

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 衰变道的事例对电弱对称性破缺过程的研 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： 二〇二〇年十一月一日

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



**Studies of Electroweak Symmetry
Breaking in ZZ Production in Purely
Leptonic Decay with ATLAS Detector**

Author: Heling Zhu

Speciality: Particle and Nuclear Physics

Supervisors: Zhengguo Zhao (USTC), Hong Ma (BNL)

Finished time: Nov. 1st, 2020

中国科学技术大学学位论文原创性声明

本人声明所呈交的学位论文，是本人在导师指导下进行研究工作所取得的成果。除已特别加以标注和致谢的地方外，论文中不包含任何他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的贡献均已在论文中作了明确的说明。

作者签名：_____

签字日期：_____

中国科学技术大学学位论文授权使用声明

作为申请学位的条件之一，学位论文著作权拥有者授权中国科学技术大学拥有学位论文的部分使用权，即：学校有权按有关规定向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅，可以将学位论文编入《中国学位论文全文数据库》等有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。本人提交的电子文档的内容和纸质论文的内容相一致。

保密的学位论文在解密后也遵守此规定。

公开 保密（____ 年）

作者签名：_____

导师签名：_____

签字日期：_____

签字日期：_____

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。LHC 是当今世界上最大的、能量最高的对撞机，在其上发生的质子-质子对撞实验质心系能量目前已达 13 TeV，其设计质心系能量最高可达 14 TeV。LHC 凭借其空前的能量，在理论与实验之间架起了一座重要桥梁。ATLAS 实验是 LHC 上的一个通用粒子探测器实验，同时也是体积最大的粒子探测器。该实验旨在利用 LHC 提供的超高能量和亮度来观测涉及大质量粒子的物理现象，而这些粒子是使用早期的低能加速器所无法观察到的。同时，ATLAS 实验也是 2012 年 7 月 LHC 上参与发现希格斯粒子的两个重要实验之一。

本文基于 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，对 ZZ 玻色子衰变到四轻子末态的过程进行研究。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。首先， ZZ 到四轻子过程截面的测量结果为 $\sigma_{ZZjj}^{tot} = 1.27 \pm 0.14 [fb]$ ，总体相对误差为 11%，其中 9% 的误差来源于数据的统计涨落。在误差范围内，该结果和标准模型预言值 $1.14 \pm 0.20 [fb]$ 相吻合。同时，在 ZZ 玻色子伴随着两个喷注 (jet) 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 倍标准差的明显偏差。在此基础上，本文也讨论了使用亮度为 3000 fb^{-1} 模拟数据，在质心能量为 14 TeV 的高亮度大型强子对撞机 (HL-LHC) 上产生矢量玻色子散射过程在 ZZ 到四轻子末态的研究前景，预言将可以观测到大于 7 倍标准差的明显偏差。

本文还介绍了在 ZZ 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。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 模型的引力子的理论质量下限，为 1500 GeV。

ZZ 玻色子衰变到四轻子的过程具有很少的本底和极高的信噪比，是进行诸如检验标准模型与精确测量希格斯粒子等粒子物理前沿研究的重要过程。本文所介绍的 ZZ 玻色子衰变到四轻子末态过程的结果都与标准模型预测的一致。该结果

摘 要

³⁴ 完善了矢量玻色子散射过程在各个衰变末态的观测，是对电弱对称性破坏研究的
³⁵ 新的里程碑。同时，没有发现明显的超出标准模型的新物理过程。我们期待，大大
³⁶ 提高了亮度和能量的下一代高亮度大型强子对撞机（HL-LHC）实验能够为这些新
³⁷ 物理过程的寻找带来更高的灵敏度，和对如四轻子末态这样的稀有物理过程更精
³⁸ 确的测量。

39

40

ABSTRACT

41 This dissertation presents my research in the field of Particle Physics with the ATLAS
42 experiment at the Large Hadron Collider (LHC). The LHC is the largest and most powerful
43 collider in the world, and it is the proton-proton collider with the centre-of-mass energy
44 currently at 13 TeV. With its unprecedented energy, the LHC was built as a bridge between
45 the theories and the experiment. The ATLAS experiment is a general-purpose particle
46 detector experiment with the largest volume at the LHC. The experiment was designed
47 to take advantage of the extremely high energy and luminosity available at the LHC and
48 observe phenomena that involve highly massive particles which were not observable using
49 earlier lower-energy accelerators. And it was one of the two LHC experiments involved
50 in the discovery of the Higgs boson in July 2012.

51 This dissertation focus on the studies with two Z bosons production decaying into
52 $\ell\ell\ell'\ell'$ final state, where ℓ stands for electron or muon, using 139 fb^{-1} of 13 TeV proton-
53 proton (pp) collision data collected by ATLAS experiment at the LHC. With small amount
54 of backgrounds and very high signal-to-background ratio, the ZZ production in $\ell\ell\ell'\ell'$
55 channel provides a most clean and sensitive tool to test the Standard Model (SM) at the
56 energy frontier and to study the *Higgs* physics. First of all, studies including the measure-
57 ment on SM $ZZjj$ production cross section and the observation of Vector Boson Scat-
58 tering (VBS) process. The fiducial cross section for SM $ZZjj$ production is measured
59 to be $\sigma_{ZZjj}^{tot} = 1.27 \pm 0.14 [\text{fb}]$, where 9% out of 11% uncertainty is from data statistic,
60 and found to be compatible with the SM prediction of $1.14 \pm 0.20 [\text{fb}]$. The electroweak
61 production of two jets in association with a Z -boson pair (EW- $ZZjj$) is observed with a
62 significant deviation from the background-only hypothesis corresponding to a statistical
63 significance of 5.5σ . Following with the observation, the prospect study for the EW-
64 $ZZjj$ production at the High luminosity LHC (HL-LHC) using 3000 fb^{-1} simulated pp
65 collision data at a centre-of-mass energy of 14 TeV is presented, with a expected signifi-
66 cance of around 7σ .

67 Meanwhile, a search for heavy resonances decaying into a pair of Z bosons to $\ell\ell\ell'\ell'$
68 final state is also conducted in this dissertation. Different mass ranges for the hypothet-
69 ical resonances are considered, depending on the signal models and spanning between
70 200 GeV and 2000 GeV. Data is found to agree with a background-only hypothesis, thus,
71 the results are interpreted as upper limits on production cross section for sevaral differ-
72 ent models, including heavy Higgs like (spin-0) narrow-width approximation (NWA) and

Abstract

73 large-width approximation (LWA), as well as the Randall–Sundrum model with a gravi-
74 ton excitation spin-2 resonance (RSG). The signal is assumed to generate dominantly via
75 gluon-gluon Fusion (ggF) production mode and Vector Boson Fusion (VBF) production
76 mode. Both ggF and VBF channels are studied in NWA, while for LWA, only ggF chan-
77 nel is studied due to worse resolution in higher mass region and the lack of statistic for
78 VBF process. In addition, mass of RS Graviton is constrained, $m(G_{KK}) < 1500$ GeV is
79 excluded at 95% CL by $ZZ \rightarrow \ell\ell\ell'\ell'$ analysis.

80 The ZZ to $\ell\ell\ell'\ell'$ production presented in this dissertation are consistent with the
81 SM prediction. This result completes the observation of weak boson scattering that is
82 a new milestone in the study of electroweak symmetry breaking. In the meantime, no
83 indication of new physics is observed. We are looking forward to the HL-LHC, with
84 greatly increased luminosity and higher centre-of-mass energy, which should enhance the
85 sensitivity for new physics search and precise measurement for rare process.

86

87

Acknowledgments

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

96 Thanks to Hong, for giving me the oppotunity to study in Brookhaven National Lab
97 (BNL) that I can work with so many senior and brilliant physicists, and teaching me a lot
98 in the details of physics. As the chair of physics department at BNL, can you imagine
99 that he managed to take time sitting with me every week, teaching me the details and
100 techniques of my analysis as well as helping me to parepre my talks at conference.

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

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

113 Moreover, I really want to give my sincere grattitude to Dr. Xu Lailin. Thank you,
114 Lailin, for all your helps during the passing five years. Thanks for teaching me in all the
115 analysis details, coding techniques, presentation skills hand by hand. You are really a
116 very good and patient teacher and give me as many knowledges as you can. Your broad
117 knowledge, your perseverance in science and your very hard working indeed affect me a
118 lot.

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

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

122 I would like to express my gratitude to many colleagues in both USTC and BNL team.
123 Thanks to Prof. Sun Yongjie, who is the supervisor of my undergraduate thesis, helped
124 me start my first detector project on MRPC. Also thanks Dr. Liu Zhen who taught me
125 in details in this project, and helped me a lot for my life at BNL too. Thanks to Prof.
126 Peng Haiping, Prof. Zhu Yingchun and Dr. Hu Qipeng for helping me all the details
127 and techniques in HWW analysis when I was a beginner. Thanks to Dr. Dai Tiesheng,
128 I have learnt quite a lot in the project of Monitored Drift Tubes (MDT) when working
129 with you at CERN and also thank you for all the help in regular life since that was my
130 first time to Europe. Thanks to Dr. Chen Hucheng, who is the supervisor of my ATLAS
131 qualification task on LAr Trigger Digitizer Boards (LTDB) and the person lead me into
132 this interesting electronic project. And my appreciation to Dr. Xu Hao, Dr. Chen Kai and
133 Dr. Liu Hongbin who taught me in patience for this project as I was really a freshman on
134 electronics. Thanks to Prof. Yuji Enari and Dr. Georges Aad for the help in LAr software
135 tasks when I moved from BNL to CERN. In the meantime, I would really like to give
136 my gratitude to Dr. Michael Begal, Dr. Marc-Andre Pleier, Dr. Alessandro Tricoli, Dr.
137 George Redlinger, Dr. Viviana Cavalieri, Dr. Gaetano Barone and many senior physicists
138 in BNL omaga group. I have learnt a lot from every chat with you and every seminar you
139 hosted. Also I want to thanks to Prof. Wu Yusheng, Prof. Qian Jianming, Prof. Liu
140 Yanwen, Dr. Ju Xiangyang for teaching me in lots of details in different physics analyses.

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

149 Last but not least, I would give my greatest gratitude to my families. Thanks my
150 parents for giving me all your endless loves and supports in my whole life. My deep
151 appreciation to my three aunts for loving and caring me so much since I was born and
152 help me to accompany with my mother when I was thousands miles far away from home.
153 And my husband, Lin, thank you for your understanding and walk through all difficulties
154 with me in these years especially when we were in a foreign country.

155		
156	Acknowledgments	V
157	Chapter 1 Introduction	1
158	Chapter 2 Theory	4
159	2.1 Elementary Particles and Interactions in the Standard Model	4
160	2.1.1 Elementary particles in the Standard Model	4
161	2.1.2 Electroweak theory	5
162	2.1.3 Higgs mechanism and Electroweak symmetry breaking	7
163	2.1.4 Beyond the SM Higgs sector	10
164	2.2 Phenomenology of Large Hadron Collider	11
165	2.2.1 Physics at hadron colliders	12
166	2.2.2 Higgs physics at the LHC	14
167	2.2.3 Diboson physics at the LHC	16
168	Chapter 3 The Large Hadron Collider and the ATLAS Detector	21
169	3.1 The Large Hadron Collider	21
170	3.1.1 Operation history and machine layout	21
171	3.1.2 Luminosity and pile-up	24
172	3.2 ATLAS detector	26
173	3.2.1 Detector overview	26
174	3.2.2 Physics requirement	28
175	3.2.3 Magnet system	28
176	3.2.4 Inner detector	29
177	3.2.5 Calorimeters	32
178	3.2.6 Muon spectrometer	33
179	3.2.7 Trigger system	35
180	Chapter 4 Simulation and Event Reconstruction for the ATLAS Experiment	39
181		
182	4.1 Event simulation	39
183	4.2 Reconstruction	41
184	4.2.1 Track	42
185	4.2.2 Primary vertex	44
186	4.2.3 Electron	45

187	4.2.4 Muon	50
188	4.2.5 Jets	54
189	4.2.6 Missing transverse energy	58
190	Chapter 5 Statistical treatment of searching for new particles or processes	61
192	5.1 The likelihood function	61
193	5.2 Test statistic	62
194	5.3 The CLs upper limit	64
195	5.4 Nuisance parameters	65
196	Chapter 6 Studies of SM ZZ production in $\ell\ell\ell'\ell'$ final state using pp collision data collected by ATLAS detector from 2015 to 2018	66
199	6.1 Introduction	66
200	6.2 Data and MC samples	67
201	6.2.1 Data samples	67
202	6.2.2 MC simulations	68
203	6.3 Objects and Event selections	69
204	6.3.1 Object selections	69
205	6.3.2 Event selections	70
206	6.4 Background estimation	71
207	6.4.1 QCD backgrounds	71
208	6.4.2 Reducible backgrounds	72
209	6.5 Systematic uncertainties	76
210	6.5.1 Theoretical systematics	76
211	6.5.2 Experimental systematics	78
212	6.6 Measurement of fiducial cross section	81
213	6.6.1 Calculation of C-factor	81
214	6.6.2 Result of fiducial cross section	81
215	6.7 Search for EW- $ZZjj$	82
216	6.7.1 MD discriminant	82
217	6.7.2 Fitting procedure	82
218	6.7.3 Result of fit	83
219	6.8 Prospect study of EW- $ZZjj$ production in HL-LHC	86
220	6.8.1 The ATLAS detector at HL-LHC	87

221	6.8.2 Simulation	87
222	6.8.3 Event selection	87
223	6.8.4 Systematics	89
224	6.8.5 Results	90
225	6.9 Conclusion	93
226	Chapter 7 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	94
229	7.1 Introduction	94
230	7.2 Data and MC samples	95
231	7.2.1 Data samples	95
232	7.2.2 Background MC simulations	95
233	7.2.3 Signal MC simulations	96
234	7.3 Object and event selections	97
235	7.3.1 Object selections	97
236	7.3.2 Event selections	98
237	7.3.3 Studies of τ -channel in particle-level	100
238	7.3.4 Event categorizations	106
239	7.3.5 Signal acceptance	111
240	7.4 Background estimation	112
241	7.4.1 Irreducible backgrounds	113
242	7.4.2 Reducible backgrounds	117
243	7.5 Signal modelling	118
244	7.5.1 Modelling of narrow-width signal	119
245	7.5.2 Modelling of large-width signal	119
246	7.5.3 Modelling of interference	123
247	7.5.4 Modelling of spin-2 RS Graviton signal	127
248	7.6 Systematic uncertainties	128
249	7.6.1 Theoretical systematics	129
250	7.6.2 Experimental systematics	131
251	7.7 Results in $\ell\ell\ell'\ell'$ channel	131
252	7.7.1 Statistical procedure	132
253	7.7.2 Fit to likelihood function under background-only hypothesis for MVA-based analysis	133
255	7.7.3 Interpretations	137

Contents

256	7.8 Conclusion	144
257	Chapter 8 Summary	145
258	Bibliography	148
259	Publications	158

260	List of Figures	
261	2.1 The elementary particles of the Standard Model.	4
262	2.2 The Feynman diagrams of interactions mediated by gauge bosons that form the basis of the standard model.	6
264	2.3 The Higgs potential $V(\phi)$ with $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right).	8
265	2.4 Schematic of a hadron-hadron collision ^[14]	12
266	2.5 The PDF4LHC15 NLO PDFs at a low scale $\mu^2 = Q^2 = 4GeV^2$ (left) and at $\mu^2 = Q^2 = 100GeV^2$ (right) as a function of x.	13
268	2.6 Feynman diagrams of four Higgs production modes: (a) ggF; (b) VBF; (c) VH; (d) ttH.	15
270	2.7 The SM Higgs boson (125 GeV) production cross section for various produc- tion modes as a function of the centre-of-mass energy for pp collision.	15
272	2.8 Higgs boson production cross section for various production modes as a func- tion of the Higgs mass for $\sqrt{s} = 13$ TeV (left) and 14 TeV (right) for pp collision.	15
275	2.9 SM Higgs decay channels.	16
276	2.10 Branching ratio of Higgs decays in various channels as a function of Higgs mass ^[20]	17
278	2.11 The tree-level Feynman diagrams of diboson production at the LHC.	17
279	2.12 Total production cross-section measured by ATLAS as a function of centre- of-mass energy \sqrt{s} from 7 to 13 TeV for some selected processes. The mea- surements of diboson processes are scaled by a factor 0.1 to avoid the over- laps.	18
283	2.13 Feynman diagrams of the vector boson scattering process.	18
284	2.14 Feynman diagrams of diboson productions with vertexes involving QGC, TGC and Higgs.	19
286	3.1 Cumulative luminosity vs time in the years from 2011 to 2018 for ATLAS detector.	22
288	3.2 Layout of CERN LHC complex ^[23]	23
289	3.3 Integrated luminosity vs delivered month from 2015 to 2018 in ATLAS ex- periment.	25

List of Figures

291	3.4	Number of interactions per crossing weighted by luminosity from 2015 to 2018 in ATLAS experiment.	25
292	3.5	Coordinate system used by the ATLAS experiment at the LHC ^[25]	26
293	3.6	Layout view of the ATLAS detector ^[26]	27
294	3.7	Schematic view of the ATLAS magnet system.	29
295	3.8	Schematic of the inner detector in ATLAS ^[28]	30
296	3.9	Schematic diagram of the 4-Layer Pixel Detector in ATLAS run-2.	30
297	3.10	SCT (a) barrel module and (b) end-cap ^[32]	31
298	3.11	Layout of the ATLAS calorimeters. The scintillator-based tile hadronic calorimeters are seen outside the LAr calorimeters ^[34]	32
299	3.12	Layout of a LAr EM calorimeter barrel module ^[35]	33
300	3.13	Schematic of tile calorimeter module ^[36]	34
301	3.14	Cut-away view of the muon spectrometer in ATLAS ^[38]	34
302	3.15	Schematic diagram of the ATLAS trigger and data acquisition system in run-2.	36
303	3.16	An examples of L1 calorimeter trigger tower for electron and photon triggers ^[40]	37
304	3.17	The HLT trigger algorithm sequence ^[40]	38
305	4.1	Schematic diagram of the ATLAS simulation software ^[42]	39
306	4.2	Schematic of a hadron-hadron collision event simulated by MC generator. The red blob in center denotes the hard collision, surrounded by tree-like structures denoting Bremsstrahlung from Parton Showers. The purple blob is a secondary hard scattering event. The light green blobs represent the parton-to-hadron transitions while the dark green blobs stand for hadron decays. The yellow lines indicate soft photon radiations.	41
307	4.3	The flowchart of the ATLAS data processing ^[52]	42
308	4.4	Cut-away view of the ATLAS inner detector.	43
309	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 ^[62]	47
310	4.6	The efficiencies of three electron identification WPs from $Z \rightarrow ee$ (left) events and hadrons misidentified as electrons estimated using di-jet MC samples (right).	48
311	4.7	$E_T^{cone0.2}$ (left) and $p_T^{varcone0.2}$ (right) distribution for electrons from $ZZ \rightarrow ee$ events in data and MC simulation. The simulated events (full histograms) are normalized to data.	49

325	4.8	Muon reconstruction efficiency as functions of η for: Medium (and Loose), Tight and High- p_T working points.	53
326	4.9	Muon reconstruction efficiency for Low- p_T working point as a function of η	53
327	4.10	Distributions of the calorimeter-based (right) and the track-based (left) rela- tive isolation variables measured in $Z \rightarrow \mu\mu$ events.	54
328	4.11	A overview schematic of ATLAS jet reconstruction ^[68]	55
329	4.12	A overview schematic of ATLAS jet calibration ^[69]	56
330	4.13	MV2c10 BDT output for b- (solid blue), c- (dashed green) and light-flavour (dotted red) jets in $t\bar{t}$ events ^[70]	58
331	4.14	Measured E_T^{miss} distribution for $Z \rightarrow \mu\mu$ events (left) and $W \rightarrow e\nu$ events (right).	60
332	5.1	(a) Illustration of the relationship between the observed t_μ and its p -value. (b) The relationship between p -value and the observed significance Z , where $\phi(x)$ is the standard normal distribution ^[74]	63
333	6.1	Typical diagrams for the production of $ZZjj$, including the relevant EW VBS diagrams (first row) and QCD diagrams (second row).	67
334	6.2	Pre-fit m_{ZZ} and m_{jj} distribution in QCD-enriched CR.	73
335	6.3	Fake factor for $Z+jets$ background, constructed with additional electron, as a function of p_T (left) and η (right).	74
336	6.4	Fake factor for $Z+jets$ background, constructed with additional muon, as a function of p_T (left) and η (right).	74
337	6.5	Fake factor for $t\bar{t}$ background, constructed with additional electron, as a func- tion of p_T (left) and η (right).	75
338	6.6	Fake factor for $t\bar{t}$ background, constructed with additional muon, as a function of p_T (left) and η (right).	75
339	6.7	The theoretical uncertainties for $gg \rightarrow ZZ$ background in particle-level SR (left) and CR (right).	78
340	6.8	MD distribution for QCD- $ZZjj$ process in low and high pile-up events for SR (left) and CR (right).	80
341	6.9	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).	80

357	6.10	Observed and post-fit 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.	84
362	6.11	Observed and post-fit 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 uncer- tainty. The contributions from the QCD and EW production of $ZZjj$ events are scaled by 0.96 and 1.35, respectively, corresponding to the observed nor- malization factors in the statistical fit. The last bin includes the overflow events.	85
369	6.12	Observed and post-fit 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.	85
375	6.13	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.	86
378	6.14	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}	89
381	6.15	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.	90
384	6.16	The expected significance of EW- $ZZjj$ processes as a function of differ- ent m_{jj} cut with 3000 fb^{-1} , under conditions of different sizes of theoretical uncertainties on the QCD- $ZZjj$ background modelling. The statistical un- certainty is estimated from expected data yield at 14 TeV with 3000 fb^{-1} . Different uncertainties are summed up quadratically.	91

389	6.17	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.	91
394	7.1	$Z(\rightarrow\tau\tau)$ boson invariant mass distribution for signal at 300 GeV of (a) M_{Z_1} : the leading pairs; (b) M_{Z_2} : the sub-leading pairs. The definition of leading and sub-leading pairs can be found in section 7.3.2.	102
397	7.2	Significance scan at the m_H of (a) 300 GeV; (b) 600 GeV; (c) 1000 GeV; (d) 1200 GeV, for events in lepton flavor channel with or without τ -lepton.	104
399	7.3	$Z(\rightarrow\tau\tau)$ boson invariant mass distribution for signal at 300 GeV of (a) M_{Z_1} : the leading pairs; (b) M_{Z_2} : the sub-leading pairs. The definition of leading and sub-leading pairs can be found in section 7.3.2.	105
402	7.4	The comparison of $m_{4\ell}$ distribution in the channels of ‘have tau’: events in $2e2\tau$ and $2\mu2\tau$ final states; ‘only em’: events in $2e2\mu$, $4e$ and 4μ final states; for three simulated samples: (a) NWA signal at 300 GeV; (b) NWA signal at 1000 GeV; (c) $q\bar{q} \rightarrow ZZ$ background.	105
406	7.5	Significance scan at the m_H of (a) 300 GeV; (b) 1000 GeV; for events in lepton flavor channel with or without τ -lepton.	106
408	7.6	(a) VBF DNN architecture diagram. (b) ggF DNN architecture.	107
409	7.7	(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.	108
412	7.8	(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.	108

415	7.9	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 100 times of the observed limit for the “ggF-classifier” and 30 times of the observed limit for the “VBF -classifier”. The ZZ backgrounds are scaled by the normalisation factors shown in Table 7.12. The lower panels show the ratio of data to prediction. Statistical and experimental systematic uncertainties are included.	110
424	7.10	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.	110
431	7.11	Illustration of the MVA-based VBF and ggF event classification for events with (a) $N_{\text{jets}} < 2$ and (b) $N_{\text{jets}} \geq 2$	112
433	7.12	NWA acceptance as a function of m_H for the MVA-based categorization for the samples of (a) ggF production; (b) VBF production.	112
435	7.13	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.	113
437	7.14	Distributions of the $m_{4\ell}$ invariant mass fit projections of the $q\bar{q} \rightarrow ZZ$ background samples for the 4μ , $4e$ and $2\mu 2e$ final states in the ggF-CBA-enriched category, and the 4ℓ inclusive VBF-CBA-enriched category. Cut-based categorization is used.	114
441	7.15	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.	115
445	7.16	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.	115

449	7.17 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.	116
450		
451		
452		
453	7.18 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.	116
454		
455		
456		
457	7.19 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.	117
458		
459		
460		
461	7.20 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.	120
462		
463		
464		
465	7.21 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.	121
466		
467		
468	7.22 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.	121
469		
470		
471	7.23 The final signal shapes for the ggF production mode, interpolated from the polynomial fit parameters.	121
472		
473	7.24 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.	123
474		
475		
476	7.25 Comparison between the analytical shape convoluted with detector effects and the reconstructed $m_{2\mu2e}$ MC distribution for mass points ranging from 400 to 1800 GeV and width equal to 15% of the mass.	124
477		
478		
479	7.26 The interference (H-B) model fitted to the truth $m_{4\ell}$ MC distribution after signal region selection for $2\mu2e$ channel.	126
480		

481	7.27	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).	128
486	7.28	Fitted parameters of the graviton RBW, m_{RBW} and Γ_{RBW} , as a function of the graviton resonance mass, m_G	129
488	7.29	Reconstructed $m_{4\ell}$ distributions in the $2\mu 2e$ channel with the final sig- nal 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.	129
493	7.30	Distribution of the four-lepton invariant mass $m_{4\ell}$ 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 like- lihood fit to the data under the background-only hypothesis. The simulated signal at 600 GeV is normalized to a cross section corresponding to 50 (5) times the observed upper limit given in section 7.7.3 for the ggF (VBF) pro- duction. The error bars on the data points indicate the statistical uncertainty, and the systematic uncertainty in MC prediction is shown by the hatched band. The lower panels show the ratio of data to prediction. The red arrows indicate data points that are outside the displayed range.	135
503	7.31	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.	136
507	7.32	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.	137
510	7.33	The expected and observed upper limits at 95% CL on $\sigma \times B(H \rightarrow ZZ)$ 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.	139

514	7.34 The expected and observed upper limits at 95% CL on $\sigma \times B(H \rightarrow ZZ)$ using 515 the cut-based analysis for ggF (left) and VBF (right) production. The green 516 and yellow bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties in the expected 517 limits.	139
518	7.35 The local p0 scan as functions of m_H on $\sigma \times B(H \rightarrow ZZ)$ using the MVA- 519 based analysis for ggF (left) and VBF (right) production.	140
520	7.36 The upper limits at 95% confidence level on $\sigma_{ggF} \times B(H \rightarrow ZZ)$ as a function 521 of the heavy resonance mass m_H for the ggF production mode with an intrinsic 522 width of 1% (top left), 5% (top right), 10% (bottom left) and 15% (bottom 523 right) for both the case where interference with Standard Model processes 524 is considered. The green and yellow bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ 525 uncertainties in the expected limits.	141
526	7.37 The upper limits at 95% confidence level on $\sigma_{ggF} \times B(G_{KK} \rightarrow ZZ)$ as a 527 function of the heavy resonance mass $m(G_{KK})$ for the ggF production mode 528 in RS Graviton model. The green and yellow bands represent the $\pm 1\sigma$ and 529 $\pm 2\sigma$ uncertainties in the expected limits.	141
530	7.38 The expected (left) and observed (right) upper limits at 95% CL on $\sigma \times B(S \rightarrow$ 531 $ZZ)$ for ggF production mode at different hypothetical models including the 532 MVA- and cut- based NWA, LWA with width assumptions of 1, 5, 10, 15%, 533 and the RSG.	142
534	7.39 Comparisons of the expected upper limits at 95% CL on the cross section 535 times branching ratio as a function of the heavy resonance mass m_H for 536 the ggF production mode (left) and for the VBF production mode (right) in 537 the case of the NWA. The expected limits from the previous publication are 538 shown in the green dashed line and are projected to the 139 fb^{-1} as shown in 539 the blue dashed line. In addition, the current results based on either cut-based 540 categorisation or the multivariate-based categorisation are shown in red and 541 black lines.	142
542	7.40 Display of one candidate event in 4μ final state with the mass of 1.35 TeV.	143

List of Tables

543	3.1	Summary of design parameters of the LHC for pp collisions.	22
545	4.1	Overview of the contributions to $E_T^{\text{miss}}[73]$.	59
546	6.1	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.	70
550	6.2	Summary of selection of physics objects and candidate events at detector level in the $\ell\ell\ell'\ell'jj$ signal region.	71
552	6.3	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.	71
555	6.4	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 6.2.2, 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.	72
560	6.5	Fake background estimations in the SR. For 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.	76
563	6.6	Summary of different variations for EW- and QCD- $ZZjj$ theoretical uncertainties measurement.	77
565	6.7	Summary of theoretical uncertainties for the fiducial volume (SR) for both EW and QCD qq -initial processes.	77
567	6.8	Experimental systematic 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.	79
570	6.9	Measured and predicted fiducial cross-sections in $\ell\ell\ell'\ell'jj$ final-state. Uncertainties due to different sources are presented.	82
572	6.10	Input features for the training of MD.	83

573	6.11	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.	84
574			
575			
576	6.12	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}	88
577			
578			
579			
580			
581	6.13	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.	92
582			
583			
584			
585	7.1	Summary of the recorded integrated luminosity (lumi), average and peak pile-up (PU) of data from 2015 to 2018.	95
586			
587	7.2	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.	98
588			
589	7.3	Summary of the object and event selection requirements.	101
590	7.4	The expected number of events for signal and $q\bar{q} \rightarrow ZZ$ background in the lepton flavor channels with or without τ -lepton. The event yields are normalized to the luminosity of 80 fb^{-1}	103
591			
592			
593	7.5	After using MMC, the expected number of events for signal and $q\bar{q} \rightarrow ZZ$ background in the lepton flavor channels with or without τ -lepton. The event yields are normalized to the luminosity of 80 fb^{-1}	104
594			
595			
596	7.6	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.	109
597			
598			
599	7.7	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.	111
600			
601			
602	7.8	Summary of acceptance uncertainties of PDF, QCD scale and parton shower variations for ggF production. The MVA-based categorization is used.	130
603			
604	7.9	Summary of acceptance uncertainties of PDF, QCD scale and parton shower variations for VBF production. The MVA-based categorization is used.	130
605			

606	7.10	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.	131
607			
608			
609	7.11	Summary of CL_s upper limits on ggF production mode at 300, 600, 1000 GeV.	133
610			
611			
612	7.12	ZZ normalization factor in each category, obtained from a likelihood fit to the data under the background-only hypothesis in MVA-based analysis.	133
613			
614			
615	7.13	ZZ normalization factor in each category, obtained from a likelihood fit to the data under the background-only hypothesis in Cut-based analysis.	133
616			
617			
618			
619	7.14	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 7.12, are also taken into account.	133
620			
621	7.15	Expected and observed numbers of events for $m_{4\ell} > 200$ GeV, together with their systematic uncertainties, for two Cut-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 7.13, are also taken into account.	134
622			
623			
624			
625			
626			
627	7.16	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.	138
628			
629			
630			
631			

632 **Chapter 1 Introduction**

633 The research of particle physics is aiming to understand how our universe works at
634 its fundamental level. It can be accomplished by pursuing the mysteries of the basic
635 construction of matter and energy, probing the interactions between elementary particles,
636 and studying the nature of time and space.

637 **Elementary particles**

638 From around the 6th century BC, ancient Greek philosophers Leucippus, Democritus,
639 and Epicurus brought up a philosophical idea that everything is composed of “uncuttable”
640 elementary particles. In the 19th century, John Dalton, through his work on stoichiometry,
641 concluded that each element of nature was composed of a single, unique type of particle.
642 The particle was named as “atom” after the Greek word atomos, with the meaning of “in-
643 divisible”. However this Dalton’s atom theory was strongly challenged later. Near the end
644 of 19th century, physicists discovered that Dalton’s atoms are not, in fact, the fundamental
645 particles of nature, but conglomerates of even smaller particles. Electron was discovered
646 by J. J. Thomson in 1897^[1], and then its charge was then carefully measured by Robert
647 Andrews Millikan and Harvey Fletcher in their “oil drop experiment” of 1909^[2]. In early
648 20th-century, Rutherford’s “gold foil experiment” showed that the most mass of atom is
649 concentrated in a small positive charge nucleus^[3]. Then the discoveries of anti-particles
650 (the positron in 1932) and other particles (e.g. the muon in 1936) indicate that more dis-
651 coveries could be expected in future experiments.

652 Starting from 1950s, more accelerator facilities were put into service. During the
653 1950s and 1960s, various particles were found in particle collisions from increasingly
654 high-energy beams, informally referred to as the “particle zoo”. In 1964, the quark model
655 was independently proposed by physicists Murray Gell-Mann and George Zweig, and
656 experimentally confirmed of their existence in mid-1970s. In 1970s, the establishment
657 of quantum chromodynamics (QCD) postulated the fundamental strong interaction, ex-
658 perienced by quarks and mediated by gluons.

659 The well-known Standard model (SM) was developed during the latter half of
660 the 20th century. At that time, confirmation of the top quark (1995), the tau neu-
661 trino (2000), and the Higgs boson (2012) have added further credence to the SM.
662 Now, the quarks, leptons and gauge bosons are the elementary particles in the frame-
663 work of Standard Model of particle physics, that theoretically describes three of
664 the four known fundamental forces (the electromagnetic, weak, and strong interactions,

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

667 **Higgs mechanics and electroweak symmetry breaking**

668 In 1961, Sheldon Glashow, Steven Weinberg and Abdus Salam together brought for-
669 ward a unified electroweak theory to combine the electromagnetic and weak interactions.
670 In the standard model, under the condition that the energy is high enough but electroweak
671 symmetry is unbroken, all elementary particles are massless. But measurements show the
672 fact that the W and Z bosons actually have masses. Later on, the Higgs mechanics re-
673 solves this conundrum. The description of the mechanism adds a Higgs field in all space
674 of the Standard Model, where the field causes spontaneous symmetry breaking during in-
675 teractions, and all massive particles in the Standard Model, including the W and Z bosons,
676 interact with Higgs boson to acquire their mass.

677 Over the past few decades, with the combination of electroweak theory, the Higgs
678 mechanics has been widely accepted. But the Higgs boson, the essential part to explain
679 this mechanics of the property “mass” for gauge bosons and fermions, had been the final
680 missing piece in the Standard Model of particle physics at that time. The mass of Higgs
681 boson was not specifically predicted by the SM, and it has been searched in several large
682 experiments (eg. LEP at CERN, Tevatron at Fermilab, and LHC at CERN) with different
683 energy. In 2012, the discovery of Higgs boson was finally announced by the ATLAS and
684 CMS collaborations at the Large Hadron Collider (LHC) with its mass round 125 GeV.
685 Peter Higgs and Francois Englert were award the Nobel Prize in Physics in the year of
686 2013 for their theoretical discovery of this mechanism that contributes to our understand-
687 ing of the origin of mass.

688 **Contents of this thesis**

689 This dissertation is organized as follows. Section 2 briefly introduces the Standard
690 Model of particle physics, the Higgs mechanism and electroweak symmetry breaking re-
691 lated to the dissertation and the LHC phenomenology. Section 3 gives an overview of the
692 LHC and the ATLAS detector. The detector simulation and the reconstruction of physics
693 objects are described in section 4. Section 5 introduces the statistical treatment of search-
694 ing for new particles and processes commonly used in experimental particle physics.

695 And then section 6 focuses on the Standard model ZZ production cross section mea-
696 surement in $ZZ \rightarrow \ell\ell\ell'\ell'$ channel, where ℓ stands for electron or muon, and the observa-
697 tion of its electroweak component as well as its further prospects in High-luminosity LHC
698 (HL-LHC). The author has mainly contributed in truth-level events studies in this anal-
699 ysis, including the MadGraph5_aMC@NLO samples generations and validations, stud-

700 ies of theoretical uncertainties and experimental shape uncertainties on QCD $q\bar{q} \rightarrow ZZ$
701 background modelling, the event displays, as well as all works in prospective studies of
702 HL-LHC including event selection, background and systematic estimations, and signifi-
703 cance and differential cross-section interpretations.

704 Section 7 presents the search of possible heavy resonances in both spin-0 and spin-2
705 hypothetical models in $ZZ \rightarrow \ell\ell\ell'\ell'$ channel, and the results of search are interpreted
706 as 95% CLs upper limits on the production cross section of resonance. The author has
707 involved into this analysis starting from the events selection optimization (eg. lepton
708 isolation optimization, muon p_T cut optimization and τ -lepton channel studies), as well
709 as all theoretical and experimental systematics studies, signal modelling and acceptance
710 parameterization studies, statistical fits and the combination between $\ell\ell\ell'\ell'$ and $\ell\ell\nu\nu$
711 channels in final paper.

712 In the end, section 8 gives the summary and outlook for future physics in LHC.

713

Chapter 2 Theory

714 2.1 Elementary Particles and Interactions in the Standard Model

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

721 2.1.1 Elementary particles in the Standard Model

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

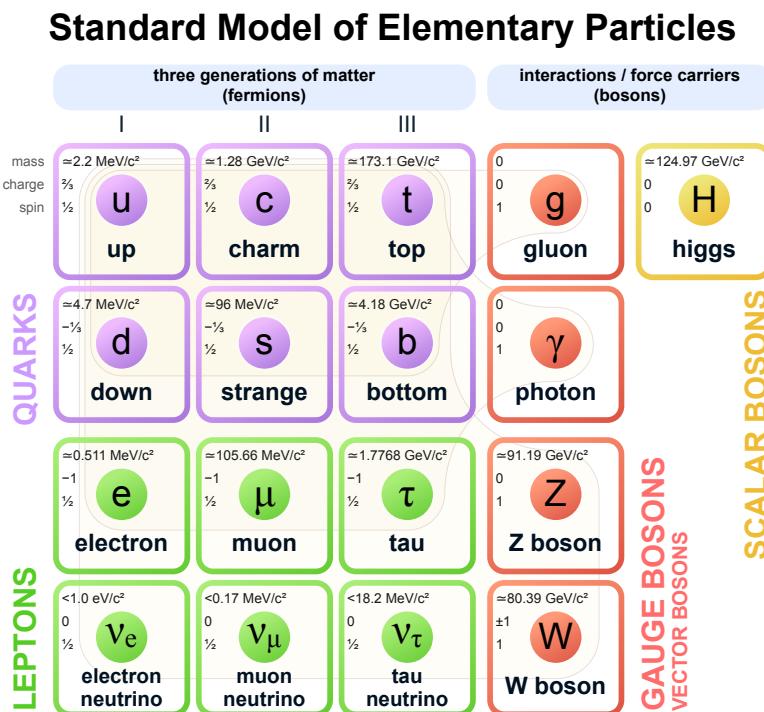


Fig. 2.1 The elementary particles of the Standard Model.

723

724 **Fermions** The Standard Model includes 12 elementary particles of spin- $\frac{1}{2}$ obeying the
 725 Fermi-Dirac statistics, known as fermions. They are classified into two types: *leptons* and
 726 *quarks* according to their interactions. The *leptons* include three generations: electron
 727 (e) and electron neutrino (ν_e); muon (μ) and muon neutrino (ν_μ); tau (τ) and tau neutrino
 728 (ν_τ). The e , μ and τ carry electric charge of -1 and three neutrinos are electrically neutral.

729 All the leptons can participate in electroweak interactions. Also there are three generations
 730 of *quarks*: up (*u*) and down (*d*); charm (*c*) and strange (*s*); top (*t*) and bottom (*b*). The
 731 defining property of the quarks is that they carry color charge (while leptons don't), and
 732 hence interact via the strong interaction, letting them to be strongly bound from one to
 733 another, forming color-neutral composite particles (known as hadrons) containing either
 734 a quark and an antiquark (mesons) or three quarks (baryons). In the meantime, *u*, *c* and
 735 *t*-quark carry electric charge of 2/3, and *d*, *s* and *b*-quark carry electric charge of -1/3.
 736 Hence they interact via all three interactions described in SM. Each fermion also has its
 737 corresponding antiparticle.

738 **Gauge bosons** act as force carriers that propagate the strong, weak, and electromag-
 739 netic interactions in SM. They are spin-1 particles obeying the Bose-Einstein statistics.
 740 There are three types of gauge bosons:

- 741 • The eight massless *gluons* propagate the strong interactions between color charged
 742 particles (quarks).
- 743 • The massless *photons* propagate the electromagnetic force between electrically
 744 charged particles.
- 745 • The W^+ , W^- and Z bosons propagate the weak interactions between both quarks
 746 and leptons. All these three bosons are massive, the W^\pm carries an electric charge
 747 of +1 and -1 and can also couple to the electromagnetic interaction while Z boson
 748 is electrically neutral.

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

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

753 2.1.2 Electroweak theory

754 The electroweak interaction is the unified description of two of the four known fun-
 755 damental interactions of nature: electromagnetism and the weak interaction. It is based
 756 on the gauge group $SU(2)_L \times SU(1)_Y$, in which L is the left-handed fields and Y is the
 757 weak hypercharge^[4]. It follows the Lagrangian of

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

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

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

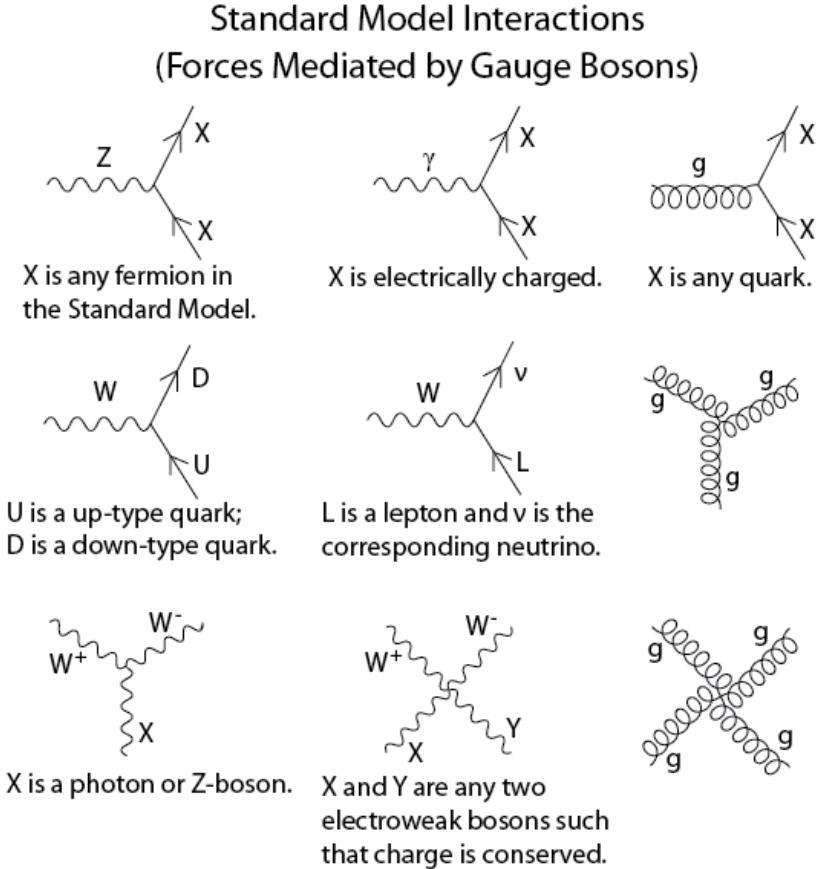


Fig. 2.2 The Feynman diagrams of interactions mediated by gauge bosons that form the basis of the standard model.

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

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

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

764 The second term is the **scalar part**:

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

765 where $\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$ denotes a complex Higgs scalar, and $V(\phi)$ is the Higgs potential which
766 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^[4]s

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

in which τ^i represents the Pauli matrices, and g' is the $U(1)_Y$ gauge coupling. The square of the covariant derivative results in three and four-point interactions between the gauge and scalar fields.

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

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

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

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

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 the right-hand singlets.

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

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

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

2.1.3 Higgs mechanism and Electroweak symmetry breaking

As shown in previous subsection, the Lagrangian L_{gauge} does not involve any mass term due to the requirement of gauge invariance. So all the W and B bosons should be massless. But experimental observations show that the gauge bosons are massive. Therefore, the gauge invariance must be broken spontaneously. The Higgs field is introduced

788 to break the $SU(2)_L \times U(1)_Y$ symmetry and gauge bosons and fermions can interact with
 789 Higgs field to acquire their masses. And this specific process is named *Higgs mechanism*
 790 in SM.

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

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

792 where $\phi_i = \phi_i^+$ stand for four Hermitian field. In this new basis, the Higgs potential in
 793 Eq. 2.5 can be expressed as:

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

794 To simplify the situation, the axis in this four-dimensional space can be chosen to satisfied
 795 $\langle 0 | \phi_i | 0 \rangle = 0$ for $i = 1, 2, 4$, and $\langle 0 | \phi_3 | 0 \rangle = v$. Thus,

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

796 The minimization of this potential depends on the sign of μ^2 as shown in figure 2.3. When
 797 $\mu^2 > 0$ the minimum occurs at $v = 0$, namely the vacuum is empty space and $SU(2)_L \times$
 798 $U(1)_Y$ symmetry is unbroken. In the case of $\mu^2 < 0$, the $v = 0$ symmetric point is no
 799 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

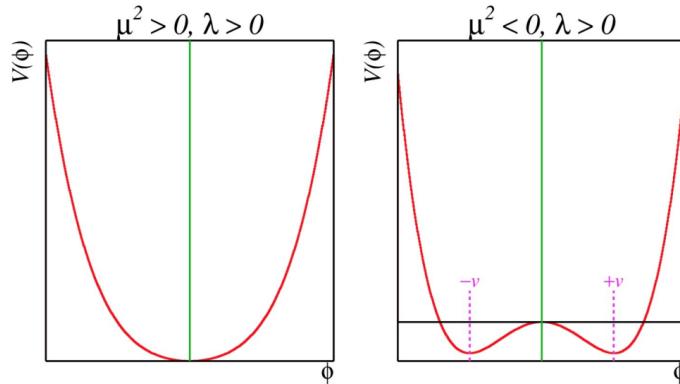


Fig. 2.3 The Higgs potential $V(\phi)$ with $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right).

800

801 expressed by

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

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

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

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

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

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

805 where the W and Z fields are

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

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

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

808 where θ_W is the weak angle defined as

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

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

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

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

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

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

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

813 Due to the unknown of quartic Higgs coupling λ , the Higgs mass is not predicted. The

814 third and fourth terms in the Higgs potential V denote the induced cubic and quartic in-
 815 teractions of the Higgs scalar.

816 Through the Higgs mechanism, fermions can also acquire their masses. In the unitary
 817 gauge, Yukawa Lagrangian (L_{Yukawa}) can be written as a simple form of^[5]

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

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

819 2.1.4 Beyond the SM Higgs sector

820 After the discovery of the Higgs boson by the ATLAS and CMS Collaborations at
 821 the LHC^[6-7] in 2012, one question comes out: if this Higgs boson at around 125 GeV is
 822 fully responsible for the unitarization of the scattering amplitudes? The possibility that
 823 this discovered particle is just a part of the extended Higgs sector by various extensions
 824 cannot be ruled out. Many models, motivated by hierarchy and naturalness arguments,
 825 predicted the extended Higgs sector, such as the electroweak-singlet model and the two-
 826 Higgs-doublet models (2HDM).

827 Singlet scalar extension of the SM

828 The electroweak singlet model can be considered as the minimal extension of the SM
 829 Higgs sector^[8], encompassing a single gauge singlet real scalar field S . In this model, a
 830 heavy, real singlet is introduced in addition to the SM one. The associated zero tempera-
 831 ture, tree-level scalar potential can be written as:

$$V = V_{SM} + V_{HS} + V_S \quad (2.23)$$

832 where

$$\begin{aligned} V_{SM} &= \mu^2 (H^\dagger H) + \bar{\lambda}_0 (H^\dagger H) \\ V_{HS} &= \frac{a_1}{2} (H^\dagger H) S + \frac{a_2}{2} (H^\dagger H) S^2 \\ V_S &= \frac{b_2}{2} S^2 + \frac{b_3}{3} S^3 + \frac{b_4}{4} S^4 \end{aligned} \quad (2.24)$$

833 where H stands for the SM scalar field of the original Higgs mechanism. After elec-
 834 troweak symmetry breaking, this model gives rise to two CP -even Higgs bosons, in which
 835 the lighter one is the Higgs boson that has been discovered at around 125 GeV. And the
 836 new heavy scalar (S) is allowed to have both SM and non-SM decays. One would expect
 837 to see suppressions of the branching ratio to SM Higgs decay modes, as the branching
 838 ratio to the pair of singlet-like scalars would be considerable.

839 **Two Higgs Doublet Model**

840 The two-Higgs-doublet model (2HDM)^[9] is another extension of SM Higgs sector carried
841 by an additional scalar doublet. In this model, through electroweak symmetry breaking,
842 there are five physical Higgs bosons: two CP-even, one CP-odd, and two charged ones.

843 The most general CP-conserving 2HDM has seven free parameters:

- 844 • The masses of five Higgs bosons: m_h , m_H , m_A and m_{H^\pm} .
845 • $\tan\beta$: v_1/v_2 , where v_1 and v_2 are the two Higgs doublets' vacuum expectation val-
846 ues.
847 • α : the two neutral CP-even Higgs bosons mixing angle .
848 • m_{12}^2 : the potential parameter mixing the two Higgs doublets.

849 where the m_h can be identified as the mass of observed Higgs boson at around 125 GeV,
850 and m_H is another heavy scalar with similar properties to h boson. The coupling of the
851 neutral Higgs bosons to either WW or ZZ follows the rules^[9]:

- 852 1. The coupling of the light Higgs (h) equals to the Standard Model coupling times
853 $\sin(\beta - \alpha)$
- 854 2. The coupling of the heavier Higgs (H) equals to the Standard Model coupling times
855 $\cos(\alpha - \beta)$.
- 856 3. The coupling of the pseudoscalar (A) to vector bosons is zero.

857 The two Higgs doublets, Φ_1 and Φ_2 , can couple to fermions (leptons and up- and down-
858 type quarks) in several ways, which leads to several types of 2HDM models:

- 859 • Type-I model: all quarks and leptons couple only to Φ_2 .
860 • Type-II model: down-type quarks and leptons couple to Φ_1 , and up-type quarks
861 couple to Φ_2 .
862 • The “lepton-specific” model: leptons couple to Φ_1 , while all quarks couple to Φ_2 .
863 • The “flipped” model: down-type quarks couple to Φ_1 , while up-type quarks and
864 leptons couple to Φ_2 .

865 **2.2 Phenomenology of Large Hadron Collider**

866 The Large Hadron Collider (LHC)^[10-12] was built as a bridge between the experiments
867 and the theories. Physicists hope that the LHC can help to answer some open questions
868 in fundamental physics, such as the basic laws of interactions, the forces among the el-
869 ementary particles, the deep structure of space and time, and the interrelation between
870 quantum mechanics and general relativity. This section will talk about firstly the general
871 introduction of Physics inside hadronic collision, then followed by two important LHC

872 phenomenologies of the Higgs physics and Diboson physics that are related closely to this
 873 dissertation.

874 2.2.1 Physics at hadron colliders

875 Protons are not elementary particle, but actually consist of quarks and gluons. So in
 876 proton-proton (pp) collision at the LHC, it is not protons themselves interact but quarks
 877 and gluons. Scattering processes can then be further classified into either *hard* or *soft*
 878 processes according to the momentum transfer during the interaction^[13]. QCD, as an un-
 879 derlying theory for both processes, its approach and level of understandings in two cases
 880 are quite different. For hard process, eg. Higgs boson, vector bosons and jets production,
 881 the perturbation theory can be used to precisely predict their rates and properties. How-
 882 ever, for soft processes like total cross-section or the underlying events, the rates and event
 883 properties are dominated by non-perturbative QCD effects that are less understood. For
 884 many hard processes, the hard interactions are accompanied by soft ones. An example of
 the hadronic collision is illustrated in figure 2.4. and the typical features are summarized

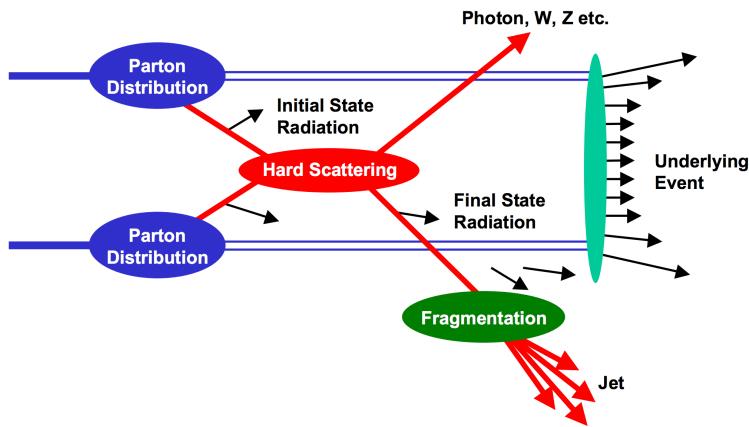


Fig. 2.4 Schematic of a hadron-hadron collision^[14].

885

886 as below:

- 887 • **Parton Distribution Function (PDF):** $f_i(x, Q^2)$ gives the probability of finding a
 888 parton with flavor i (quark or gluon), carrying a momentum fraction of x and with
 889 the scale of momentum transfer Q in a proton. Parton distribution function cannot
 890 be fully calculated by perturbative QCD because of the inherent non-perturbative
 891 nature of partons. There are many different sets of PDFs that are determined by
 892 fits to data from experimental observables in various processes. As an example,
 893 figure 2.5 shows the *PDF4LHC15 NLO PDFs*, which are based on the combination
 894 of the *CT14*, *MMHT14* and *NNPDF3.1 NNLO* PDF sets^[15].

- 895 • **Fragmentation and hadronization:** The processes to produce final state particles
 896 and jets from the partons produced in hard scattering.
- 897 • **Initial/Final state radiation:** The incoming and outgoing partons that carry color
 898 charge can emit QCD radiation, which gives rise to additional jets. Also the charged
 899 incoming and outgoing leptons can emit Quantum Electrodynamics (QED) radia-
 900 tions with photons.
- 901 • **Underlying events:** Products from soft processes (not come from the primary hard
 902 scattering) as the remnants of scattering interactions.

