention that the results
 2204 are statistically dominated and the effect of systematic uncertainties is around 2%.

2205 Since the published ATLAS and CMS results are similar, an expected upper limit and κ_λ constraint
 2206 from the statistical combination of the ATLAS and CMS full Run-2 $HH \rightarrow b\bar{b}\gamma\gamma$ results is performed.
 2207 The expected allowed 95% CL interval of κ_λ from the extrapolation is $-1.2 < \kappa_\lambda < 6.6$. Moreover,
 2208 prospect studies of the $HH \rightarrow b\bar{b}\gamma\gamma$ search at the Run-3 and HL-LHC are performed in this thesis. The
 2209 extrapolation to Run-2+Run-3 analysis is done by assuming an additional 300 fb^{-1} . This extrapolation
 2210 leads to an expected upper limit on the SM HH production cross-section of 2.6 times SM expectations
 2211 without any systematics and an expected CI of κ_λ at 95% of $[-1.5, 6.6]$. For ATLAS only, further
 2212 improvements of the sensitivity are possible to boost the $HH \rightarrow b\bar{b}\gamma\gamma$ analysis by using more advanced
 2213 techniques for events reconstruction and classification either presented in this thesis or will come over
 2214 years of R&D. At the end of HL-LHC runs, a total integrated luminosity of 3000 fb^{-1} of data at $\sqrt{s} =$
 2215 14 TeV is expected to be collected by ATLAS detector. The $HH \rightarrow b\bar{b}\gamma\gamma$ results are extrapolated to the
 2216 HL-LHC integrated luminosity. The expected κ_λ CI at 95% from the extrapolation to HL-LHC is $[-0.1,$
 2217 $4.4]$ to be compared with previous results from European strategy. At the end of HL-LHC, the signal
 2218 strength is expected to be measured with an accuracy of 44% assuming no systematic uncertainties.
 2219 The expected significance of observing the SM Higgs boson pair production at the end of HL-LHC in
 2220 its $b\bar{b}\gamma\gamma$ decay channel is 2.5σ .

Bibliography

- [1] M. Tanabashi et al. “Review of Particle Physics”. In: *Phys. Rev. D* 98 (3 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: <https://link.aps.org/doi/10.1103/PhysRevD.98.030001> (cit. on p. 6).
- [2] G. Arnison et al. “Experimental observation of isolated large transverse energy electrons with associated missing energy at $s=540$ GeV”. In: *Physics Letters B* 122.1 (1983), pp. 103–116. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(83\)91177-2](https://doi.org/10.1016/0370-2693(83)91177-2). URL: <http://www.sciencedirect.com/science/article/pii/0370269383911772> (cit. on p. 6).
- [3] P. Bagnaia et al. “Evidence for $Z \rightarrow e^+e^-$ at the CERN pp collider”. In: *Physics Letters B* 129.1 (1983), pp. 130–140. ISSN: 0370-2693. DOI: [https://doi.org/10.1016/0370-2693\(83\)90744-X](https://doi.org/10.1016/0370-2693(83)90744-X). URL: <http://www.sciencedirect.com/science/article/pii/037026938390744X> (cit. on p. 6).
- [4] Kazuhiko Nishijima. “Charge Independence Theory of V Particles*”. In: *Progress of Theoretical Physics* 13.3 (1955), pp. 285–304. ISSN: 0033-068X. DOI: 10.1143/PTP.13.285. eprint: <https://academic.oup.com/ptp/article-pdf/13/3/285/5425869/13-3-285.pdf>. URL: <https://doi.org/10.1143/PTP.13.285> (cit. on p. 8).
- [5] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. In: *Phys. Rev. Lett.* 13 (9 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.321> (cit. on p. 9).
- [6] John Ellis, Mary K. Gaillard, and Dimitri V. Nanopoulos. *A Historical Profile of the Higgs Boson*. 2012. arXiv: 1201.6045 [hep-ph] (cit. on p. 10).
- [7] David Delgove. “Mesure des propriétés du boson de Higgs avec l’expérience ATLAS au LHC”. Theses. Université Paris-Saclay, 2016. URL: <https://tel.archives-ouvertes.fr/tel-01494070> (cit. on p. 11).
- [8] In: *Physics Letters B* 565 (2003), pp. 61–75. ISSN: 0370-2693. DOI: 10.1016/S0370-2693(03)00614-2. URL: [http://dx.doi.org/10.1016/S0370-2693\(03\)00614-2](http://dx.doi.org/10.1016/S0370-2693(03)00614-2) (cit. on p. 11).
- [9] T. Aaltonen et al. “Combination of Tevatron Searches for the Standard Model Higgs Boson in the W+W-Decay Mode”. In: *Physical Review Letters* 104.6 (2010). ISSN: 1079-7114. DOI: 10.1103/physrevlett.104.061802. URL: <http://dx.doi.org/10.1103/PhysRevLett.104.061802> (cit. on p. 11).
- [10] ALEPH Collaboration et al. *Precision Electroweak Measurements and Constraints on the Standard Model*. 2009. arXiv: 0911.2604 [hep-ex] (cit. on p. 11).
- [11] Sayipjamal Dulat et al. “New parton distribution functions from a global analysis of quantum chromodynamics”. In: *Physical Review D* 93.3 (2016). ISSN: 2470-0029. DOI: 10.1103/physrevd.93.033006. URL: <http://dx.doi.org/10.1103/PhysRevD.93.033006> (cit. on p. 12).

- [12] Sayipjamal Dulat et al. “New parton distribution functions from a global analysis of quantum chromodynamics”. In: *Physical Review D* 93.3 (2016). ISSN: 2470-0029. DOI: 10.1103/physrevd.93.033006. URL: <http://dx.doi.org/10.1103/PhysRevD.93.033006> (cit. on p. 12).
- [13] *LHC Higgs Working Group*. URL: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHWG?redirectedfrom=LHCPhysics.LHCHXSWG> (cit. on p. 14).
- [14] A. M. Sirunyan et al. “Measurements of the Higgs boson width and anomalous HVV couplings from on-shell and off-shell production in the four-lepton final state”. In: *Physical Review D* 99 (June 2019). ISSN: 2470-0029. DOI: 10.1103/physrevd.99.112003. URL: <http://dx.doi.org/10.1103/PhysRevD.99.112003> (cit. on p. 14).
- [15] S. Dittmaier et al. “Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables”. In: (2011). DOI: 10.5170/CERN-2011-002. arXiv: 1101.0593 [hep-ph] (cit. on p. 15).
- [16] S. Dittmaier et al. “Handbook of LHC Higgs Cross Sections: 2. Differential Distributions”. In: (2012). DOI: 10.5170/CERN-2012-002. arXiv: 1201.3084 [hep-ph] (cit. on p. 15).
- [17] D. de Florian et al. “Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector”. In: 2/2017 (2016). DOI: 10.23731/CYRM-2017-002. arXiv: 1610.07922 [hep-ph] (cit. on p. 16).
- [18] G. Aad et al. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Physics Letters B* 716.1 (2012), pp. 1–29. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.020. URL: <http://dx.doi.org/10.1016/j.physletb.2012.08.020> (cit. on pp. 16, 17).
- [19] S. Chatrchyan et al. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Physics Letters B* 716.1 (2012), pp. 30–61. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.021. URL: <http://dx.doi.org/10.1016/j.physletb.2012.08.021> (cit. on p. 16).
- [20] Aad and *al.* “Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments”. In: *Phys. Rev. Lett.* 114 (19 2015), p. 191803. DOI: 10.1103/PhysRevLett.114.191803. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.114.191803> (cit. on pp. 17, 111).
- [21] *Measurement of the properties of Higgs boson production at $\sqrt{s}=13$ TeV in the $H \rightarrow \gamma\gamma$ channel using 139 fb^{-1} of pp collision data with the ATLAS experiment*. Tech. rep. ATLAS-CONF-2020-026. Geneva: CERN, 2020. URL: <https://cds.cern.ch/record/2725727> (cit. on pp. 17, 18).
- [22] D. de Florian et al. *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*. CERN Yellow Reports: Monographs. 869 pages, 295 figures, 248 tables and 1645 citations. Working Group web page: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG>. 2016. DOI: 10.23731/CYRM-2017-002. URL: <http://cds.cern.ch/record/2227475> (cit. on p. 19).
- [23] S. Borowka et al. “Higgs Boson Pair Production in Gluon Fusion at Next-to-Leading Order with Full Top-Quark Mass Dependence”. In: *Phys. Rev. Lett.* 117 (1 2016), p. 012001. DOI: 10.1103/PhysRevLett.117.012001. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.117.012001> (cit. on p. 19).
- [24] Aaboud et al. “Search for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state with 13 TeV pp collision data collected by the ATLAS experiment”. In: (2018). URL: <http://cds.cern.ch/record/2630823> (cit. on pp. 22, 101, 117).

- 2300 [25] Oliver Sim Brüning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-1. URL: <https://cds.cern.ch/record/782076> (cit. on p. 26).
- 2301 [26] Cinzia De Melis. “The CERN accelerator complex. Complexe des accélérateurs du CERN”. In: (2016). General Photo. URL: <https://cds.cern.ch/record/2119882> (cit. on p. 27).
- 2302 [27] The ATLAS Collaboration. *Public luminosity results*. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2#Multiple_Year_Collision_Plots (cit. on pp. 28, 30).
- 2303 [28] The ATLAS Collaboration et al. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *Journal of Instrumentation* 3.08 (2008), S08003–S08003. DOI: 10.1088/1748-0221/3/08/S08003. URL: <https://doi.org/10.1088/1748-0221/3/08/s08003> (cit. on p. 30).
- 2304 [29] *ATLAS inner detector: Technical Design Report, 1.* Technical design report. ATLAS. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/331063> (cit. on p. 32).
- 2305 [30] Kyungeon Choi. *Tracking and Vertexing with the ATLAS Inner Detector in the LHC Run-2*. Tech. rep. ATL-PHYS-PROC-2017-075. Geneva: CERN, 2017. DOI: 10.1007/978-981-13-1316-5_75. URL: <https://cds.cern.ch/record/2271033> (cit. on p. 32).
- 2306 [31] Karolos Potamianos. *The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider*. Tech. rep. ATL-PHYS-PROC-2016-104. 15 pages, EPS-HEP 2015 Proceedings. Geneva: CERN, 2016. URL: <https://cds.cern.ch/record/2209070> (cit. on pp. 32, 33).
- 2307 [32] M Capeans et al. *ATLAS Insertable B-Layer Technical Design Report*. Tech. rep. CERN-LHCC-2010-013. ATLAS-TDR-19. 2010. URL: <https://cds.cern.ch/record/1291633> (cit. on p. 32).
- 2308 [33] ATLAS Collaboration. *Impact Parameter Resolution*. URL: <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2015-007/> (cit. on p. 34).
- 2309 [34] *Timing improvements for the ATLAS offline reconstruction for run-2*. Tech. rep. ATL-SOFT-PUB-2014-004. Geneva: CERN, 2014. URL: <https://cds.cern.ch/record/1955923> (cit. on p. 33).
- 2310 [35] ATLAS Collaboration. *The Optimization of ATLAS Track Reconstruction in Dense Environments*. URL: <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-006/> (cit. on p. 34).
- 2311 [36] *Expected performance of the ATLAS b-tagging algorithms in Run-2*. Tech. rep. ATL-PHYS-PUB-2015-022. Geneva: CERN, 2015. URL: [http://cds.cern.ch/record/2037697](https://cds.cern.ch/record/2037697) (cit. on p. 33).
- 2312 [37] *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*. Tech. rep. ATLAS-CONF-2011-102. Geneva: CERN, 2011. URL: [http://cds.cern.ch/record/1369219](https://cds.cern.ch/record/1369219) (cit. on p. 33).
- 2313 [38] T G Cornelissen et al. *Updates of the ATLAS Tracking Event Data Model (Release 13)*. Tech. rep. ATL-SOFT-PUB-2007-003. ATL-COM-SOFT-2007-008. Geneva: CERN, 2007. URL: <https://cds.cern.ch/record/1038095> (cit. on p. 37).
- 2314 [39] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1996. URL: <https://cds.cern.ch/record/331061> (cit. on p. 38).
- 2315 [40] *ATLAS tile calorimeter: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1996. URL: <https://cds.cern.ch/record/331062> (cit. on p. 39).

- [41] A. Hrynevich. "Performance of the ATLAS Tile Calorimeter". In: *Journal of Instrumentation* 12.06 (2017), pp. C06021–C06021. DOI: 10.1088/1748-0221/12/06/c06021. URL: <https://doi.org/10.1088/1748-0221/12/06/c06021> (cit. on p. 40).
- [42] *ATLAS muon spectrometer: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/331068> (cit. on p. 40).
- [43] A. Airapetian et al. *ATLAS detector and physics performance: Technical Design Report, 1*. Technical design report. ATLAS. Geneva: CERN, 1999. URL: <https://cds.cern.ch/record/391176> (cit. on p. 41).
- [44] *ATLAS level-1 trigger: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1998. URL: <https://cds.cern.ch/record/381429> (cit. on p. 42).
- [45] Peter Jenni et al. *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 2003. URL: <https://cds.cern.ch/record/616089> (cit. on p. 42).
- [46] T. Golling et al. "The ATLAS data quality defect database system". In: *The European Physical Journal C* 72.4 (2012). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-012-1960-y. URL: <http://dx.doi.org/10.1140/epjc/s10052-012-1960-y> (cit. on p. 42).
- [47] Corentin Allaire. *A High-Granularity Timing Detector (HGTD) in ATLAS : Performance at the HL-LHC*. Tech. rep. Geneva: CERN, 2018. URL: <https://cds.cern.ch/record/2302827> (cit. on p. 42).
- [48] *ATLAS Phase-II Upgrade Scoping Document*. Tech. rep. Geneva: CERN, 2015. URL: <https://cds.cern.ch/record/2055248> (cit. on p. 43).
- [49] *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*. Tech. rep. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2285585> (cit. on p. 43).
- [50] *Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System*. Tech. rep. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2285584> (cit. on p. 43).
- [51] *Technical Design Report for the ATLAS Inner Tracker Strip Detector*. Tech. rep. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2257755> (cit. on p. 43).
- [52] *Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer*. Tech. rep. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2285580> (cit. on p. 43).
- [53] *ATLAS Liquid Argon Calorimeter Phase-II Upgrade: Technical Design Report*. Tech. rep. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2285582> (cit. on p. 43).
- [54] *Technical Design Report for the Phase-II Upgrade of the ATLAS Tile Calorimeter*. Tech. rep. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2285583> (cit. on p. 43).
- [55] *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*. Tech. rep. ATLAS-CONF-2012-042. Geneva: CERN, 2012. URL: <https://cds.cern.ch/record/1435196> (cit. on p. 43).
- [56] T Cornelissen et al. *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*. Tech. rep. ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002. Geneva: CERN, 2007. URL: <https://cds.cern.ch/record/1020106> (cit. on p. 43).

- [57] R. Frühwirth. “Application of Kalman filtering to track and vertex fitting”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 262.2 (1987), pp. 444–450. ISSN: 0168-9002. DOI: [https://doi.org/10.1016/0168-9002\(87\)90887-4](https://doi.org/10.1016/0168-9002(87)90887-4). URL: <http://www.sciencedirect.com/science/article/pii/0168900287908874> (cit. on p. 43).
- [58] *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*. Tech. rep. ATLAS-CONF-2012-042. Geneva: CERN, 2012. URL: <https://cds.cern.ch/record/1435196> (cit. on p. 43).
- [59] T G Cornelissen et al. “The global Ψ 2track fitter in ATLAS”. In: *Journal of Physics: Conference Series* 119.3 (2008), p. 032013. DOI: 10.1088/1742-6596/119/3/032013. URL: <https://doi.org/10.1088/1742-6596/119/3/032013> (cit. on p. 43).
- [60] W Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance*. Tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003. Geneva: CERN, 2008. URL: <https://cds.cern.ch/record/1099735> (cit. on p. 45).
- [61] G. Aad et al. “Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data”. In: *Journal of Instrumentation* 14.12 (2019), P12006–P12006. ISSN: 1748-0221. DOI: 10.1088/1748-0221/14/12/p12006. URL: <http://dx.doi.org/10.1088/1748-0221/14/12/P12006> (cit. on pp. 45, 48, 52).
- [62] G. Aad et al. “Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1”. In: *The European Physical Journal C* 77.7 (2017). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-5004-5. URL: <http://dx.doi.org/10.1140/epjc/s10052-017-5004-5> (cit. on p. 45).
- [63] *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*. Tech. rep. ATLAS-CONF-2012-047. Geneva: CERN, 2012. URL: <https://cds.cern.ch/record/1449796> (cit. on p. 45).
- [64] M. Aaboud et al. “Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data”. In: *Journal of Instrumentation* 14.03 (2019), P03017–P03017. ISSN: 1748-0221. DOI: 10.1088/1748-0221/14/03/p03017. URL: <http://dx.doi.org/10.1088/1748-0221/14/03/P03017> (cit. on p. 47).
- [65] G. Aad et al. “Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data”. In: *The European Physical Journal C* 74.10 (2014). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-014-3071-4. URL: <http://dx.doi.org/10.1140/epjc/s10052-014-3071-4> (cit. on p. 49).
- [66] M. Aaboud et al. “Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV”. In: *The European Physical Journal C* 79.8 (2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7140-6. URL: <http://dx.doi.org/10.1140/epjc/s10052-019-7140-6> (cit. on pp. 49, 51).
- [67] Matteo Cacciari and Gavin P. Salam. “Pileup subtraction using jet areas”. In: *Physics Letters B* 659.1-2 (2008), pp. 119–126. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2007.09.077. URL: <http://dx.doi.org/10.1016/j.physletb.2007.09.077> (cit. on p. 49).
- [68] G. Aad et al. “Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data”. In: *Journal of Instrumentation* 14.12 (2019), P12006–P12006. ISSN: 1748-0221. DOI: 10.1088/1748-0221/14/12/p12006. URL: <http://dx.doi.org/10.1088/1748-0221/14/12/P12006> (cit. on pp. 51, 56, 58).

- [69] David W. Scott. *Multivariate Density Estimation: Theory, Practice, and Visualization*. Wiley Series In Probability and Statistics. WILEY, 2015 (cit. on p. 51).
- [70] Andreas Hocker et al. *TMVA - Toolkit for Multivariate Data Analysis with ROOT: Users guide. TMVA - Toolkit for Multivariate Data Analysis*. Tech. rep. physics/0703039. TMVA-v4 Users Guide: 135 pages, 19 figures, numerous code examples and references. Geneva: CERN, 2007. URL: <https://cds.cern.ch/record/1019880> (cit. on pp. 51, 58).
- [71] M. Aaboud et al. “Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV”. In: *The European Physical Journal C* 79.8 (2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7140-6. URL: <http://dx.doi.org/10.1140/epjc/s10052-019-7140-6> (cit. on p. 52).
- [72] J. Illingworth and J. Kittler. “A survey of the hough transform”. In: *Computer Vision, Graphics, and Image Processing* 44.1 (1988), pp. 87–116. ISSN: 0734-189X. DOI: [https://doi.org/10.1016/S0734-189X\(88\)80033-1](https://doi.org/10.1016/S0734-189X(88)80033-1). URL: <https://www.sciencedirect.com/science/article/pii/S0734189X88800331> (cit. on p. 52).
- [73] G. Aad et al. “Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data”. In: *The European Physical Journal C* 74.11 (2014). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-014-3130-x. URL: <http://dx.doi.org/10.1140/epjc/s10052-014-3130-x> (cit. on p. 52).
- [74] G. Aad et al. “Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV”. In: *The European Physical Journal C* 76.5 (2016). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4120-y. URL: <http://dx.doi.org/10.1140/epjc/s10052-016-4120-y> (cit. on p. 52).
- [75] G. Aad et al. “Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector”. In: *Physical Review D* 91.1 (2015). ISSN: 1550-2368. DOI: 10.1103/physrevd.91.012006. URL: <http://dx.doi.org/10.1103/PhysRevD.91.012006> (cit. on p. 54).
- [76] G. Aad et al. “Search for new particles in events with one lepton and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector”. In: *Journal of High Energy Physics* 2014.9 (2014). ISSN: 1029-8479. DOI: 10.1007/jhep09(2014)037. URL: [http://dx.doi.org/10.1007/JHEP09\(2014\)037](http://dx.doi.org/10.1007/JHEP09(2014)037) (cit. on p. 54).
- [77] G. Aad et al. “Search for high-mass dilepton resonances in pp collisions at $\sqrt{s}=8$ TeV with the ATLAS detector”. In: *Physical Review D* 90.5 (2014). ISSN: 1550-2368. DOI: 10.1103/physrevd.90.052005. URL: <http://dx.doi.org/10.1103/PhysRevD.90.052005> (cit. on p. 54).
- [78] A. Hrynevich. “ATLAS jet and missing energy reconstruction, calibration and performance in LHC Run-2”. In: *Journal of Instrumentation* 12.06 (2017), pp. C06038–C06038. DOI: 10.1088/1748-0221/12/06/c06038. URL: <https://doi.org/10.1088/1748-0221/12/06/c06038> (cit. on p. 54).
- [79] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. “The anti-ktjet clustering algorithm”. In: *Journal of High Energy Physics* 2008.04 (2008), pp. 063–063. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2008/04/063. URL: <http://dx.doi.org/10.1088/1126-6708/2008/04/063> (cit. on p. 54).
- [80] Benjamin Brunt et al. *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV*. Tech. rep. Geneva: CERN, 2016. URL: <https://cds.cern.ch/record/2149445> (cit. on p. 55).

