

中国科学技术大学

博士学位论文



基于 **ATLAS** 探测器上 **ZZ** 玻色子到全轻子
通道的衰变事例对标准模型 **ZZ** 过程和寻找
重共振态衰变到双 **Z** 玻色子过程的研究

作者姓名： 祝鹤龄

学科专业： 粒子与原子核物理

导师姓名： 赵政国 马宏

完成时间： 二〇二〇年八月七日

University of Science and Technology of China
A dissertation for doctor's degree



**Studies of Standard Model ZZ
Production and Search for heavy ZZ
resonances in Purely Leptonic
Decay with ATLAS Detector**

Author: Heling Zhu

Speciality: Particle and Nuclear Physics

Supervisors: Zhengguo Zhao, Hong Ma

Finished time: August 7, 2020

摘要

论文介绍了本人在粒子物理领域基于大型强子对撞机 (LHC) 上 ATLAS 实验做的研究工作。大型强子对撞机是当今世界上最大的、能量最高的对撞机，是建立在理论和实验之间的重要桥梁。而 ATLAS 实验是 LHC 上的一个通用例子探测器实验，同时也是体积最大的探测器。基于 ATLAS 实验在 LHC 上收集到的亮度为 139 fb^{-1} 能量为 13 TeV 的质子-质子对撞数据，本文重点介绍了两个 Z 玻色子衰变到四轻子末态过程的一系列研究。包括，标准模型 (SM) 下 ZZ 到四轻子过程截面的测量、矢量玻色子散射 (VBS) 过程在 ZZ 到四轻子末态的观测，和寻找重共振态衰变到 ZZ 到四轻子末态的过程。

ZZ 到四轻子过程截面的测量结果为 $\sigma_{ZZjj}^{tot} = 1.27 \pm 0.12(\text{stat}) \pm 0.02(\text{theo}) \pm 0.07(\text{exp}) \pm 0.01(\text{bkg}) \pm 0.03(\text{lumi})$ ，总体相对误差为 11%。在误差范围内，该结果和标准模型预言值 $1.14 \pm 0.04(\text{stat}) \pm 0.20(\text{theo})$ 相吻合。同时，在两个 Z 玻色子伴随着两个喷注 (jets) 末态的电弱相互作用过程的寻找中，我们观测到偏离本底假说超过 5 倍标准差 (5.5σ) 的明显偏差。在此基础上，本文也介绍了对于下一代高亮度大型强子对撞机 (HL-LHC) 在两个 Z 玻色子伴随着两个喷注 (jets) 末态的电弱相互作用过程的模拟预言。

另一方面，本文介绍了在一对 Z 玻色子衰变至四轻子末态过程中寻找重共振态的实验。根据不同的信号模型，寻找的粒子质量区间设置在 200 GeV 到 2000 GeV 之间。基于该测量结果，没有证据可以证明重共振态的存在。因此，研究给出了基于不同信号模型的截面上限，包括在不同衰变宽度假说下自旋为 0 的共振态，以及基于 Randall–Sundrum 模型的自旋为 2 的引力子 (graviton)。在该分析中，我们认为，信号主要可通过 gluon-gluon Fusion (ggF) 和 Vector Boson Fusion (VBF) 过程产生。在自旋为 0 的窄衰变宽度模型下，我们对 ggF 和 VBF 两个过程都进行了研究。而对于大宽度模型，由于在质量很高的区间分辨率很差以及 VBF 过程的统计量太小等客观原因，只对 ggF 过程进行了研究。对于自旋为 2 的模型，实验给出了 Randall–Sundrum 模型的引力子 (graviton) 的理论质量下限，为 1500 GeV 。

28

ABSTRACT

29 This dissertation presents my research in the field of Particle Physics with the ATLAS
30 experiment at the Large Hadron Collider (LHC). The LHC is the world's largest and most
31 powerful collider, and it was built as a bridge between the theories and the experiment.
32 The ATLAS experiment is a general-purpose particle detector experiment with the largest
33 volume at the LHC. This dissertation focus on the studies with two Z bosons production
34 decaying into $\ell\ell\ell\ell$ final state, where ℓ stands for electron or muon, using 139 fb^{-1} of
35 13 TeV proton-proton (pp) collision data collected by ATLAS experiment at the LHC.
36 The ZZ production in $\ell\ell\ell\ell$ channel provides a most clean and sensitive tool to test
37 the Standard Model (SM) at the energy frontier and to study the *Higgs* physics. Studies
38 including the measurement on SM $ZZjj$ production cross section, the observation of
39 Vector Boson Scattering (VBS) process as well as the searches of heavy resonances in
40 ZZ production decaying into $\ell\ell\ell\ell$ final state are reported in this dissertation.

41 The fiducial cross section for SM $ZZjj$ production is measured to be $\sigma_{ZZjj}^{tot} =$
42 $1.27 \pm 0.12(\text{stat}) \pm 0.02(\text{theo}) \pm 0.07(\text{exp}) \pm 0.01(\text{bkg}) \pm 0.03(\text{lumi})$ with a total rela-
43 tive uncertainty of 11% for the $\ell\ell\ell\ell$ final state, and found to be compatible with the
44 SM prediction of $1.14 \pm 0.04(\text{stat}) \pm 0.20(\text{theo})$. The electroweak production of two jets
45 in association with a Z -boson pair (EW- $ZZjj$) is observed with a significant deviation
46 from the background-only hypothesis corresponding to a statistical significance of 5.5σ .
47 Following with the observation, the prospect study for the EW- $ZZjj$ production at the
48 High luminosity LHC (HL-LHC) using 3000 fb^{-1} simulated pp collision data at a centre-
49 of-mass energy of 14 TeV is presented, with a expected significance of around 7σ .

50 A search for heavy resonances decaying into a pair of Z bosons to $\ell\ell\ell\ell$ final state is
51 also conducted in this dissertation. Different mass ranges for the hypothetical resonances
52 are considered, depending on the signal models and spanning between 200 GeV and
53 2000 GeV. Data is found to agree with a background-only hypothesis, thus, the results
54 are interpreted as upper limits on production cross section for sevaral different models,
55 including heavy Higgs like (spin-0) narrow-width approximation (NWA) and large-width
56 approximation (LWA), as well as the Randall–Sundrum model with a graviton excitation
57 spin-2 resonance (RSG). The signal is assumed to generate dominatly via gluon-gluon
58 Fusion (ggF) production mode and Vector Boson Fusion (VBF) production mode. Both
59 ggF and VBF channels are studied in NWA, while for LWA, only ggF channel is studied
60 due to worse resolution in higher mass region and the lack of statistic for VBF process. In

Abstract

- 61 addition, mass of RS Graviton is constrained, $m(G_{KK}) < 1500 \text{ GeV}$ is excluded at 95%
- 62 CL by $ZZ \rightarrow \ell\ell\ell\ell$ analysis.

63

64

Acknowledgments

65 First of all, I would like to express my great gratitude to my supervisors Prof. Zhao
66 Zhengguo and Dr. Ma Hong, for their guidance and patience during my Ph.D years. It's
67 Zhengguo, who inspire me with his deep physics insight when I was an undergraduate
68 student and lead me enter the field of Particle Physics. I will never forget how I was
69 attracted by his broad knowledge and the amazing picture of particles he showed me,
70 which became the reason I choose the Particle Physics as my major. It's always relaxed
71 and benifit greatly when chating with him, which broads my version, makes me to be
72 more confident and helps me step out of so many difficulties.

73 Thanks Hong, for giving me the opportunity to study in Brookhaven National Lab
74 (BNL) that I can work with so many senior and brilliant physicist, and teaching me a lot
75 in the details of physics. As the chair of physics department at BNL, can you imagine
76 that he managed to take time sitting with me every week, teaching me the details and
77 techniques of my analysis as well as helping me to parepre my talks at conference.

78 Thank you both for leading me to the field of physics, showing me how beautiful the
79 science and the world are. And thank you for providing me so many oppotunities and
80 tremendous supports to work at physics frontier and work with people all over the world.
81 It's my greatest honor to be your student. Your strong personalitis will difinately influence
82 my future life and career.

83 I would like to give my large gratitude to Prof. Zhou Bing. It's Bing who introduced
84 me to ATLAS experiment when I was a junior. As a professor and group leader of ATLAS
85 group in University of Michigan with busy schedule, Bing still took time to teach me in
86 physics start from simple formulas and help me to prepare my first academic presentation
87 patiently when I was a undergraduate student. Also it's my great fortunate that I can have
88 opportunities to work with you and learn from you in so many analyses during these years.
89 Your kind and patience, your high standard influence me deeply in all these years.

90 Moreover, I really want to give my sincere gratitudo to Dr. Xu Lailin. Thank you,
91 Lailin, for all your helps during the passing five years. Thanks for teaching me in all
92 the analysis details, coding techniques, presentation skills hand by hand. You are really
93 a very good and patient teacher and gives me as many knowledges as you accompany.
94 Your broad knowledge, your perseverance in science and your very hard working indeed
95 affect me a lot.

96 In the meantime, I want to give my special thanks to Dr. Li Bing, who helped me

Acknowledgments

97 a lot in several different analyses (low-mass 4μ resonance search, VBSZZ analysis, Z'
98 search), he never hesitate to give his hand to me when I faced difficulties.

99 I would like to express my gratitude to many colleagues in both USTC and BNL team.
100 Thanks Prof. Sun Yongjie, who is the supervisor of my undergraduate thesis, helped me
101 start my first detector project on MRPC. Also thanks Dr. Liu Zhen who teached me in
102 details in this project, and helped me a lot for my life at BNL too. Thanks Prof. Peng Haip-
103 ing, Prof. Zhu Yingchun and Dr. Hu Qipeng for helping me all the details and techniques
104 in HWW analysis when I was a beginner. Thanks Dr. Dai Tiesheng, I have learnt quite
105 a lot in the project of Monitored Drift Tubes (MDT) when working with you at CERN
106 and also thank you for all the helps in regular life since that was my first time to Europe.
107 Thanks Dr. Chen Hucheng, who is the supervisor of my ATLAS qualification task on LAr
108 Trigger Digitizer Boards (LTDB) and the person lead me into this interesting electronic
109 project. And my appreciate to Dr. Xu Hao, Dr. Chen Kai and Dr. Liu Hongbin that
110 teached me in patience for this project as I was really a freshman on electronics. Thanks
111 Prof. Yuji Enari and Dr. Georges Aad for the helps in LAr software tasks when I moved
112 from BNL to CERN. In the meantime, I would really like to give my gratitude to Dr.
113 Michael Begal, Dr. Marc-Andre Pleier, Dr. Alessandro Tricoli, Dr. George Redlinger,
114 Dr. Viviana Cavalieri, Dr. Gaetano Barone and many senior physicist in BNL omaga
115 group. I have learnt a lot from every chat with you and every seminar you hosted. Also
116 I want to thanks Prof. Wu Yusheng, Prof. Qian Jianming, Prof. Liu Yanwen, Dr. Ju
117 Xiangyang for teaching me in lots of details in different physics analyses.

118 Moreover, I want to give my thanks to friends I met at USTC, BNL and CERN during
119 my Ph.D years. Thanks Dr. Yang Qian, Dr. Chu Xiaoxuan, Dr. Tu Biao, Dr. Gao
120 Shanshan, Dr. Liu Feng, Dr. Yuan Guangyuan, and many other friends I met at BNL.
121 Thanks Dr. Geng Cong, Dr. Li Peilian, Dr. Zhang Liqing, Dr. Guo Yicheng, Dr. Xu
122 Tairan, Dr. Wang Rongkun, Chen Jing, He Fudong, Guo Qianying, Chen Ye, Xu Hao,
123 Wang Tao, Xie Xiangyu, Liu Xiangtian and all friends I met at USTC and CERN. Thank
124 you all my friends! I will always remember all the happiness with you, and best wishes
125 to you in the future!

126 Last but not least, I would give my greatest gratitude to my families. Thanks my par-
127 ents to give me all your endless loves and supports in my whole life. My deep appreciate
128 to my three aunts for loving and caring me so much since I was born and help me to
129 accompany with my mother when I was thousands miles far away from home. And my
130 husband, Lin, thank you for your understanding and walk through all difficulties with me
131 in these years especially when we were in a foreign country.

132	Acknowledgments	IV
134	Chapter 1 Introduction	1
135	Chapter 2 Theory	3
136	2.1 The Standard Model of Particle Physics	3
137	2.1.1 Elementary particles in the Standard Model	3
138	2.1.2 Electroweak theory	5
139	2.1.3 Higgs mechanism and Electroweak symmetry breaking	6
140	2.2 Phenomenology of Large Hadron Collider	9
141	2.2.1 Physics at hadronic collision	9
142	2.2.2 Higgs physics at the LHC	12
143	2.2.3 Diboson physics	13
144	Chapter 3 The Large Hadron Collider and the ATLAS Detector	18
145	3.1 The Large Hadron Collider	18
146	3.1.1 Operation history and machine layout	18
147	3.1.2 Luminosity and pile-up	21
148	3.2 ATLAS detector	23
149	3.2.1 Detector overview	23
150	3.2.2 Physics requirement	24
151	3.2.3 Magnet system	25
152	3.2.4 Inner detector	26
153	3.2.5 Calorimeters	28
154	3.2.6 Muon spectrometer	31
155	3.2.7 Trigger system	32
156	Chapter 4 Simulation and Event Reconstruction for the ATLAS Experiment	36
157	4.1 Event simulation	36
158	4.2 Event reconstruction	39
160	4.2.1 Track	39
161	4.2.2 Primary vertex	41
162	4.2.3 Electron	42
163	4.2.4 Muon	46

164	4.2.5 Jets	52
165	4.2.6 Missing transverse energy	55
166	Chapter 5 Studies of SM ZZ production in $\ell\ell\ell\ell$ final state using pp	
167	collision data collected by ATLAS detector from 2015 to	
168	2018	58
169	5.1 Introduction	58
170	5.2 Data and MC samples	59
171	5.2.1 Data samples	59
172	5.2.2 MC simulations	60
173	5.3 Objects and Event selection	61
174	5.3.1 Objects selection	61
175	5.3.2 Event selection in reconstruction level	62
176	5.4 Background estimation	63
177	5.4.1 QCD backgrounds	64
178	5.4.2 Reducible backgrounds	64
179	5.5 Systematics	68
180	5.5.1 Theoretical systematics	69
181	5.5.2 Experimental systematics	71
182	5.6 Measurement of fiducial cross section	74
183	5.6.1 Calculation of C-factor	74
184	5.6.2 Result of fiducial cross section	75
185	5.7 Search for EW- $ZZjj$	75
186	5.7.1 MD discriminant	75
187	5.7.2 Profile likelihood ratio method	75
188	5.7.3 Fitting procedure	77
189	5.7.4 Result of statistical fit	77
190	5.8 Prospect study of EW- $ZZjj$ production in HL-LHC	78
191	5.8.1 The ATLAS detector at HL-LHC	80
192	5.8.2 Simulation	81
193	5.8.3 Event selection	81
194	5.8.4 Systematics	82
195	5.8.5 Results	84
196	5.9 Conclusion	87

197	Chapter 6	Search for heavy resonances decaying into a pair of Z bosons in $\ell\ell\ell\ell$ final state using pp collision data collected by ATLAS detector from 2015 to 2018	88
200	6.1	Introduction	88
201	6.2	Data and MC samples	89
202	6.2.1	Data samples	89
203	6.2.2	Background MC simulations	89
204	6.2.3	Signal MC simulations	90
205	6.3	Analysis selections	91
206	6.3.1	Objects selection	91
207	6.3.2	Event selection	92
208	6.3.3	Event categorizations	94
209	6.3.4	Signal acceptance	102
210	6.4	Background estimation	103
211	6.4.1	Irreducible backgrounds	104
212	6.4.2	Reducible backgrounds	108
213	6.5	Signal modelling	109
214	6.5.1	Modelling of narrow-width signal	110
215	6.5.2	Modelling of large-width signal	110
216	6.5.3	Modelling of interference	114
217	6.5.4	Modelling of spin-2 RS Graviton signal	118
218	6.6	Systematic uncertainties	119
219	6.6.1	Theoretical uncertainties	120
220	6.6.2	Experimental systematics	122
221	6.7	Results in $\ell\ell\ell\ell$ channel	122
222	6.7.1	Statistical procedure	123
223	6.7.2	Likelihood fit under background-only hypothesis for MVA-based analysis	124
224	6.7.3	Interpretations	124
225	6.8	Conclusion	134
226	Chapter 7	Summary	135
227	Bibliography		138

List of Figures

228	2.1	The elementary particles of the Standard Model.	3
230	2.2	The Feynman diagrams of interactions that form the basis of the standard model.	4
232	2.3	The Higgs potential $V(\phi)$ with $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right).	7
233	2.4	Schematic view of a hadron-hadron collision ^[4]	10
234	2.5	The PDF4LHC15 NLO PDFs at a low scale $\mu^2 = Q^2 = 4\text{GeV}^2$ (left) and at $\mu^2 = Q^2 = 100\text{GeV}^2$ (right) as a function of x.	10
236	2.6	Feynman diagrams of the Higgs production modes: (a) ggF; (b) VBF; (c) VH; (d) ttH.	12
238	2.7	The SM Higgs boson production cross sections for various production modes as a function of the centre-of-mass energy for pp collision.	13
240	2.8	Higgs boson production cross section for various production modes as a function of the Higgs mass for $\sqrt{s} = 13 \text{ TeV}$ (left) and 14 TeV (right) for pp collision.	13
243	2.9	SM Higgs decay channels.	14
244	2.10	Branching ratio of Higgs decays in various channels as a function of Higgs mass ^[10]	14
246	2.11	The tree-level Feynman diagrams of diboson production at the LHC.	15
247	2.12	Total production cross-section presented by ATLAS as a function of centre-of-mass energy \sqrt{s} from 7 to 13 TeV for some selected processes, the diboson measurements are scaled by a factor 0.1 to allow a presentation without overlaps.	15
251	2.13	Feynman diagrams of the vector boson scattering.	16
252	2.14	Feynman diagrams of vertexes involving QGC, TGC and Higgs.	16
253	3.1	Cumulative luminosity as a function of time in years from 2011 to 2018 for ATLAS detector.	19
255	3.2	CERN accelerator complex ^[16]	20
256	3.3	Integrated luminosity in ATLAS.	22
257	3.4	Number of Interactions per Crossing from 2015-2018 in ATLAS.	22
258	3.5	Coordinate system used by the ATLAS experiment at the LHC ^[18]	23
259	3.6	Cut-away view of the ATLAS detector ^[19]	24

List of Figures

260	3.7	Schematic diagram of the ATLAS magnet system.	26
261	3.8	Schematic diagram of the ATLAS inner detector ^[21] .	27
262	3.9	Schematic diagram of the ATLAS 4-Layer Pixel Detector.	27
263	3.10	SCT (a) barrel module and (b) end-cap ^[25] .	28
264	3.11	Cut-away view of the ATLAS calorimeters. The LAr calorimeters are seen 265 inside the scintillator-based tile hadronic calorimeters ^[27] .	29
266	3.12	Schematic diagram of a LAr EM calorimeter barrel module.	30
267	3.13	Schematic diagram of tile calorimeter module ^[29] .	30
268	3.14	Cut-away view of the ATLAS muon spectrometer ^[31] .	31
269	3.15	Schematic diagram of the ATLAS trigger and data acquisition system in run-2.	33
270	3.16	An examples of L1 calorimeter trigger tower for electron and photon trig- 271 gers ^[33] .	34
272	3.17	The HLT trigger algorithm sequence ^[33] .	35
273	4.1	The flow of the ATLAS simulation software.	36
274	4.2	Sketch of a hardon-hardon collision simulated by MC event generator. The 275 red blob in center denotes the hard collision, surrounded by tree-like struc- 276 tures representing Bremsstrahlung which is simulated by Parton Showers. 277 The purple blob stands for a secondary hard scattering event. The light green 278 blobs indicate the parton-to-hardon transitions and the dark green blobs rep- 279 resents hardon decays. The yellow lines are soft photon radiations.	38
280	4.3	The flowchart of the ATLAS data processing.	39
281	4.4	Schematic view of the ATLAS inner detector showing all the corresponding 282 components.	40
283	4.5	Schematic of the impact parameters of a track in the transverse plane (left) 284 and RZ-plane (right), as defined in the global ATLAS tracking frame ^[55] .	45
285	4.6	The efficiencies of three electron ID WPs from $Z \rightarrow ee$ (left) events and 286 hadrons misidentified as electrons estimated using di-jet MC samples (right).	46
287	4.7	Distributions of $E_T^{cone0.2}$ (left) and $p_T^{varcone0.2}$ (right) for electrons from $ZZ \rightarrow$ 288 ee events in data and MC simulation. The simulated events (full histograms) 289 are normalized to data.	47
290	4.8	Muon reconstruction efficiency as a function of η for: Medium (and Loose), 291 Tight and High- p_T working points.	50
292	4.9	Muon reconstruction efficiency for Low- p_T working point as a function of η .	51

293	4.10	Distributions of the calorimeter-based (right) and the track-based (left) relative isolation variables measured in $Z \rightarrow \mu\mu$ events	51
294			
295	4.11	A overview schematic of ATLAS jet reconstruction ^[61]	52
296			
297	4.12	A overview schematic of ATLAS jet calibration ^[62]	53
298			
299	4.13	MV2c10 BDT output for b- (solid blue), c- (dashed green) and light-flavour (dotted red) jets in $t\bar{t}$ events ^[63]	55
300			
301	4.14	Measured E_T^{miss} distribution for $Z \rightarrow \mu\mu$ events (left) and $W \rightarrow e\nu$ events (right)	57
302			
303	5.1	Typical diagrams for the production of $ZZjj$, including the relevant EW VBS diagrams (first row) and QCD diagrams (second row)	59
304			
305	5.2	Pre-fit m_{ZZ} and m_{jj} distribution in QCD-enriched CR	65
306			
307	5.3	Fake factor for $Z+jets$ background, constructed with additional electron, as a function of p_T (left) and η (right)	66
308			
309	5.4	Fake factor for $Z+jets$ background, constructed with additional muon, as a function of p_T (left) and η (right)	66
310			
311	5.5	Fake factor for $t\bar{t}$ background, constructed with additional electron, as a function of p_T (left) and η (right)	67
312			
313	5.6	Fake factor for $t\bar{t}$ background, constructed with additional muon, as a function of p_T (left) and η (right)	67
314			
315	5.7	m_{jj} distributions in fake control region in 4e (left), 2e2 μ (middle) and 4 μ (right) channel. The ratios between $Z+jets$ and $t\bar{t}$ ($Z+jets/t\bar{t}$) in each individual channel are: 2.59, 0.95, 0.74	68
316			
317	5.8	The theoretical uncertainties for $gg \rightarrow ZZ$ background in particle-level SR (left) and CR (right)	70
318			
319	5.9	MD distribution for QCD- $Z Z jj$ process in low and high pile-up events for SR (left) and CR (right)	73
320			
321	5.10	MD shape difference for QCD $q\bar{q} \rightarrow ZZ$ background between different Sherpa theoretical uncertainties and sample from MadGraph5_aMC@NLO on SR (left) and CR (right)	73
322			
323	5.11	Observed and expected multivariate discriminant distributions after the statistical fit in the $\ell\ell\ell\ell$ SR (left) and QCD CR (right). The error bands include the experimental and theoretical uncertainties, as well as the uncertainties in μ_{EW} and $\mu_{\text{QCD}}^{\ell\ell\ell\ell jj}$. The error bars on the data points show the statistical uncertainty on data	78
324			
325			
326			

327	5.12 Observed and expected m_{jj} distributions in SR (left) and QCD CR (right). The error bands include the expected experimental and theoretical uncertainties. The error bars on the data points show the statistical uncertainty. The contributions from the QCD and EW production of $ZZjj$ events are scaled by 0.96 and 1.35, respectively, corresponding to the observed normalization factors in the statistical fit. The last bin includes the overflow events.	79
333	5.13 Observed and expected m_{ZZ} spectrum in SR. The error bands include the expected experimental and theoretical uncertainties. The error bars on the data points show the statistical uncertainty. The contributions from the QCD and EW production of $ZZjj$ events are scaled by 0.96 and 1.35, respectively, corresponding to the observed normalization factors in the statistical fit. The last bin includes the overflow events.	79
339	5.14 Display of an event candidate of EW- $ZZjj$ production in $2e2\mu$ channel in last MD bin ($0.875 < \text{MD} < 1.0$). The invariant mass of the di-jet (four-lepton) system is 2228 (605) GeV.	80
342	5.15 Detector-level distributions of EW- and QCD- $ZZjj$ processes with selected events in defined phase space at 14 TeV of (a) m_{jj} , (b) m_{ZZ} , (c) $ \Delta\phi(ZZ) $, (d) ZZ centrality, normalized to 3000 fb^{-1}	83
345	5.16 Jet variations on m_{jj} distribution for EW- $ZZjj$ (left) and QCD- $ZZjj$ (right) processes with luminosity of 3000 fb^{-1} at 14 TeV. <i>Upgrade Performance Function</i> is used to extract the uncertainties with <i>baseline</i> setting.	84
348	5.17 The expected significance of EW- $ZZjj$ processes as a function of different m_{jj} cut with 3000 fb^{-1} , under conditions of different sizes of theoretical uncertainties on the QCD- $ZZjj$ background modelling. The statistical uncertainty is estimated from expected data yield at 14 TeV with 3000 fb^{-1} . Different uncertainties are summed up quadratically.	85
353	5.18 The projected differential cross-sections at 14 TeV for the EW- $ZZjj$ processes as a function of m_{jj} (left) and m_{ZZ} (right). The top panel shows measurement with statistical only case, where statistical uncertainty is estimated from expected data yield at 14 TeV with 3000 fb^{-1} . The bottom panel shows impact of different sizes of systematic uncertainties.	85
358	6.1 (a) VBF DNN architecture diagram. (b) ggF DNN architecture.	98

359	6.2	(a) $m_{4\ell}$ distribution of raw (unweighted) training events for VBF signal (blue) 360 and background (black); (b) $m_{4\ell}$ distribution of weighted VBF signal (blue) 361 and background (black) used at training time.	99
362	6.3	(a) $m_{4\ell}$ distribution of raw (unweighted) training events for ggF signal (blue) 363 and background (black); (b) $m_{4\ell}$ distribution of weighted ggF signal (blue) 364 and background (black) used at training time.	99
365	6.4	The output score of “ggF-classifier” (a) and “VBF-classifier” (b) with the 366 events passing the common event selections for the data, the SM backgrounds 367 and an example of a NWA signal with a mass of 600 GeV. For the “VBF- 368 classifier”, an additional requirement of at least two jets in the event is ap- 369 plied. The signals cross section are set to one hundred times of the observed 370 limit for the “ggF-classifier” and fifty times of the observed limit for the 371 “VBF -classifier”. The $Z Z$ backgrounds are scaled by the normalisation fac- 372 tors shown in Table 6.15. The lower panels show the ratio of data to predic- 373 tion. Only statistical and experimental systematic uncertainties are included. 374 101	
375	6.5	Significance improvements of the MVA-based over the cut-based categoriza- 376 tion of the VBF (ggF) category for VBF (ggF) signal samples from 300 to 377 2000 GeV for seven different cuts on the VBF (ggF) output score. The op- 378 timal cut of 0.8 (0.5) for VBF (ggF) score is chosen as the solid line, while 379 other alternative cuts are plotted with dashed lines. For VBF category, results 380 at 2000 GeV for cuts of 0.8 and 0.9 are missing due to a lack of background 381 events passing this tight selection.	101
382	6.6	Illustration of the MVA-based VBF and ggF event classification for events 383 with (a) $N_{\text{jets}} < 2$ and (b) $N_{\text{jets}} \geq 2$	103
384	6.7	NWA acceptance as a function of m_H for the MVA-based categorization for 385 the samples of (a) ggF production; (b) VBF production.	103
386	6.8	NWA acceptance as a function of m_H for the Cut-based categorization for the 387 samples of (a) ggF production mode; (b) VBF production mode.	104
388	6.9	Distributions of the $m_{4\ell}$ invariant mass fit projections of the $q\bar{q} \rightarrow ZZ$ back- 389 ground samples for the 4μ , $4e$ and $2\mu 2e$ final states in the ggF-CBA-enriched 390 category, and the 4ℓ inclusive VBF-CBA-enriched category. Cut-based cat- 391 egorization is used.	105

392	6.10	Distributions of the $m_{4\ell}$ invariant mass fit projections of the $gg \rightarrow ZZ$ background samples for the 4μ , $4e$ and $2\mu2e$ final states in the ggF-CBA-enriched category, and the 4ℓ inclusive VBF-CBA-enriched category. Cut-based categorization is used.	106
393			
394			
395			
396	6.11	Distributions of the $m_{4\ell}$ invariant mass fit projections of the $q\bar{q} \rightarrow ZZ$ (EW) background samples for the 4μ , $4e$ and $2\mu2e$ final states in the ggF-CBA-enriched category, and the 4ℓ inclusive VBF-CBA-enriched category. Cut-based categorization is used.	106
397			
398			
399			
400	6.12	Distributions of the $m_{4\ell}$ invariant mass fit projections of the $q\bar{q} \rightarrow ZZ$ background samples for the 4μ , $4e$ and $2\mu2e$ final states in the ggF-MVA-high category, the 4ℓ inclusive ggF-MVA-low category and VBF-MVA-enriched category. DNN-based categorization is used.	107
401			
402			
403			
404	6.13	Distributions of the $m_{4\ell}$ invariant mass fit projections of the $gg \rightarrow ZZ$ background samples for the 4μ , $4e$ and $2\mu2e$ final states in the ggF-MVA-high category, the 4ℓ inclusive ggF-MVA-low category and VBF-MVA-enriched category. DNN-based categorization is used.	107
405			
406			
407			
408	6.14	Distributions of the $m_{4\ell}$ invariant mass fit projections of the $q\bar{q} \rightarrow ZZ$ (EW) background samples for the 4μ , $4e$ and $2\mu2e$ final states in the ggF-MVA-high category, the 4ℓ inclusive ggF-MVA-low category and VBF-MVA-enriched category. DNN-based categorization is used.	108
409			
410			
411			
412	6.15	Distributions of the $m_{2\mu2e}$ and fit projection for signal samples between 200 to 3000 GeV for ggF production mode. Three MC campaigns, mc16a, mc16d and mc16e, are combined. The lower panel in each plot shows the pull distribution.	111
413			
414			
415			
416	6.16	Polynomial fits of the parameters μ , f_C , σ_G , σ_C , n_C and α_C for the signal $C + G$ model in the $2\mu2e$ channel as a function of m_H for the ggF production mode. The combination of the mc16a, mc16d and mc16e MC campaigns is used.	112
417			
418			
419	6.17	The difference between MC simulation and parameterization of 4μ (left), $4e$ (middle) and $2\mu2e$ (right) for the ggF production mode. The combination of the mc16a, mc16d and mc16e MC campaigns is used.	112
420			
421			
422	6.18	The final signal shapes for the ggF production mode, interpolated from the polynomial fit parameters.	112
423			
424	6.19	Comparison of the analytical shape to a truth $m_{4\ell}$ distribution of gg2VV MC samples for $m_H = 450$ GeV (top), 700 GeV (bottom) and width equal to 5% (left), 10% (middle), 15% (right) of the mass.	114
425			
426			

427	6.20	Comparison between the analytical shape convoluted with detector effects and the reconstructed $m_{2\mu 2e}$ MC distribution for mass points ranging from 400 to 1800 GeV and width equal to 15% of the mass.	115
428			
429			
430	6.21	The interference (H-B) model fitted to the truth $m_{4\ell}$ MC distribution after signal region selection for $2\mu 2e$ channel.	117
431			
432	6.22	The signal modelling for the large-width scenario at m_H of 400 GeV (top), 600 GeV (middle) and 800 GeV (bottom), as well as three different signal width: 5% (left), 10% (middle) and 15% (right). The contribution of the interference between heavy Higgs and SM Higgs (H-h) is shown together with the one between heavy Higgs and SM $gg \rightarrow ZZ$ background (H-b).	119
433			
434			
435			
436			
437	6.23	Fitted parameters of the graviton RBW, m_{RBW} and Γ_{RBW} , as a function of the graviton resonance mass, m_G	120
438			
439	6.24	Reconstructed $m_{4\ell}$ distributions in the $2\mu 2e$ channel with the final signal model superimposed for each RS graviton signal sample at masses of 600 GeV, 1600 GeV and 2000 GeV. The lower panel in each plot shows the pull distribution. The dashed green lines show the truth-level graviton signal models for reference.	120
440			
441			
442			
443			
444	6.25	Distribution of the four-lepton invariant mass $m_{4\ell}$ in the $\ell\ell\ell\ell$ search for (a), (b), (c) the ggF-MVA-high categories, (d) the ggF-MVA-low category and (e) the VBF-MVA-enriched category. The backgrounds are determined from a combined likelihood fit to the data under the background-only hypothesis. The simulated signal at 600 GeV is normalized to a cross section corresponding to one hundred times the observed upper limit given in section 6.7.3. The error bars on the data points indicate the statistical uncertainty, while the systematic uncertainty in the prediction is shown by the hatched band. The lower panels show the ratio of data to prediction.	125
445			
446			
447			
448			
449			
450			
451			
452			
453	6.26	Pulls and constraints of nuisance parameters after a background only fit to (a) Asimov data and (b) observed data in the $\ell\ell\ell\ell$ channel. The Asimov data is generated with background data only, and the observed data includes datasets from 2015 to 2018.	126
454			
455			
456			
457	6.27	Correlation of nuisance parameters after a background only fit to Asimov data in the $\ell\ell\ell\ell$ channel. The Asimov data is generated with background data only.	127
458			
459			

460	6.28	The expected and observed upper limits at 95% CL on $\sigma \times BR(H \rightarrow ZZ)$ using the MVA-based analysis for ggF (left) and VBF (right) production. The green and yellow bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties in the expected limits.	127
464	6.29	The expected and observed upper limits at 95% CL on $\sigma \times BR(H \rightarrow ZZ)$ using the cut-based analysis for ggF (left) and VBF (right) production. The green and yellow bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties in the expected limits.	128
468	6.30	The upper limits at 95% confidence level on $\sigma_{ggF} \times BR(H \rightarrow ZZ)$ as a function of the heavy resonance mass m_H for the ggF production mode with an intrinsic width of 1% (top left), 5% (top right), 10% (bottom left) and 15% (bottom right) for both the case where interference with Standard Model processes is considered. The green and yellow bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties in the expected limits.	129
474	6.31	The upper limits at 95% confidence level on $\sigma_{ggF} \times BR(G_{KK} \rightarrow ZZ)$ as a function of the heavy resonance mass $m(G_{KK})$ for the ggF production mode in RS Graviton model. The green and yellow bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties in the expected limits.	129
478	6.32	The expected (left) and observed (right) upper limits at 95% confidence level on $\sigma \times BR(S \rightarrow ZZ)$ for ggF production mode at different assumptions.	131
480	6.33	Comparisons of the expected upper limits at 95% CL on the cross section times branching ratio as a function of the heavy resonance mass m_H for the ggF production mode (left) and for the VBF production mode (right) in the case of the NWA. The expected limits from the previous publication are shown in the green dashed line and are projected to the 139 fb^{-1} as shown in the blue dashed line. In addition, the current results based on either cut-based categorisation or the multivariate-based categorisation are shown in red and black lines.	133
488	6.34	Display of one candidate event in 4μ final state with the mass of 1.35 TeV.	133

List of Tables

489	3.1	Summary of design parameters of the LHC for pp collisions.	19
491	4.1	Overview of the contributions to E_T^{miss} .	56
492	5.1	Overlap removal criteria between pre-selection objects for the $\ell\ell\ell\ell$ channel. The overlap removal follows the order shown in this table. Once an object has been marked as removed, it does not participate in the subsequent stages of the overlap removal procedure.	62
496	5.2	Summary of selection of physics objects and candidate events at detector level in the $\ell\ell\ell\ell jj$ signal region.	63
498	5.3	Observed data and expected signal and background yields in 139 fb^{-1} of lumi- nosity. Minor backgrounds are summed together as ‘Others’. Uncertainties on the predictions include both statistical and systematic components.	63
501	5.4	Observed data and expected signal and background yields in 139 fb^{-1} of lumi- nosity. Diboson background in table includes all the other diboson processes discussed in section 5.2.2, except those with four-lepton final state. Uncer- tainties include only MC statistic. No events from Z+jets and t̄t MC samples pass the selection, and are indicated as 0 in the table.	64
506	5.5	Fake background estimations in the SR. For the nominal value the 2D fake factor together with the Z+jets and t̄t combination applied. The other lines show the estimations with different uncertainty variations.	69
509	5.6	Summary of different variations for EW- and QCD-ZZjj theoretical uncer- tainties measurement.	70
511	5.7	Summary of theoretical uncertainties for the fiducial volume (SR) for both EW and QCD qq -initial processes.	70
513	5.8	Summary of theoretical uncertainties for the control region for EW and qqQCD processes.	70
515	5.9	Experimental uncertainties in $\ell\ell\ell\ell$ channel with the luminosity of 139 fb^{-1} . The ”Electron Exp”, ”Muon Exp” and ”Jet Exp” represent the quadrature of the respective sources from electron, muon, and jets.	72
518	5.10	C Factor of different ZZjj processes.	74
519	5.11	Measured and predicted fiducial cross-sections in $\ell\ell\ell\ell jj$ final-state. Uncer- tainties due to different sources are presented.	75

521	5.12	Input features for the training of MD.	76
522	5.13	Observed μ_{EW} and $\mu_{\text{QCD}}^{\ell\ell\ell\ell jj}$, as well as the observed and expected significance from the individual $\ell\ell\ell\ell$ channel. The full set of systematic uncertainties are included.	78
525	5.14	Comparison of event yields for signal ($N_{\text{EW-ZZjj}}$) and background ($N_{\text{QCD-ZZjj}}$) processes, and expected significance of EW-ZZjj processes, normalized to 3000 fb^{-1} data at 14 TeV, with baseline and alternative selections. Uncertainties in the table refer to expected data statistical uncertainty at 14 TeV with 3000 fb^{-1}	82
530	5.15	Summary of expected cross-section measured with different theoretical uncertainties. The statistical uncertainty is computed from expected data yield with 3000 fb^{-1} at 14 TeV. Different uncertainties are treated as uncorrelated and summed quadratically.	86
534	6.1	Summary of the recorded integrated luminosity (lumi), average and peak pile-up (PU) of data from 2015 to 2018.	89
536	6.2	Summary of the p_T (E_T) trigger thresholds (in GeV) employed for the muon (electron) trigger selection in the year of 2015, 2016, 2017, and 2018.	92
538	6.3	Summary of the object and event selection requirements.	95
539	6.4	Cutflow table for a narrow-width ggF signal sample at $m_H = 600 \text{ GeV}$. N_{event} denotes the number of events selected after each cut is applied, normalized to 139 fb^{-1} , according to the expected upper limit on the cross section. The acceptances (the proportion of events selected relative to the initial number of events) are also included.	96
544	6.5	Cutflow table for a narrow-width ggF signal sample at $m_H = 1000 \text{ GeV}$. N_{event} denotes the number of events selected after each cut is applied, normalized to 139 fb^{-1} , according to the expected upper limit on the cross section. The acceptances (the proportion of events selected relative to the initial number of events) are also included.	96
549	6.6	Cutflow table for a narrow-width VBF signal sample at $m_H = 600 \text{ GeV}$. N_{event} denotes the number of events selected after each cut is applied, normalized to 139 fb^{-1} , according to the expected upper limit on the cross section. The acceptances (the proportion of events selected relative to the initial number of events) are also included.	97

554	6.7	Cutflow table for a narrow-width VBF signal sample at $m_H = 1000$ GeV. N_{event} denotes the number of events selected after each cut is applied, nor- malized to 139 fb^{-1} , according to the expected upper limit on the cross sec- tion. The acceptances (the proportion of events selected relative to the initial number of events) are also included.	97
555			
556			
557			
558			
559	6.8	Input features used in the “VBF-classifier” for the $\ell\ell\ell\ell$ analysis. The RNN stands for the recurrent neural network and MLP for the multilayer percep- tron.	100
560			
561			
562	6.9	Input features used in the “ggF-classifier” for the $\ell\ell\ell\ell$ analysis. The RNN stands for the recurrent neural network and MLP for the multilayer percep- tron.	102
563			
564			
565	6.10	Summary of acceptance uncertainties of PDF, QCD scale and parton shower variations for ggF production. The MVA-based categorization is used.	121
566			
567	6.11	Summary of acceptance uncertainties of PDF, QCD scale and parton shower variations for VBF production. The MVA-based categorization is used.	121
568			
569	6.12	Summary of acceptance uncertainties of PDF, scale, and parton showering variations for QCD $q\bar{q} \rightarrow ZZ$ background. The MVA-based categorization is used.	122
570			
571			
572	6.13	A list of the experimental systematics considered in this analysis. The NPs have been separated by whether they only affect the normalisation (left col- umn) or if they affect the shape (right column) of the $m_{4\ell}$ distribution. They are further subdivided into the primary objects that they affect.	130
573			
574			
575			
576	6.14	Expected and observed numbers of events for $m_{4\ell} > 200$ GeV, together with their systematic uncertainties, for three MVA-based categories. The expected number of events, as well as their uncertainties, are obtained from a likelihood fit to the data under the background-only hypothesis. The uncertainties of the ZZ normalisation factors, presented in table 6.15, are also taken into account.	131
577			
578			
579			
580			
581			
582	6.15	ZZ normalization factor in each category, obtained from a likelihood fit to the data under the background-only hypothesis.	131
583			

584	6.16 Impact of the leading systematic uncertainties, the data statistic uncertainties,	
585	as well as the total uncertainties on the predicted signal event yield with the	
586	cross section times branching ratio being set to the expected upper limit, ex-	
587	pressed as a percentage of the signal yield for the ggF (left) and VBF (right)	
588	production modes at $m_H = 400$ and 1000 GeV.	132

Dedicated to my mother!

献给我的母亲！

593

Chapter 1 Introduction

594 The goal of particle physics is to understand how our universe works at its most funda-
595 mental level. It can be accomplished by pursuing the mysteries of the basic construction of
596 matter and energy, probing the interactions between elementary particles, and exploring
597 the basic nature of space and time itself.

598 **Elementary particles**

599 From around the 6th century BC, ancient Greek philosophers Leucippus, Democritus,
600 and Epicurus brought up a philosophical idea that everything is composed of “uncuttable”
601 elementary particles. In the 19th century, John Dalton, through his work on stoichiometry,
602 concluded that each element of nature was composed of a single, unique type of particle.
603 The particle was named as “atom” after the Greek word atomos, with the meaning of “in-
604 divisible”. However this Dalton’s atom theory was strongly challenged later. Near the
605 end of 19th century, physicists discovered that Dalton’s atoms are not, in fact, the funda-
606 mental particles of nature, but conglomerates of even smaller particles. Electron was dis-
607 covered by J. J. Thomson in 1897, and then its charge was carefully measured by Robert
608 Andrews Millikan and Harvey Fletcher in their “oil drop experiment” of 1909. In early
609 20th-century, Rutherford’s “gold foil experiment” showed that the atom is mainly empty
610 space, with almost all its mass concentrated in a tiny positive charge atomic nucleus. Then
611 the discoveries of anti-particles (the positron in 1932) and other particles (e.g. the muon
612 in 1936) show that more discoveries could be expected in future experiments.

613 Starting from 1950s, more accelerator facilities were put into service. Throughout
614 the 1950s and 1960s, a bewildering variety of particles were found in collisions of parti-
615 cles from increasingly high-energy beams. It was referred to informally as the “particle
616 zoo”. In 1964, the quark model was independently proposed by physicists Murray Gell-
617 Mann and George Zweig, and experimentally confirmed of their existence in mid-1970s.
618 In 1970s, the establishment of quantum chromodynamics (QCD) postulated the funda-
619 mental strong interaction, experienced by quarks and mediated by gluons.

620 The well-known Standard model (SM) was developed in stages throughout the lat-
621 ter half of the 20th century. Since then, confirmation of the top quark (1995), the tau
622 neutrino (2000), and the Higgs boson (2012) have added further credence to the Standard
623 Model. Now, the quarks, leptons and gauge bosons are the elementary constituents in a
624 framework of Standard Model of particle physics, which theoretically describes three of
625 the four known fundamental forces (the electromagnetic, weak, and strong interactions,

626 and not including the gravitational force) in the universe, as well as classifies all
627 known elementary particles.

628 **Higgs mechanics and electroweak symmetry breaking**

629 In 1961, Sheldon Glashow brought forward a unified electroweak theory to combine
630 the electromagnetic and weak interactions. In the standard model, if under the condition
631 that the energy high enough but electroweak symmetry is unbroken, all elementary parti-
632 cles are massless. But measurements show the fact that the W and Z bosons actually have
633 masses. Later on, the Higgs mechanics resolves this conundrum. The simplest description
634 of the mechanism adds a Higgs field that permeates all space to the Standard Model. Be-
635 low some extremely high energy, the field causes spontaneous symmetry breaking during
636 interactions. All massive particles in the Standard Model, including the W and Z bosons,
637 interact with Higgs boson to acquire their mass.

638 Over the past few decades, with the combination of electroweak theory, Higgs me-
639 chanics and strong interactions has been widely accepted. But the Higgs boson, which is
640 essential to explain the mechanics of the property "mass" for gauge bosons and fermions,
641 had been the final missing piece in the Standard Model of particle physics for the time be-
642 ing. The mass of Higgs boson was not specifically predicted, and it has been searched in
643 several large experiments (eg. LEP at CERN, Tevatron at Fermilab, and LHC at CERN)
644 with different energy. In 2012, the discovery of Higgs boson was finally announced by the
645 ATLAS and CMS collaborations at the Large Hadron Collider (LHC) with its mass round
646 125 GeV. Peter Higgs and Francois Englert were award the 2013's Nobel Prize in Physics
647 for their theoretical discovery of a mechanism that contributes to our understanding of the
648 origin of mass of subatomic particles.

649 **Contents of this thesis**

650 This thesis is organized as follows. Section 2 briefly introduces the Standard Model of
651 particle physics, the Higgs mechanism related to the thesis and the LHC phenomenology.
652 Section 3 gives an overview of the LHC and the ATLAS detector. The detector simulation
653 and the reconstruction of physics objects are described in section 4. And then section
654 5 focuses on the Standard model ZZ production cross section measurement in $ZZ \rightarrow$
655 $\ell\ell\ell\ell$ channel, and the observation of its electroweak component as well as its further
656 prospects in High luminosity LHC (HL-LHC). Section 6 present the search of possible
657 heavy resonances in $ZZ \rightarrow \ell\ell\ell\ell$ channel. In the end, section 7 gives the summary and
658 outlook for future physics in LHC.

659

Chapter 2 Theory

660 2.1 The Standard Model of Particle Physics

661 The standard model (SM) reflects our current understanding of elementary particles
 662 and several basic interactions. It is a gauge quantum field theory containing the inter-
 663 nal symmetries of the unitary product group $SU(3) \times SU(2) \times U(1)$, in which the color
 664 group $SU(3)$ presents the strong interaction, and $SU(2) \times U(1)$ describes the electroweak
 665 interactions. Over the past decades, the SM has been widely tested through various ex-
 666 periments with extremely high precision.

667 2.1.1 Elementary particles in the Standard Model

668 The elementary particles in SM can be classified into 3 class: *fermions*, *gauge bosons*
 and the *Higgs boson* as shown in Figure 2.1.

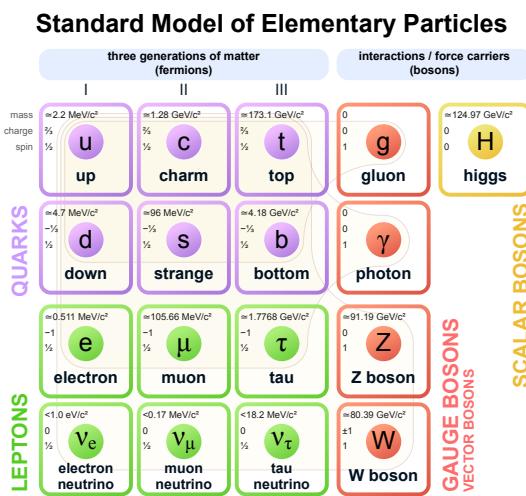


Fig. 2.1 The elementary particles of the Standard Model.

669

670 **Fermions** The Standard Model includes 12 elementary particles of spin- $\frac{1}{2}$ obeying the
 671 Fermi-Dirac statistics, known as fermions. They are classified into two types: *leptons* and
 672 *quarks* according to their interactions. The *leptons* include three generations: electron
 673 (e) and electron neutrino (ν_e); muon (μ) and muon neutrino (ν_μ); tau (τ) and tau neutrino
 674 (ν_τ). The e , μ and τ carry electric charge of -1 and three neutrinos are electrically neutral.
 675 All the leptons can participate in electroweak interactions. Also there are three generations
 676 of *quarks*: up (u) and down (d); charm (c) and strange (s); top (t) and bottom (b). The
 677 defining property of the quarks is that they carry color charge (while leptons don't), and
 678 hence interact via the strong interaction, letting them to be strongly bound from one to

another, forming color-neutral composite particles (known as hadrons) containing either a quark and an antiquark (mesons) or three quarks (baryons). In the meantime, u , c and t -quark carry electric charge of $2/3$, and d , s and b -quark carry electric charge of $-1/3$. Hence they interact via all three interactions described in SM. Each fermion also has a corresponding antiparticles.

Gauge bosons act as force carriers that mediate the strong, weak, and electromagnetic interactions in SM. They are spin-1 particles obeying the Bose-Einstein statistics. There are three types of gauge bosons:

- The eight massless *gluons* mediate the strong interactions between color charged particles (quarks).
- The massless *photons* mediate the electromagnetic force between electrically charged particles.
- The W^+ , W^- and Z bosons mediate the weak interactions between particles of different flavors (all quarks and leptons). All these three bosons are massive, the W^\pm carries an electric charge of $+1$ and -1 and couples to the electromagnetic interaction while Z boson is electrically neutral.

Figure 2.2 shows the Feynman diagrams of corresponding interactions in SM.

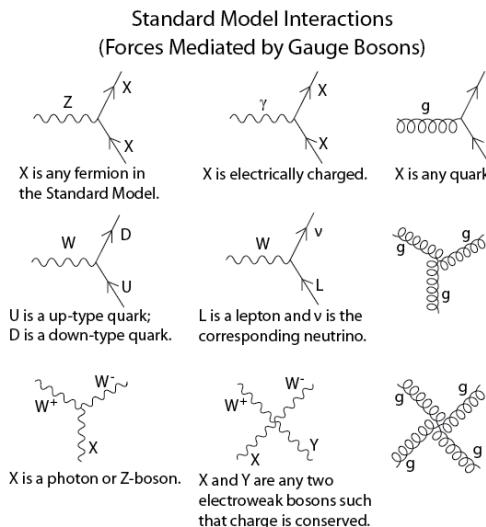


Fig. 2.2 The Feynman diagrams of interactions that form the basis of the standard model.

Higgs boson is a massive scalar elementary particle with spin-0. It plays a unique role in the SM by explaining the origin of masses of massive gauge bosons (W^\pm and Z) and fermions. And it is the last discovered particle in SM.

2.1.2 Electroweak theory

The electroweak interaction is the unified description of two of the four known fundamental interactions of nature: electromagnetism and the weak interaction. It is based on the gauge group of $SU(2)_L \times SU(1)_Y$, in which L is the left-handed fields and Y is the weak hypercharge^[1]. It follows the Lagrangian of

$$L_{EW} = L_{gauge} + L_{Higgs} + L_{fermion} + L_{Yukawa} \quad (2.1)$$

L_{gauge} is the **gauge term** part

$$L_{gauge} = -\frac{1}{4}W_{\mu\nu}^i W^{\mu\nu i} - \frac{1}{4}B_{\mu\nu} B^{\mu\nu} \quad (2.2)$$

where W_μ^i and B_μ present the $SU(2)_L$ and $SU(1)_Y$ gauge fields respectively, with the corresponding field strength tensors of

$$\begin{aligned} B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \\ W_{\mu\nu}^i &= \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g\epsilon_{ijk}W_\mu^j W_\nu^k \end{aligned} \quad (2.3)$$

In the equations above, g is the $SU(2)_L$ gauge coupling and ϵ_{ijk} is the totally antisymmetric tensor. The gauge Lagrangian has three and four-point self interactions of W^i , which result in triple and quartic gauge boson couplings.

The second term of the Lagrangian is the **scalar part**:

$$L_{Higgs} = (D^\mu \phi)^\dagger D_\mu \phi - V(\phi) \quad (2.4)$$

where $\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$ is a complex Higgs scalar, and $V(\phi)$ is the Higgs potential which is restricted into the form of

$$V(\phi) = +\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \quad (2.5)$$

due to the combination of $SU(2)_L \times SU(1)_Y$ invariance and renormalizability. In Eq. 2.5, μ is a mass-dependent parameter and λ is the quartic Higgs scalar coupling, which represents a quartic self-interaction between the scalar fields. When $\mu^2 < 0$, there will be spontaneous symmetry breaking (more details in section 2.1.3). To maintain vacuum stability, $\lambda > 0$ is required. And in Eq. 2.4, the gauge covariant derivative is defined as

$$D_\mu \phi = \left(\partial_\mu + ig \frac{\tau^i}{2} W_\mu^i + \frac{ig'}{2} B_\mu \right) \phi \quad (2.6)$$

in which τ^i represents the Pauli matrices, and g' is the $U(1)_Y$ gauge coupling. The square

719 of the covariant derivative results in three and four -point interactions between the gauge
720 and scalar fields.

721 The third term of the Lagrangian is the **fermion part**

$$L_{fermion} = \sum_{m=1}^F (\bar{q}_{mL}^0 \gamma_\mu D_\mu q_{mL}^0 + \bar{l}_{mL}^0 \gamma_\mu D_\mu l_{mL}^0 + \bar{u}_{mR}^0 \gamma_\mu D_\mu u_{mR}^0 + \bar{d}_{mR}^0 \gamma_\mu D_\mu d_{mR}^0 + \bar{e}_{mR}^0 \gamma_\mu D_\mu e_{mR}^0 + \bar{\nu}_{mR}^0 \gamma_\mu D_\mu \nu_{mR}^0) \quad (2.7)$$

722 In Eq. 2.7, m is the family index of fermions, F is the number of families. The subscripts
723 L(R) stand for the left (right) chiral projection $\psi_{L(R)} \equiv (1 \mp \gamma_5) \psi / 2$.

$$q_{mL}^0 = \begin{pmatrix} u_m^0 \\ d_m^0 \end{pmatrix}_L \quad l_{mL}^0 = \begin{pmatrix} v_m^0 \\ e_m^{-0} \end{pmatrix}_L \quad (2.8)$$

724 are the $SU(2)$ doublets of left-hand quarks and leptons, while u_{mR}^0 , d_{mR}^0 , e_{mR}^{-0} and ν_{mR}^0 are
725 the right-hand singlets.

726 The last term in Eq. 2.1 is **Yukawa term**

$$L_{Yukawa} = - \sum_{m,n=1}^F [\Gamma_{mn}^u \bar{q}_{mL}^0 \tilde{\phi} u_{nR}^0 + \Gamma_{mn}^d \bar{q}_{mL}^0 \phi d_{nR}^0 + \Gamma_{mn}^e \bar{l}_{mL}^0 \phi e_{nR}^0 + \Gamma_{mn}^\nu \bar{l}_{mL}^0 \tilde{\phi} \nu_{nR}^0] + h.c. \quad (2.9)$$

727 the matrices Γ_{mn} refer to the Yukawa couplings between single Higgs doublet (ϕ) and the
728 various flavors of quarks (m) and leptons (n).

729 2.1.3 Higgs mechanism and Electroweak symmetry breaking

730 As shown in previous subsection, the Lagrangian L_{gauge} does not involve any mass
731 term due to the requirement of gauge invariance. So all the W and B bosons should be
732 massless. But experimental observations show that the gauge bosons are massive. There-
733 fore, the gauge invariance must be broken spontaneously. The Higgs field is introduced
734 to break the $SU(2)_L \times U(1)_Y$ symmetry and gauge bosons and fermions can interact with
735 Higgs filed to acquire their masses. And this specific process is named *Higgs mechanism*
736 in SM.

737 The Higgs field ϕ is a doublet and can be written in a Hermitian basis as

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 - i\phi_2 \\ \phi_3 - i\phi_4 \end{pmatrix} \quad (2.10)$$

738 where $\phi_i = \phi_i^+$ stand for four Hermitian field. In this new basis, the Higgs potential in

⁷³⁹ Eq. 2.5 can be expressed as:

$$V(\phi) = \frac{1}{2}\mu^2 \left(\sum_{i=1}^4 \phi_i^2 \right) + \frac{1}{4}\lambda \left(\sum_{i=1}^4 \phi_i^2 \right)^2 \quad (2.11)$$

⁷⁴⁰ To simplify the situation, the axis in this four-dimensional space can be chosen to satisfied
⁷⁴¹ $\langle 0 | \phi_i | 0 \rangle = 0$ for $i = 1, 2, 4$, and $\langle 0 | \phi_3 | 0 \rangle = v$. Thus,

$$V(\phi) \rightarrow V(v) = \frac{1}{2}\mu^2 v^2 + \frac{1}{4}\lambda v^4 \quad (2.12)$$

⁷⁴² The minimization of this potential depends on the sign of μ^2 as shown in figure 2.3. When
⁷⁴³ $\mu^2 > 0$ the minimum occurs at $v = 0$, namely the vacuum is empty space and $SU(2)_L \times$
⁷⁴⁴ $U(1)_Y$ symmetry is unbroken. In the case of $\mu^2 < 0$, the $v = 0$ symmetric point is no
⁷⁴⁵ longer stable and the minimum occurs at nonzero value of $v = (-\mu^2/\lambda)^{1/2}$ which breaks
 the $SU(2)_L \times U(1)_Y$ symmetry. Thus, the classical vacuum ϕ_0 of Higgs doublet can be

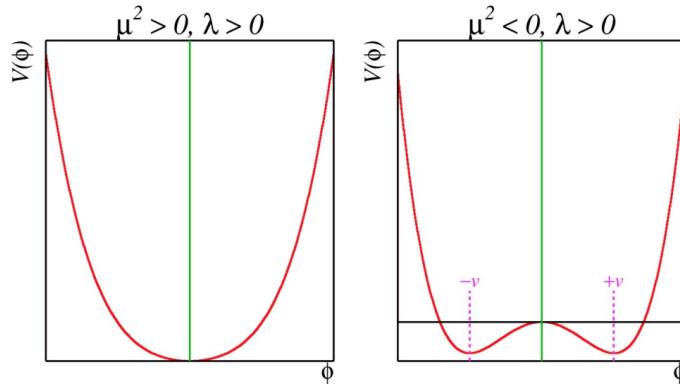


Fig. 2.3 The Higgs potential $V(\phi)$ with $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right).

