

# 中国科学技术大学

# 博士学位论文



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

作者姓名： 祝鹤龄

学科专业： 粒子与原子核物理

导师姓名： 赵政国 马宏

完成时间： 二〇二〇年九月二十四日

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



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

Author: Heling Zhu

Speciality: Particle and Nuclear Physics

Supervisors: Zhengguo Zhao, Hong Ma

Finished time: September 24, 2020

## 摘要

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

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

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

28

## ABSTRACT

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

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

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

---

Abstract

---

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

63

64

## Acknowledgments

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

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

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

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

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

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

---

## Acknowledgments

---

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

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

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

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

## Contents

## **List of Figures**

## **List of Tables**

*Dedicated to my mother!*

献给我的母亲！

139

## Chapter 1 Introduction

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

144     **Elementary particles**

145     From around the 6th century BC, ancient Greek philosophers Leucippus, Democritus,  
146     and Epicurus brought up a philosophical idea that everything is composed of “uncuttable”  
147     elementary particles. In the 19th century, John Dalton, through his work on stoichiometry,  
148     concluded that each element of nature was composed of a single, unique type of particle.  
149     The particle was named as “atom” after the Greek word atomos, with the meaning of “in-  
150     divisible”. However this Dalton’s atom theory was strongly challenged later. Near the  
151     end of 19th century, physicists discovered that Dalton’s atoms are not, in fact, the funda-  
152     mental particles of nature, but conglomerates of even smaller particles. Electron was dis-  
153     covered by J. J. Thomson in 1897, and then its charge was carefully measured by Robert  
154     Andrews Millikan and Harvey Fletcher in their “oil drop experiment” of 1909. In early  
155     20th-century, Rutherford’s “gold foil experiment” showed that the atom is mainly empty  
156     space, with almost all its mass concentrated in a tiny positively charge atomic nucleus.  
157     Then the discoveries of anti-particles (the positron in 1932) and other particles (e.g. the  
158     muon in 1936) show that more discoveries could be expected in future experiments.

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

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

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

174 **Higgs mechanics and electroweak symmetry breaking**

175 In 1961, Sheldon Glashow, Steven Weinberg and Abdus Salam together brought for-  
176 ward a unified electroweak theory to combine the electromagnetic and weak interactions.  
177 In the standard model, if under the condition that the energy is high enough but elec-  
178 troweak symmetry is unbroken, all elementary particles are massless. But measurements  
179 show the fact that the W and Z bosons actually have masses. Later on, the Higgs mechan-  
180 ics resolves this conundrum. The simplest description of the mechanism adds a Higgs  
181 field that permeates all space to the Standard Model. Below some extremely high energy,  
182 the field causes spontaneous symmetry breaking during interactions. All massive parti-  
183 cles in the Standard Model, including the W and Z bosons, interact with Higgs boson to  
184 acquire their mass.

185 Over the past few decades, with the combination of electroweak theory, Higgs me-  
186 chanics and strong interactions has been widely accepted. But the Higgs boson, which is  
187 essential to explain the mechanics of the property “mass” for gauge bosons and fermions,  
188 had been the final missing piece in the Standard Model of particle physics at that time.  
189 The mass of Higgs boson was not specifically predicted, and it has been searched in sev-  
190 eral large experiments (eg. LEP at CERN, Tevatron at Fermilab, and LHC at CERN) with  
191 different energy. In 2012, the discovery of Higgs boson was finally announced by the  
192 ATLAS and CMS collaborations at the Large Hadron Collider (LHC) with its mass round  
193 125 GeV. Peter Higgs and Francois Englert were award the 2013’s Nobel Prize in Physics  
194 for their theoretical discovery of a mechanism that contributes to our understanding of the  
195 origin of mass of subatomic particles.

196 **Contents of this thesis**

197 This dissertation is organized as follows. Section 2 briefly introduces the Standard  
198 Model of particle physics, the Higgs mechanism related to the dissertation and the LHC  
199 phenomenology. Section 3 gives an overview of the LHC and the ATLAS detector. The  
200 detector simulation and the reconstruction of physics objects are described in section 4.  
201 And then section 5 focuses on the Standard model ZZ production cross section measure-  
202 ment in  $ZZ \rightarrow \ell\ell\ell'\ell'$  channel, where  $\ell$  stands for electron or muon, and the observa-  
203 tion of its electroweak component as well as its further prospects in High luminosity LHC  
204 (HL-LHC). Section 6 present the search of possible heavy resonances in  $ZZ \rightarrow \ell\ell\ell'\ell'$   
205 channel. In the end, section 7 gives the summary and outlook for future physics in LHC.

206

# Chapter 2 Theory

## 207 2.1 The Standard Model of Particle Physics

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

### 214 2.1.1 Elementary particles in the Standard Model

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

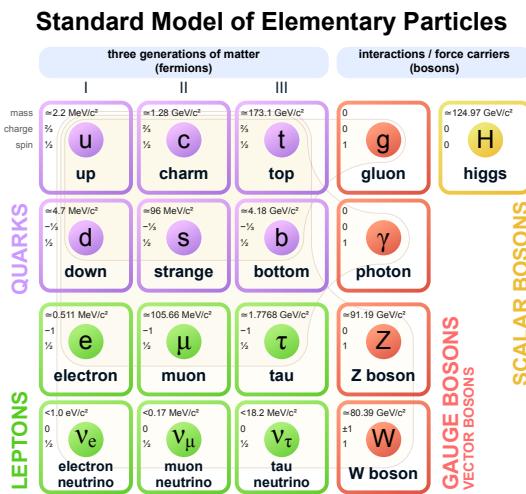


Fig. 2.1 The elementary particles of the Standard Model.

216

217 **Fermions** The Standard Model includes 12 elementary particles of spin- $\frac{1}{2}$  obeying the  
 218 Fermi-Dirac statistics, known as fermions. They are classified into two types: *leptons* and  
 219 *quarks* according to their interactions. The *leptons* include three generations: electron  
 220 ( $e$ ) and electron neutrino ( $\nu_e$ ); muon ( $\mu$ ) and muon neutrino ( $\nu_\mu$ ); tau ( $\tau$ ) and tau neutrino  
 221 ( $\nu_\tau$ ). The  $e$ ,  $\mu$  and  $\tau$  carry electric charge of -1 and three neutrinos are electrically neutral.  
 222 All the leptons can participate in electroweak interactions. Also there are three generations  
 223 of *quarks*: up ( $u$ ) and down ( $d$ ); charm ( $c$ ) and strange ( $s$ ); top ( $t$ ) and bottom ( $b$ ). The  
 224 defining property of the quarks is that they carry color charge (while leptons don't), and  
 225 hence interact via the strong interaction, letting them to be strongly bound from one to

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

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

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

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

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

### 2.1.2 Electroweak theory

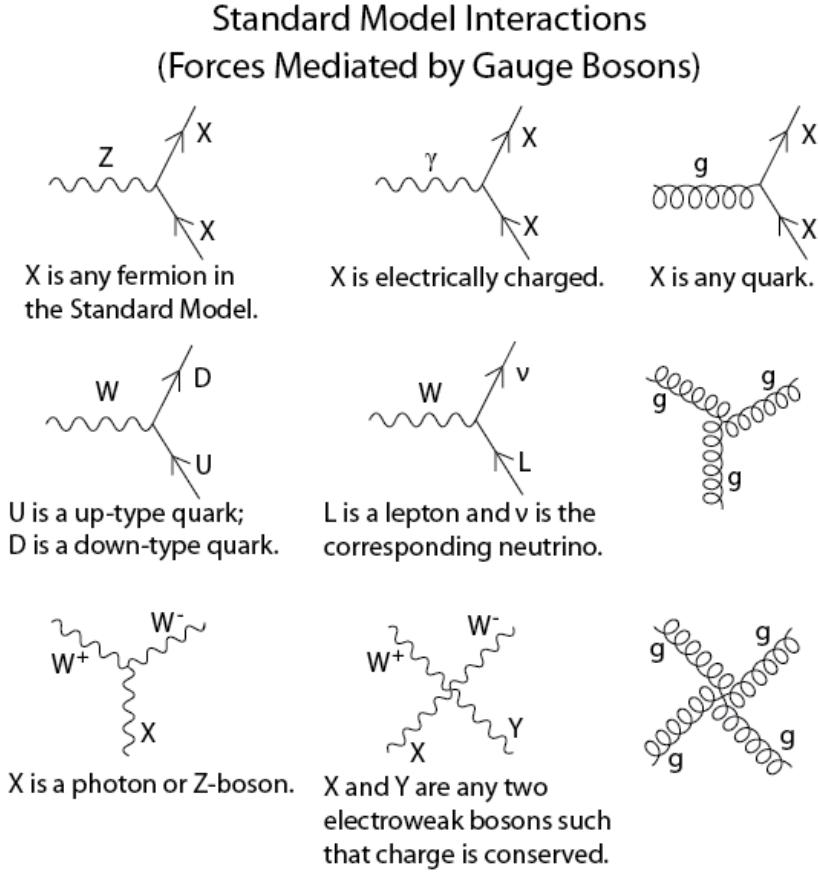
The electroweak interaction is the unified description of two of the four known fundamental interactions of nature: electromagnetism and the weak interaction. It is based on the gauge group of  $SU(2)_L \times SU(1)_Y$ , in which  $L$  is the left-handed fields and  $Y$  is the weak hypercharge<sup>[1]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

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

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

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

where  $W_\mu^i$  and  $B_\mu$  present the  $SU(2)_L$  and  $SU(1)_Y$  gauge fields respectively, with the



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

253 corresponding field strength tensors of

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

254 In the equations above,  $g$  is the  $SU(2)_L$  gauge coupling and  $\epsilon_{ijk}$  is the totally antisymmetric tensor. The gauge Lagrangian has three and four-point self interactions of  $W^i$ , which 255 result in triple and quartic gauge boson couplings.  
256

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

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

258 where  $\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$  is a complex Higgs scalar, and  $V(\phi)$  is the Higgs potential which is 259 restricted into the form of

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

260 due to the combination of  $SU(2)_L \times SU(1)_Y$  invariance and renormalizability. In Eq. 2.5,

261  $\mu$  is a mass-dependent parameter and  $\lambda$  is the quartic Higgs scalar coupling, which rep-  
 262 resents a quartic self-interaction between the scalar fields. When  $\mu^2 < 0$ , there will be  
 263 spontaneous symmetry breaking (more details in section 2.1.3). To maintain vacuum sta-  
 264 bility,  $\lambda > 0$  is required. And in Eq. 2.4, the gauge covariant derivative is defined as

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

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

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

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

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

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

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

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

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

274 the matrices  $\Gamma_{mn}$  refer to the Yukawa couplings between single Higgs doublet ( $\phi$ ) and the  
 275 various flavors of quarks (m) and leptons (n).