- [81] G. Aad et al. “Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data”. In: *The European Physical Journal C* 72.3 (2012). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-012-1909-1. URL: <http://dx.doi.org/10.1140/epjc/s10052-012-1909-1> (cit. on p. 56).
- [82] P. Colangelo, F. Fazio, and G. Nardulli. “A STUDY OF THE REACTIONS OF”. In: 1992 (cit. on p. 56).
- [83] James Saxon. “Discovery of the Higgs Boson, Measurements of its Production, and a Search for Higgs Boson Pair Production”. Presented 13 06 2014. 2014. URL: <https://cds.cern.ch/record/1746004> (cit. on p. 58).
- [84] M. Aaboud et al. “Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016”. In: *The European Physical Journal C* 79.3 (2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-6650-6. URL: <http://dx.doi.org/10.1140/epjc/s10052-019-6650-6> (cit. on pp. 58, 59).
- [85] M. Aaboud et al. “Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data”. In: *The European Physical Journal C* 76.12 (2016). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4507-9. URL: <http://dx.doi.org/10.1140/epjc/s10052-016-4507-9> (cit. on p. 59).
- [86] Tairan Xu. “Measurements of W boson properties at $\sqrt{s} = 5$ and 13 TeV with the ATLAS detector at the LHC”. Theses. Université Paris Saclay (COmUE) ; University of science and technology of China, 2019. URL: <https://tel.archives-ouvertes.fr/tel-02873298> (cit. on p. 62).
- [87] Mykola Khandoga. “Calibration of electron shower shapes, hadronic recoil reconstruction using deep learning algorithms and the measurement of W boson transverse momentum distribution with the ATLAS detector”. Theses. Université Paris-Saclay, 2020. URL: <https://tel.archives-ouvertes.fr/tel-03143181> (cit. on pp. 62, 63).
- [88] M. Egmont-Petersen, D. de Ridder, and H. Handels. “Image processing with neural networks—a review”. In: *Pattern Recognition* 35.10 (2002), pp. 2279–2301. ISSN: 0031-3203. DOI: [https://doi.org/10.1016/S0031-3203\(01\)00178-9](https://doi.org/10.1016/S0031-3203(01)00178-9). URL: <https://www.sciencedirect.com/science/article/pii/S0031320301001789> (cit. on p. 70).
- [89] F. Chollet et al. *Keras*. <https://keras.io/>. 2015 (cit. on pp. 71, 103).
- [90] M. Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, Software available from tensorflow.org*. <http://tensorflow.org/>. 2015 (cit. on pp. 71, 103).
- [91] Keras. *Pooling layers*. https://keras.io/api/layers/pooling_layers/. 2015 (cit. on p. 71).
- [92] Zhen Huang et al. *SNDCNN: Self-normalizing deep CNNs with scaled exponential linear units for speech recognition*. <https://arxiv.org/abs/1910.01992>. 2019 (cit. on p. 71).
- [93] Centre de Calcul de l’Institut National de Physique Nucléaire et de Physique des Particules, Lyon (France). 1976. URL: <https://cc.in2p3.fr/> (cit. on p. 72).
- [94] Yuan Yao, Lorenzo Rosasco, and Andrea Caponnetto. “On early stopping in gradient descent learning”. In: (2007) (cit. on p. 72).
- [95] L Asquith et al. *Performance of Jet Algorithms in the ATLAS Detector*. Tech. rep. ATL-PHYS-INT-2010-129. Geneva: CERN, 2010. URL: <https://cds.cern.ch/record/1311867> (cit. on p. 82).

- 2510 [96] M. Aaboud et al. “Jet reconstruction and performance using particle flow with the ATLAS
 2511 Detector”. In: *The European Physical Journal C* 77.7 (2017). ISSN: 1434-6052. DOI: 10.1140/
 2512 [epjc/s10052-017-5031-2](http://dx.doi.org/10.1140/epjc/s10052-017-5031-2). URL: <http://dx.doi.org/10.1140/epjc/s10052-017-5031-2>
 2513 (cit. on pp. 82, 83).
- 2514 [97] G. Aad et al. “Charged-particle multiplicities in $p\bar{p}$ interactions at $s = 900$ GeV measured with
 2515 the ATLAS detector at the LHC”. In: *Physics Letters B* 688.1 (2010), pp. 21–42. ISSN: 0370-2693.
 2516 DOI: 10.1016/j.physletb.2010.03.064. URL: <http://dx.doi.org/10.1016/j.physletb.2010.03.064> (cit. on p. 83).
- 2518 [98] M. Aaboud et al. “Jet energy scale measurements and their systematic uncertainties in proton-
 2519 proton collisions at $s=13$ TeV with the ATLAS detector”. In: *Physical Review D* 96.7 (2017).
 2520 ISSN: 2470-0029. DOI: 10.1103/physrevd.96.072002. URL: <http://dx.doi.org/10.1103/PhysRevD.96.072002> (cit. on pp. 84, 85).
- 2522 [99] Matteo Cacciari and Gavin P. Salam. “Pileup subtraction using jet areas”. In: *Physics Letters
 2523 B* 659 (2008), pp. 119–126. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2007.09.077. URL:
 2524 <http://dx.doi.org/10.1016/j.physletb.2007.09.077> (cit. on p. 83).
- 2525 [100] G. Aad et al. “Jet energy measurement with the ATLAS detector in proton-proton collisions at
 2526 $\sqrt{s} = 7$ TeV”. In: *The European Physical Journal C* 73.3 (2013). ISSN: 1434-6052. DOI: 10.1140/
 2527 [epjc/s10052-013-2304-2](http://dx.doi.org/10.1140/epjc/s10052-013-2304-2). URL: <http://dx.doi.org/10.1140/epjc/s10052-013-2304-2>
 2528 (cit. on p. 84).
- 2529 [101] G. Aad et al. “Jet energy measurement and its systematic uncertainty in proton–proton collisions
 2530 at $\sqrt{s} = 7$ TeV with the ATLAS detector”. In: *The European Physical Journal C* 75.1 (2015).
 2531 ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-014-3190-y. URL: <http://dx.doi.org/10.1140/epjc/s10052-014-3190-y> (cit. on p. 84).
- 2533 [102] *Pile-up subtraction and suppression for jets in ATLAS*. Tech. rep. ATLAS-CONF-2013-083.
 2534 Geneva: CERN, 2013. URL: <https://cds.cern.ch/record/1570994> (cit. on p. 86).
- 2535 [103] *Tagging and suppression of pileup jets*. Tech. rep. ATL-PHYS-PUB-2014-001. Geneva: CERN,
 2536 2014. URL: <https://cds.cern.ch/record/1643929> (cit. on p. 86).
- 2537 [104] *Tagging and suppression of pileup jets with the ATLAS detector*. Tech. rep. ATLAS-CONF-2014-
 2538 018. Geneva: CERN, 2014. URL: <https://cds.cern.ch/record/1700870> (cit. on p. 86).
- 2539 [105] Ariel Schwartzman, Matthew Henry Klein, and Francesco Rubbo. *Identification and rejection of
 2540 pileup jets in the forward region with the ATLAS detector*. Tech. rep. ATL-COM-PHYS-2016-710.
 2541 Geneva: CERN, 2016. URL: <https://cds.cern.ch/record/2158916> (cit. on p. 86).
- 2542 [106] G. Aad et al. “Light-quark and gluon jet discrimination in $p\bar{p}$ collisions at $\sqrt{s} = 7$ TeV with
 2543 the ATLAS detector”. In: *The European Physical Journal C* 74.8 (2014). ISSN: 1434-6052. DOI:
 2544 10.1140/epjc/s10052-014-3023-z. URL: <http://dx.doi.org/10.1140/epjc/s10052-014-3023-z> (cit. on p. 86).
- 2546 [107] G. Aad et al. “Performance of pile-up mitigation techniques for jets in $p\bar{p}$ collisions at $\sqrt{s} = 8$ TeV
 2547 using the ATLAS detector”. In: *The European Physical Journal C* 76.11 (2016). ISSN: 1434-6052.
 2548 DOI: 10.1140/epjc/s10052-016-4395-z. URL: <http://dx.doi.org/10.1140/epjc/s10052-016-4395-z> (cit. on p. 86).

- 2550 [108] N. S. Altman. "An Introduction to Kernel and Nearest-Neighbor Non parametric Regression". In: *The American Statistician* 46.3 (1992), pp. 175–185. DOI: 10.1080/00031305.1992.10475879.
 2551 eprint: <https://www.tandfonline.com/doi/pdf/10.1080/00031305.1992.10475879>. URL: <https://www.tandfonline.com/doi/abs/10.1080/00031305.1992.10475879> (cit. on p. 86).
- 2554 [109] Louis Portales. "Observation of electroweak $WZjj$ production, and studies on pile-up mitigation
 2555 with the ATLAS detector. Observation de la production $WZjj$ électrofaible, et études sur la
 2556 suppression des jets d'empilement avec le détecteur ATLAS". Presented 01 Oct 2020. 2020. URL:
 2557 <https://cds.cern.ch/record/2743491> (cit. on p. 89).
- 2558 [110] *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*. Tech. rep. ATL-PHYS-PUB-2017-003. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2255226> (cit. on p. 89).
- 2561 [111] *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*. Tech. rep. ATL-PHYS-PUB-2017-013. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2273281> (cit. on pp. 89, 109).
- 2564 [112] *Secondary vertex finding for jet flavour identification with the ATLAS detector*. Tech. rep. ATL-
 2565 PHYS-PUB-2017-011. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2270366>
 2566 (cit. on p. 89).
- 2567 [113] *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*. Tech. rep. ATL-PHYS-PUB-2018-025. Geneva: CERN,
 2568 2018. URL: <https://cds.cern.ch/record/2645405> (cit. on p. 89).
- 2570 [114] Marie Christine Lanfermann, Tobias Golling, and Andrea Coccaro. *Deep Neural Network based
 2571 higher level flavour tagging algorithm at the ATLAS experiment*. Tech. rep. ATL-COM-PHYS-
 2572 2017-1596. Geneva: CERN, 2017. URL: <https://cds.cern.ch/record/2290144> (cit. on p. 89).
- 2573 [115] ATLAS Collaboration. *Expected performance of the 2019 ATLAS b-taggers*. Geneva. URL: <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2019-005/> (cit. on p. 90).
- 2575 [116] *Search for the Standard Model Higgs boson produced in association with a vector boson and
 2576 decaying to a $b\bar{b}$ pair in pp collisions at 13 TeV using the ATLAS detector*. Tech. rep. ATLAS-
 2577 CONF-2016-091. Geneva: CERN, 2016. URL: <http://cds.cern.ch/record/2206813> (cit. on
 2578 p. 91).
- 2579 [117] ATLAS Collaboration. *Luminosity uncertainties for physics analysis*. Geneva. URL: <https://twiki.cern.ch/twiki/bin/view/Atlas/LuminosityForPhysics> (cit. on p. 96).
- 2581 [118] G. Heinrich et al. "NLO predictions for Higgs boson pair production with full top quark mass
 2582 dependence matched to parton showers". In: *Journal of High Energy Physics* 2017.8 (2017). ISSN:
 2583 1029-8479. DOI: 10.1007/jhep08(2017)088. URL: [http://dx.doi.org/10.1007/JHEP08\(2017\)
 2584 088](http://dx.doi.org/10.1007/JHEP08(2017)088) (cit. on p. 97).
- 2585 [119] G. Heinrich et al. "Probing the trilinear Higgs boson coupling in di-Higgs production at NLO QCD
 2586 including parton shower effects". In: *Journal of High Energy Physics* 2019.6 (2019). ISSN: 1029-
 2587 8479. DOI: 10.1007/jhep06(2019)066. URL: [http://dx.doi.org/10.1007/JHEP06\(2019\)066](http://dx.doi.org/10.1007/JHEP06(2019)066)
 2588 (cit. on p. 97).
- 2589 [120] Jon Butterworth et al. "PDF4LHC recommendations for LHC Run II". In: *Journal of Physics G: Nuclear and Particle Physics* 43.2 (2016), p. 023001. ISSN: 1361-6471. DOI: 10.1088/0954-
 2590 3899/43/2/023001. URL: <http://dx.doi.org/10.1088/0954-3899/43/2/023001> (cit. on
 2591 p. 97).
- 2592

- 2593 [121] Fady Bishara, Roberto Contino, and Juan Rojo. “Higgs pair production in vector-boson fusion at
 2594 the LHC and beyond”. In: *The European Physical Journal C* 77.7 (2017). ISSN: 1434-6052. DOI:
 2595 10.1140/epjc/s10052-017-5037-9. URL: <http://dx.doi.org/10.1140/epjc/s10052-017-5037-9> (cit. on p. 97).
- 2597 [122] Richard D. Ball et al. “Parton distributions with LHC data”. In: *Nuclear Physics B* 867.2 (2013),
 2598 pp. 244–289. ISSN: 0550-3213. DOI: 10.1016/j.nuclphysb.2012.10.003. URL: <http://dx.doi.org/10.1016/j.nuclphysb.2012.10.003> (cit. on p. 97).
- 2600 [123] S. Amoroso et al. *Les Houches 2019: Physics at TeV Colliders: Standard Model Working Group Report*. 2020. arXiv: 2003.01700 [hep-ph] (cit. on p. 97).
- 2602 [124] G. et al. Aad. “Measurement of Higgs boson production in the diphoton decay channel in pp
 2603 collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector”. In: *Physical Review D* 90.11 (2014). ISSN: 1550-2368. DOI: 10.1103/physrevd.90.112015. URL: <http://dx.doi.org/10.1103/PhysRevD.90.112015> (cit. on p. 98).
- 2606 [125] B. Efron. “Bootstrap Methods: Another Look at the Jackknife”. In: *The Annals of Statistics* 7.1
 2607 (1979), pp. 1–26. DOI: 10.1214/aos/1176344552. URL: <https://doi.org/10.1214/aos/1176344552> (cit. on p. 99).
- 2609 [126] Tianqi Chen and Carlos Guestrin. “XGBoost: A Scalable Tree Boosting System”. In: KDD ’16.
 2610 San Francisco, California, USA: Association for Computing Machinery, 2016, pp. 785–794. ISBN:
 2611 9781450342322. DOI: 10.1145/2939672.2939785. URL: <https://doi.org/10.1145/2939672.2939785> (cit. on p. 99).
- 2613 [127] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *The
 2614 European Physical Journal C* 71.2 (2011). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-
 2615 1554-0. URL: <http://dx.doi.org/10.1140/epjc/s10052-011-1554-0> (cit. on pp. 106, 112).
- 2616 [128] *Measurements of Higgs boson properties in the diphoton decay channel using 80 fb of pp collision
 2617 data at s = 13 TeV with the ATLAS detector*. Tech. rep. ATLAS-CONF-2018-028. Geneva: CERN,
 2618 2018. URL: <https://cds.cern.ch/record/2628771> (cit. on p. 107).
- 2619 [129] G. Aad et al. “Measurement of the isolated diphoton cross section in pp collisions at s=7 TeV
 2620 with the ATLAS detector”. In: *Physical Review D* 85.1 (2012). ISSN: 1550-2368. DOI: 10.1103/
 2621 physrevd.85.012003. URL: <http://dx.doi.org/10.1103/PhysRevD.85.012003> (cit. on
 2622 p. 107).
- 2623 [130] ATLAS Collaboration. “Electron and photon performance measurements with the ATLAS detector
 2624 using the 2015-2017 LHC proton-proton collision data”. In: *JINST* 14.1 (2019). ISSN: 1748-
 2625 0221. DOI: 10.1088/1748-0221/14/12/P12006. URL: <http://dx.doi.org/10.1088/1748-0221/14/12/P12006> (cit. on p. 109).
- 2627 [131] D. de Florian et al. “Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the
 2628 Higgs Sector”. In: 2/2017 (2016). DOI: 10.23731/CYRM-2017-002. arXiv: 1610.07922 [hep-ph]
 2629 (cit. on pp. 111, 133).
- 2630 [132] A. L. Read. “Presentation of search results: the CLs technique”. In: *JINST* 28 (2002). URL:
 2631 <http://dx.doi.org/10.1088/0954-3899/28/10/313> (cit. on p. 112).
- 2632 [133] ATLAS Collaboration. “Combined measurements of Higgs boson production and decay using up
 2633 to 80 fb of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment”.
 2634 In: *Phys. Rev. D* 101.arXiv:1909.02845 (2019), 012002. 48 p. DOI: 10.1103/PhysRevD.101.
 2635 012002. URL: <https://cds.cern.ch/record/2688596> (cit. on p. 114).

- [134] Fabian Pedregosa and al. “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12.85 (2011), pp. 2825–2830 (cit. on p. 117).
- [135] *Search for Higgs boson pair production in the two bottom quarks plus two photons final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2021-016>. Geneva: CERN, Mar. 2021. URL: <http://cds.cern.ch/record/2759683> (cit. on p. 117).
- [136] CMS Collaboration. “Search for non-resonant Higgs boson pair production in final states with two bottom quarks and two photons in proton-proton collisions at $\sqrt{s} = 13$ TeV”. In: *JHEP* 2103 (Nov. 2020). Submitted to JHEP. All figures and tables can be found at <http://cms-results.web.cern.ch/cms-results/public-results/publications/HIG-19-018> (CMS Public Pages), 257. 57 p. DOI: 10.1007/JHEP03(2021)257. arXiv: 2011.12373. URL: <https://cds.cern.ch/record/2745738> (cit. on p. 118).
- [137] Alexandra Carvalho et al. *On the reinterpretation of non-resonant searches for Higgs boson pairs*. 2017. arXiv: 1710.08261 [hep-ph] (cit. on p. 118).
- [138] G. Aad et al. “Combination of searches for Higgs boson pairs in pp collisions at $s=13\text{TeV}$ with the ATLAS detector”. In: *Physics Letters B* 800 (2020), p. 135103. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2019.135103. URL: <http://dx.doi.org/10.1016/j.physletb.2019.135103> (cit. on p. 122).
- [139] M Cepeda and al. “Report from Working Group 2: Higgs Physics at the HL-LHC and HE-LHC”. In: *CERN Yellow Rep. Monogr.* 7 (2018), 221–584. 364 p. DOI: 10.23731/CYRM-2019-007.221. arXiv: 1902.00134. URL: <https://cds.cern.ch/record/2650162> (cit. on pp. 125, 131).

2658 **Appendix A**

2659 **Cell-by-cell reweighting**

2660 In liaison with Section 4.3.3, additional control plots are shown here.

2661 **A.1 Electron reweighting applied to photons**

2662 Figures A.1, A.2, A.3 and A.4 show the energy profiles and shower shape variables for 4 $|\eta|$ bins after
2663 applying the derived reweighting for electrons on photons. Similar behaviour is observed for the rest of
2664 $|\eta|$ bins.

2665 **A.2 Specific photon reweighting**

2666 Figures A.5, A.6, A.7 and A.8 show the energy profiles and shower shape variables for 4 $|\eta|$ bins after
2667 applying the photon reweighting. Similar behaviour is observed for the rest of $|\eta|$ bins.