⁷⁴⁶

⁷⁴⁷ expressed by

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (2.13)$$

⁷⁴⁸ And to quantize around the classical vacuum in a general form:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix} \quad (2.14)$$

⁷⁴⁹ Where H is a Hermitian field for physical Higgs scalar. In this gauge, the Lagrangian

750 L_{Higgs} in Eq. 2.4 takes a simple form

$$\begin{aligned} L_{Higgs} &= (D^\mu \phi)^\dagger D_\mu \phi - V(\phi) \\ &= M_W^2 W^{\mu+} W_\mu^- \left(1 + \frac{H}{v}\right)^2 + \frac{1}{2} M_Z^2 Z^\mu Z_\mu \left(1 + \frac{H}{v}\right)^2 \\ &\quad + \frac{1}{2} (\partial_\mu H)^2 - V(\phi) \end{aligned} \quad (2.15)$$

751 where the W and Z fields are

$$\begin{aligned} W^\pm &= \frac{1}{\sqrt{2}} (W^1 \mp iW^2) \\ Z &= -\sin\theta_W B + \cos\theta_W W^3 \end{aligned} \quad (2.16)$$

752 Therefore, in Eq. 2.15 spontaneous symmetry breaking brings out masses for the W and
753 Z gauge bosons

$$\begin{aligned} M_W &= \frac{gv}{2} \\ M_Z &= \sqrt{g^2 + g'^2} \frac{v}{2} = \frac{M_W}{\cos\theta_W} \end{aligned} \quad (2.17)$$

754 where θ_W is the weak angle defined as

$$\sin\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}} \quad \cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}} \quad \tan\theta_W = \frac{g'}{g} \quad (2.18)$$

755 Then another gauge boson photon remains massless with the field of

$$A = \cos\theta_W B + \sin\theta_W W^3 \quad (2.19)$$

756 After the symmetry breaking, the Higgs potential in unitary gauge can be written into

$$V(\phi) = -\frac{\mu^4}{4\lambda} - \mu^4 H^2 + \lambda v H^3 + \frac{\lambda}{4} H^4 \quad (2.20)$$

757 The first term in V is a constant, while the second term denotes a (tree-level) mass of
758 Higgs boson

$$M_H = \sqrt{-2\mu^2} = \sqrt{2\lambda}v \quad (2.21)$$

759 Due to the unknown of quartic Higgs coupling λ , the Higgs mass is not predicted. The
760 third and fourth terms in the Higgs potential V denote the induced cubic and quartic in-
761 teractions of the Higgs scalar.

762 Through the Higgs mechanism, fermions can also acquire their masses. In the unitary

763 gauge, Yukawa Lagrangian (L_{Yukawa}) can be written as a simple form of^[2]

$$L_{Yukawa} = - \left(1 + \frac{H}{v} \right) (m_d \bar{d}d + m_u \bar{u}u + m_l \bar{l}l) \quad (2.22)$$

764 in which $m_f = \frac{y_f v}{\sqrt{2}}$ for $f = d, u, l$.

765 2.2 Phenomenology of Large Hadron Collider

766 The Large Hadron Collider (LHC) was built as a bridge between the theories and the
 767 experiment. Physicists hope that the LHC can help to answer some of the fundamental
 768 open questions in physics, concerning the basic laws of interactions and forces among the
 769 elementary particles, the deep structure of space and time, and in particular the interrela-
 770 tion between quantum mechanics and general relativity. This section will talk about firstly
 771 the general introduction of Physics inside hadronic collision, then followed by two im-
 772 portant LHC phenomenologies of the Higgs physics and Diboson physics that are related
 773 closely to this thesis.

774 2.2.1 Physics at hadronic collision

775 Protons are not the elementary particle, which actually be composed of quarks and
 776 gluons. So in proton-proton (pp) collision at the LHC, it is not protons themselves interact
 777 but quarks and gluons. Scattering processes can then be further classified into either *hard*
 778 or *soft* processes according to the momentum transfer during the interaction^[3]. QCD, as
 779 an underlying theory for both two process, its approach and level of understandings in two
 780 cases are quite different. For hard process, eg. Higgs, vector bosons and jets production,
 781 the rates and event properties can be precisely predicted based on perturbation theory.
 782 However, for soft processes like total cross-section, the underlying events, the rates and
 783 properties are dominated by non-perturbative QCD effects that are less understood. For
 784 many hard processes, the hard interactions are accompanied by soft ones. A example of
 785 the hadronic collision is illustrated in figure 2.4. and the typical features are summarized
 786 as below:

- 787 • **Parton Distribution Function (PDF):** $f_i(x, Q^2)$ gives the probability of a par-
 788 ton with flavor i (quark or gluon), carrying a momentum fraction of x and at the
 789 energy of Q in a proton. Parton distribution function cannot be fully calculated
 790 by perturbative QCD because of the inherent non-perturbative nature of partons.
 791 There are many different sets of PDFs that are determined by a fit to data from
 792 experimental observables in various processes. As an example, figure 2.5 shows

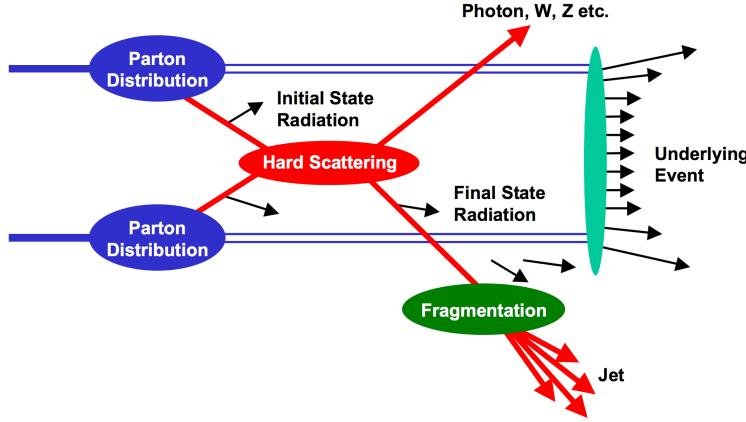


Fig. 2.4 Schematic view of a hadron-hadron collision [4].

793 the *PDF4LHC15 NLO PDFs*, which is based on the combination of the *CT14*,
 794 *MMHT14* and *NNPDF3.1 NNLO* PDF sets [5].

- 795 • **Fragmentation and hadronization:** The processes to produce final state particles
 796 (or jets) from the partons produced in hard scattering.
- 797 • **Initial/Final state radiation:** The incoming and outgoing partons that carry color
 798 charge can emit QCD radiation, which gives rise to additional jets. Also the charged
 799 incoming and outgoing particles can emit Quantum Electrodynamics (QED) radia-
 800 tions with photons.
- 801 • **Underlying events:** Products from soft processes (not come from the primary hard
 802 scattering) as the remnants of scattering interactions.

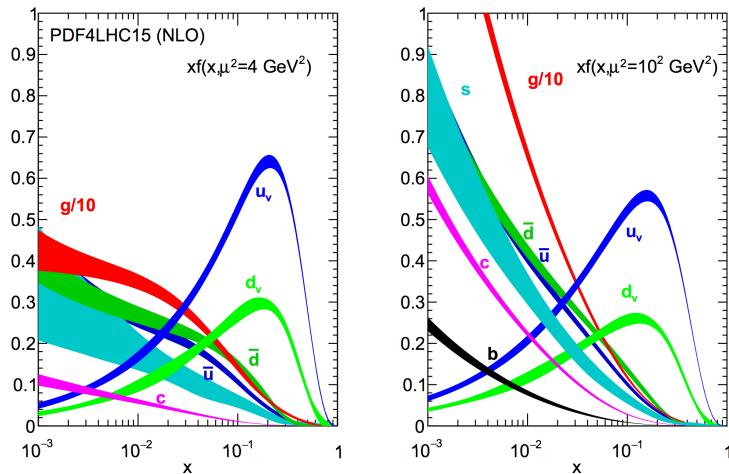


Fig. 2.5 The PDF4LHC15 NLO PDFs at a low scale $\mu^2 = Q^2 = 4 \text{ GeV}^2$ (left) and at $\mu^2 = Q^2 = 100 \text{ GeV}^2$ (right) as a function of x .

803 **Cross section of hard scattering**

804 According to *QCD factorization theorems* [6], the perturbative calculations can be ap-

plied to many important hard processes involving hadrons. The basic problem addressed by factorization theorems is how to calculate high energy cross sections. Consider the process of scattering between two hardons A and B to produce a final state X, the cross section σ can be obtained by summing over all the subprocess cross section $\hat{\sigma}$ ^[7]

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}(x_a) f_{b/B}(x_b) \hat{\sigma}_{ab \rightarrow X} \quad (2.23)$$

where $f_{q/A}(x_q)$ is the parton distribution functions of parton q . Taking into account the leading order correction:

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}(x_a Q^2) f_{b/B}(x_b Q^2) \hat{\sigma}_{ab \rightarrow X} \quad (2.24)$$

where Q^2 represents large momentum scale that characterizes the hard scattering. Later on, since the finite corrections were not universal and had to be calculated separately for each process, the perturbative $O(\alpha_S^n)$ corrections to the leading logarithm cross section in Eq. 2.24 need to be applied, one can get:

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}(x_a \mu_F^2) f_{b/B}(x_b \mu_F^2) \hat{\sigma}_{ab \rightarrow X}(\alpha_S, \mu_R, \mu_F) \quad (2.25)$$

in which μ_F is *factorization scale* which can represent the scale that separates the long- and short-distance physics, and μ_R is the *renormalization scale* for QCD running coupling. $\hat{\sigma}_{ab \rightarrow X}$ is the parton-level hard scattering cross section that can be calculated perturbatively in QCD with the form of

$$\hat{\sigma}_{ab \rightarrow X}(\alpha_S, \mu_R, \mu_F) = (\alpha_S)^n \left[\hat{\sigma}^{(0)} + (\alpha_S/2\pi) \hat{\sigma}^{(1)}(\mu_R, \mu_F) + (\alpha_S/2\pi)^2 \hat{\sigma}^{(2)}(\mu_R, \mu_F) + \dots \right] \quad (2.26)$$

where $\hat{\sigma}^{(0)}$ stands for the leading-order (LO) partonic cross section, while $\hat{\sigma}^{(1)}$ and $\hat{\sigma}^{(2)}$ are the next-to-leading-order (NLO) and next-to-next-to-leading-order (NNLO) cross section.

μ_R and μ_F depend on the order of truncation in Eq. 2.26. In principle, if cross section is calculated to all orders, it is invariant under changes in these parameters. The choices of μ_R and μ_F are arbitrary. To avoid unnaturally large logarithms reappearing in the perturbation series, it is sensible to choose μ_R and μ_F values of the order of the typical momentum scales of the hard scattering process and $\mu_R = \mu_F$ is also often assumed. Take Drell–Yan process as an example, the standard choice is $\mu_R = \mu_F = m_{ll}$, where m_{ll} is the invariant mass of di-lepton pair.

2.2.2 Higgs physics at the LHC

One important physics purpose of the LHC is searching for the Higgs boson, which was the last missing part in the SM. This section will talk about both the production and decay modes of the SM Higgs boson in proton-proton collision.

Higgs productions

The Higgs boson can be produced through several processes. There are 4 main production modes at the LHC: gluon-gluon Fusion (ggF), vector boson Fusion (VBF), associated production with vector-bosons (VH) (also called the Higgs Strahlung) and associated production with a pair of top/anti-top quarks (ttH)^[8]. Figure 2.6 shows the corresponding Feynman diagrams of each process (at LO). For pp collision, the cross section of produc-

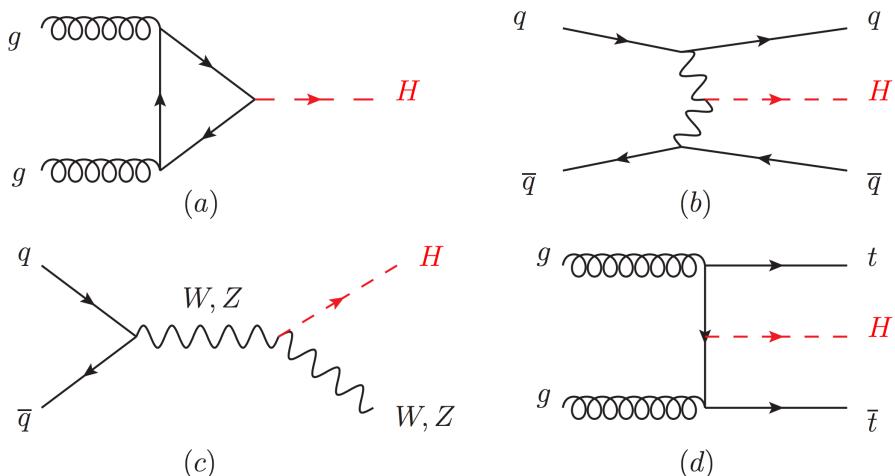


Fig. 2.6 Feynman diagrams of the Higgs production modes: (a) ggF ; (b) VBF ; (c) VH ; (d) ttH .

tions of Higgs boson is as a function of centre-of-mass energy \sqrt{s} . Figure 2.7 depicts the cross section of SM Higgs, whose mass is 125 GeV, for several different production modes when centre-of-mass energy varying from 6 to 15 TeV. Figure 2.8 shows the prospect of production cross section as a function of Higgs mass from 10 to 2000 GeV for pp collision at the centre-of-mass energy of 13 TeV and 14 TeV^[9].

Higgs decays

The Higgs boson can interact with gauge bosons and fermions through gauge coupling and the Yukawa coupling as introduced in section 2.1.3. Figure 2.9 depicts the Feynman diagrams of various possible Higgs decay channels. The branching ratio of Higgs boson decaying into different final states as a function of Higgs mass is shown in figure 2.10.

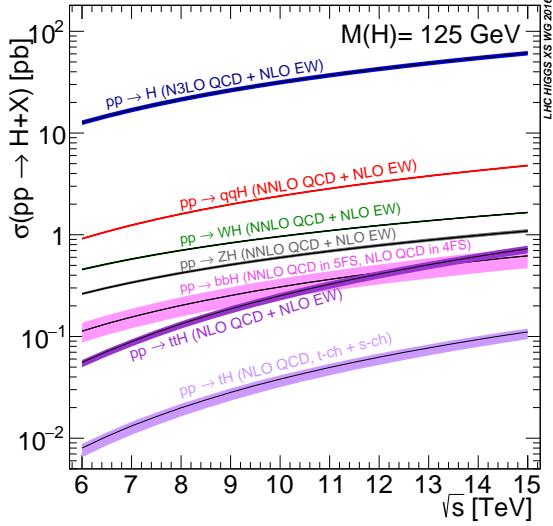


Fig. 2.7 The SM Higgs boson production cross sections for various production modes as a function of the centre-of-mass energy for pp collision.

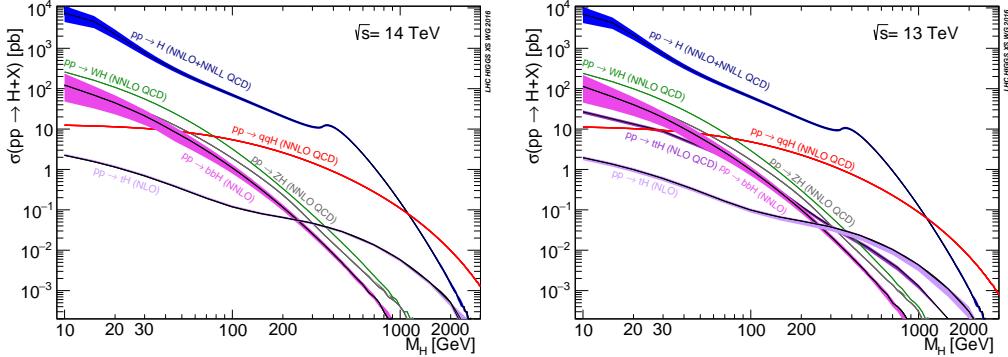


Fig. 2.8 Higgs boson production cross section for various production modes as a function of the Higgs mass for $\sqrt{s} = 13$ TeV (left) and 14 TeV (right) for pp collision.

2.2.3 Diboson physics

The study of diboson physics is another important test for SM of particle physics in electroweak sector, while the Vector Boson Scattering (VBS) is a key process for probing the mechanism of the electroweak symmetry breaking (EWSB). In the meantime, the non-resonant diboson productions are crucial backgrounds for Higgs studies at the LHC, which make the precise measurement of their cross section becomes very important.

Diboson productions

About 90% of diboson productions at hadron collider is from quark-antiquark annihilation, while others are contributed from gluon initiated process. Figure 2.11 shows the tree-level Feynman diagrams of diboson production. Then figure 2.12 illuminates the total production cross-section presented by ATLAS as a function of centre-of-mass energy \sqrt{s} from 7 to 13 TeV for several diboson processes comparing to some other major

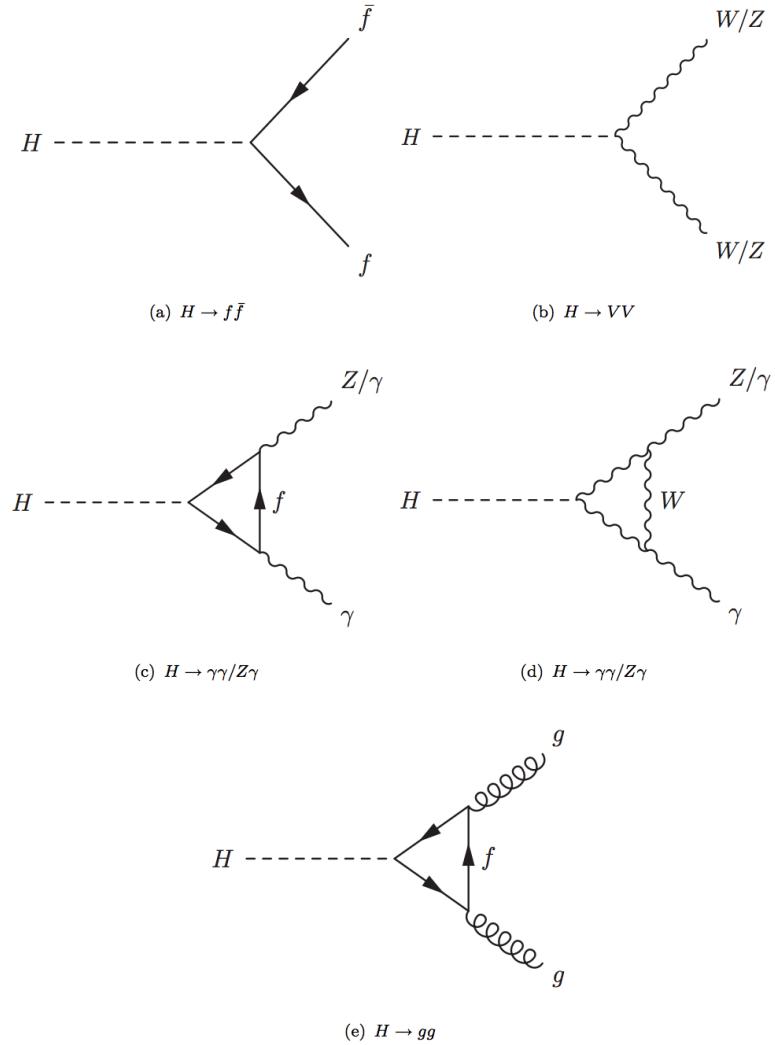
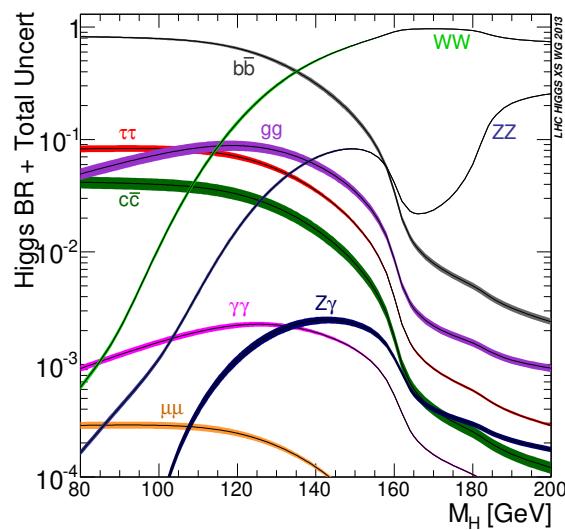


Fig. 2.9 SM Higgs decay channels.


 Fig. 2.10 Branching ratio of Higgs decays in various channels as a function of Higgs mass^[10].

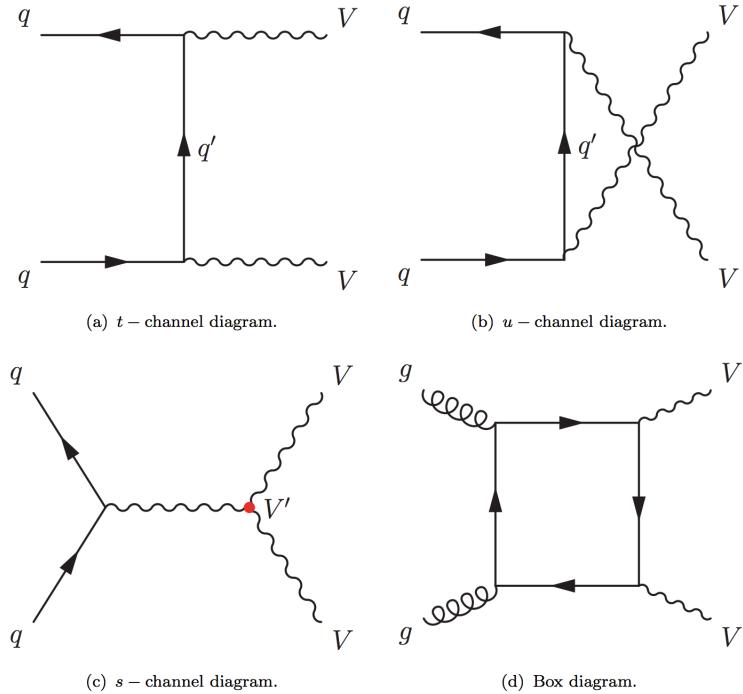


Fig. 2.11 The tree-level Feynman diagrams of diboson production at the LHC.

processes in hadron collision. The cross section for diboson processes are measured at next-to-next leading order (NNLO).

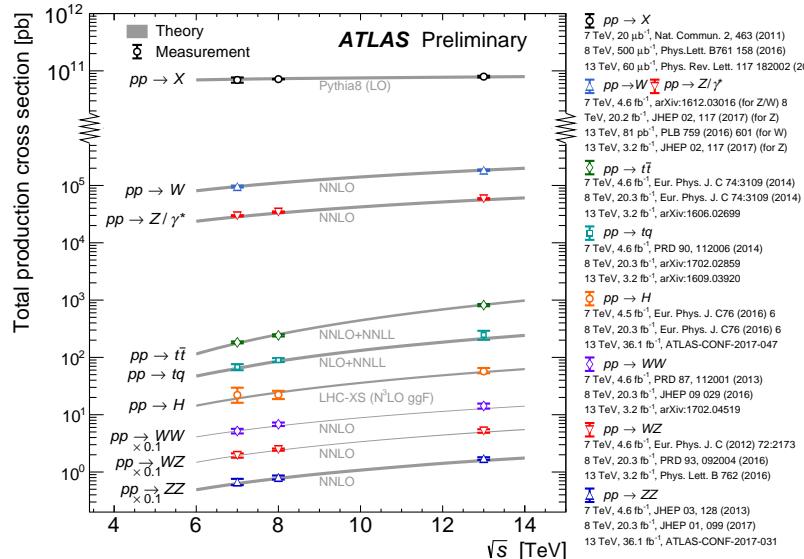


Fig. 2.12 Total production cross-section presented by ATLAS as a function of centre-of-mass energy \sqrt{s} from 7 to 13 TeV for some selected processes, the diboson measurements are scaled by a factor 0.1 to allow a presentation without overlaps.

861

862 Vector boson scattering

863 The $SU(2)_L \times U(1)_Y$ structure in SM predicts self-interactions between electroweak
864 gauge bosons. Those self-couplings can involve either three or four gauge bosons at a
865 single vertex, known as triple gauge coupling (TGC) or quartic gauge couplings (QGC),

866 respectively. Vector boson scattering or fusion (*VBS* or *VBF*) is carried out by four elec-
 867 troweak vector bosons, namely Z , W^\pm and photon (γ) as the Feynman diagrams shown
 868 in figure 2.13. And the vertexes include either those self-interactions or the interactions
 with the Higgs boson are described in figure 2.14.

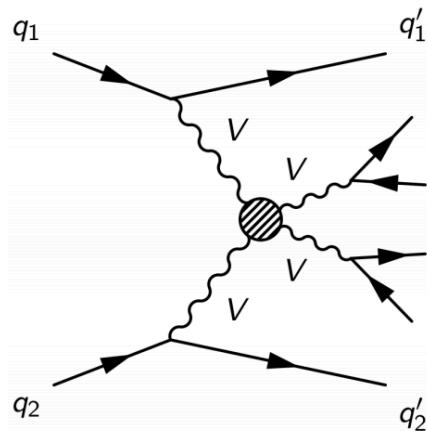


Fig. 2.13 Feynman diagrams of the vector boson scattering.

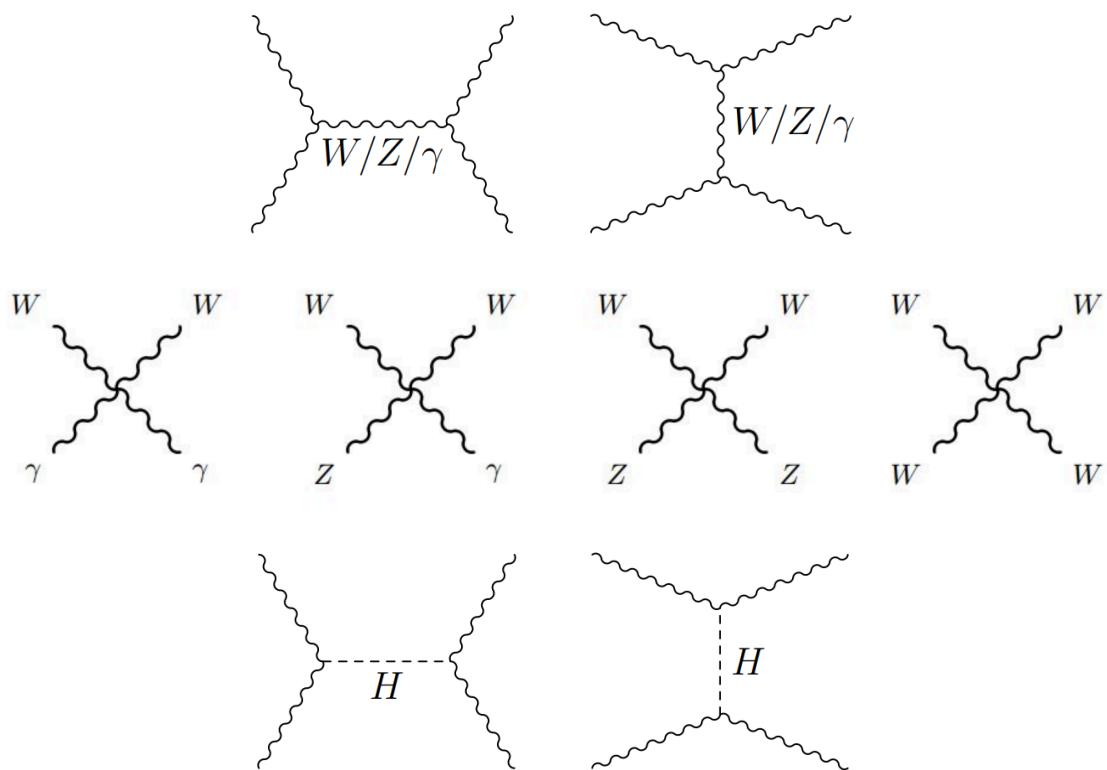


Fig. 2.14 Feynman diagrams of vertexes involving QGC, TGC and Higgs.

870 The amplitudes of leading-order (LO) VBS can be expressed as^[11]:

$$\begin{aligned} iM_{TGC}^{s-channel} &= -i \frac{g_1^2}{4m_W^4} [s(t-u) - 3m_W^2(t-u)] \\ iM_{TGC}^{t-channel} &= -i \frac{g_1^2}{4m_W^4} \left[(s-u)t - 3m_W^2(s-u) + \frac{8m_W^2}{s} u^2 \right] \end{aligned} \quad (2.27)$$

871

$$iM_{QGC} = i \frac{g_1^2}{4m_W^4} \left[s^2 + 4st + t^2 - 4m_W^2(s+t) - \frac{8m_W^2}{s} ut \right] \quad (2.28)$$

872

$$\begin{aligned} iM_{Higgs} &= -i \frac{C_v^2 g_1^2}{4m_W^2} \left[\frac{(s-2m_W^2)^2}{s-m_H^2} + \frac{(t-2m_W^2)^2}{t-m_H^2} \right] \\ &\simeq -i \frac{C_v^2 g_1^2}{4m_W^2} (s+t) \end{aligned} \quad (2.29)$$

873 Combining s- and t-channel of TGC in Eq. 2.27 and the QGC term in Eq. 2.28:

$$iM_{TGC} + iM_{QGC} = i \frac{g_1^2}{4m_W^2} (s+t) + O((s/m_W^2)^0) \quad (2.30)$$

874 In Eq. 2.30, the amplitude grows as a function of centre-of-mass energy (\sqrt{s}), which
 875 violates the unitarity in the TeV region. Considering the Higgs term in Eq. 2.29 can
 876 perfectly cancel out this growing, and the remaining term $O((s/m_W^2)^0)$ only depends on
 877 the total amplitude in SM.

878 In conclusion, the Higgs boson acts as "moderator" to unitarize high-energy longitudinal vector boson scattering as introducing the Higgs restores the unitarity of total
 879 amplitude in high energy region.
 880

881 **Chapter 3 The Large Hadron Collider and the ATLAS**
882 **Detector**

883 **3.1 The Large Hadron Collider**

884 Located near the French-Swiss border at the European Organization for Nuclear Re-
885 search (CERN), the Large Hadron Collider (LHC) is the world's largest and most pow-
886 erful particle collider. It's the proton-proton collider with the centre-of-mass energy up
887 to 14 TeV. The beams inside the LHC are made to collide at four locations around its
888 27-kilometer accelerator ring, corresponding to the positions of four particle detectors -
889 ATLAS, CMS, ALICE and LHCb. With its unprecedented energy, the LHC is designed to
890 observe physics that involve highly massive particles which have never been observable
891 in previous accelerators with lower energies.

892 **3.1.1 Operation history and machine layout**

893 **Operation history**

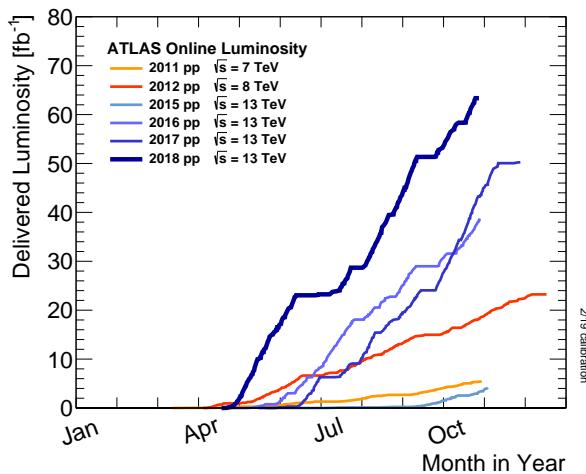
894 The LHC^[12-15] is a two-ring-superconducting-hadron accelerator and collider lies in
895 a tunnel 27 kilometres in circumference and as deep as 175 metres underground. It's
896 designed to provide proton-proton (pp) collisions at the centre-of-mass energy (\sqrt{s}) up
897 to 14 TeV with a unprecedented luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. In the meantime, it can also
898 collide heavy (Pb) ions with an energy of 2.8 TeV per nucleon and a peak luminosity of
899 $10^{27} \text{ cm}^{-2} \text{ s}^{-1}$. Table 3.1 shows the main design parameters of the LHC for proton-proton
900 collisions.

901 The LHC was built from 1998 to 2008. It started its first beam in September 2008,
902 but then was interrupted by a quench incident only after a few days running. Then it
903 resumed the operation in November 2009 with a low energy beams. From March 2010,
904 physics runs took place at the centre-of-mass energy of 7 TeV. Later on, this energy was
905 increased in 2012 to $\sqrt{s} = 8 \text{ TeV}$, with an integrated luminosity of 20.3 fb^{-1} , and this
906 period is called "run-1". After run-1, the LHC was shut down for two years for hardware
907 maintenance and upgrade, starting from February 2013.

908 The second operation period with higher centre-of-mass energy at 13 TeV started from
909 2015 called "run-2". And it continued to the end of 2018 with total integrated luminosity
910 reaching about 147 fb^{-1} for ATLAS. Figure 3.1 shows the cumulative luminosity as a
911 function of time in month delivered to ATLAS experiment during stable beams in years
912 from 2011 to 2018.

Table 3.1 Summary of design parameters of the LHC for pp collisions.

Circumference	26.7 km
Beam energy at collision	7 TeV
Beam energy at injection	0.45 TeV
Dipole field at 7 TeV	8.33 T
Luminosity	$10^{34} \text{ cm}^{-2} \text{ s}^{-1}$
Beam current	0.56 A
Protons per bunch	1.1×10^{11}
Number of bunches	2808
Nominal bunch spacing	24.95 ns
Normalized emittance	$3.75 \mu\text{m}$
Total crossing angle	$300 \mu\text{rad}$
Energy loss per turn	6.7 keV
Critical synchrotron energy	44.1 eV
Radiated power per beam	3.8 kW
Stored energy per beam	350 MJ
Stored energy in magnets	11 GJ
Operating temperature	1.9 K

**Fig. 3.1 Cumulative luminosity as a function of time in years from 2011 to 2018 for ATLAS detector.**

913

Machine layout

914

The layout of CERN accelerator complex is shown in figure 3.2. The protons are accelerated by a series of machines before being injected into the main cavity. At beginning, the 50 MeV protons are produced in the linear particle accelerator LINAC2, and then further accelerated to 1.4 GeV in Proton Synchrotron Booster (PSB). The protons are

915

916

917

918 then injected into the Proton Synchrotron (PS) to gain the energy of 26 GeV and further
 919 accelerated to 450 GeV in Super Proton Synchrotron (SPS). At the end, they are injected
 920 into the main ring, and can reach a maximum energy of 7 TeV.

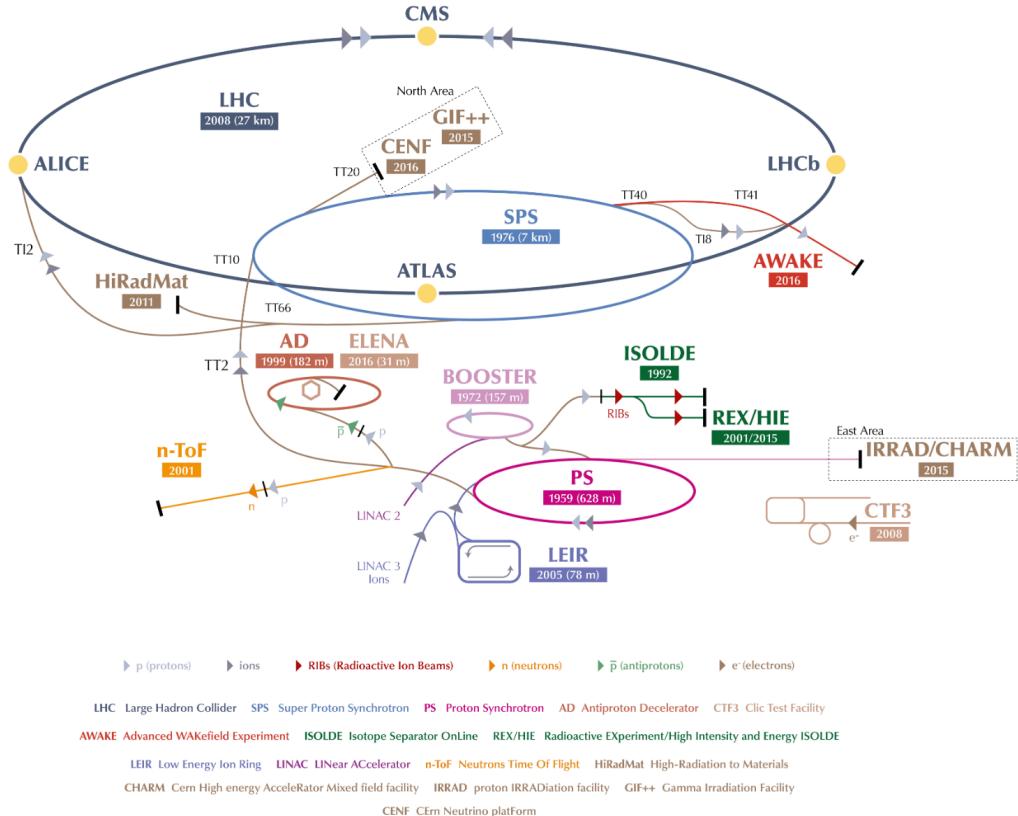


Fig. 3.2 CERN accelerator complex^[16].

921 The collisions can occur in 4 points, with corresponding 4 major detector experiments
 922 that are briefly described as follows:

- **ATLAS:** A Toroidal LHC ApparatuS, one of the two general-purpose particle detector experiments and detector with largest volume at the LHC. It is designed to search for the Higgs boson, test the stardand model of particle physics and search for possible beyond SM physics.
- **CMS:** Compact Muon Solenoid, another large general-purpose particle physics detector, with the same physics goal (also cross check) as ATLAS.
- **ALICE:** A Large Ion Collider Experiment, it is optimized to study heavy-ion (Pb-Pb nuclei) collisions at a centre-of-mass energy of 2.76 TeV per nucleon pair.
- **LHCb:** Large Hadron Collider beauty, it is a specialized b-physics experiment, designed primarily to measure the parameters of CP violation in the interactions of b-hadrons.

3.1.2 Luminosity and pile-up

Luminosity

In beam-beam collisions, the event rate for a process is given by^[15]:

$$N = \mathcal{L}\sigma \quad (3.1)$$

where σ is the cross section of the process, and \mathcal{L} is the luminosity. For the studies of rare events, \mathcal{L} must be as high as possible. The luminosity only depends on the beam parameters, and can be written as:

$$\mathcal{L} = \frac{N_b^2 n f_r \gamma}{4\pi \epsilon_n \beta^*} \quad (3.2)$$

where N_b denotes the number of particles per bunch, n is the number of bunches per beam, f_r is the revolution frequency, γ represents relativistic γ factor, ϵ_n is the normalized transverse emittance and β^* denotes the β function at the collision point. To reduce the beam-beam interaction effects, the bunches must have a crossing angle, which produces a geometrical luminosity reduction factor F :

$$F = 1/\sqrt{1 + \left(\frac{\theta_c \sigma_Z}{2\sigma^*}\right)} \quad (3.3)$$

where θ_c denotes the crossing angle at the interaction point, σ_Z is the root mean square (RMS) bunch length and σ^* is the transverse RMS beam size at crossing point.

The luminosity expressed in Eq. 3.2 is normally the instantaneous luminosity. In fact the running conditions usually vary with time, so the luminosity can change as well. To take into account the time dependence, integrated luminosity is imported, which is the integral over time:

$$L = \int \mathcal{L}(t) dt \quad (3.4)$$

The unit of integrated luminosity we commonly use is b^{-1} ($1b^{-1} = 10^{24} cm^{-2}$). Figure 3.3 shows integrated luminosity as a function of time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during run-2 pp collisions. For most physics analysis, the data with good quality (require to satisfy *Good Run List*) is used.

Pile-up

In collisions, multiple interactions can happen in one single bunch crossing, which is called “*pile-up*”. The variable $\langle \mu \rangle$, representing the average number of interactions per

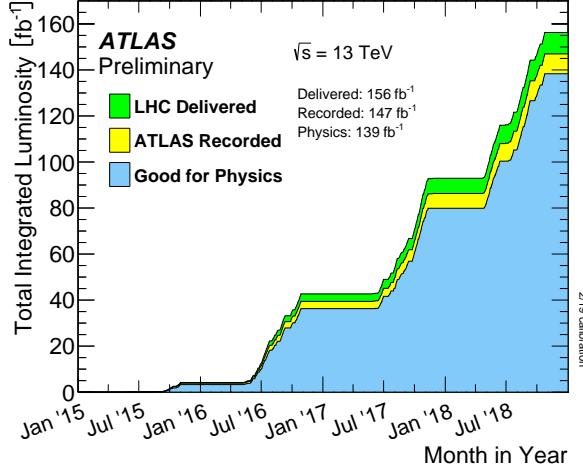


Fig. 3.3 Integrated luminosity in ATLAS.

bunch crossing, is defined to describe pile-up effect:

$$\langle \mu \rangle = \frac{L_{bunch} \sigma}{f_r n_{bunch}} \quad (3.5)$$

where L_{bunch} is the instantaneous luminosity for each bunch, σ the inelastic cross section, f_r the LHC revolution frequency and n_{bunch} the number of colliding bunches. Normally, with increasing luminosity, the pile-up becomes more significant. Figure 3.4 shows the luminosity-weighted distribution of the mean number of interactions per crossing for pp collision data from 2015 to 2018 (full run-2), the challenge of pile-up increased in each year.

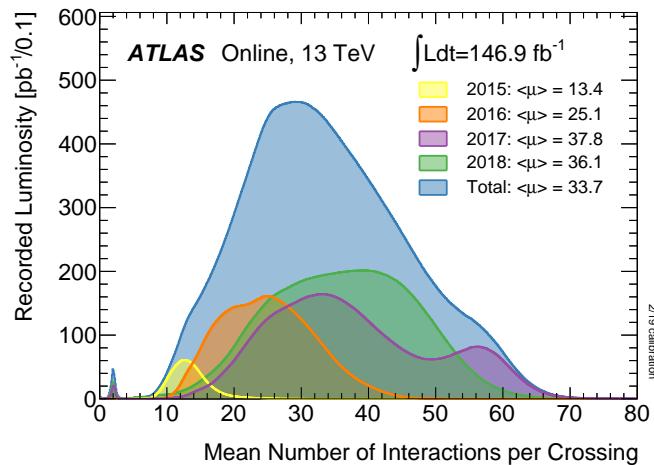


Fig. 3.4 Number of Interactions per Crossing from 2015-2018 in ATLAS.

965

966 3.2 ATLAS detector

967 3.2.1 Detector overview

968 ATLAS (A Toroidal LHC ApparatuS) is the largest volume detector ever constructed
 969 for a particle collider. It is a cylinder with 46 meters long, 25 meters in diameter, and sits
 970 in a cavern 100 meters below ground. The detector contains about 3000 km of cables and
 971 it weights 7000 tonnes.

972 This paragraph briefly summarizes the coordinate system and nomenclature used to
 973 describe the ATLAS detector^[17]. As depicted in figure 3.5, we define the nominal inter-
 974 action point as the origin of the coordinate system, the beam direction as the z -axis and the
 975 x - y plane is transverse to the beam direction. The positive x -axis is defined to be the direc-
 976 tion pointing to the center of the LHC ring, while the positive y -axis is pointing upwards.
 There are two sides of detector A and C, in which A (C) -side is defined as with positive

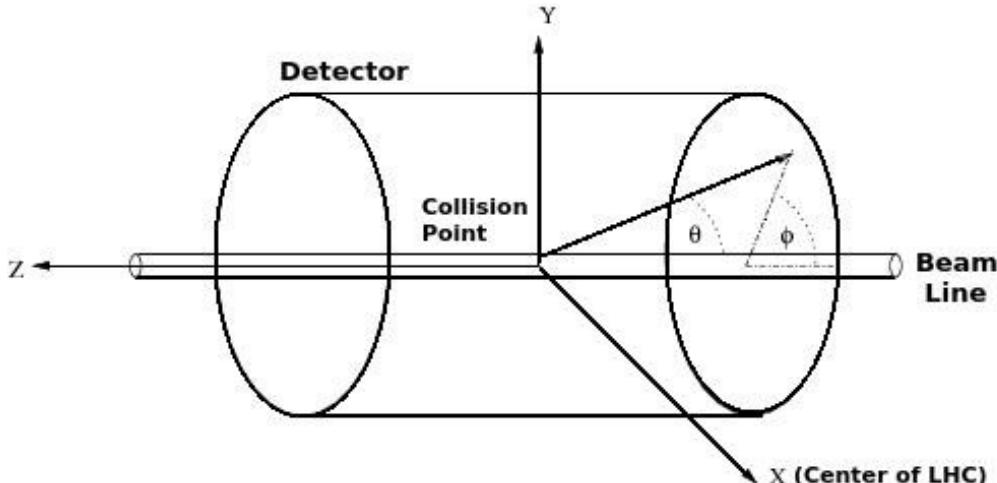


Fig. 3.5 Coordinate system used by the ATLAS experiment at the LHC^[18].

977
 978 (negative) z direction. The azimuthal angle ϕ is measured as usual around the beam axis,
 979 while the polar angle θ is the angle from the beam axis. In physics analysis, we usually
 980 use the pseudorapidity instead of θ angle, which is designed as $\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right]$.

981 For massive objects (eg. jets), the rapidity $y = \frac{1}{2} \ln \left[\frac{E+p_z}{E-p_z} \right]$ is used. In addition, the
 982 transverse momentum p_T , transverse energy E_T and the missing transverse energy E_T^{miss}
 983 are defined in x - y plane. The ΔR , a commonly used distance measurement, is defined in
 984 the pseudorapidity-azimuthal angle space as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

985 The overall ATLAS layout is shown in figure 3.6, which is forward-backward sym-
 986 metric with respect to the interaction point. The magnet configuration comprises a thin
 987 superconducting solenoid surrounding the inner-detector cavity, and three large super-
 988 conducting toroids (one barrel and two end-caps) arranged with an eight-fold azimuthal

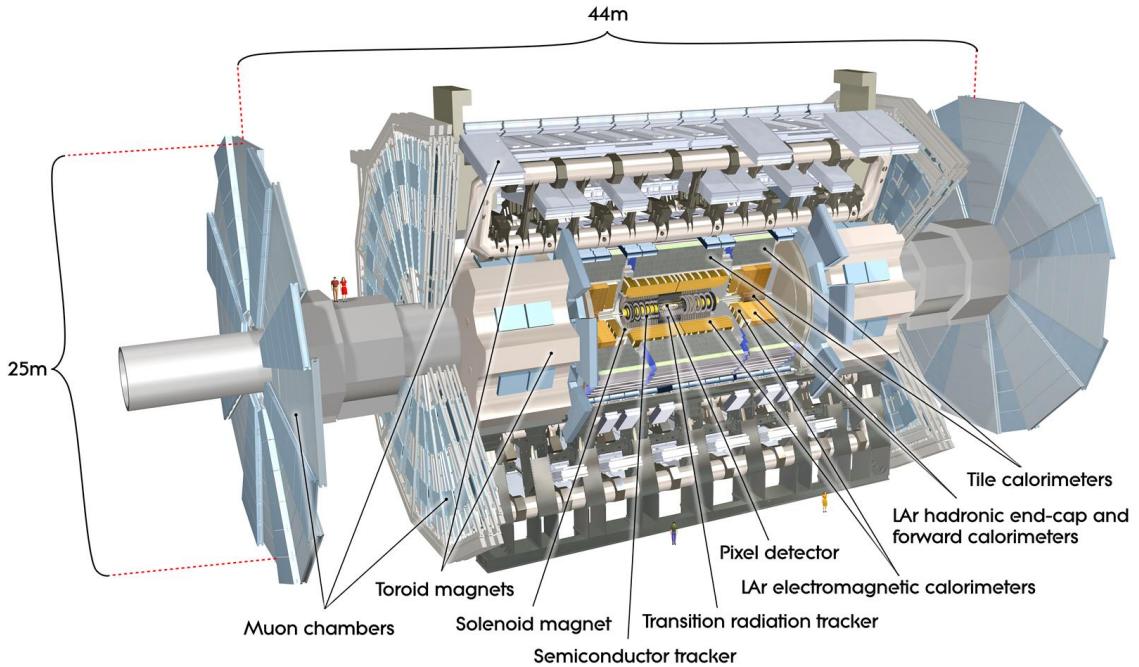


Fig. 3.6 Cut-away view of the ATLAS detector^[19].

989 symmetry around the calorimeters.

990 **The inner detector**, which is the innermost part of ATLAS, is immersed in a 2 T
 991 solenoidal magnetic field. It's used for pattern recognition, momentum and vertex mea-
 992 surements and electron identification, with the combination of tracking system.

993 **The calorimeter** is outside the inner detector, for electromagnetic and hadronic en-
 994 ergy measurements. The high granularity liquid-argon (LAr) electromagnetic sampling
 995 calorimeters is used to measure energy and position resolution with range up to $|\eta| < 3.2$
 996 for electrons and photons. For hadron, a scintillator-tile calorimeter is used in the range
 997 of $|\eta| < 1.7$. And then the LAr forward calorimeters provide both electromagnetic and
 998 hadronic energy measurements with the coverage in forward region up to $|\eta| = 4.9$.

999 **The muon spectrometer** is in the outermost side. It's a air-core toroid system, with
 1000 a long barrel and two inserted end-cap magnets that provides strong bending power in a
 1001 large volume within a light and open structure. Multiple-scattering effects are minor, and
 1002 excellent muon momentum resolution can be achieved.

1003 3.2.2 Physics requirement

1004 As mentioned previously, ATLAS is one of two general-purpose particle detector ex-
 1005 periment at the LHC. It's designed to take advantage of the unprecedented energy at the
 1006 LHC, as the discovery of Higgs boson is one of its benchmark. There are lots of pre-

1007 precise tests and measurements of SM physics are ongoing with ATLAS experiment. In the
1008 meantime, ATLAS is also designed to observe the phenomena that involve highly massive
1009 particles, which can also explore the possibility of extra dimensions proposed by several
1010 models in TeV region. To fulfil many diverse physics goals, a set of general requirements
1011 are needed:

- 1012 • The speed-fast and radiation-hard electronics are required due to the experimental
1013 conditions at the LHC.
- 1014 • High detector granularity is needed to reduce the overlapping events and handle the
1015 particle fluxes.
- 1016 • Large acceptance in pseudorapidity and azimuthal angle coverage is needed.
- 1017 • For inner detector, good charged-particle momentum resolution and reconstruction
1018 efficiency are crucial. And the vertex detectors close to the interaction region are
1019 required to be able to observe secondary vertices for offline tagging of τ -lepton and
1020 b -jets.
- 1021 • Good electromagnetic (EM) calorimetry for electron and photon, as well as full-
1022 coverage hadronic calorimetry for accurate jet and missing transverse energy mea-
1023 surements, are essentially required, since these measurements form the basis of
1024 many studies.
- 1025 • Good muon spectrometer is also required for muon identification and momentum
1026 resolution measurement over a wide range of momenta.
- 1027 • Highly efficient but with sufficient background rejection triggers are also needed
1028 and extremely important for objects with low transverse-momentum.

1029 More detailed descriptions of each sub-system will be given in the following subsec-
1030 tions.

1031 3.2.3 Magnet system

1032 A strong magnetic field is required for precise measurement of charged particle mo-
1033 mента. The ATLAS detector uses two large superconducting magnet systems, a hybrid
1034 system of a central superconducting solenoid and three outer superconducting toroids, to
1035 bend charged particles^[20]. The total magnet system is 22 m in diameter and 26 m in length
1036 as shown in figure 3.7.

1037 The central solenoid produces two Tesla (T) magnetic field surrounding the inner De-
1038 tector. When obtaining such high field strength, at the same time, the solenoid needs to
1039 be thin in order to reduce the material in front of the calorimeter.

1040 The outer toroid system comprises one barrel superconducting toroid and two end-

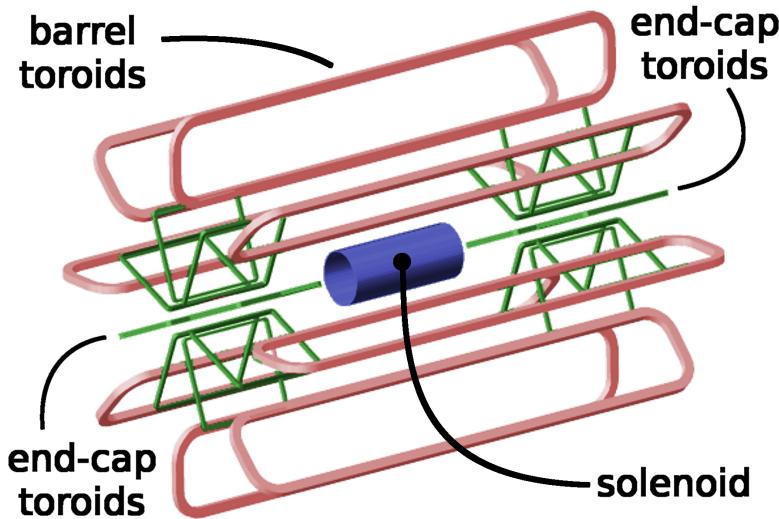


Fig. 3.7 Schematic diagram of the ATLAS magnet system.

1041 caps. The barrel one is composed of eight coils encased in individual racetrack-shaped,
 1042 stainless-steel vacuum vessels and produces the magnetic field in the cylindrical volume
 1043 surrounding the calorimeters. Each end-cap toroid consists of a single cold mass built up
 1044 from eight flat, square coil units and eight keystone wedges and provides a magnetic field
 1045 of approximately 1 T for the muon detectors in the end-cap regions.

1046 3.2.4 Inner detector

1047 The inner detector, as shown in figure 3.8, is the detector closest to beam pipe. It's
 1048 used to measure the position of charged particle tracks in high precision together with
 1049 good momentum resolution, among which the measurement of primary and secondary
 1050 vertices and electron identification are especially important. Due to the extremely high
 1051 luminosity produced by the LHC, the precise measurements of vertex and momentum
 1052 becomes tough and fine-granularity detectors are crucial. The inner detector consists of
 1053 three subdetectors described as below:

1054 **Pixel detector**

1055 The pixel detector^[22] is the innermost part of ATLAS tracking system. With finest
 1056 granularity of materials, it has the best spatial resolution and 3-dimensional space-point
 1057 measurement in inner detector. ATLAS Pixel Detector for the LHC run-2 is composed of
 1058 4 layers of barrel pixel detector and two end-caps with three pixel disks each, as shown in
 1059 figure 3.9. There are three outer layers that originally installed for run-1 and one additional
 1060 layer called Insertable B-Layer (IBL) that newly constructed in run-2^[23]. Now the 4-layer
 1061 pixel detector has very good reconstruction of primary and secondary vertices, which is
 1062 even crucial for long-lived particles like τ -lepton and b-quark.

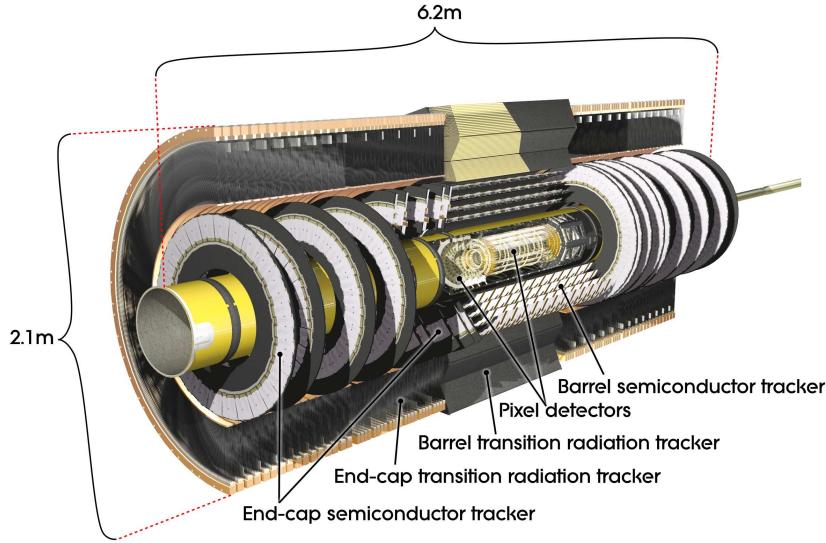


Fig. 3.8 Schematic diagram of the ATLAS inner detector^[21].

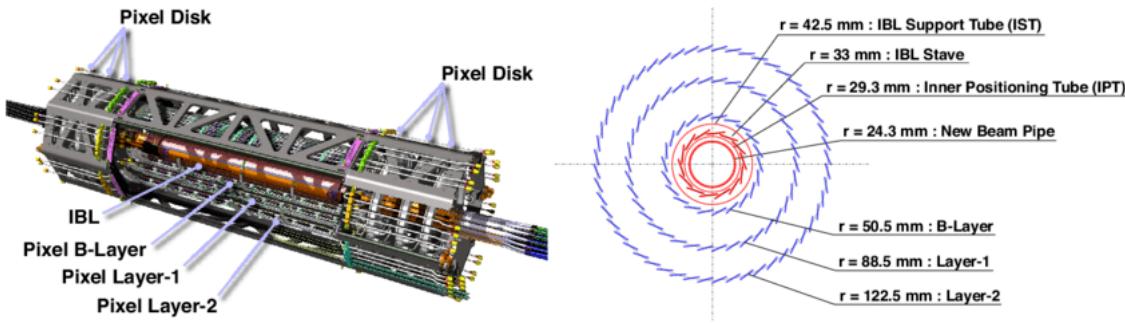


Fig. 3.9 Schematic diagram of the ATLAS 4-Layer Pixel Detector.

Semiconductor Tracker

The Semiconductor Tracker (SCT)^[24] installed outside the pixel detector is the middle component of the inner detector. It has similar function as pixel detector but with long and narrow strips instead of small pixels, which makes a much larger coverage than pixel detector. The SCT consists of 4088 modules, and contains four concentric layers in barrel (2112 modules) and nine disks in each of the two end-caps (1976 modules) as shown in figure 3.10. And it measures particles over a large area with 6.3 million readout channels and a total area of 61 square meters. The SCT is the most critical part of the inner detector for 2D track hit reconstruction. In barrel, the hit precision is $17 \mu\text{m}$ in the $r\text{-}\phi$ coordinate and $580 \mu\text{m}$ in z coordinate. In end-caps, the precision is $17 \mu\text{m}$ in the $z\text{-}\phi$ coordinate and $580 \mu\text{m}$ in r coordinate.

Transition radiation tracker

The transition radiation tracker (TRT)^[26] is the outermost part of inner detector, which has a very different design comparing to the two previously sub-detectors. It can be separated into three parts: one barrel and two end-cap regions with the $|\eta|$ coverage up to 2.0.