### 2.1.3 Higgs mechanism and Electroweak symmetry breaking

277 As shown in previous subsection, the Lagrangian  $L_{gauge}$  does not involve any mass  
 278 term due to the requirement of gauge invariance. So all the W and B bosons should be  
 279 massless. But experimental observations show that the gauge bosons are massive. There-  
 280 fore, the gauge invariance must be broken spontaneously. The Higgs field is introduced  
 281 to break the  $SU(2)_L \times U(1)_Y$  symmetry and gauge bosons and fermions can interact with

282 Higgs filed to acquire their masses. And this specific process is named *Higgs mechanism*  
 283 in SM.

284 The Higgs field  $\phi$  is a doublet and can be written in a Hermitian basis as

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

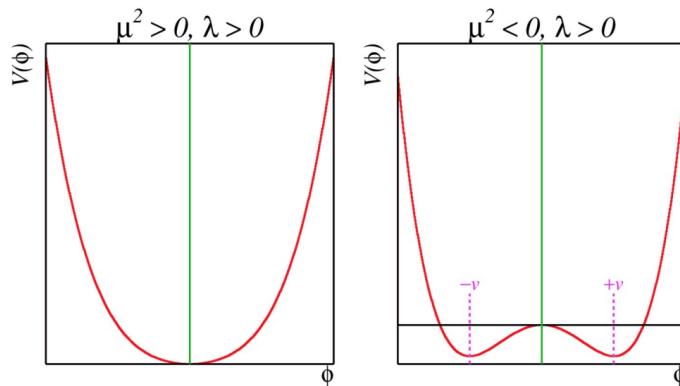
285 where  $\phi_i = \phi_i^+$  stand for four Hermitian field. In this new basis, the Higgs potential in  
 286 Eq. 2.5 can be expressed as:

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

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

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

289 The minimization of this potential depends on the sign of  $\mu^2$  as shown in figure 2.3. When  
 290  $\mu^2 > 0$  the minimum occurs at  $v = 0$ , namely the vacuum is empty space and  $SU(2)_L \times$   
 291  $U(1)_Y$  symmetry is unbroken. In the case of  $\mu^2 < 0$ , the  $v = 0$  symmetric point is no  
 292 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  $\phi_0$  of Higgs doublet can be



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

293

294 expressed by

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

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

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

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

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

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

298 where the  $W$  and  $Z$  fields are

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

299 Therefore, in Eq. 2.15 spontaneous symmetry breaking brings out masses for the  $W$  and  
300  $Z$  gauge bosons

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

301 where  $\theta_W$  is the weak angle defined as

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

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

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

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

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

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

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

306 Due to the unknown of quartic Higgs coupling  $\lambda$ , the Higgs mass is not predicted. The

307 third and fourth terms in the Higgs potential  $V$  denote the induced cubic and quartic in-  
 308 teractions of the Higgs scalar.

309 Through the Higgs mechanism, fermions can also acquire their masses. In the unitary  
 310 gauge, Yukawa Lagrangian ( $L_{Yukawa}$ ) can be written as a simple form of<sup>[2]</sup>

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

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

### 312 2.1.4 Beyond the SM Higgs sector

313 After the discovery of the Higgs boson by the ATLAS and CMS Collaborations at the  
 314 LHC<sup>[3-4]</sup> in 2012, one question comes out: if this Higgs boson at around 125 GeV is fully  
 315 responsible for the unitarization of the scattering amplitudes? The possibility that this  
 316 discovered particle is just a part of the extended Higgs sector by various extensions cannot  
 317 be ruled out. Many models, motivated by hierarchy and naturalness arguments, predicted  
 318 the extended Higgs sector, such as the electroweak-singlet model<sup>[5]</sup> and the two-Higgs-  
 319 doublet models (2HDM)<sup>[6]</sup>.

#### 320 Singlet scalar extension of the SM

321 The electroweak singlet model can be considered as the minimal extension of the SM  
 322 Higgs sector, encompassing a single gauge singlet real scalar field  $S$ . In this model, a  
 323 heavy, real singlet is introduced in addition to the SM one. The associated zero tempera-  
 324 ture, tree-level scalar potential can be written as:

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

325 where

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

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

332    **Two Higgs Doublet Model**

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

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

- 337    • The Higgs boson masses:  $m_h$ ,  $m_H$ ,  $m_A$  and  $m_{H^{\pm}}$ .  
338    •  $\tan\beta$ : the ratio of the vacuum expectation values of the two doublets.  
339    •  $\alpha$ : the mixing angle between the CP-even Higgs bosons.  
340    •  $m_{12}^2$ : the potential parameter that mixes the two Higgs doublets.

341    where the  $m_h$  can be identified as the mass of observed Higgs boson at around 125 GeV,  
342    and  $m_H$  is another heavy scalar with similar properties as  $h$  boson. The coupling of the  
343    neutral Higgs bosons to the W and Z are the same:

- 344    1. The coupling of the light Higgs,  $h$ , to either WW or ZZ is the same as the Standard  
345    Model coupling times  $\sin(\beta - \alpha)$
- 346    2. The coupling of the heavier Higgs,  $H$ , is the same as the Standard Model coupling  
347    times  $\cos(\alpha - \beta)$ .
- 348    3. The coupling of the pseudoscalar,  $A$ , to vector bosons vanishes.

349    The two Higgs doublets,  $\Phi_1$  and  $\Phi_2$ , can couple to fermions (leptons and up- and down-  
350    type quarks) in several ways, which leads to several types of 2HDM models:

- 351    • Type-I model: all quarks and leptons couple only to  $\Phi_2$ .
- 352    • Type-II model: down-type quarks and leptons couple to  $\Phi_1$ , and up-type quarks  
353    couple to  $\Phi_2$ .
- 354    • The “lepton-specific” model: leptons couple to  $\Phi_1$ , while all quarks couple to  $\Phi_2$ .
- 355    • The “flipped” model: down-type quarks couple to  $\Phi_1$ , while up-type quarks and  
356    leptons couple to  $\Phi_2$ .

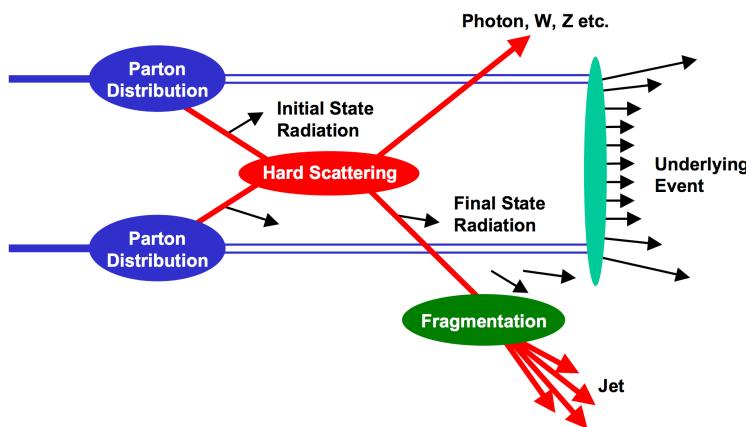
357    **2.2 Phenomenology of Large Hadron Collider**

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

365 closely to this dissertation.

366 **2.2.1 Physics at hadronic collision**

367 Protons are not the elementary particle, which actually are composed of quarks and  
 368 gluons. So in proton-proton (pp) collision at the LHC, it is not protons themselves interact  
 369 but quarks and gluons. Scattering processes can then be further classified into either *hard*  
 370 or *soft* processes according to the momentum transfer during the interaction<sup>[7]</sup>. QCD, as  
 371 an underlying theory for both processes, its approach and level of understandings in two  
 372 cases are quite different. For hard process, eg. Higgs, vector bosons and jets production,  
 373 the rates and event properties can be precisely predicted based on perturbation theory.  
 374 However, for soft processes like total cross-section, the underlying events, the rates and  
 375 properties are dominated by non-perturbative QCD effects that are less understood. For  
 376 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



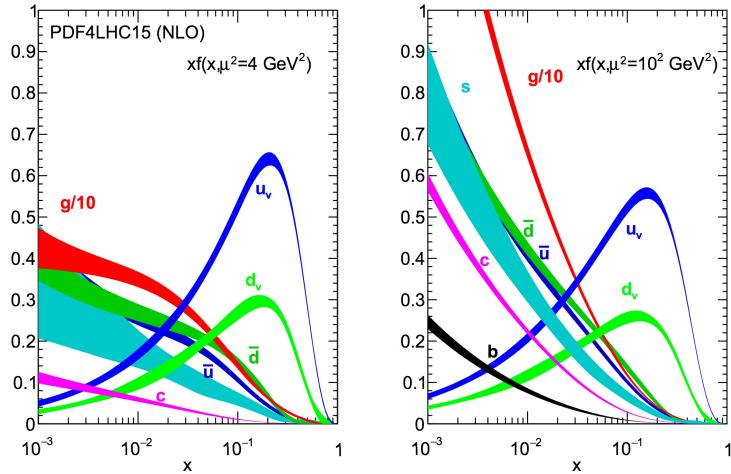
377 **Fig. 2.4 Schematic view of a hadron-hadron collision<sup>[8]</sup>.**

378 as below:

- 379 • **Parton Distribution Function (PDF):**  $f_i(x, Q^2)$  gives the probability of finding a  
 380 parton with flavor  $i$  (quark or gluon), carrying a momentum fraction of  $x$  and with  
 381 the scale of momentum transfer  $Q$  in a proton. Parton distribution function cannot  
 382 be fully calculated by perturbative QCD because of the inherent non-perturbative  
 383 nature of partons. There are many different sets of PDFs that are determined by  
 384 fits to data from experimental observables in various processes. As an example,  
 385 figure 2.5 shows the *PDF4LHC15 NLO PDFs*, which are based on the combination  
 386 of the *CT14*, *MMHT14* and *NNPDF3.1 NNLO* PDF sets<sup>[9]</sup>.
- 387 • **Fragmentation and hadronization:** The processes to produce final state particles

388 (or jets) from the partons produced in hard scattering.