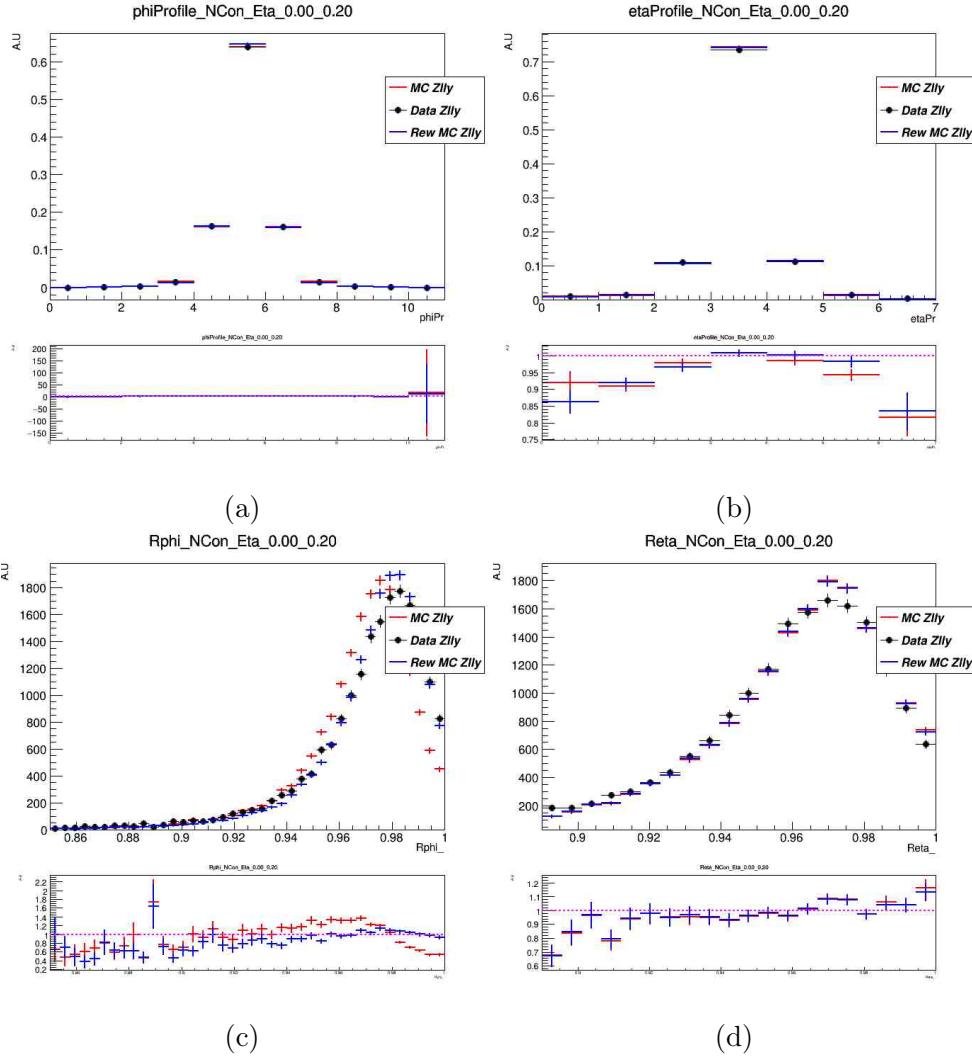


Figure A.1: The energy profile in ϕ and η directions (a, b) and the corresponding R_ϕ and R_η variables, for unconverted photons with $0.00 < |\eta| < 0.20$. The black points correspond to the pseudo data, red points to non-reweighted simulation and blue points to the reweighted simulation from $Z \rightarrow ll\gamma$ with derived electron reweighting.

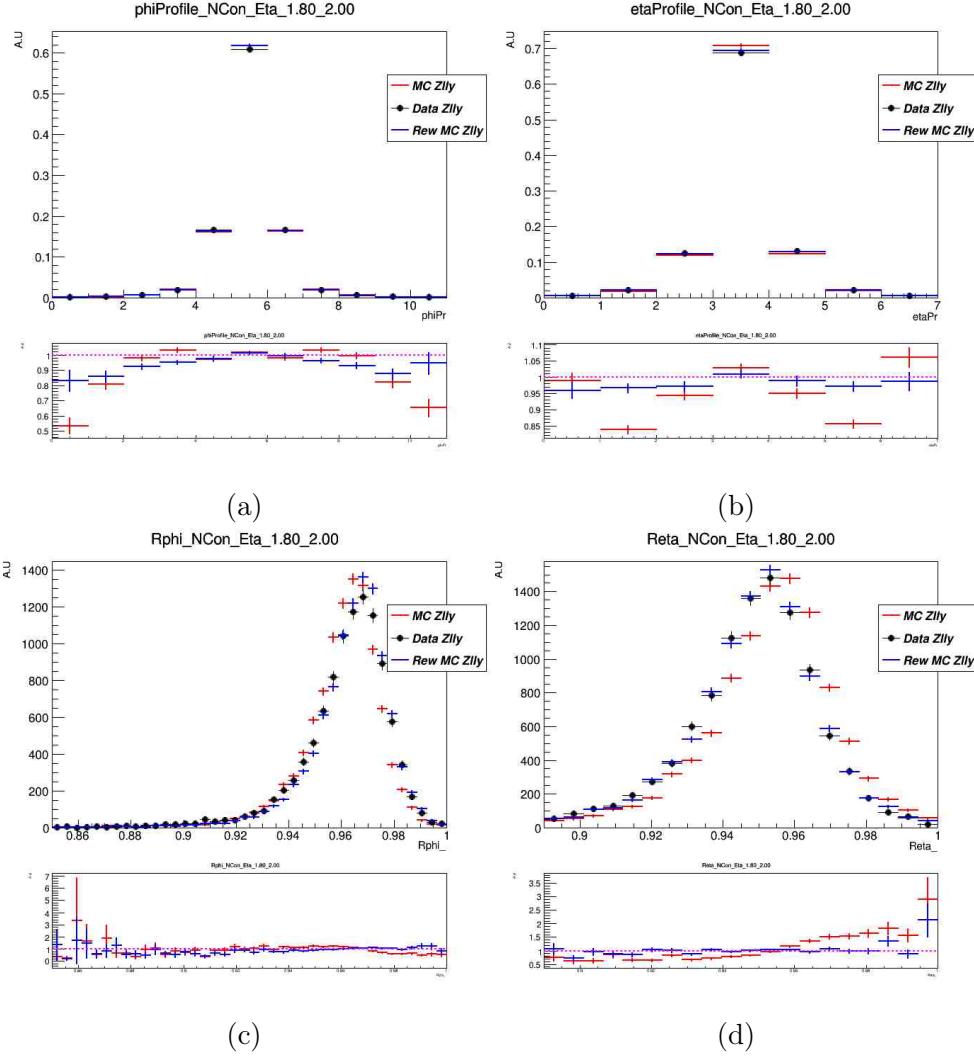


Figure A.2: The energy profile in ϕ and η directions (a,b) and the corresponding R_ϕ and R_η variables, for unconverted photons with $1.80 < |\eta| < 2.00$. The black points correspond to the pseudo data, red points to non-reweighted simulation and blue points to the reweighted simulation from $Z \rightarrow ll\gamma$ with derived electron reweighting.

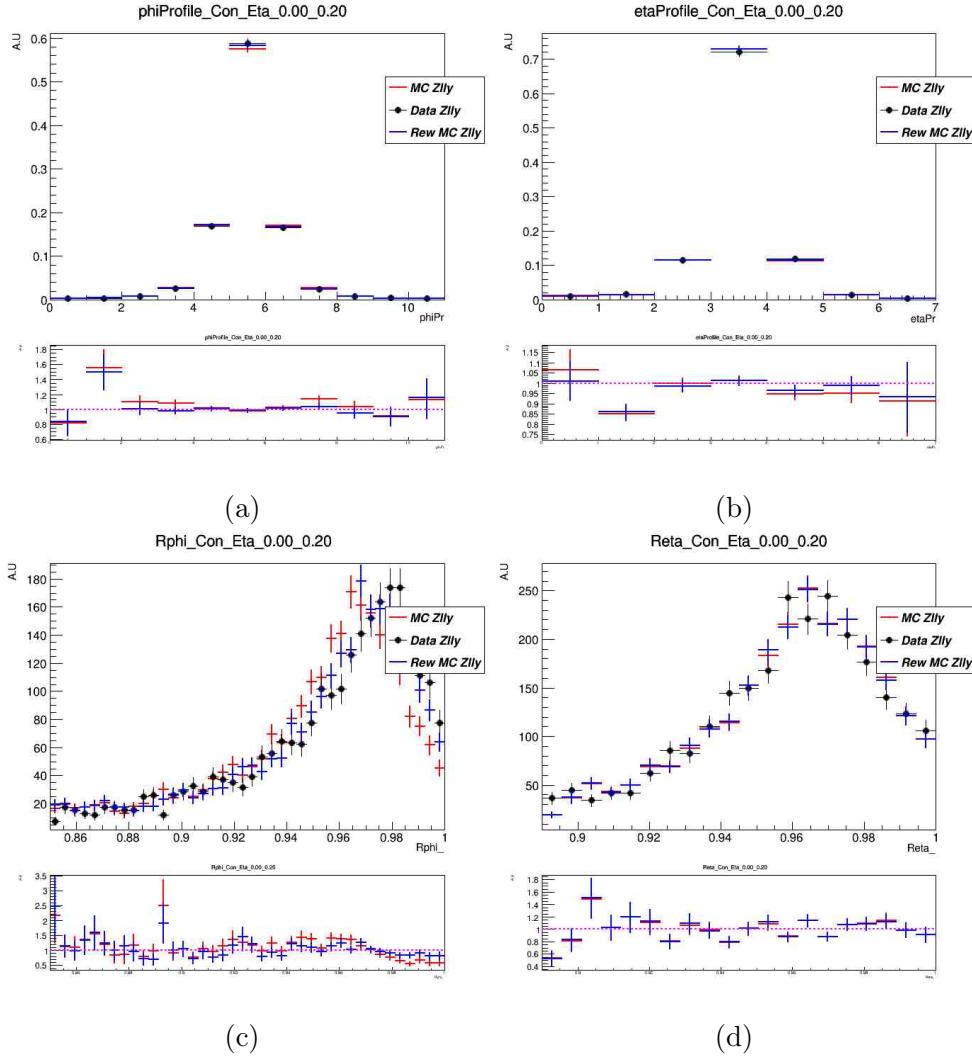


Figure A.3: The energy profile in ϕ and η directions (a,b) and the corresponding R_ϕ and R_η variables, for converted photons with $0.00 < |\eta| < 0.20$. The black points correspond to the pseudo data, red points to non-reweighted simulation and blue points to the reweighted simulation from $Z \rightarrow ll\gamma$ with derived electron reweighting.

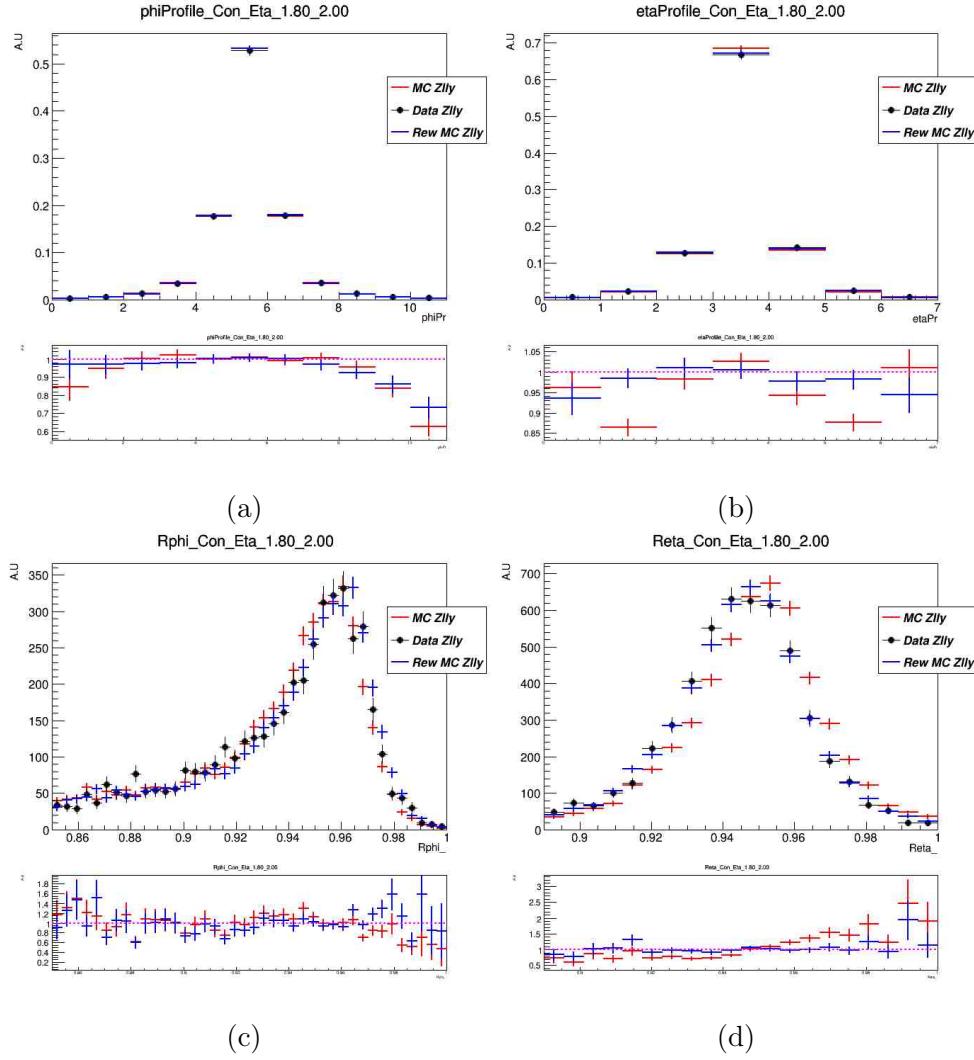


Figure A.4: The energy profile in ϕ and η directions (a,b) and the corresponding R_ϕ and R_η variables, for converted photons with $1.80 < |\eta| < 2.00$. The black points correspond to the pseudo data, red points to non-reweighted simulation and blue points to the reweighted simulation from $Z \rightarrow ll\gamma$ with derived electron reweighting.

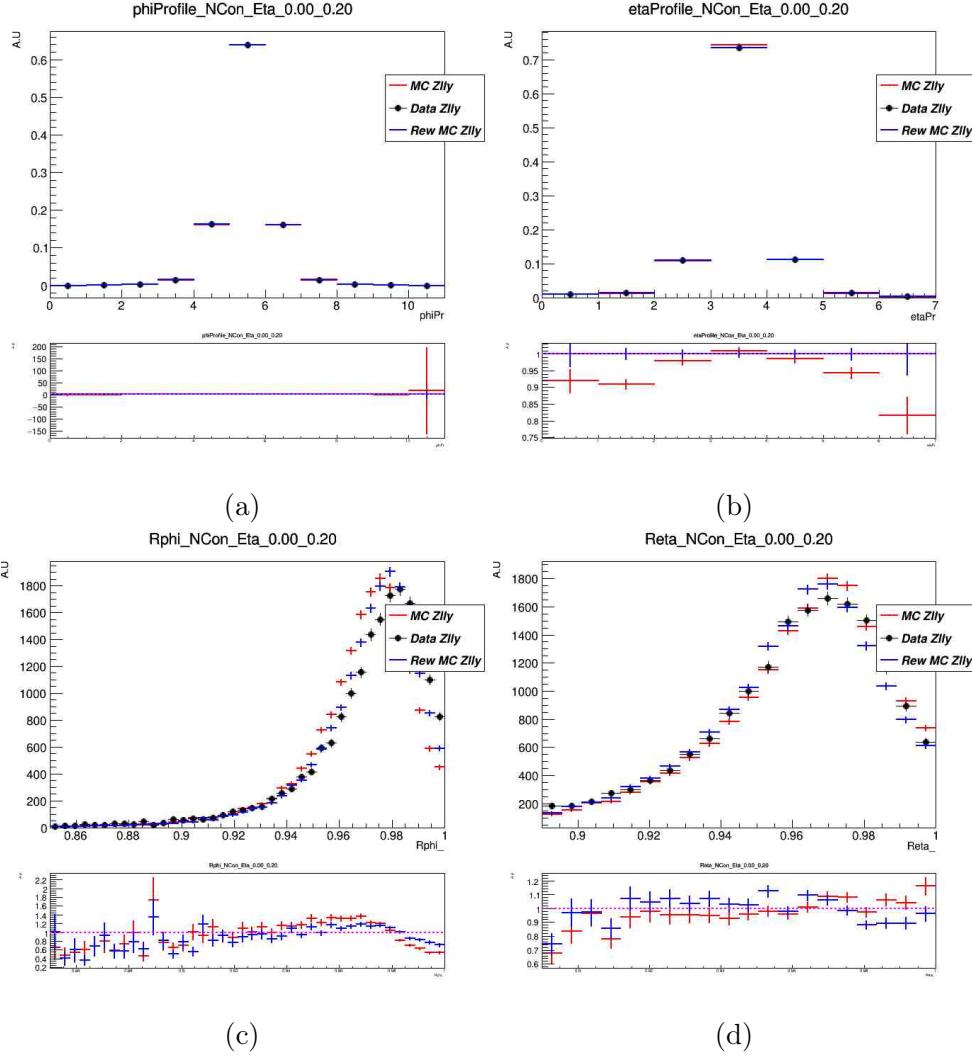


Figure A.5: The energy profile in ϕ and η directions (a,b) and the corresponding R_ϕ and R_η variables, for unconverted photons with $0.00 < |\eta| < 0.20$. The black points correspond to the pseudo data, red points to non-reweighted simulation and blue points to the reweighted simulation from $Z \rightarrow ll\gamma$ with derived photon reweighting.

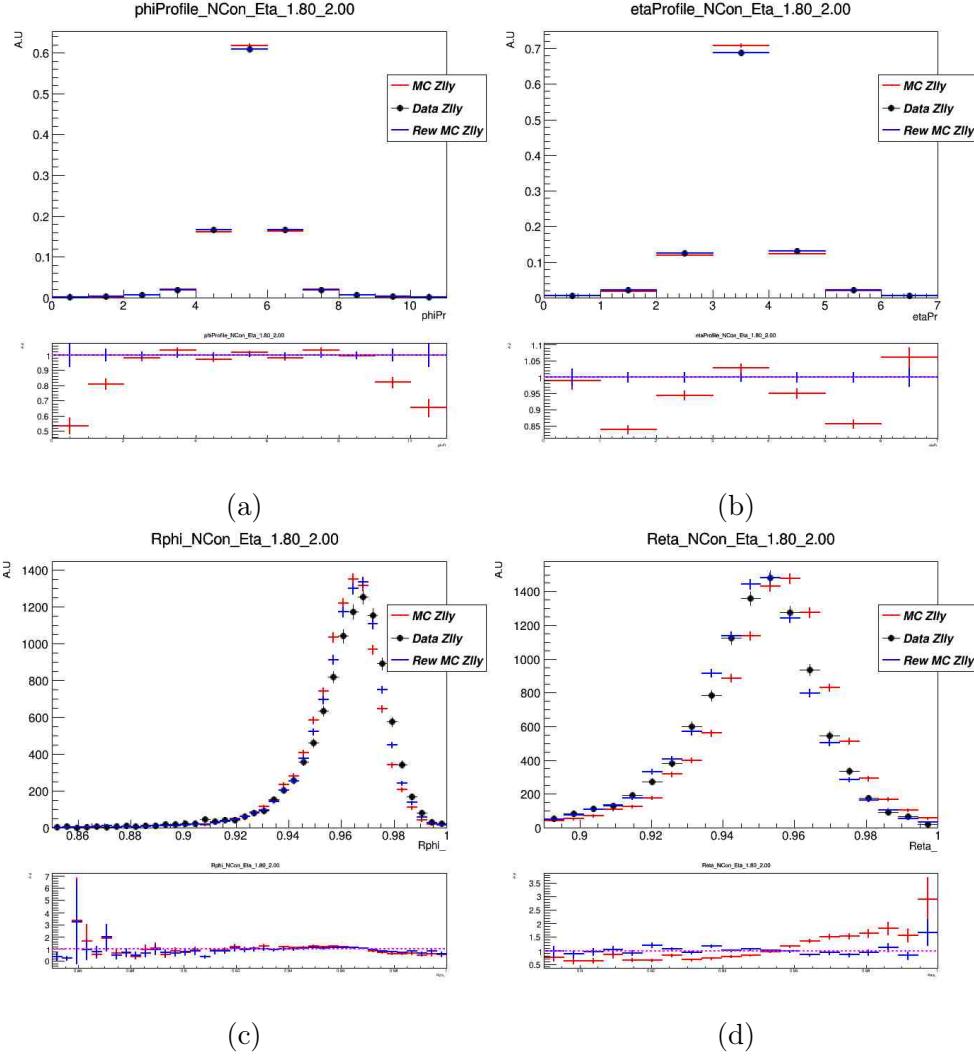


Figure A.6: The energy profile in ϕ and η directions (a,b) and the corresponding R_ϕ and R_η variables, for unconverted photons with $1.80 < |\eta| < 2.00$. The black points correspond to the pseudo data, red points to non-reweighted simulation and blue points to the reweighted simulation from $Z \rightarrow ll\gamma$ with derived photon reweighting.