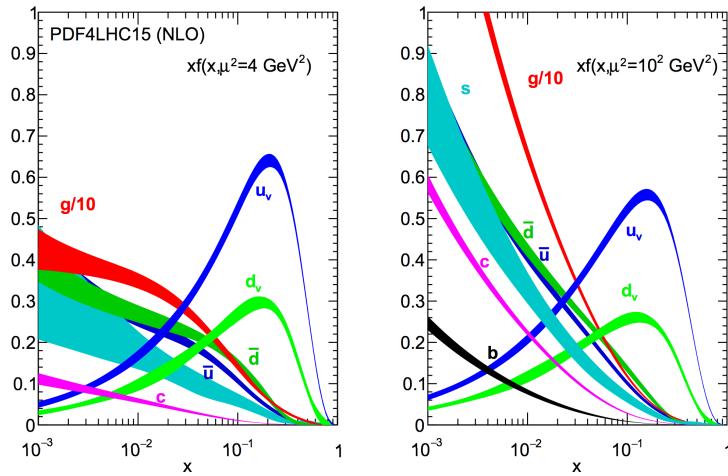


Fig. 2.5 The PDF4LHC15 NLO PDFs at a low scale $\mu^2 = Q^2 = 4\text{GeV}^2$ (left) and at $\mu^2 = Q^2 = 100\text{GeV}^2$ (right) as a function of x .

903 Cross section of hard scattering

904 According to *QCD factorization theorems*^[16], the perturbative calculations can be ap-
 905 plied to many important hard processes involving hadrons. The basic problem addressed
 906 by factorization theorems is how to calculate high energy cross sections. Consider the
 907 process of scattering between A and B hadron and producing a final state X , the cross
 908 section σ can be obtained by summing over all the subprocess cross section $\hat{\sigma}$ ^[17]

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

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

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}(x_a Q^2) f_{b/B}(x_b Q^2) \hat{\sigma}_{ab \rightarrow X} \quad (2.26)$$

911 where Q^2 represents large momentum scale that characterizes the hard scattering. Later
 912 on, since the finite corrections were not universal and had to be calculated separately for

913 each process, the perturbative $O(\alpha_S^n)$ corrections to the leading logarithm cross section
 914 in Eq. 2.26 need to be applied, one can get:

$$\sigma_{AB} = \int dx_a dx_b f_{a/A}(x_a \mu_F^2) f_{b/B}(x_b \mu_F^2) \hat{\sigma}_{ab \rightarrow X}(\alpha_S, \mu_R, \mu_F) \quad (2.27)$$

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

$$\hat{\sigma}_{ab \rightarrow X}(\alpha_S, \mu_R, \mu_F) = (\alpha_S)^n \left[\hat{\sigma}^{(0)} + (\alpha_S/2\pi) \hat{\sigma}^{(1)}(\mu_R, \mu_F) + (\alpha_S/2\pi)^2 \hat{\sigma}^{(2)}(\mu_R, \mu_F) + \dots \right] \quad (2.28)$$

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

921 μ_R and μ_F depend on the order of truncation in Eq. 2.28. In principle, if cross section
 922 is calculated to all orders, it is invariant under changes in these parameters. The choices
 923 of μ_R and μ_F are arbitrary, and $\mu_R = \mu_F$ is often assumed. It is advisable to choose the
 924 value of μ_R and μ_F on the order of the typical momentum scales of the hard scattering
 925 process to avoid unnaturally large logarithms appearing in the perturbation series. For
 926 example in Drell–Yan process, the standard choice is that both μ_R and μ_F are set to the
 927 invariant mass of di-lepton pair.

928 2.2.2 Higgs physics at the LHC

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

932 Higgs productions

933 The Higgs boson can be produced through several processes. There are 4 main produc-
 934 tion modes at the LHC: gluon-gluon Fusion (ggF), vector boson Fusion (VBF), associated
 935 production with vector-bosons (VH) (also called the Higgs Strahlung) and associated pro-
 936 duction with a pair of top/anti-top quarks ($t\bar{t}H$)^[18]. Figure 2.6 shows the corresponding
 937 Feynman diagrams of each process at leading order. For pp collisions, the cross section of
 938 production of Higgs boson is a function of centre-of-mass energy \sqrt{s} . Figure 2.7 depicts
 939 the cross section of SM Higgs, whose mass is 125 GeV, for several different produc-
 940 tion modes when centre-of-mass energy varying from 6 to 15 TeV. Figure 2.8 shows the
 941 prospect of production cross section as a function of Higgs mass from 10 to 2000 GeV for

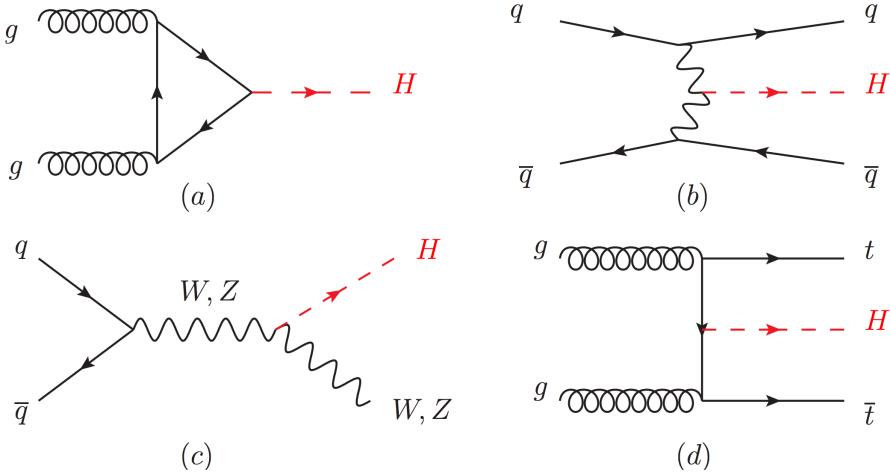


Fig. 2.6 Feynman diagrams of four Higgs production modes: (a) ggF; (b) VBF; (c) VH; (d) ttH.

pp collision at the centre-of-mass energy of 13 TeV and 14 TeV [19].

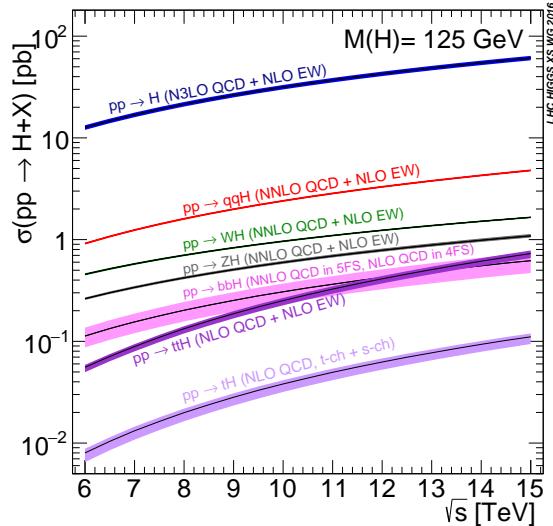


Fig. 2.7 The SM Higgs boson (125 GeV) production cross section for various production modes as a function of the centre-of-mass energy for pp collision.

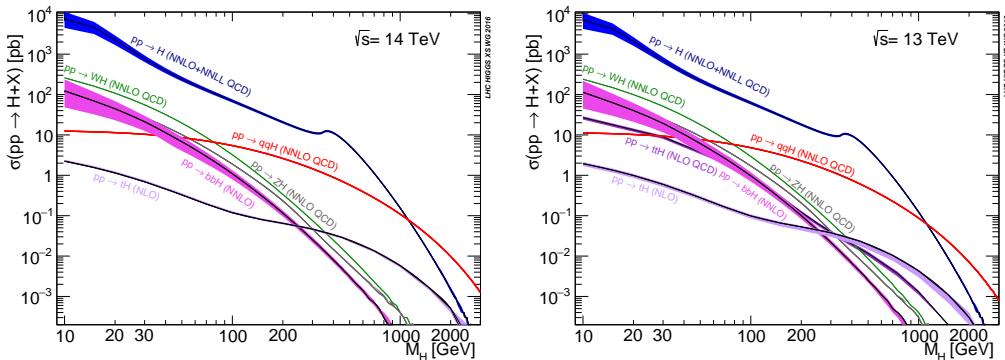


Fig. 2.8 Higgs boson production cross section for various production modes as a function of the Higgs mass for $\sqrt{s} = 13$ TeV (left) and 14 TeV (right) for pp collision.

943 **Higgs decays**

944 The Higgs boson can interact with gauge bosons and fermions through gauge coupling
 945 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

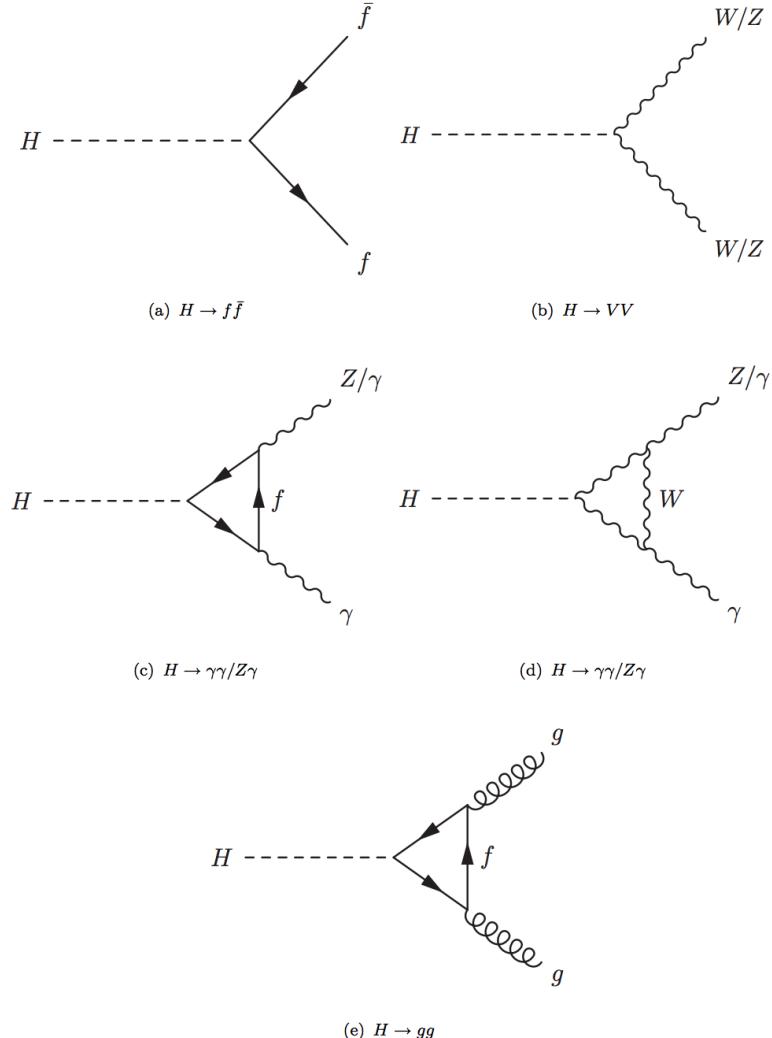


Fig. 2.9 SM Higgs decay channels.

946

947 decaying into different final states as a function of Higgs mass is shown in figure 2.10.

948 **2.2.3 Diboson physics at the LHC**

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

954 **Diboson productions**

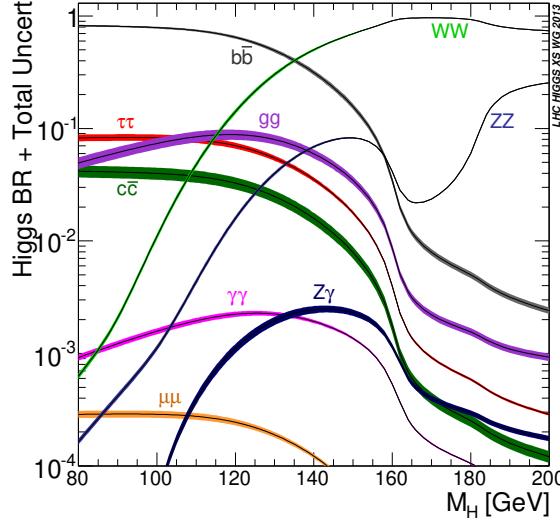


Fig. 2.10 Branching ratio of Higgs decays in various channels as a function of Higgs mass^[20].

955 About 90% of diboson productions at hadron collider is from quark-antiquark anni-
 956 hilation, while others are from gluon initiated process. Figure 2.11 shows the tree-level
 Feynman diagrams of diboson production. Then figure 2.12 illuminates the total produc-

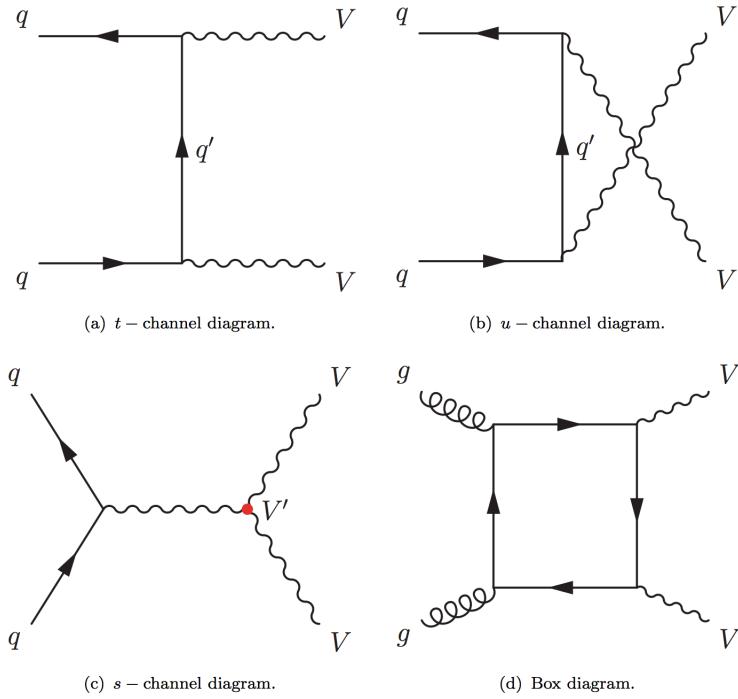


Fig. 2.11 The tree-level Feynman diagrams of diboson production at the LHC.

957
 958 tion cross-section presented by ATLAS as a function of centre-of-mass energy \sqrt{s} from
 959 7 to 13 TeV for several diboson processes compared to some other major processes in
 960 hadron collision. The cross section for diboson processes are calculated at next-to-next
 961 leading order (NNLO).

962 **Vector boson scattering**

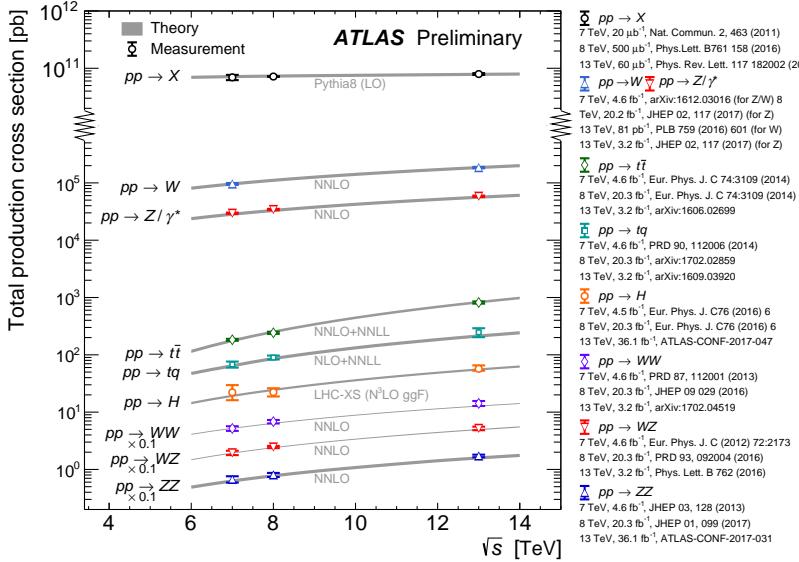


Fig. 2.12 Total production cross-section measured by ATLAS as a function of centre-of-mass energy \sqrt{s} from 7 to 13 TeV for some selected processes. The measurements of diboson processes are scaled by a factor 0.1 to avoid the overlaps.

963 The $SU(2)_L \times U(1)_Y$ structure in SM predicts self-interactions between electroweak
 964 gauge bosons. Those self-couplings can involve either three or four gauge bosons at a
 965 single vertex, known as triple gauge coupling (*TGC*) or quartic gauge couplings (*QGC*),
 966 respectively. Vector boson scattering (*VBS*) is carried out by four electroweak vector
 967 bosons, namely Z , W^\pm and photon (γ) as the Feynman diagrams shown in figure 2.13.
 968 And the vertexes include either those self-interactions or the interactions with the Higgs
 boson are described in figure 2.14.

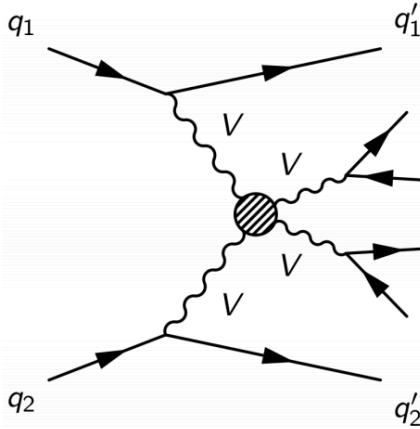


Fig. 2.13 Feynman diagrams of the vector boson scattering process.

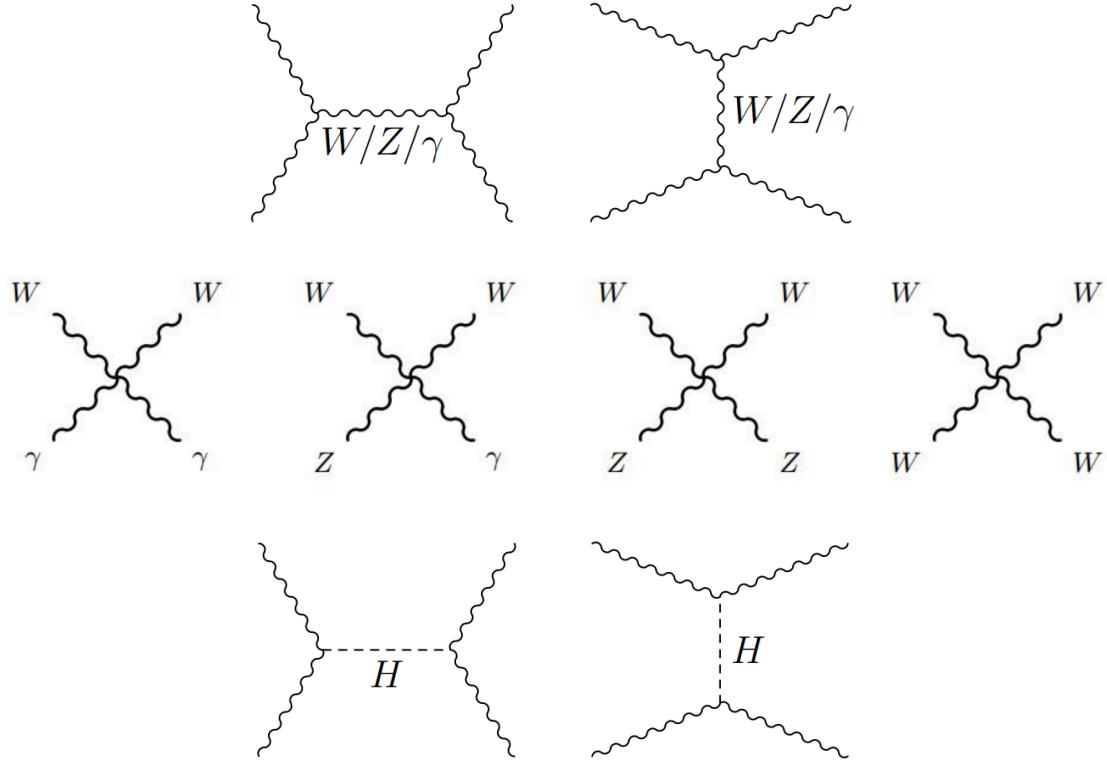


Fig. 2.14 Feynman diagrams of diboson productions with vertexes involving QGC, TGC and Higgs.

970 The amplitudes of leading-order (LO) VBS can be expressed as^[21]:

$$iM_{TGC}^{s-channel} = -i \frac{g_1^2}{4m_W^4} [s(t-u) - 3m_W^2(t-u)] \quad (2.29)$$

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

971

$$iM_{QGC} = i \frac{g_1^2}{4m_W^4} \left[s^2 + 4st + t^2 - 4m_W^2(s+t) - \frac{8m_W^2}{s} ut \right] \quad (2.30)$$

972

$$iM_{Higgs} = -i \frac{C_v^2 g_1^2}{4m_W^2} \left[\frac{(s-2m_W^2)^2}{s-m_H^2} + \frac{(t-2m_W^2)^2}{t-m_H^2} \right]$$

$$\simeq -i \frac{C_v^2 g_1^2}{4m_W^2} (s+t) \quad (2.31)$$

973 Combining s- and t-channel of TGC in Eq. 2.29 and the QGC term in Eq. 2.30:

$$iM_{TGC} + iM_{QGC} = i \frac{g_1^2}{4m_W^2} (s+t) + O((s/m_W^2)^0) \quad (2.32)$$

974 In Eq. 2.32, the amplitude grows as a function of centre-of-mass energy (\sqrt{s}), which

975 violates the unitarity in the TeV region. Considering the Higgs term in Eq. 2.31 can
976 perfectly cancel out this growing, and the remaining term $O((s/m_W^2)^0)$ only depends on
977 the total amplitude in SM.

978 In conclusion, the Higgs boson acts as “moderator” to unitarize high-energy longi-
979 tudinal vector boson scattering as introducing the Higgs restores the unitarity of total
980 amplitude in high energy region.

Chapter 3 The Large Hadron Collider and the ATLAS Detector

983 3.1 The Large Hadron Collider

Located near the French-Swiss border at the European Organization for Nuclear Research (CERN), the Large Hadron Collider (LHC) is the largest and most powerful facility for particle physics in the world. It's the proton-proton collider with the centre-of-mass energy designed up to 14 TeV. The beams inside the LHC are made to collide at four locations around its 27-kilometer accelerator ring, corresponding to four particle experiments - the ATLAS, CMS, ALICE and LHCb. With its unprecedented energy, the LHC is designed to observe physics that involve highly massive particles, which have never been observed in previous accelerators with lower energies.

3.1.1 Operation history and machine layout

Operation history

The LHC^[10-12,22] is a two-ring-superconducting-hadron accelerator and collider lies in a tunnel about 100 metres underground. It's designed to provide proton-proton collisions at the centre-of-mass energy up to 14 TeV with a unprecedented luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. In the meantime, it can also collide heavy (Pb) ions with an energy of 2.8 TeV per nucleon and a peak luminosity of $10^{27} \text{ cm}^{-2} \text{ s}^{-1}$. Table 3.1 shows the main design parameters of the LHC for proton-proton collisions.

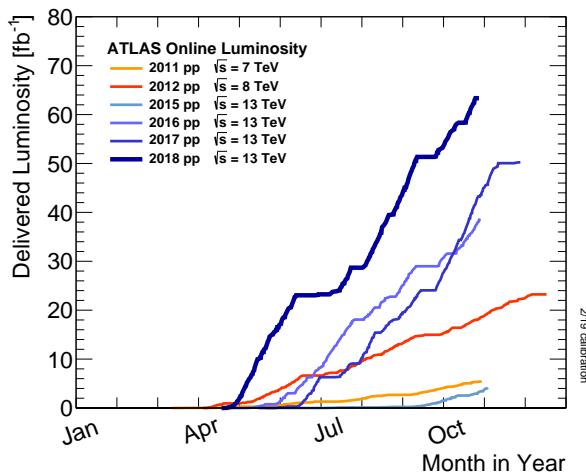
The LHC was built from 1998 to 2008. It started its first beam in September 2008, but then was interrupted by a quench incident only after a few days running. Then it resumed the operation in November 2009 with a low energy beams. From March 2010, physics runs took place at the centre-of-mass energy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

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

Machine layout

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

Circumference	27 km
Collision's beam energy	7 TeV
Injected beam energy	0.45 TeV
Dipole field at 7 TeV	8.3 T
Luminosity	$10^{34} \text{ cm}^{-2} \text{ s}^{-1}$
Beam current	0.56 A
Number of protons per bunch	1.1×10^{11}
Number of bunches	2808
Nominal bunch spacing	25 ns
Normalized emittance	$3.8 \mu\text{m}$
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 vs time in the years from 2011 to 2018 for ATLAS detector.**

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 ring. At beginning, the 50 MeV protons are produced in the linear particle accelerator LINAC2, and further accelerated to 1.4 GeV in Proton Synchrotron Booster (PSB). The protons are then injected into the Proton Synchrotron (PS) to gain the energy of 26 GeV and further

1018 accelerated to 450 GeV in Super Proton Synchrotron (SPS). At the end, they are injected
 1019 into the main ring, and can reach a maximum energy of 7 TeV.

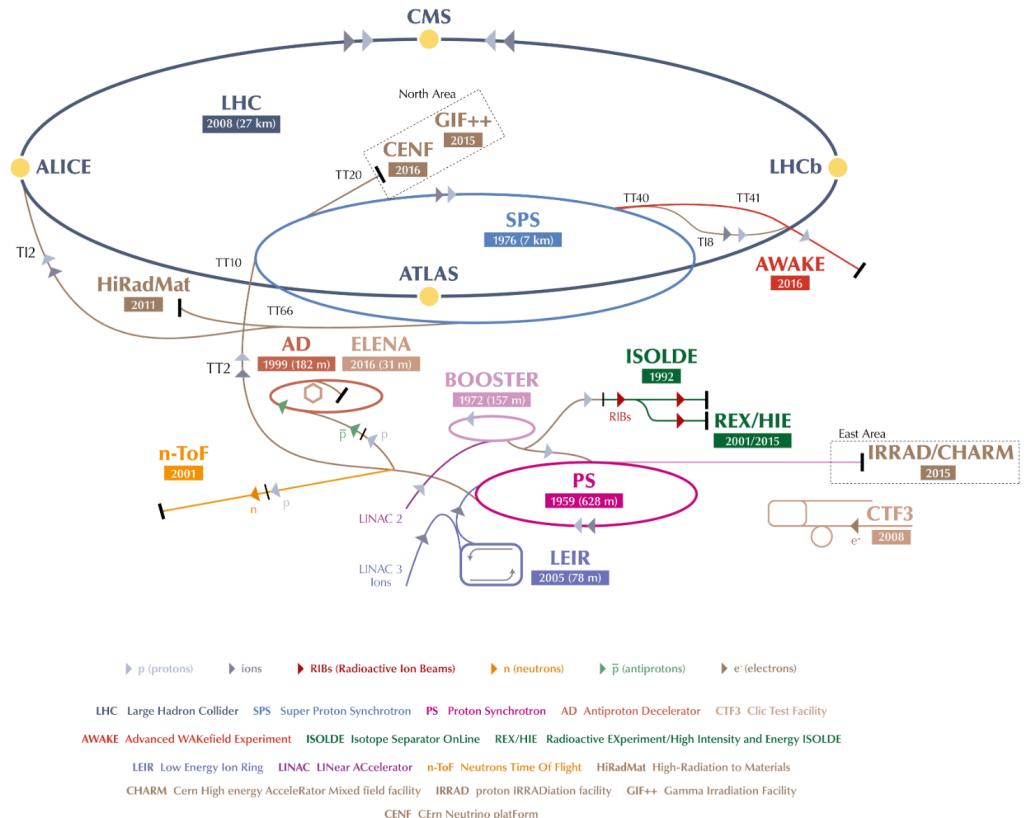


Fig. 3.2 Layout of CERN LHC complex^[23].

1020 The collisions can occur in 4 points, with corresponding 4 major detector experiments
 1021 that are briefly described as follows:

- 1022 • **ATLAS:** A Toroidal LHC ApparatuS, one general-purpose particle detector experiment
 1023 and the detector with largest volume at the LHC. It is designed to search for
 1024 the Higgs boson, test the stardand model of particle physics and search for possible
 1025 beyond SM physics.
- 1026 • **CMS:** Compact Muon Solenoid, another large general-purpose particle physics de-
 1027 tector, with the same physics goal as ATLAS and also cross check with ATLAS.
- 1028 • **ALICE:** A Large Ion Collider Experiment, it is optimized to study heavy-ion (Pb-Pb) nuclei)
 1029 collisions at a centre-of-mass energy of 2.76 TeV per nucleon pair.
- 1030 • **LHCb:** Large Hadron Collider beauty, it is a specialized b-physics experiment,
 1031 designed primarily to measure the parameters of CP violation in the interactions of
 1032 b-hadrons.

3.1.2 Luminosity and pile-up

Luminosity

In beam-beam collisions, the event rate for a process is written as^[22]:

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

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

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

in which N_b represents the number of particles per bunch, n denotes the number of bunches per beam, f_r is the revolution frequency, and γ is 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 invited, by integraling the instantaneous luminosity over time:

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

The unit of integrated luminosity we commonly use is b^{-1} that satisfying $1b^{-1} = 10^{24} cm^{-2}$. 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

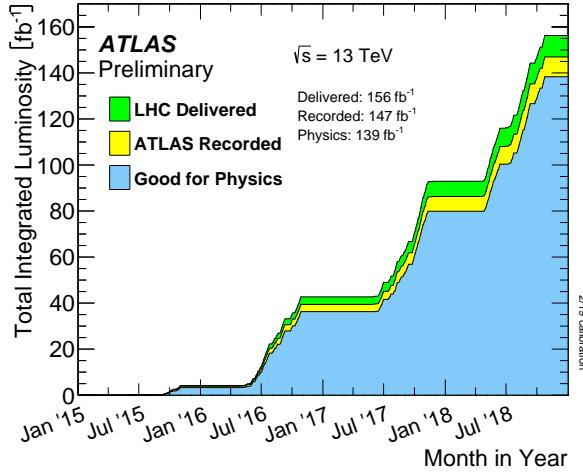


Fig. 3.3 Integrated luminosity vs delivered month from 2015 to 2018 in ATLAS experiment.

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

$$\langle \mu \rangle = \frac{\mathcal{L}_{tot} \sigma}{f_r n_{bunch}} \quad (3.5)$$

where \mathcal{L}_{tot} is the instantaneous luminosity, σ denotes the inelastic cross section, f_r represents the LHC revolution frequency and n_{bunch} is the number of colliding bunches. Usually, 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.

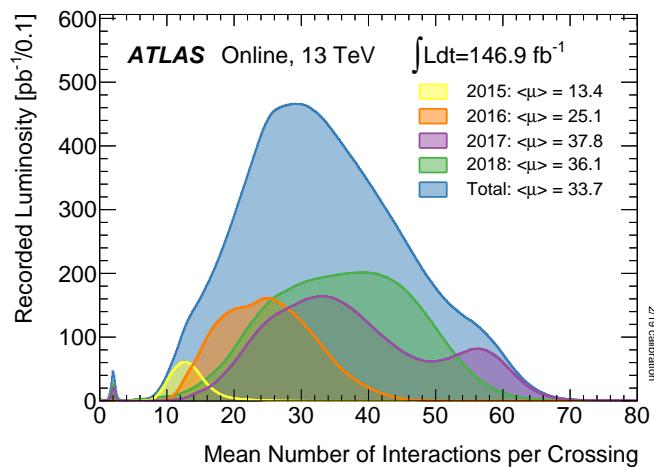


Fig. 3.4 Number of interactions per crossing weighted by luminosity from 2015 to 2018 in ATLAS experiment.

1063

1064 **3.2 ATLAS detector**

1065 **3.2.1 Detector overview**

1066 ATLAS is the world's largest volume particle detector. It is a cylinder with 46 meters
 1067 long, 25 meters in diameter, and sits in a cavern 100 meters below ground. The detector
 1068 contains about 3000 km of cables and it weights 7000 tonnes.

1069 The coordinate system and nomenclature used to describe the ATLAS detector^[24]
 1070 is depicted in figure 3.5. We define the nominal interaction point as the origin of the
 1071 coordinate system, the beam direction as the z -axis and the x - y plane is transverse to the
 1072 beam direction. The positive x -axis is given as the direction from interaction point to the
 centre of the LHC ring, while the positive y -axis points upward. There are two sides of

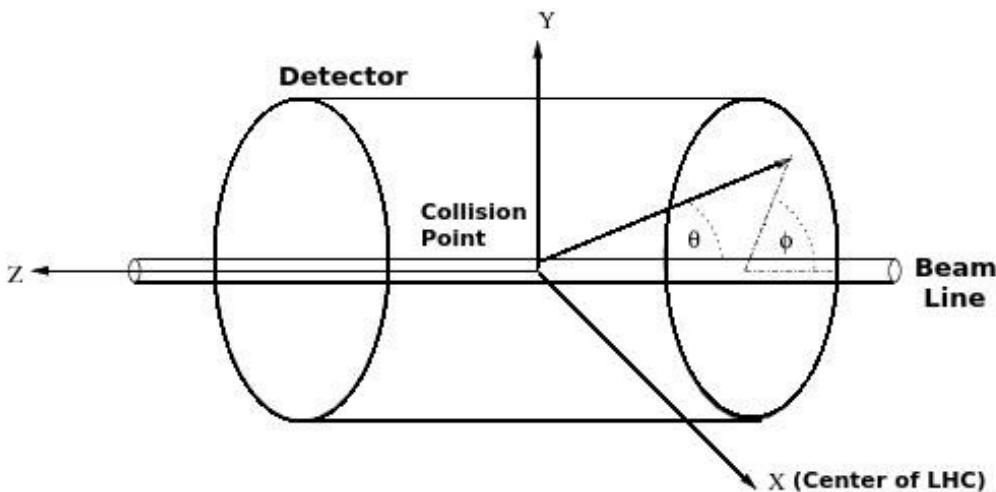


Fig. 3.5 Coordinate system used by the ATLAS experiment at the LHC^[25].

1073
 1074 detector A and C, in which A (C) -side is in the positive (negative) z direction. The polar
 1075 angle θ is measured from the beam axis, while the azimuthal angle ϕ is obtained around
 1076 the beam axis. In physics analysis, we usually use the pseudorapidity η designed as:

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] \quad (3.6)$$

1077 instead of θ angle. And for massive objects (eg. jets), the rapidity is used:

$$y = \frac{1}{2} \ln \left[\frac{E + p_z}{E - p_z} \right] \quad (3.7)$$

1078 In addition, the transverse momentum p_T , transverse energy E_T and the missing trans-
 1079 verse energy E_T^{miss} are defined in x - y plane. The ΔR , a commonly used distance measure-
 1080 ment, is defined in the pseudorapidity-azimuthal angle space as:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \quad (3.8)$$

The overall ATLAS layout is shown in figure 3.6, which is forward-backward symmetric with respect to the interaction point. The magnet configuration has

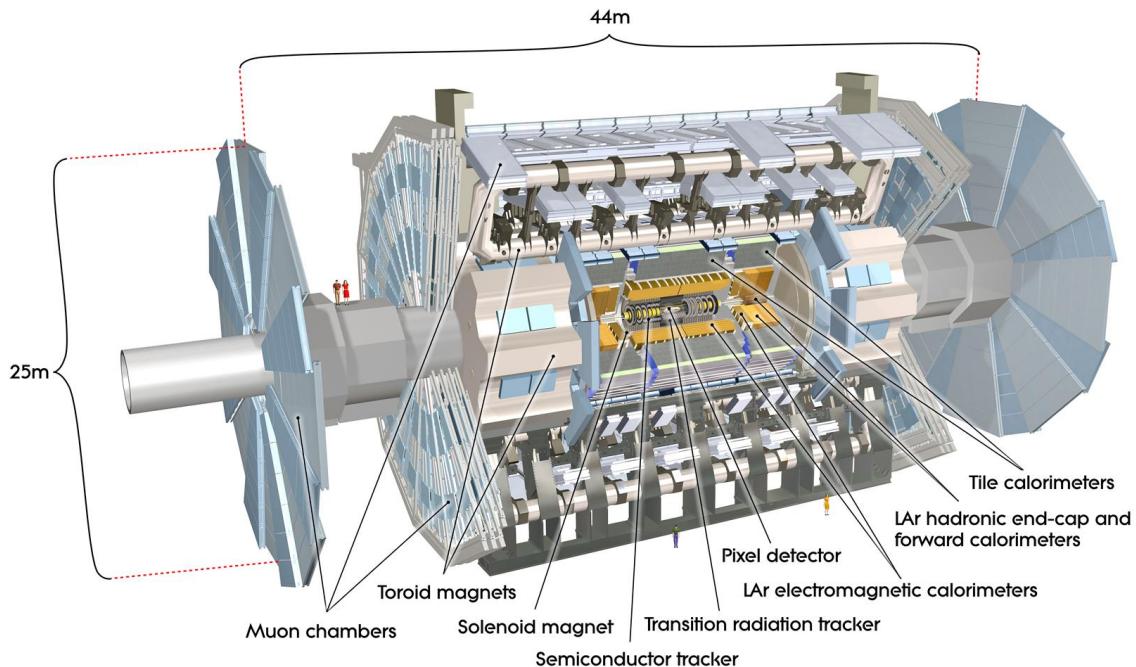


Fig. 3.6 Layout view of the ATLAS detector^[26].

a thin superconducting-solenoid surrounding the inner-detector, and three large superconducting-toroids (one barrel and two end-caps) around the calorimeters.

The inner detector, which is the innermost part of ATLAS, is surrounded by a 2 T solenoidal magnetic field. It's used for pattern recognition, momentum and vertex measurements and electron identification, with the combination of tracking system in the region of η up to 2.5.

The calorimeter is outside the solenoid, for electromagnetic and hadronic energy measurements. The high granularity liquid-argon (LAr) electromagnetic sampling calorimeters is used to measure energy and position with range up to $|\eta| < 3.2$ for electrons and photons. For hadron, a scintillator-tile calorimeter is used in the range of $|\eta| < 1.7$, and the liquid-argon hadronic endcap calorimeters (HEC) is used in end-cap region. And then the LAr forward calorimeters provide both electromagnetic and hadronic energy measurements with the coverage in forward region up to $|\eta| = 4.9$.

The muon spectrometer is the outermost layer. It's a air-core toroid system, with a long barrel and two inserted end-cap magnets that provides strong bending power in a large volume within a light and open structure. A set of chambers measuring the tracks of muons with high spatial precision and accurate time-resolution are used. Multiple-

1100 scattering effects are minor, and excellent muon momentum resolution can be achieved.

1101 **3.2.2 Physics requirement**

1102 As mentioned previously, ATLAS is one of two general-purpose particle detector ex-
1103 periment at the LHC. It's designed to take advantage of the unprecedented energy at the
1104 LHC, as the discovery of Higgs boson is one of its benchmark. Lots of precise tests and
1105 measurements of SM physics are ongoing with ATLAS experiment. while, in the mean-
1106 time, ATLAS is also designed to observe the phenomena that involve highly massive
1107 particles, which can also explore the possibility of extra dimensions proposed by several
1108 models in TeV region. To fulfil many diverse physics goals, a set of general requirements
1109 are needed:

- 1110 • The high-speed and radiation-hard electronics are required due to the experimental
1111 conditions at the LHC.
- 1112 • High detector granularity is needed to reduce the overlapping events and handle the
1113 particle fluxes.
- 1114 • Large acceptance in pseudorapidity and azimuthal angle coverage is needed.
- 1115 • For inner detector, good charged-particle momentum resolution and reconstruction
1116 efficiency are crucial. And the vertex detectors close to the interaction region are
1117 required to be able to observe secondary vertices for offline tagging of τ -lepton and
1118 b -jets.
- 1119 • Full-coverage hadronic calorimetry for accurate jet and missing transverse energy
1120 measurements, as well as good electromagnetic (EM) calorimetry for electron and
1121 photon are essentially required, since these measurements form the basis of many
1122 studies.
- 1123 • Good muon spectrometer is also required for muon identification and momentum
1124 measurement over a wide range of momenta.
- 1125 • Highly efficient but with sufficient background rejection triggers are also needed
1126 and extremely important for objects with low transverse-momentum.

1127 More detailed descriptions of each sub-system will be given in the following subsec-
1128 tions.

1129 **3.2.3 Magnet system**

1130 A strong magnetic field is required for precise measurement of charged particle mo-
1131 menta. The ATLAS detector uses two large superconducting magnet systems, a hybrid
1132 system of a central superconducting solenoid and three outer superconducting toroids, to

¹¹³³ bend charged particles^[27]. The total magnet system is 22 m in diameter and 26 m in length as shown in figure 3.7.

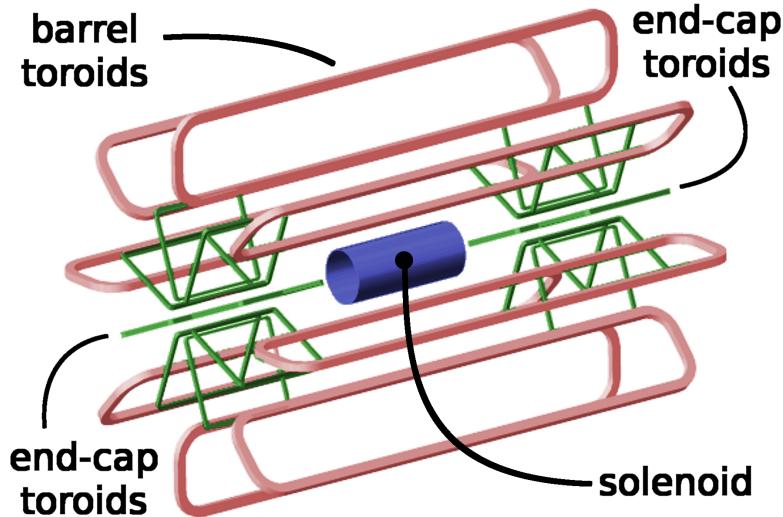


Fig. 3.7 Schematic view of the ATLAS magnet system.

¹¹³⁴

¹¹³⁵ The central solenoid produces two Tesla magnetic field surrounding the inner Detector.
¹¹³⁶ When obtaining such high field strength, at the same time, the solenoid needs to be
¹¹³⁷ thin in order to reduce the material in front of the calorimeter.

¹¹³⁸ The outer toroid system comprises one barrel superconducting toroid and two end-
¹¹³⁹ caps. The barrel one is composed of eight coils encased in individual racetrack-shaped,
¹¹⁴⁰ stainless-steel vacuum vessels and produces the magnetic field in the cylindrical volume
¹¹⁴¹ surrounding the calorimeters. Each end-cap toroid has one single cold mass built up from
¹¹⁴² eight flat, square coil units and eight keystone wedges and provides a magnetic field of
¹¹⁴³ approximately 4 T for the muon detectors in the end-cap regions.

¹¹⁴⁴ 3.2.4 Inner detector

¹¹⁴⁵ The inner detector, as shown in figure 3.8, is the detector closest to beam pipe. It's used
¹¹⁴⁶ to measure the position of charged particle tracks in high precision together with good
¹¹⁴⁷ momentum resolution within the η range of $|\eta| < 2.5$, among which the measurement
¹¹⁴⁸ of primary and secondary vertices and electron identification are especially important.
¹¹⁴⁹ Due to the extremely high luminosity produced by the LHC, the precise measurements
¹¹⁵⁰ of vertex and momentum becomes tough and fine-granularity detectors are crucial. The
¹¹⁵¹ inner detector consists of three subdetectors described as below:

¹¹⁵² Pixel detector

¹¹⁵³ The innermost part of ATLAS tracking system is the pixel detector^[29]. With finest
¹¹⁵⁴ granularity of materials, it has the best spatial resolution and 3-dimensional space-point

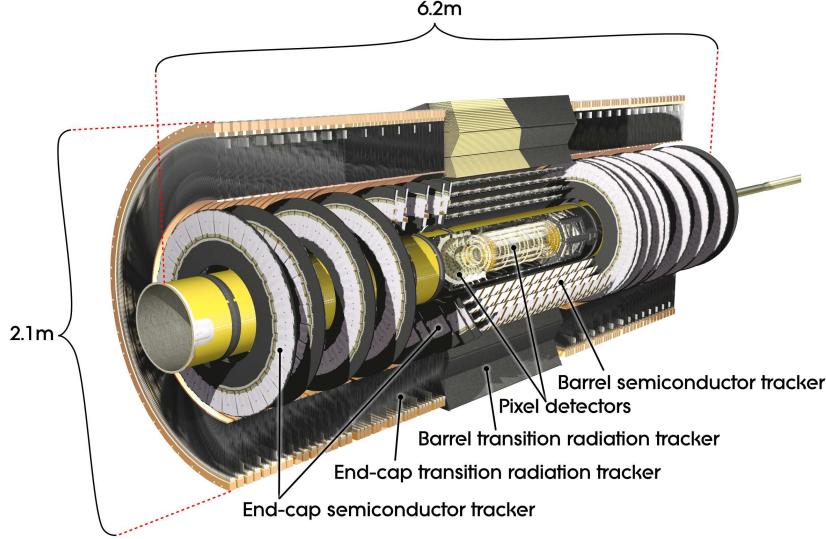


Fig. 3.8 Schematic of the inner detector in ATLAS^[28].

measurement in inner detector. ATLAS Pixel Detector for the LHC run-2 is composed of 4 layers of barrel pixel detector and two end-caps with three pixel disks each, as shown in figure 3.9. There are three outer layers that originally installed for run-1 and one additional layer called Insertable B-Layer (IBL) that newly constructed in run-2^[30]. Now the 4-layer pixel detector has very good reconstruction of primary and secondary vertices, which is even crucial for long-lived particles like τ -lepton and b-quark.

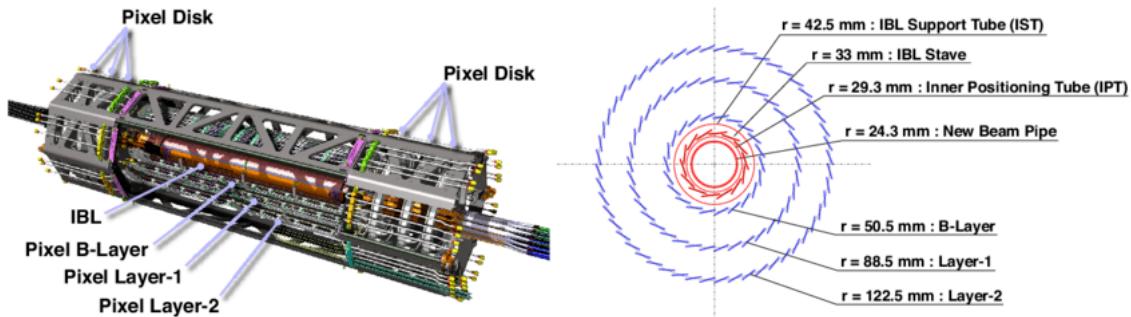


Fig. 3.9 Schematic diagram of the 4-Layer Pixel Detector in ATLAS run-2.

1160

Semiconductor Tracker

1161 The Semiconductor Tracker (SCT)^[31] installed outside the pixel detector is the mid-
 1162 dle part of the inner detector. It has similar function as pixel detector but with long and
 1163 narrow strips instead of small pixels, which makes a much larger coverage than pixel de-
 1164 tector. The SCT consists of 4088 modules, and contains four concentric layers in barrel
 1165 (2112 modules) and nine disks in each of the two end-caps (1976 modules) as shown in
 1166 figure 3.10. And it measures particles over a large area with 6.3 million readout channels
 1167 and a total area of 61 square meters. The SCT is the most critical part of the inner detector
 1168 for 2D track hit reconstruction. In barrel, the hit precision is 17 μm in the $r\text{-}\phi$ coordinate
 1169

and $580 \mu\text{m}$ in z coordinate. In end-caps, the precision is $17 \mu\text{m}$ in the $z\phi$ coordinate and $580 \mu\text{m}$ in r coordinate.

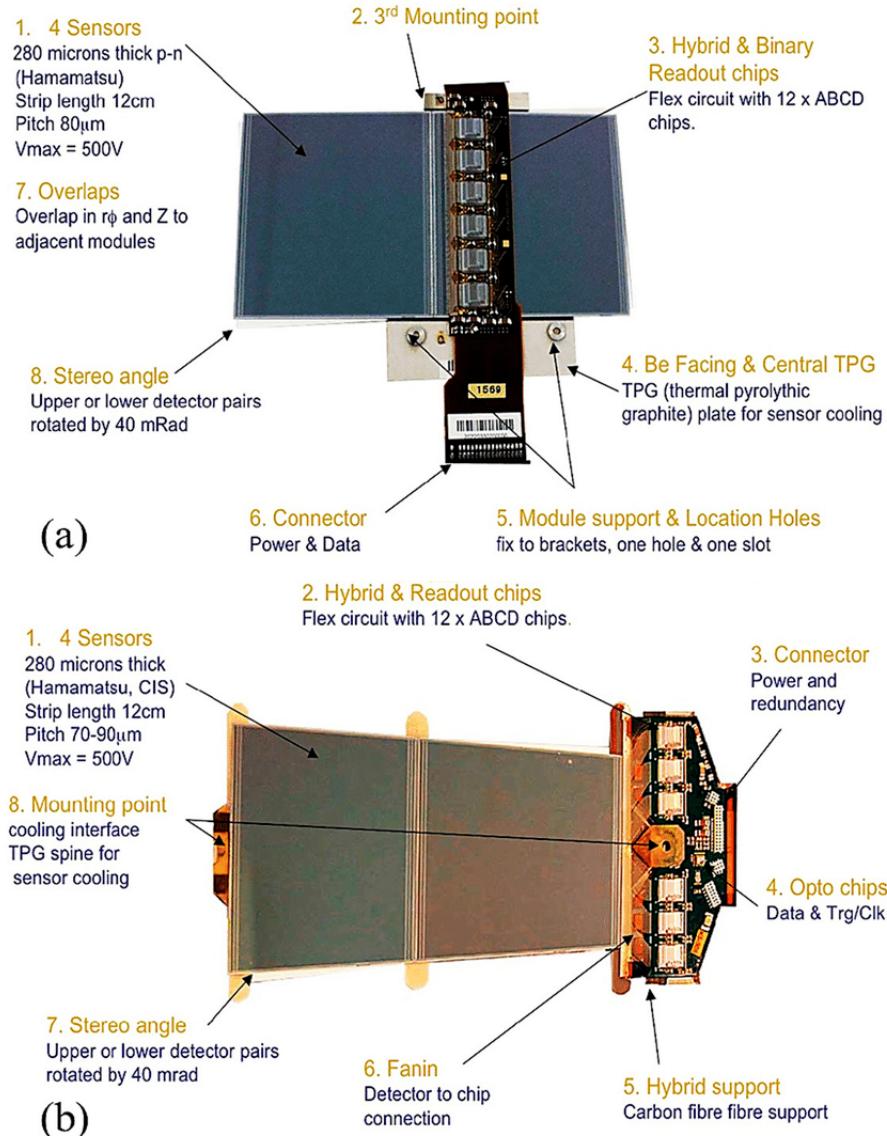


Fig. 3.10 SCT (a) barrel module and (b) end-cap^[32].

1171

1172 Transition radiation tracker

1173 The transition radiation tracker (TRT)^[33] is the outermost part of inner detector, which
 1174 has a very different design comparing to the two previously described sub-detectors. It
 1175 can be separated into three parts: one barrel and two end-cap regions with the $|\eta|$ cov-
 1176 erage up to 2.0. There are 73 barrel layers and 224 end-cap layers (112 in each) with
 1177 372000 straws in total, and about 351000 readout channels for TRT. The TRT provides
 1178 better z resolution but much worse $r\phi$ resolution (about $130 \mu\text{m}$) comparing to the pixel
 1179 detector and SCT per straw. But the straw hits still make significant contributions to mo-
 1180 mentum measurement, since its lower precision per point (compared to silicon) can be

1181 compensated by the large number of measurements and long track length.

1182 **3.2.5 Calorimeters**

1183 The calorimeters are designed to measure the energy from particles by absorbing them.
 1184 They are located outside the solenoidal magnet that surrounds the inner detector. The AT-
 1185 LAS calorimeters are comprised of a series of full ϕ -symmetrical sampling calorimeters
 1186 with the pseudorapidity range of $|\eta| < 4.9$. Figure 3.11 shows the layout of the ATLAS
 1187 calorimeter system. There are two basic calorimeter systems: an inner electromagnetic
 1188 (EM) calorimeter and an outer hadronic calorimeter. The EM calorimeter is designed for
 1189 precise measurements of electrons and photons with fine granularity; while the hadronic
 1190 one has relative coarser granularity but satisfies the physics requirements for jets recon-
 1191 structions and E_T^{miss} measurements. Two different sampling techniques are used, the EM
 1192 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:

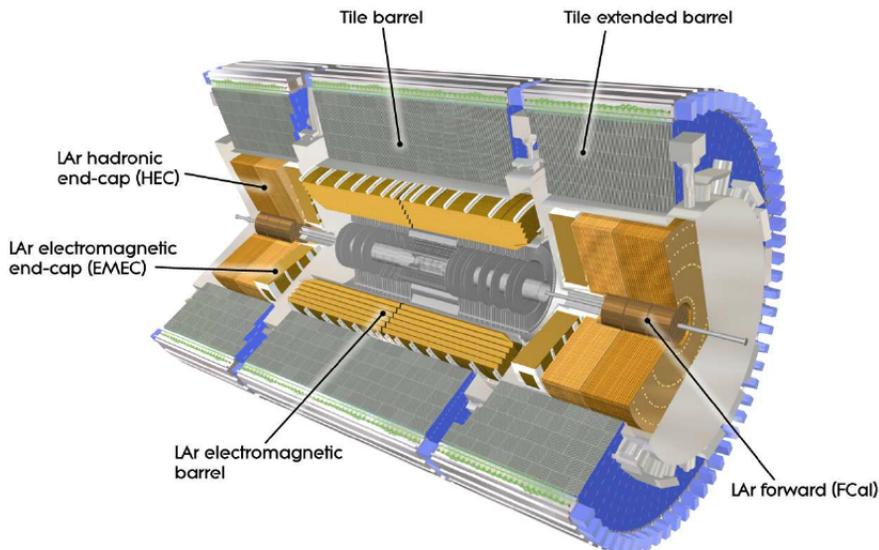


Fig. 3.11 Layout of the ATLAS calorimeters. The scintillator-based tile hadronic calorimeters are seen outside the LAr calorimeters^[34].