- 389 • **Initial/Final state radiation:** The incoming and outgoing partons that carry color  
 390 charge can emit QCD radiation, which gives rise to additional jets. Also the charged  
 391 incoming and outgoing particles can emit Quantum Electrodynamics (QED) radia-  
 392 tions with photons.
- 393 • **Underlying events:** Products from soft processes (not come from the primary hard  
 394 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  $x$ .

### 395 Cross section of hard scattering

396 According to *QCD factorization theorems*<sup>[10]</sup>, the perturbative calculations can be ap-  
 397 plied to many important hard processes involving hadrons. The basic problem addressed  
 398 by factorization theorems is how to calculate high energy cross sections. Consider the  
 399 process of scattering between two hardons A and B to produce a final state X, the cross  
 400 section  $\sigma$  can be obtained by summing over all the subprocess cross section  $\hat{\sigma}$ <sup>[11]</sup>

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

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

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

403 where  $Q^2$  represents large momentum scale that characterizes the hard scattering. Later  
 404 on, since the finite corrections were not universal and had to be calculated separately for  
 405 each process, the perturbative  $O(\alpha_S^n)$  corrections to the leading logarithm cross section

406 in Eq. 2.26 need to be applied, one can get:

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

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

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

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

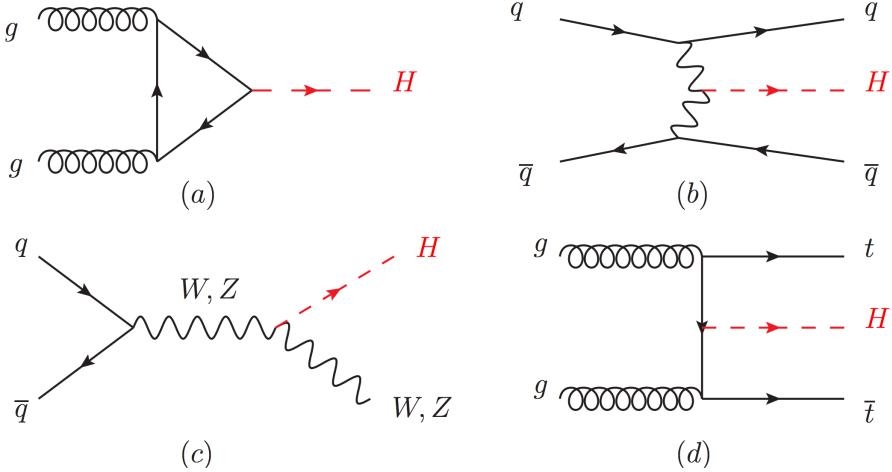
413  $\mu_R$  and  $\mu_F$  depend on the order of truncation in Eq. 2.28. In principle, if cross sec-  
 414 tion is calculated to all orders, it is invariant under changes in these parameters. The  
 415 choices of  $\mu_R$  and  $\mu_F$  are arbitrary. To avoid unnaturally large logarithms reappearing in  
 416 the perturbation series, it is sensible to choose  $\mu_R$  and  $\mu_F$  values of the order of the typical  
 417 momentum scales of the hard scattering process and  $\mu_R = \mu_F$  is also often assumed. Take  
 418 Drell–Yan process as an example, the standard choice is  $\mu_R = \mu_F = m_{ll}$ , where  $m_{ll}$  is the  
 419 invariant mass of di-lepton pair.

## 420 2.2.2 Higgs physics at the LHC

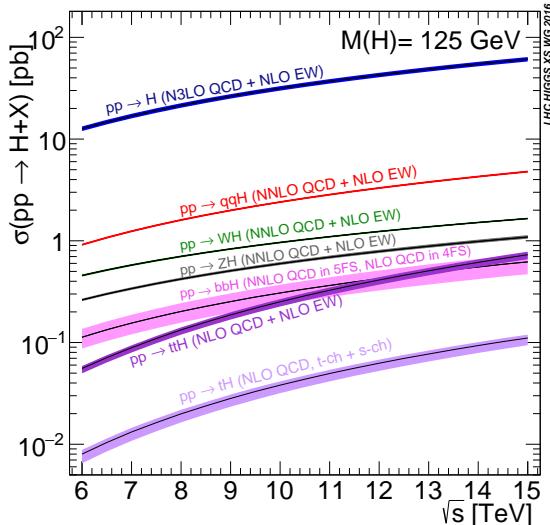
421 One important physics purpose of the LHC is searching for the Higgs boson, which  
 422 was the last missing part in the SM. This section will discuss both the production and  
 423 decay modes of the SM Higgs boson in proton-proton collision.

### 424 Higgs productions

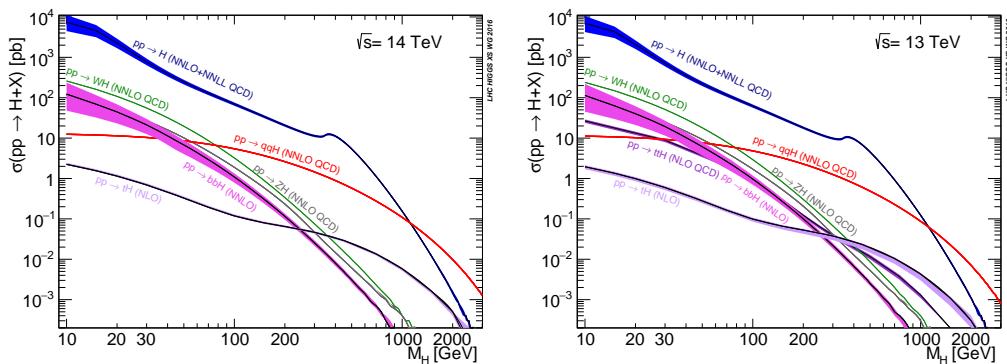
425 The Higgs boson can be produced through several processes. There are 4 main produc-  
 426 tion modes at the LHC: gluon-gluon Fusion ( $ggF$ ), vector boson Fusion ( $VBF$ ), associated  
 427 production with vector-bosons ( $VH$ ) (also called the Higgs Strahlung) and associated pro-  
 428 duction with a pair of top/anti-top quarks ( $t\bar{t}H$ )<sup>[12]</sup>. Figure 2.6 shows the corresponding  
 429 Feynman diagrams of each process (at LO). For pp collisions, the cross section of pro-  
 430 duction of Higgs boson is a function of centre-of-mass energy  $\sqrt{s}$ . Figure 2.7 depicts  
 431 the cross section of SM Higgs, whose mass is 125 GeV, for several different produc-  
 432 tion modes when centre-of-mass energy varying from 6 to 15 TeV. Figure 2.8 shows the  
 433 prospect of production cross section as a function of Higgs mass from 10 to 2000 GeV for  
 434 pp collision at the centre-of-mass energy of 13 TeV and 14 TeV<sup>[13]</sup>.



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



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

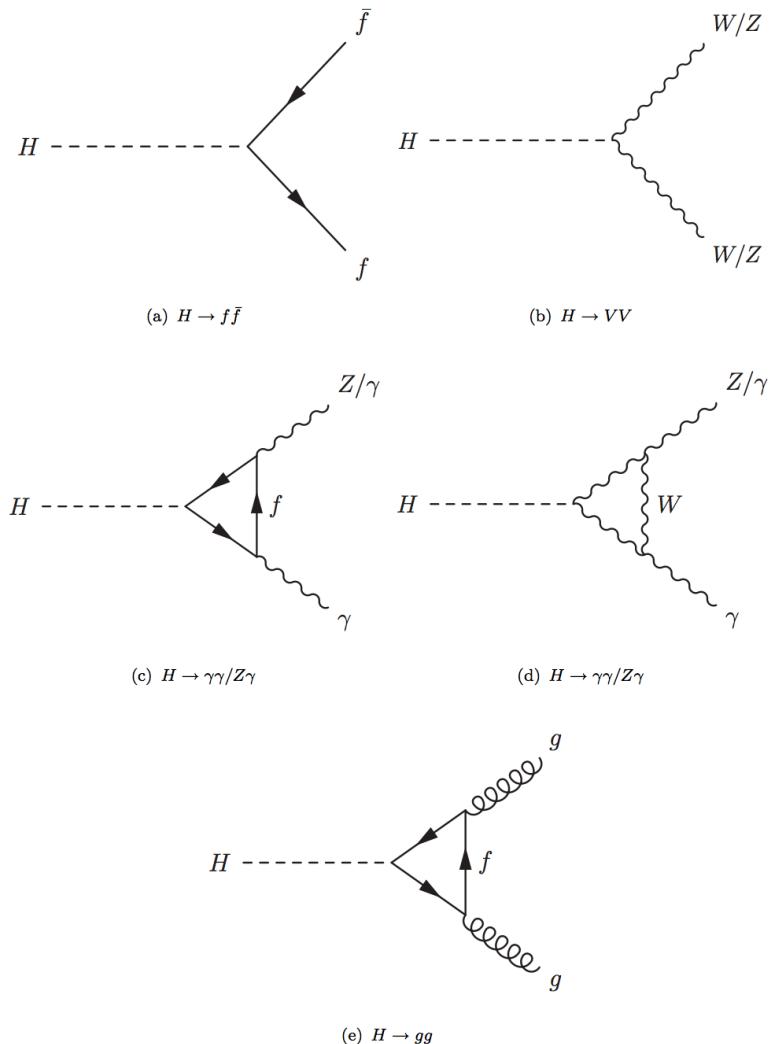


**Fig. 2.8** Higgs boson production cross section for various production modes as a function of the Higgs mass for  $\sqrt{s} = 13$  TeV (left) and 14 TeV (right) for pp collision.

### Higgs decays

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

<sup>437</sup> 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.