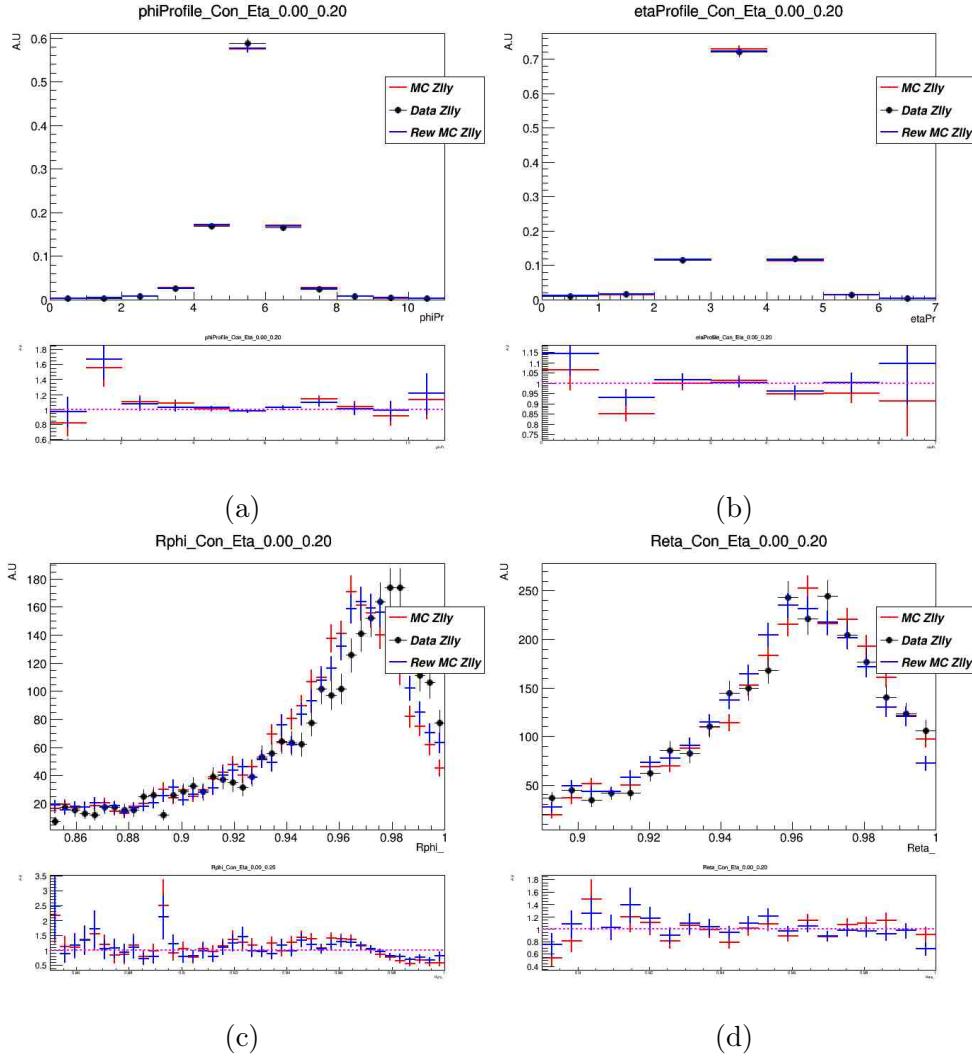


Figure A.7: The energy profile in ϕ and η directions (a,b) and the corresponding R_ϕ and R_η variables, for converted photons with $0.00 < |\eta| < 0.20$. The black points correspond to the pseudo data, red points to non-reweighted simulation and blue points to the reweighted simulation from $Z \rightarrow ll\gamma$ with derived photon reweighting.

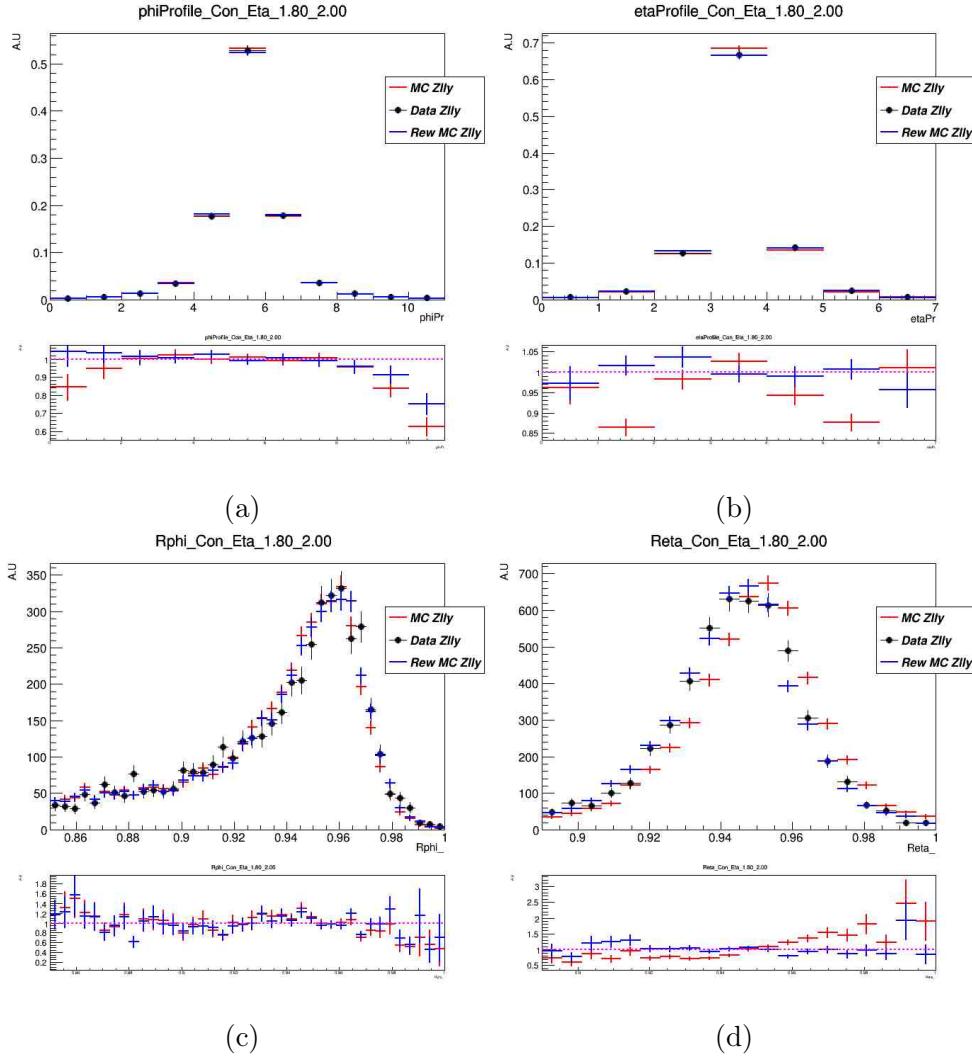


Figure A.8: The energy profile in ϕ and η directions (a,b) and the corresponding R_ϕ and R_η variables, for converted photons with $1.80 < |\eta| < 2.00$. The black points correspond to the pseudo data, red points to non-reweighted simulation and blue points to the reweighted simulation from $Z \rightarrow ll\gamma$ with derived photon reweighting.

²⁶⁶⁸ Appendix B

²⁶⁶⁹ Additional details on Convolutional Neural ²⁶⁷⁰ Network ID

²⁶⁷¹ Additional control plots for CNN ID are shown here.

²⁶⁷² B.1 Background rejection for CNN Tight and Loose WPs

²⁶⁷³ The background rejections of the CNN Tight and Loose WPs are shown in Figure B.1. Note that the
²⁶⁷⁴ two WPs are optimized to deliver similar background rejection as the cut-based in bin of $|\eta|$. The
²⁶⁷⁵ optimization is done in the $Z \rightarrow ll\gamma$ events, while figures shows the background rejection in the Z+jets
²⁶⁷⁶ sample. This serves as an additional validation of the CNN performance.

²⁶⁷⁷ B.2 CNN efficiency as a function of $|\eta|$

²⁶⁷⁸ The CNN efficiency as a function of $|\eta|$ is shown in the following Figures B.2, B.3 and B.4. A comparison
²⁶⁷⁹ with the cut-based efficiencies is also shown.

²⁶⁸⁰ B.3 CNN efficiency in additional $|\eta|$ bins

²⁶⁸¹ Figure B.5 shows the efficiency of the CNN as evaluated using the template fit method for additional
²⁶⁸² $|\eta|$ bins not provided in the main body of the thesis.

²⁶⁸³ B.4 Purity evaluation using the template fit method

²⁶⁸⁴ The template fit results are shown in Figures B.6, B.7, B.8 and B.9.

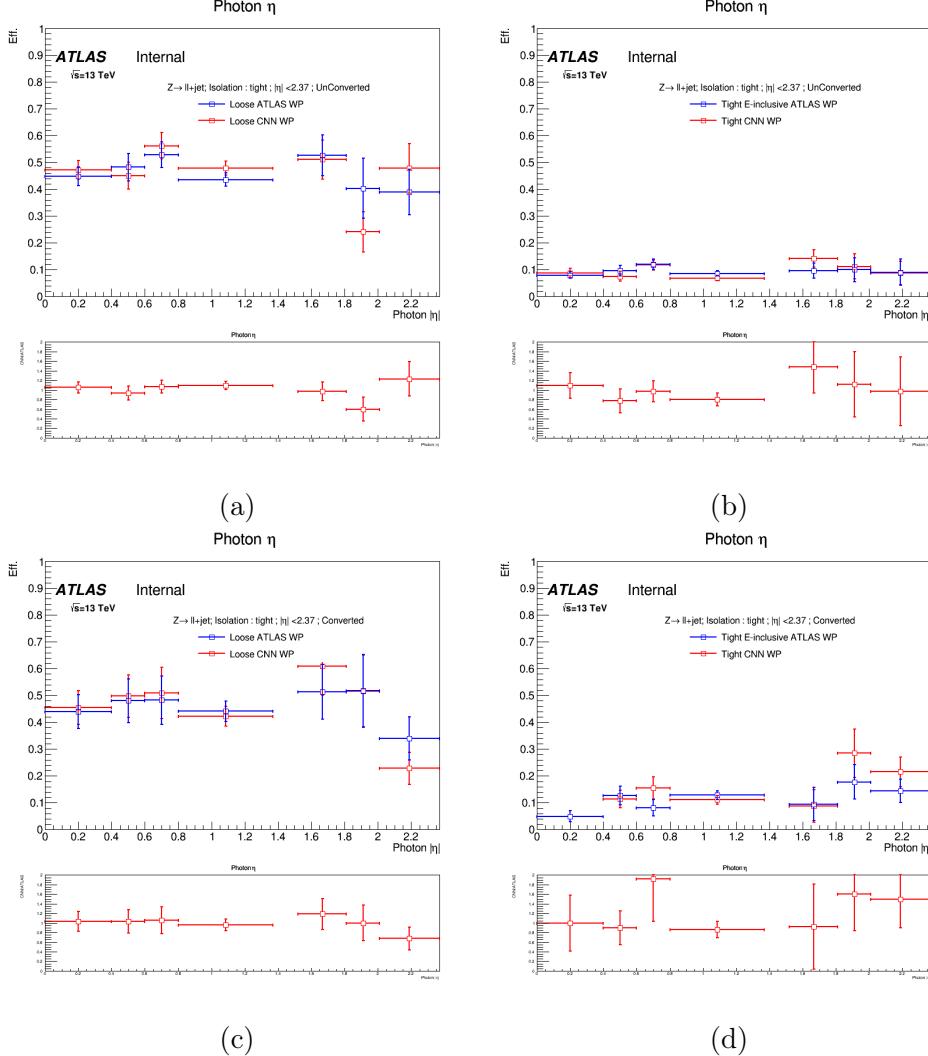


Figure B.1: Background rejection of the CNN tight (right) and Loose (left) WPs. For unconverted (top) and converted (bottom) photons in bin of $|\eta|$. The rejection is evaluated in the $Z + \text{jets}$ sample. The red line shows the rejection for CNN and blue line shows the cut-based rejection.

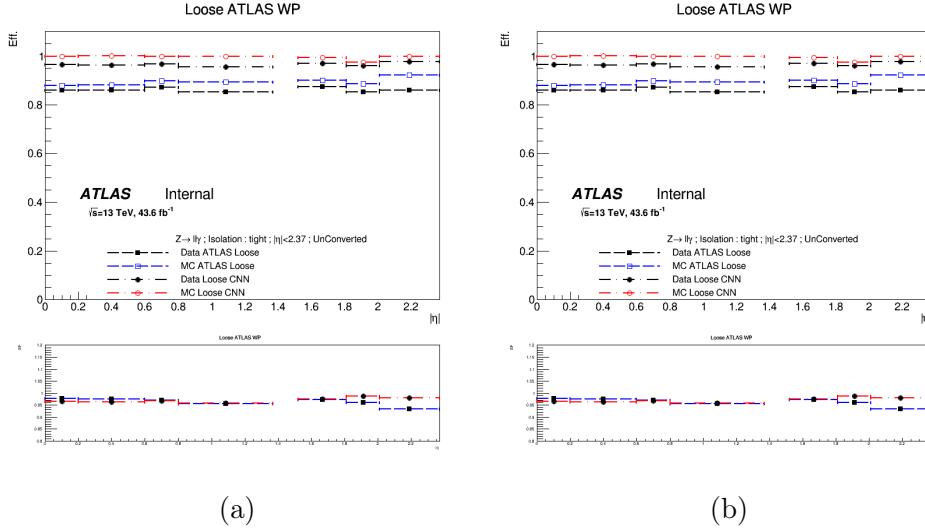


Figure B.2: A comparison between the CNN (red) and cut-based (blue) loose WPs for unconverted (right) converted (left) photons in bin of $|\eta|$. The ratio plot shows the scale factors.

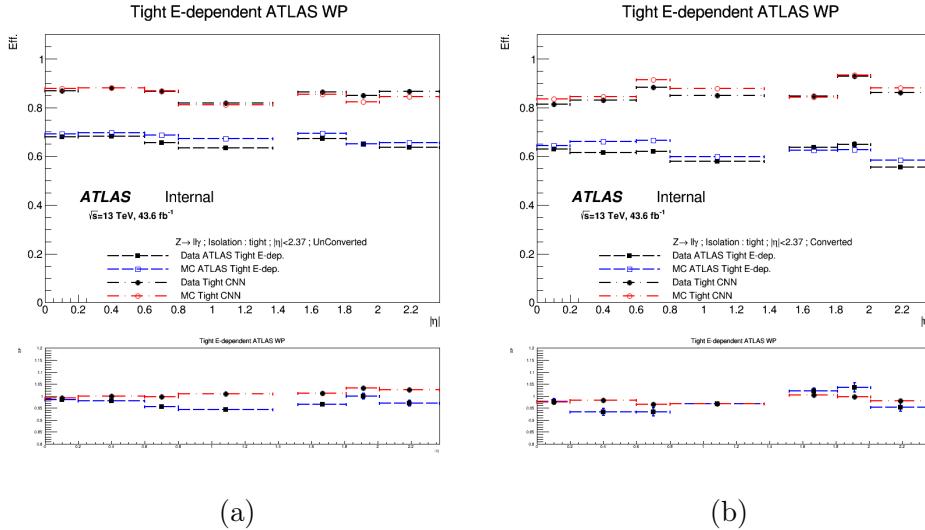


Figure B.3: A comparison between the tight CNN (red) and tight E_T -dependent cut-based (blue) WPs for unconverted (right) converted (left) photons in bin of $|\eta|$. The ratio plot shows the scale factors.

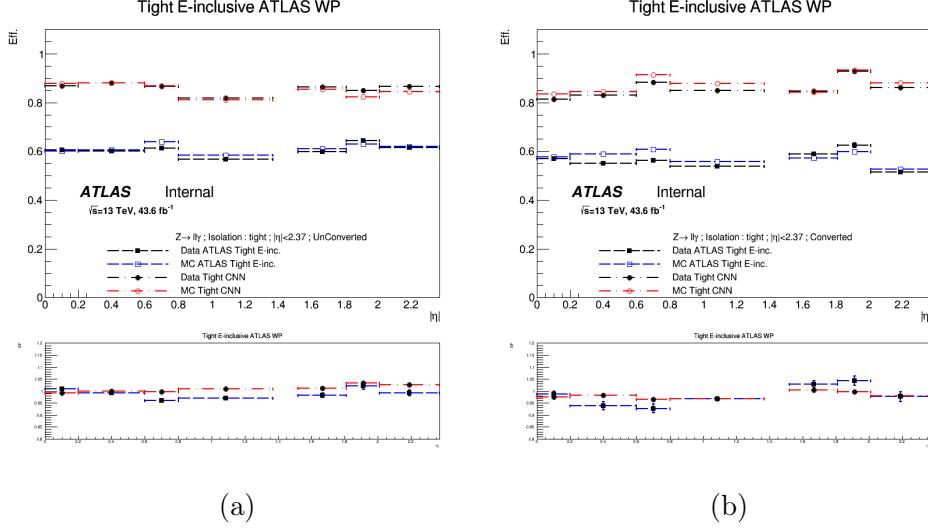


Figure B.4: A comparison between the tight CNN (red) and tight E_T -independent cut-based WPs for unconverted (right) converted (left) photons in bin of $|\eta|$. The ratio plot shows the scale factors.

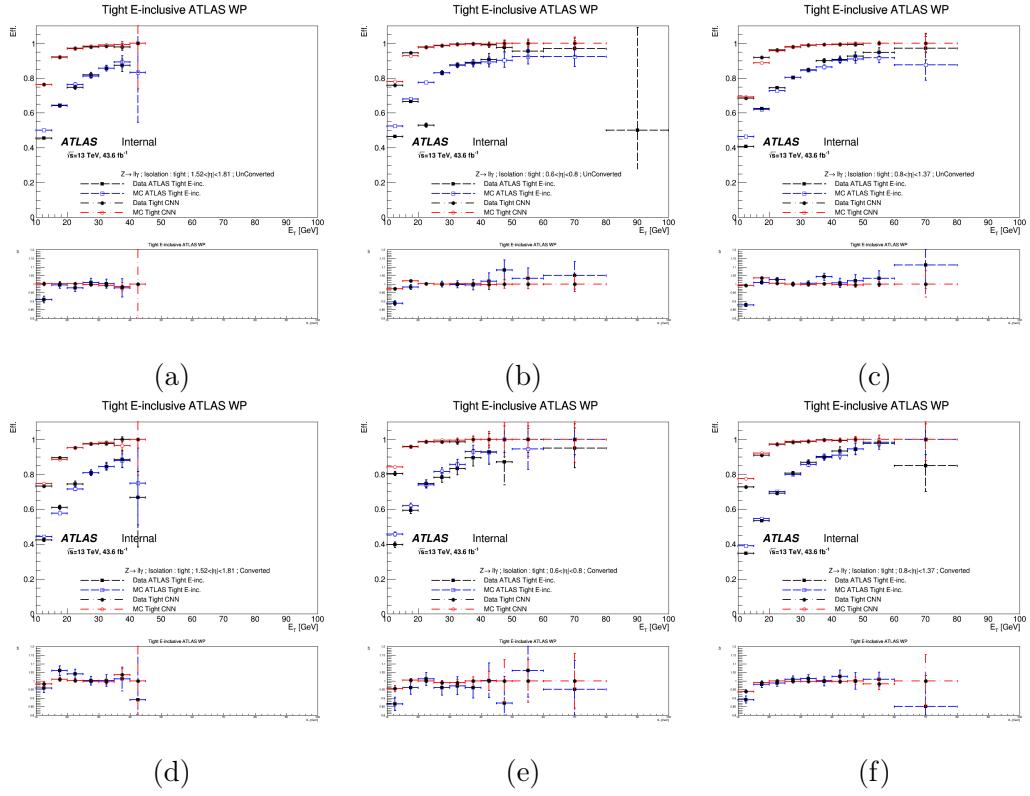


Figure B.5: A comparison between the tight CNN (red) and tight E_T -independent cut-based as a function of photons E_T in different bin of $|\eta|$ for unconverted (top) and converted (bottom) photons. The ratio plot shows the scale factors.