1193
 1194 **Liquid Argon calorimeter**

1195 The LAr calorimeter uses liquid-argon as active medium. The LAr sampling calorime-
 1196 ter technique with “accordion-shaped” electrodes is used for all electromagnetic calorime-
 1197 try covering the pseudorapidity range of $|\eta| < 3.2$; and for hadronic calorimetry with
 1198 range from $|\eta| = 1.4$ to the acceptance limit $|\eta| = 4.9$ ^[35]. Figure 3.12 depicts a segment
 1199 of the barrel calorimeter, which had “accordion-shaped” electrodes and absorber. For
 1200 barrel EM calorimeter, the absorbing material is lead-liquid argon, while the hadronic

1201 end-cap calorimeter uses copper plates. In addition, the forward calorimeter is split into
 1202 three parts, an EM sector in which copper is used as absorbing material and two hadronic
 sectors using tungsten outside the EM sector.

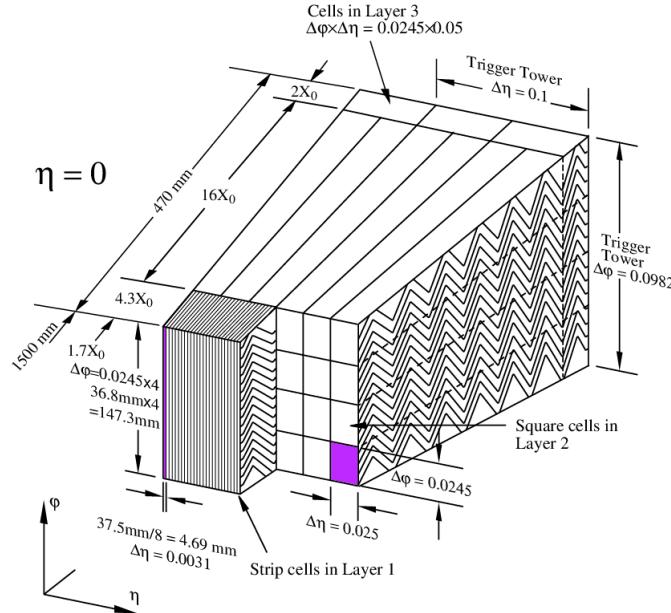


Fig. 3.12 Layout of a LAr EM calorimeter barrel module^[35].

1203

1204 Tile calorimeter

1205 Tile calorimeter is a sampling calorimeter using scintillating plates as active medium
 1206 and steel as absorber. It consists of three sections: the central barrel with the pseudorapidity
 1207 range of $|\eta| < 1.0$ and two extended barrels with $0.8 < |\eta| < 1.7$. Figure 3.13 shows
 1208 the design of one tile calorimeter module. It's used for energy reconstruction of jets and
 1209 E_T^{miss} measurement by combining the measurements with the end-cap and forward LAr
 1210 hadronic calorimeter.

1211 3.2.6 Muon spectrometer

1212 Muon spectrometer^[37] is the outermost part of the ATLAS detector with an extremely
 1213 large tracking system. It measures a large range of muon momentum, and the accuracy is
 1214 about 3% at 100 GeV and 10% at 1 TeV. The muon spectrometer comprises three main
 1215 parts: a magnetic field produced by three toroidal magnets; a set of chambers measur-
 1216 ing the tracks of muons with high spatial precision; and triggering chambers with accu-
 1217 rate time-resolution. Figure 3.14 shows the schematic of ATLAS muon spectrometer that
 1218 consists of four types of muon chambers (*MDT*, *CSC*, *RPC*, *TGC*) as well as the magnet
 1219 systems (barrel and end-cap toroid).

1220 More details of four chambers are given as below:

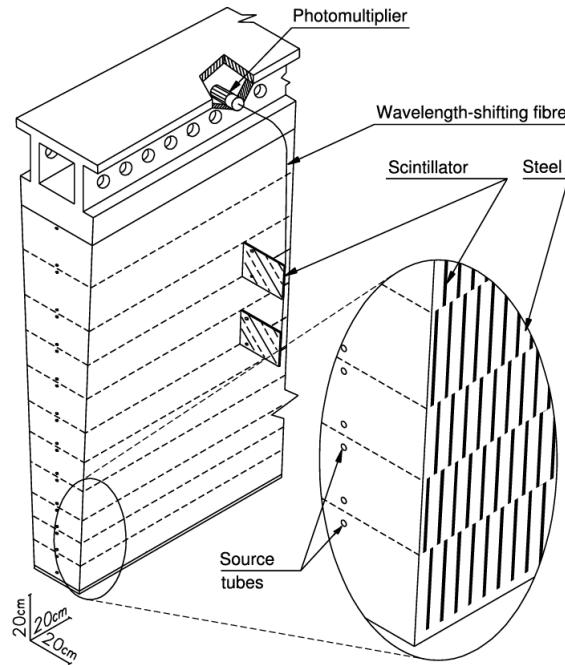


Fig. 3.13 Schematic of tile calorimeter module^[36].

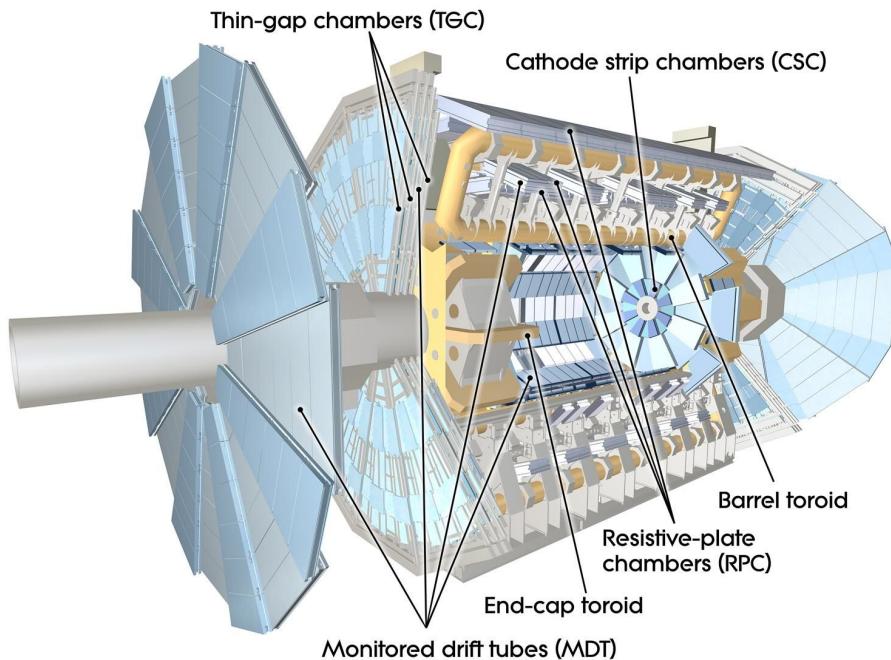


Fig. 3.14 Cut-away view of the muon spectrometer in ATLAS^[38].

- **Monitored Drift Tubes (MDT).** MDTs offer precise measurement of momentum with the $|\eta|$ range up to 2.7. The chambers include three to eight 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, because of their good time resolution and high

rate capability. They are multi-wire proportional chambers (MWPC), in which the cathode planes are segmented into strips in orthogonal directions, allowing both coordinates to be measured based on the induced-charge distribution. The resolution in the bending plane is about $40 \mu\text{m}$ and 5 mm in the transverse plane.

- **Resistive plate chambers (RPC).** The RPCs serve as fast triggers in the barrel region of $|\eta| < 1.05$ due to its high rate capability as well as its good time and spatial resolution. They are gaseous parallel electrode-plate detector without any wires. There are three concentric cylindrical layers around the beam axis working as three trigger stations, while each of them is composed of two independent layers to measure the transverse coordinates of η and ϕ .
- **Thin gap chambers (TGC).** TGCs are used as trigger system for the end-cap region of $1.05 < |\eta| < 2.4$, and works based on the same principle as multi-wire proportional chambers. In addition to the measurement of MDT in bending direction, they also offer the second azimuthal coordinate as supplement.

3.2.7 Trigger system

Trigger system in ATLAS is a very essential component, which is responsible for deciding whether to keep a given collision event for later study or not. In the LHC run-2, higher energy, luminosity and pile-up lead to an large increase of event rate by up to a factor of five comparing to run-1, which causes to a even larger challenge and more strict requirement of trigger system.

The trigger system in run-2 consists of a hardware-based first level trigger (Level-1) and a software-based high level trigger (HLT)^[39]. As depicted in figure 3.15, in Level-1, the inputs from coarse granularity calorimeter and muon detector information together with some other subsystems are sent to the Central Trigger Processor to determine Regions-of-Interest (RoIs) in the detector. The event rate can be reduced by Level-1 triggers from 40 MHz to 100 kHz. After that, the RoI information from Level-1 is sent to HLT, where more sophisticated selection algorithms are run for regional reconstruction. The HLT reduces the rate from Level-1 from 100 kHz to about 1 kHz on average. At the end, the events that accepted by HLT are transferred to local storage at experimental site for offline reconstruction. Details about Level-1 and HLT trigger systems are described as below:

Level-1 trigger

Substantial upgrades have been delivered in ATLAS Level-1 trigger system for run-2 data taking. The upgrades took place in both hardware and detector readout, allowing

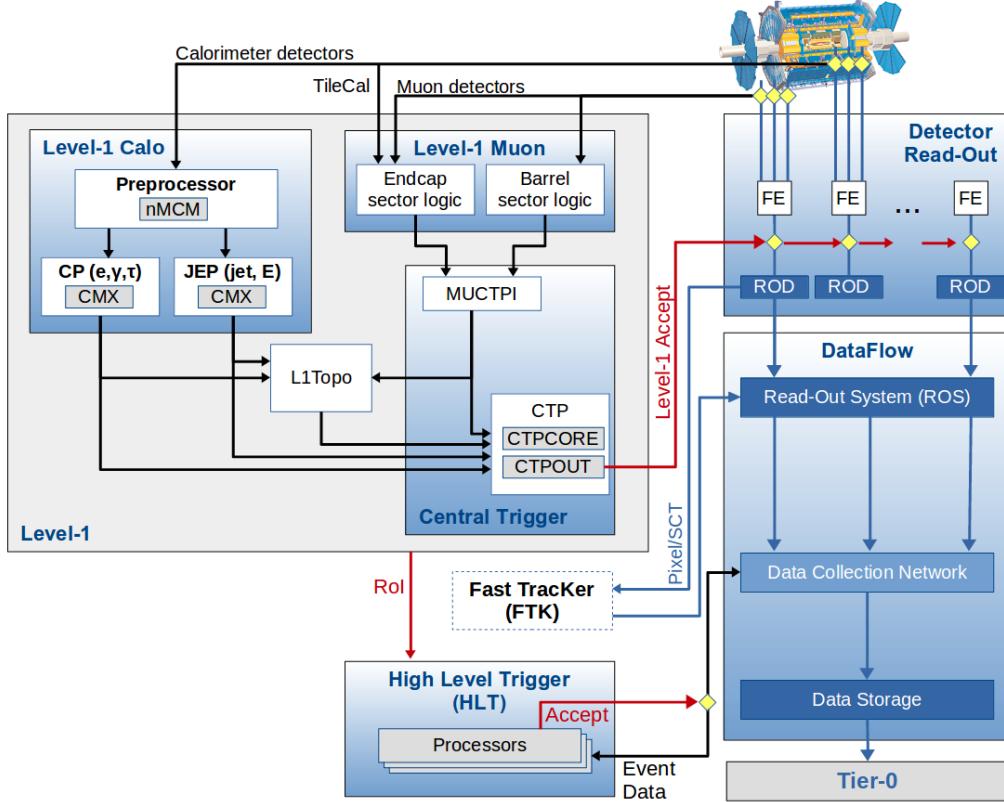


Fig. 3.15 Schematic diagram of the ATLAS trigger and data acquisition system in run-2.

the trigger rate increasing from 70 kHz (in run-1) to 100 kHz (in run-2). There are two major parts of Level-1 triggers, including Level-1 calorimeter (L1calo) trigger and Level-1 muon (L1mu) trigger.

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

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

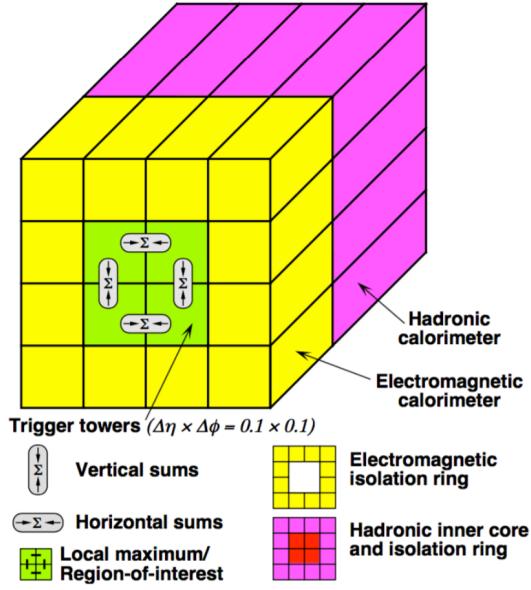


Fig. 3.16 An examples of L1 calorimeter trigger tower for electron and photon triggers^[40].

1279 High Level Trigger

1280 In run-1, the Event Filter computer clusters and Level-2 trigger system were separated,
 1281 while now in run-2, they have been merged into a single HLT event processing. The new
 1282 arrangement helps to reduce the complexity and duplication of algorithm, which leads to a
 1283 more flexible high level trigger system. During the long-shutdown between the LHC run-
 1284 1 and run-2, lots of reoptimizations have been done for trigger reconstruction algorithms
 1285 as well as the offline analysis selections, which can improve the efficiency by more than
 1286 a factor of two in some cases like hadronic tau triggers. For some triggers, the HLT
 1287 processing performed within RoIs also allows to aggregate from RoIs to single objects.
 1288 This improvement reduces the CPU processing for events with overlapping RoIs, and the
 1289 average output rate has been increased from 400 Hz to 1 kHz.

1290 The HLT reconstruction algorithm can be divided into fast and precision online re-
 1291 construction steps. As illuminated by figure 3.17, the initial fast reconstruction helps to
 1292 reduce the event rate, and to seed into precision reconstruction. Then the final online
 1293 precision reconstruction is improved and uses offline-like algorithms as much as possi-
 1294 ble. In particular, multivariate analysis techniques (based on machine learning) have been
 1295 introduced online in many aspects.

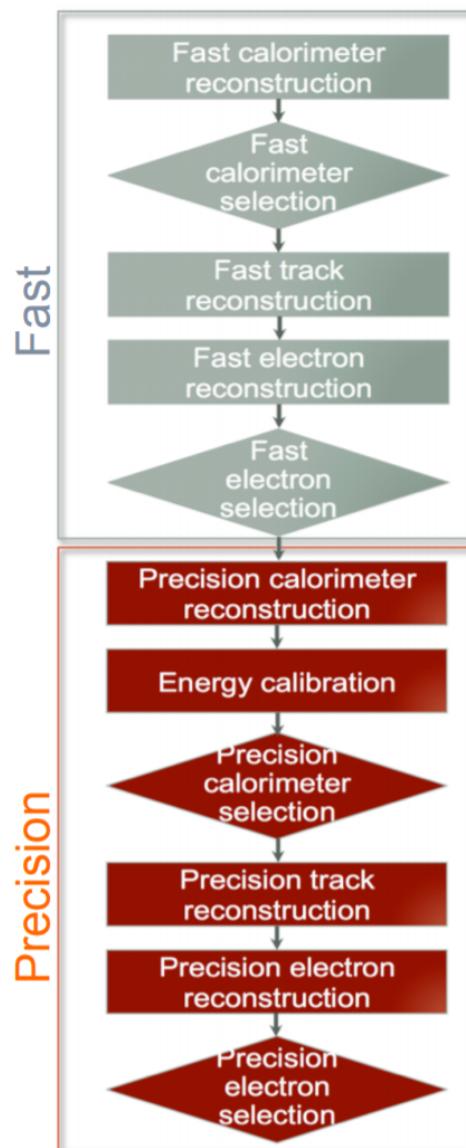


Fig. 3.17 The HLT trigger algorithm sequence^[40].

1296 Chapter 4 Simulation and Event Reconstruction for 1297 the ATLAS Experiment

1298 In current LHC pp collision, bunches of protons collide every 25 nanoseconds (ns),
1299 which gives a large challenge to event reconstruction and selections. To predict and model
1300 each process, Monte Carlo simulations of physics events are essential for high-energy
1301 physics experiments. This section will briefly discuss the event simulation and recon-
1302 struction programs based on the ATLAS software framework.

1303 4.1 Event simulation

1304 The simulation program of ATLAS experiment is integrated into the ATLAS software
1305 framework called *Athena*^[41], which uses Python as an interpreter language and object-
1306 oriented script to load C++ objects and algorithms. The ATLAS simulation data flow is
1307 shown in figure 4.1, where the square-cornered boxes represents algorithms and applica-
tions to be run and round-cornered boxes denote data objects.

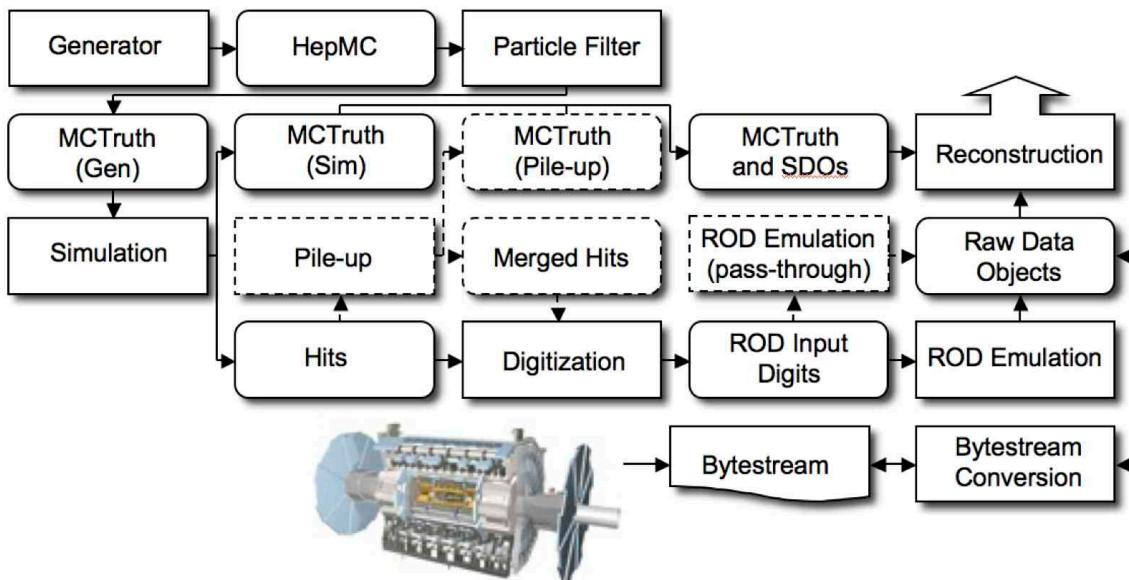


Fig. 4.1 Schematic diagram of the ATLAS simulation software^[42].

1308
1309 First of all, events are produced by MC generators in standard HepMC format and then
1310 read into the detector simulation. During the simulation, particles are propagated through
1311 the full ATLAS detector whose configurations can be set by users via GEANT4 toolkit.
1312 The *hits* informations, obtained from the energies deposited in the sensitive regions of the
1313 detector, containing the total value of energy deposition, time and position, are written

1314 into hit files. In the meantime, the events in “truth” format are also recorded to contain the
1315 history of the interactions from the generator, including incoming and outgoing particles.
1316 Simulated Data Objects (SDOs) are created from truth, which are maps between hits in
1317 sensitive portions of the detector and truth information of particles in simulation. The hit
1318 files are then sent to digitization, which constructs “digits” written into Raw Data Object
1319 (RDO) file used for reconstruction.

1320 In conclusion, there are three main parts of framework: *Generation*, *Simulation* and
1321 *Digitization*. More details are given as below:

1322 **Event generation**

1323 As shown in figure 4.2^[43], at hadron colliders, multiple scattering and rescattering
1324 effects arise, which needs to be simulated by Monte Carlo (MC) event generators to reflect
1325 the full complexity of those event structures. Several MC event generators can be used to
1326 generate events in HepMC format. The events can be filtered at generation time with some
1327 certain requirements (eg. decay channel or missing energy above a certain threshold). The
1328 generator is in charge of any prompt decays, like W and Z bosons decays, and all “stable”
1329 particles expected to propagate through the detector are stored. During the generation
1330 steps, the detector effects are ignored and only immediate decays are considered.

1331 There are several MC generators that have been widely used with general purpose,
1332 including Sherpa^[44], Herwig++^[45], PowhegBox^[46], MC@NLO^[47] and Pythia8^[48].

1333 **Simulation**

1334 GEANT4 is used as standard simulation toolkit for the ATLAS experiment, which
1335 transports physics particles through the detector’s geometry. During the generation level,
1336 the entire connected chain of the HepMC event is stored as the Monte Carlo truth. Only the
1337 stable particles are read into GEANT4 for further simulation and selection, and transfor-
1338 mations can be applied to these events to select certain processes. During the simulation,
1339 many secondary tracks can be produced, therefore only information from the interactions
1340 of interest are stored, including the incoming particles, step sequence, vertex as well as
1341 outgoing particles. The output of GEANT4 is called *hit file*, containing metadata includ-
1342 ing the simulation configuration, all requested truth information and the hit informations
1343 of each subdetector.

1344 Since the standard ATLAS detector simulation cost very large computing resources
1345 to accurately model the complex detector geometry and physics descriptions, some fast
1346 simulation programss are developed according to different user purpose. Some popular
1347 fast-sim toolkits include *Fast G4 Simulation*^[49], *ATLFAST-I*^[50] and *ATLFAST-II*^[51].

1348 **Digitization**

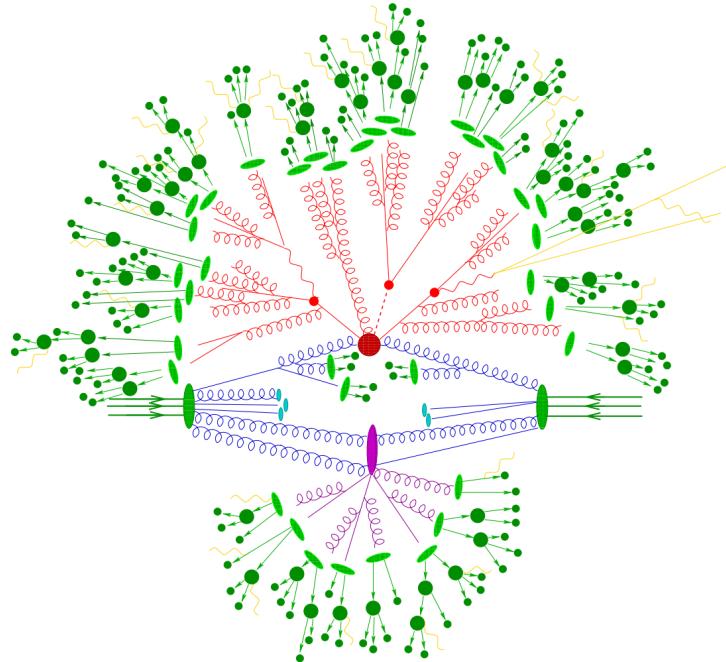


Fig. 4.2 Schematic of a hadron-hadron collision event simulated by MC generator. The red blob in center denotes the hard collision, surrounded by tree-like structures denoting Bremsstrahlung from Parton Showers. The purple blob is a secondary hard scattering event. The light green blobs represent the parton-to-hadron transitions while the dark green blobs stand for hadron decays. The yellow lines indicate soft photon radiations.

1349 The hit informations from detector simulation by GEANT4 are then sent into the digitization procedure to be converted into detector response called “digits”. Before producing
 1350 the detector signal formart, events can be overlaid at a user-specified rate, called “pile-up”. The simulation of pile-up can be done during digitization to save the CPU time. At
 1351 this stage, the detector noise and the first level trigger that implemented with hardware
 1352 on the real detector are added into events. Firstly, the “digits” inputs are constructed and
 1353 passed to the simulated readout drivers (RODs) as in the detector electronics. The output
 1354 of this step are written out as Raw Data Object (RDO) file. In addition, the digitization
 1355 algorithms can also produce Simulated Data Objects (SDOs), which contain information
 1356 about all the particles, noise and the amount of energy that contributed to the signal. Then
 1357 all information are sent into reconstruction level described in next subsection.
 1358
 1359

1360 4.2 Reconstruction

1361 The data flow of ATLAS data processing is sketched in figure 4.3. Data from detector
 1362 is firstly filtered by online trigger system before sending to the *Tier-0* (*T0*) for initial
 1363 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 around the world 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.

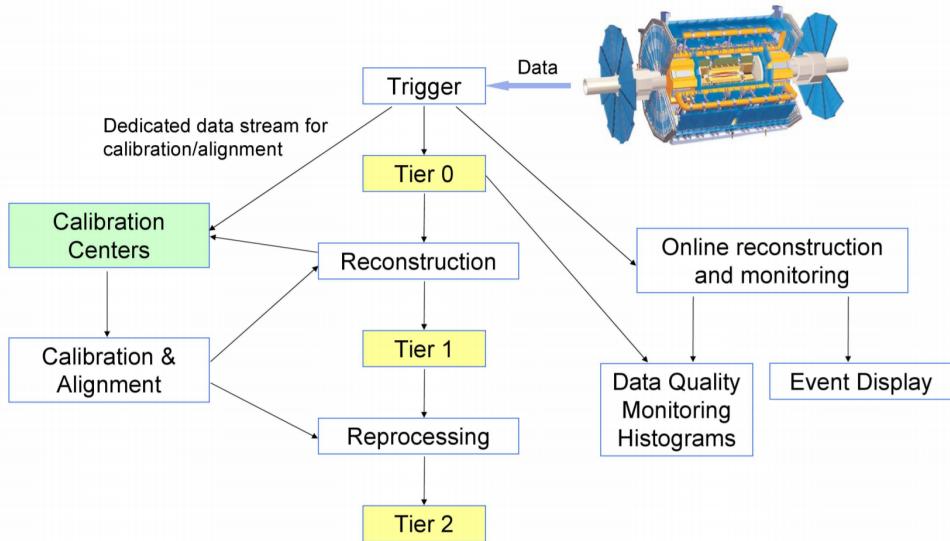


Fig. 4.3 The flowchart of the ATLAS data processing^[52].

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, namely the 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 inhomogeneous magnetic field and the large background^[53]. More details of these two types of track are given as below:

Inner detector track

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

in one.

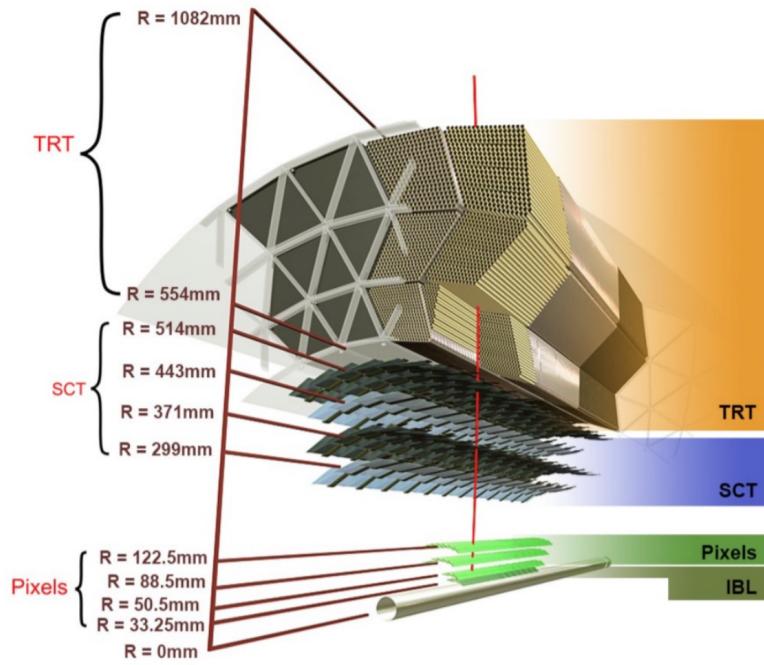


Fig. 4.4 Cut-away view of the ATLAS inner detector.

1385

1386 For inside-out tracking, it exploits the high granularity of the pixel and SCT detectors
 1387 to discover prompt tracks originating from the interaction point. In first step, the track
 1388 seeds are formed by combining the information of space-points in the three pixel layers
 1389 and the first SCT layer. Then, these seeds are extended throughout the SCT to build track
 1390 candidates. After that, these candidates are fitted with some quality cuts applied to remove
 1391 the outlier clusters, reject the fake tracks and resolve ambiguities in the cluster-to-track
 1392 association. The selected tracks are then further extended to TRT, and refitted with the
 1393 full information from pixel, SCT and TRT detectors.

1394 Another complementary approach, outside-in, searches for unused track segments
 1395 start from TRT instead. These segments are then extended into the SCT and pixel de-
 1396 tectors to improve the tracking efficiency for secondary tracks from decays of long-lived
 1397 particles or conversions.

1398 **Muon spectrometer track**

1399 First of all, the MS track reconstruction^[54] searches for hit patterns inside muon cham-
 1400 bers and forms the corresponding segments. The Hough transform method^[55] is used to
 1401 search the hits lie on a certain trajectory in the bending plane in each MDT chamber and
 1402 its nearby trigger chamber. Then one can reconstruct the MDT segments by performing
 1403 a linear fit to the hits found in each layer. In the meantime, the hits informations from
 1404 RPC or TGC can be used to measure the coordinate orthogonal to the bending plane. And

1405 the segments of CSC can be reconstructed using a separate combinatorial search in the η
 1406 and ϕ detector planes.

1407 Then by fitting the hits informations from segments in different subsystems together,
 1408 one can built the muon track candidates. The reconstruction makes use of the algorithm by
 1409 performing a segment-seeded combinatorial search, which starts by using the segments
 1410 reconstructed in the middle layers of the detector where more trigger hits are available
 1411 as seeds. The search is then extended to use the segments as seeds from the inner and
 1412 outer layers. The segments are firstly selected based on criteria such as hit multiplicity
 1413 and fit quality, and then matched using their relative positions and angles. To build a
 1414 track, at least two matching segments are required, except in the barrel-endcap transition
 1415 region where a single high-quality segment can be used. At beginning, one segment can
 1416 be used to build more than one track candidates. But then, an overlap removal algorithm
 1417 is adopted to select the best assignment to one single track, or decide whether allows the
 1418 certain segment to be shared between two tracks.

1419 The hits associated with track candidates are then fitted using a global χ^2 fit. The
 1420 algorithm accepts the track candidate if its fitting χ^2 passes the required value. Hits with
 1421 large contribution to χ^2 are removed and the track fit is repeated. In addition, the algorithm
 1422 performs a hit recovery procedure looking for additional hits consistent with the candidate
 1423 trajectory, and the track candidate is refit if additional hits are found.

1424 4.2.2 Primary vertex

1425 The reconstruction of primary vertex (PV) uses the reconstructed tracks introduced in
 1426 previous section as inputs. The tracks must satisfy the following criteria^[56]:

- 1427 • $p_T > 400$ MeV
- 1428 • $|\eta| < 2.5$
- 1429 • Number of silicon hits $\geq \begin{cases} 9 & \text{if } |\eta| \leq 1.65 \\ 11 & \text{if } |\eta| > 1.65 \end{cases}$
- 1430 • IBL hits + B-layer hits ≥ 1
- 1431 • A maximum of 1 shared module (1 shared pixel hit or 2 shared SCT hits)
- 1432 • Pixel holes = 0
- 1433 • SCT holes ≤ 1

1434 A candidate vertex is formed by requiring two tracks passing these selection criteria.

1435 The reconstruction of PV can be described into two steps^[57]: vertex finding and vertex
 1436 fitting. The first step is associating the reconstructed tracks to vertex candidates, namely
 1437 the pattern recognition process. The latter one works on the reconstruction of vertex po-

1438 sition and its covariance matrix. More details are described as below:

1439 First of all, a set of tracks passing the selection criteria mentioned above is selected.
1440 Then a seed position, determining by beam spot in the transverse plane, for the first vertex
1441 is chosen. The starting point for x- and y- coordinates are directly chosen as the centre
1442 of the beam spot, while the one for z-coordinate is computed as the mode of tracks' z-
1443 coordinates at their respective points with closest approach to the reconstructed centre of
1444 the beam spot.

1445 After determining the seed position, the iterative primary vertex finding procedure
1446 starts. An vertex fitting algorithm is adopted to find the optimal vertex position by per-
1447 forming an iterative χ^2 minimization, in which the seed position is used as the start point
1448 and the reconstructed tracks are used as input measurements. For this fitting procedure, the
1449 weights reflecting the input tracks and the vertex estimation's compatibility are assigned,
1450 and the vertex positions are re-calculated based on these weighted tracks. Then the iter-
1451 ative procedure is repeated by recalculating the track weight according to the new vertex
1452 position. After the iterations, the final weights tracks used in vertex fit are given. And
1453 those incompatible tracks ($> 7 \sigma$) are then rejected from this vertex candidate and moved
1454 back to the unused pool for next vertex finding. Then iteration procedure describes above
1455 are repeated again by using the remaining tracks in pool, until no un-associated tracks are
1456 left or no additional vertex can be found in remaining tracks.

1457 At the end, the vertices with at least two associated tracks passing through are treated
1458 as possible PV candidates. And the output of this vertex reconstruction algorithm is the in-
1459 formation of three dimensional vertex positions and their covariance matrices. In physics
1460 analysis, it's most often to choose the one with highest sum of transverse momentum
1461 ($\sum p_T^2$) as PV.

1462 4.2.3 Electron

1463 Many interesting physical processes are with the involvement of one or more elec-
1464 trons (or positrons) at the LHC. But these electrons can be subjected to large amount of
1465 backgrounds such as hadrons, non-prompt electrons from photon conversions and non-
1466 isolated electrons from heavy flavor hadon decays. It is therefore essential to efficiently
1467 reconstruct and identify electrons as well as, in the meantime, to keep high background
1468 rejection.

1469 In ATLAS, in central region, the electrons leave tracks in inner detector (ID) and
1470 deposit the energies in the electromagnetic (EM) calorimeter. Firstly the signals from
1471 calorimeter are used for L1 trigger system, and then combined with the information from

1472 ID tracks to reconstruct electron candidates that will be used for the high level trigger
1473 (HLT) decision algorithms^[58]. The backgrounds mentioned above can then be further
1474 suppressed by using several identification criteria. In addition, electrons are required to
1475 be isolated from other activities to be further distinguished from background.

1476 More details of electron *reconstruction*, *identification* and *isolation* are described as
1477 below.

1478 **Electron reconstruction**

1479 Several steps are proceeded for electron reconstruction in the region of $|\eta| < 2.47$ in
1480 ATLAS detector:

1481 1. **Seed-cluster reconstruction:** A sliding window of 3×5 in unit of $\Delta\eta^{tower} \times$
1482 $\Delta\phi^{tower} = 0.025 \times 0.025$ in $\eta \times \phi$ space is utilized to search for electron cluster
1483 seeds with total cluster transverse energy greater than 2.5 GeV. Then a clustering
1484 algorithm^[59] is applied to form the clusters around the seeds, which can take advan-
1485 tage of removing the duplications. The kinematics of clusters are then reconstructed
1486 by using an extended window depending on the cluster position. The efficiency of
1487 cluster search is from about 95% at $E_T = 7\text{GeV}$ to 99% for $E_T \geq 15\text{GeV}$.

1488 2. **Track reconstruction:** The track reconstruction can be divided into two steps: pat-
1489 tern recognition and track fit. The standard pattern recognition in ATLAS uses pion
1490 hypothesis for energy loss caused by interactions with detector material. If a track
1491 seed with $p_T > 1$ GeV cannot be successfully extended to a full track required at
1492 least seven hits using this pion hypothesis, but still falls inside one of the EM clus-
1493 ter region of interest, as a second attempt, the pattern recognition using electron
1494 hypothesis is then used to allow larger energy loss. Depending on the pattern used
1495 in previous stage, the track candidates are then fitted with either the pion hypothesis
1496 or the electron hypothesis by using ATLAS Global χ^2 Track Fitter^[60]. If a track
1497 candidate fails the fit by using pion hypothesis, it can be refit with the electron hy-
1498 pothesis again. In this method, a specific electron-oriented algorithm is integrated
1499 into the ATLAS standard track reconstruction, which improves the performance for
1500 electron and as well as maintain minimal interference with the main track recon-
1501 struction.

1502 3. **Electron specific track fit:** Once the tracks are obtained, they are loosely matched
1503 to EM cluster using the distance in η and ϕ between the position of track (after ex-
1504 trapulation) in calorimeter's middle layer and the cluster barycentre. The matching
1505 conditions take into account the energy loss of bremsstrahlung and the number of
1506 precise hits in silicon detector.

1507 **4. Electron candidate reconstruction:** The electron candidate is reconstructed by
 1508 matching the track candidate to EM cluster seed to eventually completes the electron
 1509 reconstruction procedure. If more than one track satisfy the matching condition,
 1510 one track is chosen as primary track based on the information of the cluster-track
 1511 distance R , the number of pixel hits and the presence of a hit in the first silicon
 1512 layer^[61]. In addition, the electron candidates are removed from electron pool if it's
 1513 without any associated precise hit tracks, and moved into photon candidates pool.
 1514 Then we reformed the electron clusters by using 3×7 (5×5) longitudinal towers
 1515 of cells in barrel (end-caps) in EM calorimeter. The measured energy is calibrated
 1516 to original electron energy based on MC simulated samples by using multivariate
 1517 techniques (MVA).

1518 In addition, in physics analysis, to reduce the background from photon conversions
 1519 and secondary particles, the track associated with electron is required to be compatible
 1520 with the primary vertex of the hard collision. Practically, the impact parameters cuts such
 1521 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 of
 1522 the track to the measured beam-line, z_0 is the distance along the beam-line between the
 1523 point where d_0 is measured and the beam-spot position, and θ is polar angle of the track,
 1524 σ_{d_0} denotes the estimated uncertainty of d_0 parameter. Figure 4.5 depicts the definition of
 each track impact parameter.

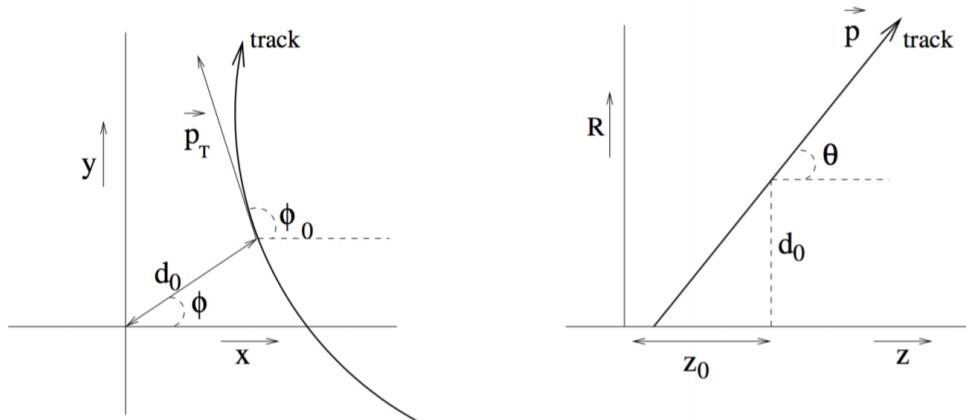


Fig. 4.5 Schematic of the impact parameters of a track in the transverse plane (left) and RZ-plane (right), as defined in the global ATLAS tracking frame^[62].

1525

1526 **Electron identification**

1527 The electron identifications are applied to determine whether the reconstructed elec-
 1528 tron candidate is more signal-like or background-like object. The identification algorithms
 1529 make use of quantities of related variables from electron cluster and track measurements
 1530 including calorimeter shower shapes, track properties, as well as variables measuring

1531 bremsstrahlung effects for distinguishing signal from background. Taking the advantage
 1532 of new IBL in run-2, the number of hits in this innermost pixel layer is utilized for discrim-
 1533 inating between electrons and converted photons. In addition, a likelihood method based
 1534 on the TRT high-threshold hits is adopted to compensate the lower transition radiation
 1535 absorption probability of the argon.

1536 The baseline identification algorithm introduced in ATLAS run-2 is the likelihood-
 1537 based (LH) method, making use of a MVA technique to simultaneously evaluate several
 1538 properties of electron candidates when making a decision. The LH method utilizes the
 1539 probability density functions (PDFs) of signal and background as the input discriminating
 1540 variables. Based on these PDFs, it can calculate overall probabilities of the object to be
 1541 signal or background. Then the probabilities of signal and background are combined
 1542 together into a discriminant $d_{\mathcal{L}}$:

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

1543 where \mathbf{x} denotes the vector of discriminating variables and $P_{s(b),i}(x_i)$ represents the value
 1544 of signal (background) PDF of the i^{th} variable as x_i .

1545 Three levels of working points (WPs) for electron identification are provided: *Loose*,
 1546 *Medium* and *Tight*, in order of increasing background rejection. Samples selected by a
 1547 looser WP are subsets of a tighter one, for example, the electrons passing Medium can
 1548 all be selected by Loose. The identification efficiency varies as function of transverse
 1549 energy (E_T) as shown in figure 4.6. For evaluations, the electron candidates from MC
 1550 simulation of $Z \rightarrow ee$ decays (di-jet) are used as signal (background). Depending on the
 1551 working point, the signal (background) efficiencies for reconstructed electron candidates
 at $E_T = 25 GeV$ are from 78 to 90% (0.3 to 0.8%), and increase (decrease) with E_T .

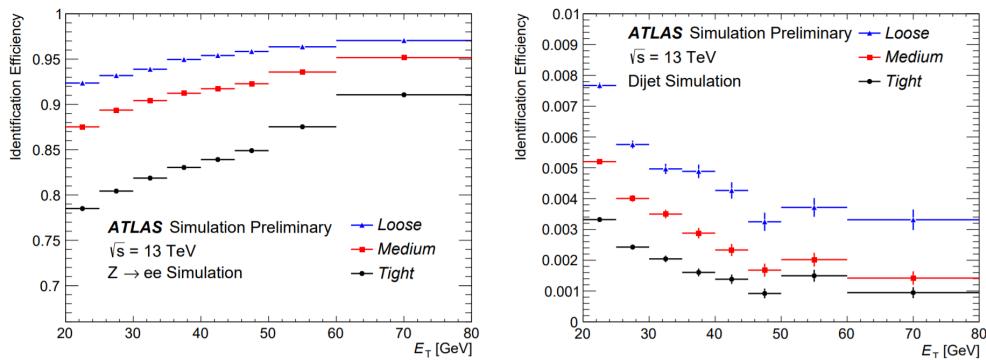


Fig. 4.6 The efficiencies of three electron identification WPs from $Z \rightarrow ee$ (left) events and hadrons misidentified as electrons estimated using di-jet MC samples (right).

1553 Electron isolation

1554 In addition to the identification criteria, most analyses have electron isolation requirement to further distinguish signal from background. To quantify the energy of particles 1555 around the electron candidate, the isolation variables can help to separate the prompt electron 1556 from other non-isolated electrons, like the electrons from converted photons or from 1557 heavy flavour hadron decays. There are two kinds of discriminating variables that have 1558 been designed:

- 1560 • **Calorimeter-based variable:** $E_T^{\text{topocone}20}$. It's computed from the sum of transverse 1561 energies of topological clusters^[63], and calibrated at EM scale in a cone of $\Delta R = 0.2$ 1562 around the candidate electron cluster. It only considers the clusters with positive 1563 reconstructed energy. In addition, a correction as a function of (E_T, η) values is 1564 applied to account for the electron energy leakage outside the cluster.
- 1565 • **Track-based variable:** $p_T^{\text{varcone}20}$. It's calculated as the sum of transverse momentum 1566 of all satisfied tracks within a cone of $\Delta R = \min(0.2, 10\text{GeV}/E_T)$ around the 1567 candidate electron track. To calculate the sum, it requires the tracks are originating 1568 from the reconstruction PV of hard collision, and exclude the associated tracks of 1569 electron itself.

1570 Based on the values of $E_T^{\text{topocone}20}/p_T$ and $p_T^{\text{varcone}20}/p_T$, a series of working points with different selection requirements are defined. The resulting WPs are divided into two kinds:

- 1572 • Efficiency targeted working points: varying requirements to obtain a certain isolation 1573 efficiency, which can either be a constant or as a function of E_T .
- 1574 • Fixed requirement working points: set the constant upper thresholds on isolation 1575 variables.

1576 The distribution of two discriminating variables are shown in figure 4.7 for $Z Z \rightarrow ee$ events with $E_T > 27\text{GeV}$ and satisfying *Tight* requirement.

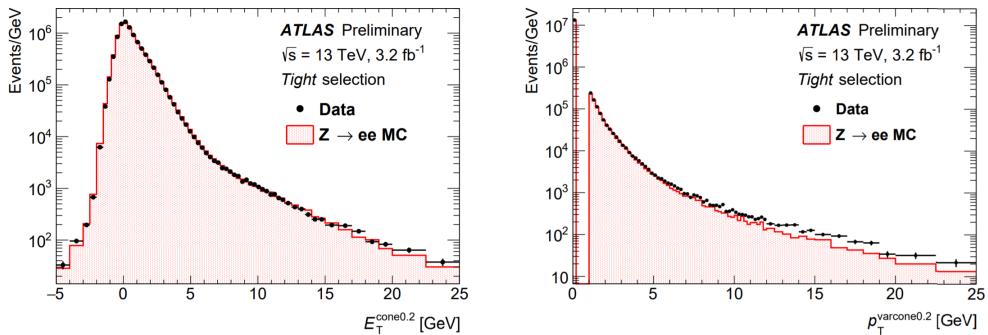


Fig. 4.7 $E_T^{\text{cone}0.2}$ (left) and $p_T^{\text{varcone}0.2}$ (right) distribution for electrons from $Z Z \rightarrow ee$ events in data and MC simulation. The simulated events (full histograms) are normalized to data.

1577 **4.2.4 Muon**

1578 Muons are important final state signatures of many physics analyses at the LHC, such
1579 as the Higgs analyses, studies of SM, BSM searches and so on. Thus, high performance of
1580 muon reconstruction and identifications are crucial. This section briefly describes some
1581 more details of the *reconstruction, identification and isolation* of muon.

1582 **Muon reconstruction**

1583 Muon reconstruction is firstly performed in inner detector (ID) and muon spectrometer (MS)
1584 independently as given in section 4.2.1. The information from each individual
1585 detector is then combined together to form the muon tracks for physics analyses. The
1586 combined ID-MS reconstruction is developed according to several algorithm based on
1587 the information from ID, MS and calorimeters. Four different muon types are defined^[54]:

- 1588 • **Combined (CB) muons:** a combined track is formed by using the reconstructed
1589 tracks performed independently in ID and MS with a global refit. To improve the fit
1590 quality, the hits from MS may be added to or removed from the track. The outside-in
1591 pattern recognition is utilized for the reconstruction of most muons, in which
1592 the muons are first reconstructed in MS and then extrapolated inward to match the
1593 ID track. In the meantime, the inside-out pattern is also used as a complementary
1594 method.
- 1595 • **Segment-tagged (ST) muons:** a reconstructed track in ID is regarded as muon, if it
1596 can be associated with at least one track segment in MDT or CSC chambers. These
1597 ST muons are used when muons can only pass across one layer of MS chambers
1598 due to their low p_T or falling into regions with less acceptance.
- 1599 • **Calorimeter-tagged (CT) muons:** a reconstructed track in ID is categorized as
1600 CT muon if it's matched to the energy deposit in calorimeter which is recognized
1601 with a minimum-ionizing particle. This CT muons have lowest purity amount all
1602 types of muons, but it covers the region where ATLAS muon spectrometer is only
1603 partially constructed. For the region of $|\eta| < 0.1$ and $15GeV < p_T < 100GeV$, the
1604 identification of CT muons is optimal.
- 1605 • **Extrapolated (ME) muons:** the ME muons are reconstructed based only on the MS
1606 track and a relative loose requirement of originating from the interaction point. In
1607 general, this type of muon needs to pass at least two (three) layers of MS chambers
1608 to provide a track measurement in barrel (forward) region. ME muons are designed
1609 to extend the reconstruction acceptance to the region of $2.5 < |\eta| < 2.7$ where ID

1610 measurement doesn't cover.

1611 Before collecting those muons for physics analyses, overlap removals are performed
1612 between different muon types with the priority of CB > ST > CT, when two types of muons
1613 share the same ID track. Besides, the overlaps with ME muons are resolved by analyzing
1614 the track hit content, and selecting the track with better fit quality and larger number of
1615 hits.

1616 **Muon identification**

1617 After reconstruction, the muon identification is then performed to further separate the
1618 signal and background, especially to suppress backgrounds from pion and kaon decays
1619 by requiring prompt muons with high efficiency and guaranteeing a robust momentum
1620 measurement. The muon identification is defined by using the fit quality of combined
1621 track. The variables utilized in judgement for CB tracks include:

- 1622 • *q/p significance*, the absolute difference between q/g (charge over momentum) of
1623 muons measured in ID and MS divided quadratic sum of their corresponding un-
1624 certainties;
- 1625 • ρ' , the absolute value of difference between the p_T measured in ID and MS, divided
1626 by the p_T of combined track;
- 1627 • *Normalized χ^2* of the combined track fit;
- 1628 • *Hit numbers in ID and MS*

1629 In addition, some new variables used for *LowPt* muon working point what will be de-
1630 scribed later^[64]:

- 1631 • *Momentum balance significance (MBS)* is computed as momentum difference be-
1632 tween the ID and MS standalone measurements with respect to the uncertainty σ
1633 on energy lost in the calorimeter system.
- 1634 • *Scattering neighbor significance (SNS)* is defined to estimated the significance of a
1635 change in trajectory along the track, expected in the presence of a hadron decaying
1636 to a muon.
- 1637 • *Scattering curvature significance (SCS)* is defined as the normalized integral of the
1638 scattering angle significances, corrected for large kinks along the trajectory.

1639 Five selection working points are developed to satisfy the different needs for different
1640 physics goals: *LowPt*, *Loose*, *Medium*, *Tight* and *HighPt*. The *Tight*, *Medium*, *Loose* are
1641 subsets from the tighter one to looser one. More detailed definition of each working point
1642 is given as follow:

- 1643 • *Loose*: this working point is designed to maximize the reconstruction efficiency
1644 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. The 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 used. In the range of $|\eta| < 1.3$, it requests muons hit at least one MS station; in $1.3 < |\eta| < 1.55$, a least two MS stations are required; while in region of $|\eta| > 1.55$, *Medium WP* is required. In addition, cuts are applied to suppress fakes as: $|\text{MBS}| < 3.0$, $|\text{SNS}| < 3.0$ and $|\text{SCS}| < 3.0$.

Figure 4.8 and 4.9 show the selection efficiency of different muon identification work-

ing points. For *Medium (Loose)*, *Tight* and *High- p_T* : $Z \rightarrow \mu\mu$ events with $p_T > 10 GeV$ are used for measurement. In the top figure of figure 4.8, the efficiency of the Loose selection (squares) is shown comparing to Medium one, where significant difference can be observed in region of $|\eta| < 0.1$. For *LowPt*, $J/\Psi \rightarrow \mu\mu$ events with $3 GeV < p_T < 10 GeV$ are used for measurement.

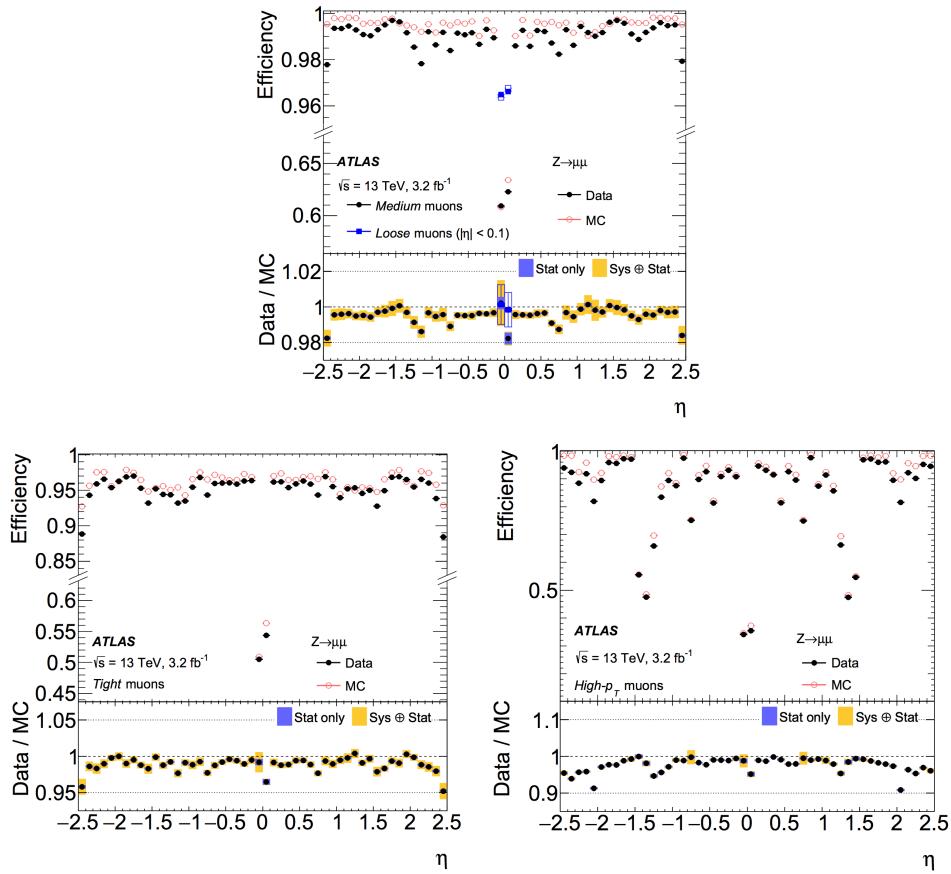


Fig. 4.8 Muon reconstruction efficiency as functions of η for: Medium (and Loose), Tight and High- p_T working points.