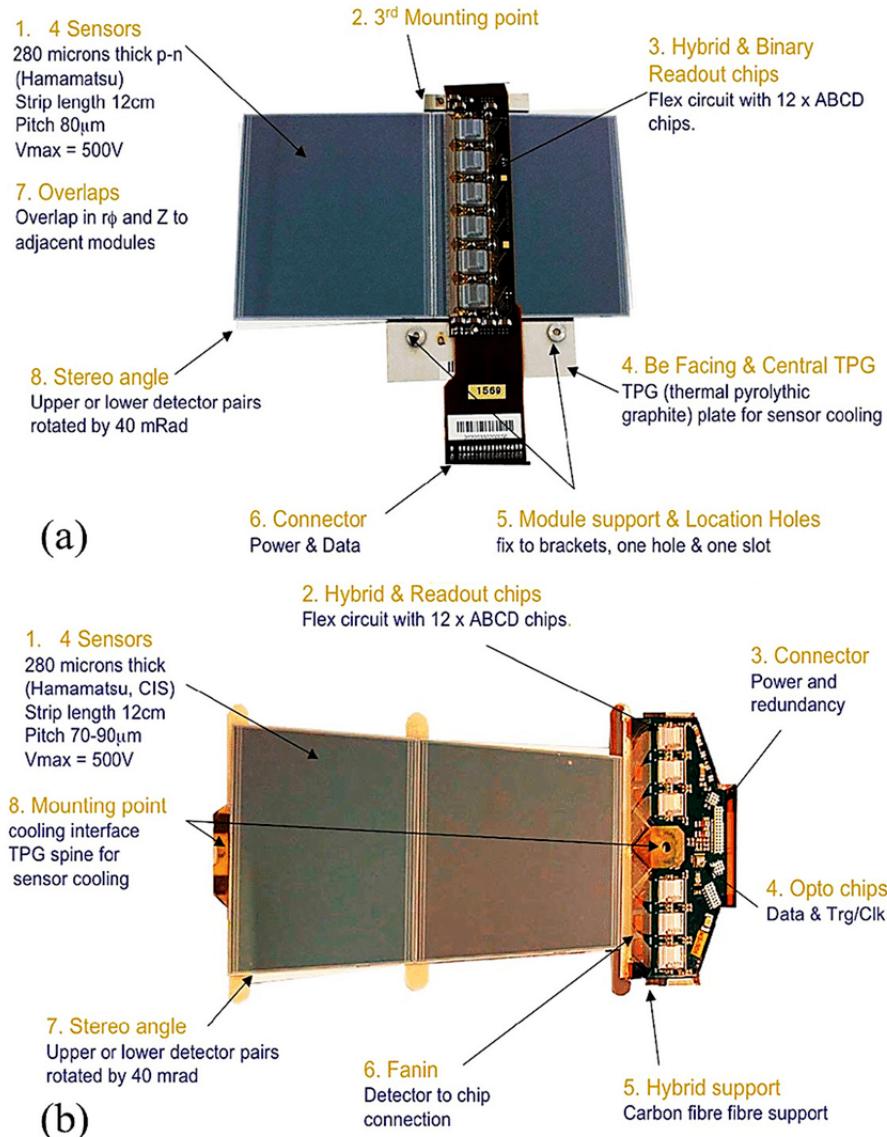


Fig. 3.10 SCT (a) barrel module and (b) end-cap [25].

1078 There are 73 barrel layers and 224 end-cap layers (112 in each) with 372000 straws in to-
 1079 tal, and about 351000 readout channels for TRT. The TRT provides better z resolution but
 1080 much worse $r\phi$ resolution (about 130 μ m) comparing to the pixel detector and SCT per
 1081 straw. But the straw hits still make significant contributions to momentum measurement,
 1082 since its lower precision per point (compared to silicon) can be compensated by the large
 1083 number of measurements and long track length.

1084 3.2.5 Calorimeters

1085 The calorimeters are designed to measure the energy from particles by absorbing them.
 1086 They are located outside the solenoidal magnet that surrounds the inner detector. The
 1087 ATLAS calorimeters are comprised of a number of sampling calorimeters with full ϕ -

symmetry and the pseudorapidity range of $|\eta| < 4.9$. Figure 3.11 shows the layout of the ATLAS calorimeter system. There are two basic calorimeter systems: an inner electromagnetic (EM) calorimeter and an outer hadronic calorimeter. The EM calorimeter is designed for precise measurements of electrons and photons with fine granularity; while the hadronic one has relative coarser granularity but satisfies the physics requirements for jets reconstructions and E_T^{miss} measurements. Two different sampling techniques are used, the EM calorimeter is purely based on liquid-argon (LAr) technology, while the hadronic one use both LAr and scintillating tiles calorimeters. More details are described as below:

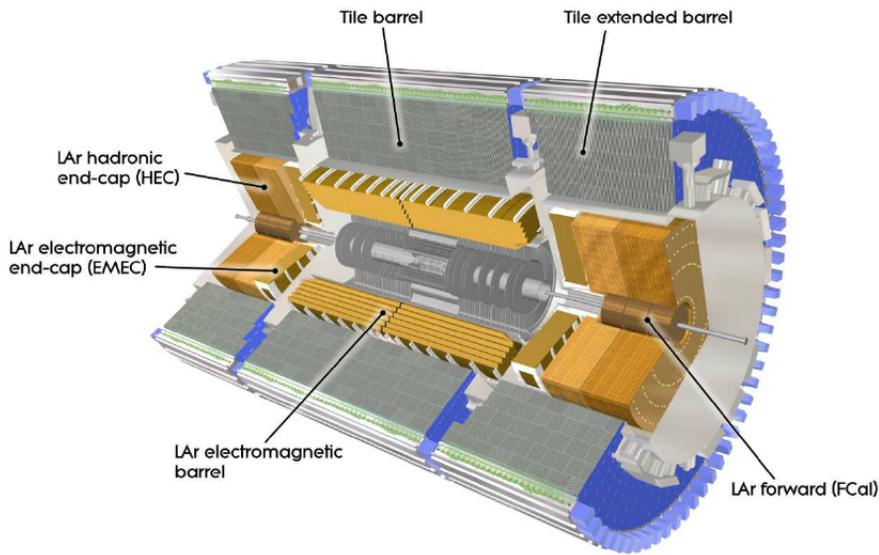


Fig. 3.11 Cut-away view of the ATLAS calorimeters. The LAr calorimeters are seen inside the scintillator-based tile hadronic calorimeters [27].

1096

1097

Liquid Argon calorimeter

1098

1099

1100

1101

1102

1103

1104

1105

1106

The LAr calorimeter uses liquid-argon as active medium. The LAR sampling calorimeter technique with “accordion-shaped” electrodes is used for all electromagnetic calorimetry covering the pseudorapidity range of $|\eta| < 3.2$; and for hadronic calorimetry with range from $|\eta| = 1.4$ to the acceptance limit $|\eta| = 4.9$ [28]. Figure 3.12 depicts the shape of a barrel module, which had an accordion geometry. For barrel EM calorimeter, the absorbing material is lead-liquid argon, while the hadronic end-cap calorimeter uses copper plates. In addition, the forward calorimeter is split into three parts, an EM sector in which copper is used as absorbing material and two hadronic sectors using tungsten outside the EM sector.

1107

Tile calorimeter

1108

Tile calorimeter is a sampling calorimeter using scintillating plates as active medium

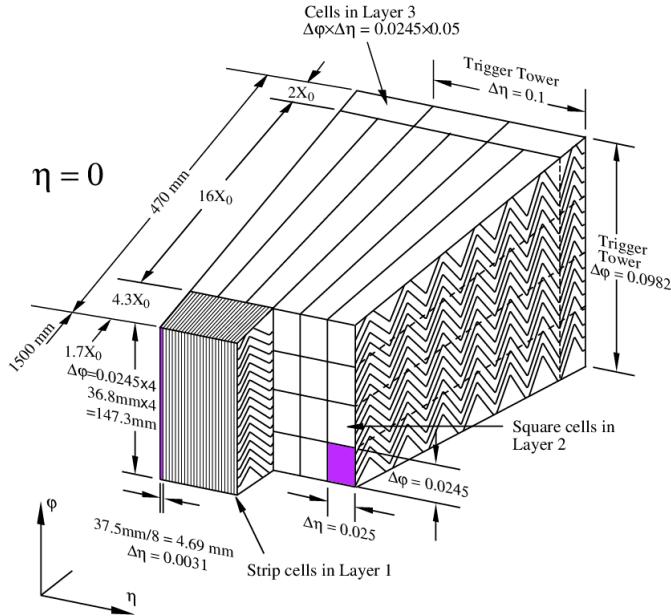


Fig. 3.12 Schematic diagram of a LAr EM calorimeter barrel module.

and steel as absorber. It consists of three sections: the central barrel with the pseudorapidity range of $|\eta| < 1.0$ and two extended barrels with $0.8 < |\eta| < 1.7$. Figure 3.13 shows the design of one tile calorimeter module. It's used for energy reconstruction of jets and E_T^{miss} measurement by combining with the forward and end-cap hadronic calorimeter.

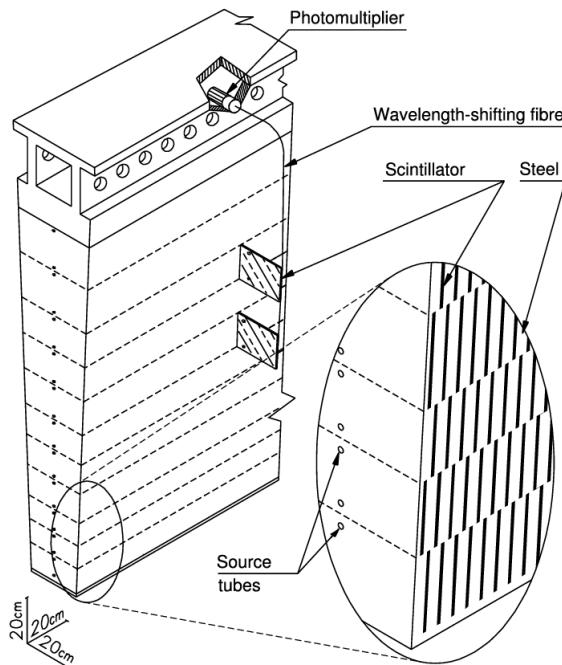


Fig. 3.13 Schematic diagram of tile calorimeter module^[29].

3.2.6 Muon spectrometer

Muon spectrometer^[30] is the outermost part of the ATLAS detector with an extremely large tracking system. It measures a large range of muon momentum, and the accuracy is about 3% at 100 GeV and 10% at 1 TeV. The muon spectrometer comprises three main parts: a magnetic field produced by three toroidal magnets; a set of chambers measuring the tracks of muons with high spatial precision; and triggering chambers with accurate time-resolution. Figure 3.14 shows the schematic of ATLAS muon spectrometer that consists of four types of muon chambers (*MDT*, *CSC*, *RPC*, *TGC*) as well as the magnet systems (barrel and end-cap toroid).

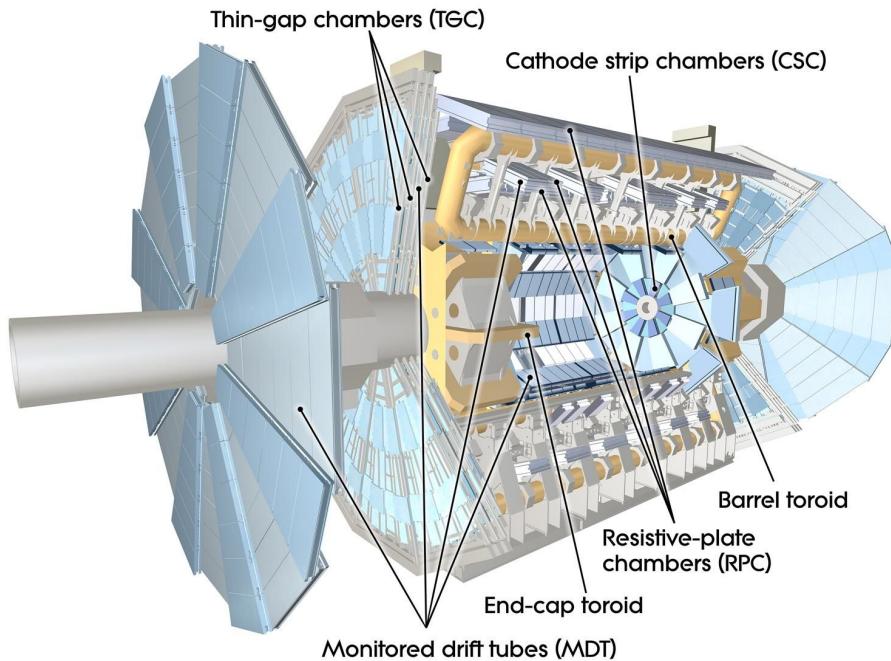


Fig. 3.14 Cut-away view of the ATLAS muon spectrometer^[31].

More details of four chambers are given as below:

- **Monitored Drift Tubes (MDT).** MDTs provide the precise momentum measurement with the $|\eta|$ range up to 2.7, except in the innermost end-cap layer where the coverage is limited to $|\eta| < 2.0$. The chambers comprises three or four layers of drift tubes, with a diameter of 29.970 mm, operated with Ar/CO₂ gas (93/7) at 3 bar. The average resolution can reach 80 μm per tube and 30 μm per chamber.
- **Cathode strip chambers (CSC).** CSCs are used in the forward region of $2 < |\eta| < 2.7$ in the innermost tracking layers, due to their good time resolution and high rate capability. The CSCs are multi-wire proportional chambers (MWPC) with the cathode planes segmented into strips in orthogonal directions, which allows both coordinates to be measured from the induced-charge distribution. The resolution of

- 1133 a chamber is about $40 \mu\text{m}$ for bending plane and 5 mm for the transverse plane.
- 1134 • **Resistive plate chambers (RPC).** The RPCs serve as fast triggers in the barrel
- 1135 region of $|\eta| < 1.05$ due to the high rate capability and good spatial and time res-
- 1136 olution. It is a gaseous parallel electrode-plate detector without any wires. There
- 1137 are three concentric cylindrical layers around the beam axis, as three trigger sta-
- 1138 tions. Each station consists of two independent layers to measure the transverse
- 1139 coordinates of η and ϕ .
- 1140 • **Thin gap chambers (TGC).** TGCs are used as trigger system for the end-cap region
- 1141 of $1.5 < |\eta| < 2.4$, and works based on the same principle as multi-wire propor-
- 1142 tional chambers. In addition, they can also provide the second azimuthal coordinate
- 1143 to complement the measurement of MDT in bending direction.

1144 3.2.7 Trigger system

1145 Trigger system in ATLAS is a very essential component, which is responsible for

1146 deciding whether to keep a given collision event for later study or not. In the LHC run-

1147 2, higher energy, luminosity and pile-up lead to a large increase of event rate by up to

1148 a factor of five, which causes to a even larger challenge and more strict requirement of

1149 trigger system.

1150 The trigger system in run-2 consists of a hardware-based first level trigger (Level-1)

1151 and a software-based high level trigger (HLT)^[32]. As depicted in figure 3.15, in Level-1,

1152 the inputs from coarse granularity calorimeter and muon detector information together

1153 with from some other subsystems are sent to the Central Trigger Processor to determine

1154 Regions-of-Interest (RoIs) in the detector. The event rate can be reduced by Level-1 trig-

1155 gers from 30 MHz to 100 kHz. After that, the RoI information from Level-1 is sent to

1156 HLT, in which more sophisticated selection algorithms are run for regional reconstruction.

1157 The HLT reduces the rate from Level-1 from 100 kHz to about 1 kHz on average. At the

1158 end, the events that accepted by HLT are transferred to local storage at experimental site

1159 for offline reconstruction. Details about Level-1 and HLT trigger systems are described

1160 as below:

1161 **Level-1 trigger**

1162 Substantial upgrades have been delivered in ATLAS Level-1 trigger system for run-2

1163 data taking. The upgrades took place in both hardware and detector readout, allow the

1164 trigger rate increasing from 70 kHz (run-1) to 100 kHz (run-2). There are two major parts

1165 of Level-1 triggers, which include Level-1 calorimeter (L1calo) trigger and Level-1 muon

1166 (L1mu) trigger.

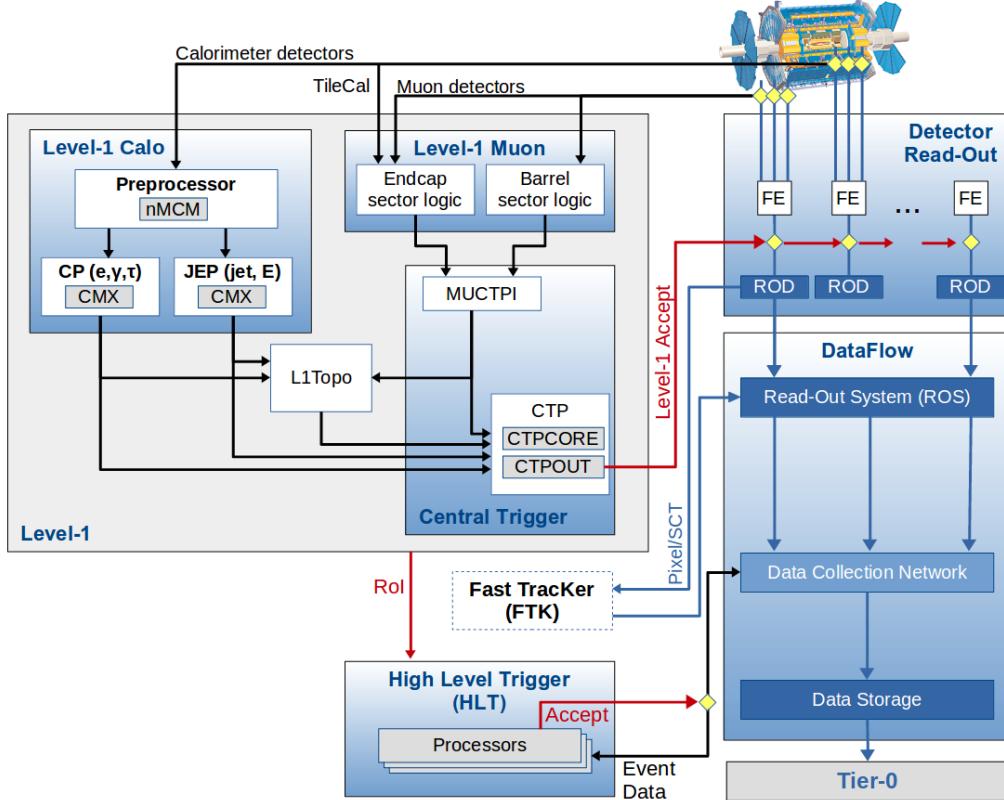


Fig. 3.15 Schematic diagram of the ATLAS trigger and data acquisition system in run-2.

1167 Level-1 Calorimeter trigger uses the reduced granularity information from the elec-
 1168 tromagnetic and hadronic calorimeters to search for electrons, photons, taus and jets and
 1169 missing transverse energy (E_T^{miss}). It can identify an Region-of-Interest (RoI) as a 2×2
 1170 trigger tower cluster in the EM calorimeter as shown in figure 3.16, and 4×4 or 8×8
 1171 trigger tower for Jet RoIs. One important upgrade was that, the new FPGA-based (field-
 1172 programmable gate array) Multi-Chip Modules are used to replace the ASICs (application-
 1173 specific integrated circuits) included in the modules used in run-1, which allows the use
 1174 of auto-correlation filters to suppress pile-up.

1175 The Level-1 Muon trigger system includes one barrel section (RPC) and two end-cap
 1176 section (TGC), which provides fast trigger signals from the muon detectors for the Level-1
 1177 trigger decision. By requiring a coincidence with hits from the innermost muon chambers,
 1178 it can reduce the L1_MU15 rate by about 50% in the region of $1.3 < |\eta| < 1.9$ while only
 1179 loss around 2% signal efficiency. In addition, the coverage was extended by around 4%
 1180 due to installing new chambers in the feet region of the muon detector.

1181 High Level Trigger

1182 In run-1, the Event Filter computer clusters and Level-2 trigger system were separated,
 1183 while now in run-2, they have been merged into a single HLT event processing. The
 1184 new arrangement helps to reduce the complexity and duplication of algorithm, which

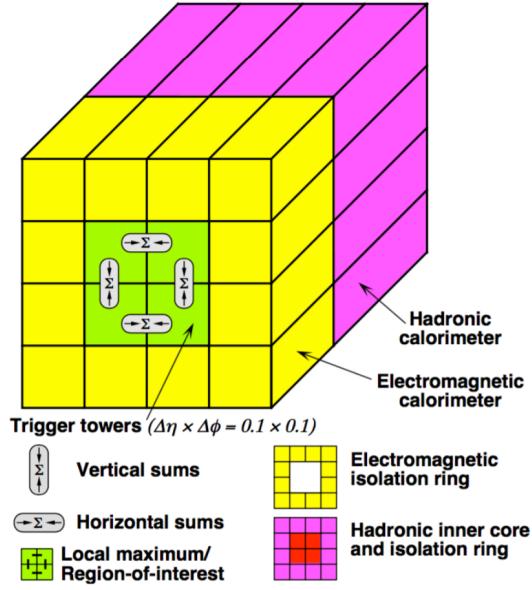


Fig. 3.16 An examples of L1 calorimeter trigger tower for electron and photon triggers^[33].

leads to a more flexible high level trigger system. During the long-shutdown between the LHC run-1 and run-2, lots of re-optimizations have been done for trigger reconstruction algorithms as well as the offline analysis selections, which can improve the efficiency by more than a factor of two in some cases like hadronic tau triggers. For some triggers, the HLT processing performed within RoIs can also allows to aggregate from RoIs to single objects. This improvement reduces the CPU processing for events with overlapping RoIs, and the average output rate has been increased from 400 Hz to 1 kHz.

The HLT reconstruction algorithm can be divided into fast and precision online reconstruction steps. As illuminated by figure 3.17, the initial fast reconstruction helps to reduce the event rate, and be seeded into precision reconstruction. Then the final online precision reconstruction is improved and uses offline-like algorithms as much as possible. In particular, multivariate analysis techniques (based on machine learning) have been introduce online in many aspects.

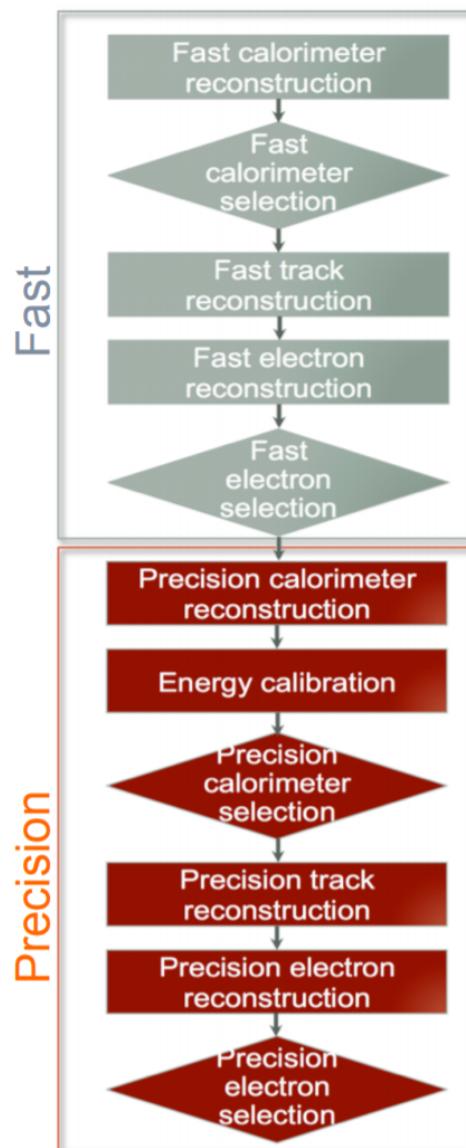


Fig. 3.17 The HLT trigger algorithm sequence^[33].

1198 Chapter 4 Simulation and Event Reconstruction for 1199 the ATLAS Experiment

1200 In current LHC pp collision, bunches of protons collide every 25 nanoseconds (ns),
1201 which gives a large challenge to event reconstruction and selections. To predict and model
1202 each process, the Monte Carlo simulations of physics events are essential for high-energy
1203 physics experiments. This section will briefly discuss the event simulation and recon-
1204 struction programs based on the ATLAS software framework.

1205 4.1 Event simulation

1206 The ATLAS simulation program is integrated into the ATLAS software framework
1207 called *Athena*^[34], which uses Python as an object-oriented scripting and interpreter lan-
1208 guage to configure and load C++ algorithms and objects. Figure 4.1 shows the overview
1209 of ATLAS simulation data flow^[35]. In the diagrams, the square-cornered boxes represents
algorithms and applications to be run and round-cornered boxes denote data objects.

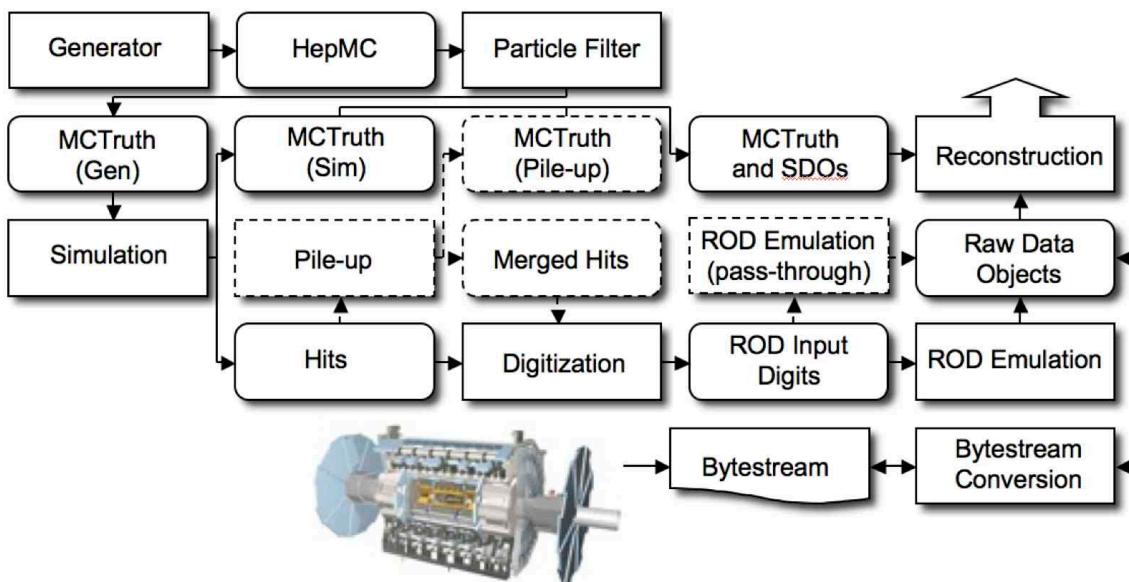


Fig. 4.1 The flow of the ATLAS simulation software.

1210
1211 First of all, events are produced by MC generators in standard HepMC format and
1212 then read into the simulation. During the simulation, particles are propagated through the
1213 full ATLAS detector whose configurations can be set by users via GEANT4 toolkit. The
1214 energies deposited in the sensitive regions of the detector are recorded as *hits* that contains
1215 the total energy deposition, position and time, and are written to a simulation hit file. In

1216 the meantime, the events in "truth" format are also recorded to contain the history of the
1217 interactions from the generator, including incoming and outgoing particles. Simulated
1218 Data Objects (SDOs) are created from truth, which are maps between hits in sensitive
1219 portions of the detector and truth information of particles in simulation. The files are then
1220 sent to digitization, with constructs "digits" inputs and be written into Raw Data Object
1221 (RDO) file used for reconstruction.

1222 In conclusion, there are three main parts of framework: *Generation*, *Simulation* and
1223 *Digitization*. More details are given as below:

1224 **Event generation**

1225 As shown in figure 4.2^[36], at hardon colliders, multiple scattering and rescattering
1226 effects arise, which needs to be simulated by Monte Carlo (MC) event generators to reflect
1227 the full complexity of those event structures. Several MC event generators can be used
1228 to generate events originally in HepMC format. The events can be filtered at generation
1229 time with some certain requirements (eg. decay channel or missing energy above a certain
1230 threshold). The generator is responsible for any prompt decays (e.g. W or Z bosons) but
1231 stores any "stable" particle expected to propagate through a part of the detector. During
1232 the generation steps, any interactions with detector are ignored and only immediate decays
1233 are considered.

1234 There are several MC generators that have been widely used with general purpose,
1235 which include Sherpa^[37], Herwig++^[38], PowhegBox^[39], MC@NLO^[40] and Pythia8^[41].

1236 **Simulation**

1237 GEANT4 is used as standard simulation toolkit for the ATLAS experiment, which
1238 transports physics particles through the detector's geometry. During the generation level,
1239 the entire connected chain of the HepMC event is stored as the Monte Carlo truth. Only the
1240 stable particles are read into GEANT4 for further simulation and selection, while transfor-
1241 mations can be applied to these events to select certain processes. During the simulation,
1242 many secondary tracks can be produced, therefore only information from the interactions
1243 of interest are stored, including the incoming particles, step sequence, vertex as well as
1244 outgoing particles. The output of GEANT4 is called *hit file*, which contains metadata de-
1245 scribing the configuration of the simulation during the run, all truth information requested
1246 and a collection of hits for each subdetector.

1247 Since the standard ATLAS detector simulation cost very large computing resources
1248 to accurately model the complex detector geometry and physics descriptions, some fast
1249 simulation programss are developed according to different user purpose. Some popular
1250 fast-sim toolkits include *Fast G4 Simulation*^[42], *ATLFAST-I*^[43] and *ATLFAST-II*^[44].

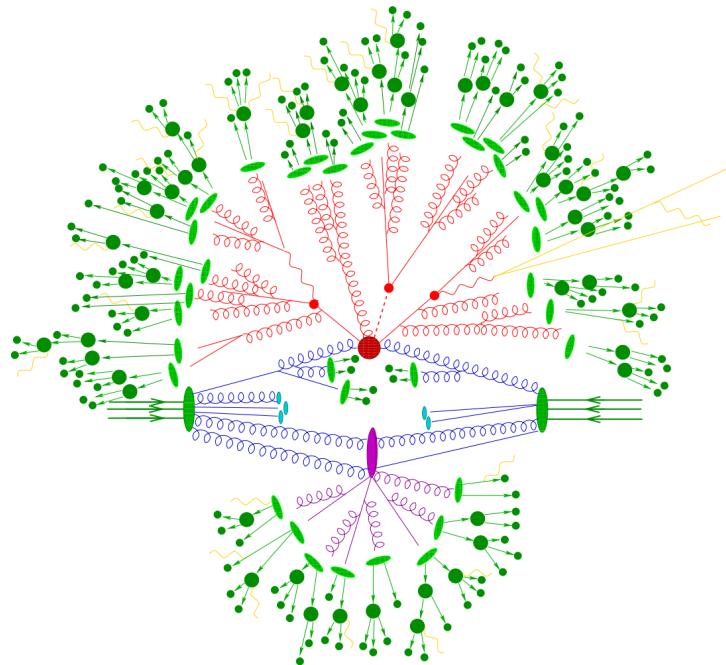


Fig. 4.2 Sketch of a hardon-hardon collision simulated by MC event generator. The red blob in center denotes the hard collision, surrounded by tree-like structures representing Bremsstrahlung which is simulated by Parton Showers. The purple blob stands for a secondary hard scattering event. The light green blobs indecate the parton-to-hardon transitions and the dark green blobs represents hadron decays. The yellow lines are soft photon radiations.

1251 Digitization

1252 The hit outputs from simulated events, including hard scattering signal, minimum bias,
 1253 beam halo, beam gas and cavern background events, are then sent into digitization pro-
 1254 cedure, converted into detector response called “digits”. Before converted into detector
 1255 signal as ‘digits’ formart, each type of event can be overlaid at a user-specified rate. Those
 1256 overlay, called “pile-up”, can be done during degitization to save the CPU time. At this
 1257 stage, the detector noise and the first level trigger that implemented with hardware on the
 1258 real detector are added into events. The digitization firstly constructs “digits” inputs to
 1259 the readout drivers (RODs) in the detector electronics. Then the ROD functionality is em-
 1260 ulated, and the output digits are written out as Raw Data Object (RDO) file. In addition,
 1261 the digitization algorithms can also produce Simulated Data Objects (SDOs), which con-
 1262 tain information about all the particles, noise and the amount of energy that contributed
 1263 to the signal. Then all information are sent into reconstruction level described in next
 1264 subsection.

4.2 Event reconstruction

The data flow of ATLAS data processing is sketched in figure 4.3^[45]. Data from detector is firstly filtered by online trigger system before sending to the *Tier-0* (*T0*) for initial processing by offline reconstruction software based on Athena. A small amount of data named “express stream” is processed in almost real time in *T0* for online data quality monitoring. In addition, some other dedicated data streams are sent out at trigger level for detector alignment and calibration. These calibration and alignment information are then used for bulk reconstruction in *T0*. At the end of the reconstruction chain, the data are delivered into *Tier-1* (*T1*) and *Tier-2* (*T2*) centers for further analysis and production of simulated data. *T1* centers are also responsible for data reprocessing by re-running data reconstruction with improved calibration and alignment constants and with improved reconstruction algorithms.

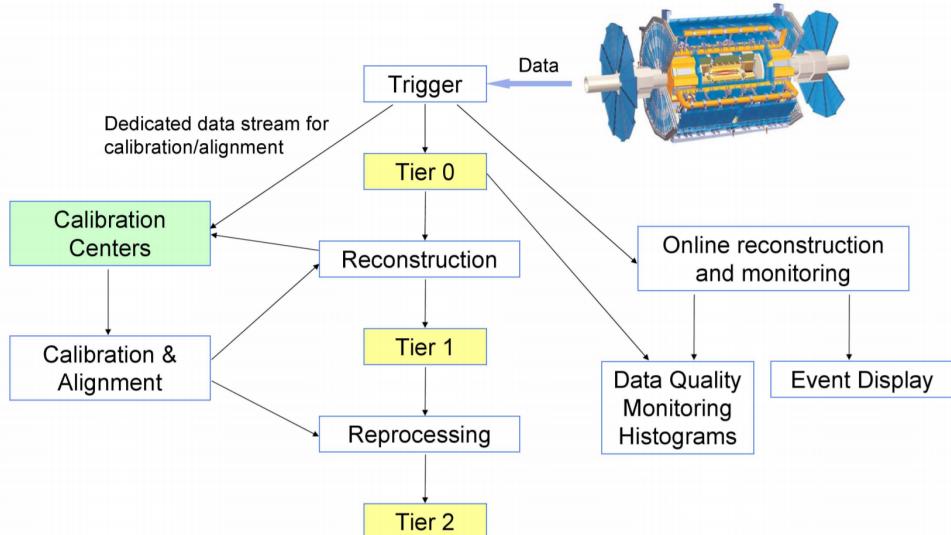


Fig. 4.3 The flowchart of the ATLAS data processing.

This section describes the reconstruction of some important physics objects in ATLAS experiment: tracks, vertices, electrons, muons, jets, and missing energies.

4.2.1 Track

The ATLAS detector is composed of two independent tracking systems: the Inner Detector (ID) close to the interaction point, and the Muon Spectrometer (MS) located in the outermost region. The reconstructed charged-particle trajectories in the ID and MS are referred to as ID tracks and MS tracks respectively. The challenge of ID reconstruction is that it needs to handle high track density that imposes a large number of combinatorial track candidates, while the MS reconstruction is however largely limited by the huge

amount of inert material, the large background and the highly inhomogeneous magnetic field^[46]. More details of these two types of track are given as below:

1288 **Inner detector track**

1289 Figure 4.4 sketches the ID system used for detecting charge-particle tracks. The ID
1290 track reconstructions contains two sequences: *inside-out* track reconstruction and *outside-in* one.

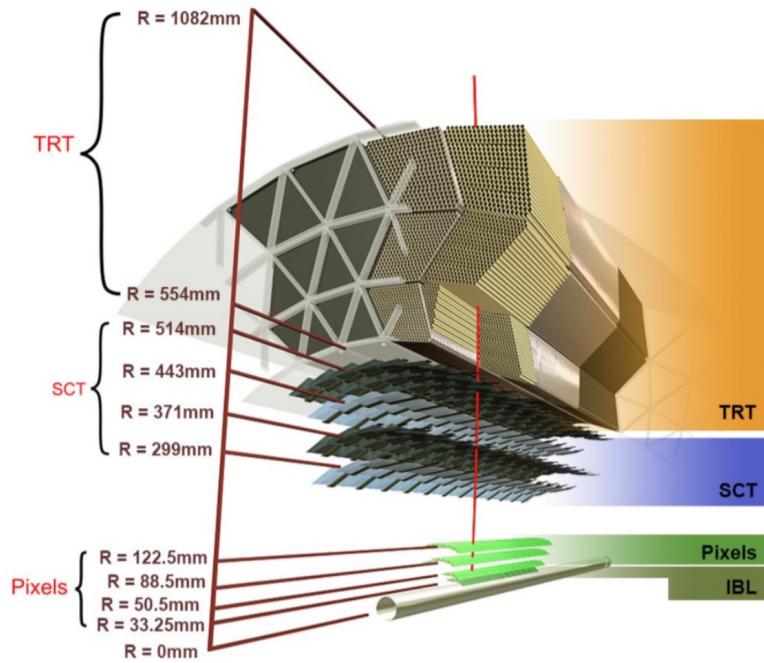


Fig. 4.4 Schematic view of the ATLAS inner detector showing all the corresponding components.

1291
1292 For inside-out tracking, it exploits the high granularity of the pixel and SCT detectors
1293 to discover prompt tracks originating from the interaction point. In first step, the track
1294 seeds are formed by combining the information of space-points in the three pixel layers
1295 and the first SCT layer. Then, these seeds are extended throughout the SCT to build track
1296 candidates. After that, these candidates are fitted with some quality cuts applied to remove
1297 the outlier clusters, reject the fake tracks and resolve ambiguities in the cluster-to-track
1298 association. The selected tracks are then further extended to TRT, and refitted with the
1299 full information from pixel, SCT and TRT detectors.

1300 Another complementary approach, outside-in, searches for unused track segments
1301 start from TRT instead. These segments are then extended into the SCT and pixel detec-
1302 tors to improve the tracking efficiency for secondary tracks from conversions or decays
1303 of long-lived particles.

1304 **Muon spectrometer track**

1305 The MS track reconstruction^[47] starts from searching hit patterns inside each muon
 1306 chamber to form segments. In each MDT chamber and nearby trigger chamber, a Hough
 1307 transform^[48] is used to search the hits lie on a certain trajectory in the bending plane of the
 1308 detector. The MDT segments are reconstructed by performing a linear fit to the hits found
 1309 in each layer. The RPC or TGC hits can be built by measuring the coordinate orthogonal to
 1310 the bending plane. And the segments of CSC can be built using a separate combinatorial
 1311 search in the η and ϕ detector planes.

1312 Then muon track candidates are built by fitting hits from segments in different layers
 1313 together. This task makes use of the algorithm by performing a segment-seeded combi-
 1314 natorial search, which starts by using the segments generated in the middle layers of the
 1315 detector where more trigger hits are available as seeds. The search is then extended to use
 1316 the segments as seeds from the inner and outer layers. The segments are selected based
 1317 on criteria of hit multiplicity and fit quality, and are matched using their relative positions
 1318 and angles. To build a track, at least two matching segments are required, except in the
 1319 barrel-endcap transition region where a single high-quality segment with η and ϕ infor-
 1320 mation can be used to build a track. At beginning, the same segment can be used to build
 1321 more than one track candidates. Later on, an overlap removal algorithm is performed to
 1322 select the best assignment to a single track, or decide whether allows the certain segment
 1323 to be shared between two tracks.

1324 The hits associated with each track candidate are then fitted using a global χ^2 fit.
 1325 The algorithm accepts the track candidate if its fitting χ^2 passes the selection criteria.
 1326 Hits contribute largely to χ^2 are removed and the track fit is repeated. In addition, the
 1327 algorithm performs a hit recovery procedure that looks for additional hits consistent with
 1328 the candidate trajectory, and the track candidate is refit if additional hits are found.

1329 4.2.2 Primary vertex

1330 The reconstruction of primary vertex (PV) uses the reconstructed tracks introduced in
 1331 previous section as inputs. The tracks must satisfy the following criteria^[49]:

- 1332 • $p_T > 400$ MeV
- 1333 • $|\eta| < 2.5$
- 1334 • Number of silicon hits $\geq \begin{cases} 9 & \text{if } |\eta| \leq 1.65 \\ 11 & \text{if } |\eta| > 1.65 \end{cases}$
- 1335 • IBL hits + B-layer hits ≥ 1
- 1336 • A maximum of 1 shared module (1 shared pixel hit or 2 shared SCT hits)
- 1337 • Pixel holes = 0

1338 • SCT holes ≤ 1

1339 A candidate vertex is formed by requiring two tracks passing these selection criteria.

1340 The reconstruction of PV can be divided into two steps^[50]: vertex finding and vertex
1341 fitting. The first step is the pattern recognition process, namely the association of recon-
1342 structed tracks to vertex candidates. The latter one works on the reconstruction of the
1343 actual vertex position and its covariance matrix. More details are described as below:

1344 First of all, a set of tracks passing the selection criteria mentioned above is selected.
1345 Then a seed position for the first vertex is chosen. This seed position is determined by
1346 beam spot in the transverse plane. The starting point for x- and y- coordinates are di-
1347 rectly from the centre of the beam spot, while the one for z-coordinate is calculated as
1348 the mode of z-coordinates of tracks at their respective points with closest approach to the
1349 reconstructed centre of the beam spot.

1350 After determining the seed position, the iterative primary vertex finding procedure
1351 starts. An vertex fitting algorithm is adopted to find the optimal vertex position by per-
1352 forming an iterative χ^2 minimization, in which the seed position is used as the start point
1353 and the reconstructed tracks are used as input measurements. For this fitting procedure,
1354 the input tracks are assigned weights to reflect their compatibility with the vertex estima-
1355 tion, and the vertex position is re-calculated based on these weighted tracks. Then the
1356 iterative procedure is repeated by re-calculating the track weight according to the new
1357 vertex position. After the last iteration, the final weight of each track used in vertex fit is
1358 estimated. And those incompatible tracks ($> 7\sigma$) are then rejected from this vertex candi-
1359 date and moved back to the unused pool for next vertex finding. Then iteration procedure
1360 describes above are repeated again by using the remaining tracks, until no un-associated
1361 tracks are left or no additional vertex can be found in remaining tracks.

1362 At the end, the vertices with at least two associated tracks passing through are treated
1363 as possible PV candidates. And the output of this vertex reconstruction algorithm is the in-
1364 formation of three dimensional vertex positions and their covariance matrices. In physics
1365 analysis, it's most often to choose the one with highest sum of transverse momentum
1366 ($\sum p_T^2$) as PV.

1367 4.2.3 Electron

1368 Many interesting physical processes are with the involvement of one or more elec-
1369 trons (or positrons) at the LHC. But these electrons can be subjected to large amount of
1370 backgrounds such as hadrons, non-prompt electrons from photon conversions and non-
1371 isolated electrons from heavy flavor hadon decays. It is therefore essential to efficiently

1372 reconstruct and identify electrons as well as, in the meantime, to keep high background
1373 rejection.

1374 In ATLAS, in central region, the electrons leave tracks in inner detector (ID) and
1375 deposit the energies in the electromagnetic (EM) calorimeter. Firstly the signals from
1376 calorimeter are used for L1 trigger system, and them combined with the information from
1377 ID tracks to reconstruct electron candidates that will be used for the high level trigger
1378 (HLT) decision algorithms^[51]. The backgrounds mentioned above can then be further
1379 suppressed by using several identification criteria. In addition, electrons are required to
1380 be isolated from other activities to be further distinguished from background.

1381 More details of electron *reconstruction*, *identification* and *isolation* are described as
1382 below.

1383 Electron reconstruction

1384 Several steps are proceeded for electron reconstruction in the central region of ATLAS
1385 detector ($|\eta| < 2.47$):

1386 1. **Seed-cluster reconstruction:** A sliding window with size of 3×5 in unit of
1387 $\Delta\eta^{tower} \times \Delta\phi^{tower} = 0.025 \times 0.025$ in $\eta \times \phi$ space is utilized to search for elec-
1388 tron cluster seeds with total cluster transverse energy greater than 2.5 GeV. Then
1389 a clustering algorithm^[52] is applied to form the clusters around the seeds, which
1390 can take advantage of removing the duplications. The kinematics of clusters are
1391 then reconstructed by using an extended window depending on the cluster position.
1392 The efficiency of cluster search is from about 95% at $E_T = 7GeV$ to 99% for
1393 $E_T \geq 15GeV$.

1394 2. **Track reconstruction:** The track reconstruction can be divided into two steps: pat-
1395 tern recognition and track fit. The standard pattern recognition in ATLAS uses pion
1396 hypothesis for energy loss caused by interactions with detector material. If a track
1397 seed with $p_T > 1$ GeV cannot be successfully extended to a full track required at
1398 least seven hits using this pion hypothesis, but still falls inside one of the EM clus-
1399 ter region of interest, as a second attempt, the pattern recognition using electron
1400 hypothesis is then used to allow larger energy loss. Depending on the pattern used
1401 in previous stage, the track candidates are then fitted with either the pion hypothesis
1402 or the electron hypothesis by using ATLAS Global χ^2 Track Fitter^[53]. If a track
1403 candidate fails the fit by using pion hypothesis, it can be refit with the electron hy-
1404 pothesis again. In this method, a specific electron-oriented algorithm is integrated
1405 into the ATLAS standard track reconstruction, which improves the performance for
1406 electron and as well as maintain minimal interference with the main track recon-

1407 struction.

1408 **3. Electron specific track fit:** Once the tracks are obtained, they are loosely matched
1409 to EM cluster using the distance in η and ϕ between the position of track (after ex-
1410 trapolation) in calorimeter's middle layer and the cluster barycentre. The matching
1411 conditions take into account the energy loss of bremsstrahlung and the number of
1412 precise hits in silicon detector.

1413 **4. Electron candidate reconstruction:** The electron candidate is reconstructed by
1414 matching the track candidate to EM cluster seed to eventually completes the electron
1415 reconstruction procedure. If more than one track satisfy the matching condition,
1416 one track is chosen as primary track based on the information of the cluster-track
1417 distance R, the number of pixel hits and the presence of a hit in the first silicon
1418 layer^[54]. In addition, the electron candidates are removed from electron pool if it's
1419 without any associated precise hit tracks, and moved into photon candidates pool.
1420 Then we reformed the electron clusters by using 3×7 (5×5) longitudinal towers
1421 of cells in barrel (end-caps) in EM calorimeter. The measured energy is calibrated
1422 to original electron energy based on MC simulated samples by using multivariate
1423 techniques (MVA).

1424 In addition, in physics analysis, to reduce the background from photon conversions
1425 and secondary particles, the track associated with electron is required to be compatible
1426 with the primary vertex of the hard collision. Practically, the impact parameters cuts such
1427 as $d_0/\sigma_{d_0} < 5$ and $z_0 \sin\theta < 0.5$ mm are usually applied, where d_0 is the closest distance
1428 of the track to the measured beam-line, z_0 is the distance along the beam-line between
1429 the point where d_0 is measured and the beam-spot position, and the θ is polar angle of
1430 the track, σ_{d_0} denotes the estimated uncertainty of d_0 parameter. Figure 4.5 depicts the
1431 definition of each track impact parameter.

1432 **Electron identification**

1433 The electron identifications (ID) are applied to determine whether the reconstructed
1434 electron candidate is more signal-like or background-like object. The ID algorithms
1435 make use of quantities of related variables from electron cluster and track measurements
1436 including calorimeter shower shapes, track properties, as well as variables measuring
1437 bremsstrahlung effects for distinguishing signal from background. Taking the advantage
1438 of new IBL in run-2, the number of hits in this innermost pixel layer is utilized for discrim-
1439 inating between electrons and converted photons. In addition, a likelihood method based
1440 on the TRT high-threshold hits is adopted to compensate the lower transition radiation
1441 absorption probability of the argon.

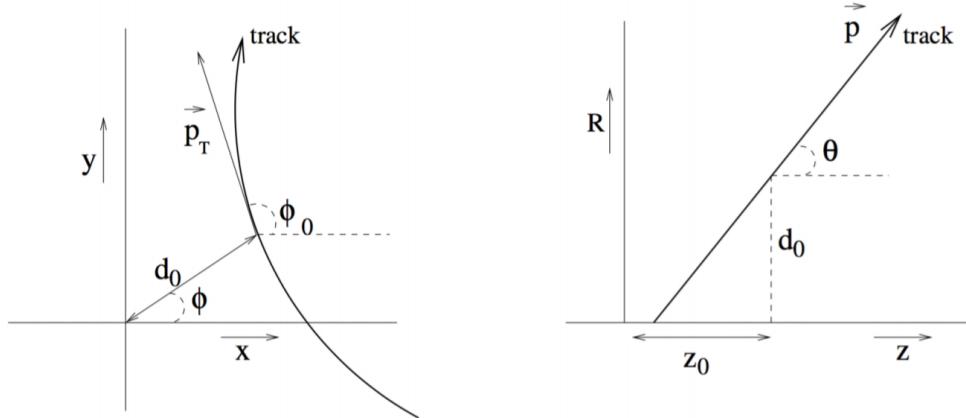


Fig. 4.5 Schematic of the impact parameters of a track in the transverse plane (left) and RZ-plane (right), as defined in the global ATLAS tracking frame^[55].

1442 The baseline ID algorithm introduced for ATLAS run-2 data analysis is the likelihood-
 1443 based (LH) method, which uses a MVA technique to simultaneously evaluate several
 1444 properties of electron candidates when making a decision. The LH method utilizes the
 1445 probability density functions (PDFs) of signal and background as the input discriminat-
 1446 ing variables. Based on these PDFs, it can calculate overall probabilities of the object to
 1447 be signal or background. Then the probabilities of signal and background are combined
 1448 together into a discriminant $d_{\mathcal{L}}$:

$$d_{\mathcal{L}} = \frac{\mathcal{L}_S}{\mathcal{L}_S + \mathcal{L}_B}, \quad \mathcal{L}_{S(B)}(\mathbf{x}) = \prod_{i=1}^n P_{s(b),i}(x_i) \quad (4.1)$$

1449 where \mathbf{x} denotes the vector of discriminating variables and $P_{s(b),i}(x_i)$ represents the value
 1450 of signal (background) PDF of the i^{th} variable as x_i .

1451 Three levels of working points (WPs) for electron ID are provided: *Loose*, *Medium*
 1452 and *Tight*, in order of increasing background rejection. Samples selected by a looser WP are
 1453 subsets of a tighter one, for example, the electrons passing Medium can all be selected by
 1454 Loose. The ID efficiency varies as function of electron energy (E_T) as shown in figure 4.6.
 1455 For evaluations, the electron candidates from MC simulation of $Z \rightarrow ee$ decays (di-jet)
 1456 are used as signal (background). Depending on the working point, the signal (background)
 1457 efficiencies for reconstructed electron candidates at $E_T = 25 GeV$ are in the range of 78
 1458 to 90% (0.3 to 0.8%), and increase (decrease) with E_T .

1459 Electron isolation

1460 In addition to the ID criteria, most analyses have electron isolation requirement to
 1461 further distinguish signal from background. To quantify the energy of particles around the
 1462 electron candidate, the isolation variables can help to separate the prompt electron from

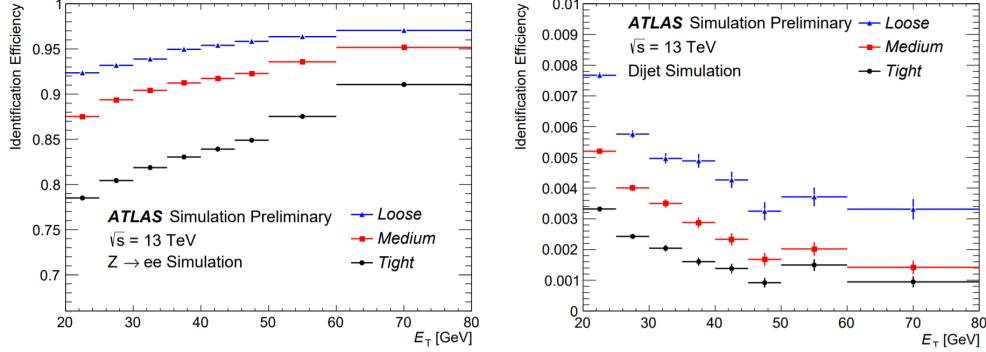


Fig. 4.6 The efficiencies of three electron ID WPs from $Z \rightarrow ee$ (left) events and hadrons misidentified as electrons estimated using di-jet MC samples (right).

other non-isolated electrons, like the electrons from converted photons or from heavy flavour hadron decays. There are two kinds of discriminating variables that have been designed:

- **Calorimeter-based variable:** $E_T^{topocone20}$. It's defined as the sum of transverse energies of topological clusters^[56], calibrated at EM scale within a cone of $\Delta R = 0.2$ around the candidate electron cluster. It only consider the clusters with positive reconstructed energy. In addition, a correction as a function of (E_T, η) values is applied to account for the electron energy leakage outside the cluster.
- **Track-based variable:** $p_T^{varcone20}$. It's calculated as the sum of transverse momentum of all satisfied tracks within a cone of $\Delta R = \min(0.2, 10\text{GeV}/E_T)$ around the candidate electron track. To calculate the sum, it requires the tracks are originating from the reconstruction PV of hard collision, and exclude the associated tracks of electron itself.

Based on the values of $E_T^{cone0.2}/p_T$ and $p_T^{varcone0.2}/p_T$, a series of working points with different selection requirements are defined. The resulting WPs are divided into two kinds:

- Efficiency targeted working points: varying requirements to obtain a certain isolation efficiency, which can either be a constant or as a function of E_T .
- Fixed requirement working points: set the constant upper thresholds on isolation variables.

The distribution of two discriminating variables are shown in figure 4.7 for $ZZ \rightarrow ee$ events with $E_T > 27\text{GeV}$ and satisfying *Tight* requirement.

4.2.4 Muon

Muons are distinctive signatures in final states of many physics analyses at the LHC including the Higgs analyses, SM measurements, BSM searches and so on. High perfor-

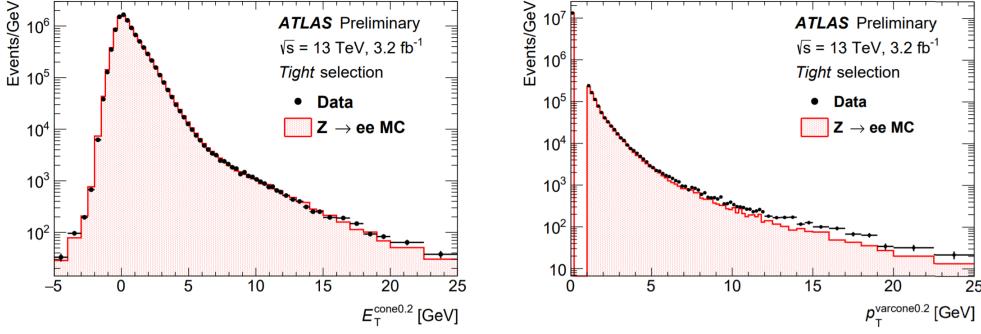


Fig. 4.7 Distributions of $E_T^{cone0.2}$ (left) and $p_T^{varcone0.2}$ (right) for electrons from $Z Z \rightarrow ee$ events in data and MC simulation. The simulated events (full histograms) are normalized to data.

1487 mance of muon reconstruction and identifications are crucial. This section briefly de-
 1488 scribes some more details of the *reconstruction, identification and isolation* of muon.

1489 Muon reconstruction

1490 Muon reconstruction is firstly performed in inner detector (ID) and muon spectrometer (MS)
 1491 independently as given in section 4.2.1. The information from each individual
 1492 detector is then combined together to form the muon tracks for physics analyses. The
 1493 combined ID-MS reconstruction is developed according to several algorithm based on
 1494 the information from ID, MS and calorimeters. Four different muon types are defined^[47]:

- 1495 • **Combined (CB) muons:** a combined track is formed by using the reconstructed
 1496 tracks performed independently in ID and MS with a global refit. To improve the fit
 1497 quality, the hits from MS may be added to or removed from the track. The outside-in
 1498 pattern recognition is utilized for the reconstruction of most muons, in which
 1499 the muons are first reconstructed in MS and then extrapolated inward to match the
 1500 ID track. In the meantime, the inside-out pattern is also used as a complementary
 1501 method.
- 1502 • **Segment-tagged (ST) muons:** a reconstructed track in ID is defined as muon, if
 1503 it can associated with at least one track segment in MDT or CSC chambers. These
 1504 ST muons are used when they can only pass across one layer of MS chambers due
 1505 to their low p_T or falling into regions with less MS acceptance.
- 1506 • **Calorimeter-tagged (CT) muons:** a reconstructed track in ID is categorized as
 1507 muon if it's matched to the energy deposit in calorimeter which is recognized with
 1508 a minimum-ionizing particle. This CT muons have lowest purity amount all types
 1509 of muons, but it covers the region where ATLAS muon spectrometer is only par-
 1510 tially constructed. For the region of $|\eta| < 0.1$ and $15 GeV < p_T < 100 GeV$, the
 1511 identification of CT muons are optimal.

- 1512 • **Extrapolated (ME) muons:** the muon is reconstructed based only on the MS track
1513 and a loose requirement of originating from the interaction point. In general, this
1514 type of muon needs to pass at least two (three) layers of MS chambers to provide
1515 a track measurement in barrel (forward) region. ME muons are designed to extend
1516 the acceptance for muon reconstruction into the region $2.5 < |\eta| < 2.7$ where ID
1517 doesn't cover.

1518 Before collecting those muons for physics analyses, overlap removals are performed
1519 between different muon types with the priority of CB > ST > CT, when two types of muons
1520 share the same ID track. Besides, the overlaps with ME muons are resolved by analyzing
1521 the track hit content, and selecting the track with better fit quality and larger number of
1522 hits.

1523 **Muon identification**

1524 After reconstruction, the muon identification is then performed to further discriminate
1525 between signal and background, especially to suppress backgrounds from pion and
1526 kaon decays by requiring prompt muons with high efficiency and guaranteeing a robust
1527 momentum measurement. The muon identification is defined by using the fit quality of
1528 combined track. The variables utilized in judgement for CB tracks include:

- 1529 • *q/p significance*, the absolute difference between q/g (charge over momentum) of
1530 muons measured in ID and MS divided quadratic sum of their corresponding un-
1531 certainties;
- 1532 • ρ' , the absolute value of difference between the p_T (transverse momentum) mea-
1533 sured in ID and MS, divided by the p_T of combined track;
- 1534 • *Normalized χ^2* of the combined track fit;
- 1535 • *Number of hits in ID and MS*

1536 In addition, some new variables used for *LowPt* muon working point what will be de-
1537 scribed later^[57]:

- 1538 • *Momentum balance significance (MBS)* is computed as momentum difference be-
1539 tween the ID and MS standalone measurements with respect to the uncertainty σ
1540 on energy lost in the calorimeter system.
- 1541 • *Scattering neighbor significance (SNS)* is defined to estimated the significance of a
1542 change in trajectory along the track, expected in the presence of a hadron decaying
1543 to a muon.
- 1544 • *Scattering curvature significance (SCS)* is defined as the normalized integral of the
1545 scattering angle significances, corrected for large kinks along the trajectory.