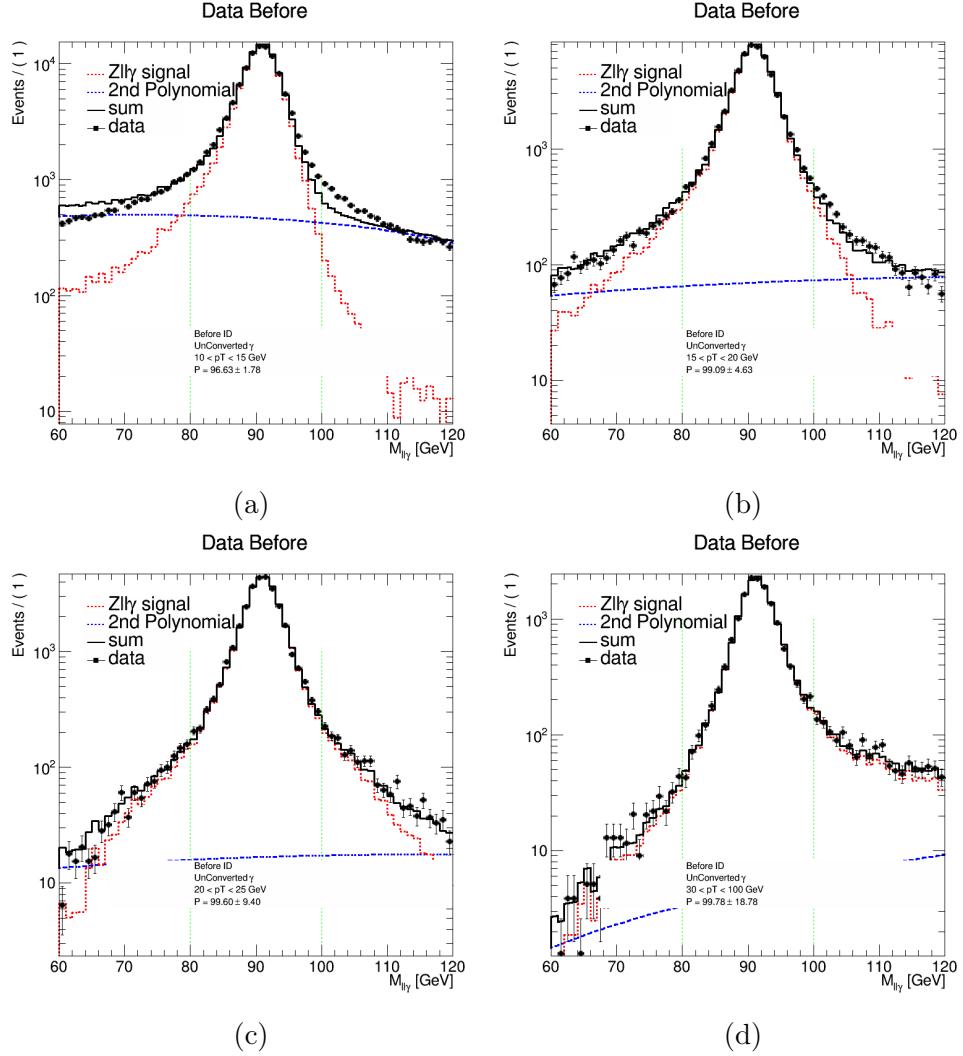


Figure B.6: Invariant mass ($m_{ll\gamma}$) distribution of events selected before tight ID in data (black dots). The solid black line represents the fitting result of the data distribution with the sum of the signal (red dashed line) and the background (blue dashed line). The photons are reconstructed as unconverted.

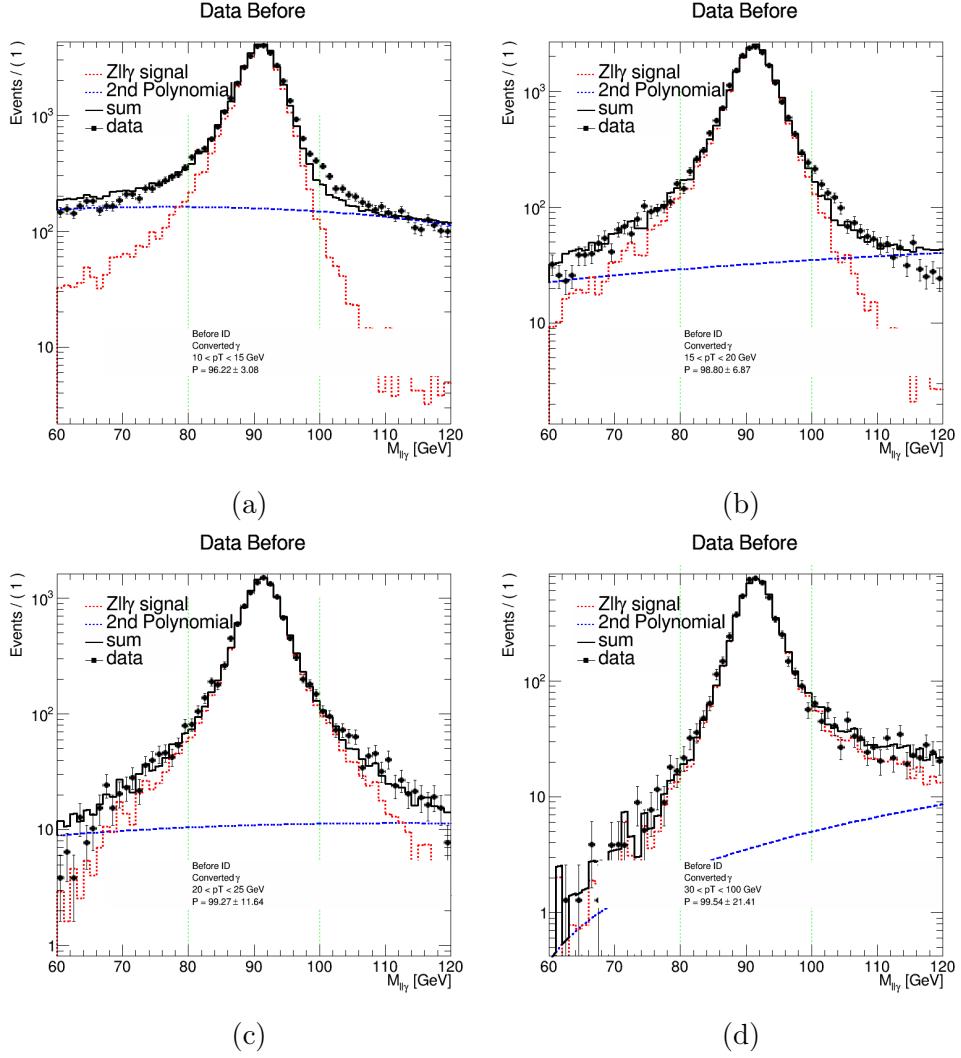


Figure B.7: Invariant mass ($m_{ll\gamma}$) distribution of events selected before tight ID in data (black dots). The solid black line represents the fitting result of the data distribution with the sum of the signal (red dashed line) and the background (blue dashed line). The photons are reconstructed as converted.

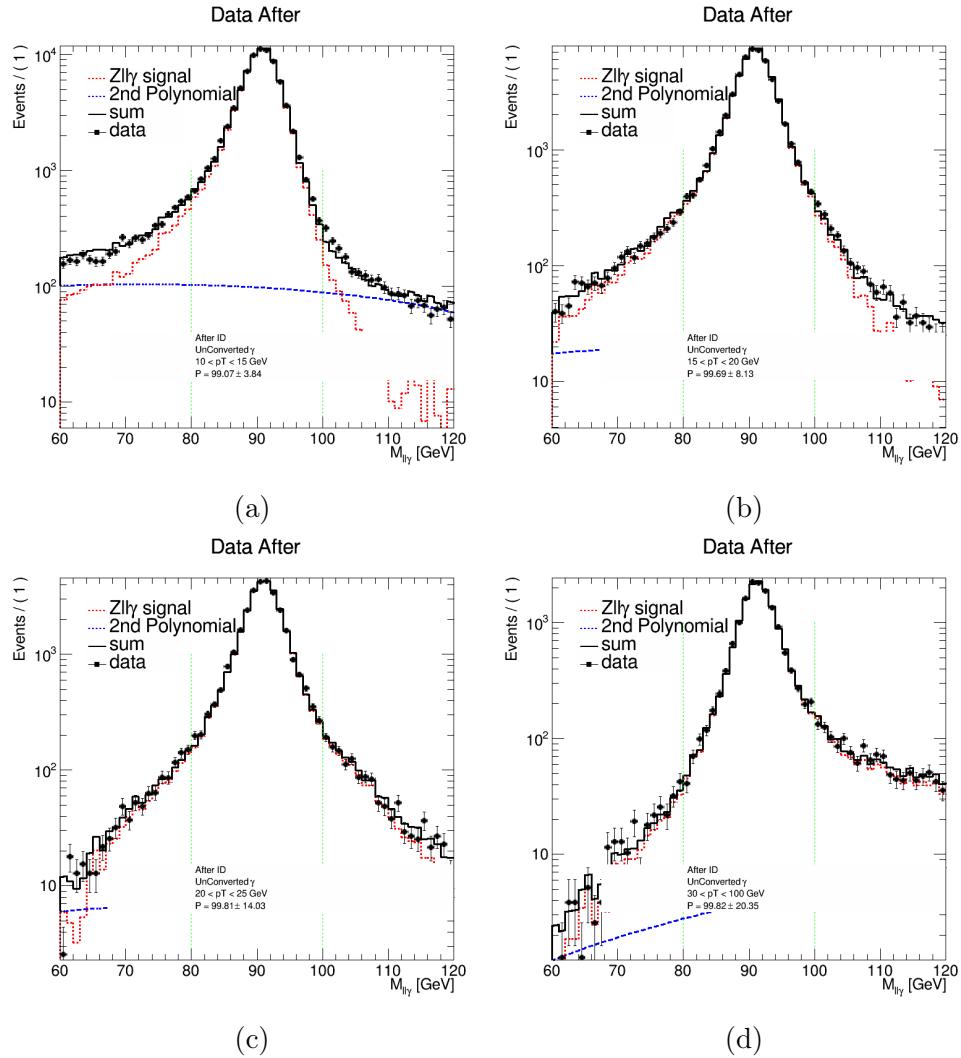


Figure B.8: Invariant mass ($m_{ll\gamma}$) distribution of events selected after tight ID in data (black dots). The solid black line represents the fitting result of the data distribution with the sum of the signal (red dashed line) and the background (blue dashed line). The photons are reconstructed as unconverted.

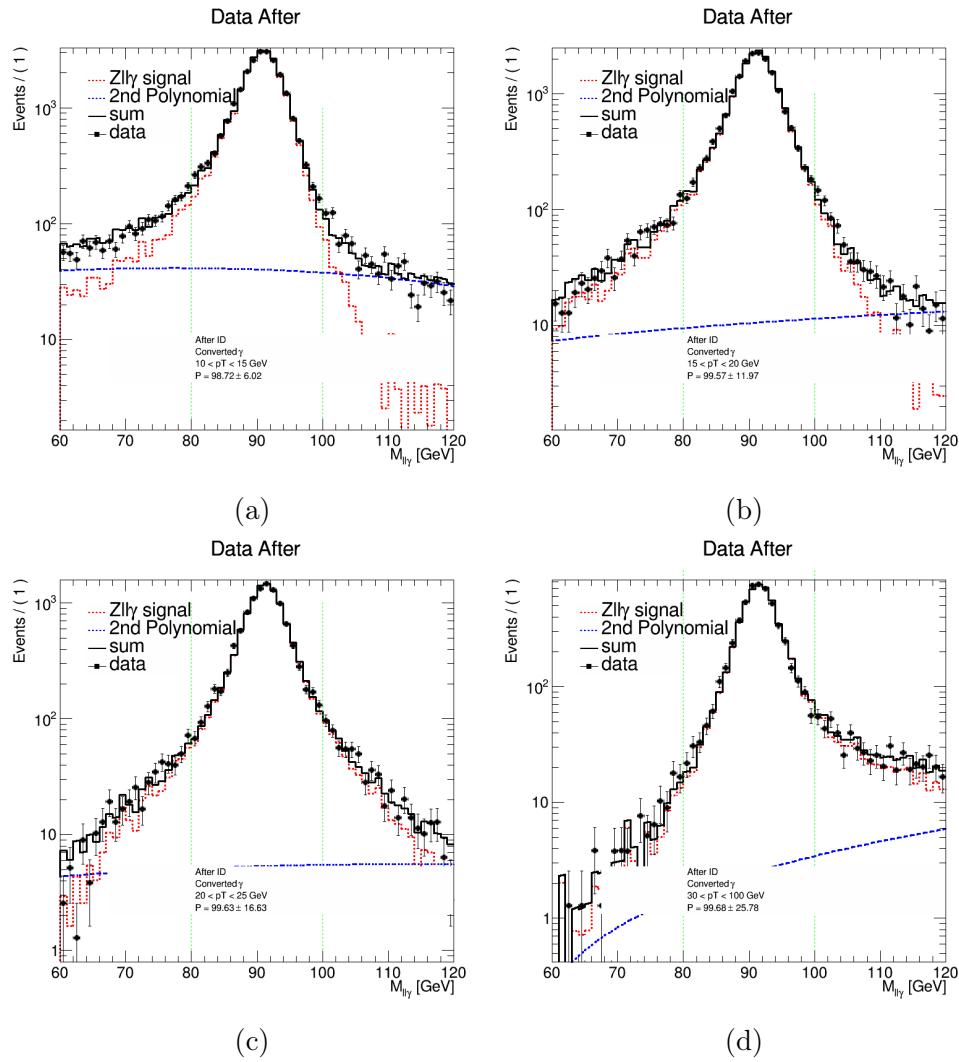


Figure B.9: Invariant mass ($m_{ll\gamma}$) distribution of events selected after tight ID in data (black dots). The solid black line represents the fitting result of the data distribution with the sum of the signal (red dashed line) and the background (blue dashed line). The photons are reconstructed as converted.

2685 Appendix C

2686 Future improvement for Convolutional Neural 2687 Network ID

2688 The developed CNN ID algorithm shows a better performance over the cut-based algorithm. But the
2689 algorithms are still not the state-of-the-art several improvements are possible to improve its performance
2690 in the future. This chapter discusses my propositions to improve the CNN algorithm.

2691 C.1 Low efficiency at high η

2692 As discussed in Section 4.4.2, the CNN performance degrades significantly for low p_T and high η photons.
2693 This is explained by the fact that low energy photons deposit a large fraction of their energy in the first
2694 EM sampling, and given the size of the first sampling images specially at high η , the CNN efficiency for
2695 those photons is very low. This is illustrated in Figure C.1. Note that the training does not take any
2696 information about the angles of the photons.

2697 There are several discussions on how to include η as an input variable to the CNN, avoiding take η
2698 as a discriminant between the signal and background. A reweighting technique to re-weight the η distri-
2699 bution to be similar between signal and background can be used in this case. Personally, developing a
2700 specific convolution layer which takes into account the direction of the photon is the best solution, but
2701 technically complicated. Such convolution layer will be able to define the size of the convolutional filter
2702 needed to extract the shower form first layer images for a given η region.

2703

2704 C.2 Multi-Task Cascade Convolutional Network (MTCCN)

2705 The Multi-Task Cascade Convolutional Network (MTCCN) is an alternative, much easier, solution to
2706 take photon direction into account. The MTCCN is a technique used for face detection and alignment
2707 to detect objects in large images without losing efficiency. The MTCCN consists mainly of three stage.
2708 In the first stage, it produces candidate windows quickly through a shallow CNN. Then, it refines the
2709 windows to reject a large number of non-photon windows through a more complex CNN. Finally, it uses
2710 a more powerful CNN to refine the result and output positions. This technique will be able to define the
2711 convolutional frame for each photon, in this way the network will be able to catch the shower of small
2712 images without be affected by the whitens in the image. The MTCCN is able to work with any size of
2713 image and being unaffected. Thus the motivation to be used with topological cluster as well. The fact
2714 that MTCCN is able to define its convolutional filter for each photon to extract the shower, makes it a

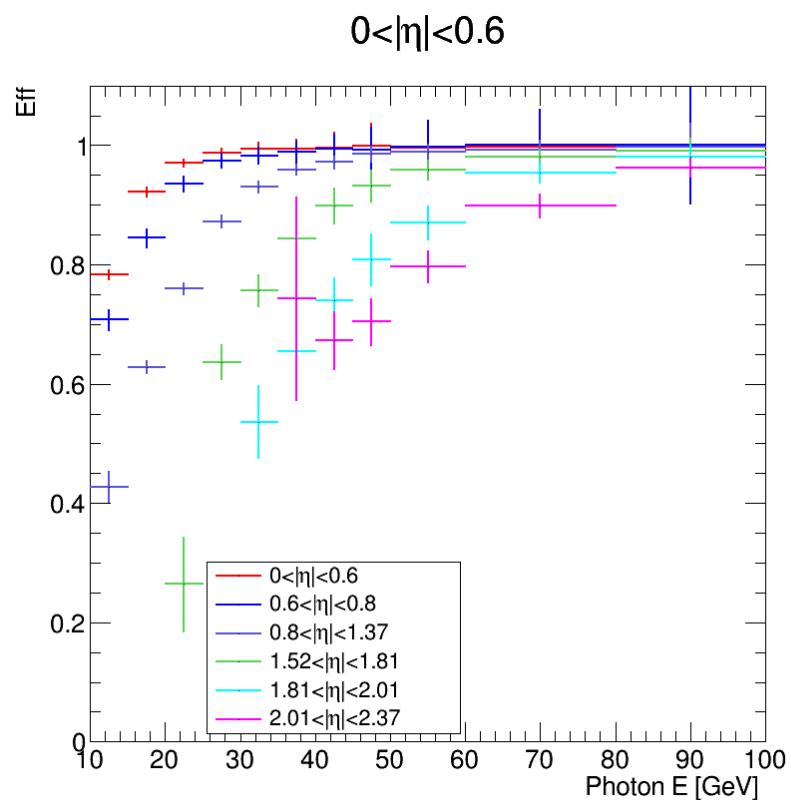


Figure C.1: The CNN efficiency as a function of photon energy for different η regions.

good candidate to handle photon isolation in order to reduce closure uncertainties. Figure C.2 shows an illustration of the MTCCN applied to a topological cluster. The green rectangle shows the global image size, the dashed rectangles show the candidate windows defined by the CNN at the first stage, the solid rectangle is the windows photons defined at stage 2. Dots are the position used to describe the shower of the photon in more details.

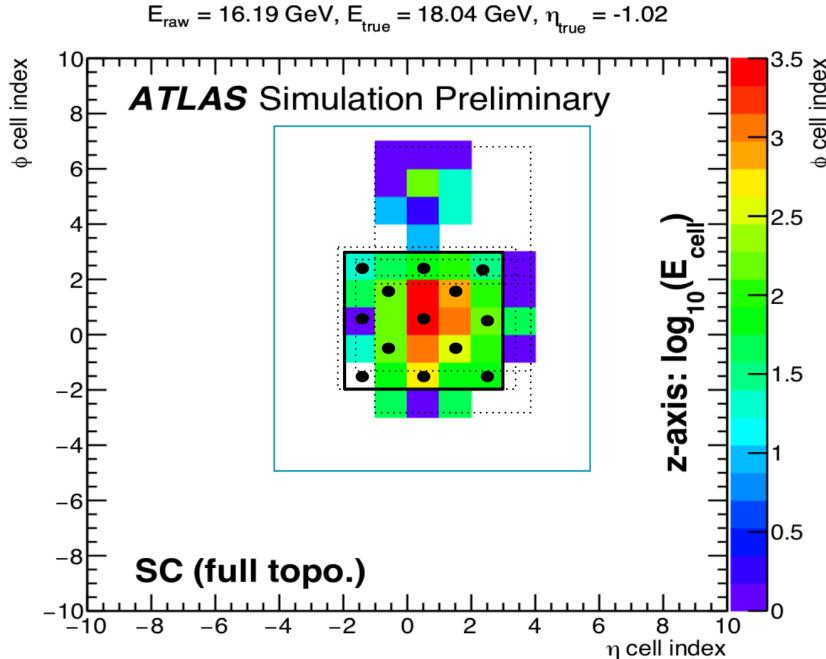


Figure C.2: Schema for MTCCN applied to a topological cluster.

2719

C.3 Pile-Up dependency

The second issue of the CNN is efficiency degradation for increasing the pile-up events level. This effect will be critical at HL-LHC. In medicine area, machine learning is used to extract noise from medical images using Auto-Encoder technique. The Auto-Encoder is used to learn the latent representation of the image which is used later to generate the image without noise as schematized in Figure C.3.



Figure C.3: Image de-noising with Auto-Encoder

The same idea can be implemented to reduce pile-up dependency of shower images. An auto-encoder could be trained for each layer using two exact generated sample one with pile-up and one without pile-up, the auto-encoder is trained to generate the non-pile images from the one with pile-up. The produced

2728 images are then passed to the CNN for classification task. The new Pile-Up Auto-Encoder Denoising
2729 (PAED) algorithm to remove pile-up is shown in Figure C.4.

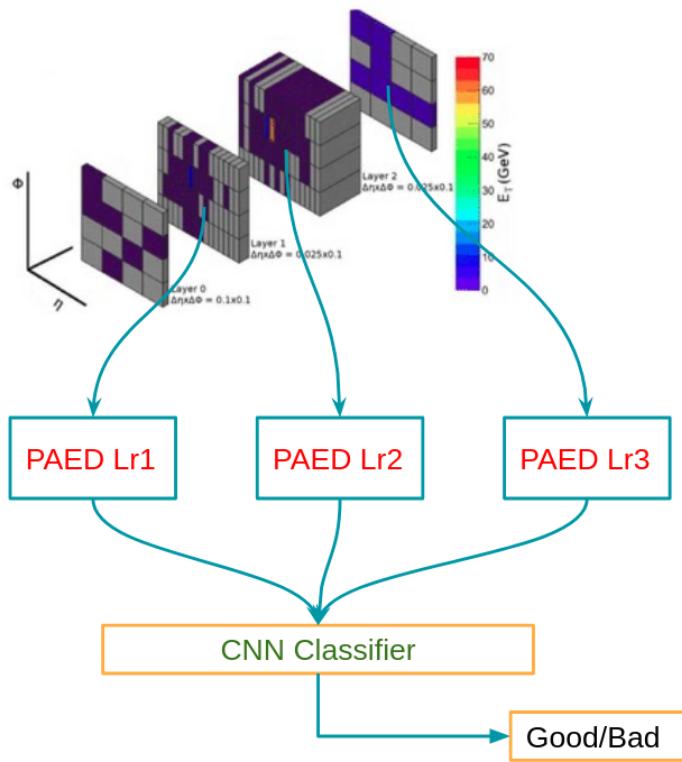


Figure C.4: The CNN classifier with PAED schema.