1684

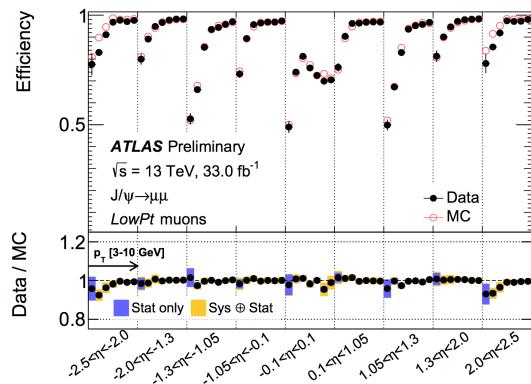


Fig. 4.9 Muon reconstruction efficiency for Low- p_T working point as a function of η .

1685

Muon isolation

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

- **Calorimeter-based variable:** $E_T^{\text{topocone}20}$. It's defined as the sum of the transverse energy of topological clusters within a cone of $\Delta R = 0.2$ around the candidate, after the correction of pile-up effects and the subtraction of the energy deposit contribution from muon itself. The contributions from pile-up and underlying events are computed using the ambient energy-density technique^[65] and are corrected on an event-by-event basis.
- **Tracked-based variable:** $p_T^{\text{varcone}30}$. It's computed from the scalar sum of the transverse momenta of the tracks with $p_T > 1\text{GeV}$ in a cone with size of $\Delta R = \min(10\text{GeV}/p_T^\mu, 0.3)$ around the candidate muon whose transverse momenta is p_T^μ and exclude the muon track candidate 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 for muons passing *Medium* identification criteria.

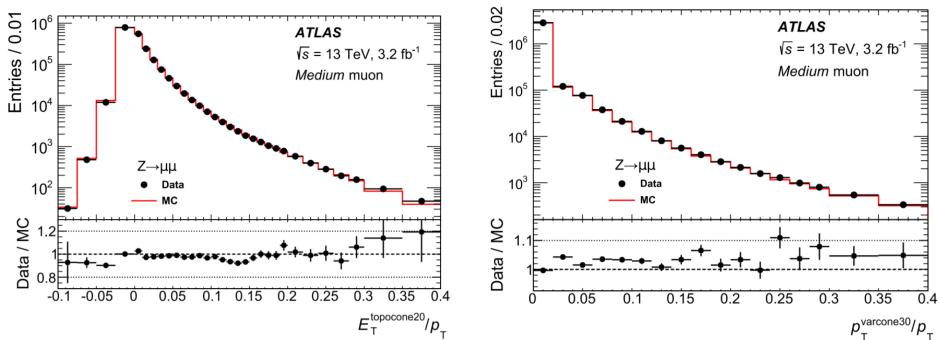


Fig. 4.10 Distributions of the calorimeter-based (right) and the track-based (left) relative isolation variables measured in $Z \rightarrow \mu\mu$ events.

1704

1705 4.2.5 Jets

1706 Jets are another important features for many physics analyses at the LHC, and es-
1707 pecially the key signatures for vector boson fusion/scattering (VBF/VBS) processes. In
1708 ATLAS detector, jets are reconstructed as groups of topologically associated energy de-
1709 posits in the calorimeters, tracks associated with charged particles measured in the inner

tacking detector, or simulated particles. This section introduces the jet reconstruction, jet energy scale (JES) calibration and the b-jet tagging techniques.

Jet reconstruction

Jets are reconstructed using anti- k_t algorithm^[66] within various radius parameter R , and usually $R = 0.4$, namely small- R jets, are used in most cases. The FastJet software package^[67] is utilized for jet finding and reconstruction. A collection of four-momentum 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*: use stable simulated particles as the inputs to jet algorithm.
- *Track jets*: use charged tracks measured from inner detector as inputs to jet algorithm.
- *Calorimeter jets*: make use of the energy deposits in calorimeters as inputs to jet algorithm.

Figure 4.11 shows the schematic of ATLAS jet reconstruction.

Jet reconstruction

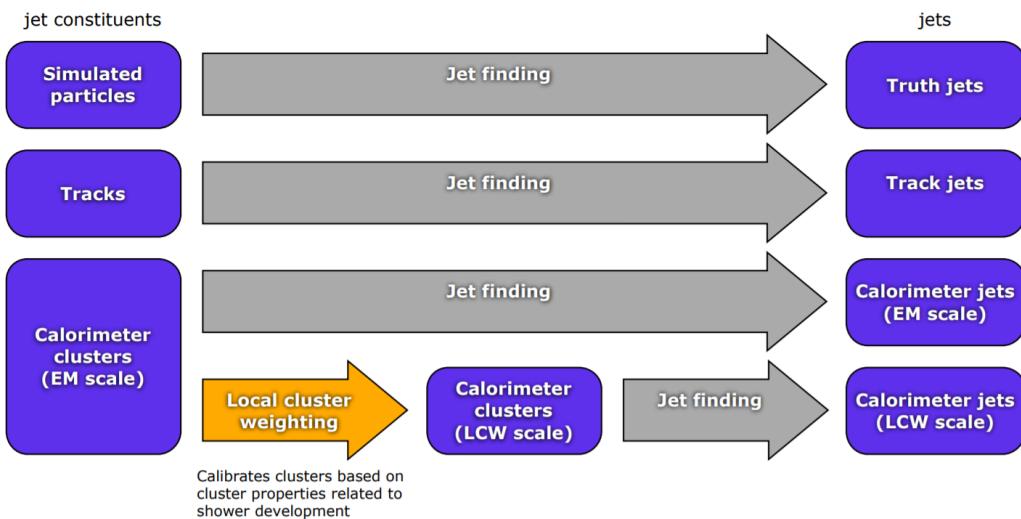


Fig. 4.11 A overview schematic of ATLAS jet reconstruction^[68].

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 se-

lected. Tracks are assigned to jets using ghost association^[65], 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^[69]. Topo-clusters are built from nearby 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 elaborates 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.

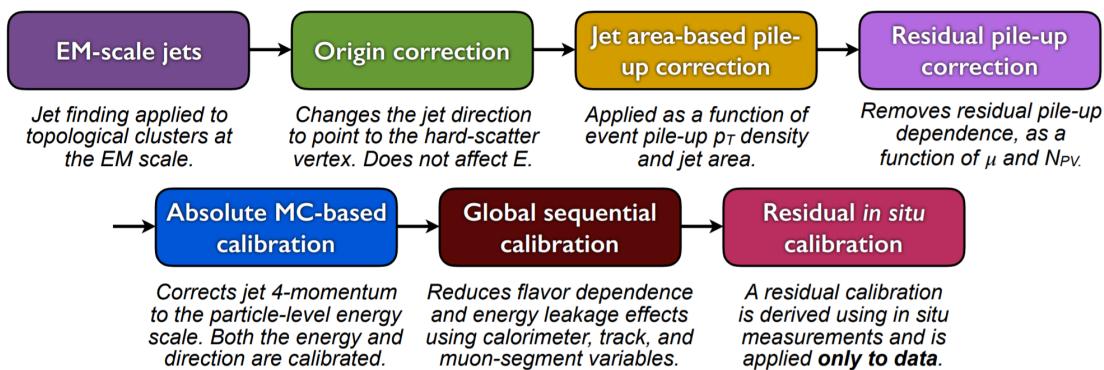


Fig. 4.12 A overview schematic of ATLAS jet calibration^[69].

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 by 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

1758 in-time and out-of-time pile-up, which consists of two processes: an area-based p_T den-
1759 sity subtraction applied on the top of each event; and a residual correction derived from
1760 the simulation. Thirdly, the absolute JES calibration corrects the jet four-momentum to
1761 the particle-level energy scale, using truth jets in di-jet MC events. Furthermore, the step
1762 of global sequential calibration uses calorimeter, track and MS-based variables to reduce
1763 the flavor dependence and energy leakage effects. Finally, the residual in situ calibration
1764 is utilized to correct jets in data by using well-measured objects such as photons, Z bosons
1765 and calibrated jets.

1766 **B-jet tagging**

1767 Tagging of b-jets plays a important role in many physics analyses involving b- or t-
1768 quark. In the meantime, lots of analyses need to apply b-jet veto to suppress $t\bar{t}$ process.
1769 There are three major types of algorithms that have been developed to distinguish b-quark
1770 jets from light-quark (u,d,s) jets^[70]:

1771 • **Impact parameter based algorithms (IP2D and IP3D):** b-hadrons usually have
1772 long lifetime (~ 1.5 ps, $c_\tau \sim 450$ μm), which leads to large impact parameter for
1773 tracks produced from b-hadron decay. The impact parameter taggers are devel-
1774 oped based on these variables. The IP2D tagger makes use of the transverse im-
1775 pact parameter significance $d_0/\sigma(d_0)$ as discriminant, while IP3D tagger uses two-
1776 dimensional discriminant of both transverse and longitudinal impact parameter sig-
1777 nificances: $d_0/\sigma(d_0)$ and $z_0 \sin\theta/\sigma(z_0)$.

1778 • **Secondary vertex finding algorithm (SV1)** makes use of the secondary vertex
1779 formed by decay products of b-hadron within the jet. All track pairs within a jet are
1780 tested for a two-track vertex hypothesis, and removed if they are likely to originate
1781 from a long-live particle decay (eg. K_s or Λ), hadronic interactions or photon con-
1782 versions. After that, a new vertex is fitted with all tracks from remaining two-track
1783 vertices, and the outliers are removed from this set of tracks.

1784 • **Decay chain multi-vertex algorithm (JetFitter)**^[71] exploits the topological struc-
1785 ture of weak b- and c- hadron decays inside the jet and tries to reconstruct the full
1786 b-hadron decay chain. A Kalman filter is adopted to find a common line between
1787 primary vertex and b-/c- vertices, as well as their position in this line, which gives
1788 a approximated flight path for the b-hadron. In this approach, the b- and c-hadron
1789 vertices, whenever resolution allows, can be resolved, even when there is only a
1790 single track associated to them.

1791 The final discrimination commonly used in many physics analyses is called **Multivariate**
1792 **Algorithm (MV2)**, which is based on Boosted Decision Tree (BDT) implemented in the

1793 TMVA package [72] by combining the outputs from underlying taggers mentioned above.
 1794 The MV2 was trained using jets in $t\bar{t}$ sample, where the b-jets are treated as signal while
 1795 the c- and light-flavor jets are treated as backgrounds. There are three kinds of MV2
 1796 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.

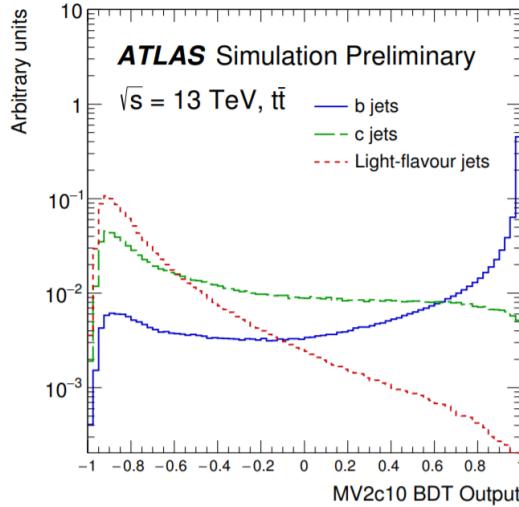


Fig. 4.13 MV2c10 BDT output for b- (solid blue), c- (dashed green) and light-flavour (dotted red) jets in $t\bar{t}$ events [70].

1797

1798 4.2.6 Missing transverse energy

1799 Many interesting physics processes are with the involvement of neutrinos. Since they
 1800 do not interact with any materials in the detector, neutrinos cannot be detected directly;
 1801 but instead, they can result in imbalance in the plane transverse to the beam axis, where
 1802 momentum conservation is assumed. It is known as the missing transverse momentum
 1803 denoted as E_T^{miss} , which is obtained from the negative vector sum of the momenta of all
 1804 particles detected in a proton-proton collision event.

1805 The E_T^{miss} is measured using selected, reconstructed and calibrated hard objects in an
 1806 event. Its x- and y- components can be calculated as follow:

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

1807 where each object term is given by the negative vectorial sum of the momenta of the
 1808 respective calibrated objects. The calorimeter signals are associated with the recon-
 1809 structed objects in the following order: electrons, photons, hadronically decaying taus,
 1810 jets, muons. The soft term is reconstructed from detected objects not match any hard
 1811 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} [73].

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$ $\Delta R(\text{track}, e/\gamma\text{cluster}) > 0.05$ $\Delta R(\text{track}, \tau_{had}) > 0.2$	$E_T^{miss,soft}$	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)

1812

1813 Based on $E_{x(y)}^{miss}$, the magnitude of E_T^{miss} and the azimuthal angle ϕ^{miss} are computed:

$$\begin{aligned} E_T^{miss} &= \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \\ \phi^{miss} &= \arctan(E_y^{miss}/E_x^{miss}) \end{aligned} \quad (4.3)$$

1814 In equation 4.2, each objects are required to pass certain reconstruction and calibrated
 1815 criteria and selections mentioned above before taken as inputs.

1816 In figure 4.14, left plot shows the observed E_T^{miss} distribution for data and MC of $Z \rightarrow$
 1817 $\mu\mu$ events without genuine missing transverse momentum; and right plot shows the E_T^{miss}
 1818 distribution for $W \rightarrow ev$ events that has genuine (true) missing transverse momentum
 due to real neutrino.

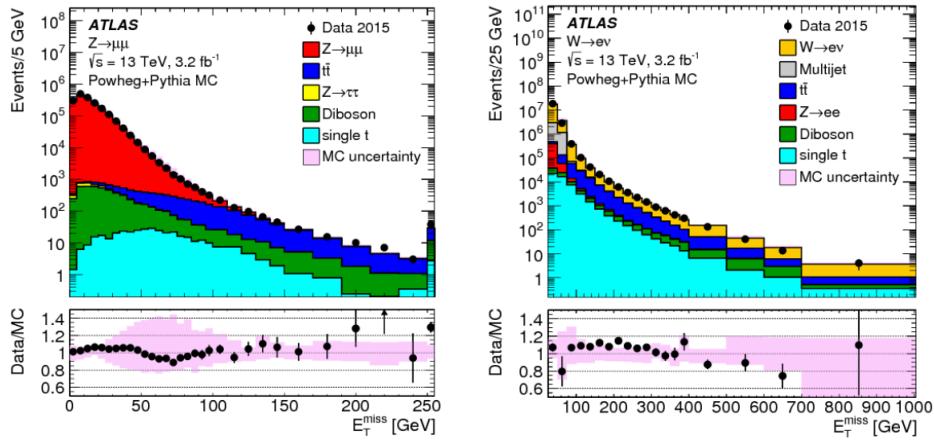


Fig. 4.14 Measured E_T^{miss} distribution for $Z \rightarrow \mu\mu$ events (left) and $W \rightarrow ev$ events (right).

1819

1820 **Chapter 5 Statistical treatment of searching for new 1821 particles or processes**

1822 In the experiments of particle physics, one often searches for particles or processes
 1823 that have been predicted but not yet observed, such as the two analysis presented in this
 1824 dissertation: searching for the vector boson scattering process and searching for the heavy
 1825 resonance(s). Usually two hypotheses are defined:

- 1826 • H_0 : null hypothesis, in most cases are designated as background-only hypothesis.
- 1827 • H_1 : signal plus background hypothesis, where signal is a new model one would
 1828 like to search for.

1829 For the purpose of discovering a new signal process, the H_0 hypothesis is tested against
 1830 the alternative H_1 . When setting limits, the H_1 hypotheses with different signal strengths
 1831 are tested against the H_0 .

1832 The level of agreement between observed data and a given hypothesis can be quan-
 1833 tified by computing the p -value, the probability under this hypothesis assumption, or
 1834 its equivalent Gaussian significance. This section describes the statistical treatment for
 1835 searches related to this dissertation.

1836 5.1 The likelihood function

1837 The likelihood function is defined as the product of a set of the probability den-
 1838 sity functions (pdfs) of variables x , that used to evaluate the probability of the observed
 1839 dataset:

$$\mathcal{L}(x_1, \dots, x_N; \theta_1, \dots, \theta_M) = \prod_i^N f(x_i; \theta_1, \dots, \theta_M) \quad (5.1)$$

1840 where $\theta_1, \dots, \theta_M$ are the nuisance parameters that can be written as $\boldsymbol{\theta}$, and x_1, \dots, x_N
 1841 denote the observables of dataset. Usually one measures the variable x by constructing a
 1842 histogram $\mathbf{n} = (n_1, \dots, n_N)$. The expectation value of the i th bin n_i is written as^[74]:

$$E[n_i] = \mu s_i + b_i \quad (5.2)$$

1843 where μ is the signal strength, s_i and b_i are the number of signal and background events
 1844 in that bin. In addition to the histogram \mathbf{n} , in some cases, one would like to use subsidiary
 1845 measurements to further constrain the nuisance parameters. For instance, due to the lack
 1846 of background simulation or the mismodelling issue of one MC sample, one can choose

1847 a control region and construct another histogram $\mathbf{m} = (m_1, \dots, m_M)$ to constrain the con-
 1848 tribution of one certain background in data. For this measurement, the expectation value
 1849 of the i th bin m_i is written as:

$$E[m_i] = u_i(\boldsymbol{\theta}) \quad (5.3)$$

1850 In most particle experiments, the number of these events observed in one bin follows
 1851 the Poisson distribution, by combining the equation 5.2 and 5.3, one can get the likelihood
 1852 function for all bins as:

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_{i=1}^N \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)} \prod_{k=1}^M \frac{u_k^{m_k}}{m_k!} e^{-u_k} \quad (5.4)$$

1853 Then the profile likelihood ratio is defined to test the hypothesized value of μ :

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\boldsymbol{\theta}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})} \quad (5.5)$$

1854 where numerator denotes the local maximum-likelihood for a specific μ , $\hat{\boldsymbol{\theta}}$ is the value of
 1855 $\boldsymbol{\theta}$ that maximizes the numerator. And the denominator is the global maximum-likelihood
 1856 with the $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ as their best fit value.

1857 5.2 Test statistic

1858 To test the level of agreement between the data and the hypothesized value μ , a test
 1859 statistic t_μ can be defined as^[74]:

$$t_\mu = -2 \ln \lambda(\mu) \quad (5.6)$$

1860 From the definition of $\lambda(\mu)$ in equation 5.5, one can see that $0 \leq \lambda \leq 1$, while a λ
 1861 with value close to 1 implies good agreement between data and μ . Thus, smaller value
 1862 of t_μ means the increase of compatibility between data and μ . To quantify the level of
 1863 disagreement, one can calculate the p -value as:

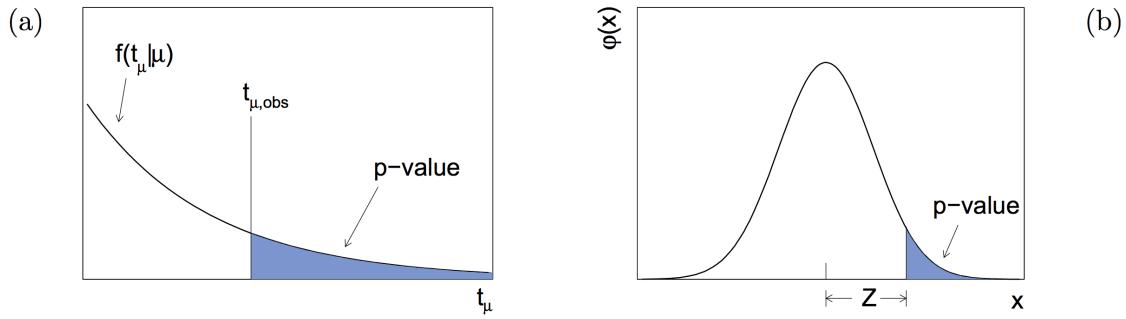
$$p_\mu = \int_{t_{\mu,obs}}^{\infty} f(t_\mu | \mu) dt_\mu \quad (5.7)$$

1864 in which $t_{\mu,obs}$ is the value of test statistic from observed data, and $f(t_\mu | \mu)$ is the pdf
 1865 of t_μ under the assumption of hypothesized value μ . This is a one-side p -value with its

1866 corresponding observed significance, Z , is defined as:

$$Z = \Phi^{-1}(1 - 2p_\mu) \quad (5.8)$$

1867 The relationship between the t_μ , p -value and significance Z are depicted in figure 5.1.
 1868 When searching for a signal process, such as Higgs boson, the particle physics commu-
 1869 nity tends to claim a discovery when the rejection of background-only hypothesis has a
 1870 significance of at least $Z = 5$.



**Fig. 5.1 (a) Illustration of the relationship between the observed t_μ and its p -value. (b) The re-
lation between p -value and the observed significance Z , where $\phi(x)$ is the standard
normal distribution^[74].**

1871 In most cases, one assumes that the presence of a new signal can only increase the
 1872 event rate comparing to the background only model, which means the signal strength
 1873 cannot be a negative value ($\mu \geq 0$). And for the case of discovery, the hypothesis of a
 1874 positive signal strength should be tested against to the background-only (null) hypothesis
 1875 by using the test statistic called q_0 :

$$q_0 = \begin{cases} -2\ln(\lambda(0)) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \quad (5.9)$$

1876 which corresponds to the p -value called p_0 :

$$p_0 = \int_{q_{0, \text{obs}}}^{\infty} f(q_0 | 0) dq_0 \quad (5.10)$$

1877 to quantify the level of disagreement between the data and the null hypothesis ($\mu = 0$).

1878 5.3 The CLs upper limit

1879 For a signal hypothesized value μ , one can compute the probability that this hypothesis
 1880 (called S+B hypothesis) gives a **greater** test statistic value than the observed one q_{obs} as:

$$1881 p_{s+b} = \int_{q_{obs}}^{\infty} f(q_{\mu}|\mu) dq_{\mu} \quad (5.11)$$

1881 In the meantime, the probability that the background-only hypothesis gives a **smaller** test
 1882 statistic than observed data can also be calculated as:

$$1883 1 - p_b = \int_{-\infty}^{q_{obs}} f(q_{\mu}|0) dq_{\mu} \quad (5.12)$$

1883 Then we define the CLs^[75] of a hypothesized value μ as:

$$CLs = \frac{p_{s+b}}{1 - p_b} \quad (5.13)$$

1884 For purpose of excluding a signal hypothesis, a threshold CLs of 0.05 is often used. For
 1885 this reason, usually under the circumstance that no significant derivation between data and
 1886 background-only hypothesis is found, one would like to find the value of hypothesized
 1887 signal strength μ by requiring its $CLs = 0.05$ (called 95% CLs upper limit) for exclusion.

1888 The sensitivity of an experiment to exclude a new signal process is quantified by *me-*
 1889 *dian upper limit*^[76], which is obtained using “Asimov dataset”. The Asimov dataset is
 1890 defined such that when one uses it to evaluate the estimators for all parameters, one obtains
 1891 the true parameter values. Moreover, it is useful to use Asimov dataset to compute how
 1892 much the sensitivity is expected to vary, given the expected fluctuations in the data. The $\hat{\mu}$
 1893 is assumed to follow a Gaussian distribution with a mean value of μ' and the standard de-
 1894 viation of σ . First of all, the test statistic from profile likelihood ratio can be approximated
 1895 as^[74]:

$$-2\ln\lambda(\mu) = \frac{(\mu - \hat{\mu})^2}{\sigma^2} + \mathcal{O}(1/\sqrt{N}) \quad (5.14)$$

1896 Given that the Asimov dataset corresponding to a signal strength μ' , one finds:

$$-2\ln\lambda_A(\mu) \approx \frac{(\mu - \mu')^2}{\sigma^2} = q_{\mu,A} \quad (5.15)$$

1897 where $q_{\mu,A} = -2\ln\lambda_A(\mu)$ is the observed test statistic of Asimov dataset. Then the stan-

1898 dard derivation can be computed as:

$$\sigma_A^2 = \frac{(\mu - \hat{\mu})^2}{q_{\mu,A}} \quad (5.16)$$

1899 In a special situation where one wants to find the median exclusion significance for the
1900 hypothesis μ assuming that there is no signal ($\mu' = 0$), one gets:

$$\sigma_A^2 = \frac{\hat{\mu}^2}{q_{0,A}} \quad (5.17)$$

1901 **5.4 Nuisance parameters**

1902 The expected numbers and pdf shapes of signal and background events also depend on
1903 a series of systematic uncertainties, which are described as a set of nuisance parameters
1904 (NPs). As showed in equation 5.1, θ is a set of NPs that plays as an additional “penalty”
1905 term to likelihood function, which will increase the negative log likelihood when any
1906 nuisance parameter is shifted from its nominal value. Usually those NPs are constrained
1907 by using Gaussian function with their estimated uncertainties provided by the experiment
1908 condition.

1909 **Chapter 6 Studies of SM ZZ production in $\ell\ell\ell'\ell'$ final** 1910 **state using pp collision data collected by ATLAS** 1911 **detector from 2015 to 2018**

1912 6.1 Introduction

1913 After the discovery of Higgs boson^[6-7], the examination of electroweak symmetry
1914 breaking (EWSB) becomes a main focus at the LHC. In addition to measuring the prop-
1915 erties of Higgs boson directly, the vector boson scattering (VBS) process is another key
1916 avenue to probe EWSB^[77-79]. As introduced in section 2.1.3, in Standard Model (SM),
1917 the Higgs boson acts as “moderator” to unitarize high-energy longitudinal VBS ampli-
1918 tudes at the TeV scale. Therefore, studying high-energy behaviours of VBS is crucial to
1919 understand the mechanism of EWSB.

1920 Since no VBS process was observed prior to the LHC era, LHC provides an exception-
1921 able opportunity to study them due to its unprecedented high energy and luminosity. At
1922 the LHC, the VBS process is typically studied through the measurements of electroweak
1923 (EW) production of two vector bosons radiated from quark-quark initial state, plus a pair
1924 of hadronic jets with high energy in the back and forward regions (denoted as EW- $VVjj$).
1925 The quantum chromodynamics (QCD) production of $VVjj$ containing two QCD vertices
1926 at the lowest order (denoted as QCD- $VVjj$) is an irreducible background to the search
1927 of EW- $VVjj$ production. The features of EW- $VVjj$ production including a large invari-
1928 ant mass of jet pair (m_{jj}) and a significant separation of rapidity between two jets (Δy_{jj}).
1929 Figure 6.1 presents some typical Feynman diagrams of EW- and QCD- $ZZjj$ processes.
1930

1931 The first evidence of the EW- $VVjj$ process was seen in same-sign WW channel
1932 (EW- $W^\pm W^\pm jj$) by ATLAS collaboration with 20.3 fb^{-1} 8 TeV data^[80], in which a 3.6σ
1933 excess was observed in data over the background-only prediction. In the LHC run-2, the
1934 observation (with $> 5 \sigma$ statistical significance) of EW- $W^\pm W^\pm jj$ process has been re-
1935 ported in both ATLAS and CMS collaboration with 36 fb^{-1} 13 TeV data^[81-82]. In WZ
1936 channel (EW- $WZjj$), an observation with 5.3σ excess was also reported by the AT-
1937 LAS collaboration recently^[83]. As for the EW- $ZZjj$ production, it was searched by
1938 CMS using 35.9 fb^{-1} 13 TeV data but no evidence was found^[84]. The EW production in
1939 ZZ final state (EW- $ZZjj$) is typically rare, whose fiducial cross section has an order of
1940 $O(0.1) \text{ fb}^{-1}$ in the final state where both Z bosons decay leptonically. But in the mean-

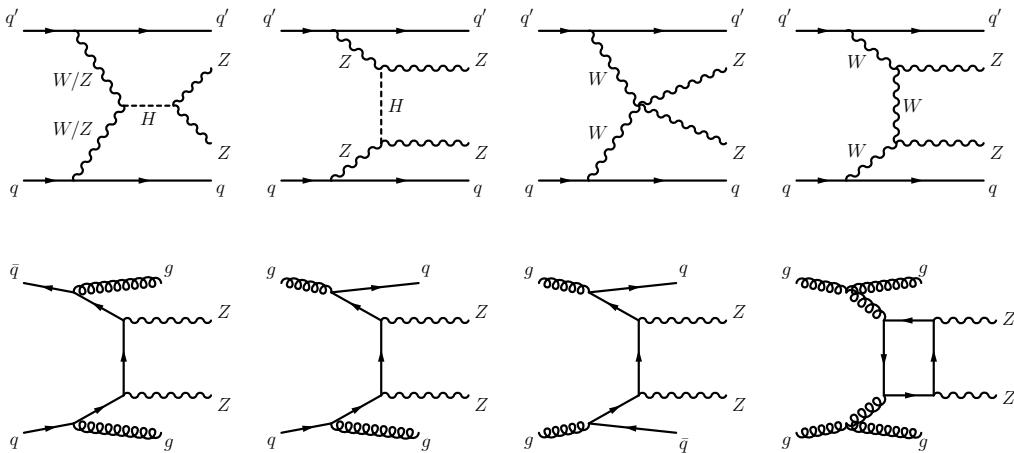


Fig. 6.1 Typical diagrams for the production of $ZZjj$, including the relevant EW VBS diagrams (first row) and QCD diagrams (second row).

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

This section presents the first observation of EW- $ZZjj$ production decaying to four charged leptons with two jets ($\ell\ell\ell'\ell' jj$) by ATLAS collaboration using the complete set of the LHC run-2 data with 139 fb^{-1} luminosity^[85-86]. 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 ZZ production including both EW and QCD processes is also reported. The $ZZjj$ 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 $\ell\ell\ell'\ell' jj$ 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.

6.2 Data and MC samples

6.2.1 Data samples

The datasets 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, corresponding to an integrated luminosity of $139.0 \pm 2.4 \text{ fb}^{-1}$.

6.2.2 MC simulations

The EW- $Z Z jj$ production is modelled using MadGraph5_aMC@NLO 2.6.1^[87] with the matrix elements (ME) calculated in the leading-order (LO) approximation in perturbative QCD (pQCD) and with the NNPDF2.3LO^[88] parton distribution functions (PDF). The VBF Higgs process is also included.

The QCD- $Z Z jj$ production is modelled using Sherpa 2.2.2^[89] with the NNPDF3.0NNLO^[90] PDF, where events with up to one (three) outgoing partons are generated at NLO (LO) in pQCD. The production of $Z Z jj$ from the gluon-gluon initial state with a four-fermion loop or with an exchange of the Higgs boson has an order of α_S^4 in QCD, and is not included in the Sherpa simulation. A separate gg induced $Z Z + 2\text{jets}$ sample is modelled using Sherpa 2.2.2 with the NNPDF3.0NNLO PDF and with an additional 1.7 k-factor^[91] being applied. Then the interference between EW- and QCD- $Z Z jj$ is modelled with MadGraph5_aMC@NLO 2.6.1 calculated at LO.

The diboson productions from QCD $WW \rightarrow \ell\nu qq$ as well as QCD and EW $WZ \rightarrow \ell\ell qq$ are modelled using Sherpa 2.2.2 with the NNPDF3.0NNLO PDF. The productions of semileptonic decays ($WW \rightarrow \ell\nu qq$ and $WZ \rightarrow q\bar{q}\ell\ell$) are modelled using Powheg-Box v2^[92] with the CT10 PDF^[93]. The triboson production is modelled using Sherpa 2.2.2 with the NNPDF3.0NNLO PDF.

For top-quark pair ($t\bar{t}$) production, the Powheg-Box v2 is used with the CT10 PDF. The single top-quark production in t -channel, s -channel and Wt -channel are simulated using the Powheg-Box v1 event generator^[94-96]. The productions of $t\bar{t}$ in association with vector boson(s) ($t\bar{t}V$) are modelled with MadGraph5_aMC@NLO 2.3.3 for $t\bar{t}W$ and $t\bar{t}Z$ with $Z \rightarrow \nu\nu/\bar{q}q$ decays, with Sherpa 2.2.1 for $t\bar{t}Z$ where the Z decays to dilepton, and with MadGraph5_aMC@NLO 2.2.2 for $t\bar{t}WW$ respectively.

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

For all the samples except those from Sherpa, the parton showering is modelled with Pythia8^[48] using the NNPDF2.3^[88] PDF set, and the A14 set of tuned parameters^[97]. While for Sherpa samples, the parton showering is simulated within the programme.

All simulated events are processed with detector response simulation 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

1996 jets reconstruction, energy scale and resolution, and the leptons identification, isolation,
 1997 trigger efficiencies for simulated events, as described in section 4.2, are all corrected to
 1998 match the data measurements.

1999 6.3 Objects and Event selections

2000 6.3.1 Object selections

2001 The selection of analysis relies on the definition of multiple objects: *electrons*, *Muons*,
 2002 and *jets*. Details of definition for each object are described as below:

2003 **Muon:** To increase the acceptance range in reconstruction (reco) -level for $\ell\ell\ell'\ell'$
 2004 channel, all four types of muons (CB, ST, CT, ME muons, described in section 4.2.4)
 2005 are used. The identified muons are then required to pass $p_T > 7$ GeV and $|\eta| < 2.7$,
 2006 and satisfy the *Loose* identification criterion (see definition in sec 4.2.4). The impact
 2007 parameter cuts are further applied to suppress the contribution from cosmic muons and
 2008 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
 2009 is the transverse impact parameter relative to the beam line, $\sigma(d_0)$ is its uncertainty, and
 2010 z_0 is the longitudinal impact parameter relative to the primary vertex. In order to avoid
 2011 muons associated with jets, all muons are required to be isolated and pass *FixedCutLoose*
 2012 isolation criteria of $E_T^{\text{topocone}20}/p_T < 0.3$ and $p_T^{\text{varcone}30}/p_T < 0.15$.

2013 **Electron:** As described in section 4.2.3, electrons are reconstructed from energy de-
 2014 posits in the EM calorimeter matched to a track in the inner detector. The electron candi-
 2015 dates must satisfy the *Loose* criterion defined by the likelihood-based (LH) method. And
 2016 electrons are required to have $p_T > 7$ GeV and $|\eta| < 2.47$. Moreover, the impact param-
 2017 eter requirements of $|d_0/\sigma(d_0)| < 5.0$ and $|z_0 \sin\theta| < 0.5$ mm are applied. Same as muon,
 2018 all electrons are required to satisfy *FixedCutLoose* isolation criteria, which, for electrons,
 2019 is $E_T^{\text{topocone}20}/p_T < 0.2$ and $p_T^{\text{varcone}20}/p_T < 0.15$.

2020 **Jets:** Jet are key signatures for VBS processes. This analysis use the jets clustered
 2021 using the anti- k_t algorithm with radius parameter $R = 0.4$, more details of jets' reconstruc-
 2022 tion can be found in section 4.2.5. The jets are required to satisfy $p_T > 30$ (40) GeV in
 2023 the $|\eta| < 2.4$ ($2.4 < |\eta| < 4.5$) region. To further reduce the effects of pile-up jets, a jet
 2024 vertex tagger (JVT) is applied to jets with $p_T < 60$ GeV and $|\eta| < 2.4$ to select jets from
 2025 hard-scattering vertex^[98].

2026 **Overlap removal:** An overlap-removal procedure is applied to selected leptons and
 2027 jets. To enhance the selection efficiency, leptons are given higher priority to be kept when
 2028 overlapping with jets. More details of the strategy is summarized in table 6.1.

	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 AND ($N_{Trk}(jet) < 3$ OR ($p_T^{jet}/p_T^\mu < 2$ and $p_T^\mu/\Sigma_{TrkP_t} > 0.7$))

Table 6.1 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.

2029 6.3.2 Event selections

2030 The events are required to additionally be recorded by single or multi-lepton triggers,
 2031 with transverse momentum (p_T) thresholds varying from 8 to 26 GeV. The overall trigger
 2032 efficiency for selected inclusive $\ell\ell\ell'\ell' jj$ signal events in the analysis region are from
 2033 95 to 99%.

2034 The $\ell\ell\ell'\ell'$ quadruplets are formed by two opposite-sign, same-flavour (OSSF) lepton pairs ($\ell^+\ell^-$), in which leptons are required to be separated by $\Delta R > 0.2$ in table 6.1.
 2035 At most one muon is allowed to be ME or CT muon. The p_T threshold of first three
 2036 leading leptons are 20, 20 and 10 GeV. If more than one quadruplets are found, the one
 2037 with minimum sum of difference between two dilepton pair masses and Z boson mass
 2039 ($|m_{l_1^+l_1^-} - m_Z| + |m_{l_2^+l_2^-} - m_Z|$) is selected. Both two dilepton pair masses are required to
 2040 be between 66 to 116 GeV. In addition, the invariant masses of all possible OSSF pairs
 2041 are required to be greater than 10 GeV to reject events from J/ϕ or Y decay.

2042 For VBS topology, the two most energetic jets in different detector side ($y_{j1} \times y_{j2} < 0$)
 2043 are selected. Furthermore, the invariant mass of two jets (m_{jj}) is required to be greater
 2044 than 300 GeV, while Δy_{jj} is required to be larger than 2. Table 6.2 summarizes the above
 2045 selection requirements, which is defined as signal region (SR).

Electrons	$p_T > 7 \text{ GeV}$, $ \eta < 2.47$ $ d_0/\sigma_{d_0} < 5$ 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$ and $ z_0 \times \sin \theta < 0.5 \text{ mm}$
Jets	$p_T > 30$ (40) GeV for $ \eta < 2.4$ ($2.4 < \eta < 4.5$)
$Z Z$ 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 all OSSF 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 6.2 Summary of selection of physics objects and candidate events at detector level in the $\ell\ell\ell'\ell' jj$ signal region.

6.4 Background estimation

Table 6.3 summarizes the background yields for $Z Z jj \rightarrow \ell\ell\ell'\ell' jj$ process in 139 fb^{-1} . Uncertainties on the predictions include both statistical and systematic components. ‘‘Others’’ includes minor contributions from non- $Z Z$ 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- $Z Z jj$	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 6.3 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.

6.4.1 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

2056 by reverting either the m_{jj} or Δy_{jj} requirements as:

$$m_{jj} < 300 \text{ GeV} \text{ or } \Delta y_{jj} < 2 \quad (6.1)$$

2057 Then this normalization factor is included into statistical fit as a free parameter to properly
 2058 treat the uncertainty correlations between SR and CR, while the shapes are taken from MC
 2059 simulation. Table 6.4 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-ZZjj	3.9 ± 0
QCD-ZZjj	136.9 ± 0.6
QCD-ggZZjj	16.8 ± 0.1
Diboson	0.3 ± 0.1
Triboson	1.6 ± 0.1
Z+jets	0
t̄t	0
Total	159.5 ± 0.62
Data	152

Table 6.4 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 6.2.2, except those with four-lepton final state. Uncertainties include only MC statistic. No events from Z+jets and t̄t MC samples pass the selection, and are indicated as 0 in the table.

2060

2061 dijet in QCD CR are shown in figure 6.2.

2062 6.4.2 Reducible backgrounds

2063 Backgrounds from Z+jets, top-quark and WZ processes called reducible back-
 2064 grounds can be estimated by data-driven method. These events usually contain two or
 2065 three leptons from Z/W decays, together with heavy-flavor jets or misidentified compo-
 2066 nents of jets reconstructed as leptons called “fake leptons”. A *fake factor* method is used
 2067 to estimate these backgrounds, where the lepton misidentification is measured in data with
 2068 the region enhances contributions from Z+jets and top-quark processes. The method is
 2069 described as below:

2070 1. Define a dedicated background dominant region to derive the fake factor for this

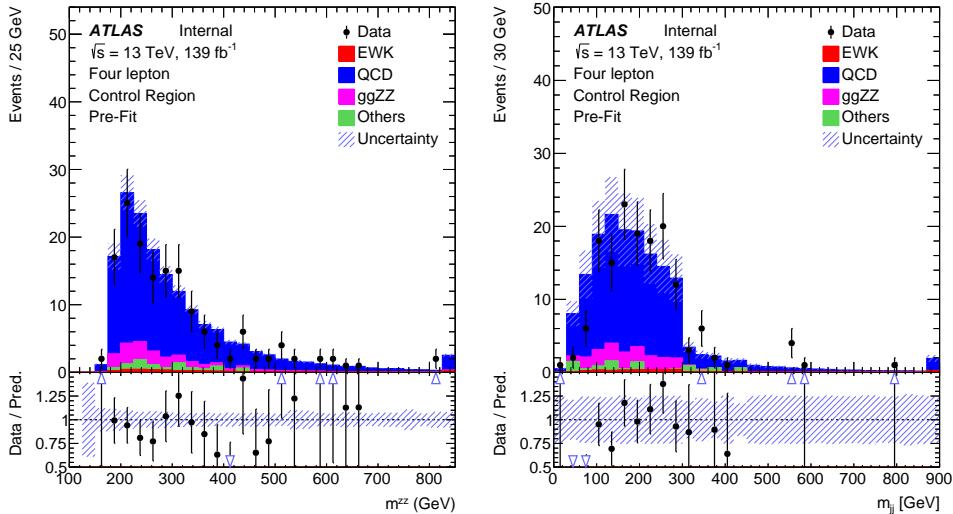


Fig. 6.2 Pre-fit m_{ZZ} and m_{ll} distribution in QCD-enriched CR.

background. The *fake factor* is defined as:

$$\mathcal{F} = \mathcal{N}_{good}/\mathcal{N}_{poor} \quad (6.2)$$

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.

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

$$\mathcal{N}_{fake} = (\mathcal{N}_{ggpp} - \mathcal{N}_{ggp}^{ZZ}) \times \mathcal{F} - (\mathcal{N}_{gpp} - \mathcal{N}_{gp}^{ZZ}) \times \mathcal{F}^2 \quad (6.3)$$

with the subtraction of ZZ contribution, and the double counting between N_{ggpp} and N_{gpp} .

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.

1. Fake factor for $Z+jets$

Fake factor for $Z+jets$ background is calculated in $Z+jets$ enriched region, where

events with one OSSF lepton pair around Z mass associated with two jets are selected.

The value of fake factor is driven from data, and is a function of p_T and η as shown in

2090 figure 6.3 for electrons and figure 6.4 for muons. During calculation, the contributions
 2091 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.

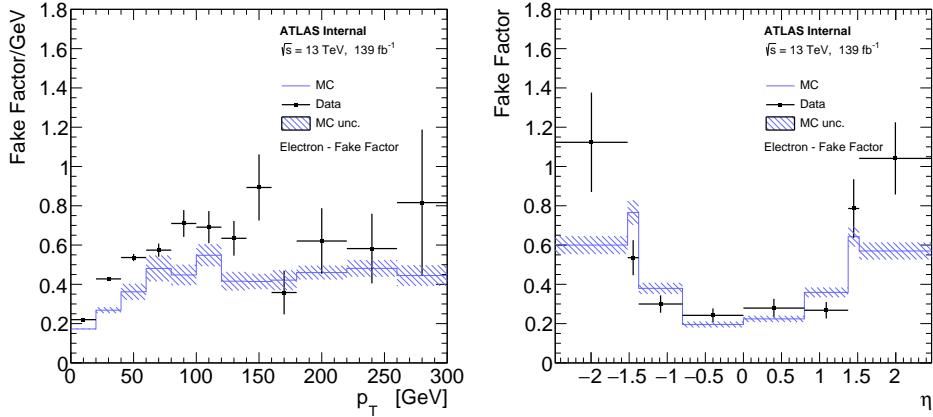


Fig. 6.3 Fake factor for $Z+jets$ background, constructed with additional electron, as a function of p_T (left) and η (right).

2092

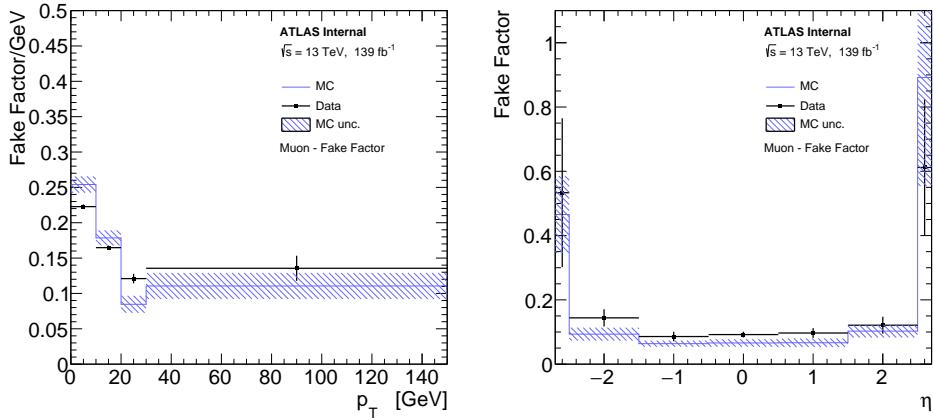


Fig. 6.4 Fake factor for $Z+jets$ background, constructed with additional muon, as a function of p_T (left) and η (right).

2093

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

2094 The fake factor for $t\bar{t}$ are calculated in $t\bar{t}$ dominant region by selecting the events that
 2095 have one $e\mu$ -pair with additional two jets. For events with three leptons, $m_T^W < 60 \text{ GeV}$ cut
 2096 is applied to reject the contribution from $t\bar{t} + W$ events. The m_T^W is defined as below:

$$m_T^W = \sqrt{2p_T^{l_3}E_T^{\text{miss}} \left[1 - \cos(\Delta\phi(p_T^{l_3}, E_T^{\text{miss}})) \right]} \quad (6.4)$$

2097 In addition, at least one b-jet is required to enhance the top component. The fake factors of
 2098 $t\bar{t}$ calculated from data as the function of p_T and η are shown in figure 6.5 for electrons and
 2099 6.6 for muons. The non- $t\bar{t}$ contributions including $Z+jets$, ZZ and WZ , are subtracted

from data.

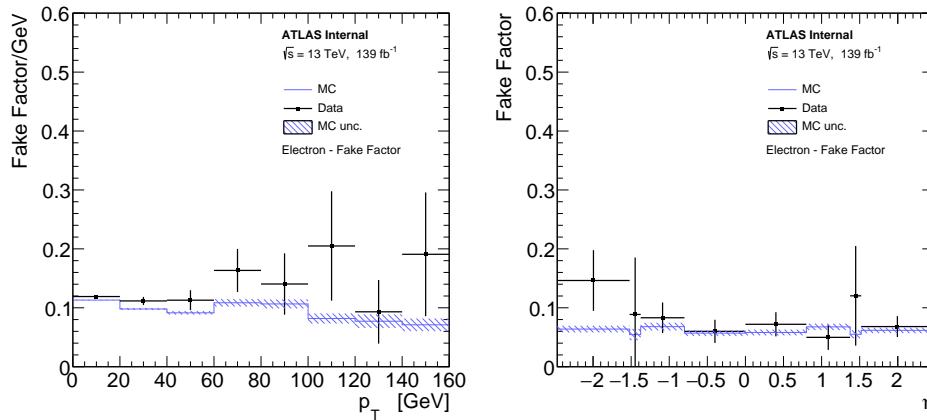


Fig. 6.5 Fake factor for $t\bar{t}$ background, constructed with additional electron, as a function of p_T (left) and η (right).

2100

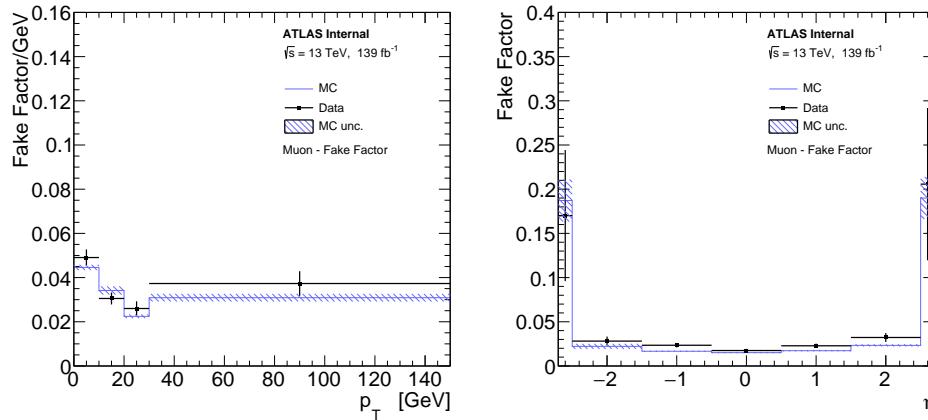


Fig. 6.6 Fake factor for $t\bar{t}$ background, constructed with additional muon, as a function of p_T (left) and η (right).

2101

3. Systematics of fake estimation and results

2102

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 additional systematics component.

2105

2106 The sources of systematics that have been included are listed as below:

2107

- 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.

2108

2109

2110

2111

- 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 6.5 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.

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 6.5 Fake background estimations in the SR. For 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.

2117

2118 6.5 Systematic uncertainties

2119 The analysis includes both the statistical fit to MD distribution to search the EW-
2120 $Z Z jj$ process, as well as the cross section measurement of inclusive EW and QCD $Z Z jj$
2121 process in fiducial volume. Therefore, theoretical and experimental uncertainties may
2122 affect the predicted background yields and shapes, the correction factors from detector-
2123 level to particle-level measurement, as well as the $Z Z jj$ MD shapes and so on. Moreover,
2124 the statistical uncertainties of simulated samples are also taken into account. Due to the
2125 extremely low cross section of $\ell\ell\ell'\ell'$ channel, the analysis is still data statistic dominant.
2126 This section describes the measurement of both theoretical and experimental systematics
2127 for $Z Z jj$ productions. The systematics for fake backgrounds have been elaborated in
2128 section 3.

2129 6.5.1 Theoretical systematics

2130 The theoretical systematics on EW- and QCD- $Z Z jj$ processes including the uncer-
2131 tainties from PDF, QCD scale, α_S and parton showering variations are summarized in ta-

ble 6.6. 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^[99-100] 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.

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

Table 6.7 summarizes the normalization uncertainties of each theoretical components in fiducial volume of SR. For QCD process, the uncertainty is QCD scale variations dominant. The normalization uncertainties are only taken into account in the measurement of $Z Z$ process cross section. As in the search of EW- $Z Z jj$ process, the QCD- $Z Z jj$ backgrounds are constrained by data during statistical fit, and only shape variations are considered.

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

The uncertainties of QCD gg -induced process ($gg \rightarrow ZZ$) as the function of MD discriminant is shown in figure 6.7 for both fiducial volume (SR) and QCD CR.

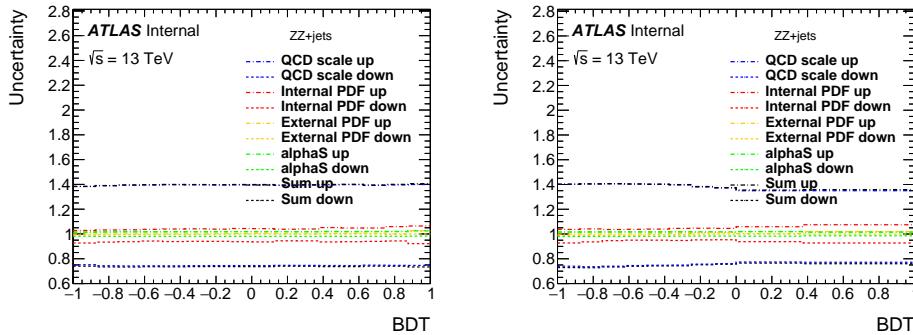


Fig. 6.7 The theoretical uncertainties for $gg \rightarrow ZZ$ background in particle-level SR (left) and CR (right).

2151 6.5.2 Experimental systematics

2152 The dominant experimental uncertainties are from the luminosity uncertainty, the mo-
 2153 mentum scale and resolution of leptons and jets, as well as the lepton reconstruction and
 2154 selection efficiency. Some smaller uncertainties, such as trigger efficiency and pile-up
 2155 correction, are also considered. Table 6.8 lists the major systematic components from
 2156 leptons and jets for signal and major background processes in $\ell\ell\ell'\ell'$ channel. The to-
 2157 tal uncertainties for sources from electron, muon and jet respectively, as well as the sum
 2158 (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 6.8 Experimental systematic 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.

2159 In addition, the uncertainty of the combined 2015 to 2018 integrated luminosity is
 2160 1.7%^[101] in ATLAS experiment, obtained using the LUCID-2 detector^[102] for the primary
 2161 luminosity measurements.

2162 On top of them, a systematic uncertainty for MD distribution with different pile-up
 2163 ($\langle\mu\rangle$) 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 6.8 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.

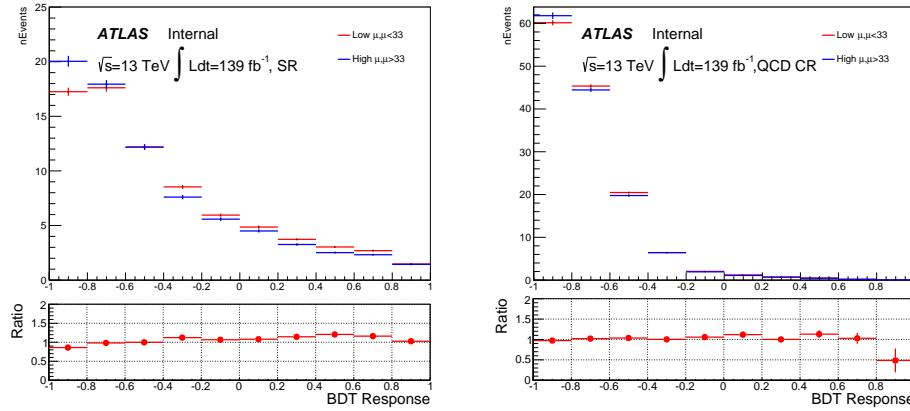


Fig. 6.8 MD distribution for QCD- $Z Z jj$ process in low and high pile-up events for SR (left) and CR (right).

Moreover, a conservative uncertainty is assigned 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 6.9. The modelling uncertainty is then calculated from the envelop between nominal and alternative samples as function of MD as one additional shape uncertainty.