1546 Five selection working points are developed to satisfy the different needs for different

physics goals: *LowPt*, *Loose*, *Medium*, *Tight* and *HighPt*. The *Tight*, *Medium*, *Loose* are subsets from the tighter one to looser one. More detailed definition of each working point is given as follow:

- *Loose*: this working point is designed to maximize the reconstruction efficiency while keeping good-quality of muon tracks. And they are specifically developed for reconstructing the Higgs boson candidates from four-lepton final states. All four muon types are used for this selection level. The CB and ME muons passing Medium WP that will mentioned below are all included into Loose category. In addition, the CT and ST muons are restricted to $|\eta| < 0.1$ region. In the range of $|\eta| < 2.5$, around 97.5% Loose muons are CB muons, and about 1.5% are CT while remaining 1% are ST muons.
- *Medium*: this working point is the default criteria of muon identification in ATLAS. This selection minimizes the systematic uncertainties of muon reconstruction and calibration. In this category, we only use CB and ME muons. For CB muons, at least 3 hits in at least two layers of MDT are required, except $|\eta| < 0.1$ region, in which tracks with ≥ 1 MDT layer but ≤ 1 MDT hole layer are allowed. For ME muons, at least 3 MDT/CSC layers are required. Furthermore, a loose cut on the compatibility between measured momentum in ID and MS is applied to reduce the fake muons from hadrons misidentification. Besides, the q/p-significance is required to be less than 7.
- *Tight*: this working point is used to maximize the purity of muons but with sacrifice of some selection efficiency. Only CB muons with hits in ≥ 2 stations of MS and passing Medium criteria are selected. In addition, the normalized χ^2 of combined track fit should be smaller than 8. Then, a two-dimensional cut of q/p-significance and ρ' is adopted as a function of muon p_T to ensure tighter background rejection for momentum below 20 GeV, in which the fake rate is usually higher.
- *High- p_T* : this set of selections aims to maximize the momentum resolution for tracks with $p_T > 100\text{GeV}$ region. The selection is especially optimized for searching high-mass Z' and W' resonances. CB muons satisfying Medium selection and with ≥ 3 hits in 3 MS stations are chosen. The specific region in MS where alignment is suboptimal are removed as a precaution.
- *Low- p_T* : this type of muon is newly designed for physics analyses with ATLAS software release version 21. It's designed to obtain a optimal muon identification with very low transverse momentum of $3\text{GeV} < p_T < 5\text{GeV}$, which is crucial for B-physics measurement in ATLAS. In this muon requirement, only CB muons are

1582 used. In the range of $|\eta| < 1.3$, it requests muons hit at least one MS station; in
 1583 $1.3 < |\eta| < 1.55$, a least two MS stations are required; while in region of $|\eta| > 1.55$,
 1584 *Medium WP* is required. In addition, cuts are applied to suppress fakes as: $|\text{MBS}|$
 1585 < 3.0 , $|\text{SNS}| < 3.0$ and $|\text{SCS}| < 3.0$.

1586 Figure 4.8 and 4.9 show the selection efficiency of different muon identification work-
 1587 ing points. For *Medium (Loose)*, *Tight* and *High- p_T* : $Z \rightarrow \mu\mu$ events with $p_T > 10\text{GeV}$
 1588 are used for measurement. In the top figure of figure 4.8, the efficiency of the Loose selec-
 1589 tion (squares) is shown comparing to Medium one, where significant difference can be ob-
 1590 served in region of $|\eta| < 0.1$. For *LowPt*, $J/\Psi \rightarrow \mu\mu$ events with $3\text{GeV} < p_T < 10\text{GeV}$
 are used for measurement.

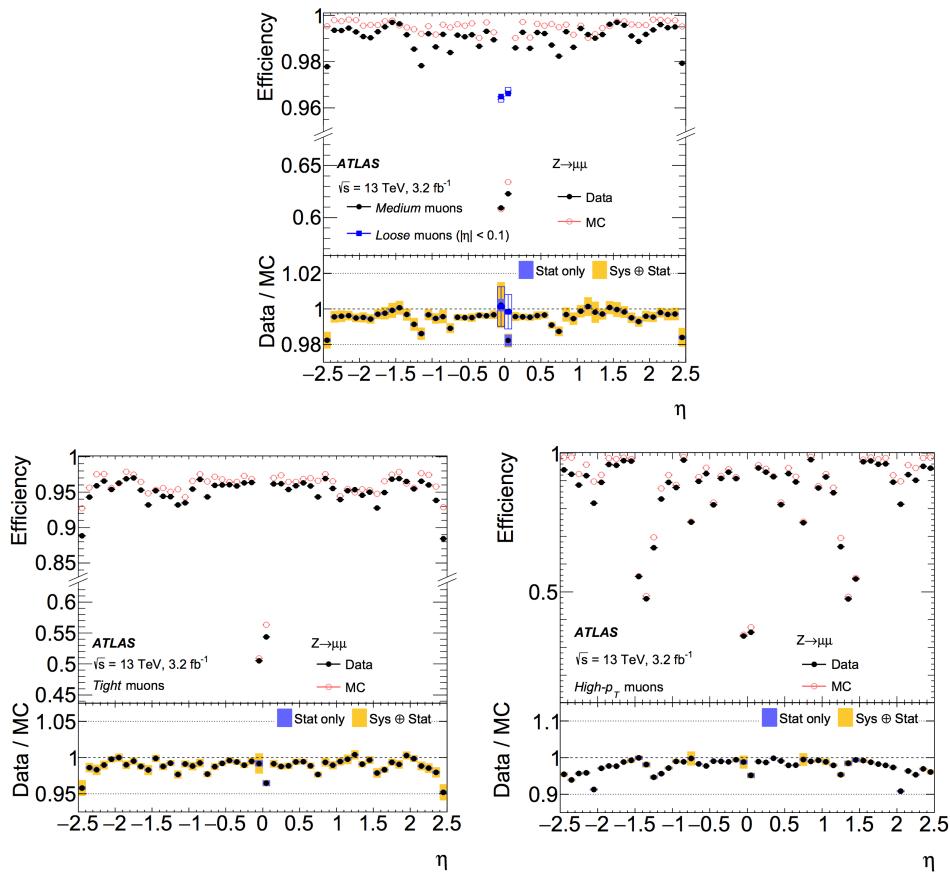


Fig. 4.8 Muon reconstruction efficiency as a function of η for: Medium (and Loose), Tight and High- p_T working points.

1591

1592 Muon isolation

1593 Similar as electron, the muon isolation is used to further distinguish the prompt muon
 1594 from non-prompt backgrounds. There are also two types of isolation variables for muon:

- 1595 • **Calorimeter-based variable:** $E_T^{\text{topocone}20}$. It's defined as the sum of the transverse
 1596 energy of topological clusters within a cone of size $\Delta R = 0.2$ around the candi-

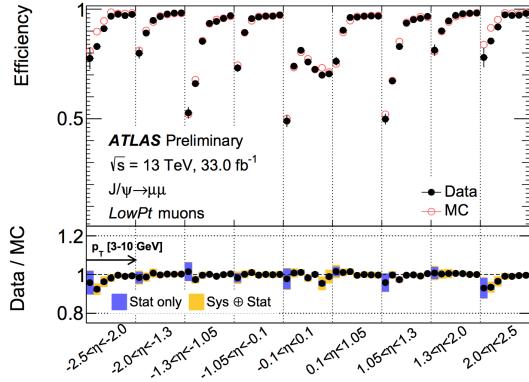


Fig. 4.9 Muon reconstruction efficiency for Low- p_T working point as a function of η .

date muon, after subtracting the contribution from the energy deposit of the muon itself and correcting for pile-up effects. The contributions from pile-up and underlying events are computed using the ambient energy-density technique^[58] and are corrected on an event-by-event basis.

- **Tracked-based variable:** $p_T^{varcone30}$. It's computed as the scalar sum of the transverse momenta of the tracks with $p_T > 1GeV$ in a cone size of $\Delta R = \min(10GeV/p_T^\mu, 0.3)$ around the candidate muon whose transverse momenta is p_T^μ after excluding the muon track itself. This p_T -dependent cone size can help to improve the performance for muons produced in the decay of particles with a large transverse momentum.

Then the isolation selections are applied based on *relative isolation variables*, which are computed as the ratio of the track- or calorimeter-based isolation variables to the transverse momentum of the muon. Figure 4.10 shows the distribution of those relative isolation variables by using $Z \rightarrow \mu\mu$ events and passing *Medium* identification criteria.

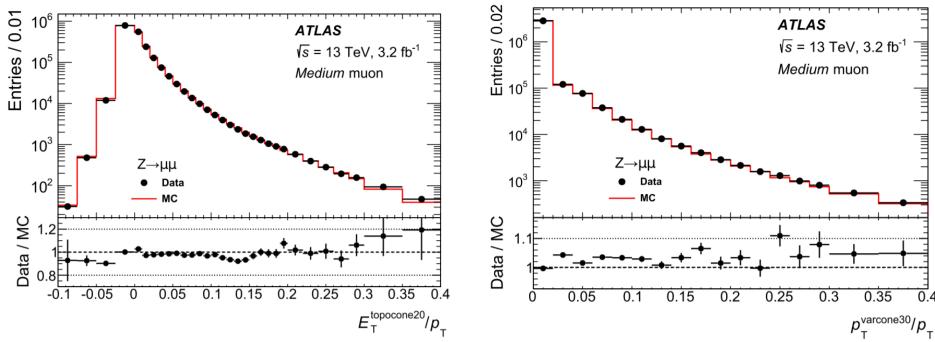


Fig. 4.10 Distributions of the calorimeter-based (right) and the track-based (left) relative isolation variables measured in $Z \rightarrow \mu\mu$ events.

4.2.5 Jets

Jets are another important features for many physics analyses at the LHC, and especially the key signatures for vector boson fusion/scattering (VBF/VBS) processes. In ATLAS detector, jets are reconstructed as groups of topologically associated energy deposits in the calorimeters, tracks associated with charged particles measured in the inner tracking detector, or simulated particles. This section introduces the jet reconstruction, jet energy scale (JES) calibration and the b-jet tagging technical.

Jet reconstruction

Jets are reconstructed using anti- k_t algorithm^[59] and with radius parameter of $R = 0.4$ in most cases. The FastJet software package^[60] is utilized for jet finding and reconstruction. A collection of four-vectors are used as inputs at each combination step in jet clustering, the total four-momentum is therefore computed as the sum of four-vector of all its constituents. There are three types of jets in ATLAS:

- *Truth jets*: the inputs to jet algorithm are simulated particles.
- *Track jets*: the inputs are charged tracks measured from inner detector.
- *Calorimeter jets*: the inputs are energy deposits in calorimeters.

Figure 4.11 shows the schematic of ATLAS jet reconstruction.

Jet reconstruction

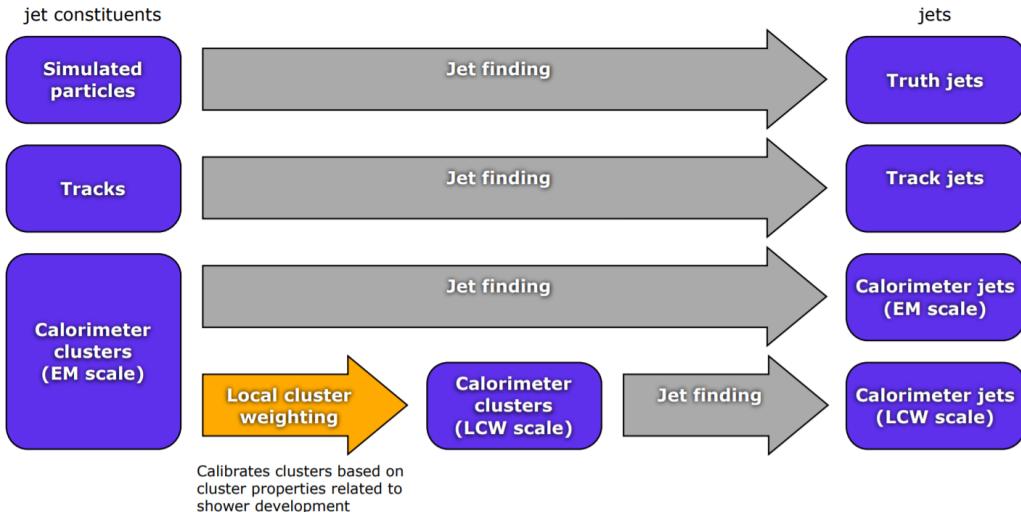


Fig. 4.11 A overview schematic of ATLAS jet reconstruction^[61].

The *truth jets* are reconstructed using anti- k_t algorithm with $R = 0.4$ by using final-state, stable particles from MC simulation as inputs. It requires the candidate particles with lifetime $c_\tau > 10$ mm and excludes the particles from pile-up. Truth jets with $p_T > 7\text{GeV}$ and $|\eta| < 4.5$ are then used for jets calibration described later.

The *track jets* are reconstructed from charged particles within the full acceptance of

inner detector ($|\eta| < 2.5$). The track reconstruction has been introduced in section 4.2.1.
 Reconstructed jets with $p_T > 500\text{MeV}$ and associated with primary vertex are then selected. Tracks are assigned to jets using ghost association^[58], a procedure that treats selected tracks as four-vectors of infinitesimal magnitude during the jet reconstruction and assigns them to the jet which they are clustered with. In addition, muon track segments are used as a compensation for those uncaptured jet energy carried by energetic particles passing through the calorimeters without being completely absorbed. Similar to the ID track, muon segments are assigned to jets using the method of ghost association mentioned above as well.

The *calorimeter jets* are reconstructed using a set of three-dimensional, positive-energy topological clusters (topo-clusters) made of calorimeter cell energies as input to the anti- k_t algorithm^[62]. Topo-clusters are built from near-by calorimeter cells that contains a significant energy above a noise threshold, which is estimated from measurements of calorimeter electronic noise and simulated pile-up noise. Those calorimeter cell energies are measured at electromagnetic energy scale (EM scale) corresponding to the energy deposited by electromagnetically interacting particles. And jets passing a p_T threshold of 7 GeV are reconstructed with the anti- k_t algorithm.

Jet energy scale calibration

Figure 4.12 depicts an overview of ATLAS jet calibration scheme for EM-scale calorimeter jets. In this procedure, the jet energies are scaled to truth jets, which is reconstructed at the particle-level. Each step of the calibration corrects the full four-momentum unless otherwise stated, scaling the jet p_T , energy, and mass.

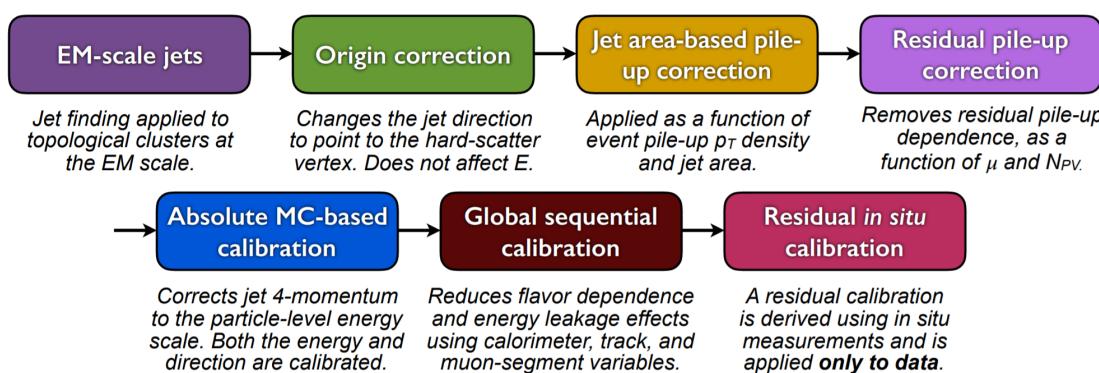


Fig. 4.12 A overview schematic of ATLAS jet calibration^[62].

First of all, the origin correction recompute the four-momentum of jets to point them to the hard-scatter primary vertex instead of the centre of detector, and in the meantime keep the jet energy unchanged. This correction improves the η resolution of jets for roughly 25% at a jet p_T of 20 GeV and > 5 times improvement for jet with p_T above 200 GeV,

as measured from the difference between reconstructed jets and truth jets in MC simulation. Secondly, the pile-up correction is adopted to remove the excess energy due to in-time and out-of-time pile-up, which consists of two processes: an area-based p_T density subtraction applied on the top of each event; and a residual correction derived from the simulation. Thirdly, the absolute JES calibration corrects the jet four-momentum to the particle-level energy scale, using truth jets in di-jet MC events. Furthermore, the step of global sequential calibration uses calorimeter, track and MS-based variables to reduce the flavor dependence and energy leakage effects. Finally, the residual in situ calibration is adopted to correct jets in data by using well-measured objects eg. photons, Z bosons and calibrated jets.

1669 **B-jet tagging**

Tagging of b-jets plays a important role in many physics analyses involving b- or t-quark. In the meantime, lots of analyses need to apply b-jet veto to suppress $t\bar{t}$ process. There are three major types of algorithms that have been developed to distinguish b-quark jets from light-quark (u,d,s) jets^[63]:

- **Impact parameter based algorithms (IP2D and IP3D):** b-hadrons usually have long lifetime (~ 1.5 ps, $c_\tau \sim 450 \mu\text{m}$), which leads to large impact parameter for tracks produced from b-hadron decay. The impact parameter taggers are developed based on these variables. The IP2D tagger makes use of the transverse impact parameter significance $d_0/\sigma(d_0)$ as discriminant, while IP3D tagger uses two-dimensional discriminant of both transverse and longitudinal impact parameter significances: $d_0/\sigma(d_0)$ and $z_0 \sin\theta/\sigma(z_0)$.
- **Secondary vertex finding algorithm (SV1)** makes use of the secondary vertex formed by decay products of b-hadron within the jet. All track pairs within a jet are tested for a two-track vertex hypothesis, and removed if they are likely to originate from a long-live particle decay (eg. K_s or Λ), hadronic interactions or photon conversions. After that, a new vertex is fitted with all tracks from remaining two-track vertices, and the outliers are removed from this set of tracks.
- **Decay chain multi-vertex algorithm (JetFitter)**^[64] exploits the topological structure of weak b- and c- hadron decays inside the jet and tries to reconstruct the full b-hadron decay chain. A Kalman filter is adopted to find a common line between primary vertex and b-/c- vertices, as well as their position in this line, which gives an approximated flight path for the b-hadron. In this approach, the b- and c-hadron vertices, whenever resolution allows, can be resolved, even when there is only a single track associated to them.

1694 The final discrimination commonly used in many physics analyses is called **Multivariate**
 1695 **Algorithm (MV2)**, which is based on Boosted Decision Tree (BDT) implemented in the
 1696 TMVA package^[65] by combining the outputs from underlying taggers mentioned above.
 1697 The MV2 was trained using jets in $t\bar{t}$ sample, where the b-jets are treated as signal while
 1698 the c- and light-flavor jets are treated as backgrounds. There are three kinds of MV2
 1699 depending on the fraction of c-jets in background for training: *MV2c00*, *MV2c10* and
MV2c20. Figure 4.13 presents the output score of MV2c10 for different flavor jets.

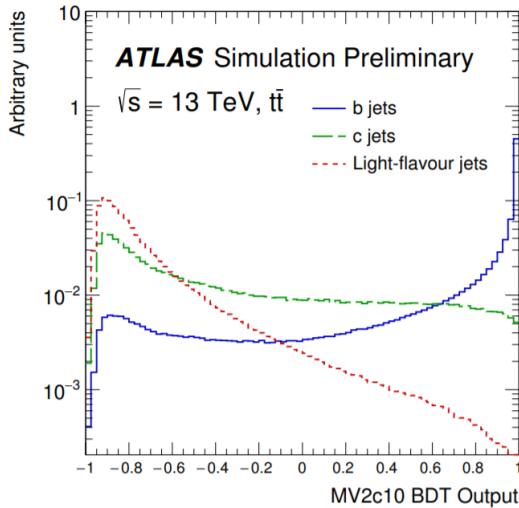


Fig. 4.13 MV2c10 BDT output for b- (solid blue), c- (dashed green) and light-flavour (dotted red) jets in $t\bar{t}$ events^[63].

1700

1701 4.2.6 Missing transverse energy

1702 Many interesting physics processes are with the involvement of neutrinos. Since they
 1703 do not interact with any materials in the detector, neutrinos cannot be detected directly;
 1704 but instead, they can result in imbalance in the plane transverse to the beam axis, where
 1705 momentum conservation is assumed. It is known as the missing transverse momentum
 1706 denoted as E_T^{miss} , which is obtained from the negative vector sum of the momenta of all
 1707 particles detected in a proton-proton collision event.

1708 The E_T^{miss} is measured using selected, reconstructed and calibrated hard objects in an
 1709 event. Its x- and y- components can be calculated as follow^[66]:

$$E_{x(y)}^{miss} = E_{x(y)}^{miss,e} + E_{x(y)}^{miss,\gamma} + E_{x(y)}^{miss,\tau} + E_{x(y)}^{miss,jets} + E_{x(y)}^{miss,\mu} + E_{x(y)}^{miss,soft} \quad (4.2)$$

1710 where each object term is given by the negative vectorial sum of the momenta of the
 1711 respective calibrated objects. The calorimeter signals are associated with the recon-
 1712 structed objects in the following order: electrons, photons, hadronically decaying taus,

¹⁷¹³ jets, muons. The soft term is reconstructed from detected objects not match any hard
¹⁷¹⁴ object passing the selections, but associated with the primary vertex. Details of applied
 selections for each term are summarized in table 4.1.

Table 4.1 Overview of the contributions to E_T^{miss} .

Objects contributing to E_T^{miss}				
Priority	Type	Selections	Variables	Comments
(1)	e	$ \eta < 1.37 \text{ or } 1.52 < \eta < 2.47$ $p_T > 10GeV$	$E_T^{miss,e}$	all e^\pm passing kinematic selections and medium reconstruction quality
(2)	γ	$ \eta < 1.37 \text{ or } 1.52 < \eta < 2.47$ $p_T > 25GeV$	$E_T^{miss,\gamma}$	all γ passing kinematic selections and tight reconstruction quality, and without overlapping with (1)
(3)	τ_{had}	$ \eta < 1.37 \text{ or } 1.52 < \eta < 2.47$ $p_T > 20GeV$	$E_T^{miss,\tau}$	all τ_{had} passing kinematic selections and medium reconstruction quality, and without overlapping with (1) and (2)
(4)	μ	$ \eta < 2.7$ $p_T > 10GeV$	$E_T^{miss,\mu}$	all μ passing kinematic selections and medium reconstruction quality
(5)	jet	$ \eta < 4.5$ $p_T > 60GeV$ --- or --- $2.4 < \eta < 4.5$ $20GeV < p_T < 60GeV$ --- or --- $ \eta < 2.4$ $20GeV < p_T < 60GeV$ $JVT > 0.59$	$E_T^{miss,jet}$	all jets passing kinematic selections and reconstruction quality (jet cleaning), and without overlap with (1)–(4)
(6)	ID track	$p_T > 400MeV$ $ d_0 < 1.5mm$ $ z_0 \sin\theta < 1.5mm$ $\Delta R(track, e/\gamma cluster) > 0.05$ $\Delta R(track, \tau_{had}) > 0.2$	$E_T^{miss,soft}$	all ID tracks from the hard-scattering vertex passing kinematic selections and reconstruction quality, and not associated with any particle from (1), (3) or (4), or associated with a jet from (5)

¹⁷¹⁵

1716 Based on $E_{x(y)}^{miss}$, the magnitude of E_T^{miss} and the azimuthal angle ϕ^{miss} are computed:

$$E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \quad (4.3)$$

$$\phi^{miss} = \arctan(E_y^{miss}/E_x^{miss})$$

1717 In equation 4.2, each objects are required to pass certain reconstruction and calibrated
1718 criteria and selections mentioned above before taken as inputs.

1719 In figure 4.14, left plot shows the observed E_T^{miss} distribution for data and MC of
1720 $Z \rightarrow \mu\mu$ events that without genuine missing transverse momentum; and right plot shows
1721 the E_T^{miss} distribution for $W \rightarrow e\nu$ events that has genuine (true) missing transverse mo-
mentum due to real neutrino.

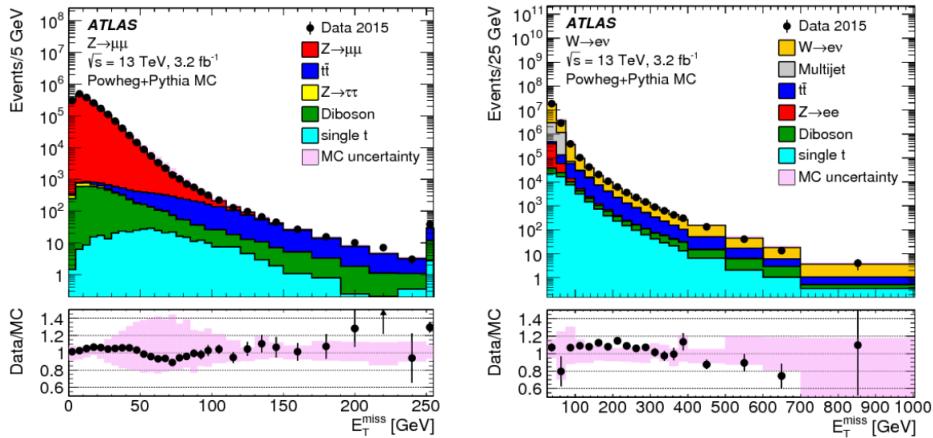


Fig. 4.14 Measured E_T^{miss} distribution for $Z \rightarrow \mu\mu$ events (left) and $W \rightarrow e\nu$ events (right).

1722

1723 **Chapter 5 Studies of SM ZZ production in $\ell\ell\ell\ell$ final** 1724 **state using pp collision data collected by ATLAS** 1725 **detector from 2015 to 2018**

1726 5.1 Introduction

1727 After the discovery of Higgs boson^[67-68], the examine of electroweak symmetry
1728 breaking (EWSB) becomes a main focus at the LHC. In addition to measuring the prop-
1729 erties of Higgs boson directly, the vector boson scattering (VBS) process is another key
1730 avenue to probe EWSB^[69-71]. As introduced in section 2.1.3, in Standard Model (SM),
1731 the Higgs boson acts as “moderator” to unitarize high-energy longitudinal VBS ampli-
1732 tudes at the TeV scale. Therefore, studying high-energy behaviours of VBS is crucial to
1733 understand the mechanism of EWSB.

1734 Since there was no VBS process was observed prior to the LHC era, LHC provides
1735 an unexceptionable opportunity to study them due to its unprecedented high energy and
1736 luminosity. At LHC, the VBS process is typically studied through the measurements of
1737 electroweak (EW) production of two vector bosons radiated from quark-quark initial state,
1738 plus a pair hadronic jets with high energy in the back and forward regions (denoted as EW-
1739 $VVjj$). The quantum chromodynamics (QCD) production of $VVjj$ contains two QCD
1740 vertices at the lowest order (denoted as QCD- $VVjj$) is an irreducible background to the
1741 search of EW- $VVjj$ production. The features of EW- $VVjj$ production include a large
1742 invariant mass of jet pair (m_{jj}) and a significant separation of rapidity between two jets
1743 (Δy_{jj}). Figure 5.1 presents some typical Feynman diagrams of EW- and QCD- $ZZjj$
1744 processes.

1745 The first evidence of the EW- $VVjj$ process was seen in same-sign WW channel
1746 (EW- $W^\pm W^\pm jj$) by ATLAS collaboration with 20.3 fb^{-1} 8 TeV data^[72], in which a 3.6σ
1747 excess was observed in data over the background-only prediction. In the LHC run-2, the
1748 observation (with $> 5 \sigma$ statistical significance) of EW- $W^\pm W^\pm jj$ process has been re-
1749 ported in both ATLAS and CMS collaboration with 36 fb^{-1} 13 TeV data^[73-74]. In WZ
1750 channel (EW- $WZjj$), an observation with 5.3σ excess was also reported by the AT-
1751 LAS collaboration recently^[75]. As for the EW- $ZZjj$ production, it was searched by
1752 CMS using 35.9 fb^{-1} 13 TeV data but no evidence was found^[76]. The EW production in
1753 ZZ final state (EW- $ZZjj$) is typically rare, whose fiducial cross section has an order of
1754 $O(0.1) \text{ fb}^{-1}$ in the final state where both Z bosons decay leptonically. But in the mean-

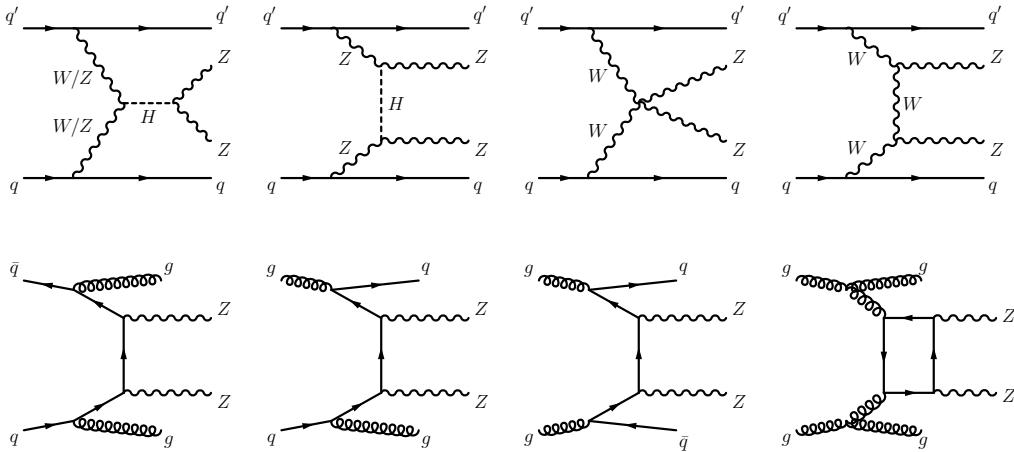


Fig. 5.1 Typical diagrams for the production of $Z Z j j$, including the relevant EW VBS diagrams (first row) and QCD diagrams (second row).

time, $Z Z \rightarrow \ell\ell\ell\ell$ process offers a extremely clean channel than all the others. So with more data collected in the LHC, the observation of EW- $Z Z j j$ becomes possible.

This section presents the first observation of EW- $Z Z j j$ production decaying to four charged leptons with two jets ($\ell\ell\ell\ell j j$) by ATLAS collaboration using the complete set of the LHC run-2 data with 139 fb^{-1} luminosity. It is a new milestone in the study of EWSB at the LHC, and completes the last missing part of observation of weak boson scattering for massive bosons. In the meantime, the measurement of fiducial cross-sections for SM $Z Z$ production including both EW and QCD processes is also reported. The $Z Z j j$ production involving intermediate τ -leptons from Z decays is considered as signal but has a negligible contribution to the selected events. Reducible backgrounds give minor contributions in the $\ell\ell\ell\ell j j$ channel are also studied. To further separate the EW signal and the QCD background, multivariate discriminant (MD) is trained using event kinematic information from simulated samples. The MD distribution is then used as discriminant in statistical fit to evaluate the signal strength of EW process.

5.2 Data and MC samples

5.2.1 Data samples

The data sets for this analysis include the full run-2 pp collision data collected by the ATLAS experiment during the years from 2015 to 2018. Data event is only used if it passed the latest Good Run List (GRL) released by the Data Quality group from ATLAS experiment as listed below, corresponding to an integrated luminosity of $139.0 \pm 2.4 \text{ fb}^{-1}$.
`data15_13TeV.periodAllYear_DetStatus-v89-pro21-02_Unknown_PHYS_StandardGRL_All_Good_25ns.xml`

1776 data16_13TeV.periodAllYear_DetStatus-v89-pro21-01_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns.xml
1777 data17_13TeV.periodAllYear_DetStatus-v99-pro22-01_Unknown_PHYS_StandardGRL_All_Good_25ns_Triggerno17e33prim.xml
1778 data18_13TeV.periodAllYear_DetStatus-v102-pro22-04_Unknown_PHYS_StandardGRL_All_Good_25ns_Triggerno17e33prim.xml

1779 **5.2.2 MC simulations**

1780 The EW- $ZZjj$ production (signal) is modelled using Mad-
1781 Graph5_aMC@NLO 2.6.1^[77] with the matrix elements (ME) calculated in the leading-
1782 order (LO) approximation in perturbative QCD (pQCD) and with the NNPDF2.3LO^[78]
1783 parton distribution functions (PDF). The VBF Higgs process is also included.

1784 The QCD- $ZZjj$ production is modelled using Sherpa 2.2.2^[79] with the
1785 NNPDF3.0NNLO^[80] PDF, where events with up to one (three) outgoing partons
1786 are generated at NLO (LO) in pQCD. The production of $ZZjj$ from the gluon-gluon
1787 initial state with a four-fermion loop or with an exchange of the Higgs boson has an
1788 order of α_S^4 in QCD, and is not included in the Sherpa simulation. A separate gg induced
1789 $ZZ + 2\text{jets}$ sample is modelled using Sherpa 2.2.2 with the NNPDF3.0NNLO PDF and
1790 with an additional 1.7 k-factor^[81] being applied. Then the interference between EW- and
1791 QCD- $ZZjj$ is modelled with MadGraph5_aMC@NLO 2.6.1 calculated at LO.

1792 The diboson productions from QCD $WW \rightarrow \ell\nu qq$ as well as QCD and EW $WZ \rightarrow$
1793 $\ell\ell qq$ are modelled using Sherpa 2.2.2 with the NNPDF3.0NNLO PDF. The productions
1794 of semileptonic decays ($WW \rightarrow \ell\nu qq$ and $WZ \rightarrow qql\ell$) are modelled using Powheg-
1795 Box v2^[82] with the CT10 PDF^[83]. The triboson production is modelled using Sherpa
1796 2.2.2 with the NNPDF3.0NNLO PDF.

1797 For top-quark pair ($t\bar{t}$) production, the Powheg-Box v2 is used with the CT10 PDF.
1798 The single top-quark production in t -channel, s -channel and Wt -channel are simulated
1799 using the Powheg-Box v1 event generator^[84-86]. The productions of $t\bar{t}$ in association with
1800 vector boson(s) ($t\bar{t}V$) are modelled with MadGraph5_aMC@NLO 2.3.3 for $t\bar{t}W$ and $t\bar{t}Z$
1801 with $Z \rightarrow \nu\nu/qq$ decays, with Sherpa 2.2.1 for $t\bar{t}Z$ where the Z decays to dilepton, and
1802 with MadGraph5_aMC@NLO 2.2.2 for $t\bar{t}WW$ respectively.

1803 The $Z+jets$ processes are modelled using Sherpa 2.2.1 with the NNPDF3.0NNLO
1804 PDF, in which the ME is calculated for up to two partons with next-to-leading-order
1805 (NLO) accuracy in pQCD and up to four partons with LO accuracy.

1806 For all the samples except those from Sherpa, the parton showering is modelled with
1807 Pythia8^[41] using the NNPDF2.3^[78] PDF set, and the A14 set of tuned parameters^[85].
1808 While for Sherpa samples, the parton showering is simulated within the programme.

1809 All simulated events are processed with detector response simulated based on Geant4

described in section 4.1. In addition, simulated inelastic pp collisions are overlaid to model additional pp collision in the same and neighbouring bunch crossings (pile-up), and reweighted to match the pile-up conditions in data. Moreover, all simulated events are processed using the same reconstruction algorithms as data. And the leptons and jets reconstruction, energy scale and resolution, and the leptons identification, isolation, trigger efficiencies for simulated events, as described in section 4.2, are all corrected to match the data measurements.

5.3 Objects and Event selection

5.3.1 Objects selection

The selection of analysis relies on the definition of multiple objects: *electrons*, *Muons*, and *jets*. Details of definition for each object are described as below:

Muon: To increase the acceptance range in reconstruction (reco) -level for $\ell\ell\ell\ell$ channel, all four types of muons (CB, ST, CT, ME muons, described in section 4.2.4) are used. The identified muons are then required to pass $p_T > 7$ GeV and $|\eta| < 2.7$, and satisfy the *Loose* identification criterion (see definition in sec 4.2.4). The impact parameter cuts are further applied to suppress the contribution from cosmic muons and non-prompt muons, with the value of: $|d_0/\sigma(d_0)| < 3.0$ and $|z_0 \sin\theta| < 0.5$ mm, where d_0 is the transverse impact parameter relative to the beam line, $\sigma(d_0)$ is its uncertainty, and z_0 is the longitudinal impact parameter relative to the primary vertex. In order to avoid muons associated with jets, all muons are required to be isolated and pass *FixedCutLoose* isolation criteria, which required $E_T^{\text{topocone}20}/p_T < 0.3$ and $p_T^{\text{varcone}30}/p_T < 0.15$.

Electron: As described in section 4.2.3, electrons are reconstructed from energy deposits in the EM calorimeter matched to a track in the inner detector. The electron candidates must satisfy the *Loose* criterion valuing by the likelihood-based (LH) method. And electrons are required to have $p_T > 7$ GeV and $|\eta| < 2.47$. Moreover, the impact parameter requirements of $|d_0/\sigma(d_0)| < 5.0$ and $|z_0 \sin\theta| < 0.5$ mm are applied. Same as muon, all electrons are required to satisfy *FixedCutLoose* isolation criteria of $E_T^{\text{topocone}20}/p_T < 0.2$ and $p_T^{\text{varcone}20}/p_T < 0.15$.

Jets: Jet are key signatures for VBS processes. This analysis use the jets clustered using the anti- k_t algorithm with radius parameter $R = 0.4$, more details of jets' reconstruction can be found in section 4.2.5. The jets are required to satisfy $p_T > 30$ (40) GeV in the $|\eta| < 2.4$ ($2.4 < |\eta| < 4.5$) region. To further reduce the effects of pile-up jets, a jet vertex tagger (JVT) is applied to jets with $p_T < 60$ GeV and $|\eta| < 2.4$ to select jets from

1843 hard-scattering vertex^[88].

1844 **Overlap removal:** An overlap-removal procedure is applied to selected leptons and
 1845 jets in this analysis. To enhance the selection efficiency, leptons are given higher priority
 1846 to be kept when overlapping with jets. With this lepton preferred method, the events of
 1847 EW signal after selection increases about 19% while background only increases about
 14%. More details of the strategy is summarized in table 5.1.

	Reference objects	Criteria
Remove electrons	electrons	Share a track or have overlapping calorimeter cluster. Keep higher p_T electron
Remove muons	electrons	Share track and muon is calo-tagged
Remove electrons	muons	Share track
Remove jets	electrons	$\Delta R_{e-jet} < 0.2$
	muons	$\Delta R_{\mu-jet} < 0.2$ OR muon track is ghost-associated to jet $\text{AND } (N_{Trk}(jet) < 3 \text{ OR } (p_T^{jet}/p_T^{\mu} < 2 \text{ and } p_T^{\mu}/\sum_{Trk} p_t > 0.7))$

Table 5.1 Overlap removal criteria between pre-selection objects for the $\ell\ell\ell\ell$ channel. The overlap removal follows the order shown in this table. Once an object has been marked as removed, it does not participate in the subsequent stages of the overlap removal procedure.

1848

1849 5.3.2 Event selection in reconstruction level

1850 The events are required to additionally be recorded by single or multi-lepton triggers,
 1851 with transverse momentum (p_T) thresholds varying from 8 to 26 GeV. The overall trigger
 1852 efficiency for selected inclusive $\ell\ell\ell\ell jj$ signal events in the analysis region are from 95
 1853 to 99%.

1854 The $\ell\ell\ell\ell$ quadruplets are formed by two opposite-sign, same-flavour (OSSF) lepton
 1855 pairs ($\ell^+\ell^-$), in which leptons are required to be separated by $\Delta R > 0.2$ in table 5.1.
 1856 At most one muon is allowed to be ME or CT muon. The p_T threshold of first three
 1857 leading leptons are 20, 20 and 10 GeV. If more than one quadruplets are found, the one
 1858 with minimum sum of difference between two dilepton pair masses and Z boson mass
 1859 ($|m_{l_1^+ l_1^-} - m_Z| + |m_{l_2^+ l_2^-} - m_Z|$) is selected. Both two dilepton pair masses are required to
 1860 be between 66 to 116 GeV. In addition, the invariant masses of all possible OSSF pairs
 1861 are required to be greater than 10 GeV to reject events from J/ψ or Υ decay.

1862 For VBS topology, the two most energetic jets in different detector side ($y_{j1} \times y_{j2} < 0$)

are selected. Furthermore, the invariant mass of two jets (m_{jj}) is required to be greater than 300 GeV, while Δy_{jj} is required to be larger than 2. Table 5.2 summarizes the above selection requirements, which is defined as signal region (SR).

Electrons	$p_T > 7 \text{ GeV}, \eta < 2.47$ $ d_0/\sigma_{d_0} < 5 \text{ and } z_0 \times \sin \theta < 0.5 \text{ mm}$
Muons	$p_T > 7 \text{ GeV}, \eta < 2.7$ $ d_0/\sigma_{d_0} < 3 \text{ and } z_0 \times \sin \theta < 0.5 \text{ mm}$
Jets	$p_T > 30 (40) \text{ GeV for } \eta < 2.4 (2.4 < \eta < 4.5)$
ZZ selection	$p_T > 20, 20, 10 \text{ GeV for the leading, sub-leading and third leptons}$ Two OSSF lepton pairs with smallest $ m_{\ell^+\ell^-} - m_Z + m_{\ell'^+\ell'^-} - m_Z $ $m_{\ell^+\ell^-} > 10 \text{ GeV for lepton pairs}$ $\Delta R(\ell, \ell') > 0.2$ $66 < m_{\ell^+\ell^-} < 116 \text{ GeV}$
Dijet selection	Two most energetic jets with $y_{j_1} \times y_{j_2} < 0$ $m_{jj} > 300 \text{ GeV and } \Delta y_{jj} > 2$

Table 5.2 Summary of selection of physics objects and candidate events at detector level in the $\ell\ell\ell\ell jj$ signal region.

5.4 Background estimation

Table 5.3 summarizes the background yields for $ZZjj \rightarrow \ell\ell\ell\ell jj$ process in 139 fb^{-1} . Uncertainties on the predictions include both statistical and systematic components. "Others" includes minor contributions from non- ZZ processes including $Z+jets$, top-quark, triboson and $t\bar{t}V$ processes. Details of estimation for each source are described as below.

Process	$\ell\ell\ell\ell jj$
EW- $ZZjj$	20.6 ± 2.5
QCD- $q\bar{q} \rightarrow ZZ$	77 ± 25
QCD- $gg \rightarrow ZZ$	13.1 ± 4.4
Others	3.2 ± 2.1
Total	114 ± 26
Data	127

Table 5.3 Observed data and expected signal and background yields in 139 fb^{-1} of luminosity. Minor backgrounds are summed together as 'Others'. Uncertainties on the predictions include both statistical and systematic components.

5.4.1 QCD backgrounds

The QCD- $Z Z jj$ production, which include both qq and gg initial processes, is the irreducible background in the search of EW- $Z Z jj$ production. A QCD-enriched control region, named as QCD CR, is defined to constrain the normalization of $Z Z$ background by reverting either the m_{jj} or Δy_{jj} requirements as:

$$m_{jj} < 300 \text{ GeV} \text{ or } \Delta y_{jj} < 2 \quad (5.1)$$

Then this normalization factor is included into statistical fit as a float parameter to properly treat the uncertainty correlations between SR and CR, while the shapes are taken from MC simulation. Table 5.4 shows the event yields of each background components in this CR. Uncertainties are statistical one only. The distributions of invariant mass of $\ell\ell\ell\ell$ and

Process	$\ell\ell\ell\ell jj$
EW- $Z Z jj$	3.9 ± 0
QCD- $Z Z jj$	136.9 ± 0.6
QCD- $ggZ Z jj$	16.8 ± 0.1
Diboson	0.3 ± 0.1
Triboson	1.6 ± 0.1
$Z+jets$	0
$t\bar{t}$	0
Total	159.5 ± 0.62
Data	152

Table 5.4 Observed data and expected signal and background yields in 139 fb^{-1} of luminosity. Diboson background in table includes all the other diboson processes discussed in section 5.2.2, except those with four-lepton final state. Uncertainties include only MC statistic. No events from $Z+jets$ and $t\bar{t}$ MC samples pass the selection, and are indicated as 0 in the table.

1880

dijet in QCD CR are shown in figure 5.2.

5.4.2 Reducible backgrounds

Backgrounds from $Z+jets$, top-quark and WZ processes called reducible backgrounds can be estimated by data-driven method. These events usually contain two or three leptons from Z/W decays, together with heavy-flavor jets or misidentified components of jets reconstructed as leptons called "fake leptons". A *fake factor* method is used to estimate these backgrounds, where the lepton misidentification is measured in data with

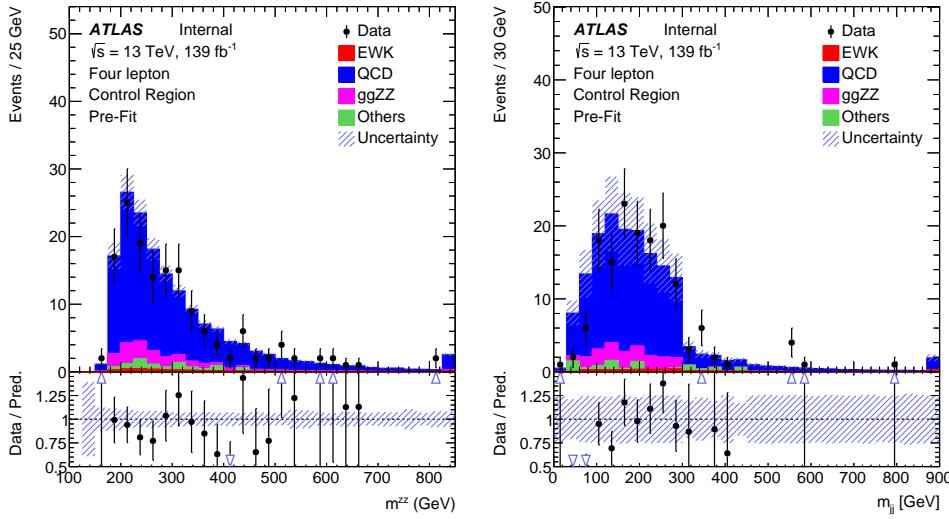


Fig. 5.2 Pre-fit m_{ZZ} and m_{jj} distribution in QCD-enriched CR.

the region enhances contributions from $Z+jets$ and top-quark processes. The method is described as below:

- Define a dedicated background dominant region to derive the fake factor for this background. The *fake factor* is defined as:

$$\mathcal{F} = \mathcal{N}_{good}/\mathcal{N}_{pool} \quad (5.2)$$

where \mathcal{N}_{good} refers to the number of good leptons passing all SR selection, while \mathcal{N}_{pool} denotes the number of poor leptons passing most SR selection but fail one certain requirement.

- Define a fake control region, where one or two lepton(s) pass *poor* requirement while all the other leptons are required to have SR selection.
- The number of fake events are calculated as:

$$\mathcal{N}_{fake} = (N_{gggp} - N_{ggpp}^{ZZ}) \times \mathcal{F} - (N_{gppp} - N_{ggpp}^{ZZ}) \times \mathcal{F}^2 \quad (5.3)$$

with the subtraction of ZZ contribution, and the double counting between N_{gggp} and N_{gppp} .

For the definition of *poor* leptons: The poor electrons are defined as failing “Fixed-CutLoose” isolation requirement, or failing “LooseLH” electron ID requirement but satisfying “VeryLooseLH” WP. The poor muons are required to fail the “FixedCutLoose” isolation requirement or invert the impact parameter cut to be $3 < d_0/\sigma(d_0) < 10$. The dedicated $Z+jets$ and $t\bar{t}$ dominant regions are defined to calculate the fake factor respectively in the following subsections.

1906 **1. Fake factor for $Z+jets$**

1907 Fake factor for $Z+jets$ background is calculated in $Z+jets$ enriched region, where
 1908 events with one SFOS lepton pair around Z mass associated with two jets are selected.
 1909 The value of fake factor is driven from data, and as a function of p_T and η as shown in
 1910 figure 5.3 for electrons and figure 5.4 for muons. During calculation, the contributions
 1911 from non- $Z+jets$ backgrounds ($t\bar{t}$, ZZ , WZ) have been subtracted from data. The values
 calculated directly from $Z+jets$ MC are also shown in plots for comparison.

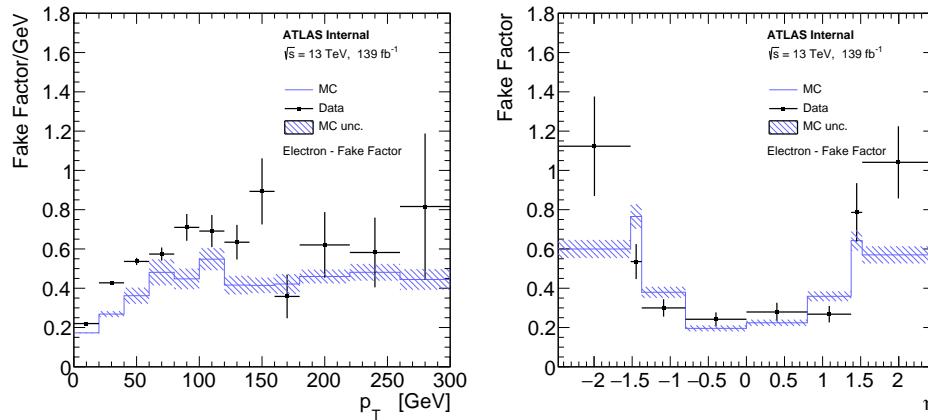


Fig. 5.3 Fake factor for $Z+jets$ background, constructed with additional electron, as a function of p_T (left) and η (right).

1912

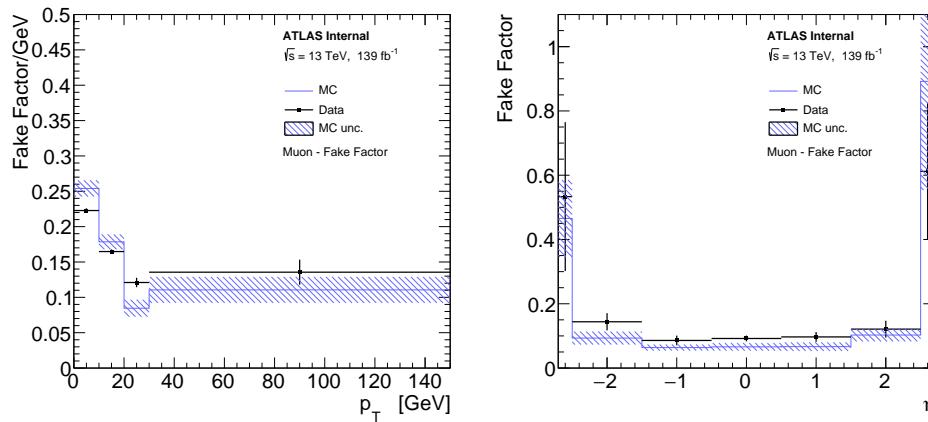


Fig. 5.4 Fake factor for $Z+jets$ background, constructed with additional muon, as a function of p_T (left) and η (right).

1913 **2. Fake factor for $t\bar{t}$**

1914 The fake factor for $t\bar{t}$ are calculated in $t\bar{t}$ dominant region by selecting one $e\mu$ -pair
 1915 with additional two jets. For events with three leptons, $m_T^W < 60 \text{ GeV}$ cut is applied to

1916 reject the contribution from $t\bar{t} + W$ events. The m_T^W is defined as below:

$$m_T^W = \sqrt{2p_T^{l_3}E_T^{miss} \left[1 - \cos(\Delta\phi(p_T^{l_3}, E_T^{miss})) \right]} \quad (5.4)$$

1917 In addition, at least one b-jet is required to enhance the top component. The fake factors of
 1918 $t\bar{t}$ calculated from data as the function of p_T and η are shown in figure 5.5 for electrons and
 1919 5.6 for muons. The non- $t\bar{t}$ contributions including $Z+jets$, ZZ and WZ , are subtracted
 from data.

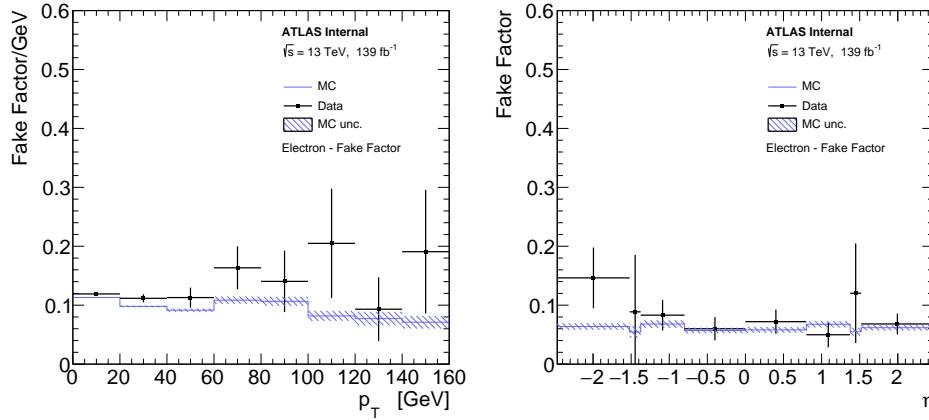


Fig. 5.5 Fake factor for $t\bar{t}$ background, constructed with additional electron, as a function of p_T (left) and η (right).

1920

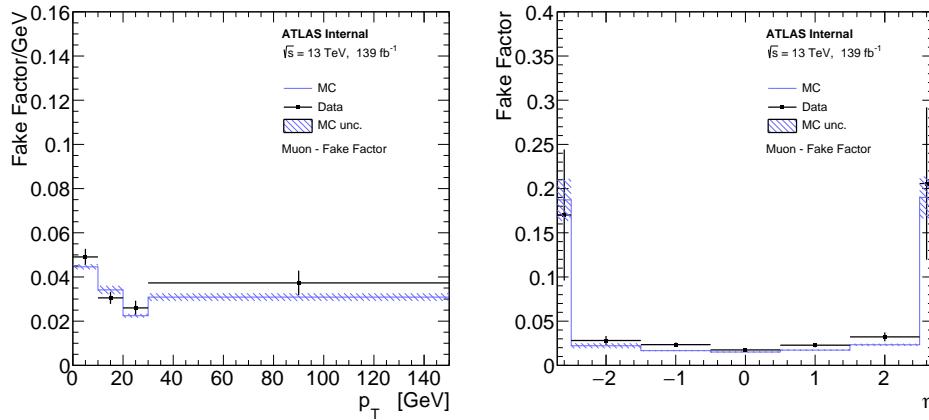


Fig. 5.6 Fake factor for $t\bar{t}$ background, constructed with additional muon, as a function of p_T (left) and η (right).

1921

3. Combination

1922 The fake factors calculated from each dedicated region are then combined together
 1923 according to their contributions in fake control region described previously. Figure 5.7
 1924 shows the m_{jj} distribution with data and major fake backgrounds in three different 41
 1925 channels.

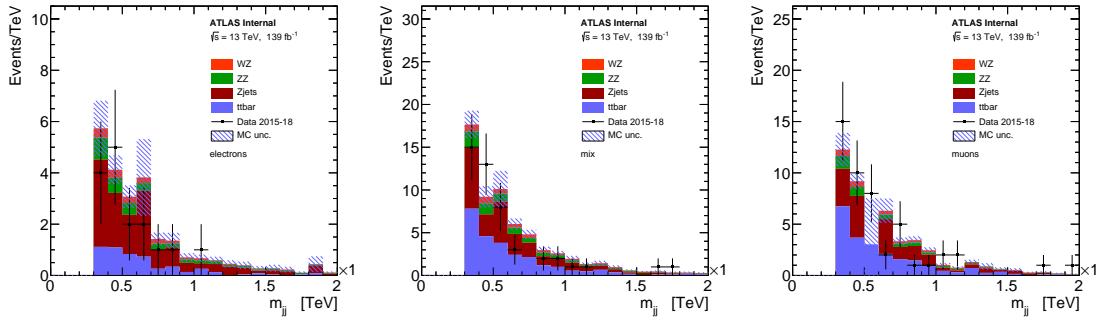


Fig. 5.7 m_{jj} distributions in fake control region in 4e (left), 2e2 μ (middle) and 4 μ (right) channel. The ratios between $Z+jets$ and $t\bar{t}$ ($Z+jets/t\bar{t}$) in each individual channel are: 2.59, 0.95, 0.74.

4. Systematics of fake estimation and results

The systematics of fake factor method can be measured by varying the parameters and selection requirements in fake factor calculation. In addition, due to the very limited data statistic in $\ell\ell\ell\ell$ channel, to be more conservative, the difference between data measurement and MC simulation are also considered as another systematics component. The sources of systematics that have been included are listed as below:

- Variations of isolation cut for the poor lepton definition up and down scaled by a factor of two.
- Variations of the yields of those subtracted MC in fake control region scaled by 30% up and down.
- The difference of fake factors between driven from data and from MC simulation.
- The difference of fake factors when changing to one bin measurement (instead of p_T or η dependent).
- The statistical uncertainties on fake factor in fake control region.

Table 5.5 summarizes the contribution of fake backgrounds in signal region under different systematic conditions mentioned above as well as the nominal one, together with their statistical uncertainties.

5.5 Systematics

The analysis includes both the statistical fit to MD distribution to search the EW- $Z Z jj$ process, as well as the cross section measurement of inclusive EW and QCD $Z Z jj$ process in fiducial volume. Therefore, theoretical and experimental uncertainties may affect the predictions background yields and shapes, correction factors from detector-level to particle-level measurement, as well as the $Z Z jj$ MD shapes and so on. Moreover,

channel	4e	2e2 μ	4 μ	inclusive
Nominal estimate	0.678 ± 0.652	1.023 ± 0.740	0.566 ± 0.240	2.268 ± 1.015
F stat. uncertainty varied down	0.698 ± 0.622	0.872 ± 0.652	0.509 ± 0.214	2.079 ± 0.926
F stat. uncertainty varied up	0.657 ± 0.685	1.173 ± 0.840	0.622 ± 0.267	2.452 ± 1.116
One bin F	0.653 ± 0.590	0.594 ± 0.558	0.646 ± 0.313	1.892 ± 0.870
MC F	0.534 ± 0.471	1.415 ± 0.993	0.439 ± 0.184	2.389 ± 1.114
Isolation varied down	0.938 ± 0.686	0.552 ± 0.466	0.215 ± 0.107	1.704 ± 0.837
Isolation varied up	0.723 ± 0.646	1.104 ± 0.739	0.559 ± 0.237	2.386 ± 1.010
MC corr. varied down	0.697 ± 0.695	1.048 ± 0.811	0.832 ± 0.385	2.577 ± 1.136
MC corr. varied up	0.660 ± 0.614	0.984 ± 0.687	0.316 ± 0.159	1.961 ± 0.935

Table 5.5 Fake background estimations in the SR. For the nominal value the 2D fake factor together with the $Z+jets$ and $t\bar{t}$ combination applied. The other lines show the estimations with different uncertainty variations.

the statistical uncertainties of simulated samples are also taken into account. Due to the extremely low cross section of $\ell\ell\ell\ell$ channel, the analysis is still data statistic dominant. This section describes the measurement of both theoretical and experimental systematics for $Z Z jj$ productions. The systematics for fake backgrounds have been elaborated in section 4.

5.5.1 Theoretical systematics

The theoretical systematics on EW- and QCD- $Z Z jj$ processes include the uncertainties from PDF, QCD scale, α_S and parton showering variations as summarized in table 5.6. The PDF uncertainty is estimated from the envelop of NNPDF internal variations and the difference between nominal and alternative PDF sets, following the PDF4LHC as introduced in section 2.2.1. The QCD scale uncertainty is estimated by varying the nominal renormalization scale (μ_R) and factorisation scale (μ_F) by a factor of 0.5 or 2.0. There are seven different configurations being considered, where the maximum of variations is chosen as final uncertainty. The parton showering uncertainty is estimated by comparing events with different parton showering setting between the nominal Pythia8 and the alternative Herwig7^[89-90] algorithm. The α_S uncertainty is estimated by varying the value of α_S within ± 0.001 . Due to the lack of simulation sample for alternative parton showering on QCD- $Z Z jj$ process, the value of parton showering component is taken from the measurement of EW process.