438

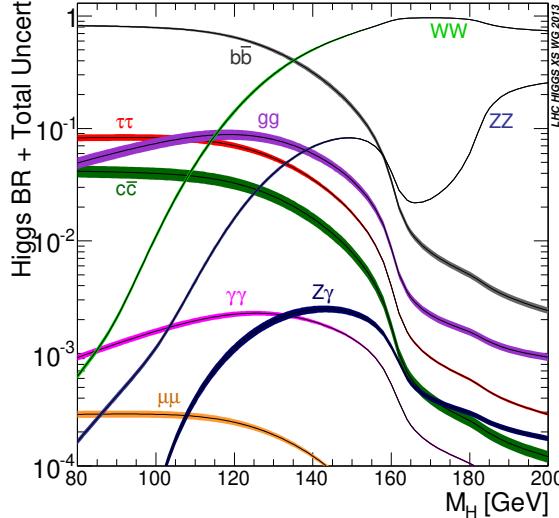
439 decaying into different final states as a function of Higgs mass is shown in figure 2.10.

### 2.2.3 Diboson physics at the LHC

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

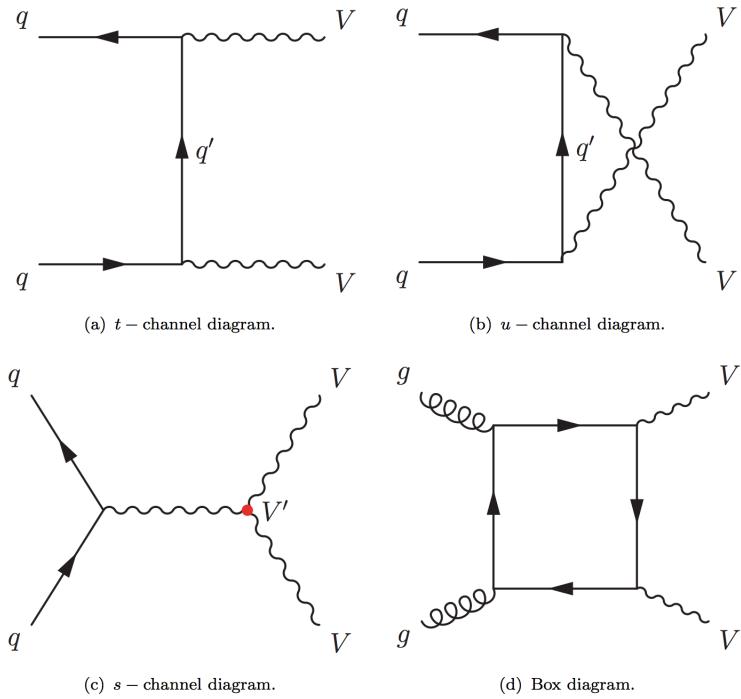
446 Diboson productions

About 90% of diboson productions at hadron collider is from quark-antiquark annihilation, while others are from gluon initiated process. Figure 2.11 shows the tree-level



**Fig. 2.10 Branching ratio of Higgs decays in various channels as a function of Higgs mass<sup>[14]</sup>.**

Feynman diagrams of diboson production. Then figure 2.12 illuminates the total produc-



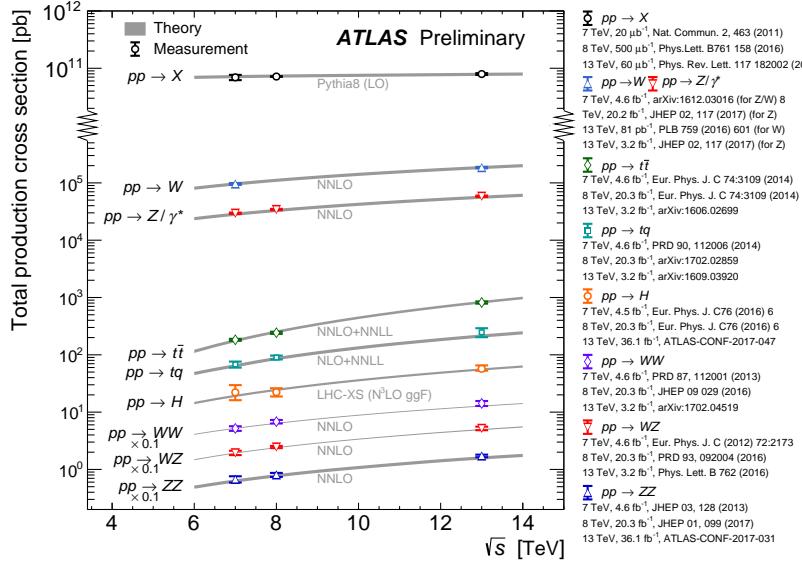
**Fig. 2.11 The tree-level Feynman diagrams of diboson production at the LHC.**

449

450   tion cross-section presented by ATLAS as a function of centre-of-mass energy  $\sqrt{s}$  from  
 451   7 to 13 TeV for several diboson processes compared to some other major processes in  
 452   hadron collision. The cross section for diboson processes are calculated at next-to-next  
 453   leading order (NNLO).

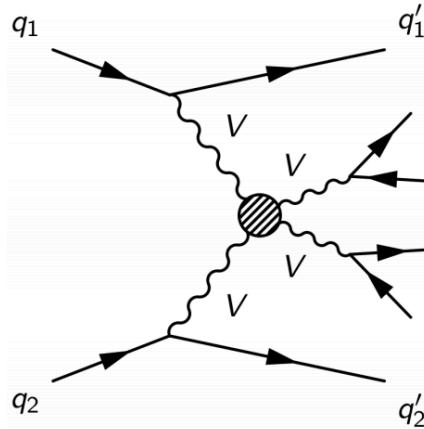
#### 454   **Vector boson scattering**

455   The  $SU(2)_L \times U(1)_Y$  structure in SM predicts self-interactions between electroweak  
 456   gauge bosons. Those self-couplings can involve either three or four gauge bosons at a



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

single vertex, known as triple gauge coupling (*TGC*) or quartic gauge couplings (*QGC*), respectively. Vector boson scattering (*VBS*) is carried out by four electroweak vector bosons, namely  $Z$ ,  $W^\pm$  and photon ( $\gamma$ ) as the Feynman diagrams shown in figure 2.13. And the vertexes include either those self-interactions or the interactions with the Higgs boson are described in figure 2.14.

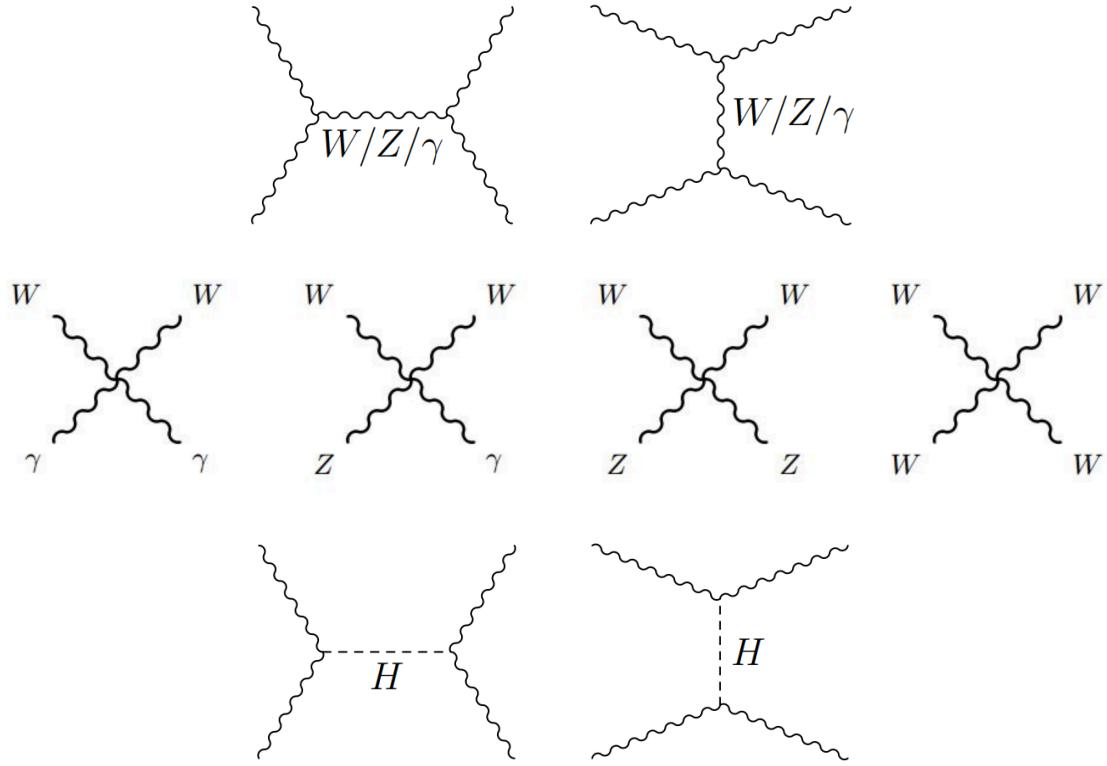


**Fig. 2.13** Feynman diagrams of the vector boson scattering.

461

462 The amplitudes of leading-order (LO) VBS can be expressed as<sup>[15]</sup>:

$$\begin{aligned}
 iM_{TGC}^{s-channel} &= -i \frac{g_1^2}{4m_W^4} [s(t-u) - 3m_W^2(t-u)] \\
 iM_{TGC}^{t-channel} &= -i \frac{g_1^2}{4m_W^4} \left[ (s-u)t - 3m_W^2(s-u) + \frac{8m_W^2}{s} u^2 \right]
 \end{aligned} \tag{2.29}$$



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

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

463

$$\begin{aligned} iM_{Higgs} &= -i \frac{C_v g_1^2}{4m_W^2} \left[ \frac{(s-2m_W^2)^2}{s-m_H^2} + \frac{(t-2m_W^2)^2}{t-m_H^2} \right] \\ &\simeq -i \frac{C_v g_1^2}{4m_W^2} (s+t) \end{aligned} \quad (2.31)$$

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

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

465 In Eq. 2.32, the amplitude grows as a function of centre-of-mass energy ( $\sqrt{s}$ ), which  
 466 violates the unitarity in the TeV region. Considering the Higgs term in Eq. 2.31 can  
 467 perfectly cancel out this growing, and the remaining term  $O((s/m_W^2)^0)$  only depends on  
 468 the total amplitude in SM.