²⁷³⁰ Appendix D

²⁷³¹ Kinematic Fit for $HH \rightarrow b\bar{b}\gamma\gamma$

²⁷³² In the VH(bb) analysis, the good energy resolution of the $Z \rightarrow ll$ part is used to constraint the $H \rightarrow b\bar{b}$
²⁷³³ through a log likelihood fit used to constraint the balance of the VH system, thus improving the resolution
²⁷³⁴ of the $H \rightarrow b\bar{b}$ component. Similarly to VH(bb), the $HH \rightarrow b\bar{b}\gamma\gamma$ analysis benefits from the likelihood
²⁷³⁵ fit to improve the b -jet resolution. Results obtained with the likelihood fit to $HH \rightarrow b\bar{b}\gamma\gamma$ is discussed
²⁷³⁶ here.

²⁷³⁷ D.1 Kinematic Fit (KF)

²⁷³⁸ In the $HH \rightarrow b\bar{b}\gamma\gamma$ channel, the 4-momenta of the $H \rightarrow \gamma\gamma$ component are reconstructed with a few
²⁷³⁹ % precision, good photon reconstruction with the LAr calorimeter. On the opposite, the $H \rightarrow b\bar{b}$ part
²⁷⁴⁰ suffers from a 20% invariant mass resolution. Assuming that the two components are correlated through
²⁷⁴¹ a good overall balance in the transverse plane as illustrated in Figure D.1, an event-per-event kinematic
²⁷⁴² fit (KF) can be used to calibrate the HH system in the transverse plane, thus additional improvement
²⁷⁴³ of $H \rightarrow b\bar{b}$ resolution on top of the b -jet calibration.

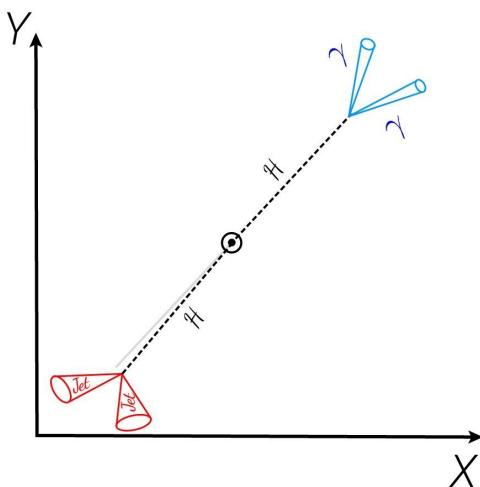


Figure D.1: $HH \rightarrow b\bar{b}\gamma\gamma$ in the transverse plane.

2744 D.2 HH system balance

2745 The per-event fit uses a log-likelihood minimization to balance the HH system. Additional jets in the
 2746 event are degrading the balance of the HH system significantly, making the possible gain achieved with
 2747 the KF significant only in the cases of zero or one additional jet denote 0-jet and 1-jet respectively.
 2748 Figure D.2 shows the Initial State Radiation (ISR) at parton level for ZH process in both qq and ggF
 2749 production mode and for HH process in ggF mode. Due to the dominant ggF mode of HH production
 2750 leading to large jet multiplicity, only $\sim 50\%$ of the events could benefit from this fit compared to ZH
 2751 which is mainly initiated by quark anti-quark. In the following the likelihood fit will be defined only for
 0-jet and 1-jet events.

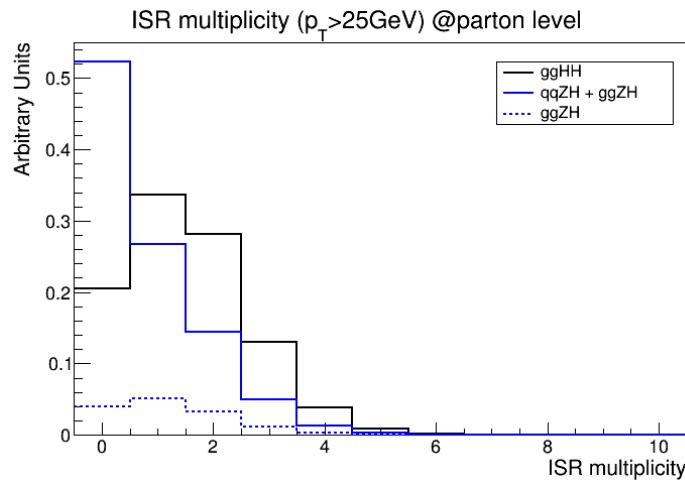


Figure D.2: Number of Initial State Radiation for qqF ZH, ggF ZH and ggF HH.

2752

2753 D.3 Likelihood

2754 The most general expression of the likelihood includes the 3-momenta of the 4 final states which translates
 2755 into 12 inputs as free variables:

- 2756 • The energies of the two photons and the two b -jets,
- 2757 • The pseudo-rapidities and azimuthal angles of the 2 photons and 2 b -jets,
- 2758 • The energy, pseudo-rapidity and azimuthal angles of the additional jet (only for 1-jet event)

2759 and implements 2 constraint terms:

- 2760 • The transverse momentum of the $b\bar{b}\gamma\gamma$ system is constraint to be zero with a width $\sigma_{b\bar{b}\gamma\gamma}$ obtained
 2761 from simulated $HH \rightarrow b\bar{b}\gamma\gamma$ events and optimized for each events type (0-jet or 1-jet) to maximize
 2762 the improvement of $m_{b\bar{b}}$ resolution.
- 2763 • Assume parameters to follow Gaussian distributions.

Given the fit parameters and constraints defined above, the log-likelihood is defined as :

$$-2 \log(\mathcal{L}) = \sum_i \left(\frac{\Omega_i^* - \Omega_i}{\sigma_{\Omega_i}} \right)^2 - 2 \log(L(p_T)) + \left(\frac{\sum p_x^* - 0}{\sigma_{\sigma_{b\bar{b}\gamma\gamma}}} \right)^2 + \left(\frac{\sum p_y^* - 0}{\sigma_{\sigma_{b\bar{b}\gamma\gamma}}} \right)^2. \quad (\text{D.1})$$

This is a product of Gaussian terms of $\Omega = (E, \eta, \phi)$ and a p_T -dependent jet response likelihood $L(p_T)$. The likelihood also includes a term constraining the sum over all $P_{x,y}$ momenta of the system. The energy and angular resolutions of photons are fixed to 1% as well as the angular resolutions of the jets. Jets energy resolution is p_T -depend. Jet response and resolution are evaluated in the $t\bar{t}$ MC samples used to drive the p_T Reco function using parton level as target. Finally, this statistics is used to minimize and balance the reconstructed HH system on an event-by-event basis. It was checked that fixing the direction of the particles did not imply a loss in final resolution while it reduces significantly the computing time to converge. For that, angles are fixed in the fit to their nominal value to improve computational cost as well as to preserve the $m_{\gamma\gamma}$ distribution.

D.4 $\sigma_{b\bar{b}\gamma\gamma}$ determination

An empirical scan is done to find the minimum value of $\sigma_{b\bar{b}\gamma\gamma}$ which correspond to the best $m_{b\bar{b}}$ resolution improvement. For $HH \rightarrow b\bar{b}\gamma\gamma$, the event kinematic is changing with different κ_λ value as illustrated in Figure 2.19, thus changing in the HH system balance. Therefore, for high κ_λ values the event kinematic gets harder to balance and large $\sigma_{b\bar{b}\gamma\gamma}$ are favoured by the fit, thus small improvement in $m_{b\bar{b}}$ is expected for high κ_λ values. Figure D.3 shows the empirical scan of $\sigma_{b\bar{b}\gamma\gamma}$ for different κ_λ for both 0-jet and 1-jet.

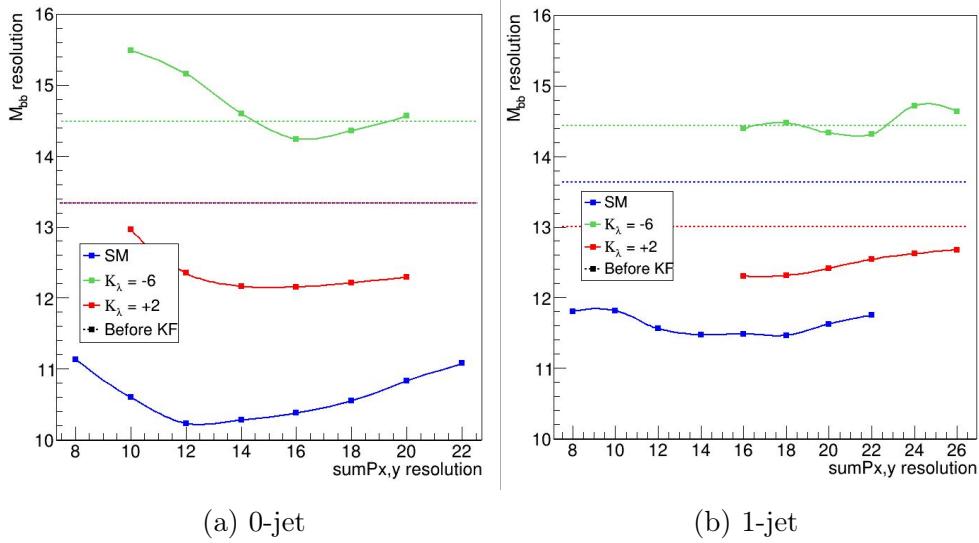


Figure D.3: $m_{b\bar{b}}$ resolution in GeV as a function of $\sigma_{b\bar{b}\gamma\gamma}$ for (a) 0-jet and (b) 1-jet. Blue shows the scan for SM case ($\kappa_\lambda = 1$), red is BSM case of $\kappa_\lambda = 2$ and green is for $\kappa_\lambda = -6$. Lines show $m_{b\bar{b}}$ resolution before the kinematic fit for each case (b -jet calibration).

For 0-jet, the maximum $m_{b\bar{b}}$ resolution improvement is obtained for $\sigma_{b\bar{b}\gamma\gamma} = 12$ GeV for SM, while this value get bigger for large κ_λ . As expected the $\sigma_{b\bar{b}\gamma\gamma}$ is worse for 1-jet than 0-jet given the additional jet in the event. For simplicity, the optimal value of $\sigma_{b\bar{b}\gamma\gamma}$ chosen are 14 and 16 GeV for 0-jet and 1-jet

2783 respectively. This choice is motivated from the di-Higgs combination result which excludes already value
 2784 of κ_λ less than -6.

2785 D.5 Kinematic Fit results

2786 D.5.1 HH system balance

2787 Figures D.4 and D.5 show the p_x and p_y of the HH system before and after the KF for both 0-jet
 2788 and 1-jet events. After the KF, the HH system balance is more constrained to zero as expected and
 2789 significance improvement in p_x and p_y is achieved.

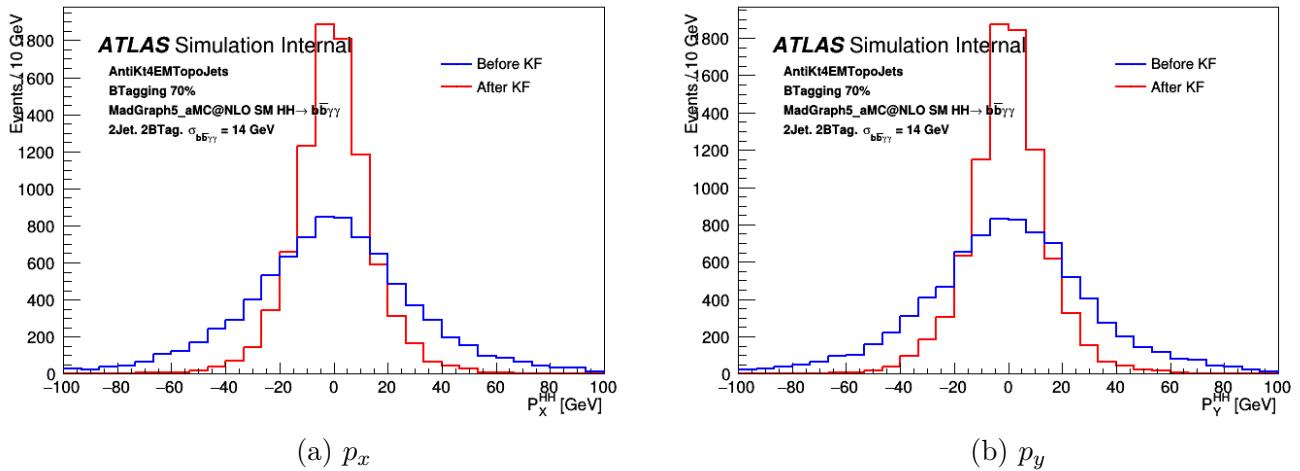


Figure D.4: p_x and p_y distributions of HH after (red) and before (blue) applying the kinematic fit for 0-jet events. The $\sigma_{b\bar{b}\gamma\gamma} = 14$ GeV is used. The kinematic fit is applied on top of the b -jet calibration.

2790 D.5.2 $m_{b\bar{b}}$ resolution

2791 Figure D.6 shows the $m_{b\bar{b}}$ distribution after the kinematic fit for both 0-jet and 1-jet. An additional
 2792 10% (5%) improvement is achieved in $m_{b\bar{b}}$ resolution for 0-jet (1-jet) events. Figure D.7 shows the $m_{\gamma\gamma}$
 2793 distribution after and before the kinematic fit for both 0-jet and 1-jet: no significant change is observed
 2794 as expected.

2795 D.6 Conclusion

2796 An additional improvement in $m_{b\bar{b}}$ resolution is achieved with the kinematic fit, and no changes in $m_{\gamma\gamma}$ is
 2797 observed. In addition to that, no artificial peak is created in the $m_{b\bar{b}}$ distribution for the continuum or the
 2798 $t\bar{t}H$ backgrounds and an improvement in the Z peak resolution is obtained for the ZH background. The
 2799 developed kinematic fit is not used in the final analysis because a bug was found in its implementation
 2800 after samples production, and fixing the bug for this round would have delayed the analysis. The large

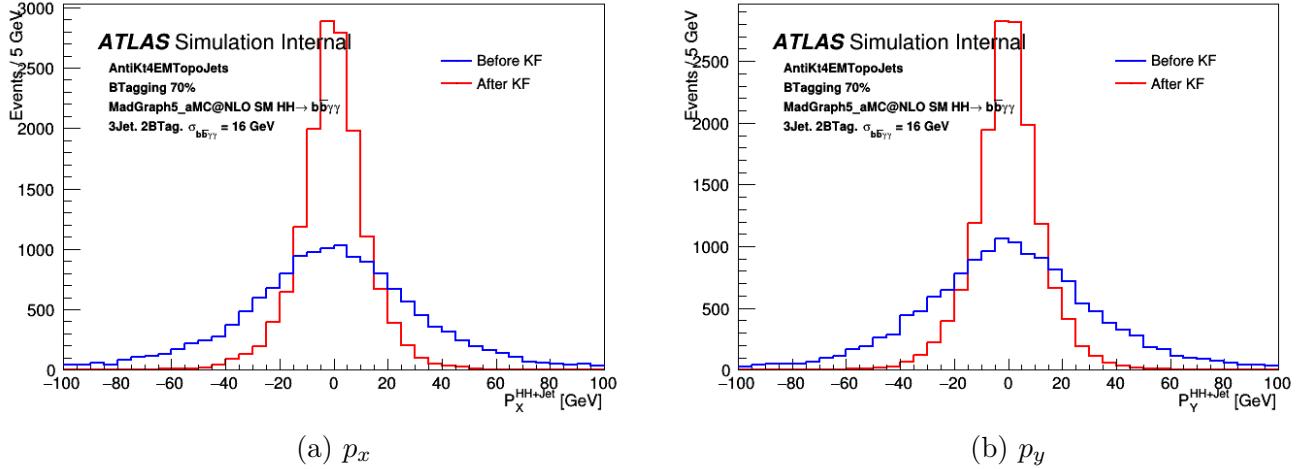


Figure D.5: p_x and p_y distributions of HH after (red) and before (blue) applying the kinematic fit for 1-jet events. The $\sigma_{b\bar{b}\gamma\gamma} = 14$ GeV is used. The kinematic fit is applied on top of the b -jet calibration.

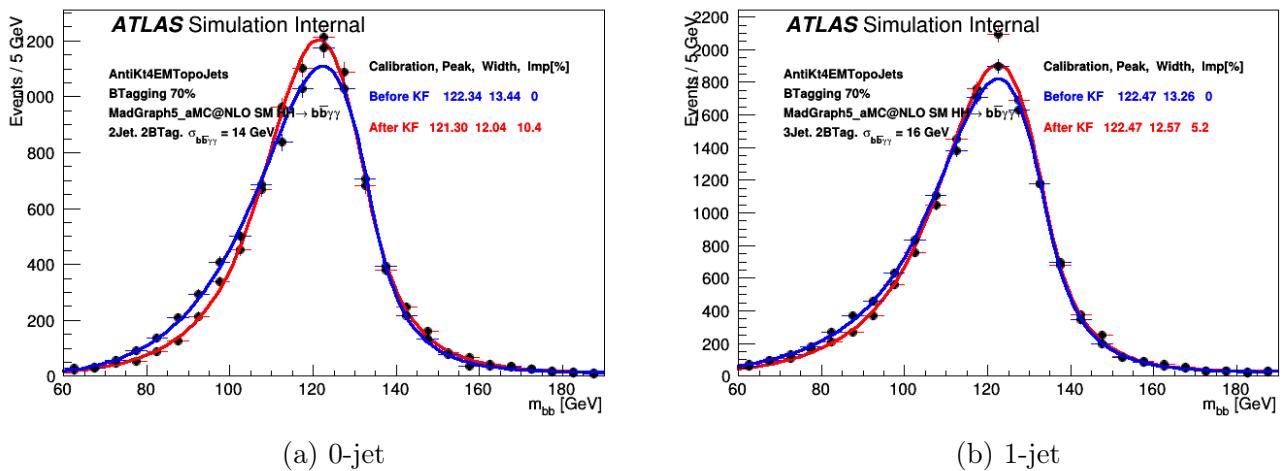


Figure D.6: m_{bb} distributions after (red) and before (blue) applying the kinematic fit for (a) 0-jet and (b) 1-jet events. The kinematic fit is applied on top of the b -jet calibration.

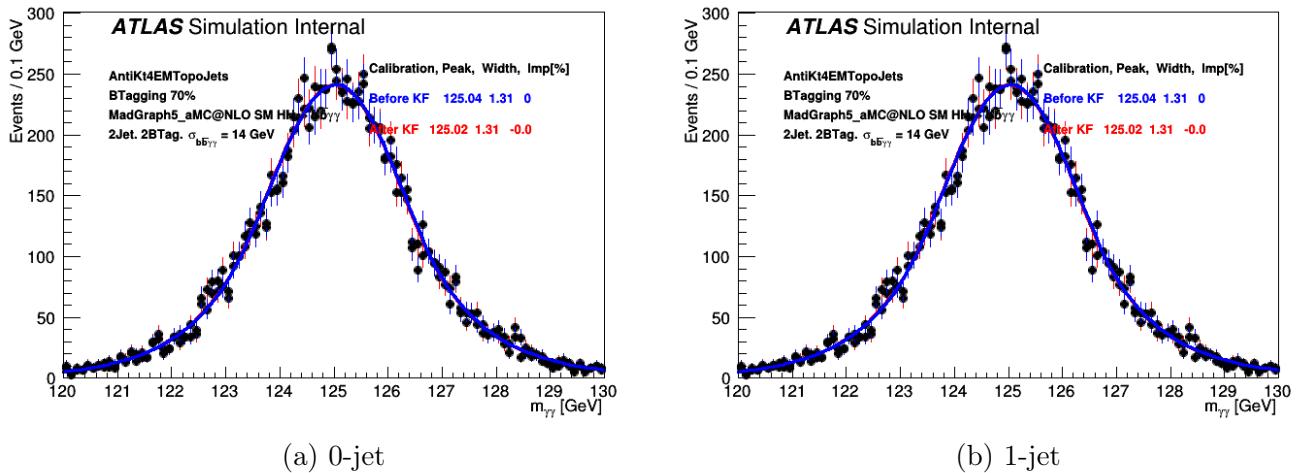


Figure D.7: $m_{\gamma\gamma}$ distributions after (red) and before (blue) applying the kinematic fit for (a) 0-jet and (b) 1-jet events.

jet ISR is critical for $HH \rightarrow b\bar{b}\gamma\gamma$ KF, including information about the event soft radiation in the fit is mandatory to improve its performance. The soft-term MET describes the soft radiation in the event. It could be included to improve the HH system balance. Figure D.8 shows the p_x of the HH system for qqZH, ggZH and ggHH for different events ISR multiplicity with and without including the soft radiation. Similar p_x resolution is achieved for both qq and ggF initiated process, thus the motivation to improve the KF by including soft-term MET in the likelihood.