Table 5.7 summarizes the uncertainties of each theoretical components in fiducial volume of SR, while table 5.8 shows the numbers in QCD-enriched CR region. For QCD process, the uncertainty is QCD scale dominant. Both of them are taken as inputs for

Process	EW- $Z Z jj$	QCD- $Z Z jj$
PDFs	NNPDF30lo (nominal), CT14lo	NNPDF30nnlo (nominal), MMHT2014nnlo68cl, CT14nnlo
α_S	0.118	0.117, 0.118 (nominal), 0.119
QCD scale ($[\mu_R, \mu_F]$)	[0.5,0.5], [0.5,1], [1,0.5], [1,1], [1,2], [2,1], [2,2]	[0.5,0.5], [0.5,1], [1,0.5], [1,1], [1,2], [2,1], [2,2]
Parton showering algorithm	Pythia8, Herwig7	-

Table 5.6 Summary of different variations for EW- and QCD- $Z Z jj$ theoretical uncertainties measurement.

statistical fit.

Process	PDF (%)	α_S (%)	QCD scale (%)	Parton shower (%)
EW	+5.9 -5.9		+6.1 -5.6	+3.3 -3.3
qqQCD	+2.0 -1.0	+2.6 -2.6	+34.2 -22.8	

Table 5.7 Summary of theoretical uncertainties for the fiducial volume (SR) for both EW and QCD $q q$ -initial processes.

1971

Process	PDF (%)	α_S (%)	QCD scale (%)	Parton shower (%)
EW $\ell\ell\ell\ell$	+6.1 -6.1		+0.8 -1.1	+10.1 -10.1
qqQCD $\ell\ell\ell\ell$	+2.0 -1.0	+2.6 -2.6	+31.5 -22.0	

Table 5.8 Summary of theoretical uncertainties for the control region for EW and qqQCD processes.

1972

The uncertainties of QCD gg -induced process ($gg \rightarrow ZZ$) as the function of MD discriminant is shown in figure 5.8 for both fiducial volume (SR) and QCD CR.

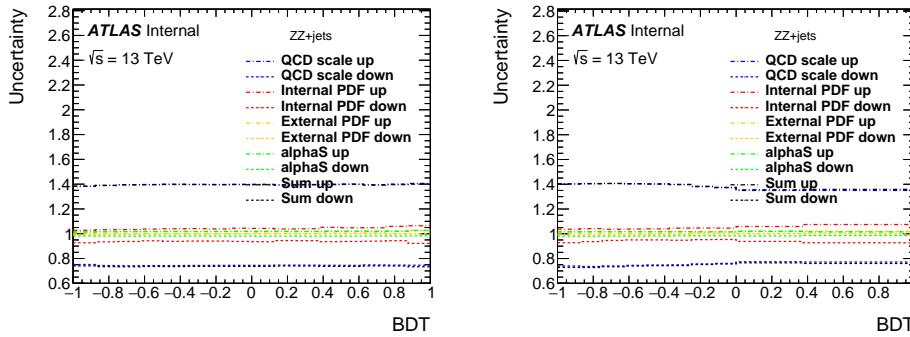


Fig. 5.8 The theoretical uncertainties for $gg \rightarrow ZZ$ background in particle-level SR (left) and CR (right).

1973

1974 **5.5.2 Experimental systematics**

1975 The dominant experimental uncertainties are from the luminosity uncertainty, the mo-
1976 mentum scale and resolution of leptons and jets, as well as the lepton reconstruction and
1977 selection efficiency. Some smaller uncertainties, such as trigger efficiency and pile-up
1978 correction, are also considered. Table 5.9 lists the major systematic components from
1979 leptons and jets for signal and major background processes in $\ell\ell\ell\ell$ channel. The total
1980 uncertainties for sources from electron, muon and jet respectively, as well as the sum
1981 (quadratic sum) of them are also summarized in this table.

name	EW- $ZZjj$	QCD qq -initial	QCD gg
nominal yield	20.61	76.69	13.10
EG_RESOLUTION_ALL	$\pm^{0.00\%}_{0.03\%}$	$\pm^{0.02\%}_{0.04\%}$	$\pm^{0.01\%}_{1.41\%}$
EG_SCALE_ALL	$\pm^{0.03\%}_{0.05\%}$	-0.04%	$\pm^{0.01\%}_{0.06\%}$
EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	$\pm^{2.66\%}_{2.58\%}$	$\pm^{2.60\%}_{2.53\%}$	$\pm^{2.65\%}_{2.57\%}$
EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	$\pm 0.70\%$	$\pm 0.47\%$	$\pm 0.42\%$
EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	$\pm 0.55\%$	$\pm 0.55\%$	$\pm 0.63\%$
JET_EtaIntercalibration_NonClosure	-0.01%	-0.03%	0%
JET_GroupedNP_1	$\pm 1.97\%$	$\pm^{11.82\%}_{10.14\%}$	$\pm^{16.21\%}_{12.92\%}$
JET_GroupedNP_2	$\pm 0.23\%$	$\pm 1.26\%$	+5.3%
JET_GroupedNP_3	$\pm 0.55\%$	$\pm 2.94\%$	$\pm^{3.14\%}_{0.12\%}$
JET_JER_SINGLE_NP	0.11%	+5.47%	+6.31%
JET_JvtEfficiency	$\pm 0.04\%$	$\pm 0.12\%$	$\pm 0.15\%$
MUON_EFF_ISO_STAT	$\pm 0.09\%$	$\pm 0.08\%$	$\pm 0.07\%$
MUON_EFF_ISO_SYS	$\pm 0.54\%$	$\pm 0.55\%$	$\pm 0.56\%$
MUON_EFF_RECO_STAT	$\pm 0.15\%$	$\pm 0.19\%$	$\pm 0.15\%$
MUON_EFF_RECO_STAT_LOWPT	$\pm 0.06\%$	$\pm 0.02\%$	$\pm 0.03\%$
MUON_EFF_TTVA_STAT	$\pm 0.06\%$	$\pm 0.07\%$	$\pm 0.06\%$
MUON_EFF_TTVA_SYS	$\pm 0.03\%$	$\pm 0.4\%$	$\pm 0.03\%$
MUON_ID	$\pm 0.03\%$	$\pm 0.02\%$	<0.001%
MUON_MS	-0.05%	$\pm^{0.04\%}_{0.01\%}$	<0.001%
MUON_SAGITTA_RESBIAS	$\pm 0.01\%$	$\pm 0.02\%$	<0.001%
MUON_SAGITTA_RHO	+1.13%	-0.73%	$\pm 1.00\%$
MUON_SCALE	$\pm 0.02\%$	$\pm^{0.03\%}_{0.02\%}$	<0.001%
PRW_DATASF	$\pm 0.5\%$	$\pm^{0.42\%}_{1.02\%}$	$\pm^{2.17\%}_{1.46\%}$
Electron Exp	$\pm^{2.8\%}_{2.7\%}$	$\pm^{2.70\%}_{2.62\%}$	$\pm^{2.75\%}_{2.64\%}$
Muon Exp	$\pm 1.3\%$	$\pm 1.3\%$	$\pm 1.04\%$
Jet Exp	$\pm 2.0\%$	$\pm^{13.39\%}_{10.64\%}$	$\pm^{18.54\%}_{13.57\%}$
Total experimental uncertainties	$\pm^{3.7\%}_{4.0\%}$	$\pm^{13.72\%}_{11.11\%}$	$\pm^{18.90\%}_{13.57\%}$

Table 5.9 Experimental uncertainties in $\ell\ell\ell\ell$ channel with the luminosity of 139 fb^{-1} . The "Electron Exp", "Muon Exp" and "Jet Exp" represent the quadrature of the respective sources from electron, muon, and jets.

In addition, the uncertainty of the combined 2015 to 2018 integrated luminosity is 1.7%^[91] in ATLAS experiment, obtained using the LUCID-2 detector^[92] for the primary luminosity measurements.

On top of them, a systematic uncertainty for MD distribution with different pile-up ($\langle\mu\rangle$) is also considered for QCD- $ZZjj$ background by comparing the distributions be-

tween events with low and high pile-up conditions. A boundary of $\langle\mu\rangle = 33$ is used to defined low/high pile-up according to the average $\langle\mu\rangle$ for signal (about 34.5) and QCD background (about 33). Figure 5.9 shows the MD distribution in SR (left) and QCD CR (right) in two different PU conditions, the difference as function of MD is then taken into account as additional shape uncertainty for statistical fit.

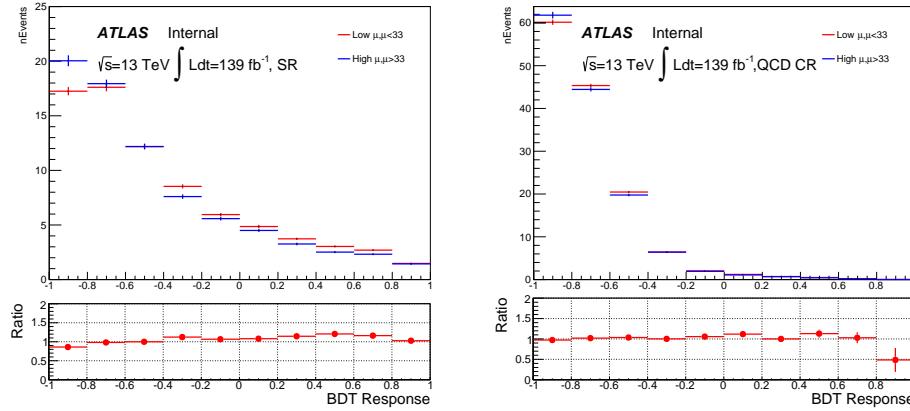


Fig. 5.9 MD distribution for QCD- $Z Z jj$ process in low and high pile-up events for SR (left) and CR (right).

Moreover, a conservative uncertainty is signed to QCD- $Z Z jj$ process by comparing the sample modelled by Sherpa generator (nominal) with MadGraph5_aMC@NLO. The MD shape difference for both SR (left) and QCD CR (right) are shown in figure 5.10. The modelling uncertainty is then calculated from the envelop between nominal and alternative samples as function of MD as one additional shape uncertainty.

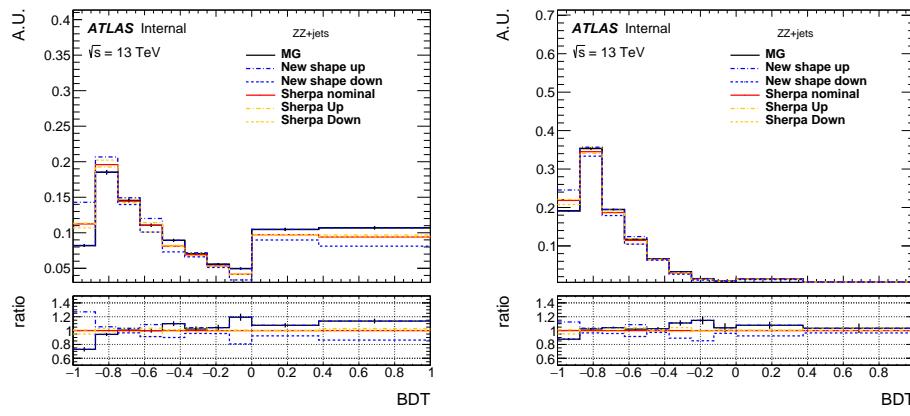


Fig. 5.10 MD shape difference for QCD $q\bar{q} \rightarrow ZZ$ background between different Sherpa theoretical uncertainties and sample from MadGraph5_aMC@NLO on SR (left) and CR (right).

5.6 Measurement of fiducial cross section

The fiducial cross section for inclusive $Z Z j j$ production, which includes both EW and QCD components, is then measured. The definition of fiducial volume, which is used for cross section measurement, follows closely to the detector-level selection but use physics objects in particle-level, which are reconstructed in simulation from stable final-state particles, prior to their interactions with the detector.

For electrons and muons, QED final-state radiation is for the most part recovered by adding the four-momenta of surrounding photons that are not originating from hadrons and within an angular distance $\Delta R < 0.1$ to the lepton four-momentum, called lepton “dressing” in truth level. Particle-level jets are built with anti- k_T algorithm with radius parameter $R = 0.4$ using all final-state particles except leptons and neutrinos as inputs. Comparing to the events selection in detector-level in section 5.3, in particle-level, the selected dilepton pair mass required is relaxed to be within 60 to 120 GeV for the reasons of reducing the migration effect, as well as being more compatibility with CMS publication^[76]. All other kinematic selections are the same as the definition in detector-level.

5.6.1 Calculation of C-factor

C-factor is defined as the ratio between the number of selected events in detector-level and the number of particle-level events in fiducial volume (FV):

$$C = \frac{N_{\text{detector-level}}}{N_{\text{FV}}} \quad (5.5)$$

The value of C-factor for each $Z Z j j$ process are calculated from each individual simulation samples as listed in table 5.10 together with their systematics.

Process	C	ΔC (stats.)	ΔC (sys.)	ΔC (theo.)
EWK $Z Z j j$	0.663	± 0.002	$\pm^{0.032}_{0.031}$	NA
QCD $q\bar{q} \rightarrow Z Z$	0.702	± 0.003	$\pm^{0.061}_{0.051}$	$\pm^{0.015}_{0.018}$
QCD $gg \rightarrow Z Z$	0.741	± 0.021	$\pm^{0.143}_{0.072}$	± 0.002

Table 5.10 C Factor of different $Z Z j j$ processes.

Then the C from different processes are combined together to be used as inputs for cross section calculation:

$$C = \sum_i \frac{N_{\text{FV}}^i}{\sum_j N_{\text{FV}}^j} \times C_i = 0.699 \pm 0.003(\text{stats.}) \pm^{0.011}_{0.013} (\text{theo.}) \pm 0.028(\text{exp.}) \quad (5.6)$$

The stats. refers to the statistical uncertainty from MC simulation statistics. The theo.

2020 and exp. denote the theoretical and experimental uncertainties described in section 5.5.

2021 **5.6.2 Result of fiducial cross section**

2022 The cross section in fiducial volume is computed as:

$$\sigma^{FV.} = \frac{N_{data} - N_{bkg}}{C \times Lumi} \quad (5.7)$$

2023 where N_{data} and N_{bkg} denote the number of events selected from detector-level selection
2024 from data and sum of backgrounds, and C is the C-factor calculated above, Lumi repre-
2025 sents the integrated luminosity of data from 2015 to 2018 of 139 fb^{-1} . Table 5.11 shows
2026 the fiducial cross section for $\ell\ell\ell\ell$ final state measured from equation 5.7, as well as the
2027 predicted cross section measured from MC simulation directly.

Measured fiducial σ [fb]	Predicted fiducial σ [fb]
$1.27 \pm 0.12(\text{stat}) \pm 0.02(\text{theo}) \pm 0.07(\text{exp}) \pm 0.01(\text{bkg}) \pm 0.03(\text{lumi})$	$1.14 \pm 0.04(\text{stat}) \pm 0.20(\text{theo})$

Table 5.11 Measured and predicted fiducial cross-sections in $\ell\ell\ell\ell jj$ final-state. Uncertainties due to different sources are presented.

2028 The measured cross section has a total uncertainty of 11%, and is found to be compat-
2029 ible with SM prediction. This measurement is still dominant by data statistic.

2030 **5.7 Search for EW- $ZZjj$**

2031 **5.7.1 MD discriminant**

2032 To further separate the EW- $ZZjj$ component from QCD- $ZZjj$, a MD based on Gra-
2033 dient Boosted Decision Tree (BDTG) algorithm^[93] is trained with simulated events via
2034 TMVA framework^[65]. For $\ell\ell\ell\ell$ channel, training is performed between EW (signal) and
2035 QCD (background) processes. Twelve event kinematic variables sensitive to the charac-
2036 teristics of the EW signal is used as input features in training. Table 5.12 lists those input
2037 variables with the order of their ranking provided by TMVA tool. The jet-related infor-
2038 mation provides larger sensitivity. Then the MD distributions in both SR and QCD CR
2039 region are used for statistical fit.

2040 **5.7.2 Profile likelihood ratio method**

2041 To examine the compatibility between data and the signal-plus-background hypoth-
2042 esis, a test statistic is driven by using the profile likelihood ratio method. The binned

Rank	Variables	Description
1	m_{jj}	Dijet invariant mass
2	p_T^{j1}	p_T of the leading jet
3	p_T^{j2}	p_T of the sub-leading jet
4	$\frac{p_T(ZZjj)}{H_T(ZZjj)}$	p_T of the $ZZjj$ system divided by the scalar p_T sum of Z bosons and two jets
5	$y_{j1} \times y_{j2}$	Product of jet rapidities
6	Δy_{jj}	Rapidity difference between two jets
7	Y_{Z2}^*	Rapidity of the second Z boson
8	Y_{Z1}^*	Rapidity of the Z boson reconstructed from the lepton pair with the mass closer to the Z boson mass
9	p_T^{ZZ}	p_T of 4l system
10	m_{ZZ}	Invariant mass of 4l system
11	p_T^{Z1}	p_T of the Z boson reconstructed from the lepton pair with the mass closer to the Z boson mass
12	$p_T^{\ell3}$	p_T of the third lepton

Table 5.12 Input features for the training of MD.

2043 likelihood function is given as

$$\mathcal{L}(\mu, \sigma) = \prod_i^{\text{bins}} \mathcal{L}_{\text{poiss}}(N_{\text{data}} | \mu s(\theta) + b(\theta))_i \times \mathcal{L}_{\text{gauss}}(\theta)_i \quad (5.8)$$

2044 where the Poisson term presents the statistical fluctuations of the data and a Gaussian
 2045 term models the pdf of auxiliary measurement to constrain the systematics. μ denotes the
 2046 signal strength of EW- $ZZjj$ process, computed as the ratio between measured (expected)
 2047 cross section to the SM prediction. θ presents the nuisance parameter, which is the set of
 2048 parameters that parameterize the effect of systematic uncertainties described in section 5.5
 2049 following the Gaussian distribution. N_{data} is the number of selected data events, while the
 2050 $s(\theta)$ is the expected signal yield and $b(\theta)$ is the expected background yield as the function

2051 of nuisance parameters.

2052 The test statistic q_μ is defined as:

$$q_\mu = -2 \ln \left(\frac{\mathcal{L}(\mu, \hat{\theta}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \right) \quad (5.9)$$

2053 in which $\mathcal{L}(\hat{\mu}, \hat{\theta})$ is the unconditional likelihood with respect to both μ and θ , and $\mathcal{L}(\mu, \hat{\theta}_\mu)$
 2054 is the conditional likelihood for a constant μ . Signal-like data distributions are more likely
 2055 to have a low test-statistic (q_μ close to 0) while the contributions of background-like data
 2056 have a larger q_μ . Under the background-only hypothesis, the compatibility of the ob-
 2057 served (Asimov) data with the prediction is calculated to obtain the observed (expected)
 2058 significance respectively.

2059 **5.7.3 Fitting procedure**

2060 A profile likelihood fit is performed on MD discriminant to extract the EW- $Z Z jj$ sig-
 2061 nal from backgrounds. The binning of MD distributions in SR is optimized to maximize
 2062 the sensitivity for detecting EW signal. The normalization of QCD- $Z Z jj$ production
 2063 (μ_{QCD}^{III}) in $\ell\ell\ell\ell$ channel is determined by data from simultaneously fit in SR and QCD
 2064 CR as described in section 5.4. The signal strength of EW- $Z Z jj$ production (μ_{EW}) is
 2065 taken as parameter of interest and floated in the fit. The effects of the uncertainties related
 2066 to normalizations and shapes described previously in section 5.5 of background processes
 2067 in the MD distribution are all taken into account.

2068 In most case, a common nuisance parameter is used for each source of systematic in
 2069 all bins and all channels. The statistical uncertainties for simulated samples are uncorre-
 2070 lated among all bins, and the background uncertainties only applied to their corresponding
 2071 backgrounds. Furthermore, to be more conservative, the generator modelling uncertainty
 2072 for QCD- $Z Z jj$ production mentioned in section 5.5 is separated to be two nuisance pa-
 2073 rameters in low and high MD region.

2074 **5.7.4 Result of statistical fit**

2075 The results of statistical fit for $\ell\ell\ell\ell$ final state are presented in table 5.13. To drive
 2076 expected results, the observed data is used for QCD CR to extract normalization factor of
 2077 QCD component (μ_{QCD}^{III}), while in SR, asimov data built from background prediction and
 2078 signal model with SM assumed cross section is used.

2079 As a conclusion, the background-only hypothesis is rejected at 5.5σ (3.9σ) for data
 2080 (expectation), which leads to the observation of EW- $Z Z jj$ production.

	μ_{EW}	$\mu_{\text{QCD}}^{\ell\ell\ell jj}$	Significance Obs. (Exp.)
$\ell\ell\ell jj$	1.54 ± 0.42	0.95 ± 0.22	$5.48 (3.90) \sigma$

Table 5.13 Observed μ_{EW} and $\mu_{\text{QCD}}^{\ell\ell\ell jj}$, as well as the observed and expected significance from the individual $\ell\ell\ell\ell$ channel. The full set of systematic uncertainties are included.

Figure 5.11 shows the post-fit MD distributions for $\ell\ell\ell\ell$ channel in SR (left) and QCD CR (right). The EW- $ZZjj$ cross section measured in $\ell\ell\ell\ell$ channel is extracted to be $0.94 \pm 0.26 \text{ fb}$.

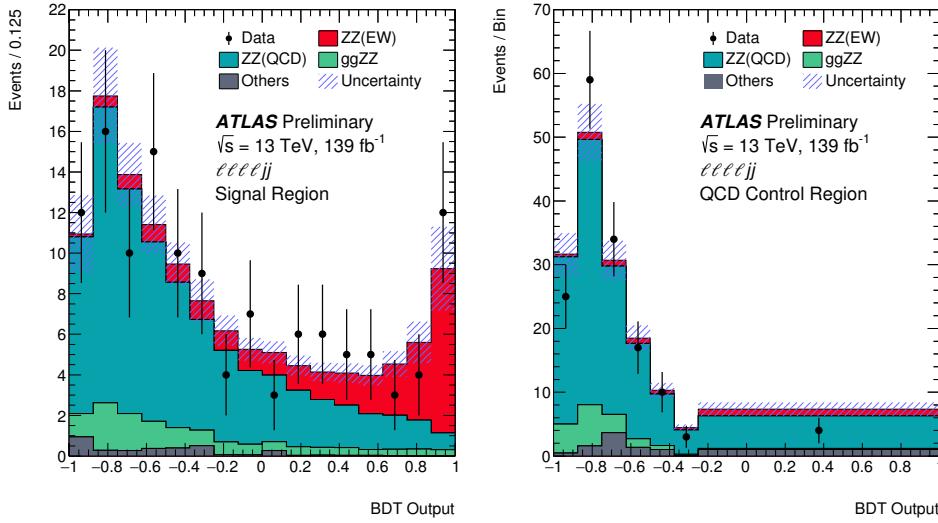


Fig. 5.11 Observed and expected multivariate discriminant distributions after the statistical fit in the $\ell\ell\ell\ell$ SR (left) and QCD CR (right). The error bands include the experimental and theoretical uncertainties, as well as the uncertainties in μ_{EW} and $\mu_{\text{QCD}}^{\ell\ell\ell jj}$. The error bars on the data points show the statistical uncertainty on data.

Figure 5.12 shows the m_{jj} distribution in SR (left) and QCD CR (right), where the normalization of EW and QCD processes are scaled according to their observed value in table 5.13. High m_{jj} region is more sensitive for EW- $ZZjj$ events detection from this figure. Figure 5.13 shows the spectrum of invariant mass of $\ell\ell\ell\ell$ system (m_{ZZ}) in SR also with the normalization of EW and QCD processes scaled.

Figure 5.14 is the display of one event candidate of EW- $ZZjj$ production in $2e2\mu$ final state with two jets in forward and backward region.

5.8 Prospect study of EW- $ZZjj$ production in HL-LHC

The High-Luminosity Large Hadron Collider (HL-LHC) project aims to increase the luminosity by a factor of 10 beyond the LHC's design value to increase the potential for discoveries after 2025. The designed luminosity will reach 3000 fb^{-1} with the centre-of-

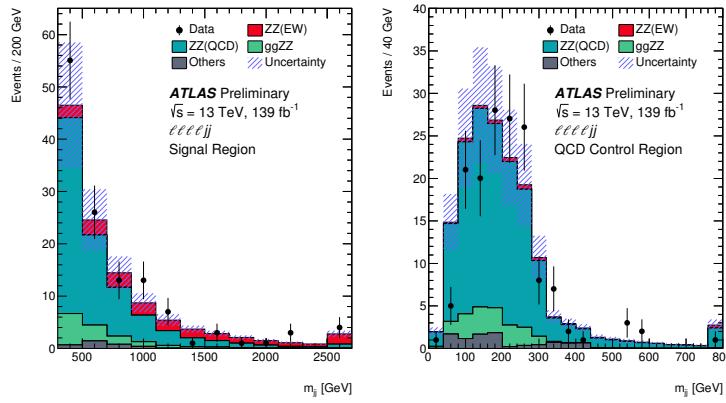


Fig. 5.12 Observed and expected m_{jj} distributions in SR (left) and QCD CR (right). The error bands include the expected experimental and theoretical uncertainties. The error bars on the data points show the statistical uncertainty. The contributions from the QCD and EW production of $Z Z j j$ events are scaled by 0.96 and 1.35, respectively, corresponding to the observed normalization factors in the statistical fit. The last bin includes the overflow events.

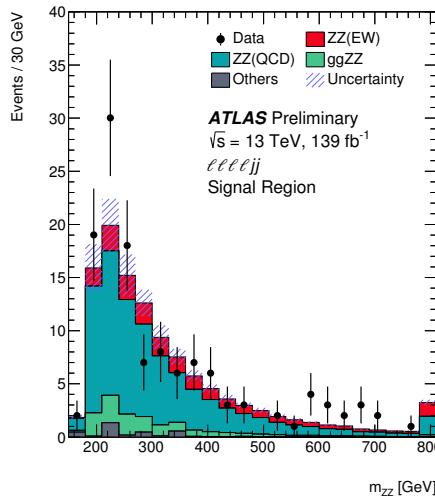


Fig. 5.13 Observed and expected m_{ZZ} spectrum in SR. The error bands include the expected experimental and theoretical uncertainties. The error bars on the data points show the statistical uncertainty. The contributions from the QCD and EW production of $Z Z j j$ events are scaled by 0.96 and 1.35, respectively, corresponding to the observed normalization factors in the statistical fit. The last bin includes the overflow events.

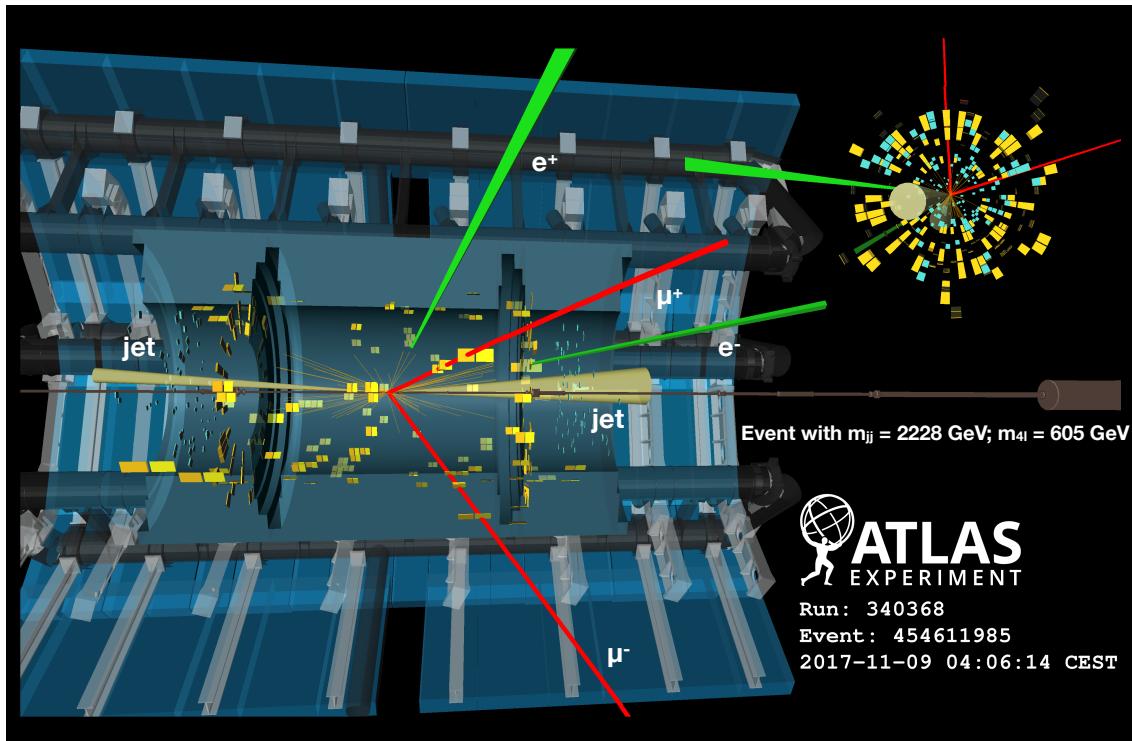


Fig. 5.14 Display of an event candidate of EW- $ZZjj$ production in $2e2\mu$ channel in last MD bin ($0.875 < \text{MD} < 1.0$). The invariant mass of the di-jet (four-lepton) system is 2228 (605) GeV.

mass energy of 14 TeV.

As introduced in previous sections, with full run-2 data of 139 fb^{-1} collected by ATLAS detector at the LHC, the EW- $ZZjj$ production is the last channel of observation for VBS processes with massive boson due to its very low cross section in ZZ decay. So we expect that this channel will benefit significantly from the increased luminosity at the HL-LHC, and can be studied in great details for this known mechanism.

In this section, a prospect study is performed for EW- $ZZjj$ production at the HL-LHC in the $\ell\ell\ell\ell$ channel. The study uses 3000 fb^{-1} of simulated pp collision data at a centre-of-mass energy of 14 TeV as expected to be recorded by the ATLAS detector at HL-LHC. All simulated events are produced at particle-level, and the detector effects of leptons and jets reconstruction and identification are estimated by corrections assuming the mean number of interactions per bunch crossing ($\langle\mu\rangle$) of 200.

5.8.1 The ATLAS detector at HL-LHC

As the expectation of HL-LHC, the new Inner Tracker (ITk)^[94] will extend the tracking acceptance capability of ATLAS detector to pseudorapidity ($|\eta|$) up to 4.0. By including a forward muon trigger, the upgraded Muon Spectrometer^[95] is also expected to provide muon identification capabilities to $|\eta|$ up to 4.0. In addition, the new high gran-

ularity timing detector (HGTD)^[96] that designed to mitigate the pile-up (PU) effects is also expected to be installed in the forward region of $2.4 < |\eta| < 4.0$. More details of expected performance of the upgraded ATLAS detector at the HL-LHC has been reported in Ref.^[97].

5.8.2 Simulation

The analysis is performed using particle-level events. The samples are generated at $\sqrt{s} = 14$ TeV and with a fast simulation based on the setting for ATLAS detector at the HL-LHC. The signal in this analysis is EW- $ZZjj$ process, while only the dominant irreducible background of QCD- $ZZjj$ is considered. Both signal and background are generated using Sherpa with the NNPDF3.0NNLO PDF set. The signal sample is modelled with two jets at Matrix Element (ME) level. The background is generated with up to one (three) outgoing partons at NLO (LO) in pQCD. As a quick study, other minor backgrounds such as fake backgrounds from $Z+jets$ and top-quark processes, as well as Diboson without 4l final-state and Triboson processes are not considered into this analysis. Furthermore, for hard scattering events, the pile-up collisions are set with a mean value of 200 interactions per bunch crossing. Signal and background yields are then scaled to an integrated luminosity of 3000 fb^{-1} as expected at the HL-LHC.

5.8.3 Event selection

The analysis selection follows closely to the one in ATLAS run-2 analysis as described in section 5.3. Here are some changes according to the expectation of the HL-LHC scenario for ATLAS detector:

- Extend the lepton (both electron and muon) identification in forward with $|\eta| < 4.0$
- Pile-up (PU) jet suppression is applied with a PU rejection factor of 50 for all PU jets in the region of $|\eta| < 3.8$, based on the expected ATLAS detector performance at the HL-LHC.
- The jets are required to have $p_T > 30$ (70) GeV in the $|\eta| < 3.8$ ($3.8 < |\eta| < 4.5$) region.
- For two selected jets, tight the m_{jj} requirement to > 600 GeV, and require $\Delta\eta_{jj} > 2$.

In addition, a fiducial volume, which is used to study the expected precision of the cross-section measurements, is defined at particle-level with the same kinematic requirements listed above.

Table 5.14 summarized the number of selected signal and background events normalized to 3000 fb^{-1} . In addition to the *baseline* selection listed above, to compare the

2145 different detector scenarios at the HL-LHC, two alternative selections are also studied:
 2146 • Reduce the lepton η region to 2.7, to understand the effect due to forward lepton
 2147 reconstruction and identification with the upgraded ATLAS detector.
 2148 • Only apply the PU jet suppression with region $|\eta| < 2.4$, to measure the improve-
 2149 ment of *baseline* by extending the rejection range of PU jets at the HL-LHC with
 2150 the installation of HGTD.

Selection	$N_{\text{EW-ZZjj}}$	$N_{\text{QCD-ZZjj}}$	$N_{\text{EW-ZZjj}} / \sqrt{N_{\text{QCD-ZZjj}}}$
Baseline	432 ± 21	1402 ± 37	11.54 ± 0.58
Leptons with $ \eta < 2.7$	373 ± 19	1058 ± 33	11.46 ± 0.62
PU jet suppression only in $ \eta < 2.4$	536 ± 23	15470 ± 120	4.31 ± 0.19

**Table 5.14 Comparison of event yields for signal ($N_{\text{EW-ZZjj}}$) and background ($N_{\text{QCD-ZZjj}}$) pro-
cesses, and expected significance of EW- $ZZjj$ processes, normalized to 3000 fb⁻¹
data at 14 TeV, with baseline and alternative selections. Uncertainties in the table
refer to expected data statistical uncertainty at 14 TeV with 3000 fb⁻¹.**

2151 From this table, one can see the extended track coverage increases the $\ell\ell\ell\ell jj$ events
 2152 by 15 to 30%, via improving the lepton efficiency. But the significance of searching for
 2153 EW- $ZZjj$ process does not improve so much due to the large increment of background
 2154 events.

2155 Figure 5.15 shows the kinematic distributions of di-jet invariant mass (m_{jj}), the ZZ
 2156 invariant mass (m_{ZZ}) and the ϕ separation of two Z bosons ($|\Delta\phi(ZZ)|$) as well as the
 2157 centrality of the ZZ system. The ZZ centrality is defined as:

$$ZZ \text{ centrality} = \frac{|y_{ZZ} - (y_{j1} + y_{j2})/2|}{|y_{j1} - y_{j2}|} \quad (5.10)$$

2158 To measure the event yield, the top panel shows the stack distribution for EW- and QCD-
 2159 $ZZjj$ processes, while bottom panel is the ratio between two processes.

2160 5.8.4 Systematics

2161 According to studies in section 5.5, the dominant systematic in $\ell\ell\ell\ell$ channel is from
 2162 theoretical systematic for QCD- $ZZjj$ background process. Different sizes of systematics
 2163 have been studied, at a factor of 5, 10 and 30% on background modelling. The 5% un-
 2164 certainty is an optimal estimation when there is enough data events from QCD-enriched
 2165 control region at the HL-LHC that can be used to constrain the theoretical normaliza-
 2166 tion on QCD- $ZZjj$ process. The 30% one is a conservative estimation, in which the
 2167 uncertainties are directly calculated from different PDF sets and QCD renormalization

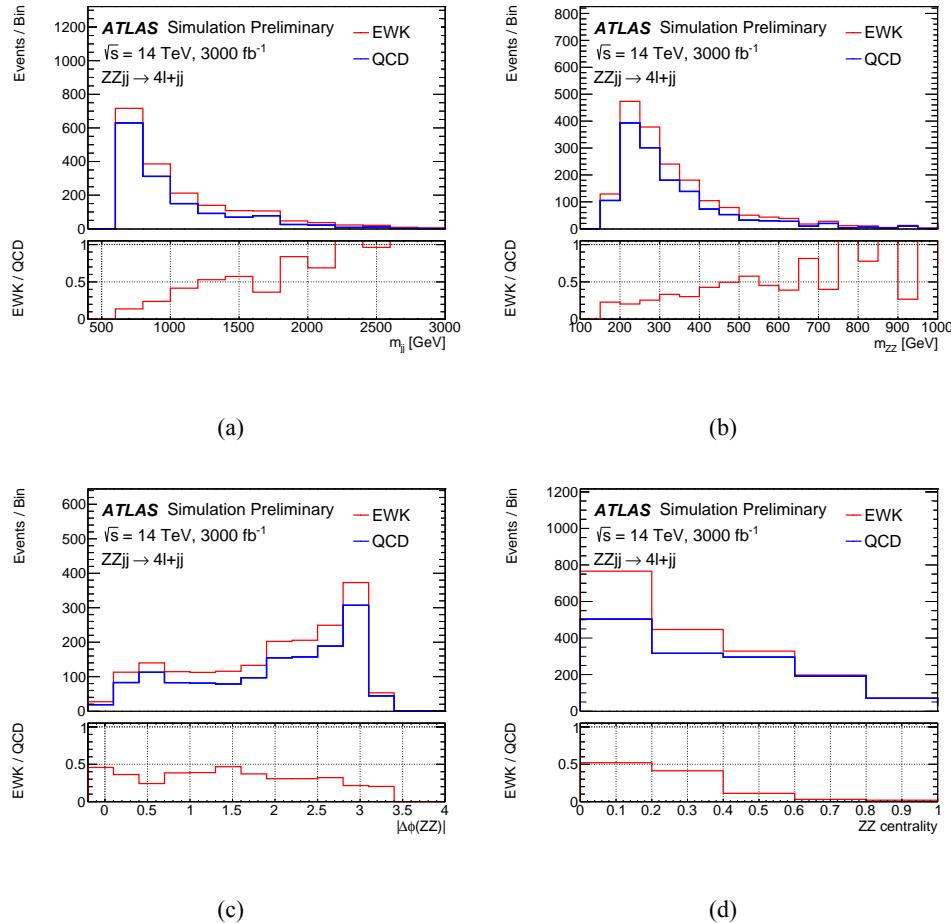


Fig. 5.15 Detector-level distributions of EW- and QCD- $Z Z jj$ processes with selected events in defined phase space at 14 TeV of (a) m_{jj} , (b) m_{ZZ} , (c) $|\Delta\phi(ZZ)|$, (d) ZZ centrality, normalized to 3000 fb^{-1} .

and factorization scales, following recommendation from the PDF4LHC mentioned in section 5.5.

For experimental sources, the jet systematics have been checked following the setting provided by the HL-LHC in Ref.^[97], and the uncertainties are within 5% level, which is smaller than run-2 measurement at 10%. Figure 5.16 depicts the up and down variations for jet uncertainty provided by the HL-LHC performance tool as function of dijet invariant mass (m_{jj}). Therefore, a conservative 5% uncertainty is used as experimental uncertainty.

Since the final result relies greatly on the uncertainties, especially the theoretical uncertainties on QCD- $Z Z jj$ production. So results with different uncertainty conditions are shown as below:

- The case with statistical uncertainty of simulated samples only.
- The case with statistical and experimental uncertainties (5%)
- The case with statistical, experimental and additional theoretical uncertainties at

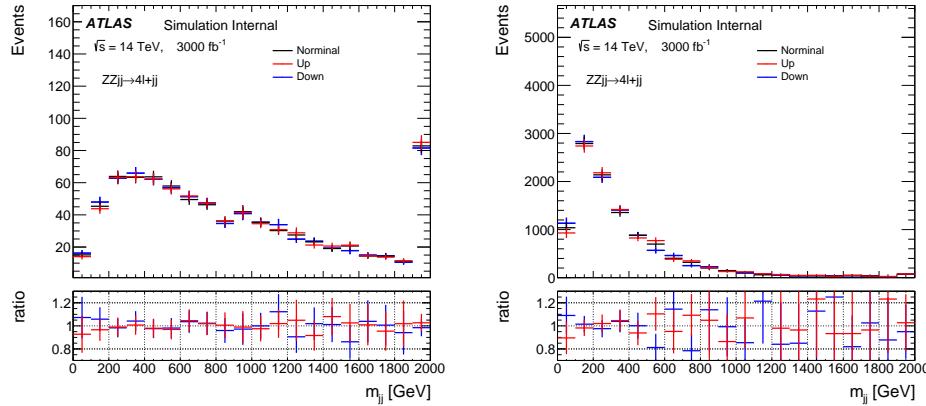


Fig. 5.16 Jet variations on m_{jj} distribution for EW- $Z Z jj$ (left) and QCD- $Z Z jj$ (right) processes with luminosity of 3000 fb^{-1} at 14 TeV. *Upgrade Performance Function* is used to extract the uncertainties with *baseline* setting.

2181 5%, 10% and 30% levels respectively.

2182 Three different sources of uncertainties are treated as uncorrelated and summed up
2183 quadratically.

2184 5.8.5 Results

2185 In this analysis, instead of a statistical fit, the expected significance of EW- $Z Z jj$
2186 production is calculated as:

$$\text{Significance} = \frac{S}{\sqrt{\sigma(B)_{\text{stat.}}^2 + \sigma(B)_{\text{syst.}}^2}}, \quad (5.11)$$

2187 where S presents the number of selected signal events, and $\sigma(B)_{\text{stat.}}$ and $\sigma(B)_{\text{syst.}}$ denote
2188 the statistical and systematic (exp. + theo.) uncertainties from background processes. The
2189 statistical uncertainty is computed from expected data yield with an integrated luminosity
2190 of 3000 fb^{-1} .

2191 Base on baseline selection of $m_{jj} > 600 \text{ GeV}$, a additional scan over different m_{jj} cuts
2192 are also performed with a step of 50 GeV under different systematic conditions, as shown
2193 in figure 5.17.

2194 In addition, the expected differential cross section of EW- $Z Z jj$ process is measured

2195 in the defined phase space at 14 TeV, as a function of m_{ZZ} and m_{jj} , shown in figure 5.18.

2196 The expected differential cross sections are calculated as:

$$\sigma = \frac{N_{\text{pseudo-data}} - N_{\text{QCD-ZZjj}}}{L * C_{\text{EW-ZZjj}}} \quad (5.12)$$

$$C_{\text{EW-ZZjj}} = \frac{N_{\text{EW-ZZjj}}^{\text{det.}}}{N_{\text{EW-ZZjj}}^{\text{part.}}}$$

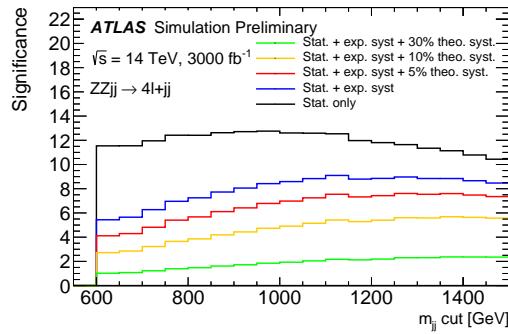


Fig. 5.17 The expected significance of EW- $Z Z j j$ processes as a function of different m_{jj} cut with 3000 fb^{-1} , under conditions of different sizes of theoretical uncertainties on the QCD- $Z Z j j$ background modelling. The statistical uncertainty is estimated from expected data yield at 14 TeV with 3000 fb^{-1} . Different uncertainties are summed up quadratically.

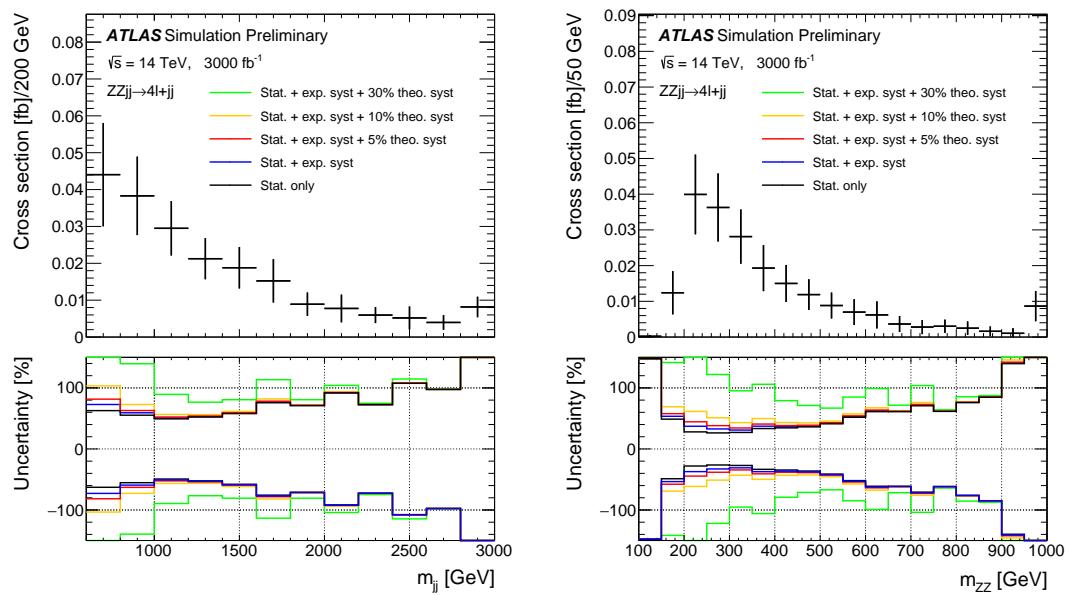


Fig. 5.18 The projected differential cross-sections at 14 TeV for the EW- $Z Z j j$ processes as a function of m_{jj} (left) and m_{ZZ} (right). The top panel shows measurement with statistical only case, where statistical uncertainty is estimated from expected data yield at 14 TeV with 3000 fb^{-1} . The bottom panel shows impact of different sizes of systematic uncertainties.

2197 where $N_{pseudo-data}$ denotes the expected number of data events with 3000 fb^{-1} luminosity
 2198 at 14 TeV, and $N_{QCD-ZZjj}$ and $N_{EW-ZZjj}$ are the number of predicted events of QCD-
 2199 $Z Z jj$ and EW- $Z Z jj$ processes in particle-level. The $C_{EW-ZZjj}$ factor represents the
 2200 detector efficiency for EW- $Z Z jj$ processes introduced in section 5.6.1. The interference
 2201 between EW- and QCD- $Z Z jj$ processes is ignored due to its minor contribution.

2202 The number of expected integrated cross section as well as its uncertainty under dif-
 2203 ferent systematic conditions are shown in table 5.15 with 3000 fb^{-1} luminosity at 14 TeV.
 2204 The statistical uncertainty is at 10% level when with such large luminosity. The result is
 2205 dominated by systematics and can reach 100% level when theoretical modelling uncer-
 tainty is 30% for QCD- $Z Z jj$ processes.

	Cross section [fb]	Stat. only	Plus exp.	Plus 5% theo.	Plus 10% theo.	Plus 30% theo.
EW- $Z Z jj$	0.21	± 0.02	± 0.04	± 0.05	± 0.08	± 0.21

**Table 5.15 Summary of expected cross-section measured with different theoretical uncer-
 tainties. The statistical uncertainty is computed from expected data yield with
 3000 fb^{-1} at 14 TeV. Different uncertainties are treated as uncorrelated and
 summed quadratically.**

2206

2207 5.9 Conclusion

2208 The fiducial cross section for inclusive $ZZjj$ production is measured in this section,
2209 with a total relative uncertainty of 11% for the $\ell\ell\ell\ell$ final state, and found to be com-
2210 patible with the SM prediction. The observation of electroweak production of two jets
2211 in association with a Z -boson pair decay to $\ell\ell\ell\ell$ final state using 139 fb^{-1} of 13 TeV pp
2212 collision data collected by ATLAS experiment at the LHC is presented in this section.
2213 The search for electroweak production of two jets in association with a Z -boson pair is
2214 based on multivariate discriminants (MD) to enhance the separation between the signal
2215 and backgrounds. In $\ell\ell\ell\ell$ final state, the background-only hypothesis is rejected with
2216 an observed (expected) significance of 5.5 (3.9) σ , which gives the first observation of
2217 electroweak production in $ZZjj$ channel.

2218 In addition, the prospect study for the EW- $ZZjj$ production at the HL-LHC in the
2219 $\ell\ell\ell\ell$ channel, using 3000 fb^{-1} simulated pp collision data at a centre-of-mass energy of
2220 14 TeV has been presented. The precision of the expected measurements of the integrated
2221 and differential cross sections as a function of dijet or $\ell\ell\ell\ell$ invariant mass are shown.
2222 Under the assumption of theoretical uncertainty for the QCD- $ZZjj$ processes and ex-
2223 perimental uncertainty for jets being constraint at 5% level respectively, with statistical
2224 uncertainty in 3000 fb^{-1} being considered, the observation of the EW- $ZZjj$ process can
2225 reach a significance of 7σ .

2226 **Chapter 6 Search for heavy resonances decaying into**
2227 **a pair of Z bosons in $\ell\ell\ell\ell$ final state using pp collision**
2228 **data collected by ATLAS detector from 2015 to 2018**

2229 **6.1 Introduction**

2230 A new particle was discovered by the ATLAS and CMS Collaborations at the
2231 LHC^[67-68] in 2012. Both two experiments have confirmed that the properties including
2232 spin, couplings and parity of this new particle are consistent with Higgs boson predicted in
2233 the Standard Model (SM), which is an important milestone in understanding of the mecha-
2234 nism of EWSB. Nevertheless, the possibility that this newly discoved particle is just a part
2235 of the extended Higgs sector as predicted by various extensions in the SM cannot be ruled
2236 out. There are many models predicted the existence of new heavy resonances decaying
2237 into dibosons, such as a heavy spin-0 neutral Higgs boson^[98] and the two-Higgs-doublet
2238 models (2HDM)^[99], as well as the spin-2 Kaluza–Klein (KK) excitations of the graviton
2239 (G_{KK})^[100].

2240 Though with smaller branching ratio comparing to semileptonic or fully hadronic de-
2241 cay channels, the $\ell\ell\ell\ell$ final state has its unique sensitivity in mass range smaller than
2242 1 TeV region due to its good mass resolution and relative smaller experimental and the-
2243 oretical systematics. This section presents the search for heavy resonance decaying into
2244 a pair of Z bosons to the $\ell\ell\ell\ell$ final state, in which ℓ denotes to either an electron or
2245 a muon. Several signal hypotheses are considered. The first hypothesis is a heavy
2246 Higgs boson (spin-0 resonance) under the narrow-width approximation (NWA). Then as
2247 several theoretical models prefer non-negligible natural widths, the models under large-
2248 width approximation (LWA), assuming widths of 1%, 5%, 10% and 15% of the resonance
2249 mass, are also studied. In addition, the graviton excitations (spin-2 resonance) under the
2250 Randall–Sundrum model are also searched. It is assumed that the heavy resonance is
2251 produced predominantly via the gluon-gluon Fusion (ggF) and the Vector Boson Fusion
2252 (VBF) productions, but with the unknown ratio of two production rates. So the results are
2253 separated for ggF and VBF production modes. To gain more sensitivity, the $\ell\ell\ell\ell$ events
2254 are classified into ggF- and VBF-enriched categories. Moreover, for the NWA model, the
2255 categorizations are studied under both cut-based and multivariate (MVA) -based methods,
2256 the details of categorization are shown in following sections.

2257 The search uses the four-lepton invariant mass in the range of 200 GeV to 2000 GeV for

signal hypothesis of spin-0 resonance under the NWA model, and from 400 GeV to 2000 GeV for the one under the LWA models. And the spin-2 graviton signals are searched in the mass range from 600 GeV to 2000 GeV. The data collected by ATLAS detector at the LHC from 2015 to 2018 at the centre-of-mass energy of 13 TeV is used. In case of no excess, upper limits on the production rate of different signal hypotheses are computed from statistical fits to m_{4l} distribution.

6.2 Data and MC samples

6.2.1 Data samples

The data used in this analysis are collected by ATLAS detector at the centre-of-mass energy of 13 TeV during the years of 2015 to 2018. Only events passing the latest Good Run List (GRL) released by the Data Quality group from ATLAS experiment as listed in section 5.2.1 corresponding to an integrated luminosity of $139.0 \pm 2.4 \text{ fb}^{-1}$ are used. Table 6.1 listed the recorded integrated luminosity, average and peak pile-up of each year's data.

Table 6.1 Summary of the recorded integrated luminosity (lumi), average and peak pile-up (PU) of data from 2015 to 2018.

Year	recorded integrated lumi	lumi after GRL	average PU	peak PU
2015	3.86 fb^{-1}	36.2 fb^{-1}	13.4	28.1
2016	35.6 fb^{-1}		25.1	52.2
2017	46.9 fb^{-1}	44.3 fb^{-1}	37.8	79.8
2018	60.6 fb^{-1}	58.5 fb^{-1}	36.1	88.6

6.2.2 Background MC simulations

Background processes considered in this analysis include ZZ ($q\bar{q} \rightarrow ZZ$, $gg \rightarrow ZZ$), triboson (WWZ , WZZ , ZZZ), $Z+jets$ and top-quark ($t\bar{t}$, $t\bar{t}V$) processes.

The QCD $q\bar{q} \rightarrow ZZ$ process is modelled using Sherpa 2.2.2^[79] with the NNPDF3.0NNLO^[80] PDF, where events with up to one (three) outgoing partons are generated at NLO (LO) in pQCD. The production of ZZ from the gluon-gluon initial state with a four-fermion loop or with an exchange of the Higgs boson, which has an order of α_S^4 in QCD, is not included in this Sherpa simulation. So a separate gg induced ZZ sample including the continuum background, the SM Higgs boson, and the interference contribution is modelled using Sherpa 2.2.2 with the NNPDF3.0NNLO PDF set, and with an additional 1.7 k-factor^[81] being applied. The EW- $ZZjj$ production is simulated using

2283 Sherpa 2.2.2 with the NNPDF3.0NNLO PDF, and the $ZZZ \rightarrow \ell\ell\ell\ell qq$ process is also
2284 taken into account in this sample.

2285 The $Z+jets$ events are generated using Sherpa 2.2.2 with the NNPDF3.0NNLO PDF,
2286 in which the ME is calculated for up to two partons with next-to-leading-order (NLO)
2287 accuracy in pQCD and up to four partons with LO accuracy. The $Z+jets$ events are
2288 normalized using the next-to-next-to-leading-order (NNLO) cross section. The triboson
2289 processes with full leptonic decays and at least four prompt charged leptons are generated
2290 using Sherpa 2.1.1. For top-quark pair ($t\bar{t}$) production and the single top-quark produc-
2291 tions in t -channel, s -channel and Wt -channel, the Powheg-Box v2 is used with the CT10
2292 PDF. The productions of $t\bar{t}$ in association with Z boson(s) ($t\bar{t}Z$) is modelled with Mad-
2293 Graph5_aMC@NLO.

2294

6.2.3 Signal MC simulations

2295 One model considered in this analysis is heavy spin-0 resonance under the Narrow
2296 Width Approximation (NWA) simulated using Powheg-Box v2 MC event generator with
2297 the CT10 PDF. The gluon-gluon fusion (ggF) production mode and vector-boson fusion
2298 (VBF) production mode are calculated separately with matrix elements up to NLO in
2299 QCD. The Powheg-Box is interfaced to Pythia8 for parton showering, and for decaying
2300 the Higgs boson into the $H \rightarrow ZZ \rightarrow \ell\ell\ell\ell$ final states. Events of NWA signal are
2301 generated at mass points between 200 GeV to 2000 GeV using the step of 100 (200) GeV up
2302 to (above) 1 TeV in both ggF and VBF production modes.

2303 In addition, heavy Higgs boson events under the Large Width Approximation (LWA)
2304 with widths of 1%, 5%, 10% and 15% of the boson mass are generated using Mad-
2305 Graph5_aMC@NLO 2.3.2 interfaced to Pythia8. Only ggF production is consid-
2306 ered. Mass points between 400 GeV to 2000 GeV are simulated with the step of 100
2307 (200) GeV up to (above) 1 TeV. To describe jet multiplicity, MadGraph5_aMC@NLO is
2308 used to simulated process of $pp \rightarrow H + \geq 2\text{jets}$ at NLO in QCD with the FxFx merging
2309 scheme^[101].

2310 Spin-2 Kaluza–Klein (KK) gravitons (G_{KK}) from the Bulk Randall–Sundrum
2311 model^[102] are also studies in this analysis. Events are generated by Mad-
2312 Graph5_aMC@NLO at LO in QCD, which is then interfaced to Pythia8 for parton show-
2313 ering. The G_{KK} -gluon coupling $k/\overline{M}_{\text{Planck}}$, where k is the curvature scale of the extra
2314 dimension and $\overline{M}_{\text{Planck}}$ is the reduced Planck mass, is set to 1. The width of the resonance
2315 is correlated with the coupling $k/\overline{M}_{\text{Planck}}$ and in this configuration is around $\square 6\%$ of its
2316 mass. The mass of the G_{KK} is the only free parameter in this simplified model. Mass

2317 points between 600 GeV to 2 TeV with 200 GeV spacing were generated.

2318 **6.3 Analysis selections**

2319 **6.3.1 Objects selection**

2320 Similar to VBSZZ analysis in section 5.3, the selection of this analysis relies on the
2321 definition of multiple objects: *electrons*, *Muons*, and *jets*. Details of definitions for each
2322 object are described as below:

2323 **Electron:** As described in section 4.2.3, electrons are reconstructed from energy
2324 deposits in the EM calorimeter matched to a track in the inner detector. The electron
2325 candidates satisfying the *Loose* criterion valuing by the likelihood-based (LH) method
2326 are selected, with a selection efficiency ranging from 90% for transverse momentum
2327 $p_T = 20$ GeV to 96% for $p_T > 60$ GeV. In addition, the electrons are required to have
2328 $p_T > 7$ GeV, $|\eta| < 2.47$ and $|z_0 \sin\theta| < 0.5$ mm.

2329 **Muon:** To increase the acceptance range in reco-level for $\ell\ell\ell\ell$ channel, all four types
2330 of muons (CB, ST, CT, ME muons, described in section 4.2.4) are used. The CT muons
2331 are required to pass $p_T > 15$ GeV and $|\eta| < 0.1$, while the ST muons are also limited in
2332 $|\eta| < 0.1$ region. The ME muons are only used in the region of $2.5 < |\eta| < 2.7$. And at
2333 most one CT, ST or ME muon is allowed in one $\ell\ell\ell\ell$ quadruplet. The Muon candidates
2334 are required to pass $p_T > 5$ GeV and $|\eta| < 2.7$, and satisfy the *Loose* identification
2335 criterion with an efficiency of at least 98.5%. The impact parameter requirements of $|d_0|$
2336 < 1 mm and $|z_0 \sin\theta| < 0.5$ mm are further applied.

2337 **Jets:** Jets are clustered using the anti- k_t algorithm with radius parameter $R = 0.4$ im-
2338 plemented in the FastJet package as described in section 4.2.5. The ‘particle flow’ (PFlow)
2339 objects^[103], which combines measurements from both the tracker and the calorimeter,
2340 are used as inputs to the FastJet package. The energy deposited in the calorimeter by
2341 all charged particles is removed, and the jet reconstruction is performed on an ensem-
2342 ble of PFlow objects consisting of the remaining calorimeter energy and tracks which
2343 are matched to the hard interaction. This improves the accuracy of the charged-hadron
2344 measurement, while retaining the calorimeter measurements of neutral-particle energies.
2345 Compared to only using topological clusters, jets reconstructed with the particle flow algo-
2346 rithm with $p_T > 30$ GeV have approximately 10% better transverse momentum resolution.
2347 The jets used in this analysis are then required to have $p_T > 30$ GeV and $|\eta| < 4.5$. To
2348 further reduce the effects of pile-up jets, a jet vertex tagger (JVT) is applied to jets with
2349 $p_T < 60$ GeV and $|\eta| < 2.4$.