469 In conclusion, the Higgs boson acts as "moderator" to unitarize high-energy longi-  
 470 tudinal vector boson scattering as introducing the Higgs restores the unitarity of total  
 471 amplitude in high energy region.

472    **Chapter 3 The Large Hadron Collider and the ATLAS**  
473    **Detector**

474    **3.1 The Large Hadron Collider**

475    Located near the French-Swiss border at the European Organization for Nuclear Re-  
476    search (CERN), the Large Hadron Collider (LHC) is the world’s largest and most pow-  
477    erful particle collider. It’s the proton-proton collider with the centre-of-mass energy up  
478    to 14 TeV. The beams inside the LHC are made to collide at four locations around its  
479    27-kilometer accelerator ring, corresponding to the positions of four particle detectors -  
480    ATLAS, CMS, ALICE and LHCb. With its unprecedented energy, the LHC is designed  
481    to observe physics that involve highly massive particles, which have never been observed  
482    in previous accelerators with lower energies.

483    **3.1.1 Operation history and machine layout**

484    **Operation history**

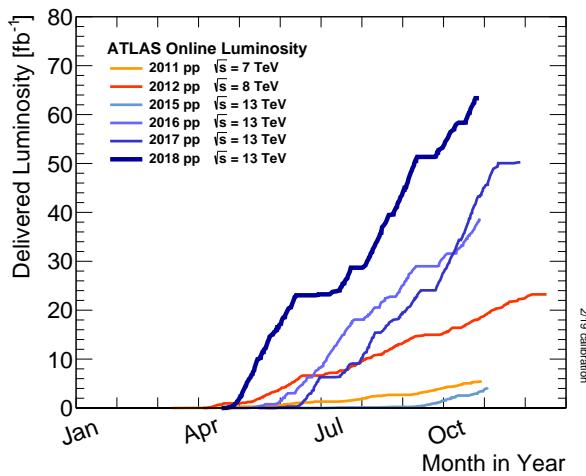
485    The LHC<sup>[16-19]</sup> is a two-ring-superconducting-hadron accelerator and collider lies in  
486    a tunnel 27 kilometres in circumference and as deep as 175 metres underground. It’s  
487    designed to provide proton-proton (pp) collisions at the centre-of-mass energy ( $\sqrt{s}$ ) up  
488    to 14 TeV with a unprecedented luminosity of  $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ . In the meantime, it can also  
489    collide heavy (Pb) ions with an energy of 2.8 TeV per nucleon and a peak luminosity of  
490     $10^{27} \text{ cm}^{-2} \text{ s}^{-1}$ . Table 3.1 shows the main design parameters of the LHC for proton-proton  
491    collisions.

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

499    The second operation period with higher centre-of-mass energy at 13 TeV started from  
500    2015 called “run-2”. And it continued to the end of 2018 with total integrated luminosity  
501    reaching about  $147 \text{ fb}^{-1}$  for ATLAS experiment. Figure 3.1 shows the cumulative lumi-  
502    nosity as a function of time in month delivered to ATLAS experiment during stable beams  
503    in years from 2011 to 2018.

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

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

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

### 504 Machine layout

505 The layout of CERN accelerator complex is shown in figure 3.2. The protons are  
 506 accelerated by a series of machines before being injected into the main ring. At begin-  
 507 ning, the 50 MeV protons are produced in the linear particle accelerator LINAC2, and  
 508 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 accelerated to 450 GeV in Super Proton Synchrotron (SPS). At the end, they are injected into the main ring, and can reach a maximum energy of 7 TeV.

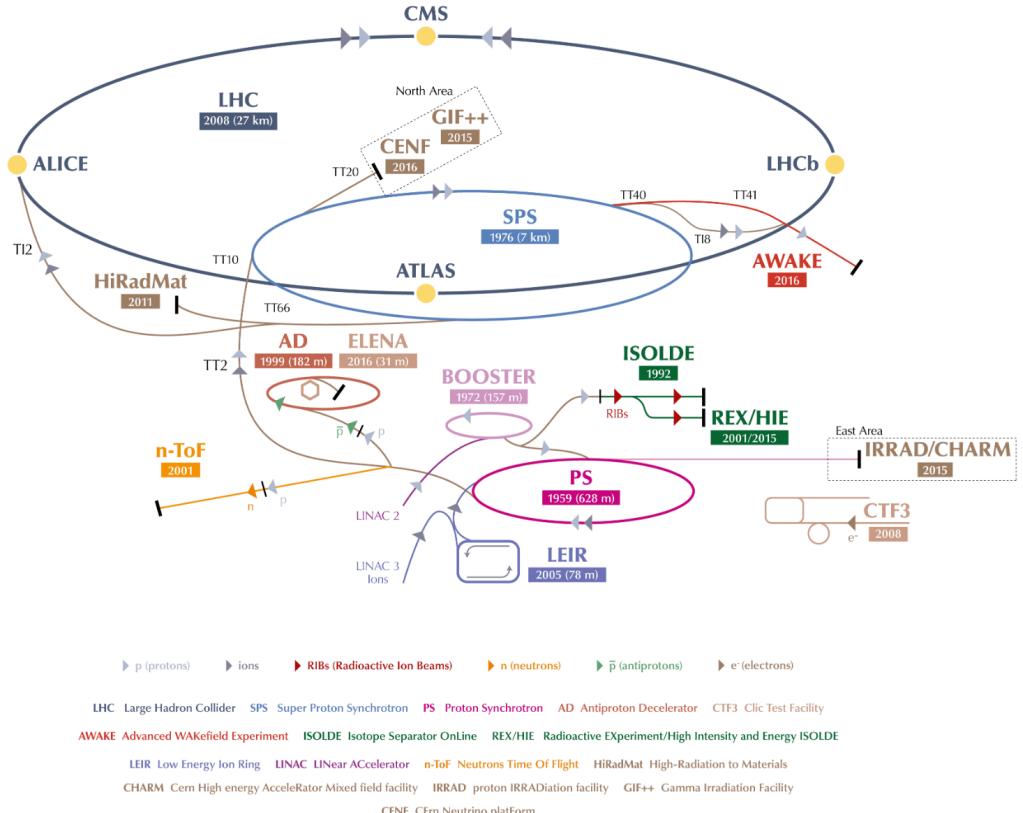


Fig. 3.2 CERN accelerator complex<sup>[20]</sup>.

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

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

### 3.1.2 Luminosity and pile-up

#### Luminosity

In beam-beam collisions, the event rate for a process is given by<sup>[19]</sup>:

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

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

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

where  $N_b$  denotes the number of particles per bunch,  $n$  is the number of bunches per beam,  $f_r$  is the revolution frequency,  $\gamma$  represents relativistic  $\gamma$  factor,  $\epsilon_n$  is the normalized transverse emittance and  $\beta^*$  denotes the  $\beta$  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  $\theta_c$  denotes the crossing angle at the interaction point,  $\sigma_Z$  is the root mean square (RMS) bunch length and  $\sigma^*$  is the transverse RMS beam size at crossing point.

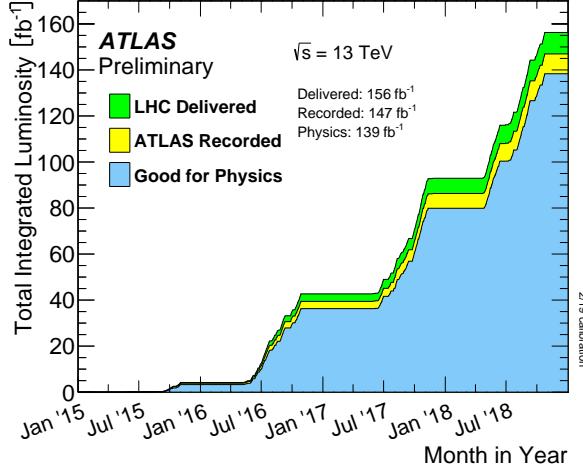
The luminosity expressed in Eq. 3.2 is normally the instantaneous luminosity. In fact the running conditions usually vary with time, so the luminosity can change as well. To take into account the time dependence, integrated luminosity is invited, by integrating the instantaneous luminosity over time:

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

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

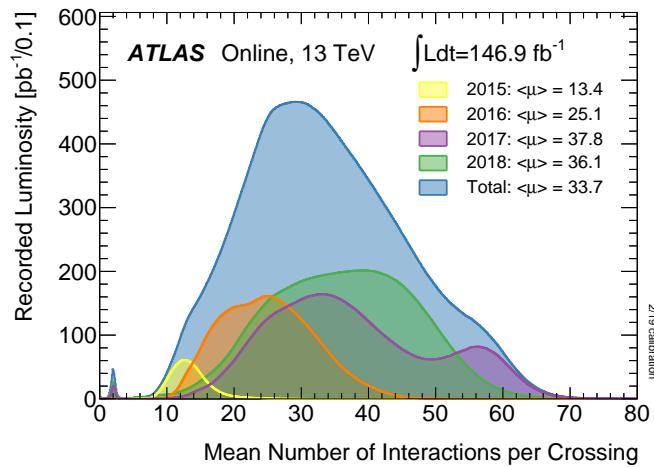


**Fig. 3.3 Integrated luminosity in ATLAS.**

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

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

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



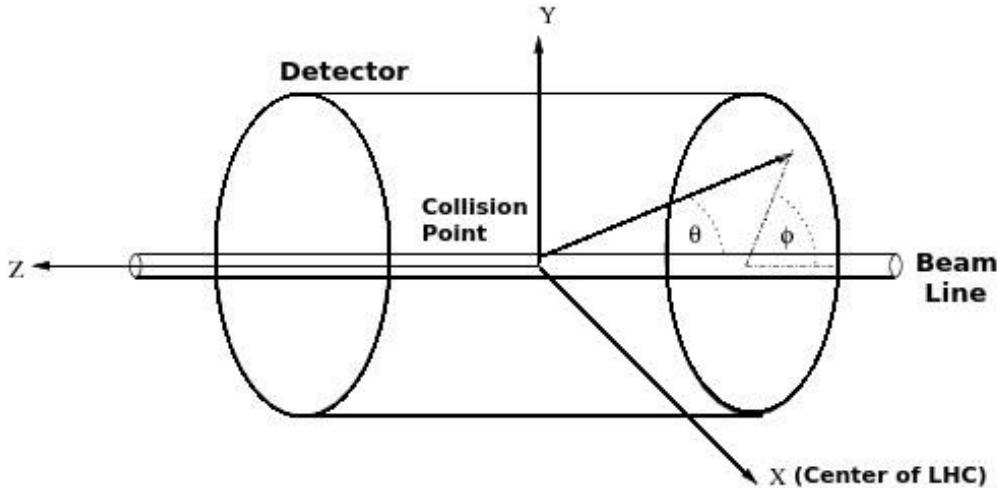
**Fig. 3.4 Number of Interactions per Crossing from 2015-2018 in ATLAS.**