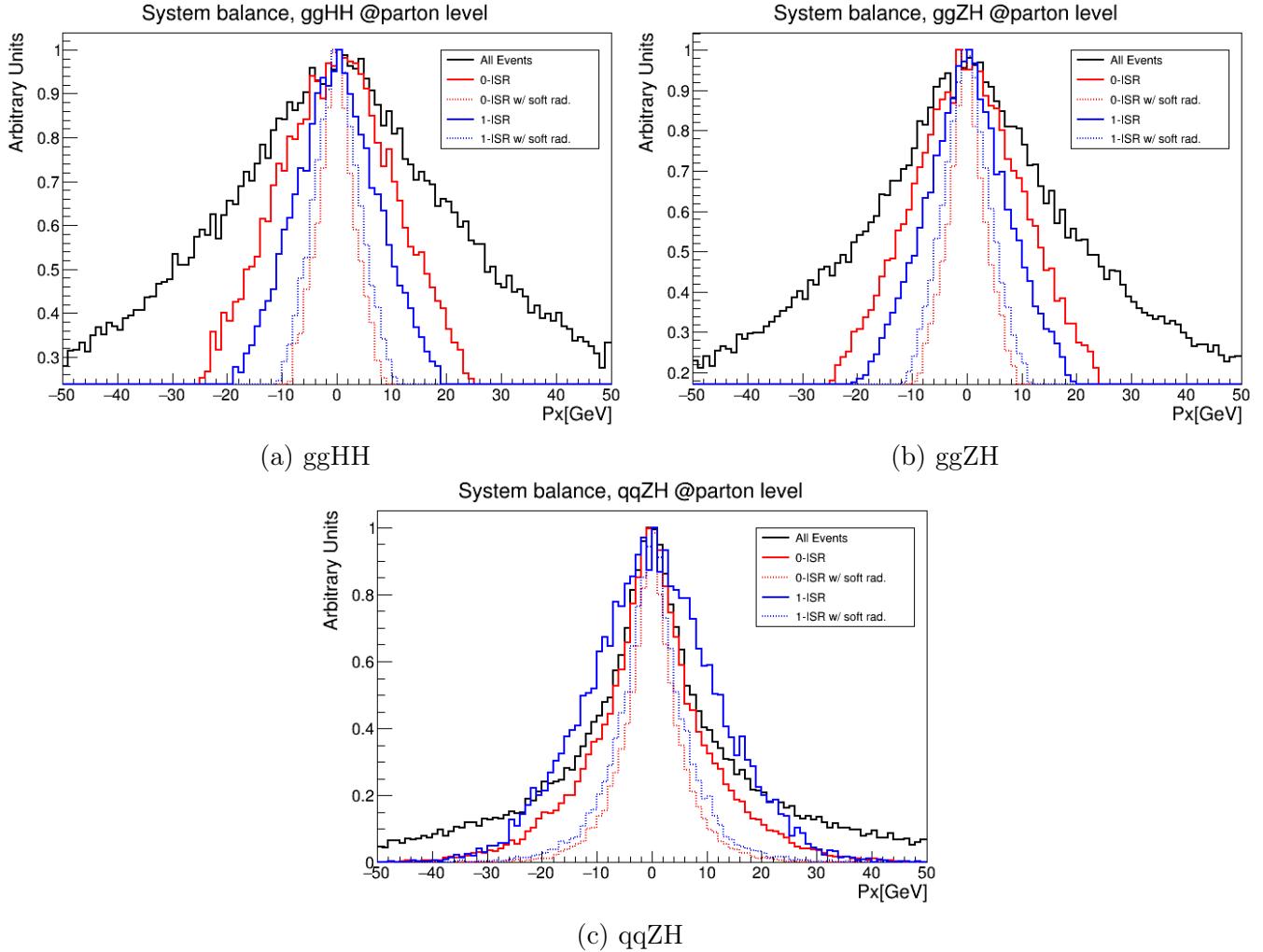


Figure D.8: p_x distribution for different ISR multiplicity before (solid line) and after (dashed line) including the soft radiation to compute the p_x for different processes.

2807 List of Figures

2808	2.1	Shape of the Higgs potential for $\mu^2 < 0$. This potential has an infinite amount of minima.	10
2809		[6]	
2810	2.2	The χ^2 function for the SM as a function of the Higgs boson mass, LEP and Tevatron	11
2811		exclusions limits are shown.	
2812	2.3	The CT14 PDFs at $Q=2$ GeV (left) and $Q = 100$ GeV (right) for different partons. These	12
2813		PDFs are computed at the next-to-next-to-leading order.	
2814	2.4	Feynman diagrams for the main production modes of Higgs boson.	13
2815	2.5	Cross-sections for different Higgs boson production modes at a proton-proton collider	
2816		with a centre-of-mass energy of 6-15 TeV (a). A Higgs boson mass of 125 GeV is assumed	
2817		in this plot. Additionally, cross-sections as a function of Higgs boson mass are shown (b)	
2818		[13].	14
2819	2.6	Higgs boson total decay width as a function of Higgs boson mass [15].	15
2820	2.7	Higgs boson branching ratio of the possible Higgs boson decay channels as a function of	15
2821		the Higgs boson mass [16]	
2822	2.8	Feynman diagrams of the $H \rightarrow \gamma\gamma$ decay. Other particles can also contribute to the loop.	16
2823	2.9	Distributions of the invariant mass of di-photon system [18].	16
2824	2.10	The observed local p-value as a function of m_H (solid line) and the expectation with its	
2825		$\pm 1\sigma$ band assuming the presence of a Standard Model Higgs boson at that mass (dashed	
2826		line) from the combination of the ZZ^* and $\gamma\gamma$ channels by the ATLAS experiment. The	
2827		horizontal dashed lines indicate the p-values corresponding to significance of 1 to 6σ [18].	17
2828	2.11	The inclusive diphoton invariant mass distribution [21].	17
2829	2.12	Cross-section times branching fraction for $ggF+b\bar{b}H$, VBF , VH and $t\bar{t}H+tH$ production,	18
2830		normalized to their SM predictions [21].	
2831	2.13	The particle content of the standard model of particle physics.	18
2832	2.14	Leading Order (LO) Feynman diagrams contributing to ggF Higgs boson pair production	
2833		through a top-quark loop (box) and through the triple self-coupling of the Higgs boson	
2834		(triangle).	19
2835	2.15	HH production modes, in decreasing order of cross-section. λ referees to Higgs boson	
2836		trilinear coupling: (a) ggF mode, (b) VBF mode, (c) Vector boson association (VH) and	
2837		(d) top associated production.	20
2838	2.16	Total cross-sections at the NLO for HH production as a function of centre-of-mass energy.	
2839		Assuming κ_λ and κ_t equals to their SM values.	21
2840	2.17	Di-Higgs boson system decay branching ratios assuming SM Higgs bosons with $m_H =$	21
2841		125.09 GeV.	
2842	2.18	Illustration of both box and triangle diagrams contribution to di-Higgs boson invariant	
2843		mass spectrum.	23

2844	2.19 Higgs boson pair invariant-mass distribution in ggF for various values of κ_λ assuming κ_t	23
2845	= 1.	
2846	2.20 The total LO and NLO cross-sections for HH production at $\sqrt{s} = 14$ TeV, as a function	24
2847	of the self-interaction coupling κ_λ	
2848	2.21 The expected and observed 95% CL limits on the non-resonant production cross-section	24
2849	$\sigma_{gg \rightarrow HH}$ as a function of κ_λ . The red line indicates the predicted HH cross-section as a	
2850	function of κ_λ with all other couplings fixed at their SM values. The red band indicates	
2851	the theoretical uncertainty of this prediction.	
2852	3.1 Overview of the CERN accelerator complex, including the LHC and its pre-accelerators	27
2853	[26]. The four main LHC experiments are depicted, too.	
2854	3.2 Cumulative luminosity versus time delivered to (green), recorded by ATLAS (yellow),	29
2855	and certified to be good quality data (blue) during stable beams and for pp collisions at	
2856	7 and 8 TeV centre-of-mass energy in 2011 and 2012 (Run-1).	
2857	3.3 Luminosity delivered by the LHC during the Run-2 data taking. ATLAS recorded this	29
2858	data with an efficiency above 90%.	
2859	3.4 Recorded integrated luminosity as function of the mean number of interaction per bunch	30
2860	crossing in pp collisions recorded by the ATLAS detector during Run-2 [27].	
2861	3.5 Sketch of the ATLAS detector with its different sub-detectors.	31
2862	3.6 Coordinate system used by the ATLAS and CMS experiments at the LHC.	32
2863	3.7 Layout of the Inner Detector (ID) [31].	32
2864	3.8 (a) Transverse view of 3 of the Insertable-B-Layer (IBL) staves, located directly on the	33
2865	beam pipe. (b) The layout of one of the 14 IBL staves [31].	
2866	3.9 Unfolded longitudinal impact parameter resolution measured from data in 2015, $\sqrt{s} = 13$	34
2867	TeV, with the Inner Detector including the IBL, as a function of η for values of $0.4 <$	
2868	$p_T < 0.5$ GeV compared to that measured from data in 2012 [33].	
2869	3.10 The average efficiency to reconstruct primary tracks with a production vertex before the	34
2870	first layer in jets as a function of jet p_T . The same sample generation, with limited	
2871	statistics, is used for both reconstruction algorithms, resulting in correlated features [35].	
2872	3.11 Rejection factor against light jets as a function of b-jet efficiency for the combined	35
2873	IP3D+SV1 tagger. Run-1 and Run-2 performances are compared.	
2874	3.12 The Perigee representation of the track [38].	37
2875	3.13 ATLAS Calorimeter system.	37
2876	3.14 Accordion shape of the EM calorimeter.	39
2877	3.15 Side view of one quadrant (<i>right</i>) and transverse view (<i>left</i>) of the muon spectrometer. .	40
2878	3.16 Transverse cross-section of the MDT tube.	41
2879	3.17 Operation plan for the LHC scientific program.	42
2880	3.18 The primary track reconstruction efficiency in minimum bias simulated samples containing	44
2881	exactly one and on average 21 or 41 interactions. The distributions are shown for tracks	
2882	passing the default(dashed) and robust (solid) requirements.	
2883	3.19 The vertex reconstruction efficiency (<i>left</i>) and fake probability (<i>right</i>) as a function of	44
2884	the average number of interactions.	
2885	3.20 Diagram of the super clustering algorithm for electrons and photons. Seed clusters are	45
2886	shown in red, satellite clusters in blue.	
2887	3.21 The cluster, track, cluster and track, and electron reconstruction efficiencies as a function	
2888	of the generated electron E_T	46

2889	3.22	The converted photon reconstruction efficiency and contributions of the different conversion types as a function of E_T^{true}	46
2890	3.23	Flowchart showing the logic of the ambiguity resolution for particles initially reconstructed both as electrons and photons.	47
2891	3.24	Schematic overview of the calibration chain for the electron and photon energies [65].	49
2892	3.25	Schema of the calorimeter isolation method: the grid represents the second-layer calorimeter cells in the η and ϕ directions. The candidate electron is located in the centre of the purple circle representing the isolation cone. All topological clusters, represented in red, for which the barycentres fall within the isolation cone are included in the computation of the isolation variable. The 5×7 cells represented by the yellow rectangle correspond to the subtracted cells in the core subtraction method.	50
2893	3.26	Efficiency of the different isolation WPs for electrons from inclusive $Z \rightarrow e^+e^-$ events as a function of the electron E_T . The lower panel shows the ratio of the efficiencies measured in data and in simulations [68].	51
2894	3.27	The transformed LH-based identification discriminant d_L for reconstructed electron candidates with good quality tracks with $30 < E_T < 35$ GeV and $ \eta < 0.6$ [71].	52
2895	3.28	The electron identification efficiency in $Z \rightarrow e^+e^-$ events in data as a function of E_T for the Loose, Medium, and Tight WPs.	53
2896	3.29	Muon identification efficiency for the Loose, Medium and Tight WPs as function of p_T . The panel at the bottom shows the ratio of the measured to predicted efficiencies, with statistical and systematic uncertainties.	54
2897	4.1	Distribution of E_T^{cone40} in data and simulation, in the central region of the detector ($ \eta < 0.6$), separately for converted (left) and unconverted (right) photons after the data-driven shifts are applied.	57
2898	4.2	Efficiency of the isolation working points for converted(left) and unconverted (right) photons as a function of photon E_T . The lower panel shows the ratio of the efficiencies measured in data and in simulation (scale factors). TThe error bars account for the systematic and statistical errors.	57
2899	4.3	Schematic representation of the photon identification discriminating variable [83].	58
2900	4.4	Distributions of the R_η and w_{s3} for converted and unconverted photon candidates with E_T in [10, 50] GeV and $ \eta < 2.37$ selected from radiative Z events for data (black), uncorrected simulation (dashed red line) and corrected simulation (solid blue line).	59
2901	4.5	The photon identification efficiency with E_T -dependent selection, and the ratio of data to simulation efficiencies, for unconverted photons with a Loose isolation requirement applied as pre-selection, as a function of E_T in four different $ \eta $ regions. The combined scale factor, obtained using a weighted average of scale factors from the individual measurements, is also presented; the band represents the total uncertainty.	60
2902	4.6	The photon identification efficiency with E_T -dependent selection, and the ratio of data to simulation efficiencies, for converted photons with a Loose isolation requirement applied as pre-selection, as a function of E_T in four different $ \eta $ regions. The combined scale factor, obtained using a weighted average of scale factors from the individual measurements, is also presented; the band represents the total uncertainty.	61
2903	4.7	Efficiencies of the tight photon identification for unconverted (left) and converted (right) signal photons, plotted as a function of photon E_T . The loose isolation is applied as a pre-selection. For both plots, the bottom panel shows the ratios between the E_T -dependent and the E_T -independent identification efficiencies.	61

2935	4.8 (a) Energy profile in η direction, (b) Energy profile in ϕ direction, (c) R_η variable and 2936 (d) R_ϕ variable before (red) and after reweighting for electrons from $Z \rightarrow e^+e^-$ for 1.80 2937 $< \eta < 2.00$ first column and $0.40 < \eta < 0.60$ in the second column (the bottom plot is 2938 the data/simulation ratio) [87].	63
2939	4.9 The energy profile in ϕ and η directions (a, b) and the corresponding R_ϕ and R_η variables, 2940 for unconverted photons with $0.40 < \eta < 0.60$. The black points correspond to Data 2941 2017, red points to non-reweighted simulation and blue points to the reweighted simulation 2942 from $Z \rightarrow ll\gamma$	65
2943	4.10 The energy profile in ϕ and η directions (a, b) and the corresponding R_ϕ and R_η variables, 2944 for unconverted photons with $0.40 < \eta < 0.60$. The black points correspond to Data 2945 2017, red points to non-reweighted simulation and blue points to the reweighted simulation 2946 from $Z \rightarrow ll\gamma$	66
2947	4.11 The energy profile in ϕ and η directions (a, b) and the corresponding R_ϕ and R_η variables, 2948 for unconverted photons with $0.40 < \eta < 0.60$. The black points correspond to the 2949 pseudo data, red points to non-reweighted simulation and blue points to the reweighted 2950 simulation from $Z \rightarrow ll\gamma$	67
2951	4.12 Average R_η (a) and R_ϕ (b) in real data (black), simulation (blue) and after reweighting 2952 (red) using photon reweighting values.	68
2953	4.13 The energy profile in ϕ and η directions (a, b) and the corresponding R_ϕ and R_η variables, 2954 for unconverted photons with $1.00 < \eta < 1.20$. The black points correspond to the 2955 pseudo data, red points to non-reweighted simulation and blue points to the reweighted 2956 simulation from $Z \rightarrow ll\gamma$	69
2957	4.14 CNN classifier schema.	71
2958	4.15 Illustration of the global CNN classifier graph.	72
2959	4.16 Neural network loss and accuracy as function of epoch number.	73
2960	4.17 Network ROC curve for inclusive (a) unconverted photons (b) converted photons, the 2961 orange line show the ROC curve and the blue dot show the tight WP, x-axis is the ϵ_{eff} 2962 and y-axis show the ϵ_{rej} as defined in Eq. 4.5.	74
2963	4.18 Relative signal efficiency improvement for similar tight WP background rejection (a) for 2964 unconverted photons and (b) converted photons.	75
2965	4.19 The unconverted photon identification efficiency as a function of photon E_T for CNN 2966 (red) and cut-based (blue) indicated with ATLAS, for data (full marks) and simulation 2967 (open marks) for 4 $ \eta $ regions. Bottom ratios show the scale factors.	77
2968	4.20 The converted photon identification efficiency as a function of photon E_T for CNN (red) 2969 and cut-based (blue) indicated with ATLAS, for data (full marks) and simulation (open 2970 marks) for 4 $ \eta $ regions. Bottom ratios show the scale factors.	78
2971	4.21 CNN output distribution as computed for data (black) and simulation (blue) for (a) for 2972 unconverted photons and (b) converted photons with $ \eta < 0.4$	79
2973	4.22 Photon identification efficiency for CNN (red) and cut-based (blue) indicated with AT- 2974 LAS, for data (full marks) and simulation (open marks). Unconverted photons (right) 2975 and converted photons (left). As a function of pile up $\langle \mu \rangle$ (top) and number of primary 2976 vertices (nPV) (bottom). Bottom ratios show the scale factors.	81
2977	5.1 A flow chart of how the particle flow algorithm proceeds, starting with track selection 2978 and continuing until the energy associated with the selected tracks have been removed 2979 from the calorimeter [96].	83

2980	5.2	Calibration stages for jets. Other than the origin correction, each stage of the calibration is applied to the four-momentum of the jet [98].	84
2981	5.3	The relative jet energy resolution in simulation for fully calibrated PFlow+JES jets (green curve) and EM+JES jets (blue curve) as a function of p_T jet.	85
2982	5.4	Ratio of the PFlow+JES (blue curve) and EM+JES (green curve) jet response in data to that in the nominal event generators as a function of p_T jet.	86
2983	5.5	(a) Comparative distribution, scaled to unity, of JVT for truth-matched HS and pile- up. (b) pile-up jet fake rate as a function of the HS-jet efficiency for cuts on the JVT discriminant and JVT input variables. The performance is compared to that of the JVT inputs, corrJVF and R_{p_T} , as well as that of the no corrected JV.	87
2984	5.6	The fJVT distribution for hard-scatter (blue) and pile-up (green) forward jets.	88
2985	5.7	Efficiency for hard-scatter jets as a function of the forward jet p_T	88
2986	5.8	ROC curves for 2018 recommended versions of MV2 and DL1, and 2019 DL1r opti- mization. The x-axis corresponds to the b-jet efficiency, while the y-axis corresponds to light-flavor jets rejection. The shaded bands represent the statistical uncertainty.	90
2987	5.9	Efficiency of identifying b-jets as a function of jet p_T for an inclusive 77% efficiency requirement, for 2018 recommended versions of MV2 and DL1, and 2019 DL1r optimization.	90
2988	5.10	Typical schema of a b-jet.	91
2989	5.11	Calibration stages for jets including energy correction for tagged b-jet.	92
2990	5.12	p_T -Reco correction distribution for (a) topological cluster jets (b) particle flow jets, b- tagging 77% WP is applied with DL1r algorithm.	93
2991	5.13	Relative energy resolution improvement after b-jet energy calibration in bin of truth energy for $t\bar{t}$ using (a) topological cluster jets (b) particle flow jets.	94
2992	5.14	Distribution of $m_{b\bar{b}}$ before calibration (blue) and after calibration (red) in $HH \rightarrow bb\gamma\gamma$ for (a) topological cluster jets (b) particle flow jets.	94
3001	6.1	Di-photon trigger efficiency in data 2015-2018 as a function of (a) E_T and (b) η . The ratios of data to MC simulation efficiencies are also shown.	97
3002	6.2	$m_{bb\gamma\gamma}^*$ distributions for the ggF HH signal with several κ_λ values. $m_{bb\gamma\gamma}^* = 350$ GeV is chosen as the separating boundary between categories targeting the SM and BSM κ_λ signals.	100
3003	6.3	The BDT score for the benchmark signals and the main backgrounds in the low- (a) and high- (b) mass categories. Distributions are normalized to unit area. The dotted lines denote the category boundaries. Events with a BDT score below 0.881 in the low mass region or below 0.857 in the high mass region are rejected.	101
3004	6.4	Distributions of $m_{\gamma\gamma}$ in all signal categories. The simulated continuum background is scaled by the $\gamma\gamma$, γ -jet/jet- γ , and jet-jet fractions and normalized to the data side-band ($m_{\gamma\gamma} \in [105, 120] \cup [130, 160]$ GeV). The background decomposition is described in Section 6.3.2.	102
3005	6.5	Evaluation of the loss function for high mass (red) and low mass (blue) DNNs as a function of epoch number.	104
3006	6.6	Normalized (to unity) confusion matrix for (a) high mass and (b) low mass models. JJ here denote the continuum $\gamma\gamma$ +jets.	105
3007	6.7	d_{HH} discriminant distribution in (a) high mass and (b) low mass. The black line shows the real data in the sideband region ($m_{\gamma\gamma} \in [105, 120] \cup [130, 160]$).	105