2350 **Overlap removal:** As the selected jet and lepton candidates can be reconstructed
 2351 from same detector information, an overlap-removal procedure is applied. For electron
 2352 and muon sharing the same ID track, the electron is selected in the case that the muon is
 2353 calorimeter-tagged and does not have a MS track, or is a segment-tagged muon, otherwise
 2354 the muon is selected. The jet overlapping with electron (muon) within a cone of size of
 2355 $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.2(0.1)$ are removed.

2356

6.3.2 Event selection

2357 First of all, the four-lepton events are required to pass single or multi-lepton triggers.
 2358 Due to the increasing of peak luminosity and pile-up, the p_T and E_T thresholds of triggers
 2359 increase slightly during the data-taking periods from 2015 to 2018. Table 6.2 summarizes
 2360 the triggers used for $\ell\ell\ell\ell$ channel. The overall trigger efficiency for selected signal
 2361 events passing final selection is around 98%.

**Table 6.2 Summary of the p_T (E_T) trigger thresholds (in GeV) employed for the muon (electron)
 trigger selection in the year of 2015, 2016, 2017, and 2018.**

Trigger item	Trigger threshold			
	2015	2016	2017	2018
single muon	$\mu 20; \mu 50; \mu 60$	$\mu 24; \mu 26; \mu 40; \mu 50$	$\mu 26; \mu 50; \mu 60$	$\mu 26; \mu 50; \mu 60$
single electron	$e24; e60; e120$	$e26; e60; e140; e300$	$e26; e60; e140; e300$	$e26; e60; e140; e300$
dimuon	$2\mu 10; \mu 18_{-\mu 8}$	$2\mu 10; 2\mu 14; \mu 22_{-\mu 8}$	$2\mu 14; \mu 22_{-\mu 8}$	$2\mu 14; \mu 22_{-\mu 8}$
dielectron	$2e12$	$2e15; 2e17$	$2e17; 2e24$	$2e17; 2e24$
electron-muon	$e24_{-\mu 8}$	$e24_{-\mu 8}; e26_{-\mu 8}$ $e17_{-\mu 14}; e7_{-\mu 24}; 2e12_{-\mu 10}; e12_{-\mu 10}$	$e26_{-\mu 8}$	$e26_{-\mu 8}$
trimuon	$\mu 18_{-\mu 4}$	$\mu 11_{-\mu 4}; \mu 6_{-\mu 4}; \mu 20_{-\mu 4}; 3\mu 4$ $3\mu 6$	$4\mu 4; \mu 20_{-\mu 4}; 3\mu 4$	$\mu 20_{-\mu 4}$
trilepton	$e17_{-2e9}$	$e17_{-2e9}; e17_{-2e10}$	$e24_{-2e12}$	$e24_{-2e12}$

2362 The $\ell\ell\ell\ell$ quadruplets are formed by two opposite-sign, same-flavour (OSSF) lepton
 2363 pairs ($\ell^+\ell^-$). The p_T threshold of first three leading leptons are required to be 20, 15
 2364 and 10 GeV. If there are more than one combination of lepton pairing in quadruplet, the
 2365 pairing is selected by keeping it with the mass of lepton pairs closest (leading pair, refers
 2366 as m_{12}) and second closest (sub-leading pair, refers as m_{34}) to Z boson mass. The mass
 2367 of leading pair is required to satisfy $50 < m_{12} < 106$ GeV, while the sub-leading pair is
 2368 required to be less than 115 GeV and larger than 50 GeV.

2369 The two lepton pairs in quadruplet are required to have angular separation with $\Delta R >$
 2370 0.1. To suppress the contribution from $J/\psi \rightarrow \ell\ell$ decays, for 4 μ and 4 e quadruplets,
 2371 the events are rejected if any opposite-sign same-flavour lepton pair is found with mass
 2372 below 5 GeV. If there are more than one quadruplets from different channels in event at

2373 this point, the one with highest expected signal rate is selected in the order of 4μ , $2e2\mu$,
 2374 $4e$. The transverse impact-parameter significance ($|d_0|/\sigma_{d_0}$) for muons (electrons) is than
 2375 required to be smaller than 3 (5) to suppress the backgrounds from heavy-flavour hadrons.

2376 In addition, the track- and calorimeter- based isolation criteria is required for all elec-
 2377 trons and muons to further suppress the reducible backgrounds of $Z+jets$ and $t\bar{t}$. For
 2378 lepton isolation selection, the two track- and calorimeter- based variables, $E_T^{topocone}$ and
 2379 $p_T^{varcone}$ as described in section 4.2.4 (section 4.2.3) for muons (electrons), are vulnerable
 2380 to pileup. For track-based variable, this is because of additional tracks in the event. The
 2381 definition of $p_T^{varcone}$ attempts to limit the tracks used in the calculation to those from the
 2382 vertex via a loose cut of $|z_0 \sin(\theta)| < 3$, which proved to be too loose in new pile-up
 2383 regime 2017 and 2018 datasets. So new track-based variable is used, by adding a require-
 2384 ment that the track be used in determining the vertex, or that, if not, it both pass the cut on
 2385 $|z_0 \sin(\theta)|$ and not be used in determining any other vertex, which makes the track-based
 2386 variable to be more isolation-robust in the high pile-up regime. The new variable is named
 2387 as $ptvarcone[cone]_TightTTVA_pt[p_T \text{ cut}]$, where [cone] is the cone size and [p_T cut] is
 2388 the cutoff for including tracks in the calculation.

2389 For calorimeter-based variable, the calculation of $E_T^{topocone}$ corrects the pile-up effects
 2390 by subtracting an average pileup contribution computed over the whole detector. But with
 2391 the increasing of energy density of pile-up events, the root mean square (RMS) of $E_T^{topocone}$
 2392 variable increases, which leads to the increment of possibility that the pile-up fluctuations
 2393 are not be accounted for correctly. One possible solution is that use particle-flow (PFlow)
 2394 method to calculate the calorimeter isolation. As part of PFlow reconstruction process, it
 2395 assigns the clusters to tracks which improves the track-cluster association for better de-
 2396 termination of the raw value of the E_T in the cone. And using PFlow jets to calculate the
 2397 pileup correction provides a further improvement. So a resulting variable named $neflow-$
 2398 $isol[cone]$ is used. Finally, a requirement of isolation, called *FixedCutPFlowLoose*, which
 2399 gives better performance in high piup-up condition is applied to electrons and muons as:
 2400 $(\max(ptcone20_TightTTVA_pt500, ptvarcone30_TightTTVA_pt500) + 0.4 \times neflow-$
 2401 $isol20) / p_T < 0.16$

2402 On the top of impact parameter cut and lepton isolation cut, the four-lepton candidates
 2403 are also required to originate from a common vertex to reduce $Z+jets$ and $t\bar{t}$ backgrounds.
 2404 This is ensured by applying a vertex fit χ^2 cut of 4 ID tracks of lepton candidates satisfying
 2405 $\chi^2/N_{dof} < 6$ (9) for events in 4μ ($4e$ and $2e2\mu$) channel(s).

2406 To improve the mass resolution, the QED process of final state radiation (FSR) pho-
 2407 tons in Z boson decays are taken into account in the reconstruction of Z bosons. The

2408 four-momentum of any reconstructed photon that is consistent with having been radiated
2409 from lepton(s) in leading pair are added into final state. Moreover, the four-momenta of
2410 leptons in both (leading and sub-leading) pairs are recomputed by performing a Z -mass-
2411 constrained kinematic fit, which uses a Breit–Wigner Z boson line-shape and Gaussian
2412 function with width set to the expected lepton resolution per lepton to model the momen-
2413 tum response function. The Z -mass-constrained mass improves the $m_{4\ell}$ resolution by up
2414 to 15% depending on m_H .

2415 In summary, table 6.3 lists a comprehensive object and event level selection as de-
2416 scribed above. Table 6.4 to 6.7 shows the cutflow of NWA ggF and VBF signal at the
2417 mass points of 600 and 1000 GeV as examples.

2418

6.3.3 Event categorizations

2419 To improve the sensitivity of search in both VBF and ggF production mode in NWA
2420 model, events are classified into the VBF- and ggF- enriched categories. With the statistic
2421 increasing in full run-2 data, a multivariate (MVA) based classifier has been studied for
2422 NWA signal, while in the meantime the traditional cut-based classifier is also used as a
2423 model-independent result for all three (NWA, LWA, graviton) models.

2424

1. Cut-based categorization

2425 There are four categories in total: one VBF-enriched category and three ggF-enriched
2426 categories. The categorization is defined based on kinematic cuts:

- 2427 • VBF-CBA-enriched category: Events have at least two selected jets as defined in
2428 section 6.3.1, with the two leading jets being separated by $|\Delta\eta_{jj}| > 3.3$ and invariant
2429 mass satisfying $m_{jj} > 400$ GeV;
2430 • ggF-CBA-enriched categories: The remaining events that are not classified into
2431 VBF-enriched category. Then events are categorized into three channels based on
2432 lepton-flavor, namely ggF_2e2μ, ggF_4e and ggF_4μ.

2433

2. MVA-based categorization

2434 In order to target different production modes, two types of classifiers, one dedicate to
2435 VBF production while the other one for ggF, have been trained using deep neural network
2436 technique. Details of two classifiers are described as below:

2437

DNN models

2438 Figure 6.1 shows the architecture of VBF (left) and ggF (right) network. The VBF
2439 network includes three parts: two recurrent neural networks (RNNs) and one multilayer
2440 perceptron (MLP). The ggF network consists of one RNN and one MLP.

2441 For training, the VBF and ggF signal samples at the masses of 200, 300, 400, 500, 600,

Table 6.3 Summary of the object and event selection requirements.

Physics Objects	
Electrons	
Loose Likelihood quality electrons with hit in innermost layer, $E_T > 7$ GeV and $ \eta < 2.47$	
Interaction point constraint: $ z_0 \cdot \sin \theta < 0.5$ mm (if ID track is available)	
Muons	
Loose identification with $p_T > 5$ GeV and $ \eta < 2.7$	
Calo-tagged muons with $p_T > 15$ GeV and $ \eta < 0.1$, segment-tagged muons with $ \eta < 0.1$	
Stand-alone and silicon-associated forward restricted to the $2.5 < \eta < 2.7$ region	
Combined, stand-alone (with ID hits if available) and segment-tagged muons with $p_T > 5$ GeV	
Interaction point constraint: $ d_0 < 1$ mm and $ z_0 \cdot \sin \theta < 0.5$ mm (if ID track is available)	
Jets	
anti- k_T jets with <i>bad-lose</i> identification, $p_T > 30$ GeV and $ \eta < 4.5$	
Overlap removal	
Jets within $\Delta R < 0.2$ of an electron or $\Delta R < 0.1$ of a muon are removed	
Vertex	
At least one collision vertex with at least two associated track	
Primary vertex	
Vertex with the largest p_T^2 sum	
Event Selection	
Quadruplet Selection	<ul style="list-style-type: none"> - Require at least one quadruplet of leptons consisting of two pairs of same-flavour opposite-charge leptons fulfilling the following requirements: - p_T thresholds for three leading leptons in the quadruplet: 20, 15 and 10 GeV - Maximum one calo-tagged or stand-alone muon or silicon-associated forward per quadruplet - Leading di-lepton mass requirement: $50 < m_{12} < 106$ GeV - Sub-leading di-lepton mass requirement: $50 < m_{34} < 115$ GeV - $\Delta R(\ell, \ell') > 0.10$ for all leptons in the quadruplet - Remove quadruplet if alternative same-flavour opposite-charge di-lepton gives $m_{\ell\ell} < 5$ GeV - Keep all quadruplets passing the above selection
Isolation	<ul style="list-style-type: none"> - Contribution from the other leptons of the quadruplet is subtracted - FixedCutPFlowLoose WP for all leptons
Impact Parameter	<ul style="list-style-type: none"> - Apply impact parameter significance cut to all leptons of the quadruplet
Significance	<ul style="list-style-type: none"> - For electrons: $d_0/\sigma_{d_0} < 5$ - For muons: $d_0/\sigma_{d_0} < 3$
Best Quadruplet Selection	<ul style="list-style-type: none"> - If more than one quadruplet has been selected, choose the quadruplet with highest Higgs decay ME according to channel: 4μ, $2e2\mu$, $2\mu2e$ and $4e$
Vertex Selection	<ul style="list-style-type: none"> - Require a common vertex for the leptons: - $\chi^2/\text{ndof} < 5$ for 4μ and < 9 for others decay channels

Table 6.4 Cutflow table for a narrow-width ggF signal sample at $m_H = 600$ GeV. N_{event} denotes the number of events selected after each cut is applied, normalized to 139 fb^{-1} , according to the expected upper limit on the cross section. The acceptances (the proportion of events selected relative to the initial number of events) are also included.

	N_{event}	$N_{\text{event}}/\text{BR}(ZZ \rightarrow \ell\ell\ell\ell)$	Acc. [%]	Acc. · $\text{BR}(ZZ \rightarrow \ell\ell\ell\ell) \cdot 1000$
Initial	17.902	3 964.3	100.00	4.516
Lepton selection	6.247	1 383.4	34.90	1.576
SFOS	5.758	1 275.1	32.16	1.453
Kinematic cuts	5.754	1 274.2	32.14	1.452
Z_1 Mass	5.726	1 267.9	31.98	1.444
Z_2 Mass	5.112	1 132.0	28.56	1.290
J/ψ Veto	5.111	1 131.9	28.55	1.289
ΔR	5.111	1 131.7	28.55	1.289
Isolation	4.864	1 077.0	27.17	1.227
Impact parameters	4.796	1 062.1	26.79	1.210
Vertex requirement	4.786	1 059.8	26.73	1.207
Trigger	4.783	1 059.1	26.72	1.207
“Badjet” veto	4.763	1 054.7	26.61	1.201

Table 6.5 Cutflow table for a narrow-width ggF signal sample at $m_H = 1000$ GeV. N_{event} denotes the number of events selected after each cut is applied, normalized to 139 fb^{-1} , according to the expected upper limit on the cross section. The acceptances (the proportion of events selected relative to the initial number of events) are also included.

	N_{event}	$N_{\text{event}}/\text{BR}(ZZ \rightarrow \ell\ell\ell\ell)$	Acc. [%]	Acc. · $\text{BR}(ZZ \rightarrow \ell\ell\ell\ell) \cdot 1000$
Initial	5.603	1 240.8	100.00	4.516
Lepton selection	2.141	474.1	38.21	1.725
SFOS	1.944	430.5	34.70	1.567
Kinematic cuts	1.943	430.3	34.68	1.566
Z_1 Mass	1.932	427.8	34.48	1.557
Z_2 Mass	1.715	379.7	30.61	1.382
J/ψ Veto	1.715	379.7	30.60	1.382
ΔR	1.714	379.6	30.60	1.382
Isolation	1.640	363.2	29.27	1.322
Impact parameters	1.620	358.6	28.90	1.305
Vertex requirement	1.616	357.8	28.84	1.302
Trigger	1.615	357.7	28.83	1.302
“Badjet” veto	1.609	356.2	28.71	1.297

Table 6.6 Cutflow table for a narrow-width VBF signal sample at $m_H = 600$ GeV. N_{event} denotes the number of events selected after each cut is applied, normalized to 139 fb^{-1} , according to the expected upper limit on the cross section. The acceptances (the proportion of events selected relative to the initial number of events) are also included.

	N_{event}	$N_{\text{event}}/\text{BR}(ZZ \rightarrow \ell\ell\ell\ell)$	Acc. [%]	Acc. · $\text{BR}(ZZ \rightarrow \ell\ell\ell\ell) \cdot 1000$
Initial	12.143	2 688.9	100.00	4.516
Lepton selection	4.307	953.7	35.47	1.602
SFOS	3.975	880.2	32.74	1.478
Kinematic cuts	3.972	879.6	32.71	1.477
Z_1 Mass	3.953	875.4	32.56	1.470
Z_2 Mass	3.545	785.0	29.19	1.318
J/ψ Veto	3.545	785.0	29.19	1.318
ΔR	3.544	784.9	29.19	1.318
Isolation	3.418	756.9	28.15	1.271
Impact parameters	3.368	745.9	27.74	1.253
Vertex requirement	3.362	744.5	27.69	1.250
Trigger	3.360	744.0	27.67	1.250
“Badjet” veto	3.340	739.7	27.51	1.242

Table 6.7 Cutflow table for a narrow-width VBF signal sample at $m_H = 1000$ GeV. N_{event} denotes the number of events selected after each cut is applied, normalized to 139 fb^{-1} , according to the expected upper limit on the cross section. The acceptances (the proportion of events selected relative to the initial number of events) are also included.

	N_{event}	$N_{\text{event}}/\text{BR}(ZZ \rightarrow \ell\ell\ell\ell)$	Acc. [%]	Acc. · $\text{BR}(ZZ \rightarrow \ell\ell\ell\ell) \cdot 1000$
Initial	3.827	847.4	100.00	4.516
Lepton selection	1.474	326.5	38.53	1.740
SFOS	1.351	299.1	35.30	1.594
Kinematic cuts	1.350	299.0	35.28	1.593
Z_1 Mass	1.341	297.0	35.04	1.583
Z_2 Mass	1.195	264.6	31.23	1.410
J/ψ Veto	1.195	264.6	31.23	1.410
ΔR	1.195	264.6	31.22	1.410
Isolation	1.161	257.1	30.34	1.370
Impact parameters	1.148	254.1	29.99	1.354
Vertex requirement	1.146	253.8	29.95	1.352
Trigger	1.145	253.6	29.93	1.352
“Badjet” veto	1.139	252.2	29.77	1.344

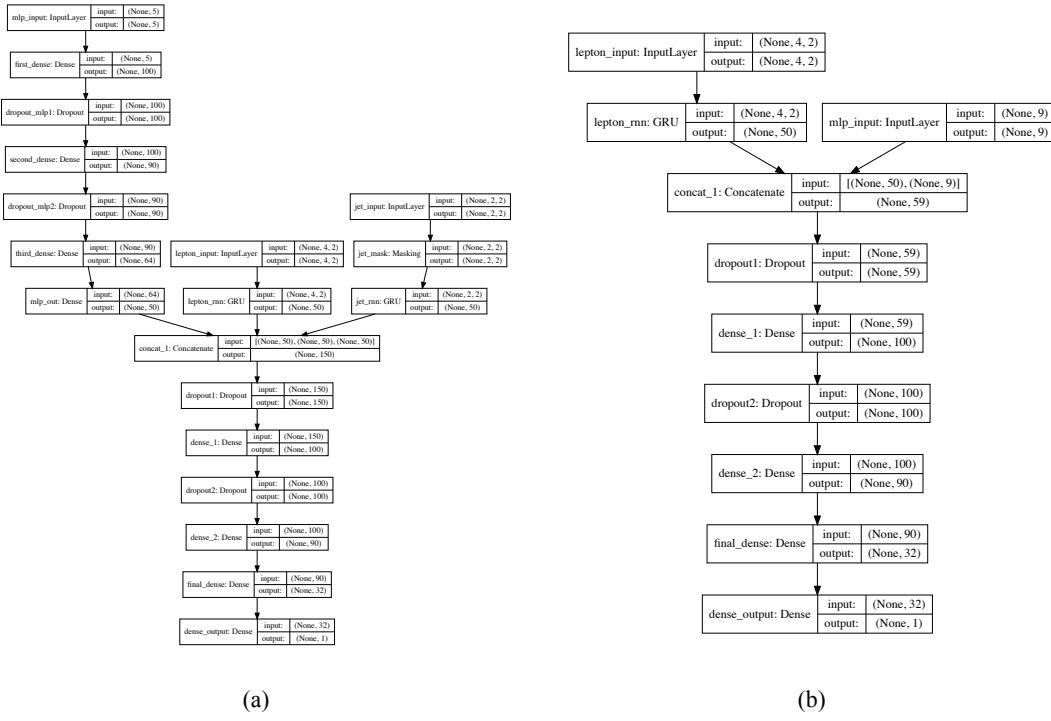


Fig. 6.1 (a) VBF DNN architecture diagram. (b) ggF DNN architecture.

700, 800, 900, 1000, 1200, 1400 GeV are used with positive label. The VBF (ggF) signals are only used for VBF (ggF) classifier. The background including simulated samples of QCD and EW $q\bar{q} \rightarrow ZZ$ processes as well as $gg \rightarrow ZZ$ process summed according to their cross section are signed with negative labels. In addition to the selections described in section 6.3.2, the events used for VBF network are required to have $N_{\text{jets}} \geq 2$, while $N_{\text{jets}} < 2$ is required for events in ggF network.

In order to assign equivalent importance to signals with different mass assumptions, during the training, signal events are reweighted to follow the $m_{4\ell}$ distribution from background, as shown in figure 6.2 (figure 6.3) before (left) and after(right) reweighting for VBF (ggF) samples.

Input features

Table 6.8 (table 6.9) lists the input features used for VBF (ggF) network during the training. For VBF network, one RNN (the other one) takes the p_T and η of p_T -ordered four leptons (two leading jets) as input features, which intends to study the time relationship from particle decay between leptons (jets). For ggF network, the only one RNN model takes the p_T and η of p_T -ordered four leptons as inputs.

Evaluation of models

Figure 6.4 shows the output of “ggF-classifier” and “VBF-classifier” for data, SM backgrounds and an example signal at 600 GeV. The ggF and VBF signals cross section

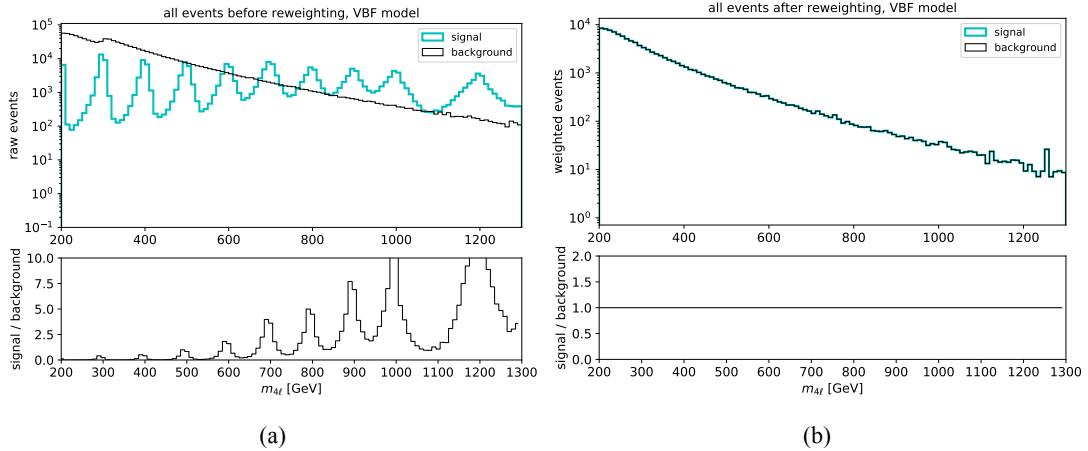


Fig. 6.2 (a) $m_{4\ell}$ distribution of raw (unweighted) training events for VBF signal (blue) and background (black); (b) $m_{4\ell}$ distribution of weighted VBF signal (blue) and background (black) used at training time.

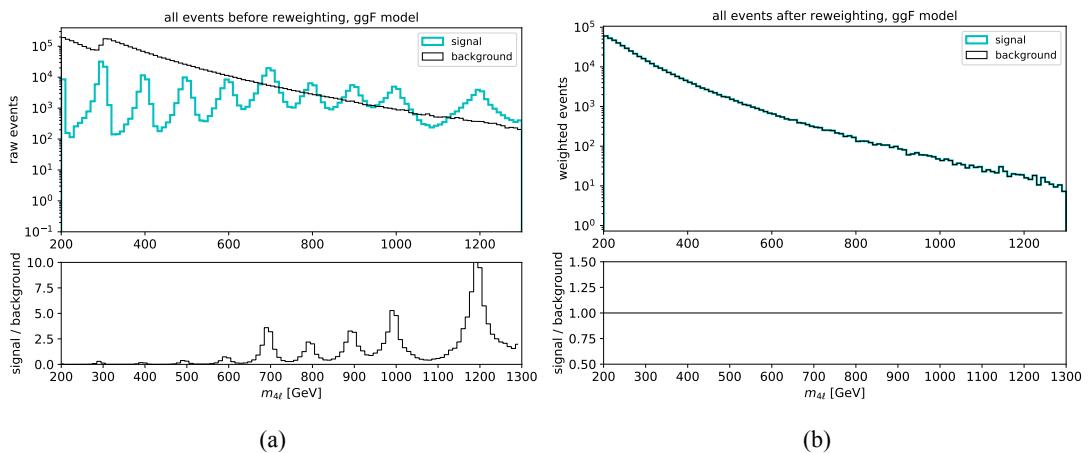


Fig. 6.3 (a) $m_{4\ell}$ distribution of raw (unweighted) training events for ggF signal (blue) and background (black); (b) $m_{4\ell}$ distribution of weighted ggF signal (blue) and background (black) used at training time.

Table 6.8 Input features used in the “VBF-classifier” for the $\ell\ell\ell\ell$ analysis. The RNN stands for the recurrent neural network and MLP for the multilayer perceptron.

Model	Inputs	Description
RNN	$p_T^{j0,j1}$	transverse momenta of the two leading jets
	$\eta^{j0,j1}$	pseudorapidity of the two leading jets
	$p_T^{\ell0,\ell1,\ell2,\ell3}$	transverse momenta of the four leptons
	$\eta^{\ell0,\ell1,\ell2,\ell3}$	pseudorapidity of the four leptons
MLP	$m_{4\ell}$	invariant mass of the four lepton system
	m_{jj}	invariant mass of the two leading jet system
	p_T^{jj}	transverse momentum of the two leading jet system
	$\Delta\eta_{H,j}$	difference in pseudorapidity between the four lepton system and the leading jet
	$\min\Delta R_{jZ}$	minimum distance between one of the two lepton pairs and a jet

2461 are set to be one hundred times of their observed upper limit described in section 6.7.3 for
 2462 ggF output and fifty times of the observed upper limit for VBF output for best visibility.

2463 Then the optimal cut at output score from each classifier is chosen based on an overall
 2464 good performance of classifier to have a large significance improvement while retaining
 2465 a high signal efficiency. Figure 6.5 shows the significance improvements of MVA-based
 2466 cuts when comparing with cut-based one at different VBF (left) and ggF (right) mass
 2467 samples, where the significance is calculated as:

$$Z = \sqrt{2 \left(n \ln \left[\frac{nb + \sigma^2}{b^2 + n\sigma^2} \right] - \frac{b^2}{\sigma^2} \ln \left[1 + \frac{\sigma^2(n-b)}{b(b+\sigma^2)} \right] \right)} \quad (6.1)$$

2468 Cut at 0.5 (0.8) for VBF (ggF) classifier is chosen as shown in solid lines.

2469 Then the events passing VBF classifier are categorized into VBF-MVA-enriched cat-
 2470 egory. Otherwise, the events failing VBF classifier but passing ggF classifier are catego-
 2471 rized into ggF-MVA-high category, which is further split into 3 channels. All remaining
 2472 events are sorted into one additional ggF-MVA-low category. Thus there are five cate-
 2473 gories defined in MVA-based categorization. In summary, cuts applied in categorization
 2474 are defined as follow, and these different phase spaces are also illustrated in figure 6.6.

- 2475 • VBF-MVA-enriched category: Events have at least two selected jets ($N_{\text{jets}} \geq 2$),
 2476 and with $DNN_{\text{VBF}} > 0.8$;
- 2477 • ggF-MVA-high categories: $(N_{\text{jets}} \geq 2 \text{ \&& } DNN_{\text{VBF}} \leq 0.8 \text{ \&& } DNN_{\text{ggF}} >$

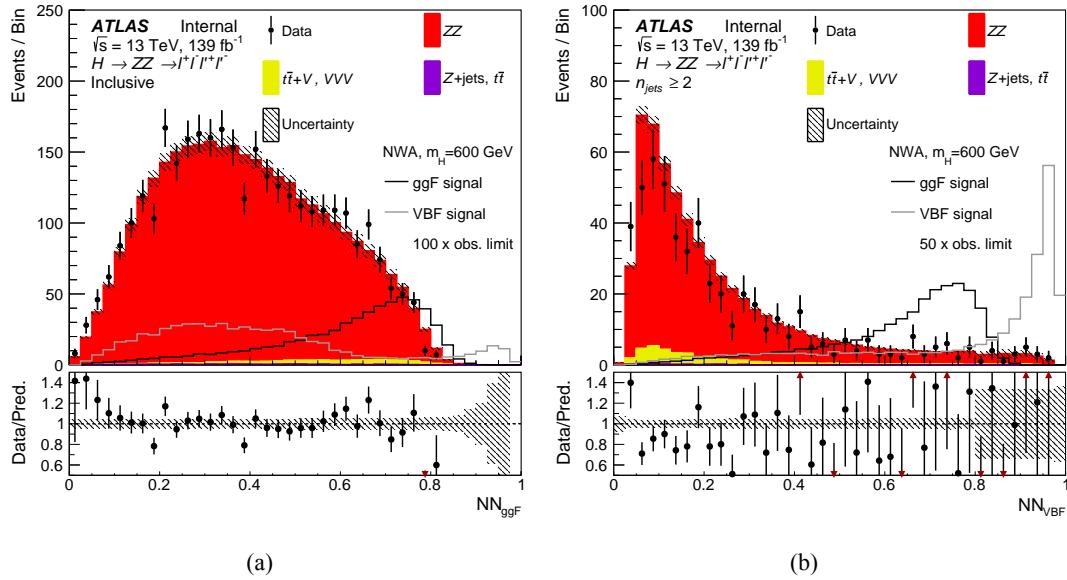


Fig. 6.4 The output score of “ggF-classifier” (a) and “VBF-classifier” (b) with the events passing the common event selections for the data, the SM backgrounds and an example of a NWA signal with a mass of 600 GeV. For the “VBF-classifier”, an additional requirement of at least two jets in the event is applied. The signals cross section are set to one hundred times of the observed limit for the “ggF-classifier” and fifty times of the observed limit for the “VBF -classifier”. The ZZ backgrounds are scaled by the normalisation factors shown in Table 6.15. The lower panels show the ratio of data to prediction. Only statistical and experimental systematic uncertainties are included.

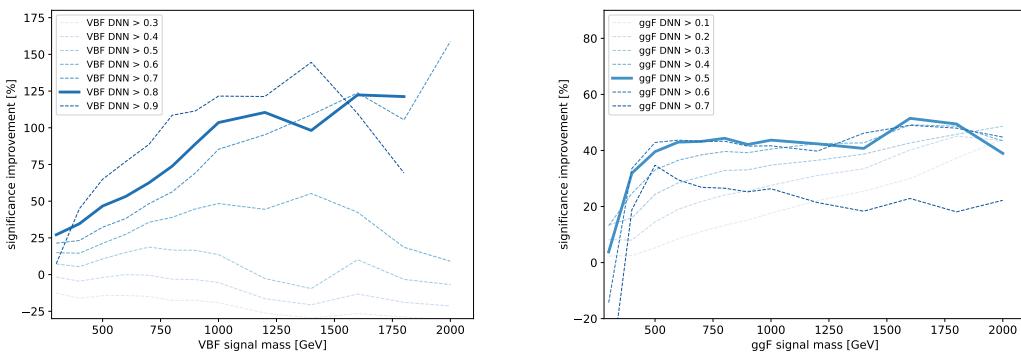


Fig. 6.5 Significance improvements of the MVA-based over the cut-based categorization of the VBF (ggF) category for VBF (ggF) signal samples from 300 to 2000 GeV for seven different cuts on the VBF (ggF) output score. The optimal cut of 0.8 (0.5) for VBF (ggF) score is chosen as the solid line, while other alternative cuts are plotted with dashed lines. For VBF category, results at 2000 GeV for cuts of 0.8 and 0.9 are missing due to a lack of background events passing this tight selection.

Table 6.9 Input features used in the “ggF-classifier” for the $\ell\ell\ell\ell$ analysis. The RNN stands for the recurrent neural network and MLP for the multilayer perceptron.

Model	Inputs	Description
RNN	$p_T^{\ell 0,\ell 1,\ell 2,\ell 3}$	transverse momenta of the four leptons
	$\eta^{\ell 0,\ell 1,\ell 2,\ell 3}$	pseudorapidity of the four leptons
MLP	$m_{4\ell}$	invariant mass of the four lepton system
	$p_T^{4\ell}$	transverse momentum of the four lepton system
	$\eta^{4\ell}$	pseudorapidity of the four lepton system
	$\cos \theta^*$	production angle of the leading Z defined in the four lepton rest frame
	$\cos \theta_1$	angle between the negative final state lepton and the direction of flight of leading Z in the Z rest frame
	$\cos \theta_2$	angle between the negative final state lepton and the direction of flight of sub-leading Z in the Z rest frame
	Φ	angle between the decay planes of the four final state leptons expressed in the four lepton rest frame
	p_T^{j0}	transverse momentum of the leading jet
	η^{j0}	pseudorapidity of the leading jet

2478 $0.5) \parallel (N_{\text{jets}} < 2 \&\& DNN_{\text{ggF}} > 0.5);$
 2479 • ggF-MVA-low category: All remaining events that fail VBF and ggF cuts men-
 2480 tioned above.

2481 **6.3.4 Signal acceptance**

2482 The signal acceptance is defined as the ratio of events passing all analysis selection in
 2483 each category to the total number of simulated events in whole phase space. In denom-
 2484 inator, the events with τ final states are not taken into account. And the contribution of
 2485 τ -lepton decay to electrons and muons final states is found to be negligible.

2486 Figure 6.7 and 6.8 show the acceptance of NWA signals in DNN- and Cut- based
 2487 categorization, estimated by merging the three signal MC campaigns, mc16a, mc16d and
 2488 mc16e. A 3-rd order polynomial fit is applied for each category.

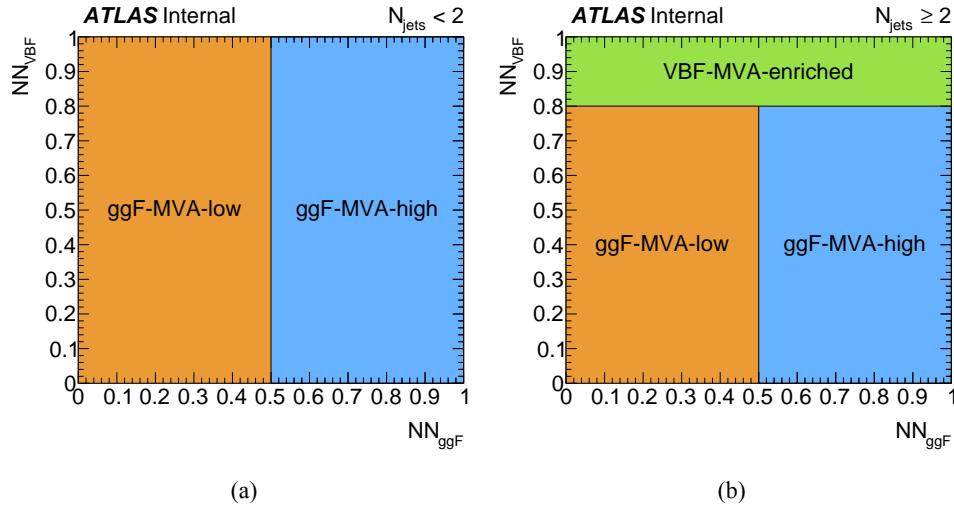


Fig. 6.6 Illustration of the MVA-based VBF and ggF event classification for events with (a) $N_{\text{jets}} < 2$ and (b) $N_{\text{jets}} \geq 2$.

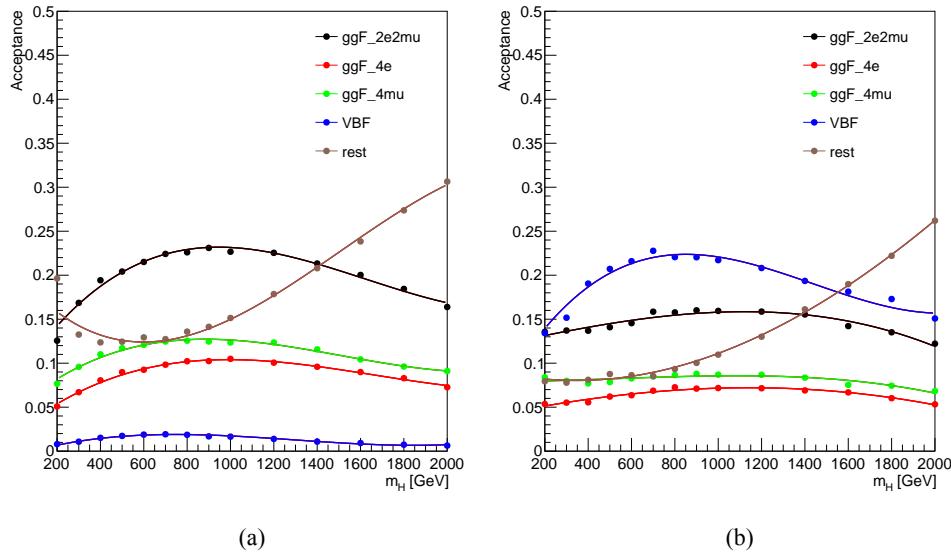


Fig. 6.7 NWA acceptance as a function of m_H for the MVA-based categorization for the samples of (a) ggF production; (b) VBF production.

2489 6.4 Background estimation

2490 In this analysis, 97% of total expected background events are from irreducible ZZ
 2491 backgrounds, which includes about 86% quark-antiquark annihilation ($q\bar{q} \rightarrow ZZ$), 10%
 2492 of gluon-induced production ($gg \rightarrow ZZ$) and around 1% of EW vector boson scattering
 2493 ($q\bar{q} \rightarrow ZZ$ EW) contribution. For $q\bar{q} \rightarrow ZZ$ EW, although it has small contribution
 2494 in total background events after analysis selection, it's important for VBF category with
 2495 about 16% contribution.

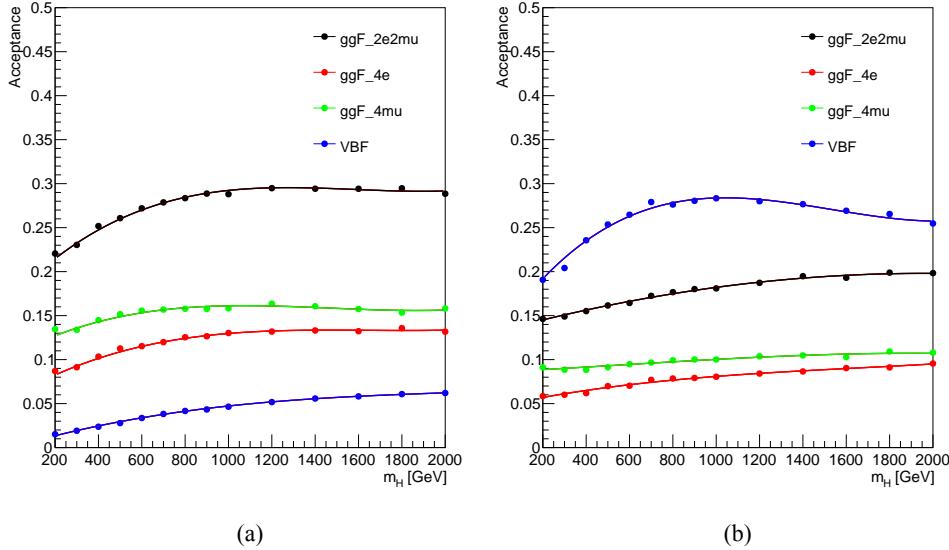


Fig. 6.8 NWA acceptance as a function of m_H for the Cut-based categorization for the samples of (a) ggF production mode; (b) VBF production mode.

2496 In addition to irreducible backgrounds, events from $Z+jets$ and $t\bar{t}$ processes, represent
 2497 as reducible backgrounds, contribute at a few percent level and can be measured using data
 2498 driven method that will be described briefly later. Additional background called ‘Others’,
 2499 including ttV and triple-V (VVV) processes, has tiny contribution and is estimated from
 2500 MC simulation directly.

2501 6.4.1 Irreducible backgrounds

2502 The Irreducible backgrounds have events with four prompt leptons. The normalization
 2503 of two dominant backgrounds $q\bar{q} \rightarrow ZZ$ and $gg \rightarrow ZZ$ are taken from data by statistical
 2504 fit, and the normalization of small $q\bar{q} \rightarrow ZZ$ EW background is measured directly from
 2505 MC simulation.

2506 The $m_{4\ell}$ shapes of all three background components are taken from MC samples and
 2507 then parameterized by an empirical function for each of them in each category respec-
 2508 tively. Details of background modellings are illustrated as below:

2509 The empirical function used for background parameterization is:

$$f(m_{4\ell}) = C_0 H(m_0 - m_{4\ell}) f_1(m_{4\ell}) + H(m_{4\ell} - m_0) f_2(m_{4\ell}), \quad (6.2)$$

where,

$$f_1(x) = \left(\frac{x - a_4}{a_3} \right)^{a_1-1} \left(1 + \frac{x - a_4}{a_3} \right)^{-a_1-a_2},$$

$$f_2(x) = \exp \left[b_0 \left(\frac{x - b_4}{b_3} \right)^{b_1-1} \left(1 + \frac{x - b_4}{b_3} \right)^{-b_1-b_2} \right],$$

$$C_0 = \frac{f_2(m_0)}{f_1(m_0)}.$$

The function consists of two parts, the first part f_1 describes the $m_{4\ell}$ spectrum in low mass region where both Z bosons decay on-shell, while the second one f_2 covers distribution at high mass tail. The transition between the low- and high- mass parts is presented in function 6.2 by the Heaviside step function $H(x)$ at the transition point m_0 . The m_0 is chosen to optimize the smoothness of the function, and practically $m_0 = 260$ (350) GeV is used for $q\bar{q} \rightarrow ZZ$ ($gg \rightarrow ZZ$ and $q\bar{q} \rightarrow ZZ$ EW). Besides, the continuity of two functions at m_0 is ensured by the factor C_0 applied to f_1 . The coefficients a_i in f_1 and b_i in f_2 are shape parameters obtained by fitting to $m_{4\ell}$ distribution from each MC simulated sample.

Figure 6.9 to 6.11 shows the fitting results of $q\bar{q} \rightarrow ZZ$, $gg \rightarrow ZZ$, $q\bar{q} \rightarrow ZZ$ EW backgrounds in four cut-based categories (ggF-CBA-enriched- $2e2\mu$, ggF-CBA-enriched- $4e$, ggF-CBA-enriched- 4μ and VBF-CBA-enriched). Figure 6.12 to 6.14 shows the fitting results of those backgrounds in five MVA-based categories (ggF-MVA-high- $2e2\mu$, ggF-MVA-high- $4e$, ggF-MVA-high- 4μ , ggF-MVA-low and VBF-MVA-enriched).

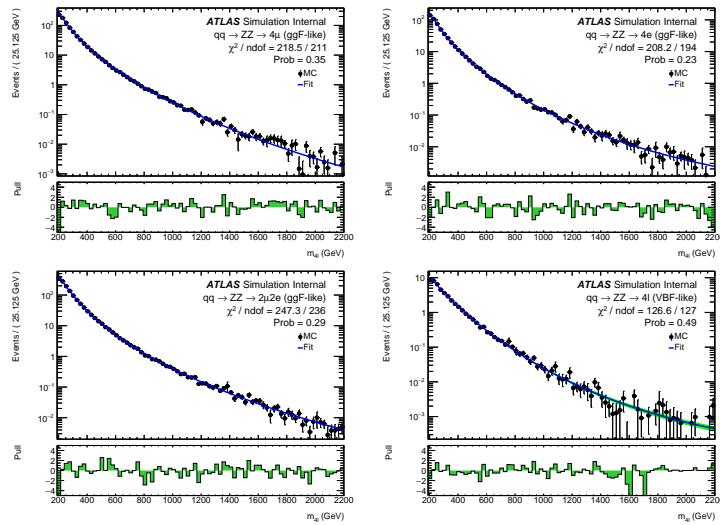


Fig. 6.9 Distributions of the $m_{4\ell}$ invariant mass fit projections of the $q\bar{q} \rightarrow ZZ$ background samples for the 4μ , $4e$ and $2\mu2e$ final states in the ggF-CBA-enriched category, and the 4ℓ inclusive VBF-CBA-enriched category. Cut-based categorization is used.

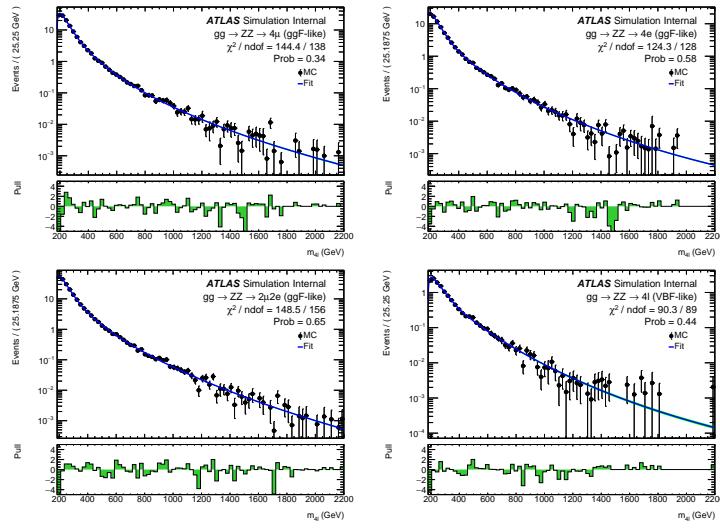


Fig. 6.10 Distributions of the $m_{4\ell}$ invariant mass fit projections of the $gg \rightarrow ZZ$ background samples for the 4μ , $4e$ and $2\mu 2e$ final states in the ggF-CBA-enriched category, and the 4ℓ inclusive VBF-CBA-enriched category. Cut-based categorization is used.

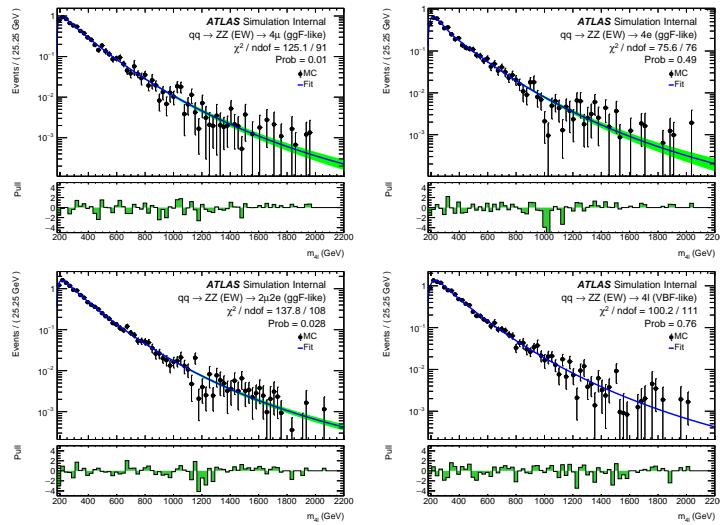


Fig. 6.11 Distributions of the $m_{4\ell}$ invariant mass fit projections of the $q\bar{q} \rightarrow ZZ$ (EW) background samples for the 4μ , $4e$ and $2\mu 2e$ final states in the ggF-CBA-enriched category, and the 4ℓ inclusive VBF-CBA-enriched category. Cut-based categorization is used.

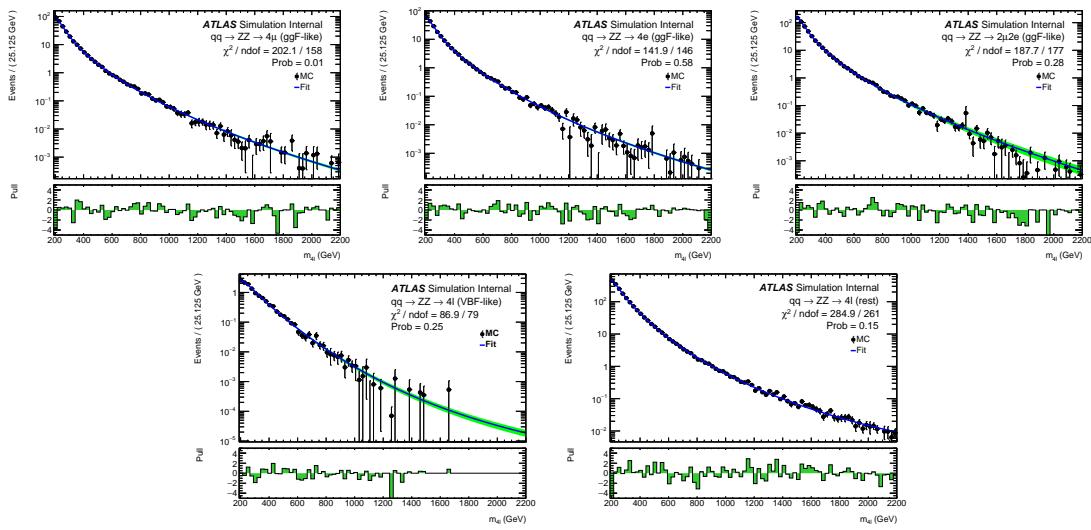


Fig. 6.12 Distributions of the $m_{4\ell}$ invariant mass fit projections of the $q\bar{q} \rightarrow ZZ$ background samples for the 4μ , $4e$ and $2\mu2e$ final states in the ggF-MVA-high category, the 4ℓ inclusive ggF-MVA-low category and VBF-MVA-enriched category. DNN-based categorization is used.

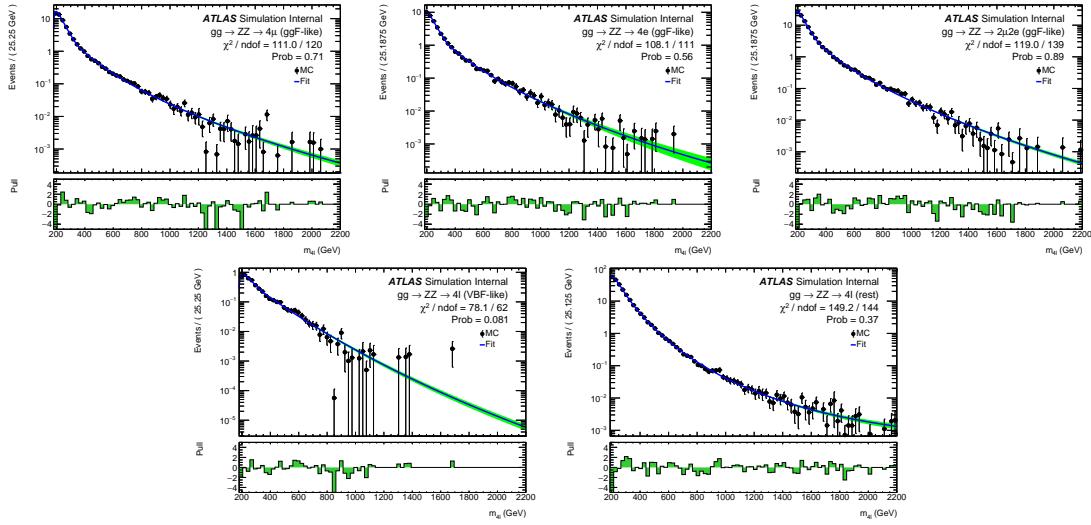


Fig. 6.13 Distributions of the $m_{4\ell}$ invariant mass fit projections of the $gg \rightarrow ZZ$ background samples for the 4μ , $4e$ and $2\mu2e$ final states in the ggF-MVA-high category, the 4ℓ inclusive ggF-MVA-low category and VBF-MVA-enriched category. DNN-based categorization is used.

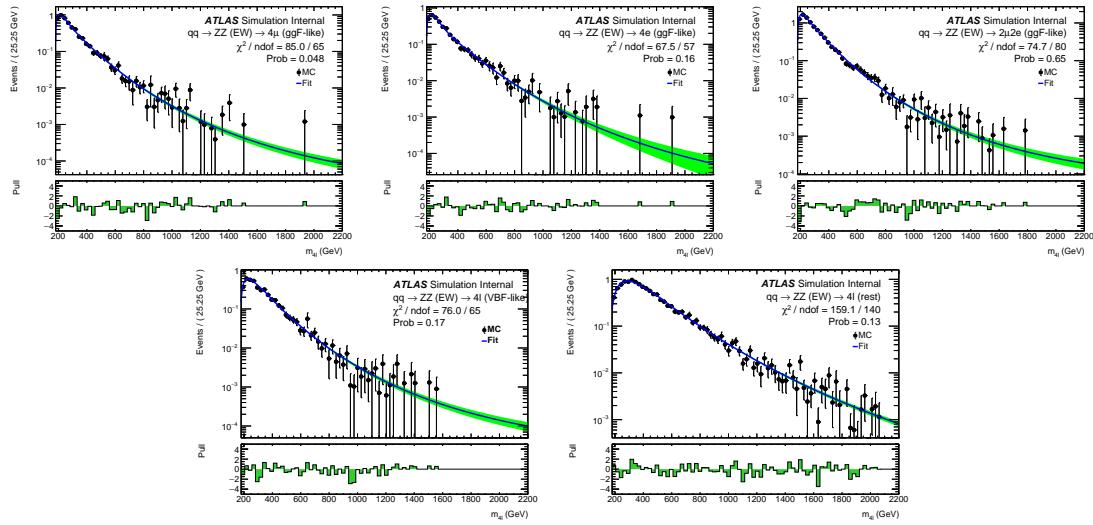


Fig. 6.14 Distributions of the $m_{4\ell}$ invariant mass fit projections of the $q\bar{q} \rightarrow ZZ$ (EW) background samples for the 4μ , $4e$ and $2\mu2e$ final states in the ggF-MVA-high category, the 4ℓ inclusive ggF-MVA-low category and VBF-MVA-enriched category. DNN-based categorization is used.

6.4.2 Reducible backgrounds

Similar as section 5.4, the reducible backgrounds include $Z+jets$ (consists of both heavy- and light-flavour jets), top quark pair, and WZ production, which contain fake and non-isolated leptons. The simulations are not very robust in terms of the selection efficiencies. Thus, the data-driven method is applied to estimated the normalization of those processes in different control regions (CRs). The estimations in this analysis are performed separately for $\ell\ell + \mu\mu$ and $\ell\ell + ee$ final states, with slightly different approaches for “muon” and “electron” backgrounds.

The “electron” backgrounds mostly come from process of a Z boson with light-flavour jets ($Z+LF$) misidentified as electrons. The large contribution of “muon” backgrounds come from heavy-flavour jets produced in association with a Z boson ($Z+HF$) or in the decays of top quark. The estimations are done following the common H4l studies without a specific $m_{4\ell}$ range requirement^[104], and then the corresponding fraction of event yield in $m_{4\ell} > 200$ GeV is calculated from MC simulation.

$\ell\ell + \mu\mu$ final states

The normalizations of “muon” backgrounds are extracted from simultaneous fits of the leading lepton pair’s invariant mass (m_{12}) in four orthogonal CRs:

- **Inverted d_0 CR:** this CR is formed by inverting the d_0 selection for at least one lepton in subleading lepton pair while the leptons in leading pair are required to pass all standard selection. This CR enhances $Z+HF$ and $t\bar{t}$ as leptons from heavy-

flavour hadronic decays are characterised by large d_0 .

- $e\mu + \mu\mu$ **CR**: this CR is formed using an opposite-charge different-flavour dilepton in leading pair. It aims to enhance $t\bar{t}$ background as the leading lepton pair cannot come from Z boson decay.
- **Inverted isolation CR**: in this CR, leptons in leading pair are required to satisfy all standard analysis selection, while for leptons in subleading pair, they are required to pass d_0 selection but have at least one of them failing isolation selection. This CR enhances the events from $Z+LF$ processes while suppress $Z+HF$ by d_0 cut.
- **Same-sign CR**: in this CR, the leptons in subleading pair are required to have same-charge, while the leading pair still passes standard selection. This CR is not dominant by any specific background since all reducible backgrounds could have sizable contribution in it.

The fit results of normalizations are then propagated to signal region (SR) by applying transfer factors to account the difference of selection efficiencies between SR and CRs.

The transfer factors are computed using $Z + \mu$ MC samples.

$\ell\ell + ee$ final states

The “electron” backgrounds are estimated in $3\ell + X$ CR, where X denotes the lower p_T electron in the subleading pair. The selection and identification criterias for X are relaxed , while other three leptons must satisfy the standard selection. In this case, X could be a light-flavour jet, a photon conversion or an electron from heavy-flavour hadron decay. Moreover, the subleading pair is required to have same charge dilepton to ensure the orthogonality to the signal region. The normalization of backgrounds are obtained based on a fit to the number of hits in the innermost ID layer in CR, and the transfer factors are computed from $Z + e$ simulated sample.

The $m_{4\ell}$ shapes of reducible backgrounds are obtained from MC simulation in signal region, and then smoothed by an one-dimensional kernel estimation, which models the input data as a superposition of Gaussian kernels, one for each data point with contributing $1/N$ to total integral N ^[105]. The difference from using different smoothing strength (ρ) in kernel estimation is taken into account as additional shape uncertainties for these reducible backgrounds.

6.5 Signal modelling

The parameterization of $m_{4\ell}$ distributions based on simulated samples for signals are described in this section. Several signal models are studied, including heavy Higgs like

2577 narrow-width signal (NWA) and large-width signal (LWA), as well as the modelling of
2578 Randall-Sundrum graviton (RSG) signal.

2579 6.5.1 Modelling of narrow-width signal

2580 For narrow-width (NWA) signal, the $m_{4\ell}$ width is totally determined by detector reso-
2581 lution, which is modelled by the sum of a Crystal Ball (C) function^[106-107] and a Gaussian
2582 (G) function:

$$P_s(m_{4\ell}) = f_C \cdot C(m_{4\ell}; \mu, \sigma_C, \alpha_C, n_C) + (1 - f_C) \cdot G(m_{4\ell}; \mu, \sigma_G) \quad (6.3)$$

2583 The two functions share the same central value μ , while the resolution parameters, σ_C
2584 and σ_G , are different. In the Crystal Ball function, the parameters α_C and n_C model the
2585 shape of non-Gaussian tail, and the fraction parameter f_C is used to ensure the relative
2586 normalization between two functions.

2587 The parameters are obtained by fitting to signal MC simulations combining the mc16a,
2588 mc16d and mc16e campaigns for each category at each mass points from 200 GeV to
2589 2000 GeV respectively, and the shape of ggF and VBF signals are found to be similar.
2590 Figure 6.15 shows the $m_{4\ell}$ distribution and fitted curves for ggF production at mass from
2591 200 GeV to 2000 GeV in $2e2\mu$ channel as examples.

2592 Then the $C + G$ parameters are fitted with a polynomial function as the function of
2593 generated mass points (m_H), as an example shown in figure 6.16 for $2e2\mu$ channel. The
2594 fitting quality can be measured by the Pearson's χ^2 , which is within 3 (2) for $2e2\mu$ ($4e$
2595 and 4μ) channel.

2596 In addition, possible difference on the signal yield extracted from parameterization and
2597 MC simulation is studied. Figure 6.17 shows this difference by computing $\frac{N_{\text{reco}} - N_{\text{fit}}}{N_{\text{fit}}}$, where
2598 N_{reco} denotes the total number of reconstructed events observed from MC simulation at
2599 that mass point and N_{fit} depicts the number of events obtained from the fitted PDF. The
2600 differences are treated as an additional systematic uncertainty with the value of 2% (1%)
2601 for $2e2\mu$ ($4e$ and 4μ) channel in statistical fit.