557 **3.2 ATLAS detector**

558 **3.2.1 Detector overview**

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

563     This paragraph briefly summarizes the coordinate system and nomenclature used to  
 564    describe the ATLAS detector<sup>[21]</sup>. As depicted in figure 3.5, we define the nominal inter-  
 565    action point as the origin of the coordinate system, the beam direction as the  $z$ -axis and  
 566    the  $x$ - $y$  plane is transverse to the beam direction. The positive  $x$ -axis is defined to be the  
 567    direction pointing to the center of the LHC ring, while the positive  $y$ -axis is pointing up-  
 568    wards. There are two sides of detector A and C, in which A (C) -side is in the positive

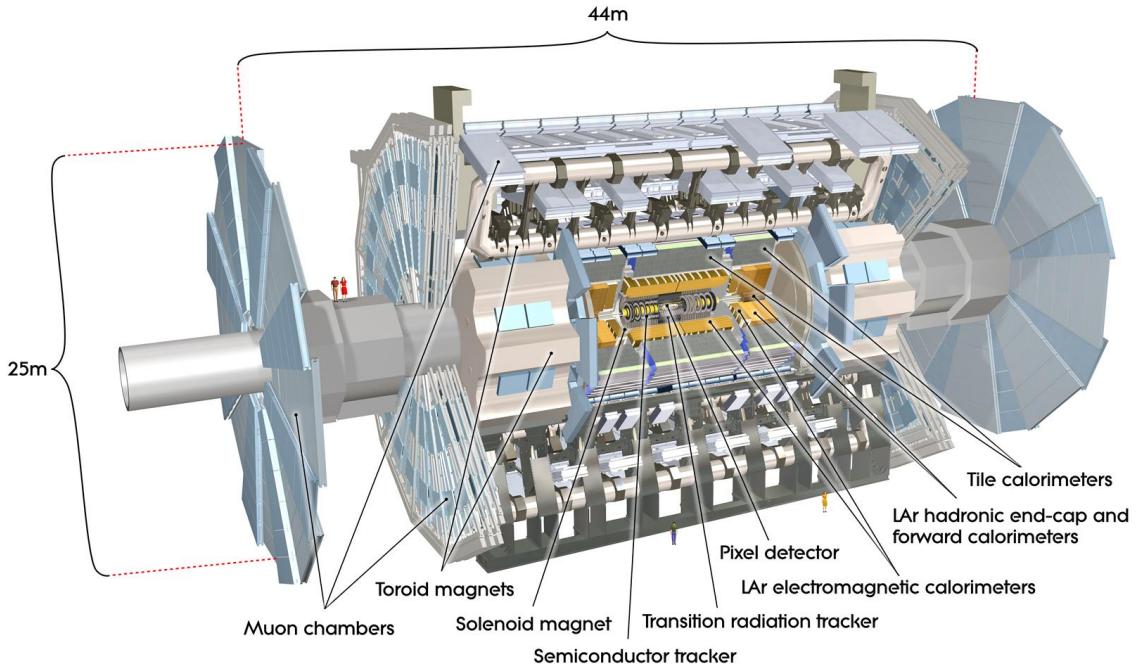


**Fig. 3.5** Coordinate system used by the ATLAS experiment at the LHC<sup>[21]</sup>.

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

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

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



**Fig. 3.6 Cut-away view of the ATLAS detector<sup>[23]</sup>.**

580 symmetry around the calorimeters.

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

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

591     **The muon spectrometer** is the outermost layer. It's a air-core toroid system, with  
 592     a long barrel and two inserted end-cap magnets that provides strong bending power in a  
 593     large volume within a light and open structure. A set of chambers measuring the tracks  
 594     of muons with high spatial precision and accurate time-resolution are used. Multiple-  
 595     scattering effects are minor, and excellent muon momentum resolution can be achieved.

### 3.2.2 Physics requirement

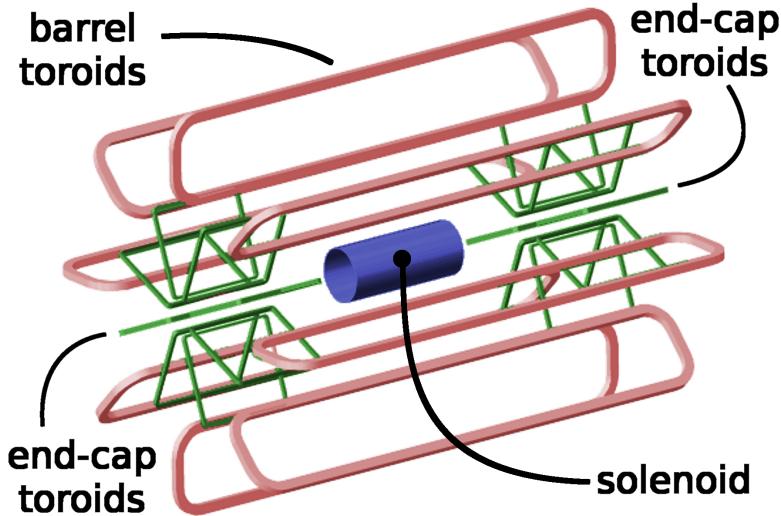
As mentioned previously, ATLAS is one of two general-purpose particle detector experiment at the LHC. It's designed to take advantage of the unprecedented energy at the LHC, as the discovery of Higgs boson is one of its benchmark. Lots of precise tests and measurements of SM physics are ongoing with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

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

More detailed descriptions of each sub-system will be given in the following subsections.

### 3.2.3 Magnet system

A strong magnetic field is required for precise measurement of charged particle momenta. The ATLAS detector uses two large superconducting magnet systems, a hybrid system of a central superconducting solenoid and three outer superconducting toroids, to bend charged particles<sup>[24]</sup>. The total magnet system is 22 m in diameter and 26 m in length as shown in figure 3.7.



**Fig. 3.7 Schematic diagram of the ATLAS magnet system.**

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

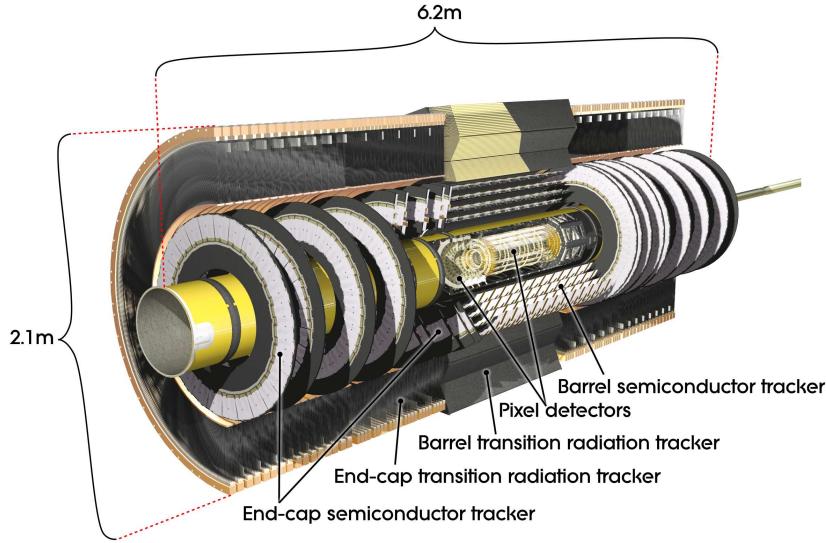
633        The outer toroid system comprises one barrel superconducting toroid and two end-  
 634        caps. The barrel one is composed of eight coils encased in individual racetrack-shaped,  
 635        stainless-steel vacuum vessels and produces the magnetic field in the cylindrical volume  
 636        surrounding the calorimeters. Each end-cap toroid consists of a single cold mass built up  
 637        from eight flat, square coil units and eight keystone wedges and provides a magnetic field  
 638        of approximately 1 T for the muon detectors in the end-cap regions.

### 639        3.2.4 Inner detector

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

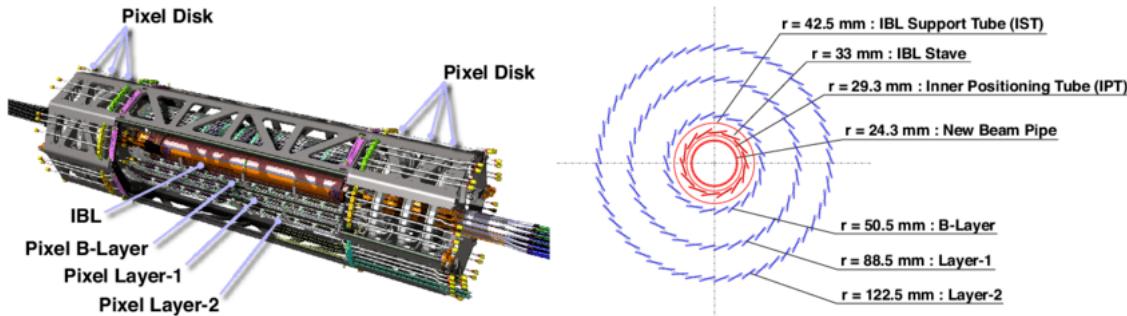
#### 647        **Pixel detector**

648        The pixel detector<sup>[26]</sup> is the innermost part of ATLAS tracking system. With finest  
 649        granularity of materials, it has the best spatial resolution and 3-dimensional space-point  
 650        measurement in inner detector. ATLAS Pixel Detector for the LHC run-2 is composed of  
 651        4 layers of barrel pixel detector and two end-caps with three pixel disks each, as shown in



**Fig. 3.8 Schematic diagram of the ATLAS inner detector<sup>[25]</sup>.**

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<sup>[27]</sup>. Now the 4-layer pixel detector has very good reconstruction of primary and secondary vertices, which is even crucial for long-lived particles like  $\tau$ -lepton and b-quark.