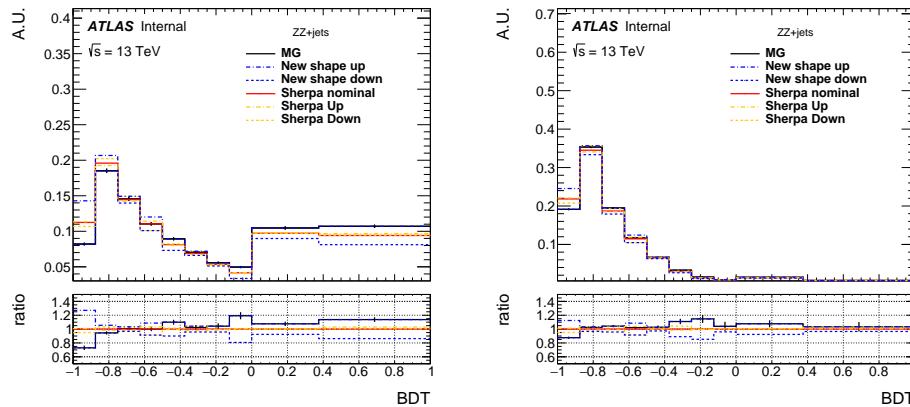


Fig. 6.9 MD shape difference for QCD $q\bar{q} \rightarrow ZZ$ background between different Sherpa theoretical uncertainties and sample from MadGraph5_aMC@NLO on SR (left) and CR (right).

2174 6.6 Measurement of fiducial cross section

2175 The fiducial cross section for inclusive $Z Z jj$ production, including both EW and
 2176 QCD components, is then measured. The definition of fiducial volume, which is used for
 2177 cross section measurement, follows closely to the detector-level selection but use physics
 2178 objects at particle-level, which are reconstructed in simulation from stable final-state par-
 2179 ticles, prior to their interactions with the detector.

2180 For electrons and muons, QED final-state radiation is for the most part recovered by
 2181 adding the four-momenta of surrounding photons that are not originating from hadrons
 2182 and within an angular distance $\Delta R < 0.1$ to the lepton four-momentum, called lepton
 2183 “dressing” at truth level. Particle-level jets are built with anti- k_T algorithm with radius
 2184 parameter $R = 0.4$ using all final-state particles except leptons and neutrinos as inputs.
 2185 Comparing to the events selection at detector-level in section 6.3, at particle-level, the
 2186 selected dilepton pair mass required is relaxed to be within 60 to 120 GeV for the reasons
 2187 of reducing the migration effect, as well as being more compatibility with CMS publica-
 2188 tion^[84]. All other kinematic selections are the same as the definition at detector-level.

2189 6.6.1 Calculation of C-factor

2190 C-factor is defined as the ratio between the number of selected events at detector-level
 2191 and the number of particle-level events in fiducial volume (FV):

$$2192 \quad C = \frac{N_{detector-level}}{N_{FV.}} \quad (6.5)$$

2193 The value of C-factor for each $Z Z jj$ process are firstly calculated from each individual
 2194 simulation samples together with their systematics. Then the C from different processes
 2195 are combined together according to their fraction of events in fiducial volume, and used
 as inputs for cross section calculation:

$$2196 \quad C = \sum_i \frac{N_{FV.}^i}{\sum_j N_{FV.}^j} \times C_i = 0.699 \pm 0.003(stats.) \pm^{0.011}_{0.013} (theo.) \pm 0.028(exp.) \quad (6.6)$$

2197 The “stats.” refers to the statistical uncertainty from MC simulation statistics. The “theo.”
 and “exp.” denote the theoretical and experimental uncertainties described in section 6.5.

2198 6.6.2 Result of fiducial cross section

2199 The cross section in fiducial volume is computed as:

$$2200 \quad \sigma^{FV.} = \frac{N_{data} - N_{bkg}}{C \times Lumi} \quad (6.7)$$

where N_{data} and N_{bkg} denote the number of events selected from detector-level selection from data and sum of backgrounds, and C is the C-factor calculated above, Lumi represents the integrated luminosity of data from 2015 to 2018 of 139 fb^{-1} . Table 6.9 shows the fiducial cross section for $\ell\ell\ell'\ell'$ final state measured from equation 6.7, as well as the predicted cross section measured from MC simulation directly.

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

Table 6.9 Measured and predicted fiducial cross-sections in $\ell\ell\ell'\ell' jj$ final-state. Uncertainties due to different sources are presented.

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

6.7 Search for EW- $ZZjj$

6.7.1 MD discriminant

To further separate the EW- $ZZjj$ component from QCD- $ZZjj$, a MD based on Gradient Boosted Decision Tree (BDT) algorithm^[103] is trained with simulated events via TMVA framework^[72]. Training is performed between EW (signal) and QCD (background) processes. Twelve event kinematic variables sensitive to the characteristics of the EW signal are used as input features in training. Table 6.10 lists those input variables with the order of their importance in BDT response provided by TMVA tool. One can see the jet-related information provides larger sensitivity. Then the MD distributions in both SR and QCD CR region are used for statistical fit.

6.7.2 Fitting procedure

A profile likelihood fit, as described in chapter 5, is performed on MD discriminant to extract the EW- $ZZjj$ signal from backgrounds. The binning of MD distributions in SR is optimized to maximize the sensitivity for detecting EW signal. The normalization of QCD- $ZZjj$ production (μ_{QCD}^{III}) in $\ell\ell\ell'\ell'$ channel is determined by data from simultaneously fit in SR and QCD CR as described in section 6.4. The signal strength of EW- $ZZjj$ production (μ_{EW}) is taken as parameter of interest and floated in the fit. The effects of the uncertainties related to normalizations and shapes described previously in section 6.5 of background processes in the MD distribution are all taken into account.

In most case, a common nuisance parameter is used for each source of systematic in

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 $Z Z jj$ 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 6.10 Input features for the training of MD.

all bins and all categories. The statistical uncertainties for simulated samples are uncorrelated among all bins, and the background uncertainties only applied to their corresponding backgrounds. Furthermore, to be more conservative, the generator modelling uncertainty for QCD- $Z Z jj$ production mentioned in section 6.5 is separated to be two nuisance parameters in low and high MD region.

6.7.3 Result of fit

The statistical fit is performed both in individual $\ell\ell\ell'\ell'$ channel, as well as the combination between $\ell\ell\ell'\ell'$ and $\ell\ell\nu\nu$ channel to gain more statistic. The results of statistical fit in $\ell\ell\ell'\ell'$ final state, and the one in combined channel are presented in table 6.11. The $\ell\ell\nu\nu$ analysis will not be described in this dissertation, but more details can refer to^[104].

To drive expected results, the observed data is used for QCD CR to extract normalization factor of QCD component ($\mu_{QCD}^{\ell\ell\ell'\ell'}$), while in SR, asimov data built from background prediction and signal model with SM assumed cross section is used.

	μ_{EW}	$\mu_{QCD}^{\ell\ell\ell'\ell'}$	Significance Obs. (Exp.)
$\ell\ell\ell'\ell'$	1.54 ± 0.42	0.95 ± 0.22	$5.48 (3.90) \sigma$
Combination of $\ell\ell\ell'\ell'$ and $\ell\ell\nu\nu$	1.35 ± 0.34	0.96 ± 0.22	$5.52 (4.30) \sigma$

Table 6.11 Observed μ_{EW} and $\mu_{QCD}^{\ell\ell\elljj}$, as well as the observed and expected significance from the individual $\ell\ell\ell'\ell'$ channel. The full set of systematic uncertainties are included.

As a conclusion, in $\ell\ell\ell'\ell'$ channel, the background-only hypothesis is rejected at 5.5σ (3.9σ) for observed (expected) data, which leads to the observation of EW- $ZZjj$ production.

Figure 6.10 shows the post-fit MD distributions for $\ell\ell\ell'\ell'$ events after performing a combined fit 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 ± 0.26 fb.

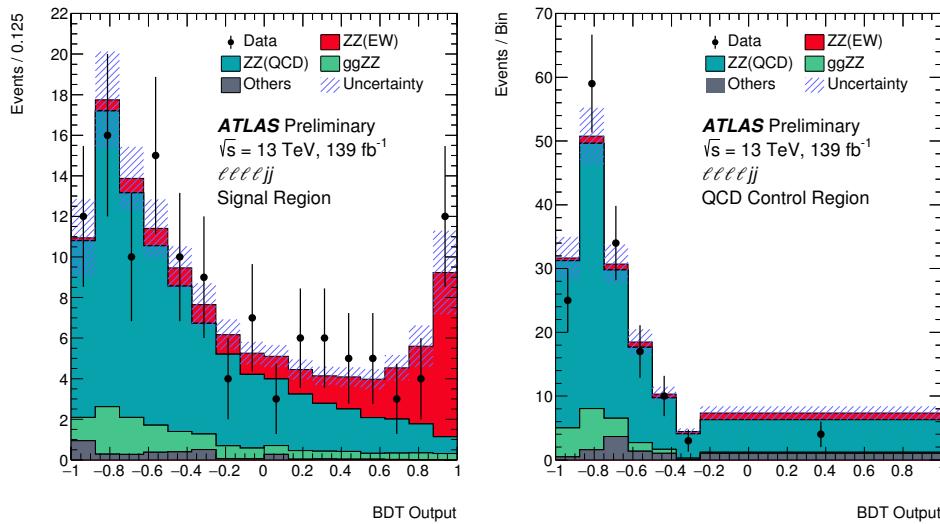


Fig. 6.10 Observed and post-fit expected multivariate discriminant distributions after the statistical fit in the $\ell\ell\ell'\ell'$ SR (left) and QCD CR (right). The error bands include the experimental and theoretical uncertainties, as well as the uncertainties in μ_{EW} and $\mu_{QCD}^{\ell\ell\elljj}$. The error bars on the data points show the statistical uncertainty on data.

Figure 6.11 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 6.11. High m_{jj} region is more sensitive for EW- $ZZjj$ events detection from this figure. Figure 6.12 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.

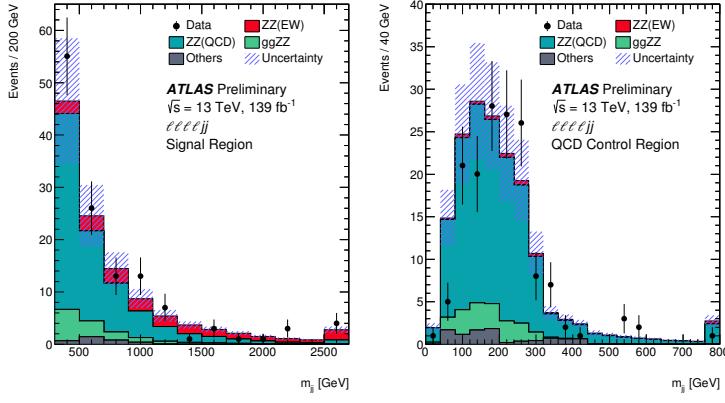


Fig. 6.11 Observed and post-fit expected m_{jj} distributions in SR (left) and QCD CR (right). The error bands include the expected experimental and theoretical uncertainties. The error bars on the data points show the statistical uncertainty. The contributions from the QCD and EW production of $Z Z j j$ 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.

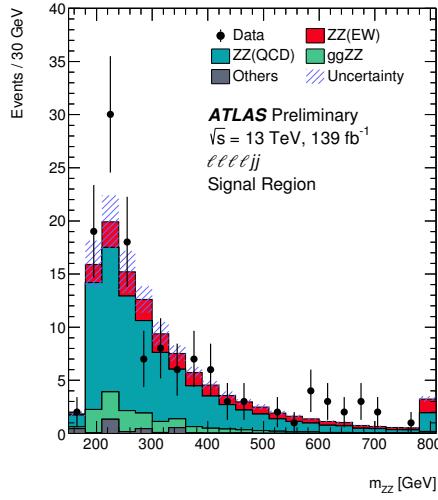


Fig. 6.12 Observed and post-fit expected m_{ZZ} spectrum in SR. The error bands include the expected experimental and theoretical uncertainties. The error bars on the data points show the statistical uncertainty. The contributions from the QCD and EW production of $Z Z j j$ events are scaled by 0.96 and 1.35, respectively, corresponding to the observed normalization factors in the statistical fit. The last bin includes the overflow events.

2251 Figure 6.13 is the display of one event candidate of EW- $Z Z jj$ production in $2e2\mu$ final state with two jets in forward and backward region.

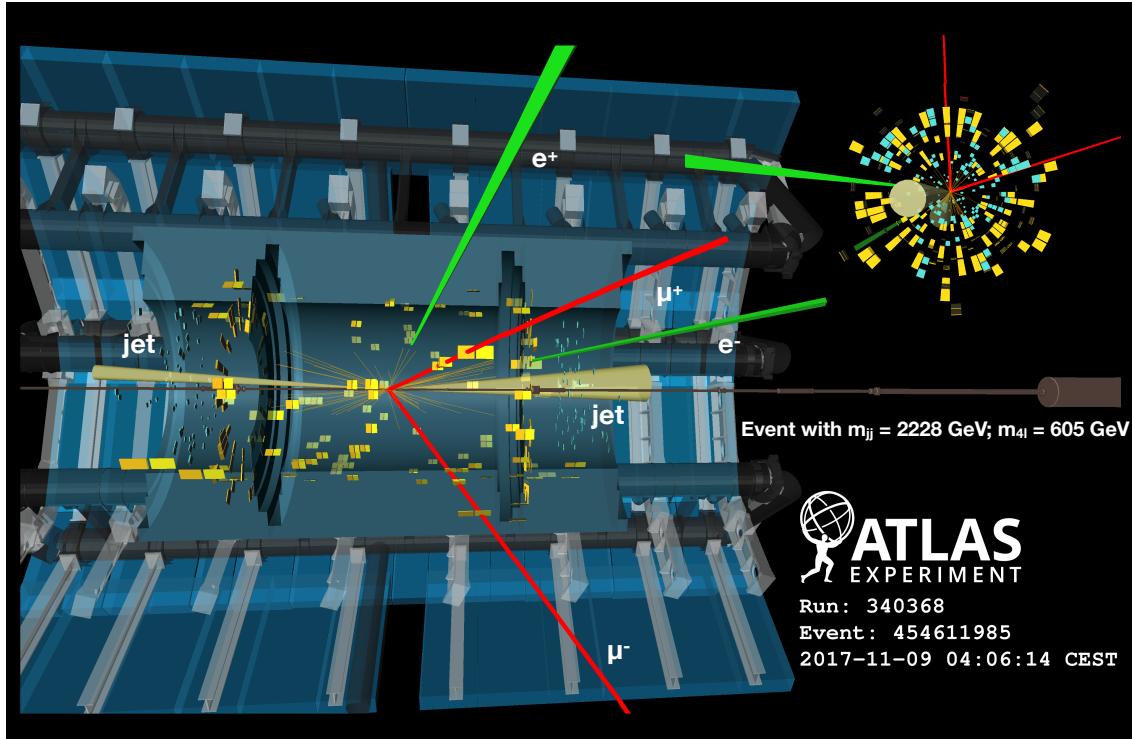


Fig. 6.13 Display of an event candidate of EW- $Z Z jj$ production in $2e2\mu$ channel in last MD bin ($0.875 < \text{MD} < 1.0$). The invariant mass of the di-jet (four-lepton) system is 2228 (605) GeV.

2252

2253 6.8 Prospect study of EW- $Z Z jj$ production in HL-LHC

2254 The High-Luminosity Large Hadron Collider (HL-LHC) project aims to increase the
 2255 luminosity by a factor of 10 beyond the LHC's design value to increase the potential for
 2256 discoveries after 2025. The expected luminosity will reach 3000 fb^{-1} with the centre-of-
 2257 mass energy of 14 TeV.

2258 As introduced in previous sections, with full run-2 data of 139 fb^{-1} collected by AT-
 2259 LAS detector at the LHC, the EW- $Z Z jj$ production is the last channel of observation for
 2260 VBS processes with massive bosons due to its very low cross section in $Z Z$ channel. So
 2261 we expect that this channel will benefit significantly from the increased luminosity at the
 2262 HL-LHC, and can be studied in great details for this known mechanism.

2263 In this section, a prospective study is performed for EW- $Z Z jj$ production at the HL-
 2264 LHC in the $\ell\ell\ell'\ell'$ channel^[105]. The study uses 3000 fb^{-1} of simulated pp collision data
 2265 at a centre-of-mass energy of 14 TeV as expected to be recorded by the ATLAS detector at
 2266 the HL-LHC. All simulated events are produced at particle-level, and the detector effects

2267 of leptons and jets reconstruction and identification are estimated by corrections assuming
2268 the mean number of interactions per bunch crossing ($\langle\mu\rangle$) of 200.

2269 6.8.1 The ATLAS detector at HL-LHC

2270 As the expectation of HL-LHC, the new Inner Tracker (ITk)^[106] will extend the track-
2271 ing acceptance capability of ATLAS detector to pseudorapidity ($|\eta|$) up to 4.0. By includ-
2272 ing a forward muon trigger, the upgraded Muon Spectrometer^[107] is also expected to
2273 provide muon identification capabilities to $|\eta|$ up to 4.0. In addition, the new high gran-
2274 ularity timing detector (HGTD)^[108] designed to mitigate the pile-up (PU) effects is also
2275 expected to be installed in the forward region of $2.4 < |\eta| < 4.0$. More details of ex-
2276 pected performance of the upgraded ATLAS detector at the HL-LHC has been reported
2277 in Ref.^[109].

2278 6.8.2 Simulation

2279 The analysis is performed using particle-level events. The samples are generated at
2280 $\sqrt{s} = 14$ TeV. The signal in this analysis is EW- $ZZjj$ process, while only the dominant
2281 irreducible background of QCD- $ZZjj$ is considered. Both signal and background are
2282 generated using Sherpa with the NNPDF3.0NNLO PDF set. The signal sample is mod-
2283 elled with two jets at Matrix Element (ME) level. The background is generated with up
2284 to one (three) outgoing partons at NLO (LO) in pQCD. As a quick study, other minor
2285 backgrounds such as fake backgrounds from $Z+jets$ and top-quark processes, as well as
2286 Diboson without 4l final-state and Triboson processes are not considered in this analysis.
2287 Furthermore, for hard scattering events, the pile-up collisions are set with a mean value
2288 of 200 interactions per bunch crossing. Signal and background yields are then scaled to
2289 an integrated luminosity of 3000 fb^{-1} as expected at the HL-LHC.

2290 6.8.3 Event selection

2291 The analysis selection follows closely to the one in ATLAS run-2 analysis as described
2292 in section 6.3. Here are some changes according to the expectation of the HL-LHC sce-
2293 nario for ATLAS detector:

- 2294 • Extend the lepton (both electron and muon) identification to $|\eta| < 4.0$
2295 • Pile-up (PU) jet suppression is applied with a PU rejection factor of 50 for all PU
2296 jets in the region of $|\eta| < 3.8$, based on the expected ATLAS detector performance
2297 at the HL-LHC.
2298 • The jets are required to have $p_T > 30$ (70) GeV in the $|\eta| < 3.8$ ($3.8 < |\eta| < 4.5$)

2299 region.

- 2300 • For two selected jets, tighten the m_{jj} requirement to be $m_{jj} > 600$ GeV, and require
2301 $\Delta\eta_{jj} > 2$.

2302 In addition, a fiducial volume, used to study the expected precision of the cross-section
2303 measurements, is defined at particle-level with the same kinematic requirements listed
2304 above.

2305 Table 6.12 summarized the number of selected signal and background events nor-
2306 malized to 3000 fb^{-1} . In addition to the *baseline* selection listed above, to compare the
2307 different detector scenarios at the HL-LHC, two alternative selections are also studied:

- 2308 1. Reduce the lepton η region to 2.7, to understand the effect due to forward lepton
2309 reconstruction and identification with the upgraded ATLAS detector.
2310 2. Only apply the PU jet suppression with region $|\eta| < 2.4$, to measure the improve-
2311 ment of *baseline* by extending the rejection range of PU jets at the HL-LHC with
2312 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 6.12 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} .**

2313 From this table, one can see the extended track coverage increases the $\ell\ell\ell'\ell'jj$ events
2314 by 15 to 30%, via improving the lepton efficiency. But the significance of searching for
2315 EW- $ZZjj$ process does not improve so much due to the large increment of background
2316 events.

2317 Figure 6.14 shows the kinematic distributions of di-jet invariant mass (m_{jj}), the ZZ
2318 invariant mass (m_{ZZ}) and the ϕ separation of two Z bosons ($|\Delta\phi(ZZ)|$) as well as the
2319 centrality of the ZZ system. The ZZ centrality is defined as:

$$ZZ \text{ centrality} = \frac{|y_{ZZ} - (y_{j1} + y_{j2})/2|}{|y_{j1} - y_{j2}|} \quad (6.8)$$

2320 To measure the event yield, the top panel shows the stack distribution for EW- and QCD-
2321 $ZZjj$ processes, while bottom panel is the ratio between two processes.

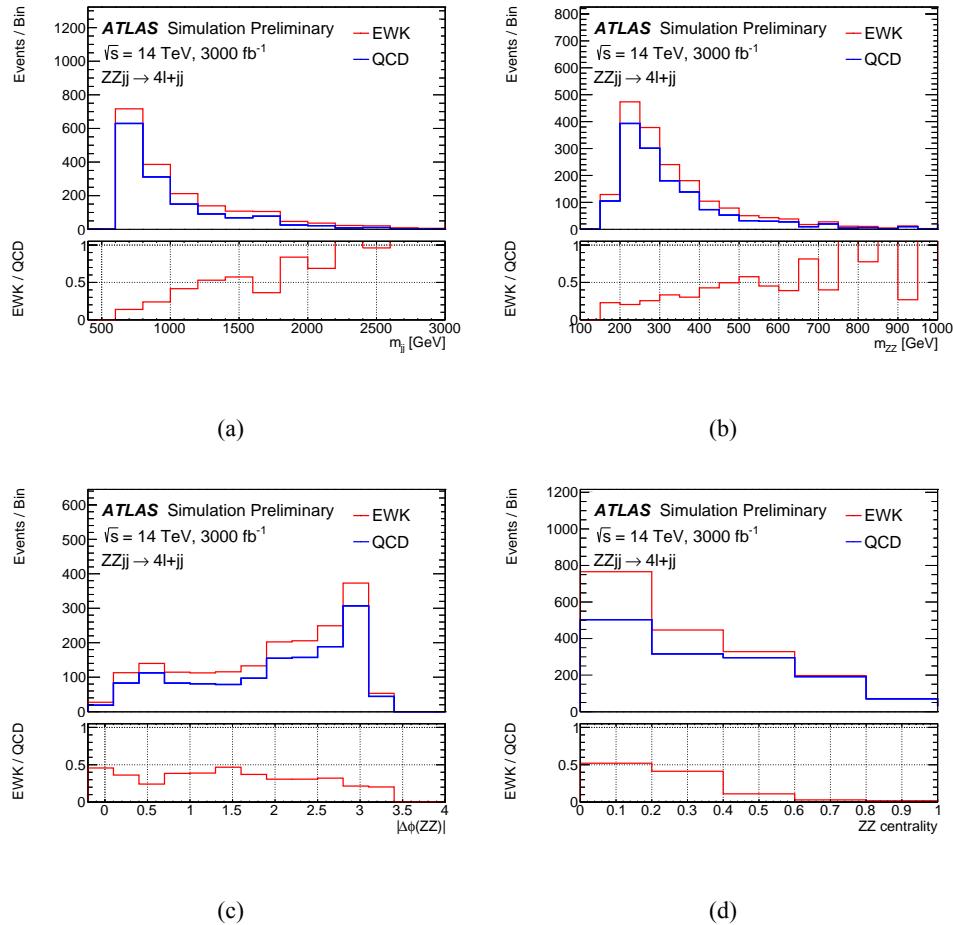


Fig. 6.14 Detector-level distributions of EW- and QCD- $Z Z j j$ processes with selected events in defined phase space at 14 TeV of (a) m_{jj} , (b) m_{ZZ} , (c) $|\Delta\phi(ZZ)|$, (d) ZZ centrality, normalized to 3000 fb^{-1} .

6.8.4 Systematics

According to studies in section 6.5, the dominant systematic in $\ell\ell\ell'\ell'$ channel is from theoretical systematic for QCD- $Z Z j j$ background process. Different sizes of systematics have been studied, at a factor of 5, 10 and 30% on background modelling. The 5% uncertainty is an optimal estimation when there is enough data events from QCD-enriched control region at the HL-LHC that can be used to constrain the theoretical normalization on QCD- $Z Z j j$ process. The 30% one is a conservative estimation, in which the uncertainties are directly calculated from different PDF sets and QCD renormalization and factorization scales, following recommendation from the PDF4LHC mentioned in section 6.5.

For experimental sources, the jet systematics have been checked following the setting provided by the HL-LHC in Ref.^[109], and the uncertainties are within 5% level, which is smaller than run-2 measurement at 10%. Figure 6.15 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.

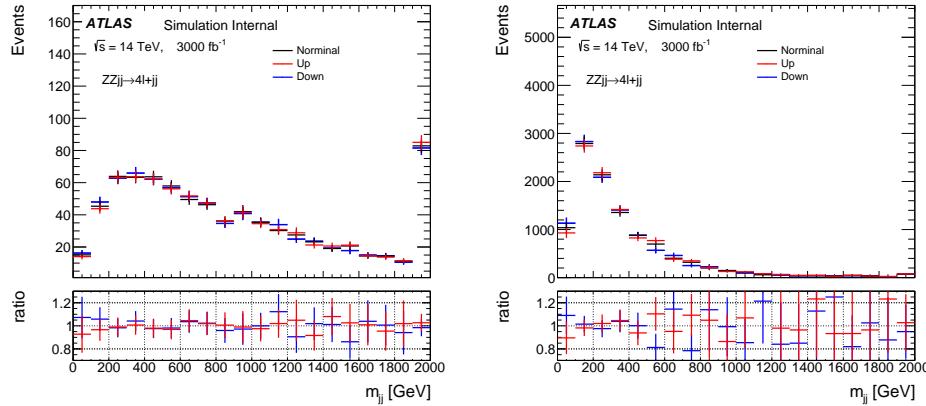


Fig. 6.15 Jet variations on m_{jj} distribution for EW- $Z Z jj$ (left) and QCD- $Z Z jj$ (right) processes with luminosity of 3000 fb^{-1} at 14 TeV. Upgrade Performance Function is used to extract the uncertainties with *baseline* setting.

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 5%, 10% and 30% levels respectively.
- Three different sources of uncertainties are treated as uncorrelated and summed up quadratically.

6.8.5 Results

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

$$\text{Significance} = \frac{S}{\sqrt{\sigma(B)_{\text{stat.}}^2 + \sigma(B)_{\text{syst.}}^2}}, \quad (6.9)$$

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

Based on baseline selection of $m_{jj} > 600 \text{ GeV}$, an additional scan over different m_{jj} cuts is also performed with a step of 50 GeV under different systematic conditions, as

shown in figure 6.16.

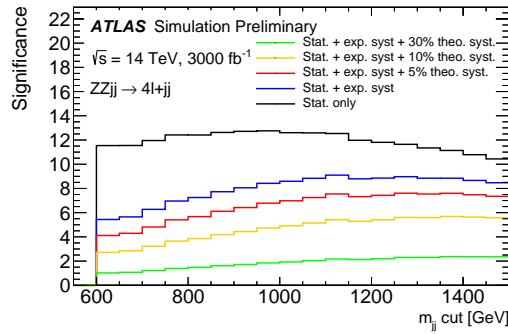


Fig. 6.16 The expected significance of EW- $Z Z j j$ processes as a function of different m_{jj} cut with 3000 fb^{-1} , under conditions of different sizes of theoretical uncertainties on the QCD- $Z Z j j$ background modelling. The statistical uncertainty is estimated from expected data yield at 14 TeV with 3000 fb^{-1} . Different uncertainties are summed up quadratically.

2355

2356 In addition, the expected differential cross section of EW- $Z Z j j$ process is measured
2357 in the defined phase space at 14 TeV, as a function of m_{ZZ} and m_{jj} , shown in figure 6.17.
The expected differential cross sections are calculated as:

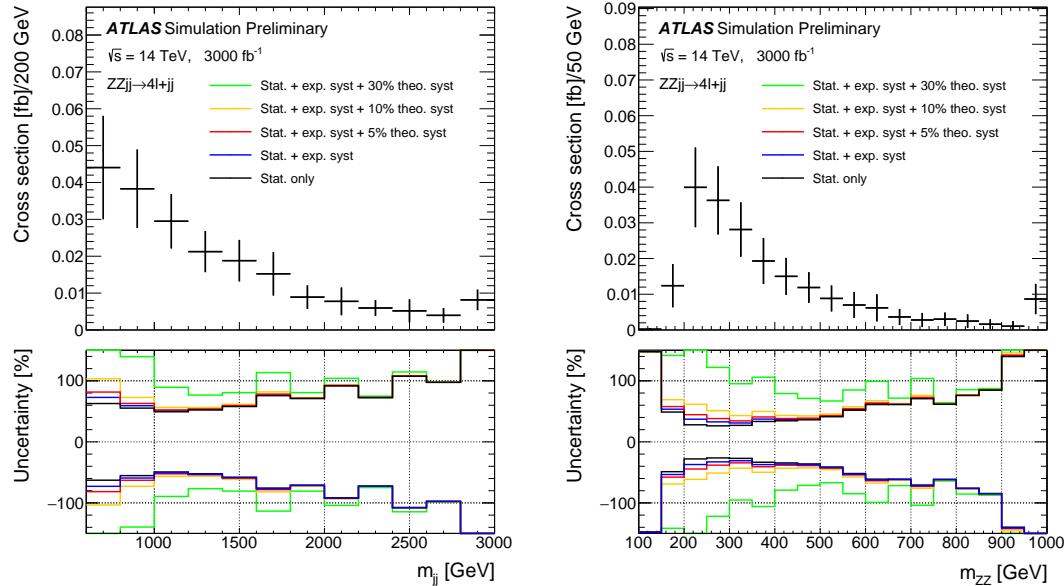


Fig. 6.17 The projected differential cross-sections at 14 TeV for the EW- $Z Z j j$ processes as a function of m_{jj} (left) and m_{ZZ} (right). The top panel shows measurement with statistical only case, where statistical uncertainty is estimated from expected data yield at 14 TeV with 3000 fb^{-1} . The bottom panel shows impact of different sizes of systematic uncertainties.

2358

$$\sigma = \frac{N_{pseudo-data} - N_{QCD-ZZjj}^{det.}}{L * C_{EW-ZZjj}} \quad (6.10)$$

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

2359 where $N_{pseudo-data}$ denotes the expected number of data events with 3000 fb^{-1} luminosity
 2360 at 14 TeV, and $N_{QCD-ZZjj}$ and $N_{EW-ZZjj}$ are the number of predicted events of QCD-
 2361 $ZZjj$ and EW- $ZZjj$ processes, in particle-level (part.) or detector-level (det.). The
 2362 $C_{EW-ZZjj}$ factor represents the detector efficiency for EW- $ZZjj$ processes introduced
 2363 in section 6.6.1. The interference between EW- and QCD- $ZZjj$ processes is ignored
 2364 due to its minor contribution.

2365 The value of expected integrated cross section as well as its uncertainty under different
 2366 systematic conditions are shown in table 6.13 with 3000 fb^{-1} luminosity at 14 TeV. The
 2367 statistical uncertainty is at 10% level when with such large luminosity. The result is dom-
 2368 inated by systematics and can reach 100% level when theoretical modelling uncertainty
 is 30% for QCD- $ZZjj$ processes.

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

Table 6.13 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.

2369

2370 6.9 Conclusion

2371 The fiducial cross section for inclusive $Z Z jj$ production is measured in this section,
2372 with a total relative uncertainty of 11% for the $\ell\ell\ell'\ell'$ final state, and found to be com-
2373 patible with the SM prediction. The observation of electroweak production of two jets in
2374 association with a Z -boson pair decay to $\ell\ell\ell'\ell'$ final state using 139 fb^{-1} of 13 TeV pp
2375 collision data collected by ATLAS experiment at the LHC is presented in this section.
2376 The search for electroweak production of two jets in association with a Z -boson pair is
2377 based on multivariate discriminants (MD) to enhance the separation between the signal
2378 and backgrounds. In $\ell\ell\ell'\ell'$ final state, the background-only hypothesis is rejected with
2379 an observed (expected) significance of 5.5 (3.9) σ , which gives the first observation of
2380 electroweak production in $Z Z jj$ channel.

2381 In addition, the prospective study for the EW- $Z Z jj$ production at the HL-LHC in
2382 the $\ell\ell\ell'\ell'$ channel, using 3000 fb^{-1} simulated pp collision data at a centre-of-mass
2383 energy of 14 TeV has been presented. The precision of the expected measurements of the
2384 integrated and differential cross sections as a function of dijet or $\ell\ell\ell'\ell'$ invariant mass are
2385 shown. Under the assumption of theoretical uncertainty for the QCD- $Z Z jj$ processes and
2386 experimental uncertainty for jets being constraint at 5% level respectively, with statistical
2387 uncertainty in 3000 fb^{-1} being considered, the observation of the EW- $Z Z jj$ process can
2388 reach a significance of 7σ .

2389 **Chapter 7 Search for heavy resonances decaying into**
2390 **a pair of Z bosons in $\ell\ell\ell'\ell'$ final state using pp**
2391 **collision data collected by ATLAS detector from 2015 to**
2392 **2018**

2393 **7.1 Introduction**

2394 A new particle was discovered by the ATLAS and CMS Collaborations at the LHC^[6-7]
2395 in 2012. Both experiments have confirmed that the properties including spin, couplings
2396 and parity of this new particle are consistent with Higgs boson predicted in the Stan-
2397 dard Model (SM), which is an important milestone in understanding of the mechanism
2398 of EWSB. Nevertheless, the possibility that this newly discovered particle is just a part
2399 of the extended Higgs sector as predicted by various extensions in the SM cannot be
2400 ruled out. Many models predicted the existence of new heavy resonances decaying into
2401 dibosons, such as a heavy spin-0 neutral Higgs boson^[110] and the two-Higgs-doublet
2402 models (2HDM)^[9], as well as the spin-2 Kaluza–Klein (KK) excitations of the graviton
2403 (G_{KK})^[111].

2404 Though with smaller branching ratio compared to semileptonic or fully hadronic de-
2405 cay channels, the $\ell\ell\ell'\ell'$ final state has its unique sensitivity in mass range smaller than
2406 1 TeV region due to its good mass resolution and relative smaller experimental and the-
2407 oretical systematics. This section presents the search for heavy resonance decaying into
2408 a pair of Z bosons to the $\ell\ell\ell'\ell'$ final state, in which ℓ denotes to either an electron or a
2409 muon^[112-113]. Several signal hypotheses are considered. The first hypothesis is a heavy
2410 Higgs boson (spin-0 resonance) under the narrow-width approximation (NWA). Then as
2411 several theoretical models prefer non-negligible natural widths, the models under large-
2412 width approximation (LWA), assuming widths of 1%, 5%, 10% and 15% of the resonance
2413 mass, are also studied. In addition, the graviton excitations (spin-2 resonance) under the
2414 Randall–Sundrum model are also searched. It is assumed that the heavy resonance is
2415 produced predominantly via the gluon-gluon Fusion (ggF) and the Vector Boson Fusion
2416 (VBF) productions, but with the unknown ratio of two production rates. So the results are
2417 separated for ggF and VBF production modes. To gain more sensitivity, the $\ell\ell\ell'\ell'$ events
2418 are classified into ggF- and VBF-enriched categories. Moreover, for the NWA model, the
2419 categorizations are studied under both cut-based and multivariate (MVA) -based methods,
2420 the details of categorization are shown in following sections.

2421 The search uses the four-lepton invariant mass in the range of 200 GeV to 2000 GeV for
 2422 signal hypothesis of spin-0 resonance under the NWA model, and from 400 GeV to
 2423 2000 GeV for the one under the LWA models. And the spin-2 graviton signals are searched
 2424 in the mass range from 600 GeV to 2000 GeV. The data collected by ATLAS detector at
 2425 the LHC from 2015 to 2018 at the centre-of-mass energy of 13 TeV is used. In case of no
 2426 excess, upper limits on the production rate of different signal hypotheses are computed
 2427 from statistical fits to m_{4l} distribution.

2428 **7.2 Data and MC samples**

2429 **7.2.1 Data samples**

2430 The data used in this analysis are collected by ATLAS detector at the centre-of-mass
 2431 energy of 13 TeV during the years of 2015 to 2018. Only events passing the latest Good
 2432 Run List (GRL) released by the Data Quality group from ATLAS experiment correspond-
 2433 ing to an integrated luminosity of $139.0 \pm 2.4 \text{ fb}^{-1}$ are used. Table 7.1 listed the recorded
 integrated luminosity, average and peak pile-up of each year's data.

Table 7.1 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

2434

2435 **7.2.2 Background MC simulations**

2436 Background processes considered in this analysis include ZZ ($q\bar{q} \rightarrow ZZ$, $gg \rightarrow ZZ$),
 2437 triboson (WWZ , WZZ , ZZZ), $Z+jets$ and top-quark ($t\bar{t}$, ttV) processes.

2438 The QCD $q\bar{q} \rightarrow ZZ$ process is modelled using Sherpa 2.2.2^[89] with the
 2439 NNPDF3.0NNLO^[90] PDF, where events with up to one (three) outgoing partons are gen-
 2440 erated at NLO (LO) in pQCD. The production of ZZ from the gluon-gluon initial state
 2441 with a four-fermion loop or with an exchange of the Higgs boson, which has an order
 2442 of α_S^4 in QCD, is not included in this Sherpa simulation. So a separate gg induced ZZ
 2443 sample including the continuum background, the SM Higgs boson, and the interference
 2444 contribution is modelled using Sherpa 2.2.2 with the NNPDF3.0NNLO PDF set, and with
 2445 an additional k-factor^[91] of 1.7 applied. The EW- $ZZjj$ production is simulated using

2446 Sherpa 2.2.2 with the NNPDF3.0NNLO PDF, and the $ZZZ \rightarrow \ell\ell\ell'\ell'qq$ process is also
2447 taken into account in this sample.

2448 The $Z+jets$ events are generated using Sherpa 2.2.2 with the NNPDF3.0NNLO PDF,
2449 in which the ME is calculated for up to two partons with NLO accuracy in pQCD and up to
2450 four partons with LO accuracy. The $Z+jets$ events are normalized using the next-to-next-
2451 to-leading-order (NNLO) cross section. The triboson processes with full leptonic decays
2452 and at least four prompt charged leptons are generated using Sherpa 2.1.1. For top-quark
2453 pair ($t\bar{t}$) production and the single top-quark productions in t -channel, s -channel and Wt -
2454 channel, Powheg-Box v2 is used with the CT10 PDF. The production of $t\bar{t}$ in association
2455 with Z boson(s) ($t\bar{t}Z$) is modelled with MadGraph5_aMC@NLO.

2456

7.2.3 Signal MC simulations

2457 One model considered in this analysis is heavy spin-0 resonance under the Narrow
2458 Width Approximation (NWA) simulated using Powheg-Box v2 MC event generator with
2459 the CT10 PDF. The gluon-gluon fusion (ggF) production mode and vector-boson fusion
2460 (VBF) production mode are calculated separately with matrix elements up to NLO in
2461 QCD. The Powheg-Box is interfaced to Pythia8 for parton showering, and for decaying
2462 the Higgs boson into the $H \rightarrow ZZ \rightarrow \ell\ell\ell'\ell'$ final states. Events of NWA signal are
2463 generated at mass points between 200 GeV to 2000 GeV using the step of 100 (200) GeV up
2464 to (above) 1 TeV in both ggF and VBF production modes.

2465 In addition, heavy Higgs boson events under the Large Width Approximation (LWA)
2466 with widths of 1%, 5%, 10% and 15% of the boson mass are generated using Mad-
2467 Graph5_aMC@NLO 2.3.2 interfaced to Pythia8. Only ggF production is consid-
2468 ered. Mass points between 400 GeV to 2000 GeV are simulated with the step of 100
2469 (200) GeV up to (above) 1 TeV. To describe jet multiplicity, MadGraph5_aMC@NLO is
2470 used to simulated process of $pp \rightarrow H + \geq 2\text{jets}$ at NLO in QCD with the FxFx merging
2471 scheme^[114].

2472 Spin-2 Kaluza–Klein (KK) gravitons (G_{KK}) from the Bulk Randall–Sundrum
2473 model^[115] are also studied in this analysis. Events are generated by Mad-
2474 Graph5_aMC@NLO at LO in QCD, which is then interfaced to Pythia8 for parton show-
2475 ering. The G_{KK} -gluon coupling $k/\overline{M}_{\text{Planck}}$, where k is the curvature scale of the extra
2476 dimension and $\overline{M}_{\text{Planck}}$ is the reduced Planck mass, is set to 1. The width of the resonance
2477 is correlated with the coupling $k/\overline{M}_{\text{Planck}}$ and in this configuration is around $\square 6\%$ of its
2478 mass. The mass of the G_{KK} is the only free parameter in this simplified model. Mass
2479 points between 600 GeV to 2 TeV with 200 GeV spacing were generated.

2480 7.3 Object and event selections

2481 7.3.1 Object selections

2482 Similar to VBSZZ analysis in section 6.3, the selection of this analysis relies on the
2483 definition of multiple objects: *electrons*, *Muons*, and *jets*. Details of definitions for each
2484 object are described as below:

2485 **Electron:** As described in section 4.2.3, electrons are reconstructed from energy de-
2486 posits in the EM calorimeter matched to a track in the inner detector. The electron candi-
2487 dates satisfying the *Loose* criterion from the likelihood-based (LH) method are selected,
2488 with a selection efficiency ranging from 90% for transverse momentum $p_T = 20$ GeV
2489 to 96% for $p_T > 60$ GeV. In addition, the electrons are required to have $p_T > 7$ GeV,
2490 $|\eta| < 2.47$ and $|z_0 \sin\theta| < 0.5$ mm.

2491 **Muon:** To increase the acceptance range in reco-level for $\ell\ell\ell'\ell'$ channel, all four
2492 types of muons (CB, ST, CT, ME muons, described in section 4.2.4) are used. The CT
2493 muons are required to pass $p_T > 15$ GeV and $|\eta| < 0.1$, while the ST muons are also
2494 limited in $|\eta| < 0.1$ region. The ME muons are only used in the region of $2.5 < |\eta| < 2.7$.
2495 And at most one CT, ST or ME muon is allowed in one $\ell\ell\ell'\ell'$ quadruplet. The Muon
2496 candidates are required to pass $p_T > 5$ GeV and $|\eta| < 2.7$, and satisfy the *Loose* identifi-
2497 cation criterion with an efficiency of at least 98.5%. The impact parameter requirements
2498 of $|d_0| < 1$ mm and $|z_0 \sin\theta| < 0.5$ mm are further applied.

2499 **Jets:** Jets are clustered using the anti- k_t algorithm with radius parameter $R = 0.4$ im-
2500 plemented in the FastJet package as described in section 4.2.5. The ‘particle flow’ (PFlow)
2501 objects^[116], which combines measurements from both the tracker and the calorimeter, are
2502 used as inputs to the FastJet package. The algorithm removes the energy deposited in
2503 the calorimeter from charged hadrons that are considered in jet reconstruction, and makes
2504 use of an ensemble of “PFlow” objects including the remaining calorimeter energy and
2505 tracks matched to the hard interaction. The accuracy of charged-particle measurements
2506 can be improved from tracker while retaining the energy measurements of neutral-particle
2507 in calorimeter. Compared to only using topological clusters, jets reconstructed with the
2508 particle flow algorithm with $p_T > 30$ GeV have approximately 10% better transverse mo-
2509 mentum resolution. The jets used in this analysis are then required to have $p_T > 30$ GeV
2510 and $|\eta| < 4.5$. To further reduce the effects of pile-up jets, a jet vertex tagger (JVT) is
2511 applied to jets with $p_T < 60$ GeV and $|\eta| < 2.4$.

2512 **Overlap removal:** As the selected jet and lepton candidates can be reconstructed
2513 from same detector information, an overlap-removal procedure is applied. For electron

and muon sharing the same ID track, the electron is selected in the case that the muon is calorimeter-tagged and does not have a MS track, or is a segment-tagged muon, otherwise the muon is selected. The jet overlapping with electron (muon) within a cone of size of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.2(0.1)$ are removed.

7.3.2 Event selections

First of all, the four-lepton events are required to pass single or multi-lepton triggers. Due to the increasing of peak luminosity and pile-up, the p_T and E_T thresholds of triggers increase slightly during the data-taking periods from 2015 to 2018. Table 7.2 summarizes the triggers used for $\ell\ell\ell'\ell'$ channel. The overall trigger efficiency for selected signal events passing final selection is around 98%.

Table 7.2 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	$\mu20; \mu50; \mu60$	$\mu24; \mu26; \mu40; \mu50$	$\mu26; \mu50; \mu60$	$\mu26; \mu50; \mu60$
single electron	$e24; e60; e120$	$e26; e60; e140; e300$	$e26; e60; e140; e300$	$e26; e60; e140; e300$
dimuon	$2\mu10; \mu18_{-\mu8}$	$2\mu10; 2\mu14; \mu22_{-\mu8}$	$2\mu14; \mu22_{-\mu8}$	$2\mu14; \mu22_{-\mu8}$
dielectron	$2e12$	$2e15; 2e17$	$2e17; 2e24$	$2e17; 2e24$
electron-muon	$e24_{-\mu8}$	$e24_{-\mu8}; e26_{-\mu8}$ $e17_{-\mu14}; e7_{-\mu24}; 2e12_{-\mu10}; e12_{-\mu10}$	$e26_{-\mu8}$	$e26_{-\mu8}$
trimuon	$\mu18_{-2\mu4}$	$\mu11_{-2\mu4}; \mu6_{-2\mu4}; \mu20_{-2\mu4}; 3\mu4$ $3\mu6$	$4\mu4; \mu20_{-2\mu4}; 3\mu4$	$\mu20_{-2\mu4}$
trilepton	$e17_{-2e9}$	$e17_{-2e9}; e17_{-2e10}$	$e24_{-2e12}$	$e24_{-2e12}$

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

The two lepton pairs in quadruplet are required to have angular separation with $\Delta R > 0.1$. To suppress the contribution from $J/\psi \rightarrow \ell\ell$ decays, for 4μ and $4e$ quadruplets, the events are rejected if any OSSF lepton pair is found with mass below 5 GeV. If there are more than one quadruplets from different channels in event at this point, the one with highest expected signal rate is selected in the order of $4\mu, 2e2\mu, 4e$. The transverse impact-parameter significance ($|d_0|/\sigma_{d_0}$) for muons (electrons) is then required to be smaller than

2537 3 (5) to suppress the backgrounds from heavy-flavour hadrons.

2538 In addition, the track- and calorimeter- based isolation criteria is required for all elec-
 2539 trons and muons to further suppress the reducible backgrounds of $Z+jets$ and $t\bar{t}$. For
 2540 lepton isolation selection, the two track- and calorimeter- based variables, $E_T^{topocone}$ and
 2541 $p_T^{varcone}$ as described in section 4.2.4 (section 4.2.3) for muons (electrons), are vulnerable
 2542 to pileup. For track-based variable, this is because of additional tracks in the event. The
 2543 definition of $p_T^{varcone}$ attempts to limit the tracks used in the calculation to those from the
 2544 vertex via a loose cut of $|z_0 \sin(\theta)| < 3$, which proved to be too loose in new pile-up
 2545 regime in 2017 and 2018 datasets. So new track-based variable is used, by adding a re-
 2546 quirement that the track be used in determining the vertex, or that, if not, it both passes the
 2547 cut on $|z_0 \sin(\theta)|$ and is not used in determining any other vertex, which makes the track-
 2548 based variable to be more isolation-robust in the high pile-up regime. The new variable is
 2549 named as $ptvarcone[\text{cone}]_{_TightTTVA_pt[p_T \text{ cut}]}$, where [cone] is the cone size and [p_T
 2550 cut] is the cutoff for including tracks in the calculation.

2551 For calorimeter-based variable, the calculation of $E_T^{topocone}$ corrects the pile-up effects
 2552 by subtracting an average pileup contribution computed over the whole detector. But with
 2553 the increasing of energy density of pile-up events, the root mean square (RMS) of $E_T^{topocone}$
 2554 variable increases, which leads to the increment of possibility that the pile-up fluctuations
 2555 are not be accounted for correctly. One possible solution is using particle-flow (PFlow)
 2556 method to calculate the calorimeter isolation. As part of PFlow reconstruction process, it
 2557 assigns the clusters to tracks which improves the track-cluster association for better de-
 2558 termination of the raw value of the E_T in the cone. And using PFlow jets to calculate the
 2559 pileup correction provides a further improvement. So a resulting variable named $neflow-$
 2560 $isol[\text{cone}]$ is used. Finally, a requirement of isolation, called *FixedCutPFlowLoose*, which
 2561 gives better performance in high pile-up condition is applied to electrons and muons as:
 2562 $(\max(ptcone20_{_TightTTVA_pt500}, ptvarcone30_{_TightTTVA_pt500}) + 0.4 \times neflow-$
 2563 $isol20) / p_T < 0.16$

2564 On the top of impact parameter cut and lepton isolation cut, the four-lepton candidates
 2565 are also required to originate from a common vertex to reduce $Z+jets$ and $t\bar{t}$ backgrounds.
 2566 This is ensured by applying a vertex fit χ^2 cut of 4 ID tracks of lepton candidates satisfying
 2567 $\chi^2/N_{dof} < 6$ (9) for events in 4μ ($4e$ and $2e2\mu$) channel(s).

2568 To improve the mass resolution, the QED process of final state radiation (FSR) pho-
 2569 tons in Z boson decays are taken into account in the reconstruction of Z bosons. The
 2570 four-momentum of any reconstructed photon that is consistent with having been radiated
 2571 from lepton(s) in leading pair are added into final state. Moreover, the four-momenta of

leptons in both (leading and sub-leading) pairs are recomputed by performing a Z -mass-constrained kinematic fit, which uses a Breit Wigner Z boson line-shape and Gaussian function with width set to the expected lepton resolution per lepton to model the momentum response function. The Z -mass-constrained mass improves the $m_{4\ell}$ resolution by up to 15% depending on m_H .

In summary, table 7.3 lists a comprehensive object and event level selection as described above.

7.3.3 Studies of τ -channel in particle-level

In addition to selecting the four-lepton events with ‘lepton’ only including electron and muon ($4e$, $2e2\mu$ and 4μ channels), a serious studies of considering $2e2\tau$ and $2\mu2\tau$ channels in final states have been done in particle-level. Without considering fake background and the rather complicated decay results of τ -lepton, the study in particle-level is a good start point.

The object selections of electron and muon and the event level selections for leptons (now including τ) are still keep the same as described in previous two sections. For the selection of τ -lepton: it must satisfy $p_T > 20$ GeV and be measured in the pseudorapidity range of $|\eta| < 2.5$. And to minic the reconstruction performance, for all leptons, a flat reconstruction efficiency is applied in object selection of electron, muon and τ -lepton. A 85% (95%) efficiency is applied to electron (muon), while to hadronic decay τ , the reconstruction efficiency of 50% times its hadronic decay rate is signed. And for its leptonic decay, the efficiency is the same as the one to electron and muon. In summary, the overall reconstruction efficiency to τ -lepton is about 64%.

Considering only the visible decay product for the reconstruction of τ -lepton

Figure 7.1 shows the invariant mass distribution of Z bosons reconstructed from two τ s using three different kinds of τ -lepton:

- The τ -lepton is reconstructed from both visible decay product (eg. electrons, muons, hardons) and the invisible neutrinos (can be extracted in particle-level from decay chain).
- In addition to the previous case, the energy of photon from decay product is also added, which is called ‘dressing’ in particle-level.
- Only consider the visible part from τ -lepton decay, as in reconstruction-level, the decay product of neutrinos are hard to be reconstructed.

Clear energy loss is seen for the thrid case with only visible τ decay product in Figure 7.1.

Table 7.3 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	
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-loose</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	<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

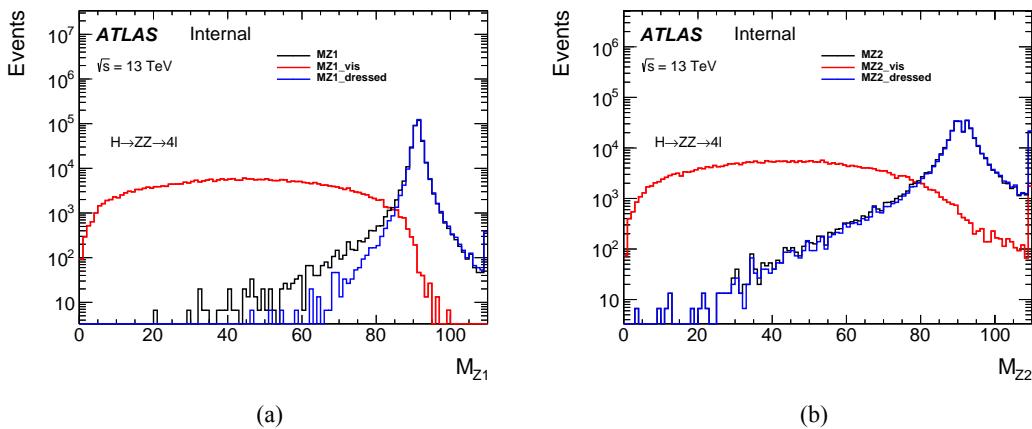


Fig. 7.1 $Z(\rightarrow\tau\tau)$ boson invariant mass distribution for signal at 300 GeV of (a) M_{Z_1} : the leading pairs; (b) M_{Z_2} : the sub-leading pairs. The definition of leading and sub-leading pairs can be found in section 7.3.2.

2606 To check the performance in reconstruction-level, only the visible decay product is used
 2607 for the reconstruction of τ in the following studies. Then the expected number of events
 2608 for signal and $q\bar{q} \rightarrow ZZ$ background in different lepton flavor channels are shown in
 2609 table 7.4 by normalizing to the luminosity of 80 fb^{-1} . Three different kind of final states
 2610 are considered:

- Inclusive: channels including $2e2\mu$, $4e$, 4μ and $2e2\tau$, $2\mu2\tau$.
- Only $e\mu$: $2e2\mu$, $4e$, 4μ channels.
- With τ : $2e2\tau$, $2\mu2\tau$ channels.

2614 The simulated narrow-width signal samples at the m_H of 300, 600, 1000 and 1200 GeV are
 2615 studied.

2616 Figure 7.2 shows the significance scan as a function of $m_{4\ell}$ at four mass points: 300,
 2617 600, 1000, 1200 GeV, for events in lepton flavor channel with or without τ -lepton. A pro-
 2618 file likelihood function is used for scan, and the background only considers the contribution
 2619 from $q\bar{q} \rightarrow ZZ$.