2602 In summary, the final interpolated signal shapes for the ggF production mode are
2603 shown together in figure 6.18 for mass points with step of 100 GeV from 200 GeV to
2604 3000 GeV.

2605 6.5.2 Modelling of large-width signal

2606 The $m_{4\ell}$ shape of heavy Higgs model in large-width (LWA) hypothesis can be de-
2607 scribed by a convolution of a truth distribution and a resolution from detector effect. The

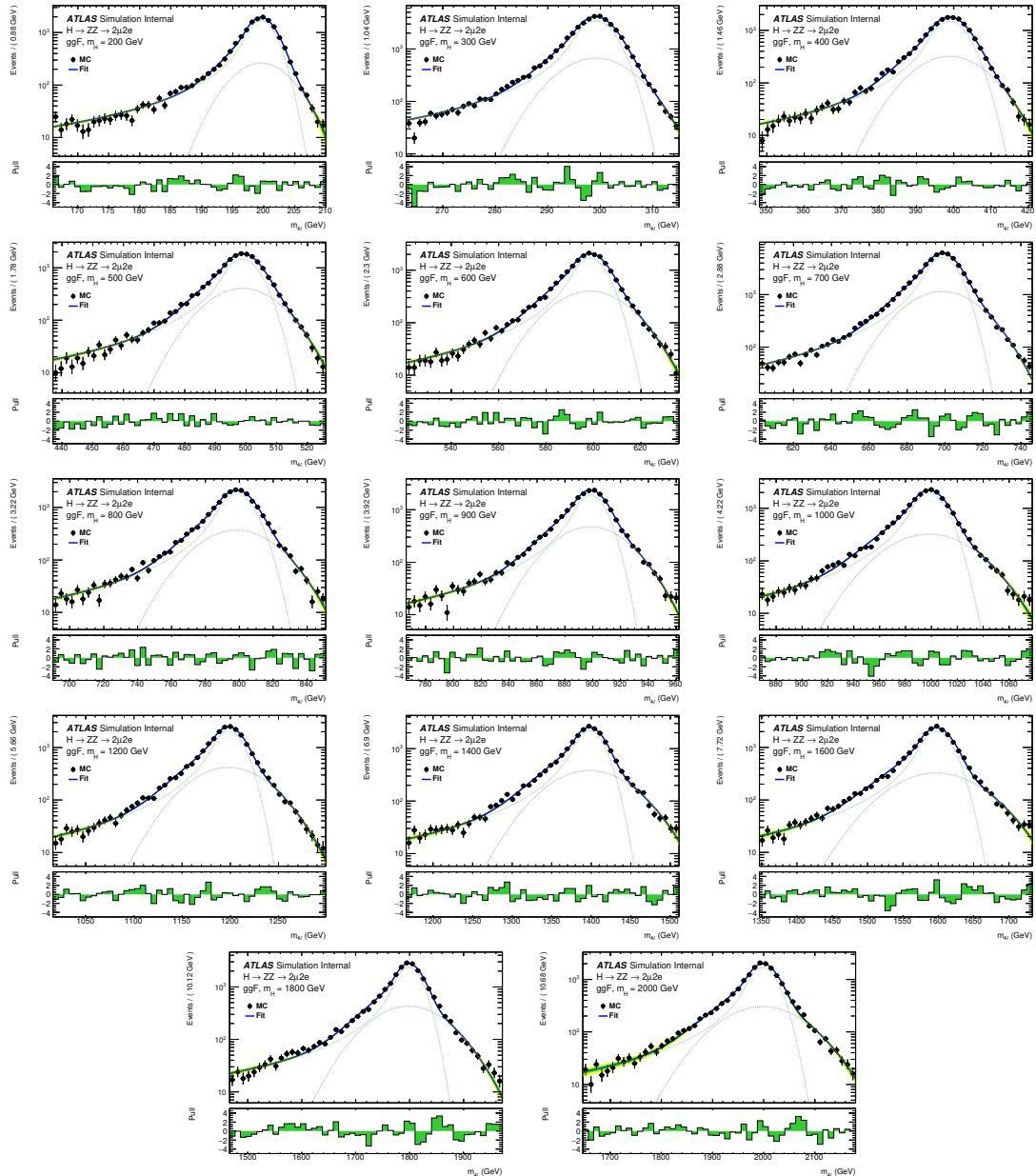


Fig. 6.15 Distributions of the $m_{2\mu 2e}$ and fit projection for signal samples between 200 to 3000 GeV for ggF production mode. Three MC campaigns, mc16a, mc16d and mc16e, are combined. The lower panel in each plot shows the pull distribution.

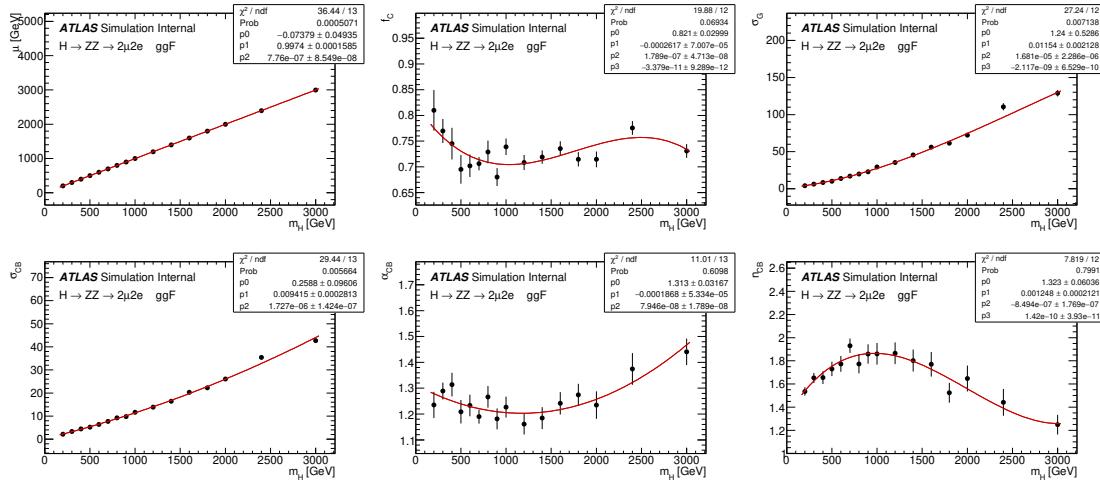


Fig. 6.16 Polynomial fits of the parameters μ , f_C , σ_G , σ_C , n_C and α_C for the signal $C + G$ model in the $2\mu 2e$ channel as a function of m_H for the ggF production mode. The combination of the mc16a, mc16d and mc16e MC campaigns is used.

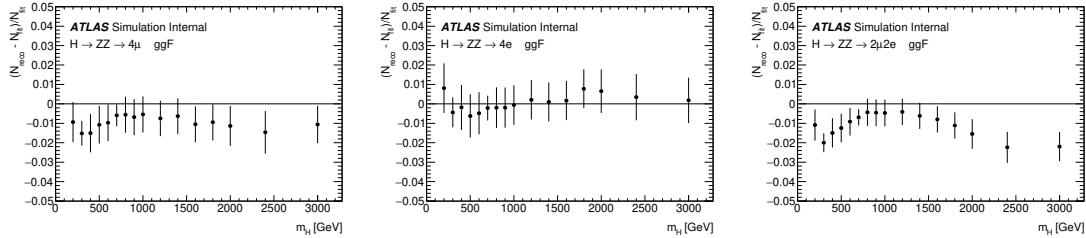


Fig. 6.17 The difference between MC simulation and parameterization of 4μ (left), $4e$ (middle) and $2\mu 2e$ (right) for the ggF production mode. The combination of the mc16a, mc16d and mc16e MC campaigns is used.

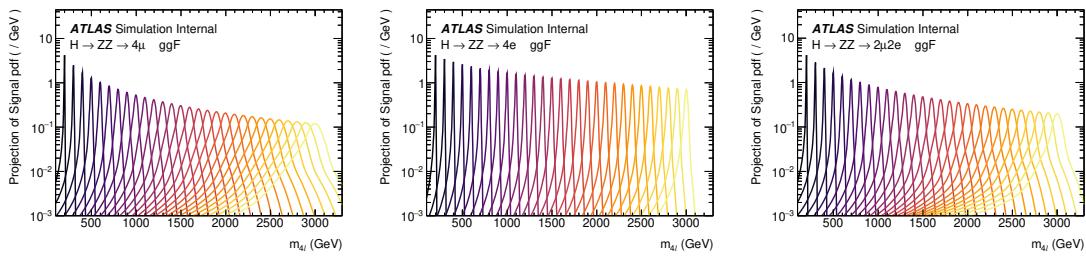


Fig. 6.18 The final signal shapes for the ggF production mode, interpolated from the polynomial fit parameters.

2608 detector resolution effect is the one modelled by the function described in NWA parameterization, as in NWA model the truth level width is negligible.

2610 The differential parton cross section for the heavy Higgs model can be written as^[108]:

$$\sigma_{gg \rightarrow H \rightarrow ZZ}(s) = \frac{1}{2s} \int d\Omega |A_{gg \rightarrow H}(s, \Omega)|^2 \frac{1}{|s - s_H|^2} |A_{H \rightarrow ZZ}(s, \Omega)|^2 \quad (6.4)$$

2611 where $A_{gg \rightarrow H}(s, \Omega)$ and $A_{H \rightarrow ZZ}(s, \Omega)$ are corresponding Higgs production and decay
2612 amplitudes, and $\frac{1}{|s - s_H|}$ denotes the Higgs propagator and Ω represents the phase space of
2613 the process.

2614 Using the definition of a partial width,

$$\Gamma_{H \rightarrow F}(s) = \frac{1}{2\sqrt{s}} \int d\Omega |A_{H \rightarrow F}(s, \Omega)|^2 \quad (6.5)$$

2615 the parton cross section can be rewritten as,

$$\sigma_{gg \rightarrow H \rightarrow ZZ}(s) = 2 \frac{1}{|s - s_H|^2} \times \Gamma_{H \rightarrow gg}(s) \times \Gamma_{H \rightarrow ZZ}(s) \quad (6.6)$$

with the components computed in Ref^[108-109]:

$$\begin{aligned} \frac{1}{s - s_H} &= \frac{1 + i \cdot \bar{\Gamma}_H / m_H}{s - \bar{m}_H^2 + i \cdot s \cdot \bar{\Gamma}_H / m_H} \\ \bar{m}_H &= \sqrt{\bar{\Gamma}_H^2 + m_H^2} \\ \bar{\Gamma}_H &= \bar{m}_H \cdot \frac{\Gamma_H}{m_H} \end{aligned} \quad (6.7)$$

$$\Gamma_{H \rightarrow ZZ}(s) = C \cdot s^{\frac{3}{2}} \cdot \left[1 - \frac{4m_Z^2}{s} + \frac{3}{4} \left(\frac{4m_Z^2}{s} \right)^2 \right] \cdot \left[1 - \frac{4m_Z^2}{s} \right]^{\frac{1}{2}} \quad (6.8)$$

$$\begin{aligned} \Gamma_{H \rightarrow gg}(s) &= C \cdot s^{\frac{3}{2}} \cdot |A_t(\tau_t)|^2 \\ A_t(\tau) &= 2 \frac{\tau + (\tau - 1)f(\tau)}{\tau^2} \\ \tau_t &= \frac{s}{4m_t^2} \\ f(\tau) &= \begin{cases} \arcsin^2(\sqrt{\tau}), & \tau \leq 1 \\ -\frac{1}{4} \left[\log \frac{1+\sqrt{1-\tau^{-1}}}{1-\sqrt{1-\tau^{-1}}} - i\pi \right]^2, & \tau > 1 \end{cases} \end{aligned} \quad (6.9)$$

2616 where m_f stands for the mass of a fermion f , and Γ_H denotes an assumed total width
2617 of the heavy Higgs boson.

At the LHC, the $m_{4\ell}$ line shape can be defined by a hadron cross section that is derived from equation 6.6 by multiplication with gluon-gluon luminosity \mathcal{L}_{gg} described in^[110]. Meanwhile, the cross section is rewritten as a function of $m_{4\ell}$ instead of s , which will give an extra power of mass dependence in the formula:

$$\sigma_{pp \rightarrow H \rightarrow ZZ}(m_{4\ell}) = 2 \cdot m_{4\ell} \cdot \mathcal{L}_{gg} \cdot \frac{1}{|s - s_H|^2} \cdot \Gamma_{H \rightarrow gg}(m_{4\ell}^2) \cdot \Gamma_{H \rightarrow ZZ}(m_{4\ell}^2) \quad (6.10)$$

The analytical shapes of truth level $m_{4\ell}$ distribution of gg2VV MC samples is shown on figure 6.19.

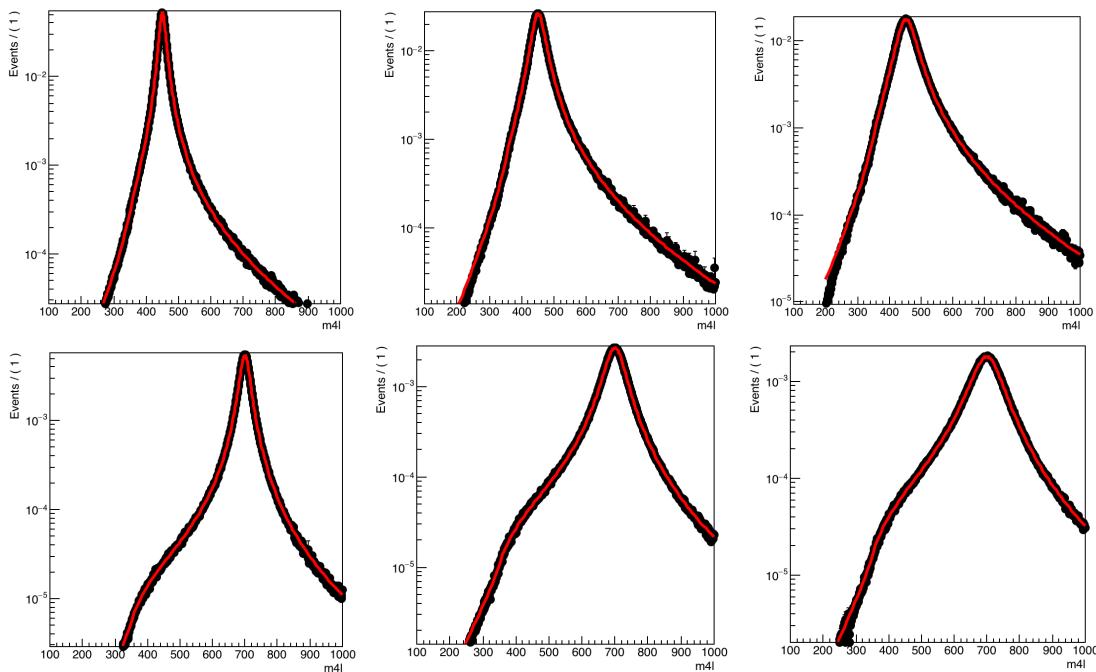


Fig. 6.19 Comparison of the analytical shape to a truth $m_{4\ell}$ distribution of gg2VV MC samples for $m_H = 450$ GeV (top), 700 GeV (bottom) and width equal to 5% (left), 10% (middle), 15% (right) of the mass.

The reconstruction level signal shape can then be modelled by the analytical truth shape convoluted with detector effects modelled in section 6.5.1. A comparison between the modelled shape and reconstruction level MC simulation for signal mass above 400 GeV (for ggF production in $2e2\mu$ channel as an example) are shown in figure 6.20, the shapes are well compatible between each other. This modelling is not valid for lower masses due to the rapid change of detector resolution.

6.5.3 Modelling of interference

There are three processes sharing the same gg initial state and ZZ final state:

- The SM $gg \rightarrow ZZ$ process with an amplitude A_B

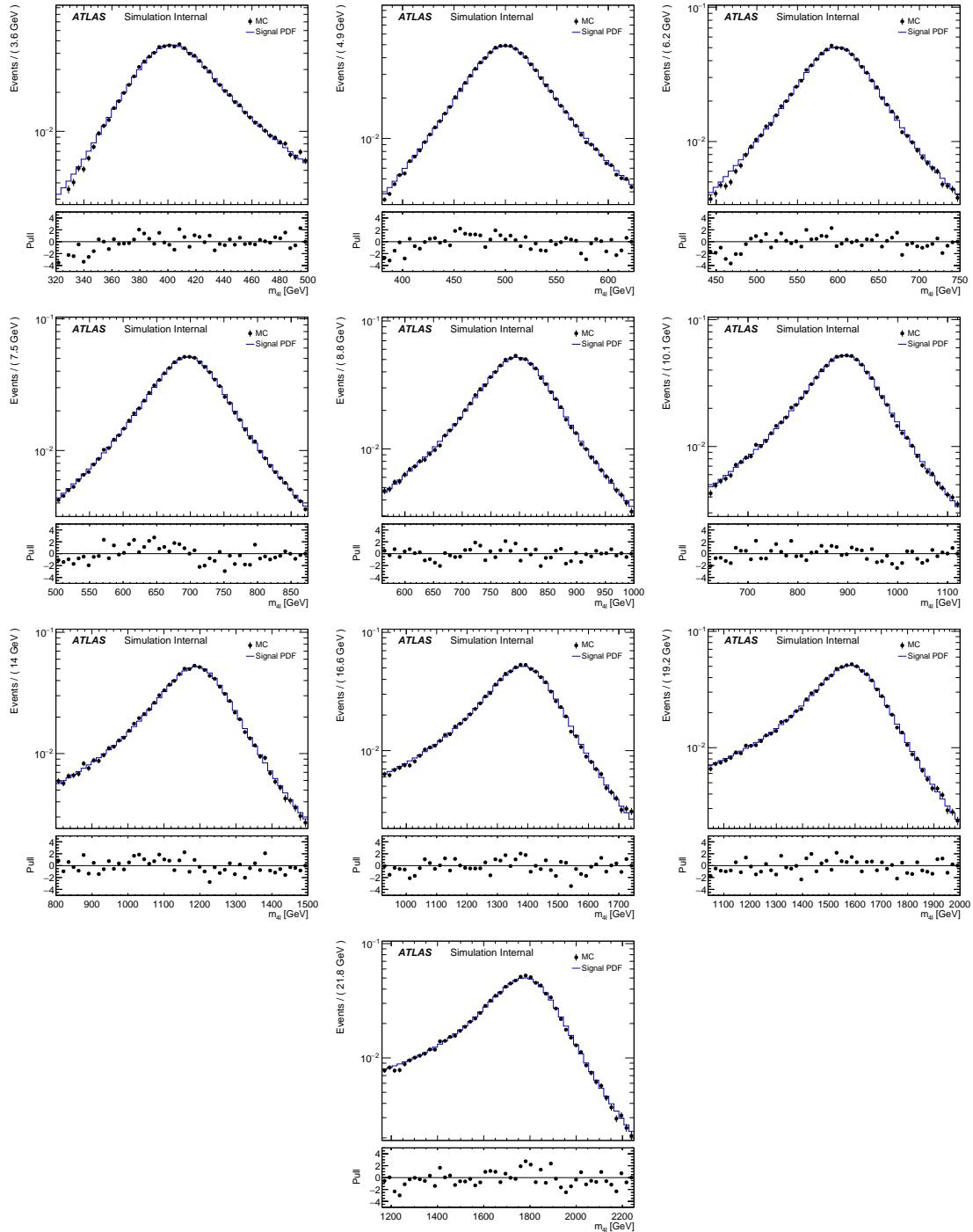


Fig. 6.20 Comparison between the analytical shape convoluted with detector effects and the reconstructed $m_{2\mu 2e}$ MC distribution for mass points ranging from 400 to 1800 GeV and width equal to 15% of the mass.

- 2633 • The SM (light) Higgs at mass of around 125 GeV with an amplitude A_h
 - 2634 • The BSM heavy Higgs we are searching in this analysis with an amplitude A_H
- 2635 The three processes can interfere with each other due to the same initial and final states.
- 2636 The parton cross section for these processes can be written as:

$$\begin{aligned}
 \sigma_{gg \rightarrow (X) \rightarrow ZZ}(s) &= \frac{1}{2s} \int d\Omega |A_h(s, \Omega) + A_H(s, \Omega) + A_B(s, \Omega)|^2 \\
 &= \frac{1}{2s} \int d\Omega \left(|A_h(s, \Omega)|^2 + |A_H(s, \Omega)|^2 + |A_B(s, \Omega)|^2 \right) + \\
 &\quad + \frac{1}{s} \int d\Omega (Re [A_h(s, \Omega) \cdot A_B^*(s, \Omega)] \\
 &\quad + Re [A_H(s, \Omega) \cdot A_B^*(s, \Omega)] + Re [A_H(s, \Omega) \cdot A_h^*(s, \Omega)]) \\
 &\quad + \frac{1}{s} \text{Re} \left[\frac{1}{s - s_H} \int d\Omega \cdot A_H^P(s, \Omega) \cdot A_H^D(s, \Omega) \cdot A_B^*(s, \Omega) \right] \\
 &\quad + \frac{1}{s} \int d\Omega \cdot \text{Re} \left[A_H^P(s, \Omega) \cdot \frac{1}{s - s_H} \cdot A_H^D(s, \Omega) \cdot A_h^{P*}(s, \Omega) \cdot \frac{1}{(s - s_h)^*} \cdot A_h^{D*}(s, \Omega) \right]
 \end{aligned} \tag{6.11}$$

2637 The first term in equation 6.11 denotes the on-shell SM Higgs contribution, which is
 2638 negligible in this analysis. The second term corresponds to the heavy Higgs contribution,
 2639 whose line shape has been described in previous section. The third term is the $gg \rightarrow ZZ$
 2640 continuum process, while the forth term is the interference between SM Higgs and $gg \rightarrow$
 2641 ZZ continuum. The fifth and sixth terms are the interferences between heavy Higgs and
 2642 $gg \rightarrow ZZ$ continuum (H-B), and between heavy Higgs and SM Higgs (H-h) that we
 2643 are interested in. More details about the parameterization of these two interferences are
 2644 described as below.

2645 **1. Interference between heavy Higgs and $gg \rightarrow ZZ$ continuum**

2646 The parton cross section of this interference term has been written down in equa-
 2647 tion 6.11. By assuming that this function has a smooth behaviour, it can be replaced with
 2648 complex polynomial:

$$\int d\Omega \cdot A_H^P(s, \Omega) \cdot A_H^D(s, \Omega) \cdot A_B^*(s, \Omega) \approx (a_0 + a_1 \cdot \sqrt{s} + \dots) + i \cdot (b_0 + b_1 \cdot \sqrt{s} + \dots) \tag{6.12}$$

2649 The parameters a_i and b_i can be extracted by fitting to the $m_{4\ell}$ distribution from truth
 2650 level MC simulation after analysis selection. Since the signal mass and width does not
 2651 enter into this function, the parameters should be independent for every tested signal hy-
 2652 pothesis.

2653 Same as description for equation 6.10, the parton cross section can be transformed

2654 into a hadron cross section as a function of $m_{4\ell}$:

$$\sigma_{pp}(m_{4\ell}) = \mathcal{L}_{gg} \cdot \frac{1}{m_{4\ell}} \cdot \text{Re} \left[\frac{1}{s - s_H} \cdot ((a_0 + a_1 \cdot m_{4\ell} + \dots) + i \cdot (b_0 + b_1 \cdot m_{4\ell} + \dots)) \right] \quad (6.13)$$

2655 where the propagators are shown in equation 6.7.

2656 Figure 6.21 shows the distributions of interference function obtained by simultaneous
2657 fitting to $m_{4\ell}$ shape from truth level H-B interference simulation at different mass in $2e2\mu$
2658 channel as an example.

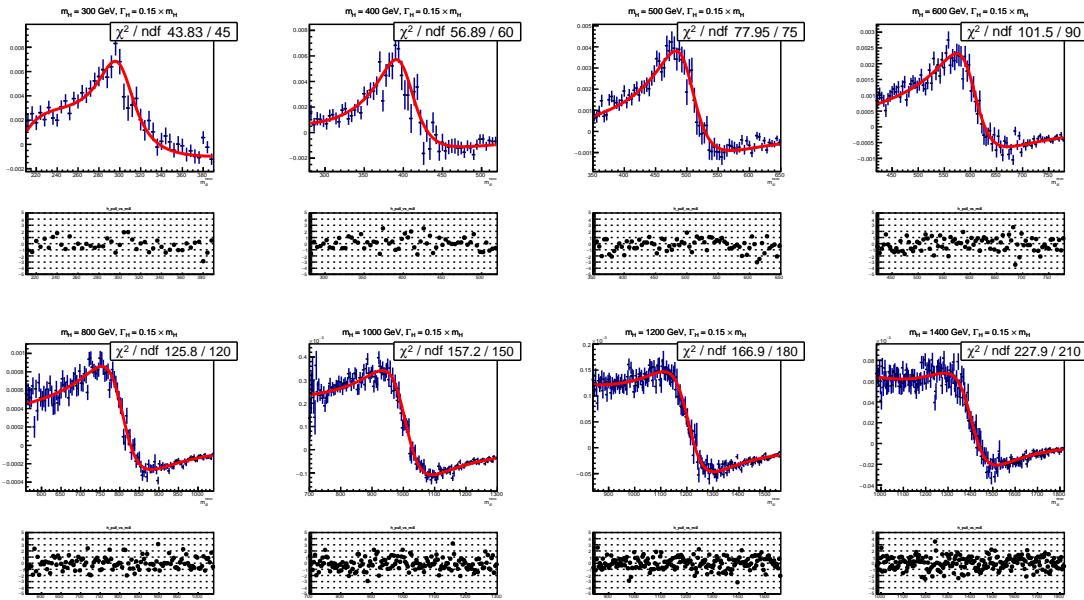


Fig. 6.21 The interference (H-B) model fitted to the truth $m_{4\ell}$ MC distribution after signal region selection for $2\mu 2e$ channel.

2. Interference between heavy Higgs and SM Higgs

2659 The parton cross section of this interference term has been written down in equa-
2660 tion 6.11. By assuming the production and decay amplitudes are the same for heavy
2661 Higgs boson and SM Higgs boson, the cross section function can be simplified to:

$$\sigma_{gg}(s) = \frac{1}{s} \int d\Omega \cdot \text{Re} \left[\frac{1}{s - s_H} \cdot \frac{1}{(s - s_h)^*} \right] \cdot |A_{gg \rightarrow H}(s, \Omega)|^2 |A_{H \rightarrow ZZ}(s, \Omega)|^2 \quad (6.14)$$

2663 Taking into account Equation 6.5:

$$\sigma_{gg}(s) = 4 \cdot \text{Re} \left[\frac{1}{s - s_H} \cdot \frac{1}{(s - s_h)^*} \right] \cdot \Gamma_{H \rightarrow gg}(s) \cdot \Gamma_{H \rightarrow ZZ}(s) \quad (6.15)$$

2664 where the propagators are described in equation 6.7, and the partial widths are de-
2665 scribed in equations 6.8 and 6.9.

2666 Same as previous procedure, the parton cross section can be transformed to a hadron
 2667 cross section as a function of $m_{4\ell}$:

$$\sigma_{pp}(m_{4\ell}) = 4 \cdot m_{4\ell} \cdot \mathcal{L}_{gg} \cdot \text{Re} \left[\frac{1}{s - s_H} \cdot \frac{1}{(s - s_h)^*} \right] \cdot \Gamma_{H \rightarrow gg}(m_{4\ell}) \cdot \Gamma_{H \rightarrow ZZ}(m_{4\ell}) \quad (6.16)$$

2668 The modelling procedure of interference is the same as the way for large-width signal
 2669 described in section 6.5.2. The truth line shape is measured as analytical function from
 2670 equation 6.16, and then convolute with detector effect from NWA parameterization to get
 2671 the reconstruction level shape.

2672 For LWA signal model, these two interferences are carefully token into account, and
 2673 the integration of the pure LWA signal with the interferences is used for further studies.
 2674 Figure 6.22 shows the signal model for large-width scenario at mass points of 400 GeV,
 2675 600 GeV, 800 GeV, for three different signal widths: 5%, 10%, 15%, with and without
 2676 interference. Additionally, the contribution of the interference between heavy Higgs and
 2677 SM Higgs (H-h) is shown together with the one between heavy Higgs and SM $gg \rightarrow ZZ$ background (H-b). One can see the interference effect on signal shape becomes less
 2678 important when going to higher mass.
 2679

2680 6.5.4 Modelling of spin-2 RS Graviton signal

2681 The search for Randall-Sundrum (RS) graviton is performed in mass region between
 2682 600 to 2000 GeV. The width of resonance is determined by the $k/\overline{M}_{\text{Planck}}$, which, as
 2683 mentioned in section 6.2.3, is set to be 1. In this configuration, the width of signal is
 2684 expected to be about 6% of its mass.

The reconstructed $m_{4\ell}$ lineshape of graviton is also built by convolving the truth-level lineshape with a detector resolution function, where the detector resolution effect is modelled by a Gaussian + Crystal Ball function, whose parameters are taken from the NWA signal parameterization in section 6.5.1. And for truth-level shape, for graviton, it's modelled as the product of a relativistic Breit-Wigner (RBW) term, a term corresponding to the squared matrix element of the production process and a parton luminosity term \mathcal{L} as given in [111]. So the truth lineshape of $m_{4\ell}$ is token from:

$$m_{4\ell}^{\text{Truth}} \sim \mathcal{L}_{gg} \cdot s^2 \cdot \frac{s(1+s)(1+2s+2s^2)}{(s^2 - m_G^2)^2 + m_G^2 \Gamma^2}$$

2685 The truth-level signal model is extracted by fitting to MC simulation at truth-level with
 2686 the mass m_G and width Γ parameters floating at each mass points respectively. And then

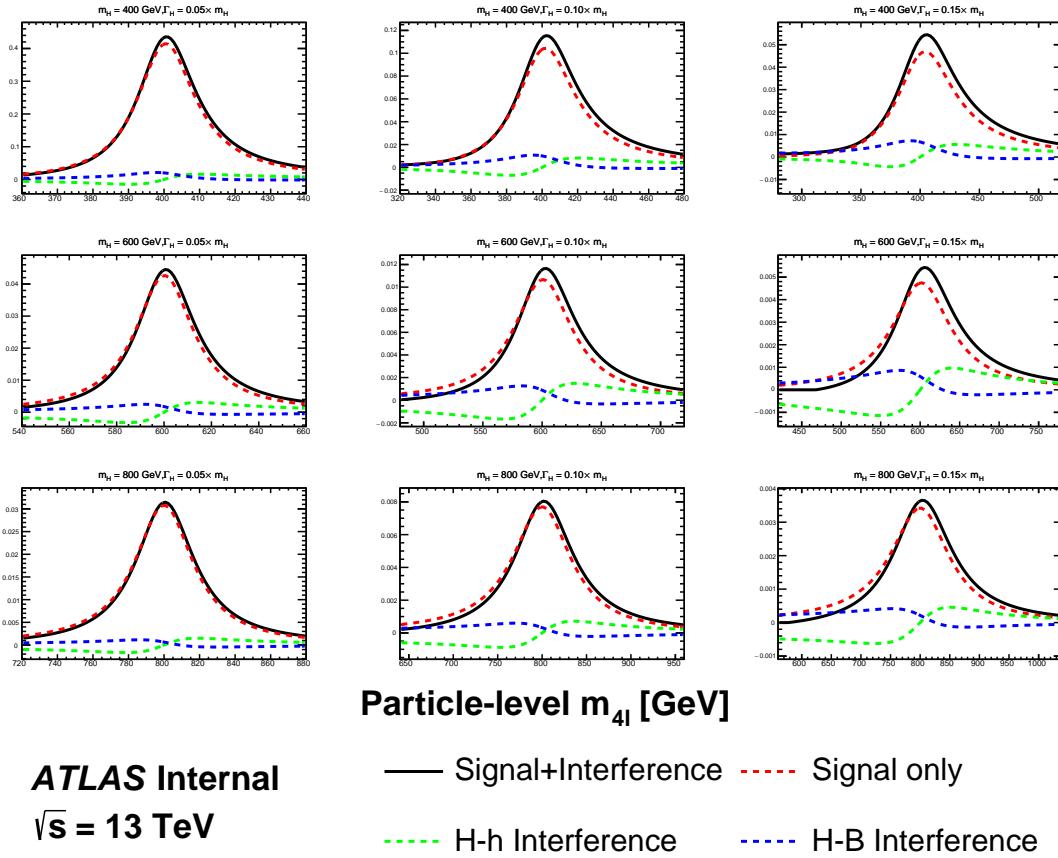


Fig. 6.22 The signal modelling for the large-width scenario at m_H of 400 GeV (top), 600 GeV (middle) and 800 GeV (bottom), as well as three different signal width: 5% (left), 10% (middle) and 15% (right). The contribution of the interference between heavy Higgs and SM Higgs (H-h) is shown together with the one between heavy Higgs and SM $gg \rightarrow ZZ$ background (H-b).

2687 the two parameters are parameterized as the function of m_H by a linear fit as shown in
 2688 figure 6.23.

2689 The final signal model is obtained by convolving the truth-level lineshape with the
 2690 detector resolution function. To verify the result, figure 6.24 compares the $m_{4\ell}$ lineshape
 2691 from parameterization with the one observed from reconstructed-level MC simulation in
 2692 $2e2\mu$ channel at masses of 600 GeV, 1600 GeV and 2000 GeV as examples.

2693 6.6 Systematic uncertainties

2694 This section describes the sources and value of theoretical and experimental system-
 2695 atic uncertainties considered in this analysis. In addition, as mentioned in previous sec-
 2696 tions, the uncertainties of irreducible background modelling, reducible background shape
 2697 smoothing procedure and signal yield difference between simulation and parameterization

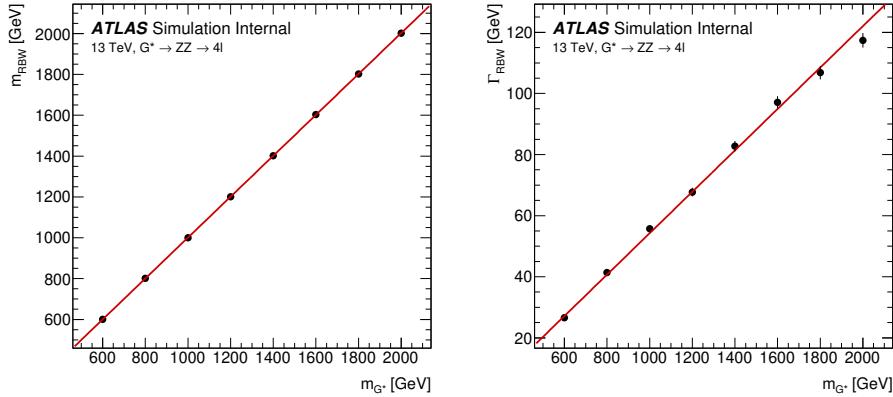


Fig. 6.23 Fitted parameters of the graviton RBW, m_{RBW} and Γ_{RBW} , as a function of the graviton resonance mass, m_G .

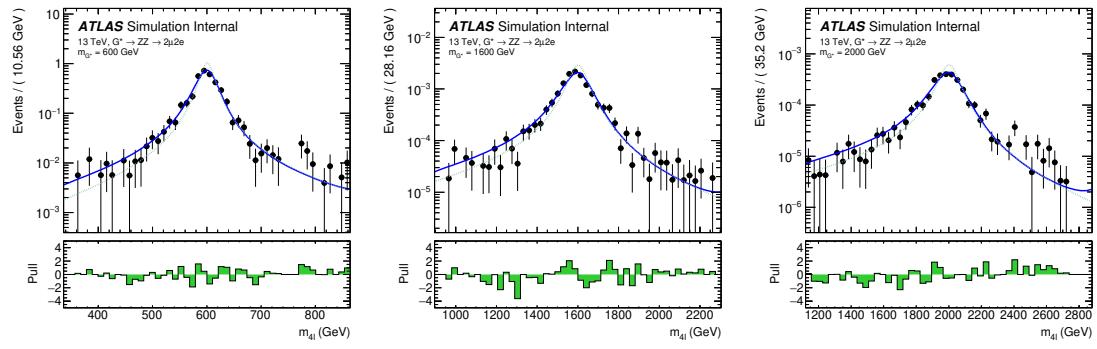


Fig. 6.24 Reconstructed $m_{4\ell}$ distributions in the $2\mu 2e$ channel with the final signal model superimposed for each RS graviton signal sample at masses of 600 GeV, 1600 GeV and 2000 GeV. The lower panel in each plot shows the pull distribution. The dashed green lines show the truth-level graviton signal models for reference.

2698 are all taken into account.

2699 6.6.1 Theoretical uncertainties

2700 The theoretical modelling uncertainties include the PDF variations, missing QCD
2701 higher-order corrections via the variations of factorisation and renormalization scales, and
2702 the parton showering uncertainties.

2703 1. Theoretical uncertainties for signal

2704 The PDF, QCD scale and parton showering uncertainties affecting the acceptance dif-
2705 ference originating from analysis selection for signal are taken into account in different
2706 categories. The acceptance uncertainties are calculated on the acceptance factor which
2707 extrapolates from the fiducial space to the full phase space by a simple ratio:

$$A = \frac{N_{\text{fiducial}}}{N_{\text{total}}} \quad (6.17)$$

2708 For PDF uncertainties, the standard derivations of 100 PDF replicas of NNPDF3.0
 2709 NNLO, as well as comparison to two external PDF sets: MMHT2014 NNLO, CT14
 2710 NNLO are considered. For missing QCD higher-order corrections, the effects are studied
 2711 with truth events by comparing weights corresponding to variations of the renormalization
 2712 and factorization scale factors, up and down by a factor of two, and the envelop of dif-
 2713 ferent variations is used. The parton showering uncertainties are estimated by comparing
 2714 events with different setting via Pythia8.

2715 Systematic uncertainties are studied for both cut- and MVA- based event catego-
 2716 rizations, in two different categories: the inclusive ggF-CBA-enriched and VBF-CBA-
 2717 enriched category for cut-based analysis and in three different categories: inclusive ggF-
 2718 MVA-high, ggF-MVA-low and VBF-MVA-enriched category for MVA-based one. This
 2719 section shows the MVA-based results as an example.

2720 Table 6.10 and 6.11 show the theoretical uncertainties mentioned above for ggF and
 2721 VBF signal respectively in MVA-based categorization.

**Table 6.10 Summary of acceptance uncertainties of PDF, QCD scale and parton shower vari-
 ations for ggF production. The MVA-based categorization is used.**

Categories	PDF	QCD Scale	Parton Shower
ggF-MVA-high	0.40%	0.06%	2.03%
ggF-MVA-low	0.56%	0.07%	4.86%
VBF-MVA-enriched	0.53%	0.09%	3.43%

**Table 6.11 Summary of acceptance uncertainties of PDF, QCD scale and parton shower vari-
 ations for VBF production. The MVA-based categorization is used.**

Categories	PDF	QCD Scale	Parton Shower
ggF-MVA-high	0.18%	1.20%	0.41%
ggF-MVA-low	0.43%	0.26%	0.36%
VBF-MVA-enriched	0.23%	3.19%	0.85%

2722 **2. Theoretical uncertainties for SM background processes**

2723 The theoretical uncertainties of irreducible ZZ backgrounds are considered in terms
 2724 of both the variations of shape of $m_{4\ell}$ distributions and the acceptance originating from
 2725 the event selection.

2726 The PDF and QCD scale uncertainties are considered by using the same method as
 2727 described for signal. The parton showering uncertainties for those Sherpa samples are
 2728 evaluated by varying the resummation scale by a factor of 2, changing the CKKW setting
 2729 and using different showering option, following the PMG recommendation in ref.^[112],
 2730 and the quadratic sum between the uncertainties in different kinds of showering option

2731 is taken as final result of uncertainties. Moreover, the shape uncertainty associated with
 2732 electroweak higher-order correction for $q\bar{q} \rightarrow ZZ$ process is also taken into account.

2733 Same as for signals, these theoretical uncertainties for irreducible backgrounds are
 2734 studied for both cut- and MVA- based event categorizations. The value of shape uncer-
 2735 tainties vary from less than 1% at low mass region to 50% at high mass tail due to large
 2736 statistic fluctuation. As for the acceptance uncertainties, the values vary from about 1%
 2737 for PDF variations to 40% for parton showering variations. The VBF category has relative
 2738 larger uncertainties.

2739 Table 6.12 summarizes the acceptance uncertainties of PDF, QCD scale, and parton
 2740 showering variations for the dominant background: $q\bar{q} \rightarrow ZZ$.

**Table 6.12 Summary of acceptance uncertainties of PDF, scale, and parton showering varia-
 tions for QCD $q\bar{q} \rightarrow ZZ$ background. The MVA-based categorization is used.**

Categories	PDF	QCD Scale	Parton showering
ggF-MVA-high	1.15%	10.16 %	3.71%
ggF-MVA-low	1.04%	3.26 %	3.80%
VBF-MVA-enriched	2.91%	27.90 %	23.82%

2741 6.6.2 Experimental systematics

2742 The signal and background predictions used in this analysis are also affected by
 2743 various sources of experimental systematic uncertainties. Similar as described in sec-
 2744 tion 5.5.2, the dominant experimental uncertainties in this analysis come from the en-
 2745 ergy/momenta scales and reconstruction and identification efficiencies of the leptons
 2746 and jets, as well as the luminosity uncertainty. The systematic uncertainties are calculated
 2747 using the recommendations from the Combined Performance (CP) groups of ATLAS ex-
 2748 periment. Table 6.13 summarizes the experimental systematics considered in this analysis
 2749 that affect either the normalization of total event yield or the shape of $m_{4\ell}$ distribution.
 2750 The impact of those systematics in statistical fit are studied in section 6.7.

2751 6.7 Results in $\ell\ell\ell\ell$ channel

2752 The statistical treatment in searching for heavy resonances in $ZZ \rightarrow \ell\ell\ell\ell$ final state
 2753 is described in this section. Results are presented in both cut- and MVA- based analysis.

2754 6.7.1 Statistical procedure

2755 The upper limits on heavy resonances are obtained using the unbinned profile likelihood fits. $m_{4\ell}$ is the discriminant. The likelihood function is a product of a Poisson term
 2756 representing the probability of observing n events and a weighted sum of both signal and
 2757 background probability distribution functions (PDFs) evaluated at all observed events.
 2758

$$L(x_1..x_n|\sigma_{ggF}, \sigma_{VBF}) = \text{Pois}(n|S_{ggF} + S_{VBF} + B) \left[\prod_{i=1}^n \frac{S_{ggF}f_{ggF}(x_i) + S_{VBF}f_{VBF}(x_i) + Bf_B(x_i)}{S_{ggF} + S_{VBF} + B} \right] \quad (6.18)$$

2759 where f_X s are the probability distribution functions of signal and backgrounds modelled
 2760 in section 6.5 and 6.4, S_X and B are the normalizations of signal and sum of backgrounds.

2761 The parameters of interest (POI) in the search is σ_{ggF} (and σ_{VBF} only for NWA signal),
 2762 which is the cross section of signal model in ggF (and VBF) production mode. In the case
 2763 of there are two POIs, when testing one POI, the other one is profiled along with other
 2764 nuisance parameters (except left unconstrained) during the minimization. These POIs
 2765 enter the likelihood inside the expected signal yields S_{ggF} and S_{VBF} as:

$$S_{ggF(VBF)} = \sigma_{ggF(VBF)} \times BR(S \rightarrow ZZ) \times A \times C \times \int \mathcal{L} \quad (6.19)$$

2766 where $A \times C$ is the signal acceptance as parameterized in 6.3.4, and $\int \mathcal{L} = 139 \text{ fb}^{-1}$ is the
 2767 integrated luminosity of the dataset.

2768 The dependence of the expected number of signal and background events (normaliza-
 2769 tions) and the shape of the PDFs on the systematic uncertainties measured in section 6.6
 2770 is described by a set of nuisance parameters (NPs) θ_i . The Gaussian constraints are ap-
 2771 plied to those NPs. The constraints are implemented as additional ‘penalty’ terms added
 2772 to the likelihood which increase the negative log-likelihood when any nuisance parameter
 2773 is shifted from its nominal value. The final likelihood function $L(\sigma_{ggF}, \sigma_{VBF}, m_H, \theta_i)$ is
 2774 therefore a function of σ_{ggF} , σ_{VBF} , m_H , and θ_i .

2775 Furthermore, the normalization of SM background $pp \rightarrow ZZ$, including both $q\bar{q} \rightarrow$
 2776 ZZ and $gg \rightarrow ZZ$, is a free parameter (μ_{ZZ}) and profiled during the minimization.
 2777 Floating ZZ normalization in fit takes the advantage of reducing the dependence on the-
 2778 ory predictions and their associated uncertainties, especially given that the increased data
 2779 luminosity would provide precise determination of the SM ZZ background rate.

2780 At the end, the upper limit on production cross-section $\sigma_{ggF(VBF)}$ at a given heavy
 2781 resonance model is obtained by setting the mass of signal m_H parameter as constant at the
 2782 desired value, and maximising the likelihood function with respect to nuisance parame-
 2783 ters. The $CL_s^{[113]}$ method is used to obtain exclusion limits.

2784 6.7.2 Likelihood fit under background-only hypothesis for MVA-based
2785 analysis

2786 Both MVA- and cut-based analysis are studied by performing likelihood fit to the
2787 (pseudo-) data under the background-only hypothesis and under different signal models.
2788 Due to the same background estimation and modelling procedures, as well as the same
2789 method of systematic measurements, this section only shows the results of background-
2790 only fits for MVA-based analysis under the model of heavy Higgs resonance with narrow-
2791 width as an example. The final results of interpretation in both MVA- and cut- based
2792 analysis in all signal models described in section 6.5 will be measured in next section.

2793 First of all, table 6.14 summarized the expected and observed number of events for
2794 region of $m_{4\ell} > 200$ GeV together with their systematic uncertainties after background-
2795 only fit. The post-fit $m_{4\ell}$ spectrum in each category is shown in figure 6.25.

2796 To inspect the likelihood model, pulls and constraints as well as the correlation matrix
2797 of NPs are studied by performing a background only fit. Figure 6.26 shows the pulls and
2798 constraints when fitting to pseudo-data (top) and observed data (bottom). Figure 6.27
2799 shows the correlation matrix, only for NPs with correlation between each others greater
2800 than 0.1 when fitting to pseudo-data. The normalization of ZZ background is taken from
2801 data for one category each, as shown in table 6.15.

2802 The impact of a systematic uncertainty on the result depends on the production mode
2803 and the mass hypothesis. To check the impact of systematic uncertainties on expected sig-
2804 nal sensitivity, a NP ranking study is performed using signal injected Asimov data with
2805 the injected cross section close to 95% CLs upper limit at the masses of 400 GeV and
2806 1000 GeV. The results are shown in table 6.16. For ggF production, at lower masses, the
2807 systematic uncertainties of parton showering variation for signal, the luminosity uncer-
2808 tainty, and the parametrization of signal acceptance dominate, while at higher masses, the
2809 shape uncertainties from PDF variation for ZZ ($q\bar{q} \rightarrow ZZ$ and $gg \rightarrow ZZ$) background
2810 become important, as also seen in VBF production mode. In addition for VBF, jet related
2811 uncertainties become more important comparing to ggF production. Moreover, the domi-
2812 nate uncertainties include the acceptance uncertainty from QCD scale variation for signal
2813 and the luminosity uncertainty.

2814 6.7.3 Interpretations

2815 1. Spin-0 resonance with NWA

2816 In the absence of a specific model, the ratio of ggF and VBF production mode is un-
2817 known for this additional heavy scalar. For this reason, the fits for ggF and VBF processes

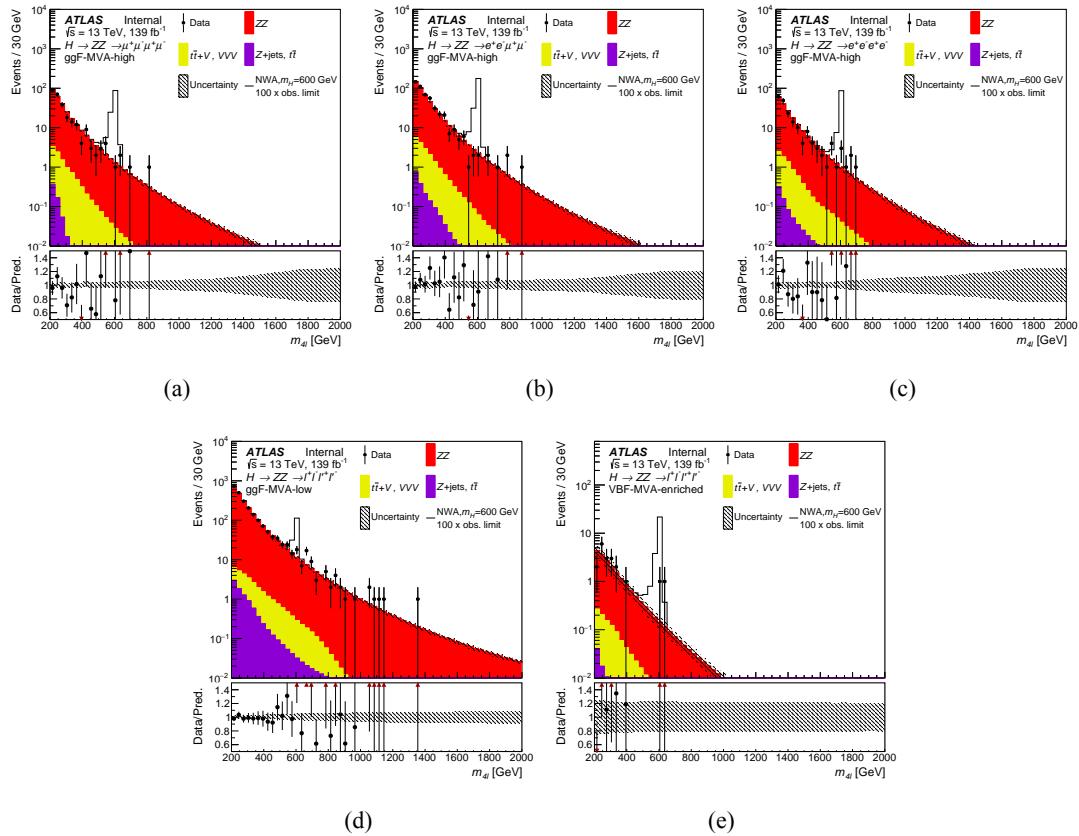


Fig. 6.25 Distribution of the four-lepton invariant mass $m_{4\ell}$ in the $\ell\ell\ell\ell$ search for (a), (b), (c) the ggF-MVA-high categories, (d) the ggF-MVA-low category and (e) the VBF-MVA-enriched category. The backgrounds are determined from a combined likelihood fit to the data under the background-only hypothesis. The simulated signal at 600 GeV is normalized to a cross section corresponding to one hundred times the observed upper limit given in section 6.7.3. The error bars on the data points indicate the statistical uncertainty, while the systematic uncertainty in the prediction is shown by the hatched band. The lower panels show the ratio of data to prediction.

are done separately, and in each case the cross section of the untested process is allowed to be a free parameter in the statistical fit. The observed and expected upper limit at 95% confidence level (CL) on the $\sigma \times BR(H \rightarrow ZZ)$ of a narrow scalar resonance for both ggF (left) and VBF (right) production mode with the integrated luminosity of 139 fb^{-1} is shown in figure 6.28 (6.29) for MVA- (cut-) based analysis. No excess over 2σ is found.

2. Spin-0 resonance with LWA

In the case of LWA model, only ggF production mode is studied. The interference between the heavy scalar and SM Higgs boson ($H-h$), as well as the heavy scalar and SM $gg \rightarrow ZZ$ continuum background ($H-B$) as modelled in section 6.5.3 are taken into account. The upper limit at 95% confidence level (CL) on ggF cross section times branch ratio ($\sigma_{ggF} \times BR(H \rightarrow ZZ)$) is shown in figure 6.30 for a width of 1, 5, 10 and 15% of

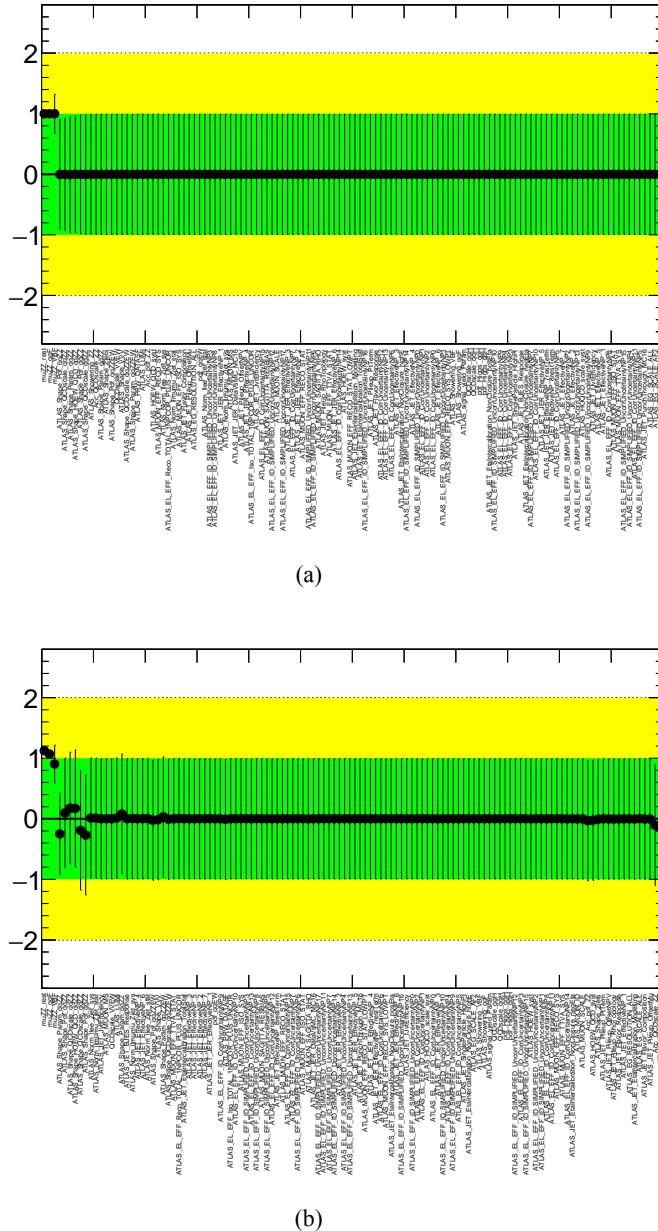
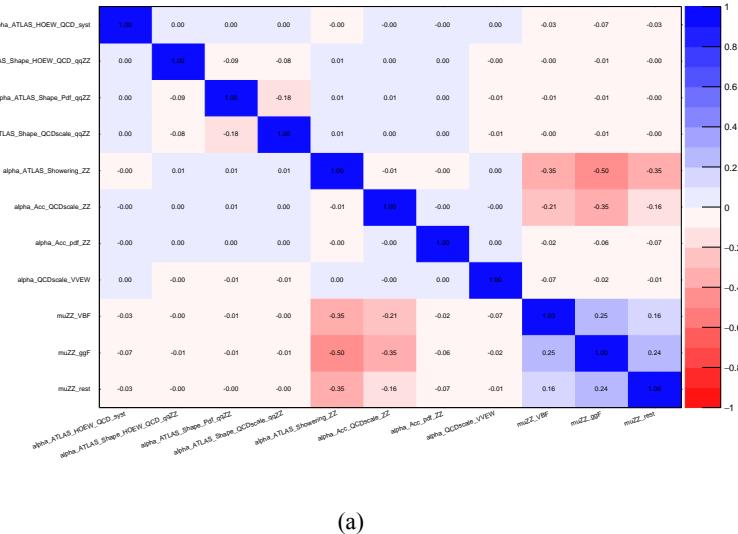


Fig. 6.26 Pulses and constraints of nuisance parameters after a background only fit to (a) Asimov data and (b) observed data in the $\ell\ell\ell\ell$ channel. The Asimov data is generated with background data only, and the observed data includes datasets from 2015 to 2018.

2829 m_H .

2830 3. Spin-2 RS Graviton resonance

2831 The observed and expected 95% upper limit on the cross section times branching ratio
 2832 for RS Graviton (RSG) scenario is shown in figure 6.31. Same as LWA case, only $4e$,
 2833 4μ and $2e2\mu$ channel of ggF production mode are used. On top of the expected and
 2834 observed upper limits in this model, a predicted cross section as function of mass provided
 2835 by theorist is also shown in the figure. Comparing with the observed result provided by
 2836 $ZZ \rightarrow \ell\ell\ell\ell$ decay, this spin-2 graviton is excluded up to a mass of 1500 GeV.



(a)

Fig. 6.27 Correlation of nuisance parameters after a background only fit to Asimov data in the $\ell\ell\ell\ell$ channel. The Asimov data is generated with background data only.

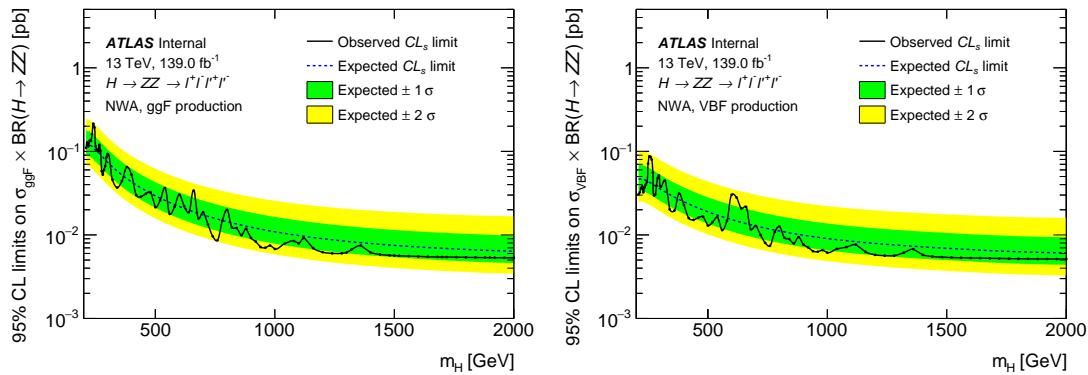


Fig. 6.28 The expected and observed upper limits at 95% CL on $\sigma \times BR(H \rightarrow ZZ)$ using the MVA-based analysis for ggF (left) and VBF (right) production. The green and yellow bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties in the expected limits.

4. Summary of interpretation

As a summary, figure 6.32 shows the comparison of expected and observed 95% CL upper limits between different models described above.

Figure 6.33 compares the expected 95% CL upper limits as a function of the NWA resonance mass in this analysis with full run-2 data and the one in previous publication^[114] with the integrated luminosity of 36.1 fb^{-1} . With a significant increase of integrated luminosity and an improved analysis strategy, comparing to the previous publication, the expected sensitivities of searching for narrow-width heavy resonance reduce by up to 70% in MVA-based analysis, where 50% of reduction is due to luminosity increase while other improvement mainly comes from inviting multivariate method.