3024	6.8 $\gamma\gamma$, γj , and jj fractions in the inclusive (a) and the BDT categories (b). In the BDT figure bin 1 corresponds to the High mass BDT tight category, bin 2 to the High mass BDT loose category, bin 3 to the Low mass BDT tight category, and finally bin 4 to the Low mass BDT loose category.	108
3025		
3026		
3027		
3028	6.9 p_T -Reco correction distribution using different DL1r b -tagging WPs for both semileptonic (a) and hadronic (b) jets. Bottom panel shows the ratio of the p_T -Reco to the nominal (77% WP).	110
3029		
3030		
3031	6.10 The data points are compared to the background-only fit for the four BDT categories. Both the continuum background and the background from single Higgs boson production are considered.	113
3032		
3033		
3034	6.11 Observed and expected limits at 95% CL on the cross-section of Higgs boson pair production as a function of the Higgs boson self-coupling modifier $\kappa_\lambda = \lambda_{HHH}/\lambda_{HHH}^{\text{SM}}$. The $\pm 1\sigma$ and $\pm 2\sigma$ bands show the expected variations on the expected limit due to statistical and systematic uncertainties. The theory prediction curve represents the scenario where all parameters and couplings are set to their SM values except for κ_λ . The uncertainty band of the theory prediction curve shows the cross-section uncertainty.	115
3035		
3036		
3037		
3038		
3039		
3040	6.12 The likelihood function as a function of κ_λ , with $\mu = 1$. The observed 68% and 95% CL ranges are [0.5, 4.7] and [-1.4, 6.5] respectively, while the expected are [-1.4, 6.4] and [-3.1, 8.2].	115
3041		
3042		
3043	6.13 Expected limits at 95% CL on the cross-section of Higgs boson pair production as a function of the Higgs boson self-coupling modifier κ_λ for statistical only and statistical+systematic configurations. The luminosity is equal to 139 fb^{-1}	116
3044		
3045		
3046	6.14 Expected and observed 95% CL upper limits on the $HH \rightarrow b\bar{b}\gamma\gamma$ cross-section obtained for different values of κ_λ assuming $\kappa_t = 1$. The green and yellow bands represent, respectively, the one and two standard deviation. The long-dashed red line shows the theoretical prediction.	119
3047		
3048		
3049		
3050	6.15 Negative log-likelihood, as a function of κ_λ , evaluated with an Asimov data set assuming the SM hypothesis. The 68 and 95% CL intervals are obtained from the intersection with the dashed grey lines. The two curves are shown for the HH (blue) and HH+ $t\bar{t}H$ (orange) analysis categories. All other couplings are set to their SM values.	119
3051		
3052		
3053		
3054	6.16 Negative log-likelihood contours at 68 and 95% CL in the $(\kappa_\lambda, \kappa_t)$ plane evaluated with an Asimov data set assuming the SM hypothesis (left) and the observed data (right). The contours obtained using the HH analysis categories only are shown in blue, and in orange when combined with the $t\bar{t}H$ categories. The best fit value for the HH categories only ($\kappa_\lambda = 0.6, \kappa_t = 1.2$) is indicated by a blue circle, for the HH + $t\bar{t}H$ categories ($\kappa_\lambda = 1.4, \kappa_\lambda = 1.3$) by an orange diamond, and the SM prediction ($\kappa_\lambda = 1.0, \kappa_t = 1.0$) by a black star. The grey band shows the region where LO parametrization of κ_t used in CMS analysis is not reliable.	120
3055		
3056		
3057		
3058		
3059		
3060		
3061		
3062	6.17 The stat. only expected 95% CL upper limits on the HH cross-section obtained for different values of κ_λ . Approximate ATLAS+CMS expected limits are obtained by scaling the ATLAS luminosity by a factor 2. The $\pm 1\sigma$ and $\pm 2\sigma$ bands are shown for the ATLAS expected limits.	121
3063		
3064		
3065		
3066	6.18 The BDT distributions in the $b\bar{b}\tau^+\tau^-$ analysis for the $\tau_{lep}\tau_{had}$ and $\tau_{had}\tau_{had}$ categories.	122
3067		
3068	6.19 Upper limits at 95% CL on the ggF SM HH production normalized to its SM expectation from the $b\bar{b}\tau^+\tau^-$, $b\bar{b}b\bar{b}$, $b\bar{b}\gamma\gamma$, $b\bar{b}W^+W^-$, $b\bar{b}\gamma\gamma$ and $W^+W^-W^+W^-$ decay channels, and their combination.	123
3069		

3070	6.20 Upper limits at 95% CL on the cross-section of the ggF SM HH production as a function of κ_λ	123
3071		
3072	7.1 Photon identification efficiency of cut-based Run-2 ID algorithm (blue) and the CNN ID algorithm (red) as function of photon energy and $ \eta $ from $HH \rightarrow b\bar{b}\gamma\gamma$ simulated events. Tight isolation is applied as baseline.	126
3073		
3074		
3075	7.2 $m_{\gamma\gamma}$ invariant mass distribution for two selected photons with cut-based algorithm (blue) and CNN algorithm (red).	127
3076		
3077	7.3 Photon identification efficiency of cut-based Run-2 ID algorithm (blue) and the CNN ID algorithm (red) as function of photon energy from $\gamma\gamma+\text{jets}$ simulated events. Tight isolation is applied as baseline.	127
3078		
3079		
3080	7.4 $m_{\gamma\gamma}$ invariant mass distribution of the continuum $\gamma\gamma+\text{jets}$ for two selected photons with cut-based algorithm (blue) and CNN algorithm (red).	128
3081		
3082	7.5 The 95% CL expected upper limits on the HH production cross-section as function of κ_λ . The ± 1 and ± 2 bands are shown for the Run-2 results. The expected allowed κ_λ intervals are also reported.	129
3083		
3084		
3085	7.6 Negative natural logarithm of the ratio of the maximum likelihood to the maximum likelihood for $\kappa_\lambda = 1$. The dashed lines at 0.5 and 2.0 indicate the values corresponding to a 1σ and 2σ confidence intervals, respectively (assuming an asymptotic χ^2 distribution of the test statistic).	129
3086		
3087		
3088	7.7 BDT response for signal and background. The vertical line denotes the optimal cut on the BDT response that maximizes the statistical-only significance,	130
3089		
3090		
3091	7.8 Negative natural logarithm of the ratio of the maximum likelihood for κ_λ to the maximum likelihood for $\kappa_\lambda = 1$ for the fit with only statistical uncertainties.	131
3092		
3093	7.9 Negative natural logarithm of the ratio of the maximum likelihood for κ_λ to the maximum likelihood for $\kappa_\lambda = 1$ for the fit with only statistical uncertainties. The black circles show the results for the combination, while the coloured markers show the values coming from the individual channels.	132
3094		
3095		
3096		
3097	7.10 Expected significance of observing Higgs boson pair production as a function of κ_λ without any systematic uncertainties. The two horizontal dashed lines show the 3σ and 5σ thresholds.	132
3098		
3099		
3100	7.11 Negative natural logarithm of the ratio of the maximum likelihood for κ_λ to the maximum likelihood for $\kappa_\lambda = 1$. The dashed lines at 0.5 and 2.0 indicate the values corresponding to a 1σ and 2σ confidence intervals, respectively (assuming an asymptotic χ^2 distribution of the test statistic).	134
3101		
3102		
3103		
3104	7.12 Expected statistic-only significance of observing Higgs boson pair production as a function of κ_λ from the European Strategy $b\bar{b}\gamma\gamma$ (blue) and the extrapolation of the full Run-2 $b\bar{b}\gamma\gamma$ to HL-LHC (black). The European Strategy numbers are taken from Figure 7.10.	135
3105		
3106		
3107	A.1 The energy profile in ϕ and η directions (a, b) and the corresponding R_ϕ and R_η variables, for unconverted photons with $0.00 < \eta < 0.20$. The black points correspond to the pseudo data, red points to non-reweighted simulation and blue points to the reweighted simulation from $Z \rightarrow ll\gamma$ with derived electron reweighting.	150
3108		
3109		
3110		
3111	A.2 The energy profile in ϕ and η directions (a,b) and the corresponding R_ϕ and R_η variables, for unconverted photons with $1.80 < \eta < 2.00$. The black points correspond to the pseudo data, red points to non-reweighted simulation and blue points to the reweighted simulation from $Z \rightarrow ll\gamma$ with derived electron reweighting.	151
3112		
3113		
3114		

3115 A.3 The energy profile in ϕ and η directions (a,b) and the corresponding R_ϕ and R_η variables, 3116 for converted photons with $0.00 < \eta < 0.20$. The black points correspond to the pseudo 3117 data, red points to non-reweighted simulation and blue points to the reweighted simulation 3118 from $Z \rightarrow ll\gamma$ with derived electron reweighting.	152
3119 A.4 The energy profile in ϕ and η directions (a,b) and the corresponding R_ϕ and R_η variables, 3120 for converted photons with $1.80 < \eta < 2.00$. The black points correspond to the pseudo 3121 data, red points to non-reweighted simulation and blue points to the reweighted simulation 3122 from $Z \rightarrow ll\gamma$ with derived electron reweighting.	153
3123 A.5 The energy profile in ϕ and η directions (a,b) and the corresponding R_ϕ and R_η variables, 3124 for unconverted photons with $0.00 < \eta < 0.20$. The black points correspond to the 3125 pseudo data, red points to non-reweighted simulation and blue points to the reweighted 3126 simulation from $Z \rightarrow ll\gamma$ with derived photon reweighting.	154
3127 A.6 The energy profile in ϕ and η directions (a,b) and the corresponding R_ϕ and R_η variables, 3128 for unconverted photons with $1.80 < \eta < 2.00$. The black points correspond to the 3129 pseudo data, red points to non-reweighted simulation and blue points to the reweighted 3130 simulation from $Z \rightarrow ll\gamma$ with derived photon reweighting.	155
3131 A.7 The energy profile in ϕ and η directions (a,b) and the corresponding R_ϕ and R_η variables, 3132 for converted photons with $0.00 < \eta < 0.20$. The black points correspond to the pseudo 3133 data, red points to non-reweighted simulation and blue points to the reweighted simulation 3134 from $Z \rightarrow ll\gamma$ with derived photon reweighting.	156
3135 A.8 The energy profile in ϕ and η directions (a,b) and the corresponding R_ϕ and R_η variables, 3136 for converted photons with $1.80 < \eta < 2.00$. The black points correspond to the pseudo 3137 data, red points to non-reweighted simulation and blue points to the reweighted simulation 3138 from $Z \rightarrow ll\gamma$ with derived photon reweighting.	157
3139 B.1 Background rejection of the CNN tight (right) and Loose (left) WPs. For unconverted 3140 (top) and converted (bottom) photons in bin of $ \eta $. The rejection is evaluated in the 3141 $Z+jets$ sample. The red line shows the rejection for CNN and blue line shows the cut- 3142 based rejection.	159
3143 B.2 A comparison between the CNN (red) and cut-based (blue) loose WPs for unconverted 3144 (right) converted (left) photons in bin of $ \eta $. The ratio plot shows the scale factors.	160
3145 B.3 A comparison between the tight CNN (red) and tight E_T -dependent cut-based (blue) 3146 WPs for unconverted (right) converted (left) photons in bin of $ \eta $. The ratio plot shows 3147 the scale factors.	160
3148 B.4 A comparison between the tight CNN (red) and tight E_T -independent cut-based (blue) 3149 WPs for unconverted (right) converted (left) photons in bin of $ \eta $. The ratio plot shows 3150 the scale factors.	161
3151 B.5 A comparison between the tight CNN (red) and tight E_T -independent cut-based as a 3152 function of photons E_T in different bin of $ \eta $ for unconverted (top) and converted (bottom) 3153 photons. The ratio plot shows the scale factors.	161
3154 B.6 Invariant mass ($m_{ll\gamma}$) distribution of events selected before tight ID in data (black dots). 3155 The solid black line represents the fitting result of the data distribution with the sum 3156 of the signal (red dashed line) and the background (blue dashed line). The photons are 3157 reconstructed as unconverted.	162

3158	B.7 Invariant mass ($m_{ll\gamma}$) distribution of events selected before tight ID in data (black dots). The solid black line represents the fitting result of the data distribution with the sum of the signal (red dashed line) and the background (blue dashed line). The photons are reconstructed as converted.	163
3159		
3160		
3161		
3162	B.8 Invariant mass ($m_{ll\gamma}$) distribution of events selected after tight ID in data (black dots). The solid black line represents the fitting result of the data distribution with the sum of the signal (red dashed line) and the background (blue dashed line). The photons are reconstructed as unconverted.	164
3163		
3164		
3165		
3166	B.9 Invariant mass ($m_{ll\gamma}$) distribution of events selected after tight ID in data (black dots). The solid black line represents the fitting result of the data distribution with the sum of the signal (red dashed line) and the background (blue dashed line). The photons are reconstructed as converted.	165
3167		
3168		
3169		
3170	C.1 The CNN efficiency as a function of photon energy for different η regions.	167
3171	C.2 Schema for MTCCN applied to a topological cluster.	168
3172	C.3 Image de-noising with Auto-Encoder	168
3173	C.4 The CNN classifier with PAED schema.	169
3174	D.1 $HH \rightarrow b\bar{b}\gamma\gamma$ in the transverse plane.	170
3175	D.2 Number of Initial State Radiation for qqF ZH, ggF ZH and ggF HH.	171
3176	D.3 $m_{b\bar{b}}$ resolution in GeV as a function of $\sigma_{b\bar{b}\gamma\gamma}$ for (a) 0-jet and (b) 1-jet. Blue shows the scan for SM case ($\kappa_\lambda = 1$), red is BSM case of $\kappa_\lambda = 2$ and green is for $\kappa_\lambda = -6$. Lines show $m_{b\bar{b}}$ resolution before the kinematic fit for each case (b -jet calibration).	172
3177		
3178		
3179	D.4 p_x and p_y distributions of HH after (red) and before (blue) applying the kinematic fit for 0-jet events. The $\sigma_{b\bar{b}\gamma\gamma} = 14$ GeV is used. The kinematic fit is applied on top of the b -jet calibration.	173
3180		
3181		
3182	D.5 p_x and p_y distributions of HH after (red) and before (blue) applying the kinematic fit for 1-jet events. The $\sigma_{b\bar{b}\gamma\gamma} = 14$ GeV is used. The kinematic fit is applied on top of the b -jet calibration.	174
3183		
3184		
3185	D.6 $m_{b\bar{b}}$ distributions after (red) and before (blue) applying the kinematic fit for (a) 0-jet and (b) 1-jet events. The kinematic fit is applied on top of the b -jet calibration.	174
3186		
3187	D.7 $m_{\gamma\gamma}$ distributions after (red) and before (blue) applying the kinematic fit for (a) 0-jet and (b) 1-jet events.	175
3188		
3189	D.8 p_x distribution for different ISR multiplicity before (solid line) and after (dashed line) including the soft radiation to compute the p_x for different processes.	176
3190		

List of Tables

3192	2.1 Generations of quarks and leptons with their masses and electrical charges	7
3193	2.2 The 95% CL observed and expected limits on the Higgs boson pair cross-section in pb and as a multiple of the SM production cross-section. The $\pm 1\sigma$ band around each 95% CL limit is also indicated.	23
3194		
3195		
3196	3.1 Granularity of the EM calorimeter ($\Delta\eta \times \Delta\phi$) [39].	38
3197	3.2 Discriminating variables used for electron and photon identification. The usage column indicates if the variables are used for the identification of electrons, photons, or both. [61].	48
3198		
3199	3.3 Definition of the electron isolation WPs.	50
3200		
3201	4.1 Definition of the photon isolation WPs.	56
3202	4.2 Discriminating variables used for Loose, Medium and Tight photon identification.	58
3203	4.3 Number of cells in 7×11 EM calorimeter windows.	70
3204	4.4 Image shape of 7×11 windows in each sampling in (η, ϕ)	71
3205	4.5 Fitted photon purity in % of all probes, before tight CNN WP. The uncertainties are only statistical.	75
3206	4.6 Fitted photon purity in % of all probes, after tight CNN WP. The uncertainties are only statistical.	76
3207		
3208	4.7 Scale factors for tight CNN WP efficiency measured with unconverted photons from $Z \rightarrow ll\gamma$ decays, in various bins of pseudorapidity and transverse energy. The uncertainty includes only statistical components.	76
3209		
3210		
3211	4.8 Scale factors for tight CNN WP efficiency measured with converted photons from $Z \rightarrow ll\gamma$ decays, in various bins of pseudorapidity and transverse energy. The uncertainty includes only statistical components.	79
3212		
3213		
3214	5.1 Muon selection requested for μ -in-jet correction	92
3215		
3216	6.1 Integrated luminosity recorded by ATLAS in Run 2 per year of data taking.	96
3217	6.2 Summary of single Higgs boson background samples, split by production modes, and continuum background samples. The generators used, and the PDF sets are also listed.	98
3218		
3219	6.3 Input variables used in the BDT.	100
3220	6.4 Definition of the analysis categories.	101
3221		
3222	6.5 Significance in each category for SM ggF HH and BSM $\kappa_\lambda = 10$ ggF signals.	106
3223		
3224	6.6 Spurious signal result for the exponential functional form for the various BDT categories. In each category, the spurious signal value, its ratio to the expected statistical error from data, and χ^2 the probability of the background-only fit.	108
	6.7 Relative p_T -Reco systematic uncertainty on the yield for given variation.	111

3225 6.8	Expected and observed numbers of events in the signal region ([120, 130] GeV) for the four BDT categories. The uncertainties on the continuum background are those arising from the fitting procedure. The uncertainties on the single Higgs boson and Higgs boson pair production are from MC statistical error.	114
3226 6.9	Summary of $HH \rightarrow b\bar{b}\gamma\gamma$ limits at 95% CL from the 139 fb^{-1} analysis.	114
3227 6.10	The expected limit on the HH cross-section at 95% CL for each analysis category for SM case. The combined limit is also shown.	116
3228 6.11	Expected upper limit at 95% CL on the cross-section of the HH pair production and allowed κ_λ interval for statistical only and statistical + systematic uncertainties	117
3229 6.12	Breakdown of the dominant systematic uncertainties. Only systematic uncertainties with an impact of at least 0.5% are shown. Uncertainties of Norm. + Shape type have effects on both the yield and the shape of the $m_{\gamma\gamma}$ distributions, the rest of uncertainties affect only the yields.	117
3230 6.13	Summary of $HH \rightarrow b\bar{b}\gamma\gamma$ results from 36 fb^{-1} and 139 fb^{-1} analyses.	117
3231 6.14	Summary of ATLAS and CMS $HH \rightarrow b\bar{b}\gamma\gamma$ results. The limit is presented as a multiple of the SM expected value.	119
3232 6.15	The 95% confidence limits of the HH production cross-section in $b\bar{b}b\bar{b}$ channel and κ_λ interval.	121
3233 6.16	The 95% confidence limits on the HH production cross-section and κ_λ for the $b\bar{b}\tau^+\tau^-$ analysis.	122
3234 6.17	The 95% CL observed and expected limits on the Higgs boson pair cross-section normalized to the SM production cross-section and the allowed κ_λ interval.	123
3235 7.1	Constraints on κ_λ from the likelihood ratio test performed on the Asimov dataset created from the background and the $\kappa_\lambda = 1$ signal, as shown in Figure 7.6.	128
3236 7.2	Constraints on κ_λ from the likelihood ratio test performed on the Asimov dataset created from the backgrounds and the SM HH signal.	131
3237 7.3	Significance of the individual $b\bar{b}b\bar{b}$, $b\bar{b}\tau^+\tau^-$ and $b\bar{b}\gamma\gamma$ channels as well as their combination.	132
3238 7.4	Signal strength measured in the individual channels and their combination using an Asimov dataset with SM HH signal injected.	133
3239 7.5	Constraints on κ_λ from the likelihood ratio test performed on the Asimov dataset created from the backgrounds and the SM HH signal. Results are presented as 1σ and 2σ CI on κ_λ	133
3240 7.6	Scale factors between the production cross-sections at 13 TeV (LHC Run 2) and 14 TeV (HL-LHC).	133
3241 7.7	Constraints on κ_λ from the likelihood ratio test performed on the Asimov dataset created from the background and the $\kappa_\lambda = 1$ signal, as shown in Figure 7.11.	134
3242 7.8	The Statistical-only significance of the individual categories as well as their combination.	134

³²⁶¹ Acknowledgements

³²⁶² To Be Added