2620 Using Missing Mass Calculator tool for the reconstruction of τ -lepton

2621 As using only visible decay product of τ -lepton leads to very large energy loss in
 2622 reconstruction, an *Missing Mass Calculator* (MMC) tool^[117] is applied to extract the in-
 2623 visible part of τ decay products. The MMC method is a technique that can significantly
 2624 improve the neutrino momentum resolutions and di- τ invariant mass, by considering a
 2625 complete reconstruction of event kinematics in two τ -lepton final state.

2626 Figure 7.3 shows the mass of Z bosons reconstructed from two τ s again by adding the
 2627 case that di- τ mass calculated from MMC. Better mass reconstruction can be observed
 2628 than using the only visible decay products.

Table 7.4 The expected number of events for signal and $q\bar{q} \rightarrow ZZ$ background in the lepton flavor channels with or without τ -lepton. The event yields are normalized to the luminosity of 80 fb^{-1} .

Mass window	Channels	signal		$q\bar{q} \rightarrow ZZ$	
		Events	Fraction	Events	Fraction
[250GeV, 350GeV]	Inclusive	3423.5	100%	539.1	100%
	Only e μ	3148.2	92.0%	493.8	91.6%
	With τ	275.4	8.0%	42.7	7.9%
[550GeV, 650GeV]	Inclusive	1312.1	100%	26.0	100%
	Only e μ	1269.9	96.8%	24.1	92.6%
	With τ	42.2	3.2%	1.7	6.5%
[950GeV, 1050GeV]	Inclusive	526.9	100%	1.20	100%
	Only e μ	529.9	98.7%	1.18	98.5%
	With τ	7.0	1.3%	0.12	10.4%
[1150GeV, 1350GeV]	Inclusive	419.1	100%	1.09	100%
	Only e μ	414.9	99.0%	1.07	98.7%
	With τ	4.1	1.0%	0	0%

Figure 7.4 shows the $m_{4\ell}$ distribution after event selections for events in $2e2\mu$, $4e$, and 4μ channels, and those in $2e2\tau$ and $2\mu2\tau$ channels. The simulated narrow-width signal samples at the m_H of 300 GeV and 1000 GeV, as well as the $q\bar{q} \rightarrow ZZ$ background are studied.

After using MMC, the expected number of events for signal and $q\bar{q} \rightarrow ZZ$ background in different lepton flavor channels are shown in table 7.5 by normalizing to the luminosity of 80 fb^{-1} .

Figure 7.5 shows the significance scan as a function of $m_{4\ell}$ at two mass points: 300 and 1000 GeV, for events in lepton flavor channel with or without τ -lepton. A profile likelihood function is used for scan, and the background only considers the contribution from $q\bar{q} \rightarrow ZZ$.

From the studies above in particle-level, the Missing Mass Calculator tool improves the mass reconstruction of di- τ pair significantly, and provides a possibility to include τ -lepton into four-lepton final state analyses in the future. But the fraction of events reconstructed with τ -lepton in background is still larger than the one in signal in higher mass region equal to or greater than 1000 GeV as shown in table 7.4 and 7.5 even in particle-level. And as the significance scan shown in figure 7.5, in the two $m_{4\ell}$ bins on peak, the significance with only τ -channel events is worse than other cases, which will

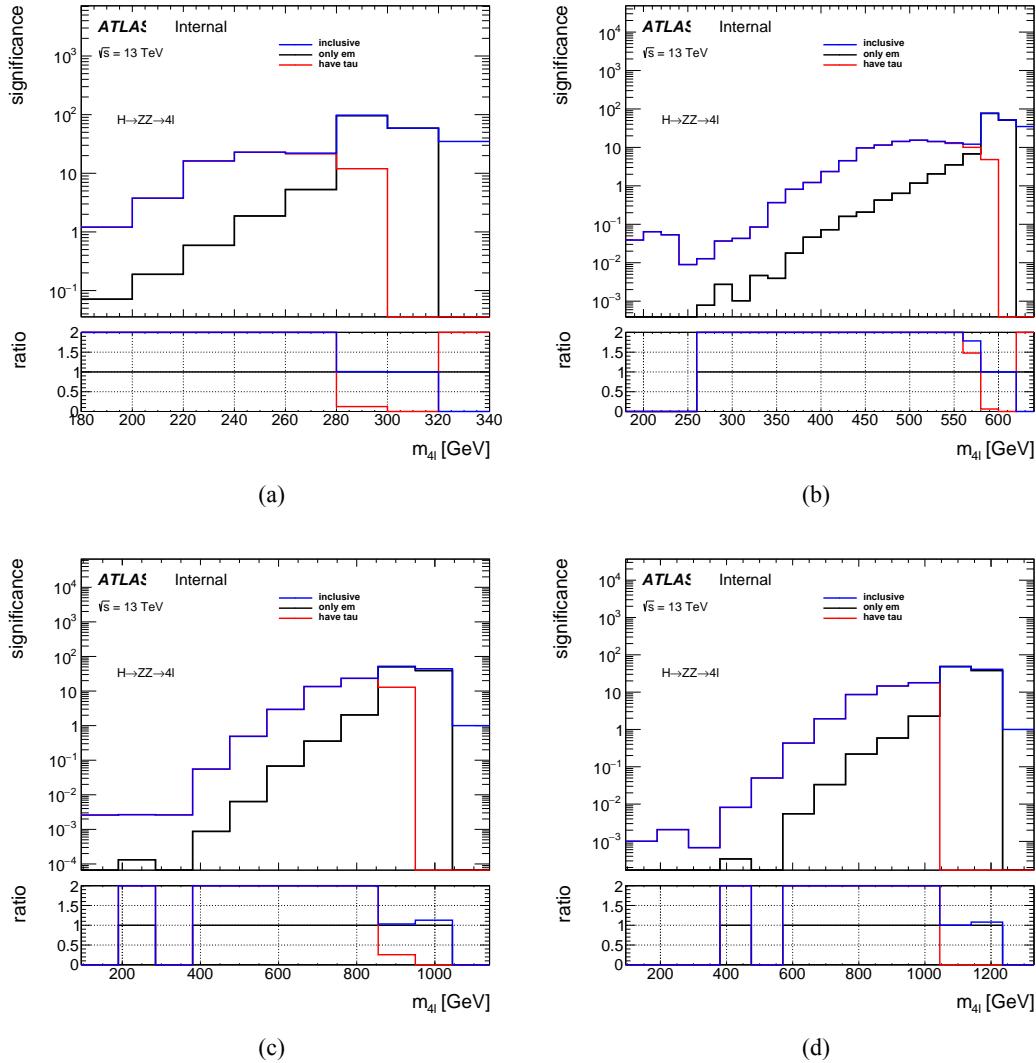


Fig. 7.2 Significance scan at the m_H of (a) 300 GeV; (b) 600 GeV; (c) 1000 GeV; (d) 1200 GeV, for events in lepton flavor channel with or without τ -lepton.

Table 7.5 After using MMC, the expected number of events for signal and $q\bar{q} \rightarrow ZZ$ background in the lepton flavor channels with or without τ -lepton. The event yields are normalized to the luminosity of 80 fb^{-1} .

Mass window	Channels	signal		$q\bar{q} \rightarrow ZZ$	
		Events	Fraction	Events	Fraction
[250GeV, 350GeV]	Inclusive	4517.2	100%	685.6	100%
	Only e\mu	3148.2	69.7%	500.3	73.0%
	With τ	1369.0	30.3%	185.0	27.0%
[950GeV, 1050GeV]	Inclusive	785.5	100%	1.69	100%
	Only e\mu	529.9	67.5%	1.26	74.4%
	With τ	255.6	32.5%	0.59	35.1%

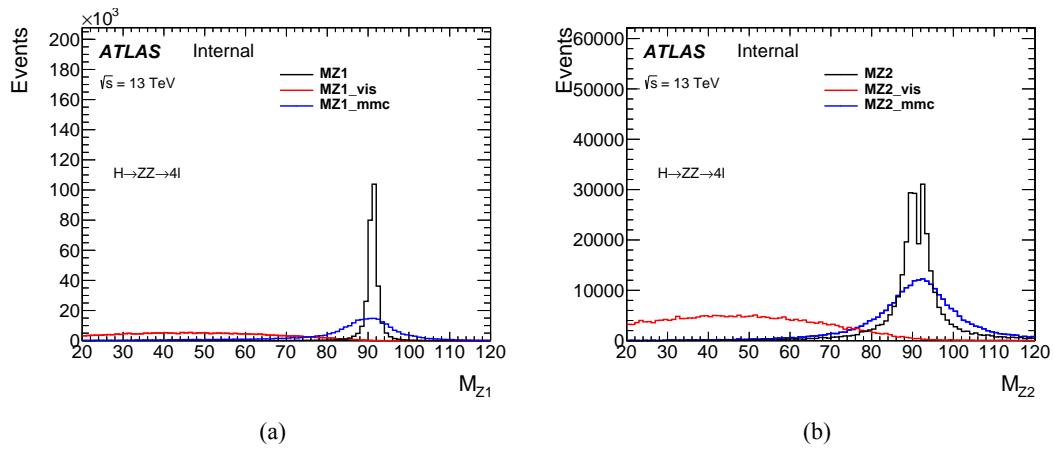


Fig. 7.3 $Z(\rightarrow\tau\tau)$ boson invariant mass distribution for signal at 300 GeV of (a) M_{Z_1} : the leading pairs; (b) M_{Z_2} : the sub-leading pairs. The definition of leading and sub-leading pairs can be found in section 7.3.2.

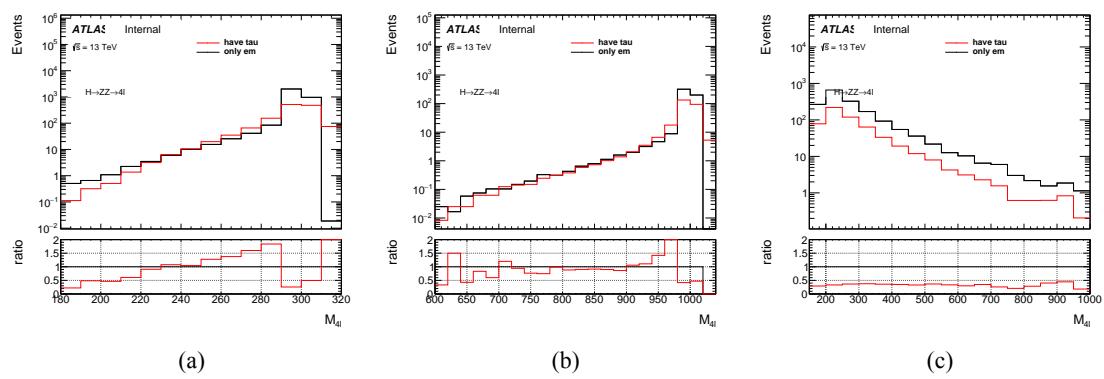


Fig. 7.4 The comparison of $m_{4\ell}$ distribution in the channels of ‘have tau’: events in $2e2\tau$ and $2\mu2\tau$ final states; ‘only em’: events in $2e2\mu$, $4e$ and 4μ final states; for three simulated samples: (a) NWA signal at 300 GeV; (b) NWA signal at 1000 GeV; (c) $q\bar{q} \rightarrow ZZ$ background.

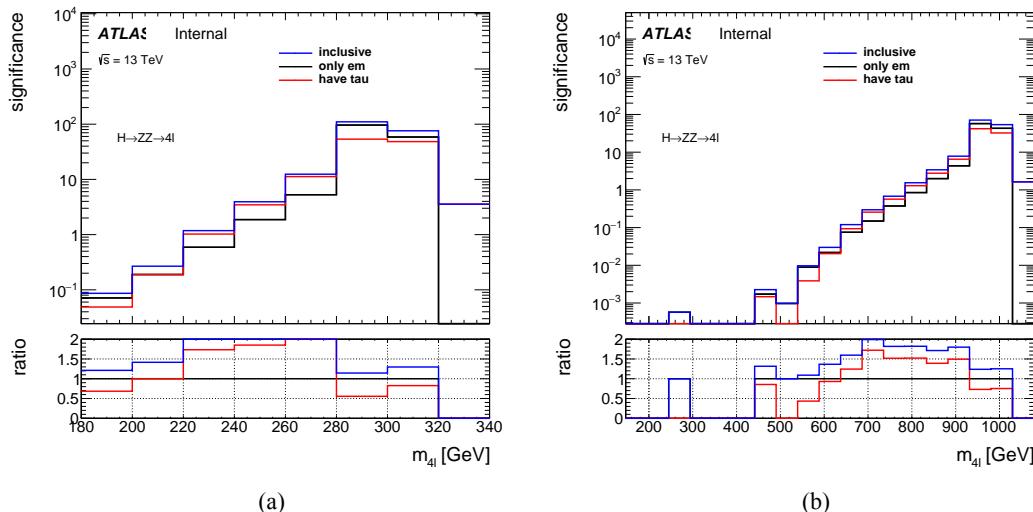


Fig. 7.5 Significance scan at the m_H of (a) 300 GeV; (b) 1000 GeV; for events in lepton flavor channel with or without τ -lepton.

reduce the separation power of $m_{4\ell}$ discriminant between signal and our backgrounds. Considering the more complicated background components and worse reconstructed mass resolution after inviting τ -lepton in reconstruction-level, the analysis team decided not to include τ -channel in this analysis.

7.3.4 Event categorizations

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

1. Cut-based categorization

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

- VBF-CBA-enriched category: Events have at least two selected jets as defined in section 7.3.1, with the two leading jets being separated by $|\Delta\eta_{jj}| > 3.3$ and invariant mass satisfying $m_{jj} > 400 \text{ GeV}$;
 - ggF-CBA-enriched categories: The remaining events that are not classified into VBF-enriched category. Then events are categorized into three channels based on lepton-flavor, namely ggF_2e2 μ , ggF_4e and ggF_4 μ .
- where ‘CBA’ stands for the cut-based categorization.

2. MVA-based categorization

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

DNN models

Figure 7.6 shows the architecture of VBF (left) and ggF (right) network. The network structure was chosen based on its AUC^[118] value from evaluation samples. The VBF network includes three parts: two recurrent neural networks (RNNs) and one multilayer perceptron (MLP). The ggF network consists of one RNN and one MLP.



Fig. 7.6 (a) VBF DNN architecture diagram. (b) ggF DNN architecture.

For training, the VBF and ggF signal samples at the masses of 200, 300, 400, 500, 600, 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 assigned with negative labels. In addition to the selections described in section 7.3.2, 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, so the training events for two network are independent.

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 back-

2686 ground (idea from Ref.^[119] with a modified reweighting procedure), as shown in figure 7.7
 2687 (figure 7.8) before and after reweighting for VBF (ggF) samples.

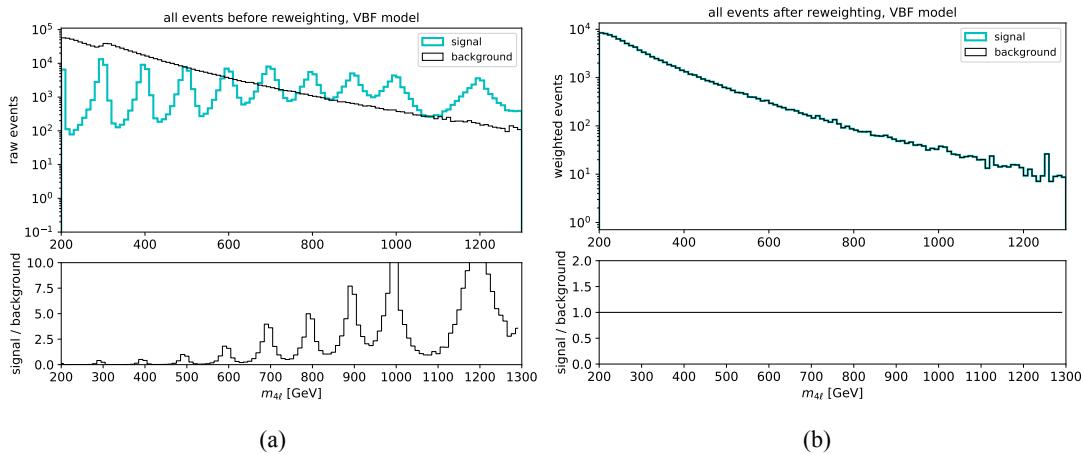


Fig. 7.7 (a) $m_{4\ell}$ distribution of raw (unweighted) training events for VBF signal (blue) and background (black); (b) $m_{4\ell}$ distribution of weighted VBF signal (blue) and background (black) used at training time.

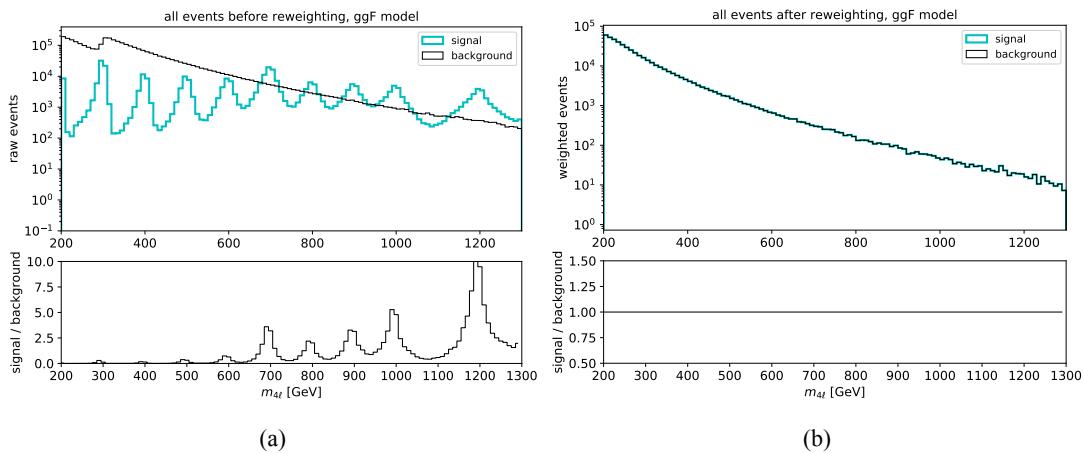


Fig. 7.8 (a) $m_{4\ell}$ distribution of raw (unweighted) training events for ggF signal (blue) and background (black); (b) $m_{4\ell}$ distribution of weighted ggF signal (blue) and background (black) used at training time.

2688 Input features

2689 Table 7.6 (table 7.7) lists the input features used for VBF (ggF) network during the
 2690 training. For VBF network, one RNN (the other one) takes the p_T and η of p_T -ordered four
 2691 leptons (two leading jets) as input features, which intends to study the time relationship
 2692 from particle decay between leptons (jets). For ggF network, the only one RNN model
 2693 takes the p_T and η of p_T -ordered four leptons as inputs.

2694 Evaluation of models

Table 7.6 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

2695 Figure 7.9 shows the output of “ggF-classifier” and “VBF-classifier” for data, SM
 2696 backgrounds and an example signal at 600 GeV. The ggF and VBF signals cross section
 2697 are set to be 100 times of their observed upper limit described in section 7.7.3 for ggF
 2698 output and 30 times of the observed upper limit for VBF output for best visibility.

2699 Then the optimal cut at output score from each classifier is chosen based on an overall
 2700 good performance of classifier to have a large significance improvement while retaining a
 2701 high signal efficiency. Figure 7.10 shows the significance improvements of MVA-based
 2702 cuts when comparing with cut-based one at different VBF (left) and ggF (right) mass
 2703 samples, where the significance is calculated from an asymptotic approximation^[120]:

$$Z = \sqrt{2 \left(n \ln \left[\frac{nb + \sigma_b^2}{b^2 + n\sigma_b^2} \right] - \frac{b^2}{\sigma_b^2} \ln \left[1 + \frac{\sigma_b^2(n-b)}{b(b+\sigma_b^2)} \right] \right)} \quad (7.1)$$

2704 where n denotes to the sum of expected signal and background, b is the background, and
 2705 σ_b is the uncertainty of background. Cut at 0.5 (0.8) for VBF (ggF) classifier is chosen as
 2706 shown in solid lines.

2707 Then the events passing VBF classifier are categorized into VBF-MVA-enriched cat-
 2708 egory. Otherwise, the events failing VBF classifier but passing ggF classifier are catego-
 2709 rized into ggF-MVA-high category, which is further split into 3 channels based on their
 2710 lepton flavor. All remaining events are sorted into one additional ggF-MVA-low cate-
 2711 gory. Thus there are five categories defined in MVA-based categorization. In summary,

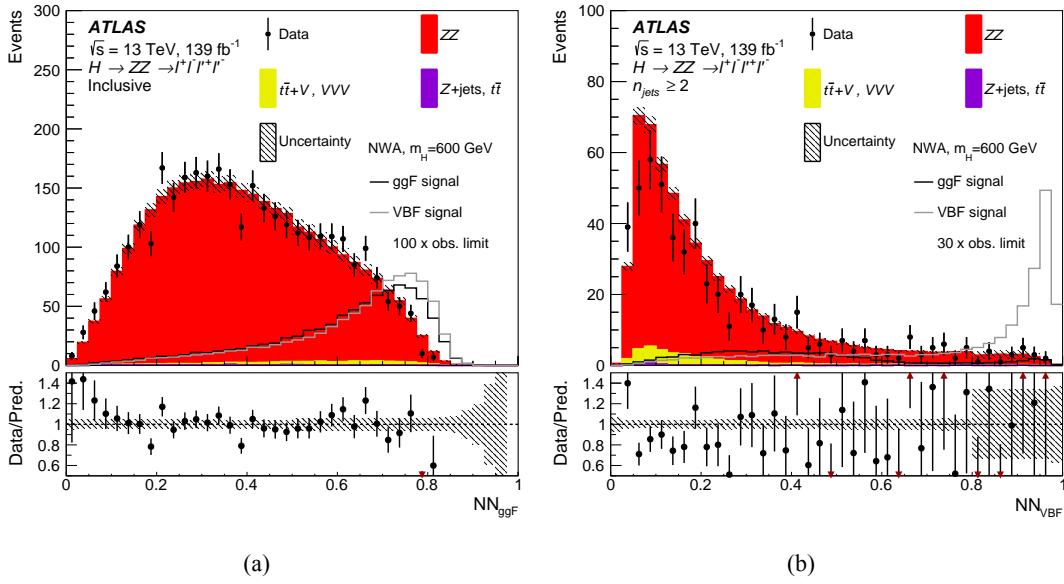


Fig. 7.9 The output score of “ggF-classifier” (a) and “VBF-classifier” (b) with the events passing the common event selections for the data, the SM backgrounds and an example of a NWA signal with a mass of 600 GeV. For the “VBF-classifier”, an additional requirement of at least two jets in the event is applied. The signals cross section are set to 100 times of the observed limit for the “ggF-classifier” and 30 times of the observed limit for the “VBF -classifier”. The $Z Z$ backgrounds are scaled by the normalisation factors shown in Table 7.12. The lower panels show the ratio of data to prediction. Statistical and experimental systematic uncertainties are included.

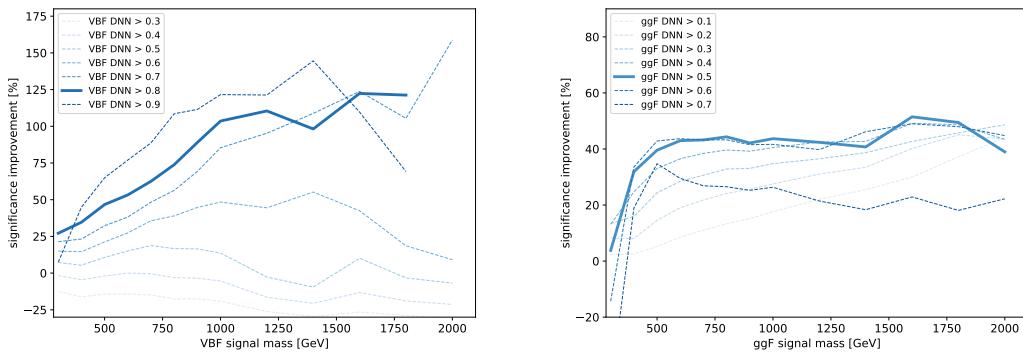


Fig. 7.10 Significance improvements of the MVA-based over the cut-based categorization of the VBF (ggF) category for VBF (ggF) signal samples from 300 to 2000 GeV for seven different cuts on the VBF (ggF) output score. The optimal cut of 0.8 (0.5) for VBF (ggF) score is chosen as the solid line, while 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 7.7 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

2712 cuts applied in categorization are defined as follow, and these different phase spaces are
 2713 also illustrated in figure 7.11.

- 2714 • VBF-MVA-enriched category: Events have at least two selected jets ($N_{\text{jets}} \geq 2$),
 2715 and with $\text{DNN}_{\text{VBF}} > 0.8$;
- 2716 • ggF-MVA-high categories: ($N_{\text{jets}} \geq 2 \ \&\& \text{DNN}_{\text{VBF}} \leq 0.8 \ \&\& \text{DNN}_{\text{ggF}} >$
 2717 $0.5) \ || (N_{\text{jets}} < 2 \ \&\& \text{DNN}_{\text{ggF}} > 0.5)$);
- 2718 • ggF-MVA-low category: All remaining events that fail VBF and ggF cuts men-
 2719 tioned above.

2720 7.3.5 Signal acceptance

2721 The signal acceptance is defined as the ratio of events passing all analysis selection in
 2722 each category to the total number of simulated events in whole phase space. In denom-
 2723 inator, the events with τ final states are not taken into account. And the contribution of
 2724 τ -lepton decay to electrons and muons final states is found to be negligible.

2725 Figure 7.12 and 7.13 show the acceptance of NWA signals in DNN- and Cut- based
 2726 categorization, estimated by merging the three signal MC campaigns, mc16a, mc16d and

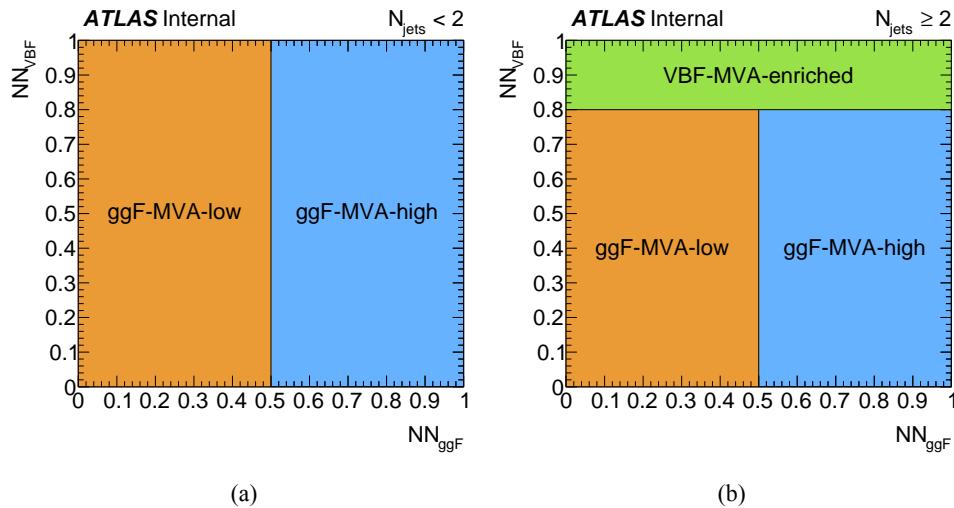


Fig. 7.11 Illustration of the MVA-based VBF and ggF event classification for events with (a) $N_{\text{jets}} < 2$ and (b) $N_{\text{jets}} \geq 2$.

2727 mc16e. A 3-rd order polynomial fit is applied for each category.

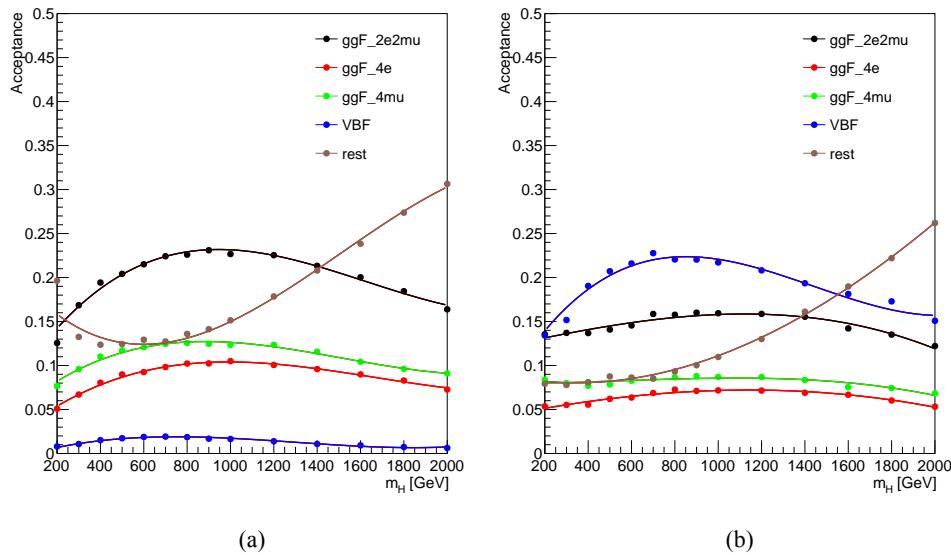


Fig. 7.12 NWA acceptance as a function of m_H for the MVA-based categorization for the samples of (a) ggF production; (b) VBF production.

2728 7.4 Background estimation

2729 In this analysis, 97% of total expected background events are from irreducible ZZ
 2730 backgrounds, which includes about 86% quark-antiquark annihilation ($q\bar{q} \rightarrow ZZ$), 10%
 2731 of gluon-induced production ($gg \rightarrow ZZ$) and around 1% of EW vector boson scattering
 2732 ($q\bar{q} \rightarrow ZZ$ EW) contribution. For $q\bar{q} \rightarrow ZZ$ EW, although it has small contribution

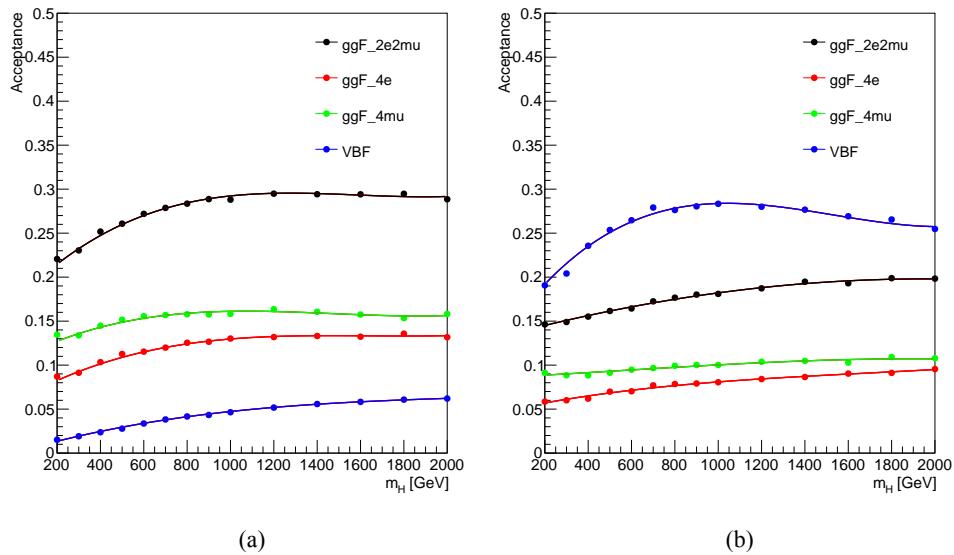


Fig. 7.13 NWA acceptance as a function of m_H for the Cut-based categorization for the samples of (a) ggF production mode; (b) VBF production mode.

in total background events after analysis selection, it's important for VBF category with about 16% contribution.

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

7.4.1 Irreducible backgrounds

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

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

The empirical function used for background parameterization is:

$$f(m_{4\ell}) = C_0 H(m_0 - m_{4\ell}) f_1(m_{4\ell}) + H(m_{4\ell} - m_0) f_2(m_{4\ell}), \quad (7.2)$$

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 7.2 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 7.14 to 7.16 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 7.17 to 7.19 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).

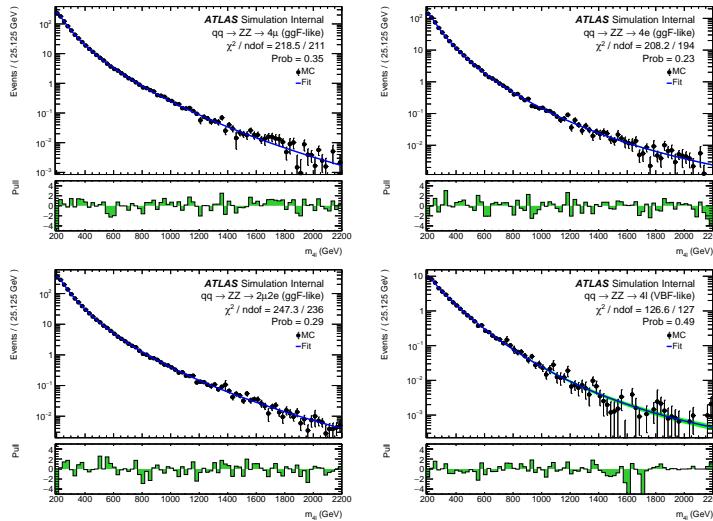


Fig. 7.14 Distributions of the $m_{4\ell}$ invariant mass fit projections of the $q\bar{q} \rightarrow ZZ$ background samples for the 4μ , $4e$ and $2\mu2e$ final states in the ggF-CBA-enriched category, and the 4ℓ inclusive VBF-CBA-enriched category. Cut-based categorization is used.

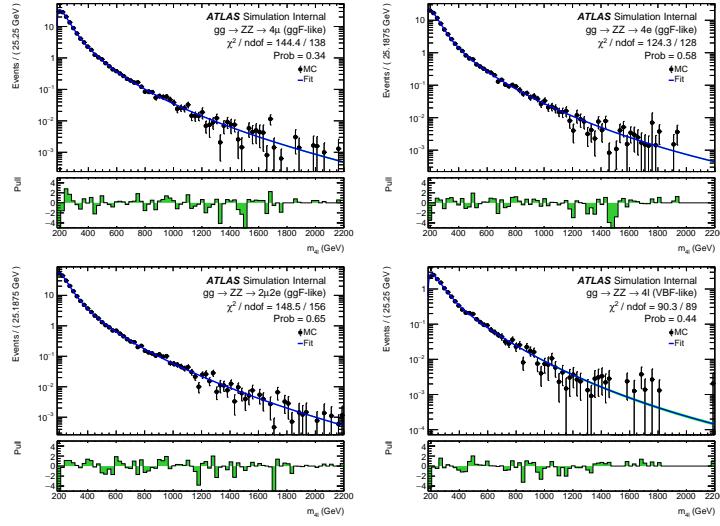


Fig. 7.15 Distributions of the $m_{4\ell}$ invariant mass fit projections of the $gg \rightarrow ZZ$ background samples for the 4μ , $4e$ and $2\mu2e$ final states in the ggF-CBA-enriched category, and the 4ℓ inclusive VBF-CBA-enriched category. Cut-based categorization is used.

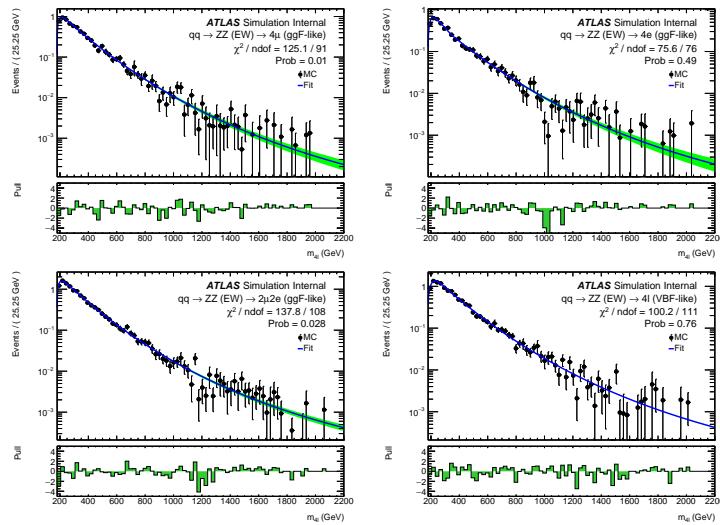


Fig. 7.16 Distributions of the $m_{4\ell}$ invariant mass fit projections of the $q\bar{q} \rightarrow ZZ$ (EW) background samples for the 4μ , $4e$ and $2\mu2e$ final states in the ggF-CBA-enriched category, and the 4ℓ inclusive VBF-CBA-enriched category. Cut-based categorization is used.

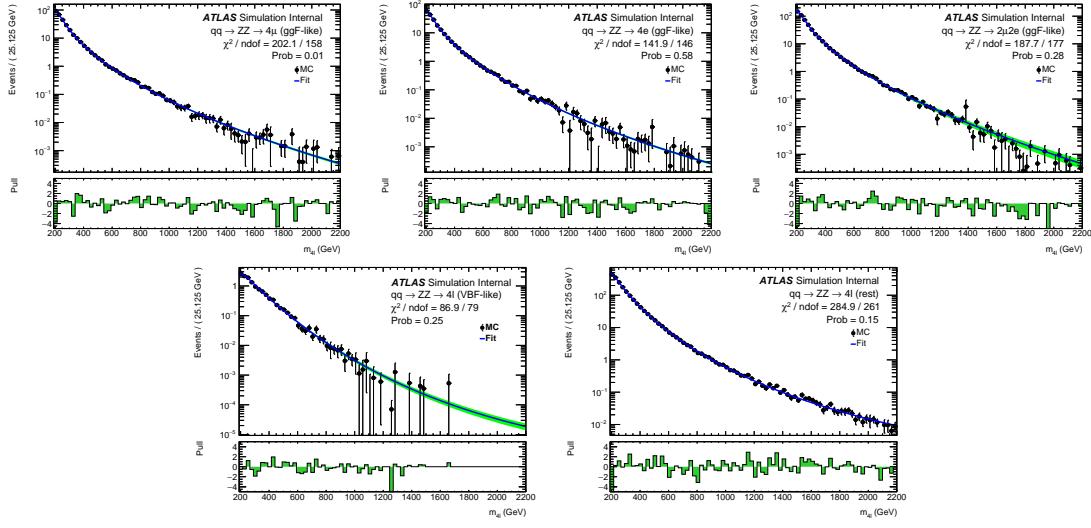


Fig. 7.17 Distributions of the $m_{4\ell}$ invariant mass fit projections of the $q\bar{q} \rightarrow ZZ$ background samples for the 4μ , $4e$ and $2\mu2e$ final states in the ggF-MVA-high category, the 4ℓ inclusive ggF-MVA-low category and VBF-MVA-enriched category. DNN-based categorization is used.

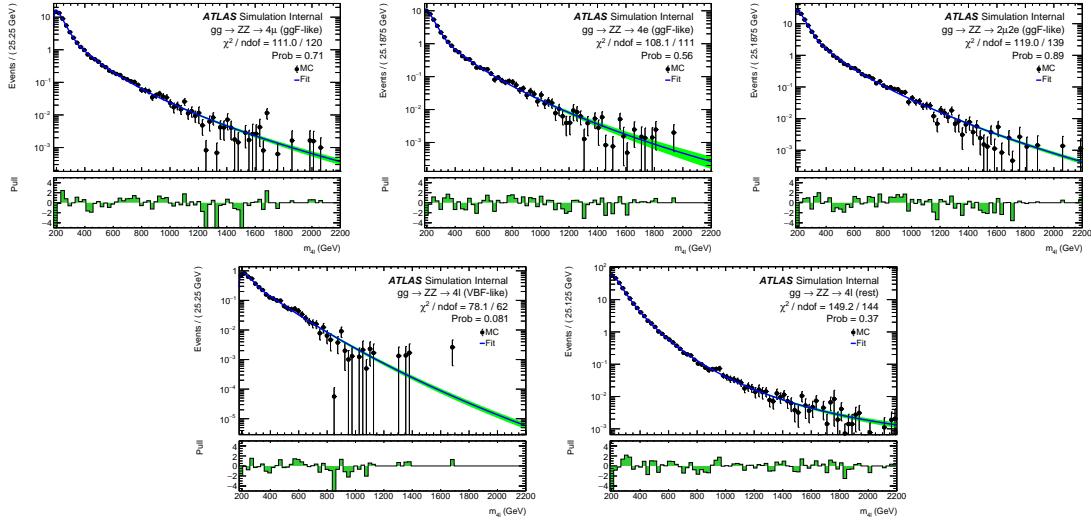


Fig. 7.18 Distributions of the $m_{4\ell}$ invariant mass fit projections of the $gg \rightarrow ZZ$ background samples for the 4μ , $4e$ and $2\mu2e$ final states in the ggF-MVA-high category, the 4ℓ inclusive ggF-MVA-low category and VBF-MVA-enriched category. DNN-based categorization is used.

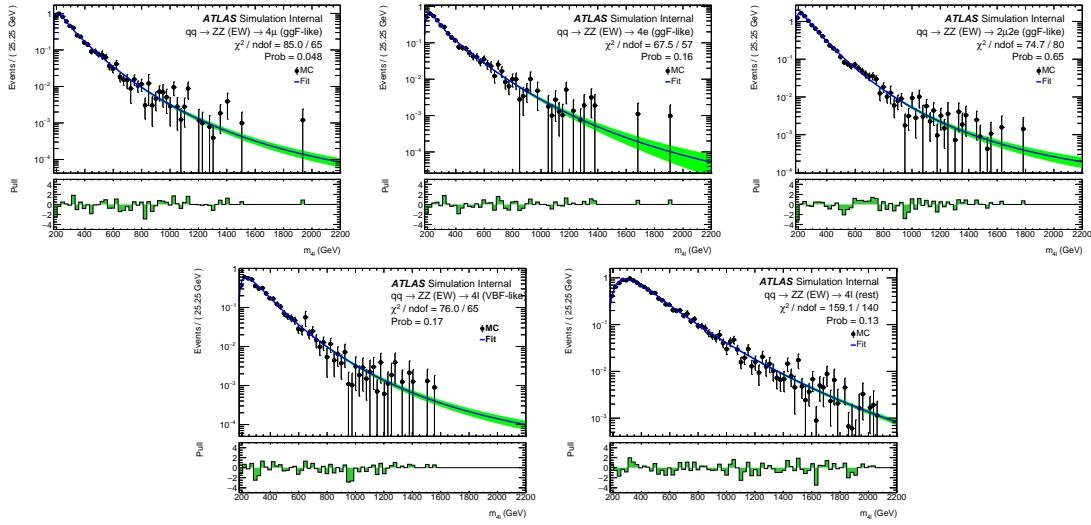


Fig. 7.19 Distributions of the $m_{4\ell}$ invariant mass fit projections of the $q\bar{q} \rightarrow ZZ$ (EW) background samples for the 4μ , $4e$ and $2\mu2e$ final states in the ggF-MVA-high category, the 4ℓ inclusive ggF-MVA-low category and VBF-MVA-enriched category. DNN-based categorization is used.

7.4.2 Reducible backgrounds

Similar to section 6.4, the reducible backgrounds include $Z+jets$ (consisting 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 estimate 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^[121], 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-sign 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, but 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 to it.

The fit results of normalizations are then propagated to signal region (SR) by applying transfer factors to account for 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 ^[122]. The difference from using different smoothing strength (ρ) in kernel estimation is taken into account as additional shape uncertainties for these reducible backgrounds.

7.5 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

2816 narrow-width signal (NWA) and large-width signal (LWA), as well as the modelling of
2817 Randall-Sundrum graviton (RSG) signal.

2818 7.5.1 Modelling of narrow-width signal

2819 For narrow-width (NWA) signal, the $m_{4\ell}$ width is totally determined by detector reso-
2820 lution, which is modelled by the sum of a Crystal Ball (C) function^[123-124] and a Gaussian
2821 (G) function:

$$P_s(m_{4\ell}) = f_C \cdot C(m_{4\ell}; \mu, \sigma_C, \alpha_C, n_C) + (1 - f_C) \cdot G(m_{4\ell}; \mu, \sigma_G) \quad (7.3)$$

2822 The two functions share the same central value μ , while the resolution parameters, σ_C
2823 and σ_G , are different. In the Crystal Ball function, the parameters α_C and n_C model the
2824 shape of non-Gaussian tail, and the fraction parameter f_C is used to ensure the relative
2825 normalization between two functions.

2826 The parameters are obtained by fitting to signal MC simulations combining the mc16a,
2827 mc16d and mc16e campaigns for each category at each mass points from 200 GeV to
2828 2000 GeV respectively, and the shape of ggF and VBF signals are found to be similar.
2829 Figure 7.20 shows the $m_{4\ell}$ distribution and fitted curves for ggF production at mass from
2830 200 GeV to 2000 GeV in $2e2\mu$ channel as examples.

2831 Then the $C + G$ parameters are fitted with a polynomial function as the function of
2832 generated mass points (m_H), as an example shown in figure 7.21 for ggF production in
2833 $2e2\mu$ channel. The fitting quality can be measured by the Pearson's χ^2 , which is within
2834 3 (2) for $2e2\mu$ ($4e$ and 4μ) channel.

2835 In addition, possible difference on the signal yield extracted from parameterization and
2836 MC simulation is studied. Figure 7.22 shows this difference by computing $\frac{N_{\text{reco}} - N_{\text{fit}}}{N_{\text{fit}}}$, where
2837 N_{reco} denotes the total number of reconstructed events observed from MC simulation at
2838 that mass point and N_{fit} depicts the number of events obtained from the fitted pdf. The
2839 differences are treated as an additional systematic uncertainty with the value of 2% (1%)
2840 for $2e2\mu$ ($4e$ and 4μ) channel in statistical fit.

2841 In summary, the final interpolated signal shapes for the ggF production mode are
2842 shown together in figure 7.23 for mass points with step of 100 GeV from 200 GeV to
2843 3000 GeV.

2844 7.5.2 Modelling of large-width signal

2845 The $m_{4\ell}$ shape of heavy Higgs model in large-width (LWA) hypothesis can be de-
2846 scribed by a convolution of a truth distribution and a resolution from detector effect. The

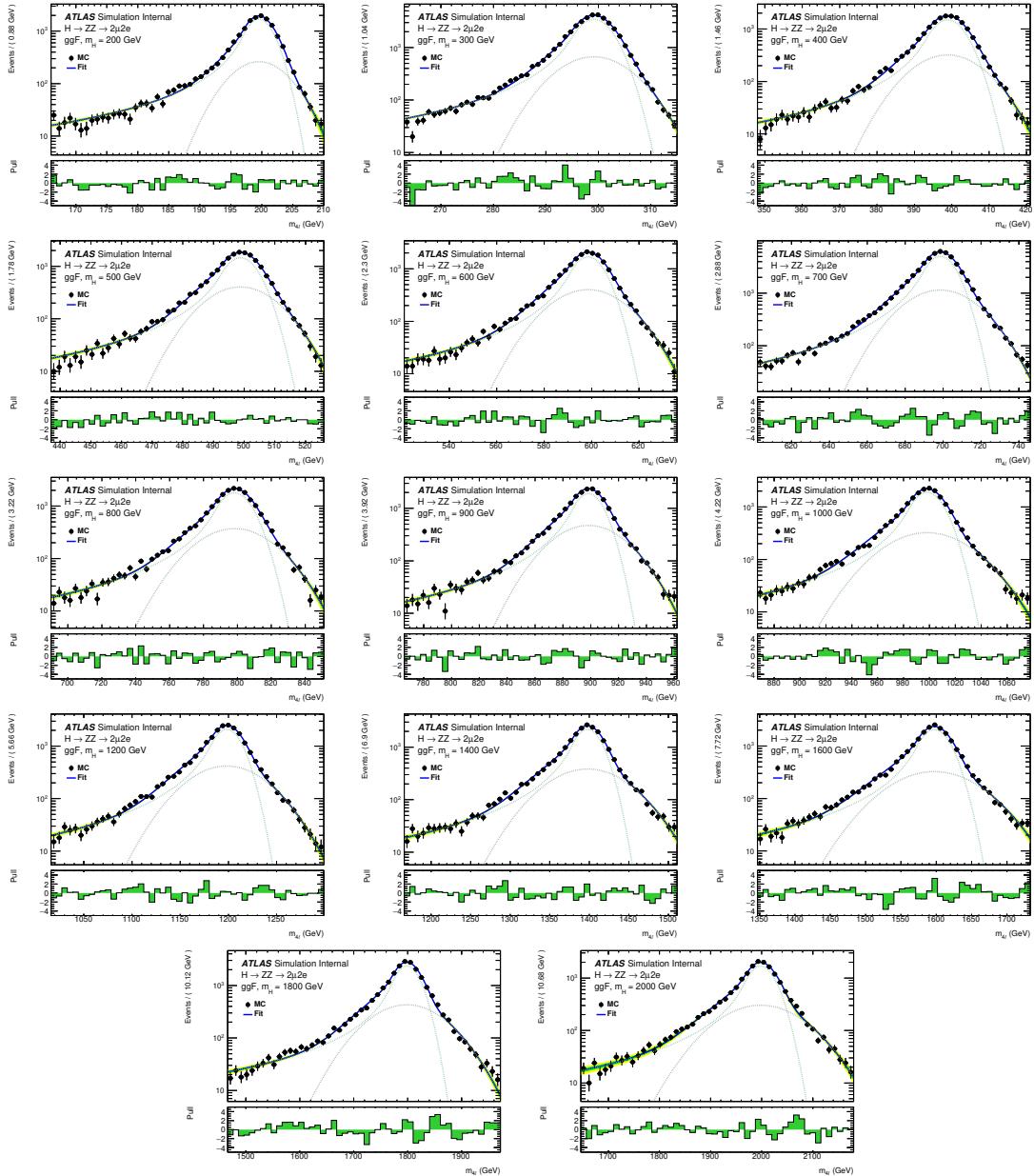


Fig. 7.20 Distributions of the $m_{2\mu 2e}$ and fit projection for signal samples between 200 to 3000 GeV for ggF production mode. Three MC campaigns, mc16a, mc16d and mc16e, are combined. The lower panel in each plot shows the pull distribution.

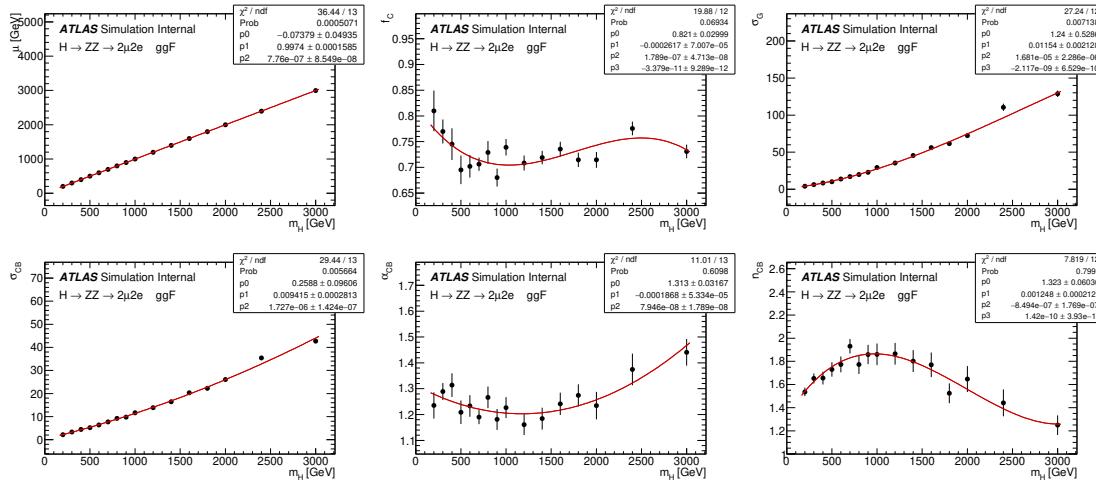


Fig. 7.21 Polynomial fits of the parameters μ , f_C , σ_G , σ_C , n_C and α_C for the signal $C + G$ model in the $2\mu 2e$ channel as a function of m_H for the ggF production mode. The combination of the mc16a, mc16d and mc16e MC campaigns is used.

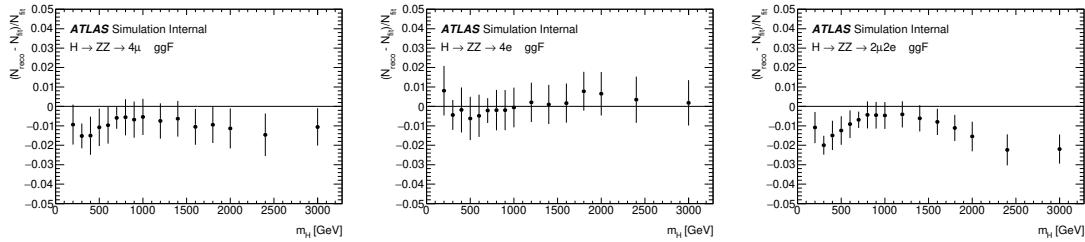


Fig. 7.22 The difference between MC simulation and parameterization of 4μ (left), $4e$ (middle) and $2\mu 2e$ (right) for the ggF production mode. The combination of the mc16a, mc16d and mc16e MC campaigns is used.

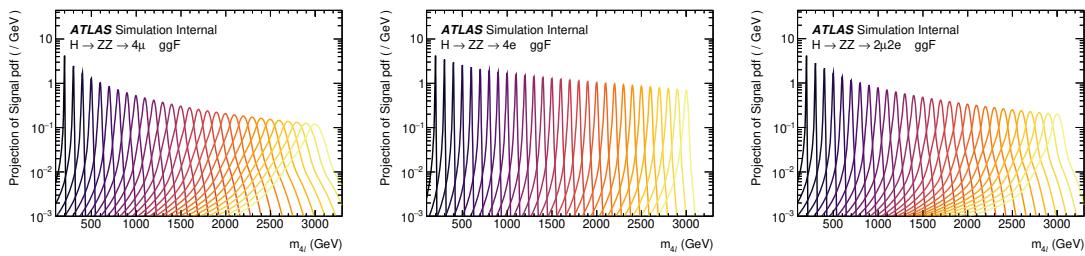


Fig. 7.23 The final signal shapes for the ggF production mode, interpolated from the polynomial fit parameters.

2847 detector resolution effect is the one modelled by the function described in NWA parame-
2848 terization, as in NWA model the truth level width is negligible.