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

655

656

## Semiconductor Tracker

657

658

659

660

661

662

663

664

665

666

The Semiconductor Tracker (SCT)<sup>[28]</sup> installed outside the pixel detector is the middle component of the inner detector. It has similar function as pixel detector but with long and narrow strips instead of small pixels, which makes a much larger coverage than pixel detector. The SCT consists of 4088 modules, and contains four concentric layers in barrel (2112 modules) and nine disks in each of the two end-caps (1976 modules) as shown in figure 3.10. And it measures particles over a large area with 6.3 million readout channels and a total area of 61 square meters. The SCT is the most critical part of the inner detector for 2D track hit reconstruction. In barrel, the hit precision is  $17 \mu\text{m}$  in the  $r\phi$  coordinate 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.

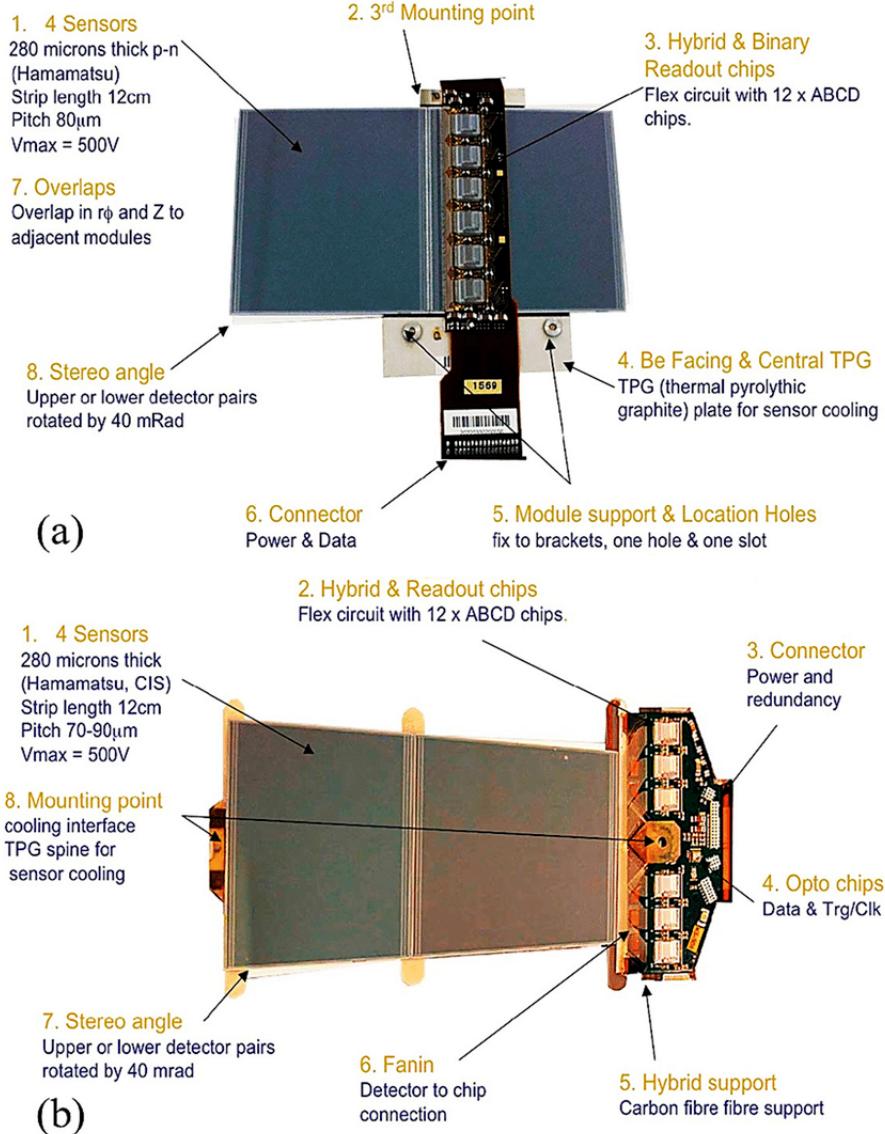


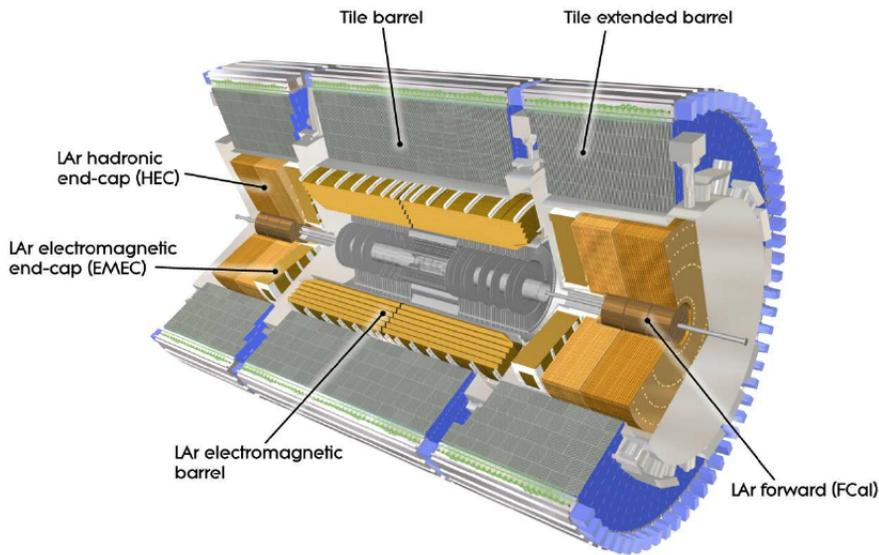
Fig. 3.10 SCT (a) barrel module and (b) end-cap<sup>[29]</sup>.

### 667 Transition radiation tracker

668 The transition radiation tracker (TRT)<sup>[30]</sup> is the outermost part of inner detector, which  
 669 has a very different design comparing to the two previously described sub-detectors. It  
 670 can be separated into three parts: one barrel and two end-cap regions with the  $|\eta|$  cov-  
 671 erage up to 2.0. There are 73 barrel layers and 224 end-cap layers (112 in each) with  
 672 372000 straws in total, and about 351000 readout channels for TRT. The TRT provides  
 673 better  $z$  resolution but much worse  $r\phi$  resolution (about 130  $\mu$ m) comparing to the pixel  
 674 detector and SCT per straw. But the straw hits still make significant contributions to mo-  
 675 mentum measurement, since its lower precision per point (compared to silicon) can be  
 676 compensated by the large number of measurements and long track length.

### 3.2.5 Calorimeters

The calorimeters are designed to measure the energy from particles by absorbing them. They are located outside the solenoidal magnet that surrounds the inner detector. The ATLAS calorimeters are comprised of a number of sampling calorimeters with full  $\phi$ -symmetry and the pseudorapidity range of  $|\eta| < 4.9$ . Figure 3.11 shows the layout of the ATLAS calorimeter system. There are two basic calorimeter systems: an inner electromagnetic (EM) calorimeter and an outer hadronic calorimeter. The EM calorimeter is designed for precise measurements of electrons and photons with fine granularity; while the hadronic one has relative coarser granularity but satisfies the physics requirements for jets reconstructions and  $E_T^{miss}$  measurements. Two different sampling techniques are used, the EM calorimeter is purely based on liquid-argon (LAr) technology, while the hadronic one uses both LAr and scintillating tiles calorimeters. More details are described as below:



**Fig. 3.11** Cut-away view of the ATLAS calorimeters. The LAr calorimeters are seen inside the scintillator-based tile hadronic calorimeters<sup>[31]</sup>.

#### Liquid Argon calorimeter

The LAr calorimeter uses liquid-argon as active medium. The LAr sampling calorimeter technique with “accordion-shaped” electrodes is used for all electromagnetic calorimetry covering the pseudorapidity range of  $|\eta| < 3.2$ ; and for hadronic calorimetry with range from  $|\eta| = 1.4$  to the acceptance limit  $|\eta| = 4.9$ <sup>[32]</sup>. Figure 3.12 depicts a segment of the barrel calorimeter, which had “accordion-shaped” electrodes and absorber. For barrel EM calorimeter, the absorbing material is lead-liquid argon, while the hadronic end-cap calorimeter uses copper plates. In addition, the forward calorimeter is split into

698 three parts, an EM sector in which copper is used as absorbing material and two hadronic  
sectors using tungsten outside the EM sector.

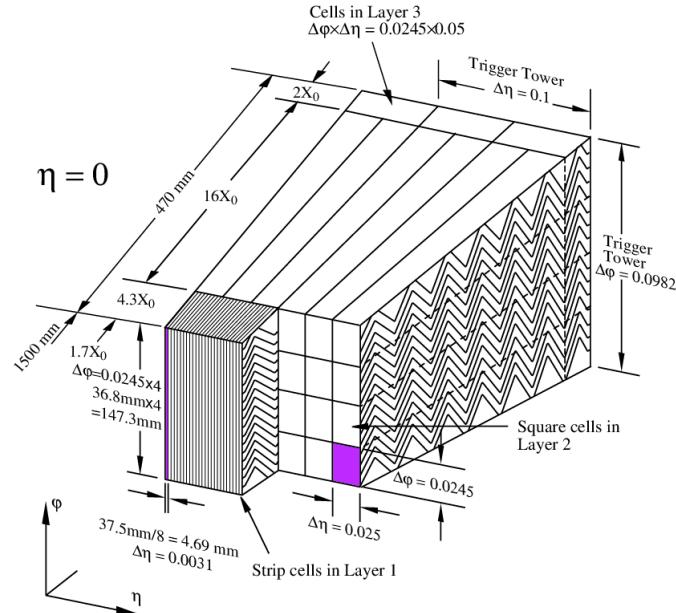


Fig. 3.12 Schematic diagram of a LAr EM calorimeter barrel module.