Figure 6.34 shows the display of one candidate event passing analysis selection in

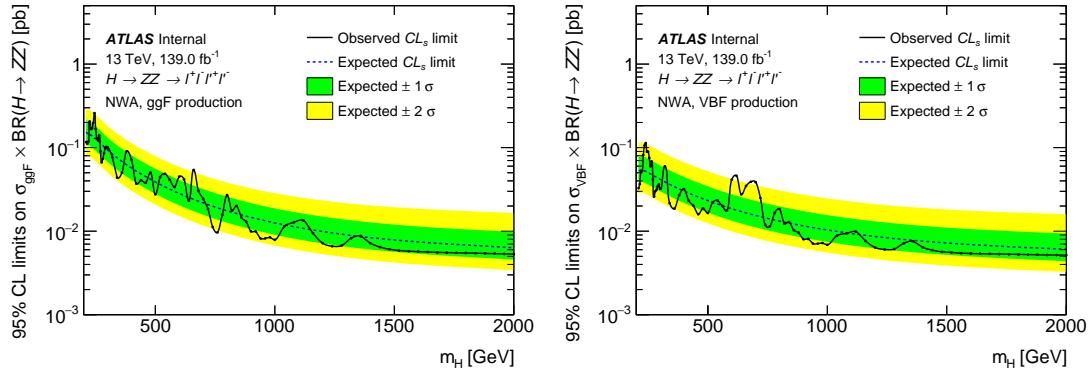


Fig. 6.29 The expected and observed upper limits at 95% CL on $\sigma \times BR(H \rightarrow ZZ)$ using the cut-based analysis for ggF (left) and VBF (right) production. The green and yellow bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties in the expected limits.

2848 four-muon final state with four-muon invariant mass of 1.34 TeV.

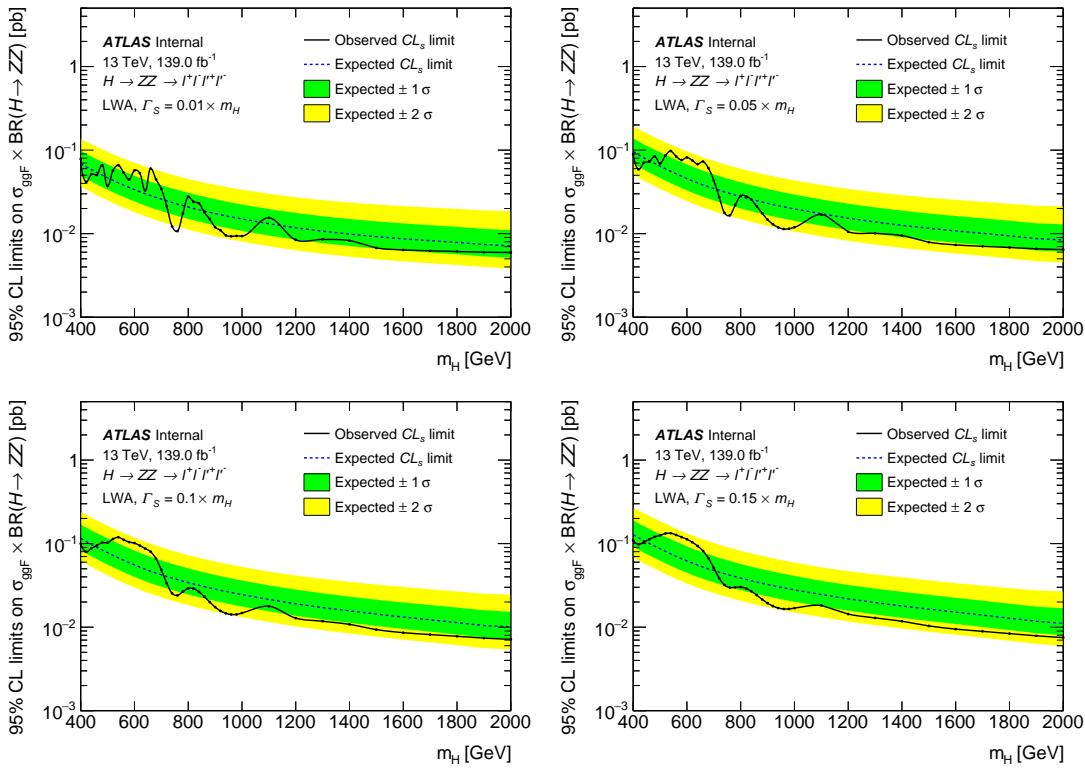


Fig. 6.30 The upper limits at 95% confidence level on $\sigma_{ggF} \times BR(H \rightarrow ZZ)$ as a function of the heavy resonance mass m_H for the ggF production mode with an intrinsic width of 1% (top left), 5% (top right), 10% (bottom left) and 15% (bottom right) for both the case where interference with Standard Model processes is considered. The green and yellow bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties in the expected limits.

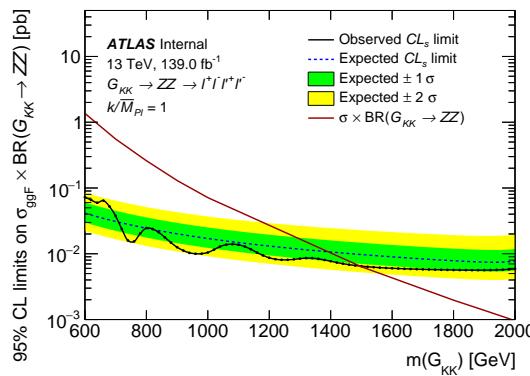


Fig. 6.31 The upper limits at 95% confidence level on $\sigma_{ggF} \times BR(G_{KK} \rightarrow ZZ)$ as a function of the heavy resonance mass $m(G_{KK})$ for the ggF production mode in RS Graviton model. The green and yellow bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties in the expected limits.

Table 6.13 A list of the experimental systematics considered in this analysis. The NPs have been separated by whether they only affect the normalisation (left column) or if they affect the shape (right column) of the $m_{4\ell}$ distribution. They are further subdivided into the primary objects that they affect.

Normalisation NPs	Shape NPs
Electrons	
EL_EFF_ID_CorrUncertaintyNP[0–15]	EG_RESOLUTION_ALL
EL_EFF_ID_SIMPLIFIED_UncorrUncertaintyNP[0–17]	EG_SCALE_ALLCORR
EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	EG_SCALE_E4SCINTILLATOR
EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	EG_SCALE_LARCALIB_EXTRA2015PRE EG_SCALE_LARTEMPERATURE_EXTRA2015PRE EG_SCALE_LARTEMPERATURE_EXTRA2016PRE
Muons	
MUON_EFF_ISO_STAT	MUON_ID
MUON_EFF_ISO_SYS	MUON_MS
MUON_EFF_RECO_STAT	MUON_SAGITTA_RESBIAS
MUON_EFF_RECO_STAT_LOWPT	MUON_SAGITTA_RHO
MUON_EFF_RECO_SYS	MUON_SCALE
MUON_EFF_RECO_SYS_LOWPT	
MUON_EFF_TTVA_STAT	
MUON_EFF_TTVA_SYS	
Jets	
	JET_BJES_Response JET_EffectiveNP_[1–7] JET_EffectiveNP_8restTerm JET_EtaIntercalibration_Modelling JET_EtaIntercalibration_NonClosure_highE JET_EtaIntercalibration_NonClosure_negEta JET_EtaIntercalibration_NonClosure_posEta JET_EtaIntercalibration_TotalStat JET_Flavor_Composition JET_Flavor_Response JET_JER_DataVsMC JET_JER_EffectiveNP_[1–6] JET_JER_EffectiveNP_7restTerm JET_Pileup_OffsetMu JET_Pileup_OffsetNPV JET_Pileup_PtTerm JET_Pileup_RhoTopology JET_PunchThrough_MC16 JET_SingleParticle_HighPt
Other	
HOEW_QCD_syst HOEW_syst HOQCD_scale_syst PRW_DATASF	

Table 6.14 Expected and observed numbers of events for $m_{4\ell} > 200$ GeV, together with their systematic uncertainties, for three MVA-based categories. The expected number of events, as well as their uncertainties, are obtained from a likelihood fit to the data under the background-only hypothesis. The uncertainties of the ZZ normalisation factors, presented in table 6.15, are also taken into account.

Process	VBF-enriched category	ggF-enriched categories			the “rest” category
		4 μ channel	2e2 μ channel	4e channel	
$q\bar{q} \rightarrow ZZ$	11 ± 4	232 ± 10	389 ± 17	154 ± 7	2008 ± 47
$gg \rightarrow ZZ$	3 ± 2	37 ± 6	64 ± 10	26 ± 4	247 ± 19
ZZ (EW)	4.1 ± 0.4	4.5 ± 0.2	7.5 ± 0.4	3 ± 0.2	14.3 ± 0.7
$Z+jets, t\bar{t}$	0.08 ± 0.02	0.6 ± 0.1	1.7 ± 0.4	0.8 ± 0.1	8.8 ± 2.1
$t\bar{t}V, VVV$	0.97 ± 0.1	9.8 ± 0.2	17.5 ± 0.4	7.8 ± 0.2	21.9 ± 0.5
Total background	19 ± 4.5	285 ± 11.7	479 ± 19.7	192 ± 8.1	2301 ± 50.7
Observed	19	271	493	191	2301

Table 6.15 ZZ normalization factor in each category, obtained from a likelihood fit to the data under the background-only hypothesis.

Normalization factor	Fitted value
$\mu_{ZZ}^{ggF-MVA-high}$	1.07 ± 0.047
$\mu_{ZZ}^{ggF-MVA-low}$	1.12 ± 0.026
$\mu_{ZZ}^{VBF-MVA-enriched}$	0.91 ± 0.314

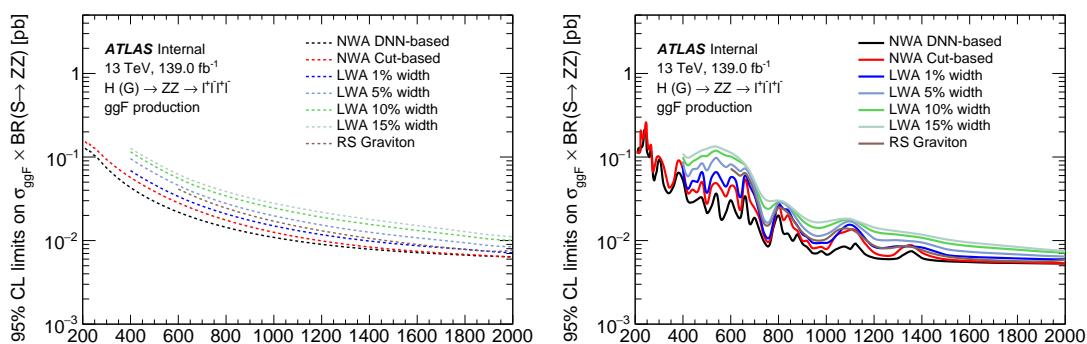


Fig. 6.32 The expected (left) and observed (right) upper limits at 95% confidence level on $\sigma \times BR(S \rightarrow ZZ)$ for ggF production mode at different assumptions.

Table 6.16 Impact of the leading systematic uncertainties, the data statistic uncertainties, as well as the total uncertainties on the predicted signal event yield with the cross section times branching ratio being set to the expected upper limit, expressed as a percentage of the signal yield for the ggF (left) and VBF (right) production modes at $m_H = 400$ and 1000 GeV.

ggF production		VBF production	
Systematic source	Impact [%]	Systematic source	Impact [%]
$m_H = 400$ GeV			
Parton showering of ggF	2.3	QCD scale of VBF	2.7
Luminosity	1.8	Jet flavor composition	2.5
PDF of $q\bar{q} \rightarrow ZZ$	1.6	Luminosity	1.8
Signal yield parameterization	1.4	Jet energy scale (in-su calibration)	1.6
Data stat. uncertainty	48	Data stat. uncertainty	57
Total Uncertainty	49	Total Uncertainty	58
$m_H = 1000$ GeV			
PDF of $q\bar{q} \rightarrow ZZ$	2.5	QCD scale of VBF	2.3
Parton showering of ggF	2.4	PDF of $q\bar{q} \rightarrow ZZ$	2.2
PDF of $gg \rightarrow ZZ$	1.9	Luminosity	1.8
Luminosity	1.8	PDF of $gg \rightarrow ZZ$	1.6
Data stat. uncertainty	84	Data stat. uncertainty	92
Total Uncertainty	86	Total Uncertainty	93

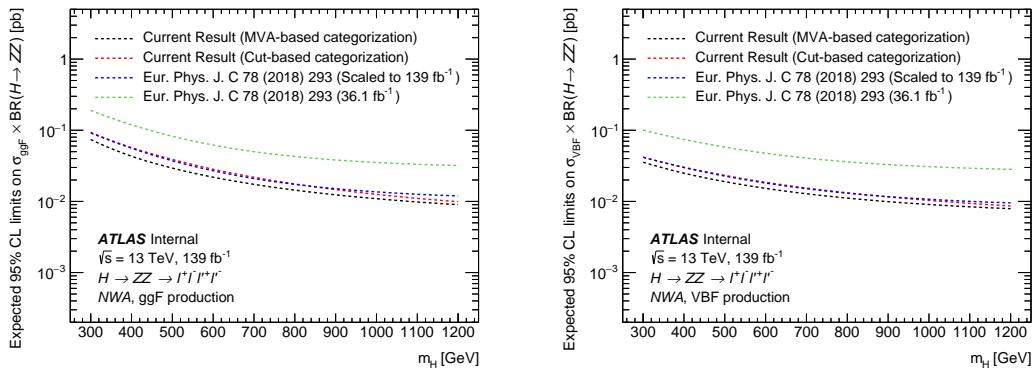


Fig. 6.33 Comparisons of the expected upper limits at 95% CL on the cross section times branching ratio as a function of the heavy resonance mass m_H for the ggF production mode (left) and for the VBF production mode (right) in the case of the NWA. The expected limits from the previous publication are shown in the green dashed line and are projected to the 139 fb^{-1} as shown in the blue dashed line. In addition, the current results based on either cut-based categorisation or the multivariate-based categorisation are shown in red and black lines.

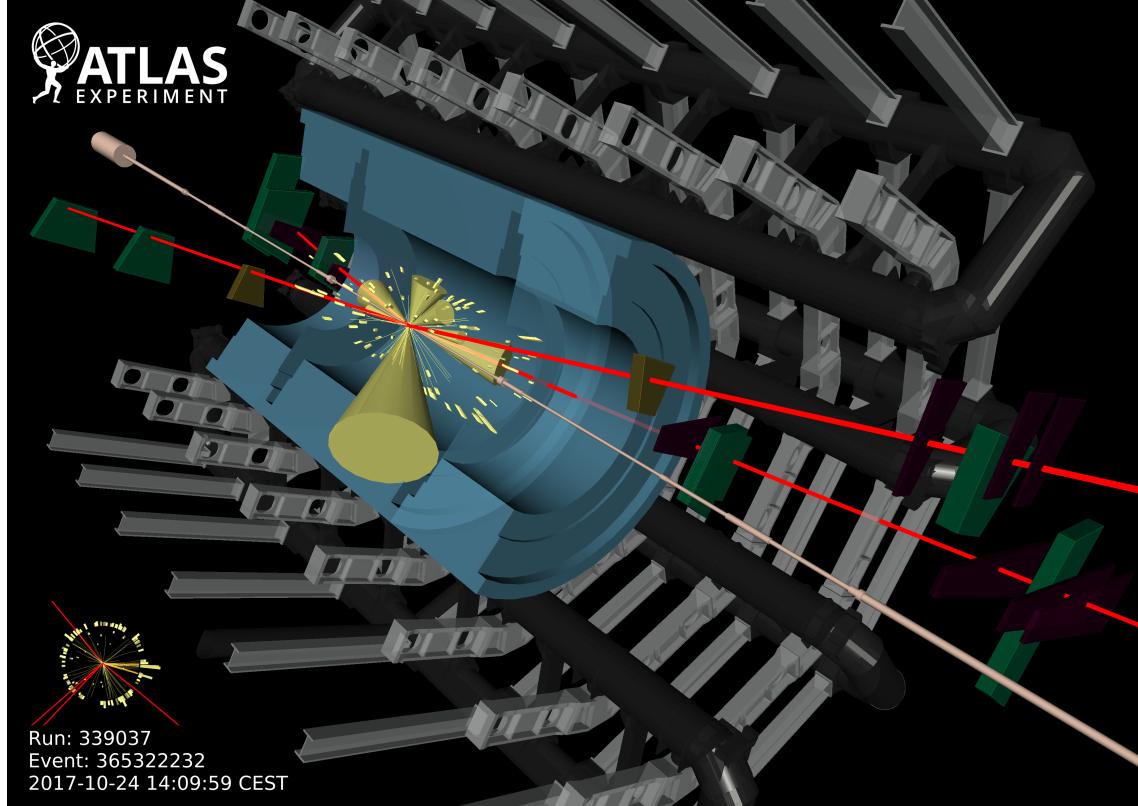


Fig. 6.34 Display of one candidate event in 4μ final state with the mass of 1.35 TeV.

2849 6.8 Conclusion

2850 Searches of heavy resonances decaying into a pair of Z boson to $\ell\ell\ell\ell$ final state
2851 are performed using 139 fb^{-1} of 13 TeV pp collision data collected by ATLAS experiment
2852 at the LHC. The results are interpreted as 95% CL upper limits on the production cross
2853 section of a spin-0 and spin-2 resonances under different theoretical models. The search
2854 range of the hypothetical resonances is between 200 GeV to 2000 GeV depending on the
2855 signal model.

2856 The spin-0 resonance is assumed to be a heavy Higgs like scalar produced predom-
2857 inantly from gluon–gluon fusion (ggF) and vector-boson fusion (VBF) decays, and it is
2858 studied under both the narrow-width approximation and with the large-width assumption.
2859 For narrow-width approximation, limits on cross section of heavy scalar decaying into
2860 two Z bosons are set separately for ggF and VBF production modes, under DNN- and
2861 cut- based analysis. In DNN-based analysis, the 95% CL upper limit range is from 215 fb
2862 at $m_H = 240 \text{ GeV}$ to 5.3 fb at $m_H = 2000 \text{ GeV}$ for ggF production mode, and from 87 fb
2863 at $m_H = 255 \text{ GeV}$ to 5.1 fb at $m_H = 1960 \text{ GeV}$ for VBF production mode. In cut-based
2864 analysis, the 95% CL upper limit range is from 259 fb at $m_H = 245 \text{ GeV}$ to 5.3 fb at
2865 $m_H = 2000 \text{ GeV}$ for ggF production mode, and from 113 fb at $m_H = 240 \text{ GeV}$ to 5.1 fb
2866 at $m_H = 2000 \text{ GeV}$ for VBF production mode. DNN-based analysis gains about 20% im-
2867 provement on upper limits at lower mass region comparing to the cut-based analysis, while
2868 for mass above 1500 GeV, both analyses perform closely. For large-width approximation,
2869 limits are studied on ggF production rate at four different widths assumptions: 1%, 5%,
2870 10% and 15% of resonance’s mass, with the interference between the heavy scalar and the
2871 SM Higgs boson as well as the heavy scalar and the SM $gg \rightarrow ZZ$ continuum background
2872 taken into account. The maximum and minimum of upper limits are obtained as 78 fb at
2873 $m_H = 400 \text{ GeV}$ to 5.9 fb at $m_H = 2000 \text{ GeV}$ for 1% width; 98 fb at $m_H = 540 \text{ GeV}$ to 6.4
2874 fb at $m_H = 2000 \text{ GeV}$ for 5% width; 119 fb at $m_H = 540 \text{ GeV}$ to 7.1 fb at $m_H = 2000 \text{ GeV}$
2875 for 10% width; 133 fb at $m_H = 540 \text{ GeV}$ to 7.5 fb at $m_H = 2000 \text{ GeV}$ for 15% width.
2876 Last but not least, the framework of the Randall–Sundrum model with a graviton excita-
2877 tion spin-2 resonance with $m(G_{KK}) < 1500 \text{ GeV}$ is excluded at 95% CL.

2878

Chapter 7 Summary

2879 On December 3rd, 2018, the LHC finished its second run (run-2) after three fantastic
2880 years. Thanks to run-2 with largely increased statistic, we now know the masses of the
2881 Higgs boson, top quark and W boson to considerably greater precision. And also confirm
2882 the Standard Model as a stable theory.

2883 In this dissertation, various physics processes in $ZZ \rightarrow \ell\ell\ell\ell$ final state are studied,
2884 taking the advantage of full run-2 pp collision data in the LHC. Using this signature, we
2885 measured the fiducial cross section of ZZ production to $\ell\ell\ell\ell jj$ channel in SM, which is
2886 an important physics process and major background in many analysis with ZZ produc-
2887 tion, eg. Higgs analysis (HZZ). In addition, we searched the electroweak ZZ produc-
2888 tion via vector boson scattering in associated with 2-jet process in $\ell\ell\ell\ell$ final state. In the
2889 meantime, the searches of heavy resonances decaying into a pair of Z bosons to $\ell\ell\ell\ell$
2890 final state for several different hypothetical resonances are conducted in this dissertation.
2891 The results of several analyses are summarized as below:

2892 **Measurement of fiducial cross section of ZZ production in $\ell\ell\ell\ell jj$ final state**

2893 The fiducial cross section of inclusive SM $ZZ \rightarrow \ell\ell\ell\ell$ production is measured to
2894 be:

$$\sigma_{ZZ \rightarrow \ell\ell\ell\ell}^{fid} = 1.27 \pm 0.12(stat) \pm 0.02(theo) \pm 0.07(exp) \pm 0.01(bkg) \pm 0.03(lumi) \quad (7.1)$$

2895 which is found to be compatible with the SM prediction. The ZZ cross section is calcu-
2896 lated with up to one (three) outgoing partons at NLO (LO) using Sherpa 2.2.2 for QCD
2897 production, and in LO using MadGraph5_aMC@NLO 2.6.1 for EW production. The to-
2898 tal uncertainty is 11%, the analysis is still data static dominant (data statistic uncertainty
2899 is about 9.5%).

2900 **Observation of electroweak ZZ production in $\ell\ell\ell\ell jj$ final state**

2901 Thanks to the largely increased data statistic collected by ATLAS experiment in the
2902 LHC run-2, the electroweak ZZ production (EW- $ZZjj$) to $\ell\ell\ell\ell$ channel in association
2903 with two jets is observed with a significant deviation from the background-only hypothe-
2904 sis. The signal strength of EW- $ZZjj$ production, the normalization of QCD- $ZZjj$ pro-
2905 duction, as well as the observed and expected statistical significance measured in $\ell\ell\ell\ell jj$

2906 channel are found to be:

$$\begin{aligned}\mu_{\text{EW}} &= 1.54 \pm 0.42 \\ \mu_{\text{QCD}} &= 0.95 \pm 0.22\end{aligned}\quad (7.2)$$

Obs. (Exp.) Significance = 5.48 (3.90) σ

2907 Then in this dissertation, the differential cross section and expected significance of
2908 EW- $ZZjj$ production, using 3000 fb^{-1} simulated pp collision data at a centre-of-mass
2909 energy of 14 TeV to be recorded by ATLAS experiment at the HL-LHC, are studied via
2910 simulations. The HL-LHC will for sure give us more opportunity to probe rare process
2911 like $ZZ \rightarrow \ell\ell\ell\ell$ in the future.

2912 **Searches of heavy ZZ resonances in $\ell\ell\ell\ell$ final state**

2913 Searches of heavy ZZ resonances are performed in four-lepton invariant mass $m_{4\ell}$
2914 range from 200 GeV to 2000 GeV. Data are found to agree with the background-only
2915 hypothesis, and 95% CL upper limits are set on the production rate under the models of:

- 2916 • Spin-0 heavy Higgs under narrow-width approximation (NWA).

2917 Search range is from 200 GeV to 2000 GeV.

2918 In DNN-based analysis, the limits are range from 215 fb at $m_H = 240$ GeV to 5.3
2919 fb at $m_H = 2000$ GeV for ggF production mode, and from 87 fb at $m_H = 255$ GeV
2920 to 5.1 fb at $m_H = 1960$ GeV for VBF production mode.

2921 The DNN-based analysis is found to be at most 20% better than cut-based results.

- 2922 • Spin-0 heavy Higgs under large-width approximation (LWA) with the width of 1,
2923 5, 10, 15% of its mass.

2924 Search range is from 400 GeV to 2000 GeV, and only ggF production is studied.

2925 The maximum and minimum of upper limits are obtained as 78 fb at $m_H = 400$ GeV
2926 to 5.9 fb at $m_H = 2000$ GeV for 1% width; 98 fb at $m_H = 540$ GeV to 6.4 fb at $m_H =$
2927 2000 GeV for 5% width; 119 fb at $m_H = 540$ GeV to 7.1 fb at $m_H = 2000$ GeV for
2928 10% width; 133 fb at $m_H = 540$ GeV to 7.5 fb at $m_H = 2000$ GeV for 15% width.

- 2929 • Spin-2 graviton excitation under the Randall–Sundrum model.

2930 Search range is from 600 GeV to 2000 GeV, and only ggF production is studied.

2931 The maximum and minimum of limits are 73 fb at $m_H = 600$ GeV and 5.6 fb
2932 at $m_H = 1880$ GeV for ggF production mode. And the mass of graviton below
2933 1500 GeV is excluded comparing the observed results with theoretical prediction.

2934

2935 In summary, the $ZZ \rightarrow \ell\ell\ell\ell$ production presented in this dissertation are consistent
2936 with SM prediction. This result completes the observation of weak boson scattering for

2937 massive bosons, which is a new milestone reached in the study of electroweak symmetry
2938 breaking. In the meantime, no indication of new physics is observed. We are looking
2939 forward the HL-LHC, with greatly increased luminosity and higher centre-of-mass energy,
2940 which should enhance the sensitivity for new physics search and precise measurement for
2941 rare process like $\ell\ell\ell\ell$ final state.

2942

Bibliography

- 2943 [1] LANGACKER P. Introduction to the Standard Model and Electroweak Physics[C/OL]//
2944 Proceedings of Theoretical Advanced Study Institute in Elementary Particle Physics on The
2945 dawn of the LHC era (TASI 2008): Boulder, USA, June 2-27, 2008. 2010: 3-48. DOI:
2946 [10.1142/9789812838360_0001](https://doi.org/10.1142/9789812838360_0001).
- 2947 [2] PICH A. Electroweak Symmetry Breaking and the Higgs Boson[J/OL]. Acta Phys. Polon.,
2948 2016, B47:151. DOI: [10.5506/APhysPolB.47.151](https://doi.org/10.5506/APhysPolB.47.151).
- 2949 [3] DREMIN I M. Soft and hard processes in QCD[J/OL]. JETP Lett., 2005, 81:307-310. DOI:
2950 [10.1134/I.1944068](https://doi.org/10.1134/I.1944068).
- 2951 [4] WOMERSLEY J. QCD at the Tevatron: Status and prospects[C/OL]//Proceedings, 5th Inter-
2952 national Symposium on Radiative Corrections - RADCOR 2000. 2000. <http://www.slac.stanford.edu/econf/C000911/>.
- 2953 [5] LIN H W, et al. Parton distributions and lattice QCD calculations: a community white paper
2954 [J/OL]. Prog. Part. Nucl. Phys., 2018, 100:107-160. DOI: [10.1016/j.ppnp.2018.01.007](https://doi.org/10.1016/j.ppnp.2018.01.007).
- 2955 [6] COLLINS J C, SOPER D E, STERMAN G F. Factorization of Hard Processes in QCD[J/OL].
2956 Adv. Ser. Direct. High Energy Phys., 1989, 5:1-91. DOI: [10.1142/9789814503266_0001](https://doi.org/10.1142/9789814503266_0001).
- 2957 [7] STIRLING W J. Perturbative QCD[J/OL]. 2000(CERN-OPEN-2000-296):40 p. <https://cds.cern.ch/record/1194745>. DOI: [10.5170/CERN-2000-007.305](https://doi.org/10.5170/CERN-2000-007.305).
- 2958 [8] GROJEAN C. Higgs Physics[J/OL]. 2017(arXiv:1708.00794):143-158. 12 p. <https://cds.cern.ch/record/2243593>. DOI: [10.5170/CERN-2016-005.143](https://doi.org/10.5170/CERN-2016-005.143).
- 2959 [9] Cern yellow reports: Monographs: Handbook of LHC Higgs Cross Sections: 4. Deciphering
2960 the Nature of the Higgs Sector[M/OL]. 2016. <https://cds.cern.ch/record/2227475>. DOI: [10.23731/CYRM-2017-002](https://doi.org/10.23731/CYRM-2017-002).
- 2961 [10] HEINEMEYER S. Cern yellow reports: Monographs: Handbook of LHC Higgs Cross Sec-
2962 tions: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group[M/OL].
2963 2013. <https://cds.cern.ch/record/1559921>. DOI: [10.5170/CERN-2013-004](https://doi.org/10.5170/CERN-2013-004).
- 2964 [11] CHANG J, CHEUNG K, LU C T, et al. $w w$ scattering in the era of post-higgs-boson discovery
2965 [J/OL]. Phys. Rev. D, 2013, 87:093005. <https://link.aps.org/doi/10.1103/PhysRevD.87.093005>.
- 2966 [12] BRUNING O S, COLLIER P, LEBRUN P, et al. LHC Design Report Vol.1: The LHC Main
2967 Ring[J]. 2004.
- 2968 [13] BUNING O, COLLIER P, LEBRUN P, et al. LHC Design Report. 2. The LHC infrastructure
2969 and general services[J]. 2004.
- 2970 [14] BENEDIKT M, COLLIER P, MERTENS V, et al. LHC Design Report. 3. The LHC injector

- 2975 chain[J]. 2004.
- 2976 [15] EVANS L, BRYANT P. LHC machine[J/OL]. Journal of Instrumentation, 2008, 3(08):S08001-
2977 S08001. <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08001>. DOI: [10.1088/1748-0221/3/08/s08001](https://doi.org/10.1088/1748-0221/3/08/s08001).
- 2979 [16] MOBS E. The CERN accelerator complex. Complexe des accélérateurs du CERN[J/OL]. 2016.
2980 <https://cds.cern.ch/record/2197559>.
- 2981 [17] COLLABORATION A. The ATLAS experiment at the CERN large hadron collider[J/OL].
2982 Journal of Instrumentation, 2008, 3(08):S08003-S08003. <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08003>. DOI: [10.1088/1748-0221/3/08/s08003](https://doi.org/10.1088/1748-0221/3/08/s08003).
- 2984 [18] PEREZ G. Unitarization models for vector boson scattering at the lhc[D/OL]. 2018. DOI:
2985 [10.5445/IR/1000082199](https://doi.org/10.5445/IR/1000082199).
- 2986 [19] PEQUENAO J. Computer generated image of the whole ATLAS detector[Z/OL]. 2008. <https://cds.cern.ch/record/1095924>.
- 2988 [20] MCFAYDEN J. The lhc and atlas detector. in: Third generation susy and $t\bar{t} + z$ productin.
2989 [D/OL]. 2014. DOI: [10.1007/978-3-319-07191-6_2](https://doi.org/10.1007/978-3-319-07191-6_2).
- 2990 [21] COLLABRATION A. Operation and performance of the ATLAS semiconductor tracker. Operation
2991 and performance of the ATLAS semiconductor tracker[J/OL]. JINST, 2014, 9(CERN-
2992 PH-EP-2014-049. CERN-PH-EP-2014-049):P08009. 80 p. <https://cds.cern.ch/record/1698966>.
2993 DOI: [10.1088/1748-0221/9/08/P08009](https://doi.org/10.1088/1748-0221/9/08/P08009).
- 2994 [22] ATLAS pixel detector electronics and sensors[J/OL]. Journal of Instrumentation, 2008, 3(07):
2995 P07007-P07007. <https://doi.org/10.1088%2F1748-0221%2F3%2F07%2Fp07007>. DOI: [10.1088/1748-0221/3/07/p07007](https://doi.org/10.1088/1748-0221/3/07/p07007).
- 2997 [23] MULLIER G. The upgraded pixel detector of the atlas experiment for run-2 at the large hadron
2998 collider[J/OL]. Journal of Instrumentation, 2016, 11:C02061-C02061. DOI: [10.1088/1748-0221/11/02/C02061](https://doi.org/10.1088/1748-0221/11/02/C02061).
- 3000 [24] The silicon microstrip sensors of the atlas semiconductor tracker[J/OL]. Nuclear Instruments
3001 and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Asso-
3002 ciated Equipment, 2007, 578(1):98 - 118. <http://www.sciencedirect.com/science/article/pii/S0168900207007644>. DOI: <https://doi.org/10.1016/j.nima.2007.04.157>.
- 3004 [25] SULTAN D M S. Development of small-pitch, thin 3d sensors for pixel detector upgrades at
3005 hl-lhc[D/OL]. 2017. DOI: [10.13140/RG.2.2.36253.82403/1](https://doi.org/10.13140/RG.2.2.36253.82403/1).
- 3006 [26] The ATLAS transition radiation tracker (TRT) proportional drift tube: design and performance
3007 [J/OL]. Journal of Instrumentation, 2008, 3(02):P02013-P02013. <https://doi.org/10.1088%2F1748-0221%2F3%2F02%2Fp02013>. DOI: [10.1088/1748-0221/3/02/p02013](https://doi.org/10.1088/1748-0221/3/02/p02013).
- 3009 [27] BUCHANAN N, CHEN L, GINGRICH D, et al. Design and implementation of the front end

- 3010 board for the readout of the atlas liquid argon calorimeters[J/OL]. Journal of Instrumentation,
3011 2008, 3:P03004. DOI: [10.1088/1748-0221/3/03/P03004](https://doi.org/10.1088/1748-0221/3/03/P03004).
- 3012 [28] Technical design report atlas: ATLAS liquid-argon calorimeter: Technical Design Report
3013 [M/OL]. Geneva: CERN, 1996. <https://cds.cern.ch/record/331061>.
- 3014 [29] AAD G, ABBOTT B, ABDALLAH J, et al. Readiness of the atlas tile calorimeter for lhc
3015 collisions[J]. European Physical Journal C, 2010.
- 3016 [30] Technical design report atlas: ATLAS muon spectrometer: Technical Design Report[M/OL].
3017 Geneva: CERN, 1997. <https://cds.cern.ch/record/331068>.
- 3018 [31] SLIWA K. "ATLAS Overview and Main Results"[C]//Proceedings, International School on
3019 High Energy Physics : Workshop on High Energy Physics in the near Future. (LISHEP 2013):
3020 Rio de Janeiro, Brazil, March 17-24, 2013. 2013.
- 3021 [32] RUIZ-MARTINEZ A, COLLABORATION A. The Run-2 ATLAS Trigger System: ATL-DAQ-
3022 PROC-2016-003[R/OL]. Geneva: CERN, 2016. <https://cds.cern.ch/record/2133909>. DOI:
3023 [10.1088/1742-6596/762/1/012003](https://doi.org/10.1088/1742-6596/762/1/012003).
- 3024 [33] PÁSZTOR G. The Upgrade of the ATLAS Electron and Photon Triggers towards LHC Run 2 and
3025 their Performance: ATL-DAQ-PROC-2015-053. arXiv:1511.00334[R/OL]. Geneva: CERN,
3026 2015. <https://cds.cern.ch/record/2063746>.
- 3027 [34] COLLABORATION A. Atlas computing: Technical design report[R/OL]. Geneva: CERN.
3028 <http://atlas-computing.web.cern.ch/atlas-computing/packages/athenaCore/athenaCore.php>.
- 3029 [35] AAD G, et al. The ATLAS Simulation Infrastructure[J/OL]. Eur. Phys. J., 2010, C70:823-874.
3030 DOI: [10.1140/epjc/s10052-010-1429-9](https://doi.org/10.1140/epjc/s10052-010-1429-9).
- 3031 [36] HÖCHE S. Introduction to parton-shower event generators[C/OL]//Proceedings, Theoretical
3032 Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Fron-
3033 tier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014. 2015: 235-
3034 295. DOI: [10.1142/9789814678766_0005](https://doi.org/10.1142/9789814678766_0005).
- 3035 [37] GLEISBERG T, HöCHE S, KRAUSS F, et al. Event generation with SHERPA 1.1[J/OL].
3036 Journal of High Energy Physics, 2009, 2009(02):007-007. <https://doi.org/10.1088%2F1126-6>
3037 [708%2F2009%2F02%2F007](https://doi.org/10.1088/1126-6708/2009/02/007). DOI: [10.1088/1126-6708/2009/02/007](https://doi.org/10.1088/1126-6708/2009/02/007).
- 3038 [38] BÄHR M, GIESEKE S, GIGG M A, et al. Herwig++ physics and manual[J/OL]. The European
3039 Physical Journal C, 2008, 58(4):639-707. <https://doi.org/10.1140/epjc/s10052-008-0798-9>.
- 3040 [39] NASON P. A New method for combining NLO QCD with shower Monte Carlo algorithms
3041 [J/OL]. JHEP, 2004, 11:040. DOI: [10.1088/1126-6708/2004/11/040](https://doi.org/10.1088/1126-6708/2004/11/040).
- 3042 [40] FRIXIONE S, WEBBER B R. Matching NLO QCD computations and parton shower simula-
3043 tions[J/OL]. Journal of High Energy Physics, 2002, 2002(06):029-029. <https://doi.org/10.108>
3044 [8%2F1126-6708%2F2002%2F06%2F029](https://doi.org/10.1088/1126-6708%2F2002%2F06%2F029). DOI: [10.1088/1126-6708/2002/06/029](https://doi.org/10.1088/1126-6708/2002/06/029).

Bibliography

- 3045 [41] SJOSTRAND T, MRENN S, SKANDS P Z. A Brief Introduction to PYTHIA 8.1[J/OL].
3046 Comput. Phys. Commun., 2008, 178:852-867. DOI: [10.1016/j.cpc.2008.01.036](https://doi.org/10.1016/j.cpc.2008.01.036).
- 3047 [42] BARBERIO E, et al. The Geant4-Based ATLAS Fast Electromagnetic Shower Simulation
3048 [C/OL]//Astroparticle, particle and space physics, detectors and medical physics applications.
3049 Proceedings, 10th Conference, ICATPP 2007, Como, Italy, October 8-12, 2007. 2008: 802-806.
3050 <http://cdsweb.cern.ch/record/1064665/files/soft-conf-2007-002.pdf>. DOI: [10.1142/9789812819093_0133](https://doi.org/10.1142/9789812819093_0133).
- 3052 [43] RICHTER-WAS E, FROIDEVAUX D, POGGIOLE L. ATLFAST 2.0 a fast simulation package
3053 for ATLAS: ATL-PHYS-98-131[R/OL]. Geneva: CERN, 1998. <https://cds.cern.ch/record/683751>.
- 3055 [44] EDMONDS K, FLEISCHMANN S, LENZ T, et al. The Fast ATLAS Track Simulation
3056 (FATRAS): ATL-SOFT-PUB-2008-001. ATL-COM-SOFT-2008-002[R/OL]. Geneva: CERN,
3057 2008. <https://cds.cern.ch/record/1091969>.
- 3058 [45] BOYD J T, COSTA M J, TONOVAN A, et al. Commissioning of the ATLAS reconstruction
3059 software with first data[J/OL]. Journal of Physics: Conference Series, 2010, 219(3):032059.
3060 <https://doi.org/10.1088%2F1742-6596%2F219%2F3%2F032059>. DOI: [10.1088/1742-6596/219/3/032059](https://doi.org/10.1088/1742-6596/219/3/032059).
- 3062 [46] CORNELISSEN T, ELSING M, FLEISCHMANN S, et al. Concepts, Design and Implemen-
3063 tation of the ATLAS New Tracking (NEWT): ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-
3064 2007-002[R/OL]. Geneva: CERN, 2007. <https://cds.cern.ch/record/1020106>.
- 3065 [47] AAD G, et al. Muon reconstruction performance of the ATLAS detector in proton-proton col-
3066 lision data at $\sqrt{s}=13$ TeV[J/OL]. Eur. Phys. J., 2016, C76(5):292. DOI: [10.1140/epjc/s10052-016-4120-y](https://doi.org/10.1140/epjc/s10052-016-4120-y).
- 3068 [48] ILLINGWORTH J, KITTLER J. A survey of the hough transform[J/OL]. Computer Vision,
3069 Graphics, and Image Processing, 1988, 44(1):87 - 116. <http://www.sciencedirect.com/science/article/pii/S0734189X88800331>. DOI: [https://doi.org/10.1016/S0734-189X\(88\)80033-1](https://doi.org/10.1016/S0734-189X(88)80033-1).
- 3071 [49] Vertex Reconstruction Performance of the ATLAS Detector at 13 TeV: ATL-PHYS-PUB-2015-
3072 026[R/OL]. Geneva: CERN, 2015. <https://cds.cern.ch/record/2037717>.
- 3073 [50] AABOUD M, et al. Reconstruction of primary vertices at the ATLAS experiment in Run 1
3074 proton-proton collisions at the LHC[J/OL]. Eur. Phys. J., 2017, C77(5):332. DOI: [10.1140/epjc/s10052-017-4887-5](https://doi.org/10.1140/epjc/s10052-017-4887-5).
- 3076 [51] Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton
3077 collision data: ATLAS-CONF-2016-024[R/OL]. Geneva: CERN, 2016. <https://cds.cern.ch/record/2157687>.
- 3079 [52] LAMPL W, LAPLACE S, LELAS D, et al. Calorimeter Clustering Algorithms: Description

Bibliography

- 3080 and Performance: ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003[R/OL]. Geneva:
3081 CERN, 2008. <https://cds.cern.ch/record/1099735>.
- 3082 [53] CORNELISSEN T G, ELSING M, GAVRILENKO I, et al. The global χ^2 track fitter in ATLAS
3083 [J/OL]. Journal of Physics: Conference Series, 2008, 119(3):032013. <https://doi.org/10.1088%2F1742-6596%2F119%2F3%2F032013>. DOI: [10.1088/1742-6596/119/3/032013](https://doi.org/10.1088/1742-6596/119/3/032013).
- 3085 [54] Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton
3086 collision data: ATLAS-CONF-2014-032[R/OL]. Geneva: CERN, 2014. <http://cds.cern.ch/re>
3087 [cord/1706245](http://cds.cern.ch/record/1706245).
- 3088 [55] LIMPER M. Track and vertex reconstruction in the ATLAS inner detector[D/OL]. 2009. <http://cds.cern.ch/record/1202457>.
- 3090 [56] AAD G, et al. Topological cell clustering in the ATLAS calorimeters and its performance in
3091 LHC Run 1[J/OL]. Eur. Phys. J., 2017, C77:490. DOI: [10.1140/epjc/s10052-017-5004-5](https://doi.org/10.1140/epjc/s10052-017-5004-5).
- 3092 [57] ZHENG Z. Identification of very-low transverse momentum muons with the ATLAS experi-
3093 ment: ATL-MUON-PROC-2018-018[R/OL]. Geneva: CERN, 2018. <https://cds.cern.ch/reco>
3094 [rd/2649299](https://cds.cern.ch/record/2649299).
- 3095 [58] CACCIARI M, SALAM G P. Pileup subtraction using jet areas[J/OL]. Physics Letters B, 2008,
3096 659(1):119 - 126. <http://www.sciencedirect.com/science/article/pii/S0370269307011094>. DOI:
3097 <https://doi.org/10.1016/j.physletb.2007.09.077>.
- 3098 [59] CACCIARI M, SALAM G P, SOYEZ G. The anti-ktjet clustering algorithm[J/OL]. Journal of
3099 High Energy Physics, 2008, 2008(04):063-063. <https://doi.org/10.1088%2F1126-6708%2F20%2F04%2F063>. DOI: [10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063).
- 3101 [60] CACCIARI M, SALAM G P, SOYEZ G. Fastjet user manual[J/OL]. The European Physical
3102 Journal C, 2012, 72(3):1896. <https://doi.org/10.1140/epjc/s10052-012-1896-2>.
- 3103 [61] AAD G, et al. Jet energy measurement and its systematic uncertainty in proton-proton collisions
3104 at $\sqrt{s} = 7$ TeV with the ATLAS detector[J/OL]. Eur. Phys. J., 2015, C75:17. DOI: [10.1140/epjc/s10052-014-3190-y](https://doi.org/10.1140/epjc/s10052-014-3190-y).
- 3106 [62] AABOUD M, et al. Jet energy scale measurements and their systematic uncertainties in proton-
3107 proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector[J/OL]. Phys. Rev., 2017, D96(7):
3108 072002. DOI: [10.1103/PhysRevD.96.072002](https://doi.org/10.1103/PhysRevD.96.072002).
- 3109 [63] Optimisation of the ATLAS b -tagging performance for the 2016 LHC Run: ATL-PHYS-PUB-
3110 2016-012[R/OL]. Geneva: CERN, 2016. <https://cds.cern.ch/record/2160731>.
- 3111 [64] PIACQUADIO G, WEISER C. A new inclusive secondary vertex algorithm for b -jet tagging
3112 in ATLAS[J/OL]. Journal of Physics: Conference Series, 2008, 119(3):032032. <https://doi.or>
3113 [g/10.1088%2F1742-6596%2F119%2F3%2F032032](https://doi.org/10.1088%2F1742-6596%2F119%2F3%2F032032). DOI: [10.1088/1742-6596/119/3/032032](https://doi.org/10.1088/1742-6596/119/3/032032).
- 3114 [65] SPECKMAYER P, HÖCKER A, STELZER J, et al. The toolkit for multivariate data analysis,

Bibliography

- 3115 TMVA 4[J/OL]. Journal of Physics: Conference Series, 2010, 219(3):032057. <https://doi.org/10.1088%2F1742-6596%2F219%2F3%2F032057>. DOI: [10.1088/1742-6596/219/3/032057](https://doi.org/10.1088/1742-6596/219/3/032057).
- 3116
- 3117 [66] Performance of missing transverse momentum reconstruction with the atlas detector using
3118 proton–proton collisions at $\sqrt{s}=13\text{ TeV}$ [J/OL]. The European Physi-
3119 cal Journal C, 2018, 78(11):903. <https://doi.org/10.1140/epjc/s10052-018-6288-9>.
- 3120 [67] Observation of a new particle in the search for the standard model higgs boson with the atlas
3121 detector at the lhc[J/OL]. Physics Letters B, 2012, 716(1):1 - 29. <http://www.sciencedirect.com/science/article/pii/S037026931200857X>. DOI: <https://doi.org/10.1016/j.physletb.2012.08.020>.
- 3123 [68] Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc[J/OL].
3124 Physics Letters B, 2012, 716(1):30 - 61. <http://www.sciencedirect.com/science/article/pii/S0370269312008581>. DOI: <https://doi.org/10.1016/j.physletb.2012.08.021>.
- 3126 [69] LEE B W, QUIGG C, THACKER H B. The Strength of Weak Interactions at Very High-
3127 Energies and the Higgs Boson Mass[J/OL]. Phys. Rev. Lett., 1977, 38:883-885. DOI: [10.1103/PhysRevLett.38.883](https://doi.org/10.1103/PhysRevLett.38.883).
- 3129 [70] CHANOWITZ M S, GAILLARD M K. The TeV Physics of Strongly Interacting W's and Z's
3130 [J/OL]. Nucl. Phys., 1985, B261:379-431. DOI: [10.1016/0550-3213\(85\)90580-2](https://doi.org/10.1016/0550-3213(85)90580-2).
- 3131 [71] SZLEPER M. The Higgs boson and the physics of WW scattering before and after Higgs
3132 discovery[J]. 2014.
- 3133 [72] Evidence for electroweak production of $W^\pm W^\pm jj$ in pp collisions at $\sqrt{s} = 8$ TeV with the
3134 atlas detector[J/OL]. Phys. Rev. Lett., 2014, 113:141803. <https://link.aps.org/doi/10.1103/PhysRevLett.113.141803>.
- 3135
- 3136 [73] Observation of electroweak production of a same-sign w boson pair in association with two jets
3137 in pp collisions at $\sqrt{s} = 13$ TeV with the atlas detector[J/OL]. Phys. Rev. Lett., 2019, 123:
3138 161801. <https://link.aps.org/doi/10.1103/PhysRevLett.123.161801>.
- 3139 [74] SIRUNYAN A M, et al. Observation of electroweak production of same-sign W boson pairs
3140 in the two jet and two same-sign lepton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV
3141 [J/OL]. Phys. Rev. Lett., 2018, 120(8):081801. DOI: [10.1103/PhysRevLett.120.081801](https://doi.org/10.1103/PhysRevLett.120.081801).
- 3142 [75] Observation of electroweak $w\pm z$ boson pair production in association with two jets in pp col-
3143 lisions at $s=13$ tev with the atlas detector[J/OL]. Physics Letters B, 2019, 793:469 - 492.
3144 <http://www.sciencedirect.com/science/article/pii/S0370269319303211>. DOI: <https://doi.org/10.1016/j.physletb.2019.05.012>.
- 3145
- 3146 [76] Measurement of vector boson scattering and constraints on anomalous quartic couplings from
3147 events with four leptons and two jets in proton–proton collisions at $s=13$ tev[J/OL]. Physics
3148 Letters B, 2017, 774:682 - 705. <http://www.sciencedirect.com/science/article/pii/S0370269317308328>. DOI: <https://doi.org/10.1016/j.physletb.2017.10.020>.
- 3149

Bibliography

- 3150 [77] ALWALL J, FREDERIX R, FRIXIONE S, et al. The automated computation of tree-level and
3151 next-to-leading order differential cross sections, and their matching to parton shower simula-
3152 tions[J/OL]. JHEP, 2014, 07:079. DOI: [10.1007/JHEP07\(2014\)079](https://doi.org/10.1007/JHEP07(2014)079).
- 3153 [78] BALL R D, et al. Parton distributions with LHC data[J/OL]. Nucl. Phys. B, 2013, 867:244-289.
3154 DOI: [10.1016/j.nuclphysb.2012.10.003](https://doi.org/10.1016/j.nuclphysb.2012.10.003).
- 3155 [79] GLEISBERG T, HÖCHE S, KRAUSS F, et al. Event generation with SHERPA 1.1[J/OL].
3156 JHEP, 2009, 02:007. DOI: [10.1088/1126-6708/2009/02/007](https://doi.org/10.1088/1126-6708/2009/02/007).
- 3157 [80] BALL R D, et al. Parton distributions for the LHC run II[J/OL]. JHEP, 2015, 04:040. DOI:
3158 [10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040).
- 3159 [81] CAOLA F, MELNIKOV K, RÖNTSCH R, et al. Qcd corrections to zz production in gluon
3160 fusion at the lhc[J/OL]. Phys. Rev. D, 2015, 92:094028. <https://link.aps.org/doi/10.1103/PhysRevD.92.094028>.
- 3162 [82] FRIXIONE S, RIDOLFI G, NASON P. A positive-weight next-to-leading-order Monte Carlo
3163 for heavy flavour hadroproduction[J/OL]. JHEP, 2007, 09:126. DOI: [10.1088/1126-6708/2007/09/126](https://doi.org/10.1088/1126-6708/2007/09/126).
- 3165 [83] LAI H L, et al. New parton distributions for collider physics[J/OL]. Phys. Rev. D, 2010, 82:
3166 074024. DOI: [10.1103/PhysRevD.82.074024](https://doi.org/10.1103/PhysRevD.82.074024).
- 3167 [84] ALIOLI S, NASON P, OLEARI C, et al. NLO single-top production matched with shower in
3168 POWHEG: s- and t-channel contributions[J/OL]. JHEP, 2009, 09:111. DOI: [10.1088/1126-6708/2009/09/111](https://doi.org/10.1088/1126-6708/2009/09/111).
- 3170 [85] FREDERIX R, RE E, TORRIELLI P. Single-top t-channel hadroproduction in the four-flavour
3171 scheme with POWHEG and aMC@NLO[J/OL]. JHEP, 2012, 09:130. DOI: [10.1007/JHEP09\(2012\)130](https://doi.org/10.1007/JHEP09(2012)130).
- 3173 [86] RE E. Single-top Wt-channel production matched with parton showers using the POWHEG
3174 method[J/OL]. Eur. Phys. J. C, 2011, 71:1547. DOI: [10.1140/epjc/s10052-011-1547-z](https://doi.org/10.1140/epjc/s10052-011-1547-z).
- 3175 [87] COLLABORATION A. ATLAS Pythia 8 tunes to 7 TeV data[M/OL]. 2014. <https://cds.cern.ch/record/1966419>.
- 3177 [88] ATLAS Collaboration. Performance of pile-up mitigation techniques for jets in pp collisions at
3178 $\sqrt{s} = 8$ TeV using the ATLAS detector[J/OL]. Eur. Phys. J. C, 2016, 76:581. DOI: [10.1140/epjc/s10052-016-4395-z](https://doi.org/10.1140/epjc/s10052-016-4395-z).
- 3180 [89] BELLM J, et al. Herwig 7.0/Herwig++ 3.0 release note[J/OL]. Eur. Phys. J. C, 2016, 76(4):
3181 196. DOI: [10.1140/epjc/s10052-016-4018-8](https://doi.org/10.1140/epjc/s10052-016-4018-8).
- 3182 [90] BAHR M, et al. Herwig++ Physics and Manual[J/OL]. Eur. Phys. J. C, 2008, 58:639-707. DOI:
3183 [10.1140/epjc/s10052-008-0798-9](https://doi.org/10.1140/epjc/s10052-008-0798-9).
- 3184 [91] Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the

Bibliography

- 3185 LHC: ATLAS-CONF-2019-021[R/OL]. Geneva: CERN, 2019. <http://cds.cern.ch/record/2677054>.
- 3186
- 3187 [92] The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS[J/OL]. Journal of Instrumentation, 2018, 13(07):P07017-P07017. <https://doi.org/10.1088%2F1748-0221%2F13%2F07%2Fp07017>. DOI: [10.1088/1748-0221/13/07/p07017](https://doi.org/10.1088/1748-0221/13/07/p07017).
- 3190 [93] COADOU Y. Boosted decision trees and applications[J/OL]. EPJ Web of Conferences, 2013, 55:02004-. DOI: [10.1051/epjconf/20135502004](https://doi.org/10.1051/epjconf/20135502004).
- 3192 [94] COLLABORATION A. Technical Design Report for the ATLAS Inner Tracker Pixel Detector: CERN-LHCC-2017-021. ATLAS-TDR-030[R/OL]. Geneva: CERN, 2017. <https://cds.cern.ch/record/2285585>.
- 3195 [95] COLLABORATION A. Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer: CERN-LHCC-2017-017. ATLAS-TDR-026[R/OL]. Geneva: CERN, 2017. <https://cds.cern.ch/record/2285580>.
- 3198 [96] COLLABORATION A. Technical Proposal: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade: CERN-LHCC-2018-023. LHCC-P-012[R/OL]. Geneva: CERN, 2018. <https://cds.cern.ch/record/2623663>.
- 3201 [97] Expected performance for an upgraded ATLAS detector at High-Luminosity LHC: ATL-PHYS-PUB-2016-026[R/OL]. Geneva: CERN, 2016. <https://cds.cern.ch/record/2223839>.
- 3203 [98] HILL A, VAN DER BIJ J J. Strongly interacting singlet-doublet higgs model[J/OL]. Phys. Rev. D, 1987, 36:3463-3473. <https://link.aps.org/doi/10.1103/PhysRevD.36.3463>.
- 3205 [99] BRANCO G, FERREIRA P, LAVOURA L, et al. Theory and phenomenology of two-higgs-doublet models[J/OL]. Physics Reports, 2012, 516(1):1 - 102. <http://www.sciencedirect.com/science/article/pii/S0370157312000695>. DOI: <https://doi.org/10.1016/j.physrep.2012.02.002>.
- 3208 [100] DAVOUDIASL H, HEWETT J, RIZZO T. Bulk gauge fields in the randall–sundrum model1work supported by the department of energy, contract de-ac03-76sf00515.1[J/OL]. Physics Letters B, 2000, 473(1):43 - 49. <http://www.sciencedirect.com/science/article/pii/S0370269399014306>. DOI: [https://doi.org/10.1016/S0370-2693\(99\)01430-6](https://doi.org/10.1016/S0370-2693(99)01430-6).
- 3211
- 3212 [101] FREDERIX R, FRIXIONE S. Merging meets matching in MC@NLO[J/OL]. JHEP, 2012, 12: 061. DOI: [10.1007/JHEP12\(2012\)061](https://doi.org/10.1007/JHEP12(2012)061).
- 3213
- 3214 [102] AGASHE K, DAVOUDIASL H, PEREZ G, et al. Warped Gravitons at the LHC and Beyond [J/OL]. Phys.Rev., 2007, D76:036006. DOI: [10.1103/PhysRevD.76.036006](https://doi.org/10.1103/PhysRevD.76.036006).
- 3216 [103] ATLAS Collaboration. Jet reconstruction and performance using particle flow with the ATLAS Detector[J/OL]. Eur. Phys. J. C, 2017, 77:466. DOI: [10.1140/epjc/s10052-017-5031-2](https://doi.org/10.1140/epjc/s10052-017-5031-2).
- 3217
- 3218 [104] 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[J/OL]. Phys. Rev. D,
- 3219

Bibliography

- 3220 2015, 91:012006. <https://link.aps.org/doi/10.1103/PhysRevD.91.012006>.
- 3221 [105] CRANMER K S. Kernel estimation in high-energy physics[J/OL]. Comput. Phys. Commun.,
3222 2001, 136:198-207. DOI: [10.1016/S0010-4655\(00\)00243-5](https://doi.org/10.1016/S0010-4655(00)00243-5).
- 3223 [106] OREGLIA M. A Study of the Reactions $\psi' \rightarrow \gamma\gamma\psi$ [M/OL]. SLAC, 1980. <https://www.slac.stanford.edu/cgi-wrap/getdoc/slac-r-236.pdf>.
- 3225 [107] GAISER J. Charmonium Spectroscopy From Radiative Decays of the J/ψ and ψ' [M/OL].
3226 SLAC, 1982. <https://www.slac.stanford.edu/cgi-wrap/getdoc/slac-r-255.pdf>.
- 3227 [108] GORIA S, PASSARINO G, ROSCO D. The Higgs Boson Lineshape[J/OL]. Nucl.Phys., 2012,
3228 B864:530-579. DOI: [10.1016/j.nuclphysb.2012.07.006](https://doi.org/10.1016/j.nuclphysb.2012.07.006).
- 3229 [109] SPIRA M, DJOUADI A, GRAUDENZ D, et al. Higgs boson production at the LHC[J]. Nucl.
3230 Phys., 1995, B 453:17-82.
- 3231 [110] BALL R D, et al. Parton Distribution Benchmarking with LHC Data[J/OL]. JHEP, 2013, 04:
3232 125. DOI: [10.1007/JHEP04\(2013\)125](https://doi.org/10.1007/JHEP04(2013)125).
- 3233 [111] BIJNENS J, EEROLA P, MAUL M, et al. Qcd signatures of narrow graviton resonances in
3234 hadron colliders[J/OL]. Physics Letters B, 2001, 503(3-4):341-348. DOI: [10.1016/s0370-2693\(01\)00238-6](https://doi.org/10.1016/s0370-2693(01)00238-6).
- 3236 [112] ATLAS Collaboration. PMG Systematic Uncertainty Recipes[EB/OL]. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PmgSystematicUncertaintyRecipes>.
- 3238 [113] READ A L. Presentation of search results: The CL(s) technique[J]. J. Phys. G, 2002, 28:
3239 2693-2704.
- 3240 [114] AABOUD M, et al. Search for heavy ZZ resonances in the $\ell^+\ell^-\ell^+\ell^-$ and $\ell^+\ell^-\nu\bar{\nu}$ final states
3241 using proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector[J/OL]. Eur. Phys. J.
3242 C, 2018, 78(4):293. DOI: [10.1140/epjc/s10052-018-5686-3](https://doi.org/10.1140/epjc/s10052-018-5686-3).