2849 The differential parton cross section for the heavy Higgs model can be written as^[125]:

$$\sigma_{gg \rightarrow H \rightarrow ZZ}(s) = \frac{1}{2s} \int d\Omega |A_{gg \rightarrow H}(s, \Omega)|^2 \frac{1}{|s - s_H|^2} |A_{H \rightarrow ZZ}(s, \Omega)|^2 \quad (7.4)$$

2850 where $A_{gg \rightarrow H}(s, \Omega)$ and $A_{H \rightarrow ZZ}(s, \Omega)$ are corresponding Higgs production and decay
2851 amplitudes, and $\frac{1}{|s - s_H|}$ denotes the Higgs propagator and Ω represents the phase space of
2852 the process.

2853 Using the definition of a partial width,

$$\Gamma_{H \rightarrow F}(s) = \frac{1}{2\sqrt{s}} \int d\Omega |A_{H \rightarrow F}(s, \Omega)|^2 \quad (7.5)$$

2854 the parton cross section can be rewritten as,

$$\sigma_{gg \rightarrow H \rightarrow ZZ}(s) = 2 \frac{1}{|s - s_H|^2} \times \Gamma_{H \rightarrow gg}(s) \times \Gamma_{H \rightarrow ZZ}(s) \quad (7.6)$$

with the components computed in Ref^[125-126]:

$$\begin{aligned} \frac{1}{s - s_H} &= \frac{1 + i \cdot \bar{\Gamma}_H / \bar{m}_H}{s - \bar{m}_H^2 + i \cdot s \cdot \bar{\Gamma}_H / \bar{m}_H} \\ \bar{m}_H &= \sqrt{\bar{\Gamma}_H^2 + m_H^2} \\ \bar{\Gamma}_H &= \bar{m}_H \cdot \frac{\Gamma_H}{m_H} \end{aligned} \quad (7.7)$$

$$\Gamma_{H \rightarrow ZZ}(s) = C \cdot s^{\frac{3}{2}} \cdot \left[1 - \frac{4m_Z^2}{s} + \frac{3}{4} \left(\frac{4m_Z^2}{s} \right)^2 \right] \cdot \left[1 - \frac{4m_Z^2}{s} \right]^{\frac{1}{2}} \quad (7.8)$$

$$\begin{aligned} \Gamma_{H \rightarrow gg}(s) &= C \cdot s^{\frac{3}{2}} \cdot |A_t(\tau_t)|^2 \\ A_t(\tau) &= 2 \frac{\tau + (\tau - 1)f(\tau)}{\tau^2} \\ \tau_t &= \frac{s}{4m_t^2} \\ f(\tau) &= \begin{cases} \arcsin^2(\sqrt{\tau}), & \tau \leq 1 \\ -\frac{1}{4} \left[\log \frac{1+\sqrt{1-\tau^{-1}}}{1-\sqrt{1-\tau^{-1}}} - i\pi \right]^2, & \tau > 1 \end{cases} \end{aligned} \quad (7.9)$$

2855 where m_f stands for the mass of a fermion f , and Γ_H denotes an assumed total width
2856 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 7.6 by multiplication with gluon-gluon luminosity \mathcal{L}_{gg} described in^[127]. 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:

$$\sigma_{pp \rightarrow H \rightarrow ZZ}(m_{4\ell}) = 2 \cdot m_{4\ell} \cdot \mathcal{L}_{gg} \cdot \frac{1}{|s - s_H|^2} \cdot \Gamma_{H \rightarrow gg}(m_{4\ell}^2) \cdot \Gamma_{H \rightarrow ZZ}(m_{4\ell}^2) \quad (7.10)$$

The analytical shapes of truth level $m_{4\ell}$ distribution of gg2VV MC samples is shown on figure 7.24.

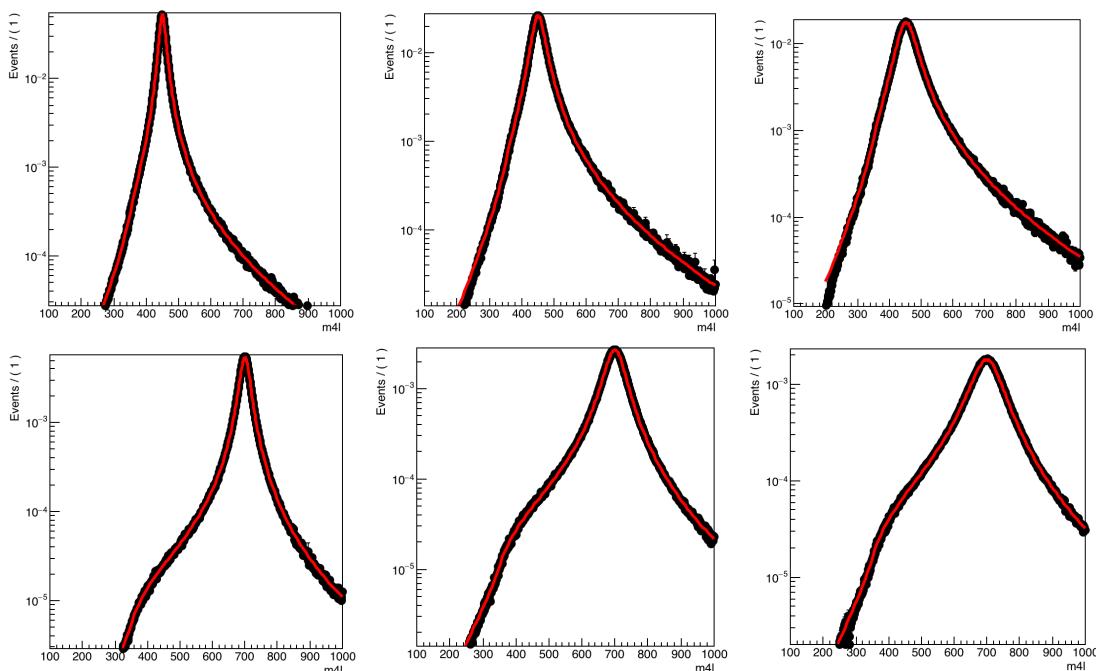


Fig. 7.24 Comparison of the analytical shape to a truth $m_{4\ell}$ distribution of gg2VV MC samples for $m_H = 450$ GeV (top), 700 GeV (bottom) and width equal to 5% (left), 10% (middle), 15% (right) of the mass.

The reconstruction level signal shape can then be modelled by the analytical truth shape convoluted with detector effects modelled in section 7.5.1. 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 7.25, the shapes are well compatible between each other. This modelling is not valid for lower masses due to the rapid change of detector resolution.

7.5.3 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

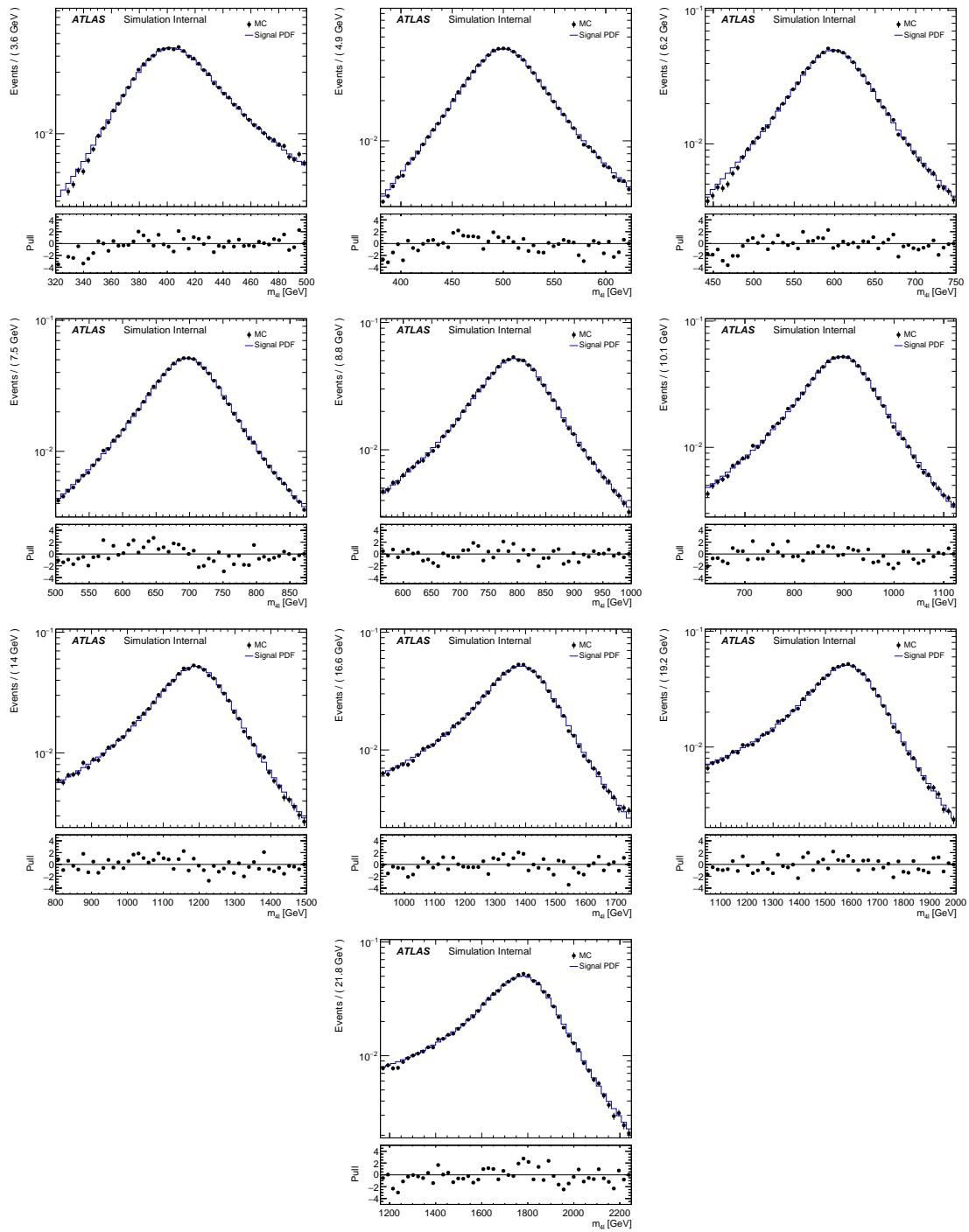


Fig. 7.25 Comparison between the analytical shape convoluted with detector effects and the reconstructed $m_{2\mu 2e}$ MC distribution for mass points ranging from 400 to 1800 GeV and width equal to 15% of the mass.

- The SM (light) Higgs at mass of around 125 GeV with an amplitude A_h
- The BSM heavy Higgs we are searching in this analysis with an amplitude A_H

The three processes can interfere with each other due to the same initial and final states.

The parton cross section for these processes can be written as:

$$\begin{aligned}\sigma_{gg \rightarrow (X) \rightarrow ZZ}(s) &= \frac{1}{2s} \int d\Omega |A_h(s, \Omega) + A_H(s, \Omega) + A_B(s, \Omega)|^2 \\ &= \frac{1}{2s} \int d\Omega (|A_h(s, \Omega)|^2 + |A_H(s, \Omega)|^2 + |A_B(s, \Omega)|^2) + \\ &\quad + \frac{1}{s} \int d\Omega (Re [A_h(s, \Omega) \cdot A_B^*(s, \Omega)] \\ &\quad + Re [A_H(s, \Omega) \cdot A_B^*(s, \Omega)] + Re [A_B(s, \Omega) \cdot A_h^*(s, \Omega)])\end{aligned}\quad (7.11)$$

The first term in equation 7.11 denotes the on-shell SM Higgs contribution, which is negligible in this analysis. The second term corresponds to the heavy Higgs contribution, whose line shape has been described in previous section. The third term is the $gg \rightarrow ZZ$ continuum process, while the forth term is the interference between SM Higgs and $gg \rightarrow ZZ$ continuum. The fifth and sixth terms are the interferences between heavy Higgs and $gg \rightarrow ZZ$ continuum (H-B), and between heavy Higgs and SM Higgs (H-h) that we are interested in. More details about the parameterization of these two interferences are described as below.

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

The parton cross section of this interference term can be expressed as:

$$\sigma_{gg}(s) = \frac{1}{s} \text{Re} \left[\frac{1}{s - s_H} \int d\Omega \cdot A_H^P(s, \Omega) \cdot A_H^D(s, \Omega) \cdot A_B^*(s, \Omega) \right] \quad (7.12)$$

By assuming that this function has a smooth behaviour, it can be replaced with complex polynomial:

$$\int d\Omega \cdot A_H^P(s, \Omega) \cdot A_H^D(s, \Omega) \cdot A_B^*(s, \Omega) \approx (a_0 + a_1 \cdot \sqrt{s} + \dots) + i \cdot (b_0 + b_1 \cdot \sqrt{s} + \dots) \quad (7.13)$$

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

Same as description for equation 7.10, the parton cross section can be transformed into a hadron cross section as a function of $m_{4\ell}$:

$$\sigma_{pp}(m_{4\ell}) = \mathcal{L}_{gg} \cdot \frac{1}{m_{4\ell}} \cdot \text{Re} \left[\frac{1}{s - s_H} \cdot ((a_0 + a_1 \cdot m_{4\ell} + \dots) + i \cdot (b_0 + b_1 \cdot m_{4\ell} + \dots)) \right] \quad (7.14)$$

where the propagators are shown in equation 7.7.

Figure 7.26 shows the distributions of interference function obtained by simultaneous fitting to $m_{4\ell}$ shape from truth level H-B interference simulation at different mass in $2e2\mu$ channel as an example.

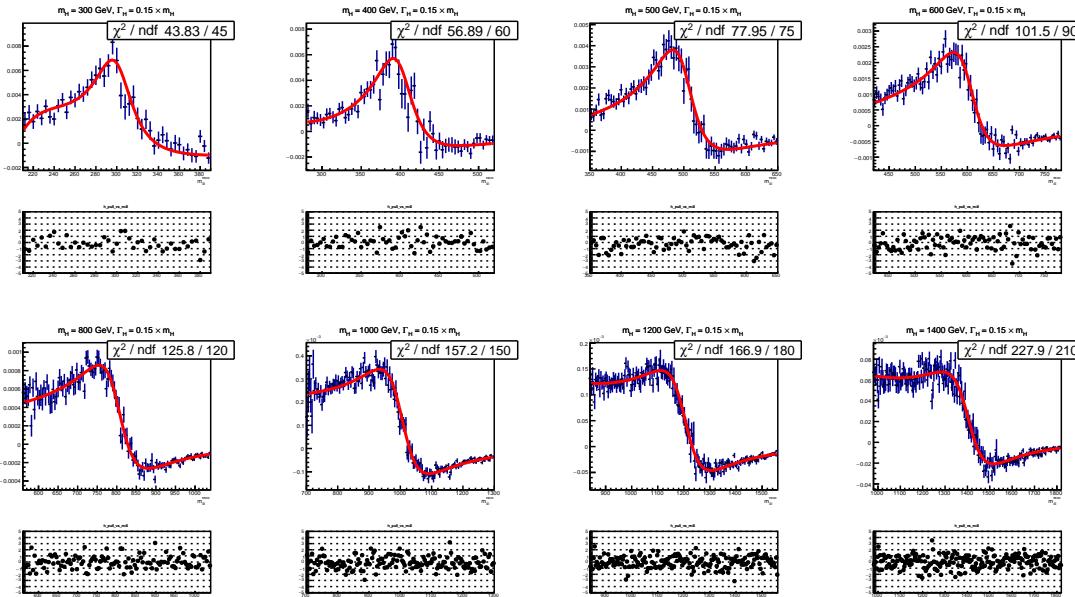


Fig. 7.26 The interference (H-B) model fitted to the truth $m_{4\ell}$ MC distribution after signal region selection for $2\mu 2e$ channel.

2. Interference between heavy Higgs and SM Higgs

The parton cross section of this interference term can be written as:

$$\sigma_{gg}(s) = \frac{1}{s} \int d\Omega \cdot \text{Re} \left[A_H^P(s, \Omega) \cdot \frac{1}{s - s_H} \cdot A_H^D(s, \Omega) \cdot A_h^{P*}(s, \Omega) \cdot \frac{1}{(s - s_h)^*} \cdot A_h^{D*}(s, \Omega) \right] \quad (7.15)$$

By assuming the production and decay amplitudes are the same for heavy Higgs boson and SM Higgs boson, the cross section function can be simplified to:

$$\sigma_{gg}(s) = \frac{1}{s} \int d\Omega \cdot \text{Re} \left[\frac{1}{s - s_H} \cdot \frac{1}{(s - s_h)^*} \right] \cdot |A_{gg \rightarrow H}(s, \Omega)|^2 |A_{H \rightarrow ZZ}(s, \Omega)|^2 \quad (7.16)$$

Taking into account Equation 7.5:

$$\sigma_{gg}(s) = 4 \cdot \text{Re} \left[\frac{1}{s - s_H} \cdot \frac{1}{(s - s_h)^*} \right] \cdot \Gamma_{H \rightarrow gg}(s) \cdot \Gamma_{H \rightarrow ZZ}(s) \quad (7.17)$$

where the propagators are described in equation 7.7, and the partial widths are described in equations 7.8 and 7.9.

2905 Same as previous procedure, the parton cross section can be transformed to a hadron
 2906 cross section as a function of $m_{4\ell}$:

$$\sigma_{pp}(m_{4\ell}) = 4 \cdot m_{4\ell} \cdot \mathcal{L}_{gg} \cdot \text{Re} \left[\frac{1}{s - s_H} \cdot \frac{1}{(s - s_h)^*} \right] \cdot \Gamma_{H \rightarrow gg}(m_{4\ell}) \cdot \Gamma_{H \rightarrow ZZ}(m_{4\ell}) \quad (7.18)$$

2907 The modelling procedure of interference is the same as the way for large-width signal
 2908 described in section 7.5.2. The truth line shape is measured as analytical function from
 2909 equation 7.18, and then convolute with detector effect from NWA parameterization to get
 2910 the reconstruction level shape.

2911 For LWA signal model, these two interferences are carefully taken into account, and
 2912 the integration of the pure LWA signal with the interferences is used for further studies.
 2913 Figure 7.27 shows the signal model for large-width scenario at mass points of 400 GeV,
 2914 600 GeV, 800 GeV, for three different signal widths: 5%, 10%, 15%, with and without
 2915 interference. Additionally, the contribution of the interference between heavy Higgs and
 2916 SM Higgs (H-h) is shown together with the one between heavy Higgs and SM $gg \rightarrow$
 2917 ZZ background (H-b). One can see the interference effect on signal shape becomes less
 2918 important when going to higher mass.

2919 7.5.4 Modelling of spin-2 RS Graviton signal

2920 The search for Randall-Sundrum (RS) graviton is performed in mass region between
 2921 600 to 2000 GeV. The width of resonance is determined by the $k/\overline{M}_{\text{Planck}}$, which, as
 2922 mentioned in section 7.2.3, is set to be 1. In this configuration, the width of signal is
 2923 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 7.5.1. And the truth-level shape is 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 [128]. So the truth lineshape of $m_{4\ell}$ is taken from:

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

2924 The truth-level signal model is extracted by fitting to MC simulation at truth-level with
 2925 the mass m_G and width Γ parameters floating at each mass points respectively. And then

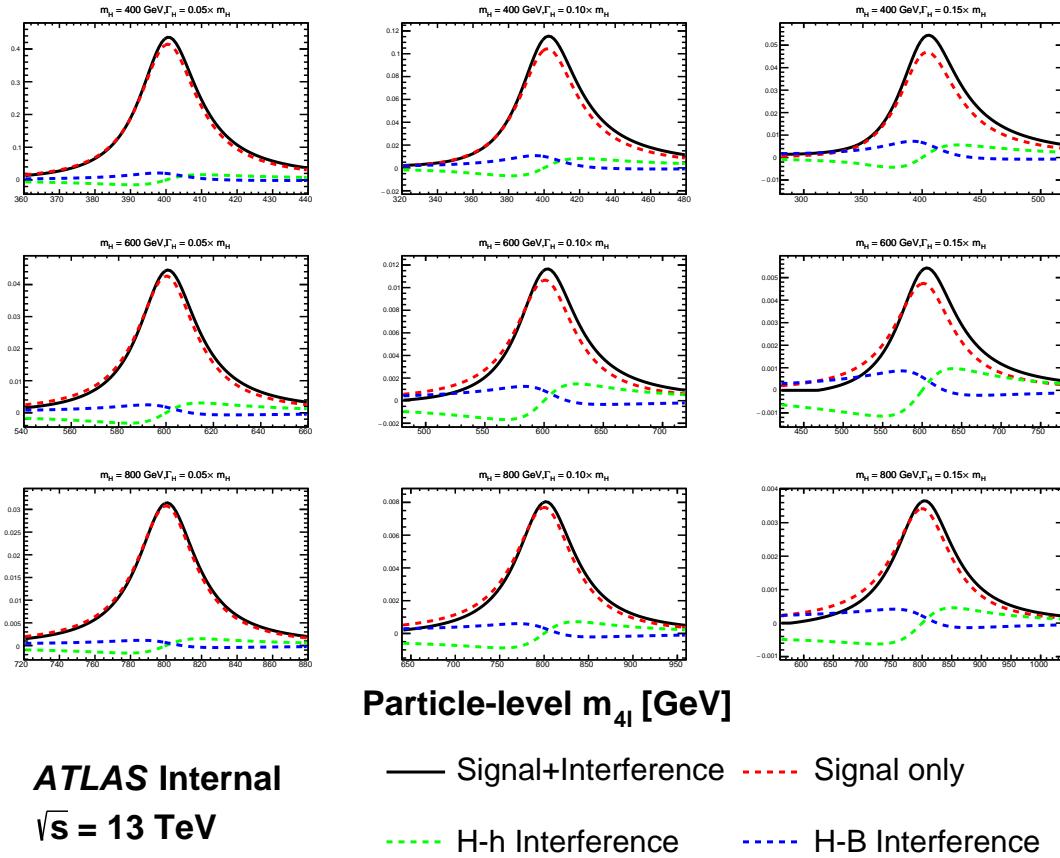


Fig. 7.27 The signal modelling for the large-width scenario at m_H of 400 GeV (top), 600 GeV (middle) and 800 GeV (bottom), as well as three different signal width: 5% (left), 10% (middle) and 15% (right). The contribution of the interference between heavy Higgs and SM Higgs (H-h) is shown together with the one between heavy Higgs and SM $gg \rightarrow ZZ$ background (H-b).

the two parameters are parameterized as the function of m_H by a linear fit as shown in figure 7.28.

The final signal model is obtained by convolving the truth-level lineshape with the detector resolution function. To verify the result, figure 7.29 compares the $m_{4\ell}$ lineshape from parameterization with the one observed from reconstructed-level MC simulation in $2e2\mu$ channel at masses of 600 GeV, 1600 GeV and 2000 GeV as examples.

7.6 Systematic uncertainties

This section describes the sources and values of theoretical and experimental systematic uncertainties considered in this analysis.

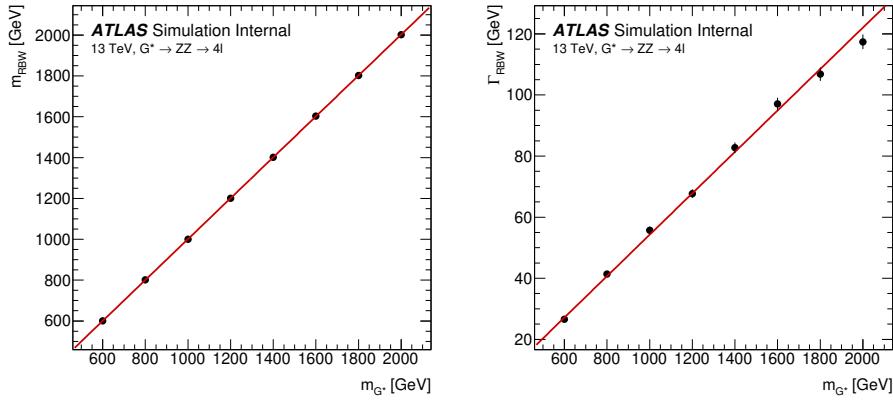


Fig. 7.28 Fitted parameters of the graviton RBW, m_{RBW} and Γ_{RBW} , as a function of the graviton resonance mass, m_G .

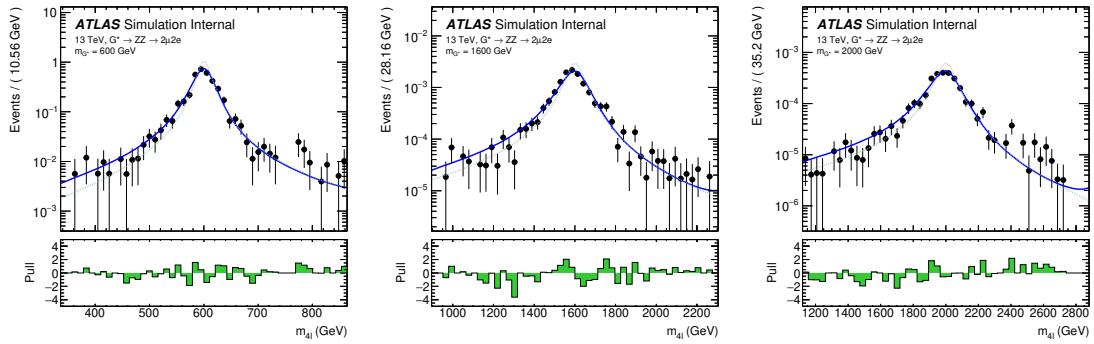


Fig. 7.29 Reconstructed $m_{4\ell}$ distributions in the $2\mu 2e$ channel with the final signal model superimposed for each RS graviton signal sample at masses of 600 GeV, 1600 GeV and 2000 GeV. The lower panel in each plot shows the pull distribution. The dashed green lines show the truth-level graviton signal models for reference.

2935 7.6.1 Theoretical systematics

2936 The theoretical modelling uncertainties include the PDF variations, missing QCD
 2937 higher-order corrections via the variations of factorisation and renormalization scales, and
 2938 the parton showering uncertainties.

2939 1. Theoretical uncertainties for signal

2940 The PDF, QCD scale and parton showering uncertainties affecting the acceptance dif-
 2941 ference originating from analysis selection for signal are taken into account in different
 2942 categories. The acceptance uncertainties are calculated on the acceptance factor which
 2943 extrapolates from the fiducial space to the full phase space by a simple ratio:

$$A = \frac{N_{\text{fiducial}}}{N_{\text{total}}} \quad (7.19)$$

2944 For PDF uncertainties, the standard derivations of 100 PDF replicas of NNPDF3.0
 2945 NNLO, as well as comparison to two external PDF sets: MMHT2014 NNLO, CT14

2946 NNLO are considered. For missing QCD higher-order corrections, the effects are studied
 2947 with truth events by comparing weights corresponding to variations of the renormalization
 2948 and factorization scale factors, up and down by a factor of two, and the envelop of dif-
 2949 ferent variations is used. The parton showering uncertainties are estimated by comparing
 2950 events with different setting via Pythia8.

2951 Systematic uncertainties are studied for both cut- and MVA- based event categoriza-
 2952 tions, for cut-based analysis in two different categories: the inclusive ggF-CBA-enriched
 2953 and VBF-CBA-enriched category, and for MVA-based one in three different categories:
 2954 inclusive ggF-MVA-high, ggF-MVA-low and VBF-MVA-enriched category. This section
 2955 shows the MVA-based results as an example.

2956 Table 7.8 and 7.9 show the theoretical uncertainties mentioned above for ggF and
 2957 VBF signal respectively in MVA-based categorization.

**Table 7.8 Summary of acceptance uncertainties of PDF, QCD scale and parton shower varia-
 tions 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 7.9 Summary of acceptance uncertainties of PDF, QCD scale and parton shower varia-
 tions 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%

2958 **2. Theoretical uncertainties for SM background processes**

2959 The theoretical uncertainties of irreducible ZZ backgrounds are considered in terms
 2960 of both the variations of shape of $m_{4\ell}$ distributions and the acceptance originating from
 2961 the event selection.

2962 The PDF and QCD scale uncertainties are considered by using the same method as
 2963 described for signal. The parton showering uncertainties for those Sherpa samples are
 2964 evaluated by varying the resummation scale by a factor of 2, changing the CKKW setting
 2965 and using different showering option, following the PMG recommendation in Ref.^[129],
 2966 and the quadratic sum between the uncertainties in different kinds of showering option
 2967 is taken as final result of uncertainties. Moreover, the shape uncertainty associated with
 2968 electroweak higher-order correction for $q\bar{q} \rightarrow ZZ$ process is also taken into account.

2969 Same as for signals, these theoretical uncertainties for irreducible backgrounds are
 2970 studied for both cut- and MVA- based event categorizations. The value of shape uncer-
 2971 tainties vary from less than 1% at low mass region to 50% at high mass tail due to large
 2972 statistic fluctuation. As for the acceptance uncertainties, the values vary from about 1%
 2973 for PDF variations to 40% for parton showering variations. The VBF category has relative
 2974 larger uncertainties.

2975 Table 7.10 summarizes the acceptance uncertainties of PDF, QCD scale, and parton
 2976 showering variations for the dominant background: $q\bar{q} \rightarrow ZZ$.

**Table 7.10 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%

2977 7.6.2 Experimental systematics

2978 The signal and background predictions used in this analysis are also affected by
 2979 various sources of experimental systematic uncertainties. Similar as described in sec-
 2980 tion 6.5.2, the dominant experimental uncertainties in this analysis also come from the
 2981 energy/momentum scales and reconstruction and identification efficiencies of the leptons
 2982 and jets, as well as the luminosity uncertainty. The systematic uncertainties are calculated
 2983 using the recommendations from the Combined Performance (CP) groups of ATLAS ex-
 2984 periment. In addition, as mentioned in previous sections, the uncertainties of irreducible
 2985 background modelling, reducible background shape smoothing procedure and signal yield
 2986 difference between simulation and parameterization are all taken into account. The impact
 2987 of a few largest systematics and their value from statistical fit are studied in section 7.7.

2988 7.7 Results in $\ell\ell\ell'\ell'$ channel

2989 The statistical treatment in searching for heavy resonances in $ZZ \rightarrow \ell\ell\ell'\ell'$ final
 2990 state is described in this section. Results are presented in both cut- and MVA- based
 2991 analysis.

2992 7.7.1 Statistical procedure

2993 The upper limits on heavy resonances are obtained using the unbinned profile likelihood fits. $m_{4\ell}$ is the discriminant. The likelihood function can be written as the product
2994 of the probability of observing n events in Poisson distribution and the sum of signal and
2995 background weighted following their probability density functions (PDFs):
2996

$$L(x_1..x_n|\sigma_{ggF}, \sigma_{VBF}) = \text{Pois}(n|S_{ggF} + S_{VBF} + B) \left[\prod_{i=1}^n \frac{S_{ggF}f_{ggF}(x_i) + S_{VBF}f_{VBF}(x_i) + Bf_B(x_i)}{S_{ggF} + S_{VBF} + B} \right] \quad (7.20)$$

2997 where f_X s are the PDFs of signal and backgrounds modelled in section 7.5 and 7.4, S_X
2998 and B are the normalizations of signal and sum of backgrounds.

2999 The parameters of interest (POI) in this search is σ_{ggF} (and σ_{VBF} only for NWA signal),
3000 presenting the cross section of signal in ggF (and VBF) production mode. In the case of
3001 two POIs, when testing one POI, the other is left unconstrained and profiled with other
3002 nuisance parameters during the minimization. The expected signal yields S_{ggF} and S_{VBF}
3003 in equation 7.20 can be written from these POIs:

$$S_{ggF(VBF)} = \sigma_{ggF(VBF)} \times B(S \rightarrow ZZ) \times A \times C \times \int \mathcal{L} \quad (7.21)$$

3004 where $B(S \rightarrow ZZ)$ is the branching ratio of signal decay to two Z bosons, $A \times C$ is the
3005 parameterized signal acceptance as described in 7.3.5, and $\int \mathcal{L} = 139 \text{ fb}^{-1}$ is the integrated
3006 luminosity of data.

3007 The systematic uncertainties of the expected number of signal and background events
3008 (normalizations) and the shape of the PDFs measured in section 7.6 is described by a
3009 set of nuisance parameters (NPs) θ_i , constrained by Gaussian distribution described in
3010 section 5.4 in fit. Thus, the final likelihood function $L(\sigma_{ggF}, \sigma_{VBF}, m_H, \theta_i)$ is modelled as
3011 a function of several variables: σ_{ggF} , σ_{VBF} , m_H , and θ_i .

3012 Furthermore, the normalization of SM background $pp \rightarrow ZZ$, including both $q\bar{q} \rightarrow$
3013 ZZ and $gg \rightarrow ZZ$, is a free parameter (μ_{ZZ}) and profiled during the minimization.
3014 Given that the increased data luminosity can now provide precise determination of the
3015 SM ZZ background with only about 2% statistical uncertainty, floating ZZ normaliza-
3016 tion in fit reduces the dependence on theory predictions and their corresponding large
3017 uncertainties (about 30%).

3018 In addition, the study on comparison of expected signal sensitivity between fixing
3019 (with Gaussian constraints on its normalization uncertainties) and floating ZZ normal-
3020 ization has been done. Table 7.11 shows the value of expected 95% CLs upper limits un-
3021 der these two cases at three different mass points. The results prove that with the excellent
3022 discrimination power of $m_{4\ell}$ distribution between the SM background and the resonant

Table 7.11 Summary of CL_s upper limits on ggF production mode at 300, 600, 1000 GeV.

$\sigma_{\text{ggF}} \times B(S \rightarrow ZZ) [\text{pb}]$	300 GeV	600 GeV	1000 GeV
Floating ZZ normalization	0.0892	0.0253	0.0125
Fixing ZZ normalization	0.0892	0.0254	0.0125

signal, floating ZZ normalization in fit would not reduce the expected signal sensitivity.

At the end, the upper limit on production cross-section $\sigma_{\text{ggF(VBF)}}$ at a given heavy resonance model is obtained by fixing the mass of signal (m_H) at the desired value, and maximising the likelihood function. The CL_s method described in section 5.3 is used to obtain exclusion limits.

7.7.2 Fit to likelihood function under background-only hypothesis for MVA-based analysis

Both MVA- and cut-based analysis are studied by performing likelihood fit to the (pseudo-) data under the background-only hypothesis and under different signal models. First of all, table 7.14 (table 7.15) summaries the expected and observed number of events in MVA (Cut) -based analysis for region of $m_{4\ell} > 200$ GeV together with their systematic uncertainties after background-only fit. The normalization of ZZ background is taken from data for one category each, as shown in table 7.12 (table 7.13) in MVA (Cut) -based analysis.

Table 7.12 ZZ normalization factor in each category, obtained from a likelihood fit to the data under the background-only hypothesis in MVA-based analysis.

Normalization factor	Fitted value
$\mu_{ZZ}^{\text{ggF-MVA-high}}$	1.07 ± 0.05
$\mu_{ZZ}^{\text{ggF-MVA-low}}$	1.12 ± 0.03
$\mu_{ZZ}^{\text{VBF-MVA-enriched}}$	0.9 ± 0.3

Table 7.13 ZZ normalization factor in each category, obtained from a likelihood fit to the data under the background-only hypothesis in Cut-based analysis.

Normalization factor	Fitted value
$\mu_{ZZ}^{\text{ggF-CBA-enriched}}$	1.10 ± 0.02
$\mu_{ZZ}^{\text{VBF-CBA-enriched}}$	1.1 ± 0.2

Then due to the same background estimation and modelling procedures, as well as the same method of systematic measurements between MVA- and Cut- based analysis, the

Table 7.14 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 7.12, are also taken into account.

Process	VBF-MVA-enriched	ggF-MVA-high			ggF-MVA-low
		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 ± 2.0
$t\bar{t}V, VVV$	0.96 ± 0.10	9.8 ± 0.2	17.5 ± 0.4	7.8 ± 0.2	21.9 ± 0.5
Total background	19 ± 4	284 ± 9	480 ± 16	192 ± 6	2300 ± 48
Observed	19	271	493	191	2301

Table 7.15 Expected and observed numbers of events for $m_{4\ell} > 200$ GeV, together with their systematic uncertainties, for two Cut-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 7.13, are also taken into account.

Process	VBF-CBA-enriched	ggF-CBA-enriched		
		4 μ channel	2e2 μ channel	4e channel
$q\bar{q} \rightarrow ZZ$	48 ± 8	860 ± 18	1360 ± 28	515 ± 11
$gg \rightarrow ZZ$	13 ± 4	114 ± 9	189 ± 14	73 ± 6
ZZ (EW)	10.9 ± 0.9	6.9 ± 0.3	11.1 ± 0.4	4.4 ± 0.2
$Z+jets, t\bar{t}$	0.3 ± 0.1	2.1 ± 0.4	6.7 ± 1.6	3.1 ± 0.4
$t\bar{t}V, VVV$	3.0 ± 0.2	16.3 ± 0.4	26.8 ± 0.6	11.8 ± 0.3
Total background	75 ± 9	1000 ± 18	1593 ± 28	607 ± 11
Observed	75	932	1656	612

following results of background-only fits in this selection are only shown in MVA-based analysis under the model of heavy Higgs resonance with narrow-width as an example. The final results of interpretation in both MVA- and cut- based analysis in different signal models described in section 7.5 will be presented in next section.

The post-fit $m_{4\ell}$ spectrum in each category is shown in figure 7.30.

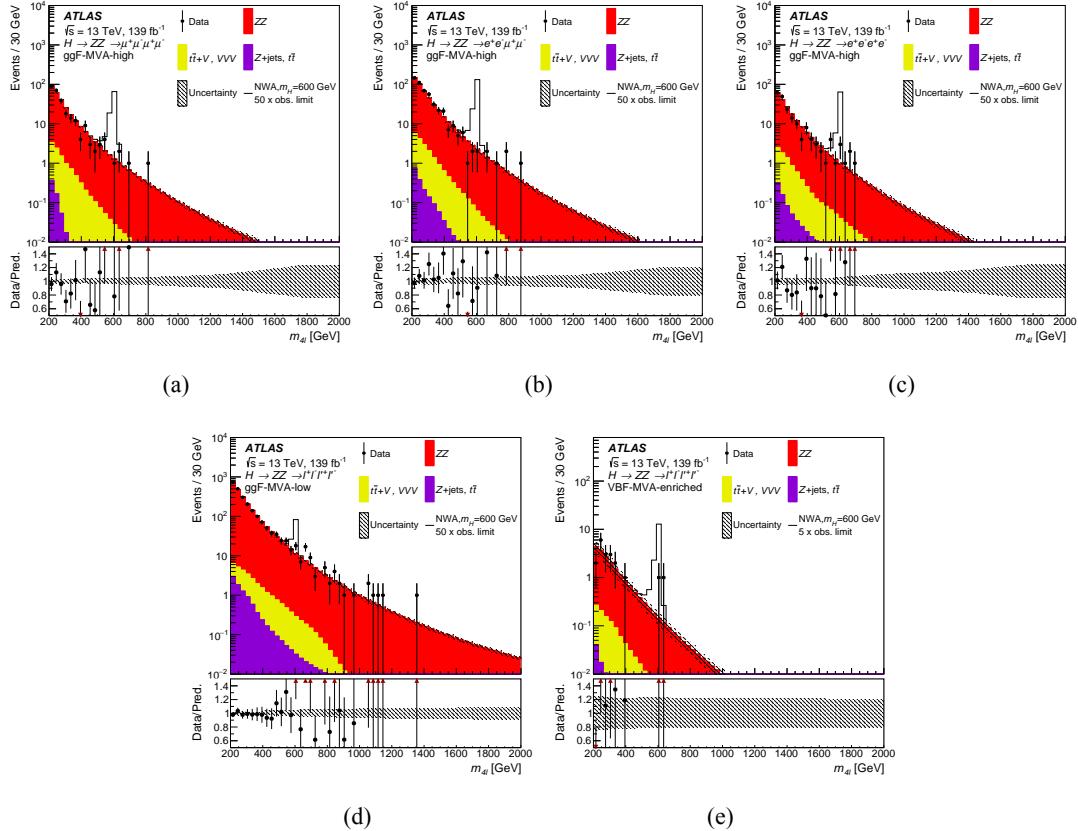


Fig. 7.30 Distribution of the four-lepton invariant mass $m_{4\ell}$ for (a), (b), (c) the ggF-MVA-high categories, (d) the ggF-MVA-low category and (e) the VBF-MVA-enriched category. The backgrounds are determined from a combined likelihood fit to the data under the background-only hypothesis. The simulated signal at 600 GeV is normalized to a cross section corresponding to 50 (5) times the observed upper limit given in section 7.7.3 for the ggF (VBF) production. The error bars on the data points indicate the statistical uncertainty, and the systematic uncertainty in MC prediction is shown by the hatched band. The lower panels show the ratio of data to prediction. The red arrows indicate data points that are outside the displayed range.

To inspect the likelihood model, pulls and constraints as well as the correlation matrix of NPs are studied by performing a background only fit. Figure 7.31 shows the pulls and constraints when fitting to background only Asimov data (top) and observed data (bottom). Figure 7.32 shows the correlation matrix, only for NPs with correlation between each others greater than 0.1 when fitting to background only Asimov data.

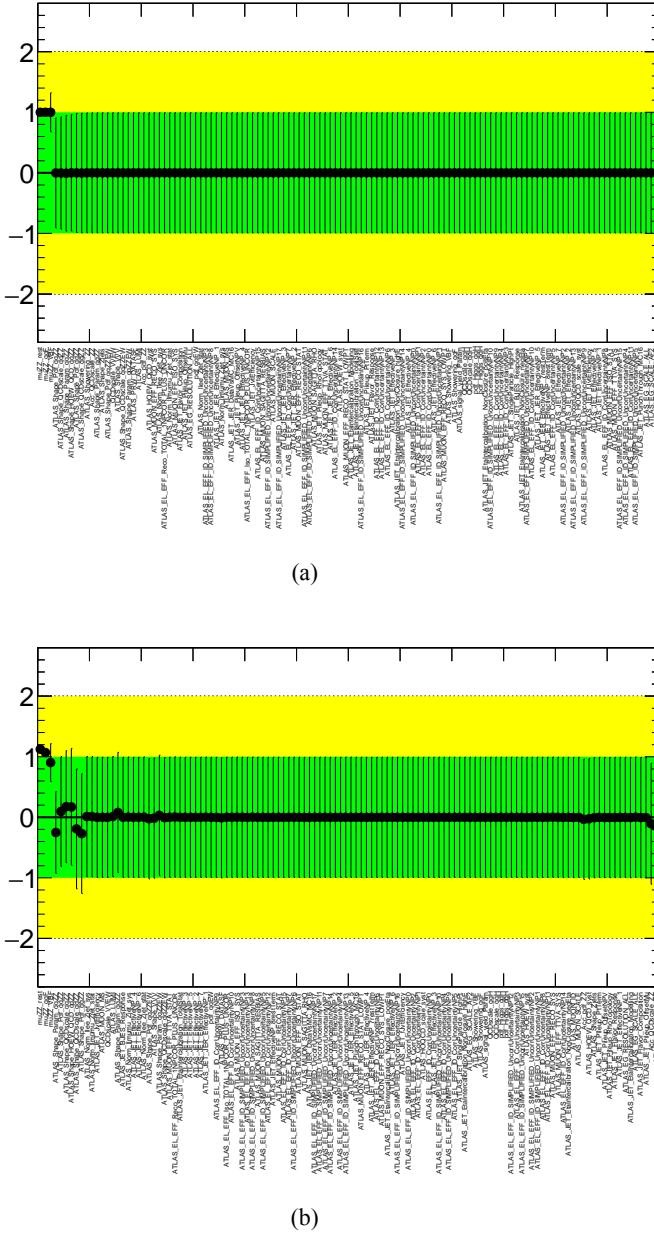
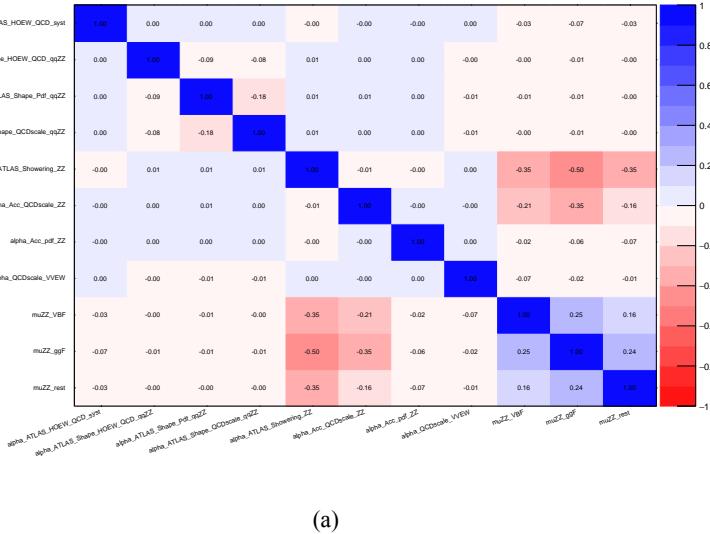


Fig. 7.31 Pulls and constraints of nuisance parameters after a background only fit to (a) Asimov data and (b) observed data in the $\ell\ell\ell'\ell'$ channel. The Asimov data is generated with background data only, and the observed data includes datasets from 2015 to 2018.

3049 The impact of a systematic uncertainty on the result depends on the production mode
 3050 and the mass hypothesis. To check the impact of systematic uncertainties on expected
 3051 signal sensitivity, a NP ranking study is performed using signal injected Asimov data
 3052 with the injected cross section close to 95% CLs upper limit in figure 7.33 at the masses
 3053 of 400 GeV and 1000 GeV. The results are shown in table 7.16. For ggF production,
 3054 at lower mass, the systematic uncertainties of parton showering variation for signal, the
 3055 luminosity uncertainty, and the parametrization of signal acceptance dominate, while at
 3056 higher mass, the shape uncertainties from PDF variation for ZZ ($q\bar{q} \rightarrow ZZ$ and $gg \rightarrow$



(a)

Fig. 7.32 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.

3057 ZZ) background become important, as also seen in VBF production mode. In addition
 3058 for VBF, jet related uncertainties become more important comparing to ggF production.
 3059 Moreover, the dominate uncertainties include the acceptance uncertainty from QCD scale
 3060 variation for signal and the luminosity uncertainty.

3061 7.7.3 Interpretations

3062 1. Spin-0 resonance with NWA

3063 In the absence of a specific model, the ratio of ggF and VBF production mode is un-
 3064 known for this additional heavy scalar. For this reason, the fits for ggF and VBF processes
 3065 are done separately, and in each case the cross section of the untested process is allowed
 3066 to be a free parameter in the statistical fit. The observed and expected upper limit at 95%
 3067 confidence level (CL) on the $\sigma \times B(H \rightarrow ZZ)$ of a narrow scalar resonance for both
 3068 ggF and VBF production mode with the integrated luminosity of 139 fb^{-1} is shown in
 3069 figure 7.33 (7.34) for MVA- (cut-) based analysis.

3070 In order to measure the compatibility of the data with the background-only hypothesis
 3071 ($\sigma_{\text{ggF}} = \sigma_{\text{VBF}} = 0$), the test statistic q_0 , as described in equation 5.9, is used. Figure 7.35
 3072 shows its corresponding p_0 scan for ggF (left) and VBF (right) production as functions of
 3073 m_H in MVA-based analysis. No excess over 2σ is found for ggF production, while for
 3074 VBF production, the maximal local significance is found to be 2.46σ at 620 GeV with
 3075 corresponding global significance of 0.88σ computed from Ref. [130].

Table 7.16 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

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 7.5.3 are token into account. The upper limit at 95% confidence level (CL) on ggF cross section times branch ratio ($\sigma_{ggF} \times B(H \rightarrow ZZ)$) is shown in figure 7.36 for a width of 1, 5, 10 and 15% of m_H .

3. Spin-2 RS Graviton resonance

The observed and expected 95% upper limit on the cross section times branching ratio for RS Graviton (RSG) scenario is shown in figure 7.37. Similar to LWA, only $4e$, 4μ and $2e2\mu$ channel of ggF production mode are used. On top of the expected and observed upper limits in this model, a predicted cross section as function of mass provided

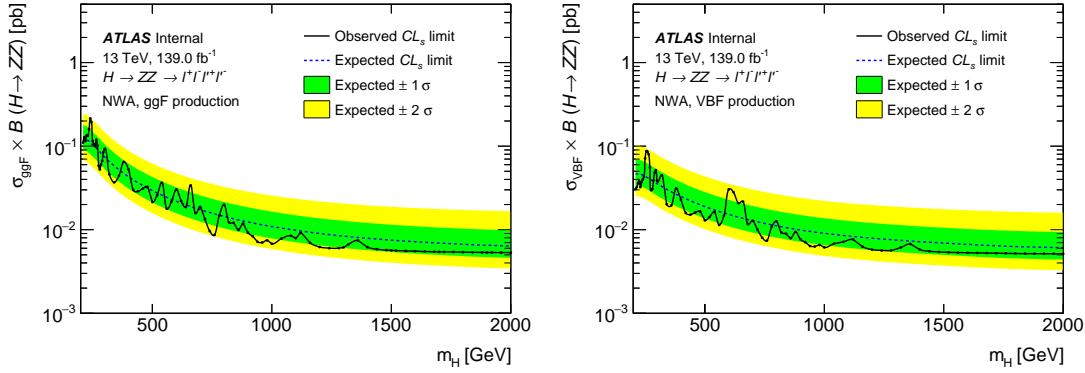


Fig. 7.33 The expected and observed upper limits at 95% CL on $\sigma \times B(H \rightarrow ZZ)$ using the MVA-based analysis for ggF (left) and VBF (right) production. The green and yellow bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties in the expected limits.

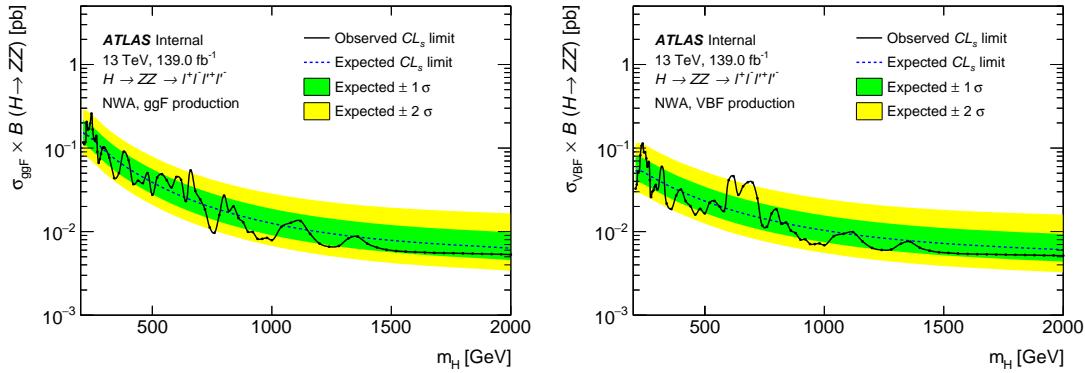


Fig. 7.34 The expected and observed upper limits at 95% CL on $\sigma \times B(H \rightarrow ZZ)$ using the cut-based analysis for ggF (left) and VBF (right) production. The green and yellow bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties in the expected limits.

3088 by theorist is also shown in the figure. Comparing with the observed result provided by
 3089 $ZZ \rightarrow \ell\ell\ell'\ell'$ decay, this spin-2 graviton is excluded up to a mass of 1500 GeV.

3090 4. Summary of interpretation

3091 As a summary, figure 7.38 shows the comparison of expected and observed 95% CL
 3092 upper limits between different models described above.

3093 Figure 7.39 compares the expected 95% CL upper limits as a function of the NWA
 3094 resonance mass in this analysis with full run-2 data and the one in previous publication^[112]
 3095 with the integrated luminosity of 36.1 fb^{-1} . With a significant increase of integrated lu-
 3096 minosity and an improved analysis strategy, comparing to the previous publication, the
 3097 expected sensitivities of searching for narrow-width heavy resonance reduced by up to
 3098 70% in MVA-based analysis, where 50% of reduction is due to luminosity increase while
 3099 other improvement mainly comes from inviting multivariate method.

3100 Figure 7.40 shows the display of one candidate event passing analysis selection in

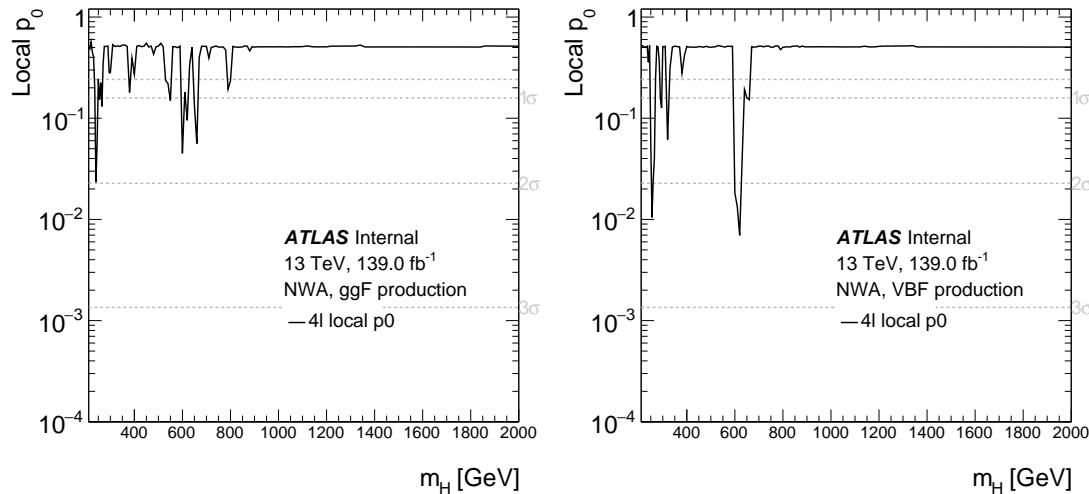


Fig. 7.35 The local p_0 scan as functions of m_H on $\sigma \times B(H \rightarrow ZZ)$ using the MVA-based analysis for ggF (left) and VBF (right) production.

3101 four-muon final state with four-muon invariant mass of 1.34 TeV.

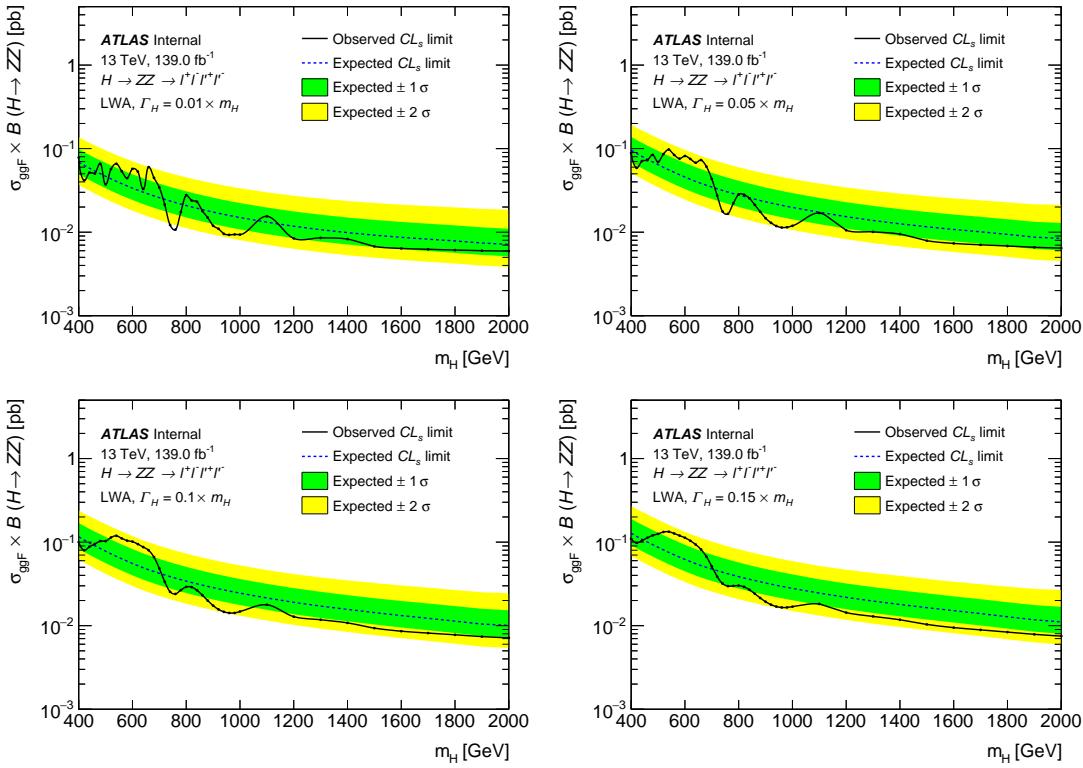


Fig. 7.36 The upper limits at 95% confidence level on $\sigma_{ggF} \times B(H \rightarrow ZZ)$ as a function of the heavy resonance mass m_H for the ggF production mode with an intrinsic width of 1% (top left), 5% (top right), 10% (bottom left) and 15% (bottom right) for both the case where interference 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.

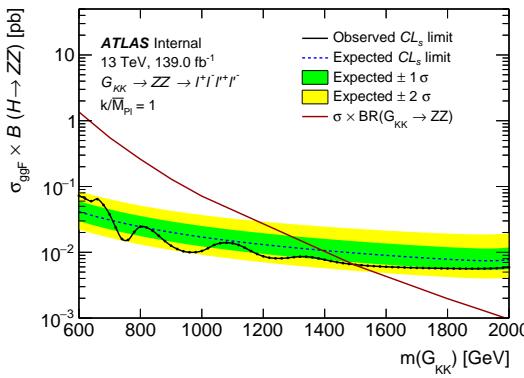


Fig. 7.37 The upper limits at 95% confidence level on $\sigma_{ggF} \times B(G_{KK} \rightarrow ZZ)$ as a function of the heavy resonance mass $m(G_{KK})$ for the ggF production mode in RS Graviton model. The green and yellow bands represent the $\pm 1 \sigma$ and $\pm 2 \sigma$ uncertainties in the expected limits.

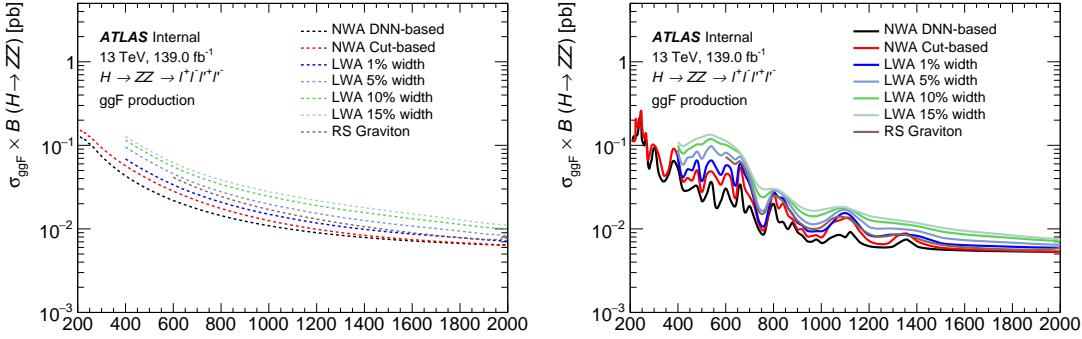


Fig. 7.38 The expected (left) and observed (right) upper limits at 95% CL on $\sigma \times B(S \rightarrow ZZ)$ for ggF production mode at different hypothetical models including the MVA- and cut- based NWA, LWA with width assumptions of 1, 5, 10, 15%, and the RSG.

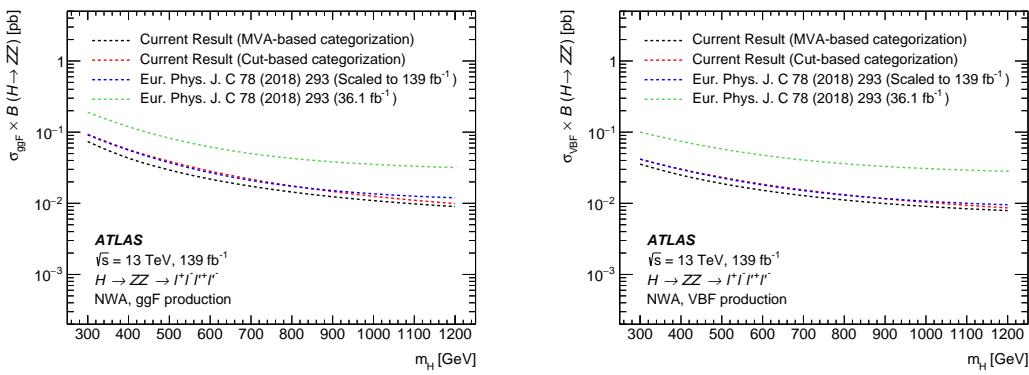


Fig. 7.39 Comparisons of the expected upper limits at 95% CL on the cross section times branching ratio as a function of the heavy resonance mass m_H for the ggF production mode (left) and for the VBF production mode (right) in the case of the NWA. The expected limits from the previous publication are 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.

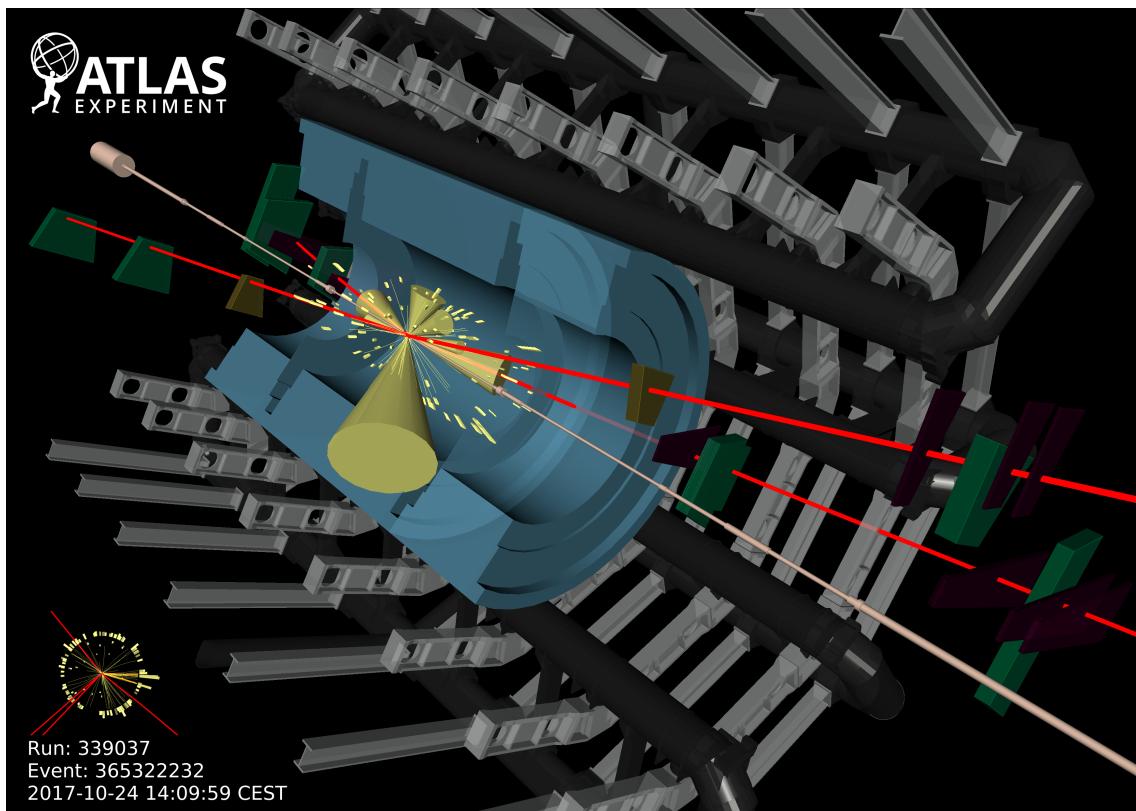


Fig. 7.40 Display of one candidate event in 4μ final state with the mass of 1.35 TeV.

3102 7.8 Conclusion

3103 Searches of heavy resonances decaying into a pair of Z boson to $\ell\ell\ell'\ell'$ final state
3104 are performed using 139 fb^{-1} of 13 TeV pp collision data collected by ATLAS experiment
3105 at the LHC. The results are interpreted as 95% CL upper limits on the production cross
3106 section of a spin-0 and spin-2 resonances under different theoretical models. The search
3107 range of the hypothetical resonances is between 200 GeV to 2000 GeV depending on the
3108 signal model.

3109 The spin-0 resonance is assumed to be a heavy Higgs like scalar produced predom-
3110 inantly from gluon–gluon fusion (ggF) and vector-boson fusion (VBF) decays, and it is
3111 studied under both the narrow-width approximation and with the large-width assump-
3112 tion. For narrow-width approximation, the exclusion limits on cross section of heavy
3113 scalar decaying into two Z bosons are set separately for ggF and VBF production modes,
3114 under MVA- and cut- based analysis. In MVA-based analysis, the 95% CL upper limit
3115 range is from 215 fb at $m_H = 240 \text{ GeV}$ to 5.3 fb at $m_H = 2000 \text{ GeV}$ for ggF produc-
3116 tion mode, and from 87 fb at $m_H = 255 \text{ GeV}$ to 5.1 fb at $m_H = 1960 \text{ GeV}$ for VBF
3117 production mode. In cut-based analysis, the 95% CL upper limit range is from 259 fb at
3118 $m_H = 245 \text{ GeV}$ to 5.3 fb at $m_H = 2000 \text{ GeV}$ for ggF production mode, and from 113 fb
3119 at $m_H = 240 \text{ GeV}$ to 5.1 fb at $m_H = 2000 \text{ GeV}$ for VBF production mode. MVA-based
3120 analysis gains about 20% improvement on upper limits at lower mass region comparing
3121 to the cut-based analysis, while for mass above 1500 GeV, both analyses perform closely.
3122 For large-width approximation, limits are studied on ggF production rate at four different
3123 width assumptions: 1%, 5%, 10% and 15% of resonance’s mass, with the interference
3124 between the heavy scalar and the SM Higgs boson as well as the heavy scalar and the SM
3125 $gg \rightarrow ZZ$ continuum background taken into account. The maximum and minimum of
3126 upper limits are obtained as 78 fb at $m_H = 400 \text{ GeV}$ to 5.9 fb at $m_H = 2000 \text{ GeV}$ for
3127 1% width; 98 fb at $m_H = 540 \text{ GeV}$ to 6.4 fb at $m_H = 2000 \text{ GeV}$ for 5% width; 119 fb at
3128 $m_H = 540 \text{ GeV}$ to 7.1 fb at $m_H = 2000 \text{ GeV}$ for 10% width; 133 fb at $m_H = 540 \text{ GeV}$ to
3129 7.5 fb at $m_H = 2000 \text{ GeV}$ for 15% width. Last but not least, the framework of the Randall–
3130 Sundrum model with a graviton excitation spin-2 resonance with $m(G_{KK}) < 1500 \text{ GeV}$ is
3131 excluded at 95% CL.

3132

Chapter 8 Summary

3133 On December 3rd, 2018, the LHC finished its second run (run-2) after three fantastic
3134 years. Thanks to run-2 with large increase in statistics, we now know the masses of the
3135 Higgs boson, top quark and W boson to considerably greater precision, and also confirm
3136 the Standard Model as a stable theory.

3137 In this dissertation, various physics processes in $ZZ \rightarrow \ell\ell\ell'\ell'$ final state are studied,
3138 taking the advantage of full run-2 pp collision data in the LHC. Using this signature,
3139 we measured the fiducial cross section of ZZ production to $\ell\ell\ell'\ell' jj$ channel in SM,
3140 which is an important physics process and major background in many analysis with ZZ
3141 production, eg. Higgs analysis (HZZ). In addition, we searched the electroweak ZZ
3142 production via vector boson scattering in associated with 2-jet process in $\ell\ell\ell'\ell'$ final
3143 state. In the meantime, the searches of heavy resonances decaying into a pair of Z bosons
3144 to $\ell\ell\ell'\ell'$ final state for several different hypothetical resonances are conducted in this
3145 dissertation. The results of several analyses are summarized as below:

3146 **Measurement of fiducial cross section of ZZ production in $\ell\ell\ell'\ell' jj$ final state**

3147 The fiducial cross section of inclusive SM $ZZ \rightarrow \ell\ell\ell'\ell'$ production is measured to
3148 be:

$$\sigma_{ZZ \rightarrow \ell\ell\ell'\ell'}^{fid} = 1.27 \pm 0.12(stat) \pm 0.02(theo) \pm 0.07(exp) \pm 0.01(bkg) \pm 0.02(lumi) [fb] \quad (8.1)$$

3149 which is found to be compatible with the SM prediction. The ZZ cross section is calcu-
3150 lated with up to one (three) outgoing partons at NLO (LO) using Sherpa 2.2.2 for QCD
3151 production, and in LO using MadGraph5_aMC@NLO 2.6.1 for EW production. The total
3152 uncertainty is 11%, the analysis is still data statistic dominant (data statistic uncertainty is
3153 about 9.5%).

3154 **Observation of electroweak ZZ production in $\ell\ell\ell'\ell' jj$ final state**

3155 Thanks to the largely increased data statistic collected by ATLAS experiment in the
3156 LHC run-2, the electroweak ZZ production (EW- $ZZjj$) to $\ell\ell\ell'\ell'$ channel in associa-
3157 tion with two jets is observed with a significant deviation from the background-only
3158 hypothesis. The signal strength of EW- $ZZjj$ production, the normalization of QCD-
3159 $ZZjj$ production, as well as the observed and expected statistical significance measured

3160 in $\ell\ell\ell'\ell' jj$ channel are found to be:

$$\begin{aligned}\mu_{\text{EW}} &= 1.54 \pm 0.42 \\ \mu_{\text{QCD}} &= 0.95 \pm 0.22\end{aligned}\quad (8.2)$$

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

3161 Then in this dissertation, the differential cross section and expected significance of
 3162 EW- $ZZjj$ production, using 3000 fb^{-1} simulated pp collision data at a centre-of-mass
 3163 energy of 14 TeV to be recorded by ATLAS experiment at the HL-LHC, are studied via
 3164 simulations. The HL-LHC will for sure give us more opportunity to probe rare process
 3165 like $ZZ \rightarrow \ell\ell\ell'\ell'$ in the future.

3166 **Searches of heavy ZZ resonances in $\ell\ell\ell'\ell'$ final state**

3167 Searches of heavy ZZ resonances are performed in four-lepton invariant mass $m_{4\ell}$
 3168 range from 200 GeV to 2000 GeV. Data are found to agree with the background-only
 3169 hypothesis, and 95% CL upper limits are set on the production rate under the models of:

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

3171 Search range is from 200 GeV to 2000 GeV.

3172 In DNN-based analysis, the limits range from 215 fb at $m_H = 240$ GeV to 5.3 fb
 3173 at $m_H = 2000$ GeV for ggF production mode, and from 87 fb at $m_H = 255$ GeV to
 3174 5.1 fb at $m_H = 1960$ GeV for VBF production mode.

3175 The DNN-based analysis is found to be at most 20% better than cut-based results.

- Spin-0 heavy Higgs under large-width approximation (LWA) with the width of 1,
 3177 5, 10, 15% of its mass.

3178 Search range is from 400 GeV to 2000 GeV, and only ggF production is studied.

3179 The maximum and minimum of upper limits are obtained as 78 fb at $m_H = 400$ GeV
 3180 to 5.9 fb at $m_H = 2000$ GeV for 1% width; 98 fb at $m_H = 540$ GeV to 6.4 fb at $m_H =$
 3181 2000 GeV for 5% width; 119 fb at $m_H = 540$ GeV to 7.1 fb at $m_H = 2000$ GeV for
 3182 10% width; 133 fb at $m_H = 540$ GeV to 7.5 fb at $m_H = 2000$ GeV for 15% width.

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

3184 Search range is from 600 GeV to 2000 GeV, and only ggF production is studied.

3185 The maximum and minimum of limits are 73 fb at $m_H = 600$ GeV and 5.6 fb
 3186 at $m_H = 1880$ GeV for ggF production mode. And the mass of graviton below
 3187 1500 GeV is excluded comparing the observed results with theoretical prediction.

3188

3189 In summary, the $ZZ \rightarrow \ell\ell\ell'\ell'$ production presented in this dissertation are
 3190 consistent with SM prediction. This result completes the observation of weak boson

3191 scattering, which is a new milestone in the study of electroweak symmetry breaking,
3192 and a starting point towards the future longitudinal polarization study in the VBS ZZ
3193 processes directly relating to the Higgs unitarization mechanism in VV scatterings. In
3194 the meantime, no indication of new physics is observed. We are looking forward to the
3195 HL-LHC, with greatly increased luminosity and higher centre-of-mass energy, which
3196 should enhance the sensitivity for new physics search and precise measurement for rare
3197 process like $\ell\ell\ell'\ell'$ final state.

3198

Bibliography

- 3199 [1] F.R.S. J J T M. XI. cathode rays[J/OL]. The London, Edinburgh, and Dublin Philosophical
3200 Magazine and Journal of Science, 1897, 44(269):293-316. DOI: [10.1080/14786449708621070](https://doi.org/10.1080/14786449708621070).
- 3201 [2] MILLIKAN. R A. On the elementary electrical charge and the avogadro constant[J/OL]. Phys.
3202 Rev., 1913, 2:109-143. DOI: [10.1103/PhysRev.2.109](https://doi.org/10.1103/PhysRev.2.109).
- 3203 [3] F.R.S. P E R. Lxxix. the scattering of α and β particles by matter and the structure of the atom
3204 [J/OL]. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science,
3205 1911, 21(125):669-688. DOI: [10.1080/14786440508637080](https://doi.org/10.1080/14786440508637080).
- 3206 [4] LANGACKER P. Introduction to the Standard Model and Electroweak Physics[C/OL]//
3207 Proceedings of Theoretical Advanced Study Institute in Elementary Particle Physics on The
3208 dawn of the LHC era (TASI 2008): Boulder, USA, June 2-27, 2008. 2010: 3-48. DOI:
3209 [10.1142/9789812838360_0001](https://doi.org/10.1142/9789812838360_0001).
- 3210 [5] PICH A. Electroweak Symmetry Breaking and the Higgs Boson[J/OL]. Acta Phys. Polon.,
3211 2016, B47:151. DOI: [10.5506/APhysPolB.47.151](https://doi.org/10.5506/APhysPolB.47.151).
- 3212 [6] Observation of a new particle in the search for the standard model higgs boson with the atlas
3213 detector at the lhc[J/OL]. Physics Letters B, 2012, 716(1):1 - 29. DOI: [10.1016/j.physletb.2012.08.020](https://doi.org/10.1016/j.physletb.2012.08.020).
- 3215 [7] Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc[J/OL].
3216 Physics Letters B, 2012, 716(1):30 - 61. DOI: [10.1016/j.physletb.2012.08.021](https://doi.org/10.1016/j.physletb.2012.08.021).
- 3217 [8] PROFUMO S, RAMSEY-MUSOLF M J, SHAUGHNESSY G. Singlet higgs phenomenology
3218 and the electroweak phase transition[J/OL]. Journal of High Energy Physics, 2007, 2007(08):
3219 010-010. DOI: [10.1088/1126-6708/2007/08/010](https://doi.org/10.1088/1126-6708/2007/08/010).
- 3220 [9] BRANCO G, FERREIRA P, LAVOURA L, et al. Theory and phenomenology of two-higgs-
3221 doublet models[J/OL]. Physics Reports, 2012, 516(1):1 - 102. DOI: [10.1016/j.physrep.2012.02.002](https://doi.org/10.1016/j.physrep.2012.02.002).
- 3223 [10] BRUNING O S, COLLIER P, LEBRUN P, et al. LHC Design Report Vol.1: The LHC Main
3224 Ring[J]. 2004.
- 3225 [11] BUNING O, COLLIER P, LEBRUN P, et al. LHC Design Report. 2. The LHC infrastructure
3226 and general services[J]. 2004.
- 3227 [12] BENEDIKT M, COLLIER P, MERTENS V, et al. LHC Design Report. 3. The LHC injector
3228 chain[J]. 2004.
- 3229 [13] DREMIN I M. Soft and hard processes in QCD[J/OL]. JETP Lett., 2005, 81:307-310. DOI:
3230 [10.1134/1.1944068](https://doi.org/10.1134/1.1944068).

Bibliography

- 3231 [14] WOMERSLEY J. QCD at the Tevatron: Status and prospects[C]//Proceedings, 5th Interna-
3232 tional Symposium on Radiative Corrections - RADCOR 2000. 2000.
- 3233 [15] LIN H W, et al. Parton distributions and lattice QCD calculations: a community white paper
3234 [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).
- 3235 [16] COLLINS J C, SOPER D E, STERMAN G F. Factorization of Hard Processes in QCD[J/OL].
3236 Adv. Ser. Direct. High Energy Phys., 1989, 5:1-91. DOI: [10.1142/9789814503266_0001](https://doi.org/10.1142/9789814503266_0001).
- 3237 [17] STIRLING W J. Perturbative QCD[J/OL]. 2000(CERN-OPEN-2000-296):40 p. <https://cds.cern.ch/record/1194745>.
- 3239 [18] GROJEAN C. Higgs Physics[J/OL]. 2017(arXiv:1708.00794):143-158. 12 p. <https://cds.cern.ch/record/2243593>.
- 3241 [19] Cern yellow reports: Monographs: Handbook of LHC Higgs Cross Sections: 4. Deciphering
3242 the Nature of the Higgs Sector[M/OL]. 2016. <https://cds.cern.ch/record/2227475>.
- 3243 [20] HEINEMEYER S. Cern yellow reports: Monographs: Handbook of LHC Higgs Cross Sec-
3244 tions: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group[M/OL].
3245 2013. <https://cds.cern.ch/record/1559921>.
- 3246 [21] CHANG J, CHEUNG K, LU C T, et al. *ww* scattering in the era of post-higgs-boson discovery
3247 [J/OL]. Phys. Rev. D, 2013, 87:093005. DOI: [10.1103/PhysRevD.87.093005](https://doi.org/10.1103/PhysRevD.87.093005).
- 3248 [22] EVANS L, BRYANT P. LHC machine[J/OL]. Journal of Instrumentation, 2008, 3(08):S08001-
3249 S08001. DOI: [10.1088/1748-0221/3/08/s08001](https://doi.org/10.1088/1748-0221/3/08/s08001).
- 3250 [23] MOBS E. The CERN accelerator complex. Complexe des accélérateurs du CERN[J/OL]. 2016.
3251 <https://cds.cern.ch/record/2197559>.
- 3252 [24] COLLABORATION A. The ATLAS experiment at the CERN large hadron collider[J/OL].
3253 Journal of Instrumentation, 2008, 3(08):S08003-S08003. DOI: [10.1088/1748-0221/3/08/s08003](https://doi.org/10.1088/1748-0221/3/08/s08003).
- 3255 [25] PEREZ G. Unitarization models for vector boson scattering at the lhc[D/OL]. 2018. DOI:
3256 [10.5445/IR/1000082199](https://doi.org/10.5445/IR/1000082199).
- 3257 [26] PEQUENAO J. Computer generated image of the whole ATLAS detector[Z/OL]. 2008. <https://cds.cern.ch/record/1095924>.
- 3259 [27] MCFAYDEN J. The lhc and atlas detector. in: Third generation susy and $t^-t^+ + z$ productin.
3260 [D/OL]. 2014. DOI: [10.1007/978-3-319-07191-6_2](https://doi.org/10.1007/978-3-319-07191-6_2).
- 3261 [28] COLLABRATION A. Operation and performance of the ATLAS semiconductor tracker. Op-
3262 eration and performance of the ATLAS semiconductor tracker[J/OL]. JINST, 2014, 9(CERN-
3263 PH-EP-2014-049. CERN-PH-EP-2014-049):P08009. 80 p. <https://cds.cern.ch/record/1698966>.
- 3264 [29] ATLAS pixel detector electronics and sensors[J/OL]. Journal of Instrumentation, 2008, 3(07):
3265 P07007-P07007. DOI: [10.1088/1748-0221/3/07/p07007](https://doi.org/10.1088/1748-0221/3/07/p07007).

- 3266 [30] MULLIER G. The upgraded pixel detector of the atlas experiment for run-2 at the large hadron
3267 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).
- 3269 [31] The silicon microstrip sensors of the atlas semiconductor tracker[J/OL]. Nuclear Instruments
3270 and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Asso-
3271 ciated Equipment, 2007, 578(1):98 - 118. DOI: [10.1016/j.nima.2007.04.157](https://doi.org/10.1016/j.nima.2007.04.157).
- 3272 [32] SULTAN D M S. Development of small-pitch, thin 3d sensors for pixel detector upgrades at
3273 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).
- 3274 [33] The ATLAS transition radiation tracker (TRT) proportional drift tube: design and performance
3275 [J/OL]. Journal of Instrumentation, 2008, 3(02):P02013-P02013. DOI: [10.1088/1748-0221/3/02/p02013](https://doi.org/10.1088/1748-0221/3/02/p02013).
- 3277 [34] BUCHANAN N, CHEN L, GINGRICH D, et al. Design and implementation of the front end
3278 board for the readout of the atlas liquid argon calorimeters[J/OL]. Journal of Instrumentation,
3279 2008, 3:P03004. DOI: [10.1088/1748-0221/3/03/P03004](https://doi.org/10.1088/1748-0221/3/03/P03004).
- 3280 [35] Technical design report atlas: ATLAS liquid-argon calorimeter: Technical Design Report
3281 [M/OL]. Geneva: CERN, 1996. <https://cds.cern.ch/record/331061>.
- 3282 [36] AAD G, ABBOTT B, ABDALLAH J, et al. Readiness of the atlas tile calorimeter for lhc
3283 collisions[J]. European Physical Journal C, 2010.
- 3284 [37] Technical design report atlas: ATLAS muon spectrometer: Technical Design Report[M/OL].
3285 Geneva: CERN, 1997. <https://cds.cern.ch/record/331068>.
- 3286 [38] SLIWA K. "ATLAS Overview and Main Results"[C]//Proceedings, International School on
3287 High Energy Physics : Workshop on High Energy Physics in the near Future. (LISHEP 2013):
3288 Rio de Janeiro, Brazil, March 17-24, 2013. 2013.
- 3289 [39] RUIZ-MARTINEZ A, COLLABORATION A. The Run-2 ATLAS Trigger System: ATL-
3290 DAQ-PROC-2016-003[R/OL]. Geneva: CERN, 2016. <https://cds.cern.ch/record/2133909>.
- 3291 [40] PÁSZTOR G. The Upgrade of the ATLAS Electron and Photon Triggers towards LHC Run 2 and
3292 their Performance: ATL-DAQ-PROC-2015-053. arXiv:1511.00334[R/OL]. Geneva: CERN,
3293 2015. <https://cds.cern.ch/record/2063746>.
- 3294 [41] COLLABORATION A. Atlas computing: Technical design report[R]. Geneva: CERN.
- 3295 [42] AAD G, et al. The ATLAS Simulation Infrastructure[J/OL]. Eur. Phys. J., 2010, C70:823-874.
3296 DOI: [10.1140/epjc/s10052-010-1429-9](https://doi.org/10.1140/epjc/s10052-010-1429-9).
- 3297 [43] HOCHÉ S. Introduction to parton-shower event generators[C/OL]//Proceedings, Theoretical
3298 Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Fron-
3299 tier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014. 2015: 235-
3300 295. DOI: [10.1142/9789814678766_0005](https://doi.org/10.1142/9789814678766_0005).

- 3301 [44] GLEISBERG T, HöCHE S, KRAUSS F, et al. Event generation with SHERPA 1.1[J/OL].
3302 Journal of High Energy Physics, 2009, 2009(02):007-007. DOI: [10.1088/1126-6708/2009/02/007](https://doi.org/10.1088/1126-6708/2009/02/007).
- 3304 [45] BÄHR M, GIESEKE S, GIGG M A, et al. Herwig++ physics and manual[J/OL]. The European
3305 Physical Journal C, 2008, 58(4):639-707. DOI: [10.1140/epjc/s10052-008-0798-9](https://doi.org/10.1140/epjc/s10052-008-0798-9).
- 3306 [46] NASON P. A New method for combining NLO QCD with shower Monte Carlo algorithms
3307 [J/OL]. JHEP, 2004, 11:040. DOI: [10.1088/1126-6708/2004/11/040](https://doi.org/10.1088/1126-6708/2004/11/040).
- 3308 [47] FRIXIONE S, WEBBER B R. Matching NLO QCD computations and parton shower simula-
3309 tions[J/OL]. Journal of High Energy Physics, 2002, 2002(06):029-029. DOI: [10.1088/1126-6708/2002/06/029](https://doi.org/10.1088/1126-6708/2002/06/029).
- 3311 [48] SJOSTRAND T, MRENNAN S, SKANDS P Z. A Brief Introduction to PYTHIA 8.1[J/OL].
3312 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).
- 3313 [49] BARBERIO E, et al. The Geant4-Based ATLAS Fast Electromagnetic Shower Simulation
3314 [C/OL]//Astroparticle, particle and space physics, detectors and medical physics applications.
3315 Proceedings, 10th Conference, ICATPP 2007, Como, Italy, October 8-12, 2007. 2008: 802-806.
3316 DOI: [10.1142/9789812819093_0133](https://doi.org/10.1142/9789812819093_0133).
- 3317 [50] RICHTER-WAS E, FROIDEVAUX D, POGGIOLI L. ATLFAST 2.0 a fast simulation package
3318 for ATLAS: ATL-PHYS-98-131[R/OL]. Geneva: CERN, 1998. <https://cds.cern.ch/record/683751>.
- 3320 [51] EDMONDS K, FLEISCHMANN S, LENZ T, et al. The Fast ATLAS Track Simulation
3321 (FATRAS): ATL-SOFT-PUB-2008-001. ATL-COM-SOFT-2008-002[R/OL]. Geneva: CERN,
3322 2008. <https://cds.cern.ch/record/1091969>.
- 3323 [52] BOYD J T, COSTA M J, TONOVAN A, et al. Commissioning of the ATLAS reconstruction
3324 software with first data[J/OL]. Journal of Physics: Conference Series, 2010, 219(3):032059.
3325 DOI: [10.1088/1742-6596/219/3/032059](https://doi.org/10.1088/1742-6596/219/3/032059).
- 3326 [53] CORNELISSEN T, ELSING M, FLEISCHMANN S, et al. Concepts, Design and Implemen-
3327 tation of the ATLAS New Tracking (NEWT): ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-
3328 2007-002[R/OL]. Geneva: CERN, 2007. <https://cds.cern.ch/record/1020106>.
- 3329 [54] AAD G, et al. Muon reconstruction performance of the ATLAS detector in proton-proton col-
3330 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).
- 3332 [55] ILLINGWORTH J, KITTNER J. A survey of the hough transform[J/OL]. Computer Vision,
3333 Graphics, and Image Processing, 1988, 44(1):87 - 116. DOI: [10.1016/S0734-189X\(88\)80033-1](https://doi.org/10.1016/S0734-189X(88)80033-1).
- 3334 [56] Vertex Reconstruction Performance of the ATLAS Detector at 13 TeV: ATL-PHYS-PUB-2015-
3335 026[R/OL]. Geneva: CERN, 2015. <https://cds.cern.ch/record/2037717>.

Bibliography

- 3336 [57] AABOUD M, et al. Reconstruction of primary vertices at the ATLAS experiment in Run 1
3337 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).
- 3339 [58] Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton
3340 collision data: ATLAS-CONF-2016-024[R/OL]. Geneva: CERN, 2016. <https://cds.cern.ch/reCORD/2157687>.
- 3342 [59] LAMPL W, LAPLACE S, LELAS D, et al. Calorimeter Clustering Algorithms: Description
3343 and Performance: ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003[R/OL]. Geneva:
3344 CERN, 2008. <https://cds.cern.ch/record/1099735>.
- 3345 [60] CORNELISSEN T G, ELSING M, GAVRILENKO I, et al. The global χ^2 track fitter in ATLAS
3346 [J/OL]. Journal of Physics: Conference Series, 2008, 119(3):032013. DOI: [10.1088/1742-6596/119/3/032013](https://doi.org/10.1088/1742-6596/119/3/032013).
- 3348 [61] Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton
3349 collision data: ATLAS-CONF-2014-032[R/OL]. Geneva: CERN, 2014. <http://cds.cern.ch/reCORD/1706245>.
- 3351 [62] LIMPER M. Track and vertex reconstruction in the ATLAS inner detector[D/OL]. 2009. <http://cds.cern.ch/record/1202457>.
- 3353 [63] AAD G, et al. Topological cell clustering in the ATLAS calorimeters and its performance in
3354 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).
- 3355 [64] ZHENG Z. Identification of very-low transverse momentum muons with the ATLAS experi-
3356 ment: ATL-MUON-PROC-2018-018[R/OL]. Geneva: CERN, 2018. <https://cds.cern.ch/reCORD/2649299>.
- 3358 [65] CACCIARI M, SALAM G P. Pileup subtraction using jet areas[J/OL]. Physics Letters B, 2008,
3359 659(1):119 - 126. DOI: [10.1016/j.physletb.2007.09.077](https://doi.org/10.1016/j.physletb.2007.09.077).
- 3360 [66] CACCIARI M, SALAM G P, SOYEZ G. The anti-ktjet clustering algorithm[J/OL]. Journal of
3361 High Energy Physics, 2008, 2008(04):063-063. DOI: [10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063).
- 3362 [67] CACCIARI M, SALAM G P, SOYEZ G. Fastjet user manual[J/OL]. The European Physical
3363 Journal C, 2012, 72(3):1896. DOI: [10.1140/epjc/s10052-012-1896-2](https://doi.org/10.1140/epjc/s10052-012-1896-2).
- 3364 [68] AAD G, et al. Jet energy measurement and its systematic uncertainty in proton-proton collisions
3365 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).
- 3367 [69] AABOUD M, et al. Jet energy scale measurements and their systematic uncertainties in proton-
3368 proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector[J/OL]. Phys. Rev., 2017, D96(7):
3369 072002. DOI: [10.1103/PhysRevD.96.072002](https://doi.org/10.1103/PhysRevD.96.072002).
- 3370 [70] Optimisation of the ATLAS b -tagging performance for the 2016 LHC Run: ATL-PHYS-PUB-

Bibliography

- 3371 2016-012[R/OL]. Geneva: CERN, 2016. <https://cds.cern.ch/record/2160731>.
- 3372 [71] PIACQUADIO G, WEISER C. A new inclusive secondary vertex algorithm for b-jet tagging
3373 in ATLAS[J/OL]. Journal of Physics: Conference Series, 2008, 119(3):032032. DOI: [10.1088/1742-6596/119/3/032032](https://doi.org/10.1088/1742-6596/119/3/032032).
- 3375 [72] SPECKMAYER P, HöCKER A, STELZER J, et al. The toolkit for multivariate data analysis,
3376 TMVA 4[J/OL]. Journal of Physics: Conference Series, 2010, 219(3):032057. DOI: [10.1088/1742-6596/219/3/032057](https://doi.org/10.1088/1742-6596/219/3/032057).
- 3378 [73] Performance of missing transverse momentum reconstruction with the atlas detector using
3379 proton–proton collisions at $\sqrt{s}=13\text{ TeV}$ [J/OL]. The European Physical Journal C, 2018, 78(11):903. DOI: [10.1140/epjc/s10052-018-6288-9](https://doi.org/10.1140/epjc/s10052-018-6288-9).
- 3381 [74] COWAN G, CRANMER K, GROSS E, et al. Asymptotic formulae for likelihood-based tests
3382 of new physics[J/OL]. Eur. Phys. J. C, 2011, 71:1554. DOI: [10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0).
- 3383 [75] READ A L. Presentation of search results: theCLstechnique[J/OL]. Journal of Physics G:
3384 Nuclear and Particle Physics, 2002, 28(10):2693-2704. DOI: [10.1088/0954-3899/28/10/313](https://doi.org/10.1088/0954-3899/28/10/313).
- 3385 [76] BELLOS P. Background estimation, search for heavy resonances and off-shell Higgs boson
3386 signal strength measurement in the 4 lepton final state with the ATLAS detector[EB/OL]. 2020.
3387 <http://cds.cern.ch/record/2725027>.
- 3388 [77] LEE B W, QUIGG C, THACKER H B. The Strength of Weak Interactions at Very High-
3389 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).
- 3391 [78] CHANOWITZ M S, GAILLARD M K. The TeV Physics of Strongly Interacting W's and Z's
3392 [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).
- 3393 [79] SZLEPER M. The Higgs boson and the physics of WW scattering before and after Higgs
3394 discovery[J]. 2014.
- 3395 [80] Evidence for electroweak production of $W^\pm W^\pm jj$ in pp collisions at $\sqrt{s} = 8$ TeV with the
3396 atlas detector[J/OL]. Phys. Rev. Lett., 2014, 113:141803. DOI: [10.1103/PhysRevLett.113.141803](https://doi.org/10.1103/PhysRevLett.113.141803).
- 3398 [81] Observation of electroweak production of a same-sign w boson pair in association with two jets
3399 in pp collisions at $\sqrt{s} = 13$ TeV with the atlas detector[J/OL]. Phys. Rev. Lett., 2019, 123:
3400 161801. DOI: [10.1103/PhysRevLett.123.161801](https://doi.org/10.1103/PhysRevLett.123.161801).
- 3401 [82] SIRUNYAN A M, et al. Observation of electroweak production of same-sign W boson pairs
3402 in the two jet and two same-sign lepton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV
3403 [J/OL]. Phys. Rev. Lett., 2018, 120(8):081801. DOI: [10.1103/PhysRevLett.120.081801](https://doi.org/10.1103/PhysRevLett.120.081801).
- 3404 [83] Observation of electroweak $w\pm z$ boson pair production in association with two jets in pp colli-
3405 sions at $s=13$ tev with the atlas detector[J/OL]. Physics Letters B, 2019, 793:469 - 492. DOI:

- 3406 [84] Measurement of vector boson scattering and constraints on anomalous quartic couplings from
3407 events with four leptons and two jets in proton–proton collisions at s=13 tev[J/OL]. Physics
3408 Letters B, 2017, 774:682 - 705. DOI: [10.1016/j.physletb.2017.10.020](https://doi.org/10.1016/j.physletb.2017.10.020).
- 3410 [85] AAD G, et al. Observation of electroweak production of two jets and a Z -boson pair with the
3411 ATLAS detector at the LHC[J]. 2020.
- 3412 [86] ZHU H. Measurement of inclusive cross section and observation of electroweak production
3413 of two jets in association with a Z -boson pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS
3414 detector: ATL-PHYS-PROC-2020-032[R/OL]. Geneva: CERN, 2020. <https://cds.cern.ch/reco>
3415 [rd/2714092](https://cds.cern.ch/record/2714092).
- 3416 [87] ALWALL J, FREDERIX R, FRIXIONE S, et al. The automated computation of tree-level and
3417 next-to-leading order differential cross sections, and their matching to parton shower simula-
3418 tions[J/OL]. JHEP, 2014, 07:079. DOI: [10.1007/JHEP07\(2014\)079](https://doi.org/10.1007/JHEP07(2014)079).
- 3419 [88] BALL R D, et al. Parton distributions with LHC data[J/OL]. Nucl. Phys. B, 2013, 867:244-289.
3420 DOI: [10.1016/j.nuclphysb.2012.10.003](https://doi.org/10.1016/j.nuclphysb.2012.10.003).
- 3421 [89] GLEISBERG T, HÖCHE S, KRAUSS F, et al. Event generation with SHERPA 1.1[J/OL].
3422 JHEP, 2009, 02:007. DOI: [10.1088/1126-6708/2009/02/007](https://doi.org/10.1088/1126-6708/2009/02/007).
- 3423 [90] BALL R D, et al. Parton distributions for the LHC run II[J/OL]. JHEP, 2015, 04:040. DOI:
3424 [10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040).
- 3425 [91] CAOLA F, MELNIKOV K, RÖNTSCH R, et al. Qcd corrections to zz production in gluon
3426 fusion at the lhc[J/OL]. Phys. Rev. D, 2015, 92:094028. DOI: [10.1103/PhysRevD.92.094028](https://doi.org/10.1103/PhysRevD.92.094028).
- 3427 [92] FRIXIONE S, RIDOLFI G, NASON P. A positive-weight next-to-leading-order Monte Carlo
3428 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).
- 3430 [93] LAI H L, et al. New parton distributions for collider physics[J/OL]. Phys. Rev. D, 2010, 82:
3431 074024. DOI: [10.1103/PhysRevD.82.074024](https://doi.org/10.1103/PhysRevD.82.074024).
- 3432 [94] ALIOLI S, NASON P, OLEARI C, et al. NLO single-top production matched with shower in
3433 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).
- 3435 [95] FREDERIX R, RE E, TORRIELLI P. Single-top t-channel hadroproduction in the four-flavour
3436 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).
- 3438 [96] RE E. Single-top Wt-channel production matched with parton showers using the POWHEG
3439 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).
- 3440 [97] COLLABORATION A. ATLAS Pythia 8 tunes to 7 TeV data[M/OL]. 2014. <https://cds.cern.c>

- 3441 [h/record/1966419](#).
- 3442 [98] ATLAS Collaboration. Performance of pile-up mitigation techniques for jets in pp collisions at
3443 $\sqrt{s} = 8$ TeV using the ATLAS detector[J/OL]. Eur. Phys. J. C, 2016, 76:581. DOI: [10.1140/epjc/s10052-016-4395-z](#).
- 3445 [99] BELLM J, et al. Herwig 7.0/Herwig++ 3.0 release note[J/OL]. Eur. Phys. J. C, 2016, 76(4):
3446 196. DOI: [10.1140/epjc/s10052-016-4018-8](#).
- 3447 [100] BAHR M, et al. Herwig++ Physics and Manual[J/OL]. Eur. Phys. J. C, 2008, 58:639-707. DOI:
3448 [10.1140/epjc/s10052-008-0798-9](#).
- 3449 [101] Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the
3450 LHC: ATLAS-CONF-2019-021[R/OL]. Geneva: CERN, 2019. <http://cds.cern.ch/record/2677054>.
- 3452 [102] The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS[J/OL]. Journal
3453 of Instrumentation, 2018, 13(07):P07017-P07017. DOI: [10.1088/1748-0221/13/07/p07017](#).
- 3454 [103] COADOU Y. Boosted decision trees and applications[J/OL]. EPJ Web of Conferences, 2013,
3455 55:02004-. DOI: [10.1051/epjconf/20135502004](#).
- 3456 [104] COLLABORATION T A. Observation of electroweak production of two jets in association
3457 with a Z -boson pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector[J]. 2019.
- 3458 [105] Prospect study of electroweak production of a Z boson pair plus two jets at the HL-LHC: ATL-
3459 PHYS-PUB-2018-029[R/OL]. Geneva: CERN, 2018. <http://cds.cern.ch/record/2647219>.
- 3460 [106] COLLABORATION A. Technical Design Report for the ATLAS Inner Tracker Pixel Detector:
3461 CERN-LHCC-2017-021. ATLAS-TDR-030[R/OL]. Geneva: CERN, 2017. <https://cds.cern.ch/record/2285585>.
- 3463 [107] COLLABORATION A. Technical Design Report for the Phase-II Upgrade of the ATLAS Muon
3464 Spectrometer: CERN-LHCC-2017-017. ATLAS-TDR-026[R/OL]. Geneva: CERN, 2017. <https://cds.cern.ch/record/2285580>.
- 3466 [108] COLLABORATION A. Technical Proposal: A High-Granularity Timing Detector for the AT-
3467 LAS Phase-II Upgrade: CERN-LHCC-2018-023. LHCC-P-012[R/OL]. Geneva: CERN, 2018.
3468 <https://cds.cern.ch/record/2623663>.
- 3469 [109] Expected performance for an upgraded ATLAS detector at High-Luminosity LHC: ATL-PHYS-
3470 PUB-2016-026[R/OL]. Geneva: CERN, 2016. <https://cds.cern.ch/record/2223839>.
- 3471 [110] HILL A, VAN DER BIJ J J. Strongly interacting singlet-doublet higgs model[J/OL]. Phys. Rev.
3472 D, 1987, 36:3463-3473. DOI: [10.1103/PhysRevD.36.3463](#).
- 3473 [111] DAVOUDIASL H, HEWETT J, RIZZO T. Bulk gauge fields in the randall–sundrum
3474 model1work supported by the department of energy, contract de-ac03-76sf00515.1[J/OL].
3475 Physics Letters B, 2000, 473(1):43 - 49. DOI: [10.1016/S0370-2693\(99\)01430-6](#).

Bibliography

- 3476 [112] AABOUD M, et al. Search for heavy ZZ resonances in the $\ell^+\ell^-\ell^+\ell^-$ and $\ell^+\ell^-\nu\bar{\nu}$ final states
3477 using proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector[J/OL]. Eur. Phys. J.
3478 C, 2018, 78(4):293. DOI: [10.1140/epjc/s10052-018-5686-3](https://doi.org/10.1140/epjc/s10052-018-5686-3).
- 3479 [113] AAD G, et al. Search for heavy resonances decaying into a pair of Z bosons in the $\ell^+\ell^-\ell^+\ell^-$
3480 and $\ell^+\ell^-\nu\bar{\nu}$ final states using 139 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV with the
3481 ATLAS detector[J]. 2020.
- 3482 [114] FREDERIX R, FRIXIONE S. Merging meets matching in MC@NLO[J/OL]. JHEP, 2012, 12:
3483 061. DOI: [10.1007/JHEP12\(2012\)061](https://doi.org/10.1007/JHEP12(2012)061).
- 3484 [115] AGASHE K, DAVOUDIASL H, PEREZ G, et al. Warped Gravitons at the LHC and Beyond
3485 [J/OL]. Phys.Rev., 2007, D76:036006. DOI: [10.1103/PhysRevD.76.036006](https://doi.org/10.1103/PhysRevD.76.036006).
- 3486 [116] ATLAS Collaboration. Jet reconstruction and performance using particle flow with the ATLAS
3487 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).
- 3488 [117] ELAGIN A, MURAT P, PRANKO A, et al. A new mass reconstruction technique for reso-
3489 nances decaying to $\tau\tau$ [J/OL]. Nuclear Instruments and Methods in Physics Research Section A:
3490 Accelerators, Spectrometers, Detectors and Associated Equipment, 2011, 654:481-489. DOI:
3491 [10.1016/j.nima.2011.07.009](https://doi.org/10.1016/j.nima.2011.07.009).
- 3492 [118] BRADLEY A P. The use of the area under the roc curve in the evaluation of machine learning
3493 algorithms[J/OL]. Pattern Recognition, 1997, 30(7):1145 - 1159. DOI: [10.1016/S0031-3203\(96\)00142-2](https://doi.org/10.1016/S0031-3203(96)00142-2).
- 3495 [119] BALDI P, CRANMER K, FAUCETT T, et al. Parameterized neural networks for high-energy
3496 physics[J/OL]. Eur. Phys. J. C, 2016, 76(5):235. DOI: [10.1140/epjc/s10052-016-4099-4](https://doi.org/10.1140/epjc/s10052-016-4099-4).
- 3497 [120] Cousins R D, Linnemann J T, Tucker J. Evaluation of three methods for calculating statistical
3498 significance when incorporating a systematic uncertainty into a test of the background-only
3499 hypothesis for a Poisson process[J/OL]. Nuclear Instruments and Methods in Physics Research
3500 A, 2008, 595(2):480-501. DOI: [10.1016/j.nima.2008.07.086](https://doi.org/10.1016/j.nima.2008.07.086).
- 3501 [121] Measurements of higgs boson production and couplings in the four-lepton channel in pp col-
3502 lisions at center-of-mass energies of 7 and 8 tev with the atlas detector[J/OL]. Phys. Rev. D,
3503 2015, 91:012006. DOI: [10.1103/PhysRevD.91.012006](https://doi.org/10.1103/PhysRevD.91.012006).
- 3504 [122] CRANMER K S. Kernel estimation in high-energy physics[J/OL]. Comput. Phys. Commun.,
3505 2001, 136:198-207. DOI: [10.1016/S0010-4655\(00\)00243-5](https://doi.org/10.1016/S0010-4655(00)00243-5).
- 3506 [123] OREGLIA M. A Study of the Reactions $\psi' \rightarrow \gamma\gamma\psi$ [M]. SLAC, 1980.
- 3507 [124] GAISER J. Charmonium Spectroscopy From Radiative Decays of the J/ψ and ψ' [M]. SLAC,
3508 1982.
- 3509 [125] GORIA S, PASSARINO G, ROSCO D. The Higgs Boson Lineshape[J/OL]. Nucl.Phys., 2012,
3510 B864:530-579. DOI: [10.1016/j.nuclphysb.2012.07.006](https://doi.org/10.1016/j.nuclphysb.2012.07.006).

Bibliography

- 3511 [126] SPIRA M, DJOUADI A, GRAUDENZ D, et al. Higgs boson production at the LHC[J]. Nucl.
3512 Phys., 1995, B 453:17-82.
- 3513 [127] BALL R D, et al. Parton Distribution Benchmarking with LHC Data[J/OL]. JHEP, 2013, 04:
3514 125. DOI: [10.1007/JHEP04\(2013\)125](https://doi.org/10.1007/JHEP04(2013)125).
- 3515 [128] BIJNENS J, EEROLA P, MAUL M, et al. Qcd signatures of narrow graviton resonances in
3516 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).
- 3518 [129] ATLAS Collaboration. PMG Systematic Uncertainty Recipes[Z].
- 3519 [130] GROSS E, VITELLS O. Trial factors for the look elsewhere effect in high energy physics
3520 [J/OL]. Eur. Phys. J. C, 2010, 70:525-530. DOI: [10.1140/epjc/s10052-010-1470-8](https://doi.org/10.1140/epjc/s10052-010-1470-8).

3521

Publications

3522 已发表论文

- 3523 1. ATLAS Collaboration. Search for electroweak diboson production in association
3524 with a high-mass dijet system in semileptonic final states in pp collisions at 13 TeV
3525 with the ATLAS detector[J]. **Phys. Rev. D** **100**, 032007.
- 3526 2. ATLAS Collaboration. Constraints on off-shell Higgs boson production and the
3527 Higgs boson total width in $ZZ \rightarrow \ell\ell\ell'\ell'$ and $ZZ \rightarrow \ell\ell\nu\nu$ final states with the
3528 ATLAS detector[J]. **Physics Letters B**, **2018**, 786.
- 3529 3. ATLAS Collaboration. Measurement of the four-lepton invariant mass spectrum
3530 in 13 TeV proton-proton collisions with the ATLAS detector[J]. **Journal of High**
3531 **Energy Physics**, **2019**, 2019(4): 48.
- 3532 4. Chen K, et al. Design and Evaluation of the LAr Trigger Digitizer Board in the
3533 ATLAS Phase-I Upgrade[J]. **IEEE Transactions on Nuclear Science**, **2019**, 66(8):
3534 2011-2016.

3535 待发表论文

- 3536 1. Observation of electroweak production of two jets and a Z-boson pair with
3537 the ATLAS detector at the LHC. **Submitted to Nature Physics**, e-print:
3538 [arXiv:2004.10612](https://arxiv.org/abs/2004.10612).
- 3539 2. Search for heavy resonances decaying into a pair of Z bosons in $ZZ \rightarrow \ell\ell\ell'\ell'$
3540 and $ZZ \rightarrow \ell\ell\nu\nu$ final states using 139 fb^{-1} of pp collisions at 13 TeV with the
3541 ATLAS detector. **Submitted to Eur. Phys. J. C**, e-print: [arXiv:2009.14791](https://arxiv.org/abs/2009.14791).

3542 会议报告

- 3543 1. Search for heavy resonances decaying into a pair of Z bosons with the ATLAS
3544 Detector. **China LHC Physics Workshop, online (Beijing, China), 2020.11**
- 3545 2. Measurement of Inclusive $\ell\ell\ell'\ell'$ and $\ell\ell\nu\nu + 2\text{-jet}$ Cross Section and Observation
3546 of EW Component with the ATLAS detector. **Workshop on Connecting Insights**
3547 **in Fundamental Physics: Standard Model and Beyond, Corfu, Greece, 2019.9**
- 3548 3. Off-shell Higgs signal strength measurement in the high-mass $ZZ \rightarrow \ell\ell\ell'\ell'$ and
3549 $ZZ \rightarrow \ell\ell\nu\nu$ final states with the ATLAS detector. **American Physics Society**
3550 **April Meeting, Columbus, USA, 2018.4**