699

700

### Tile calorimeter

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

707

### 3.2.6 Muon spectrometer

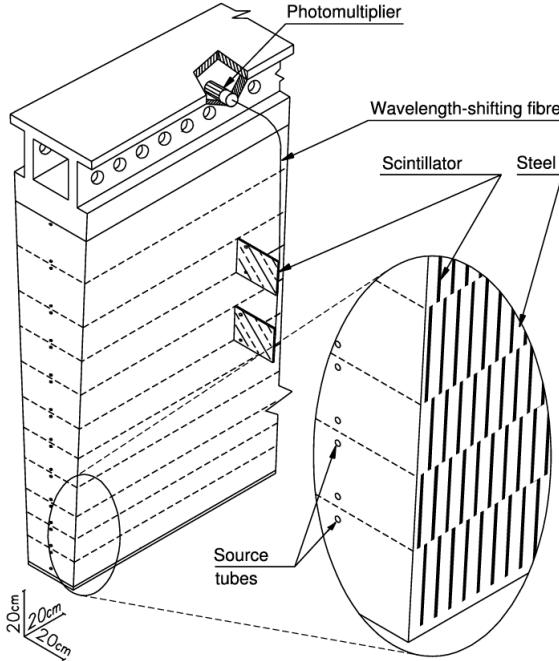
708 Muon spectrometer<sup>[34]</sup> is the outermost part of the ATLAS detector with an extremely  
709 large tracking system. It measures a large range of muon momentum, and the accuracy is  
710 about 3% at 100 GeV and 10% at 1 TeV. The muon spectrometer comprises three main  
711 parts: a magnetic field produced by three toroidal magnets; a set of chambers measur-  
712 ing the tracks of muons with high spatial precision; and triggering chambers with accu-  
713 rate time-resolution. Figure 3.14 shows the schematic of ATLAS muon spectrometer that  
714 consists of four types of muon chambers (*MDT*, *CSC*, *RPC*, *TGC*) as well as the magnet  
715 systems (barrel and end-cap toroid).

716

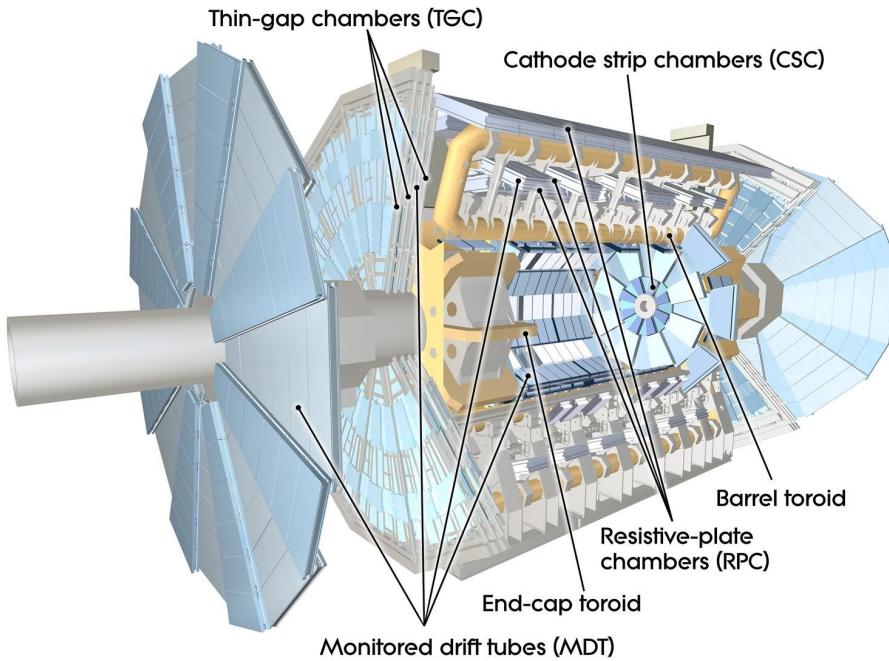
More details of four chambers are given as below:

717

- **Monitored Drift Tubes (MDT).** MDTs provide the precise momentum measure-



**Fig. 3.13 Schematic diagram of tile calorimeter module<sup>[33]</sup>.**



**Fig. 3.14 Cut-away view of the ATLAS muon spectrometer<sup>[35]</sup>.**

718      ment with the  $|\eta|$  range up to 2.7, except in the innermost end-cap layer where the  
 719      coverage is limited to  $|\eta| < 2.0$ . The chambers comprises three or four layers of  
 720      drift tubes, with a diameter of 29.970 mm, operated with Ar/CO<sub>2</sub> gas (93/7) at 3  
 721      bar. The average resolution can reach 80  $\mu\text{m}$  per tube and 30  $\mu\text{m}$  per chamber.

- 722      • **Cathode strip chambers (CSC).** CSCs are used in the forward region of  $2 < |\eta| <$   
 723      2.7 in the innermost tracking layers, due to their good time resolution and high

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

- **Resistive plate chambers (RPC).** The RPCs serve as fast triggers in the barrel region of  $|\eta| < 1.05$  due to the high rate capability and good spatial and time resolution. It is a gaseous parallel electrode-plate detector without any wires. There are three concentric cylindrical layers around the beam axis, as three trigger stations. Each station consists of two independent layers to measure the transverse coordinates of  $\eta$  and  $\phi$ .
- **Thin gap chambers (TGC).** TGCs are used as trigger system for the end-cap region of  $1.5 < |\eta| < 2.4$ , and works based on the same principle as multi-wire proportional chambers. In addition, they can also provide the second azimuthal coordinate to complement the measurement of MDT in bending direction.

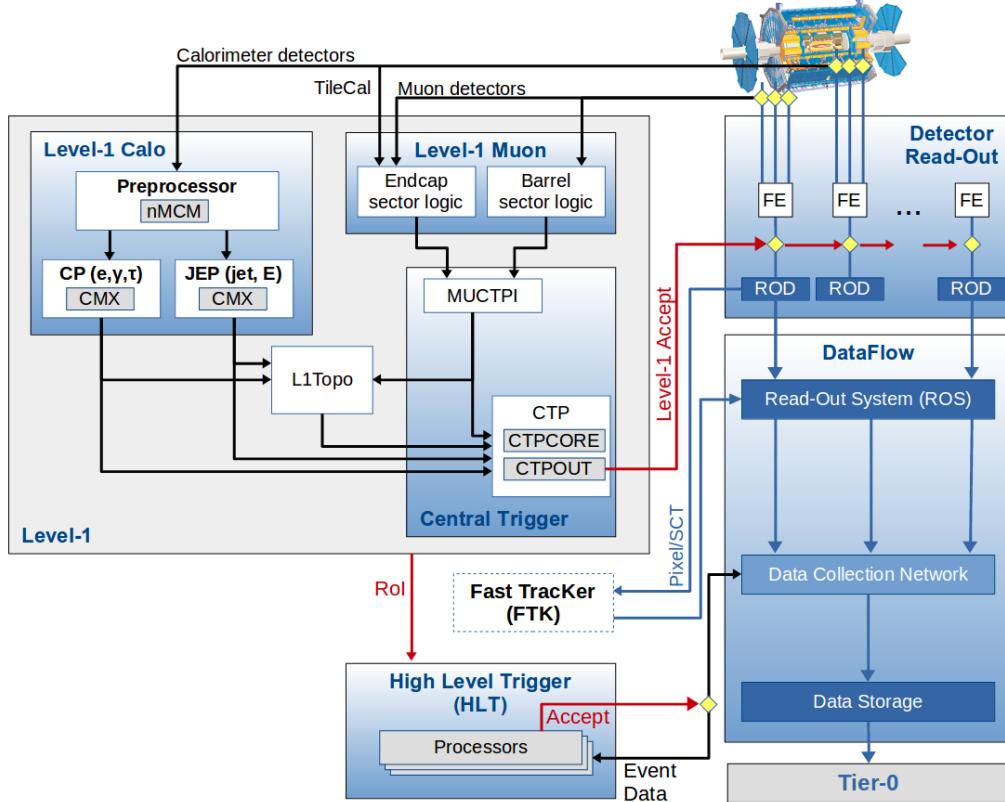
### 3.2.7 Trigger system

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 a large increase of event rate by up to a factor of five, 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)<sup>[36]</sup>. 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 30 MHz to 100 kHz. After that, the RoI information from Level-1 is sent to HLT, in which 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, allow the



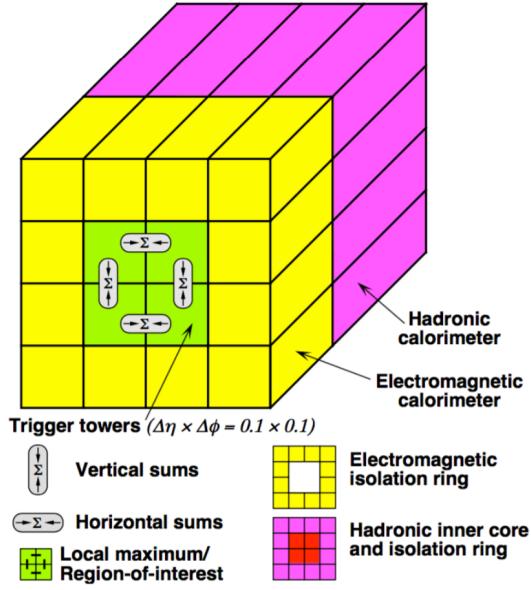
**Fig. 3.15 Schematic diagram of the ATLAS trigger and data acquisition system in run-2.**

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 reduced granularity information from the electromagnetic and hadronic calorimeters to search for electrons, photons, taus and jets and missing transverse energy ( $E_T^{miss}$ ). It can identify an Region-of-Interest (RoI) as a  $2 \times 2$  trigger tower cluster in the EM calorimeter as shown in figure 3.16, and  $4 \times 4$  or  $8 \times 8$  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) included in the modules used in run-1, which allows the use 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, 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% due to installing new chambers in the feet region of the muon detector.

### 775 High Level Trigger



**Fig. 3.16 An examples of L1 calorimeter trigger tower for electron and photon triggers<sup>[37]</sup>.**

776 In run-1, the Event Filter computer clusters and Level-2 trigger system were separated,  
 777 while now in run-2, they have been merged into a single HLT event processing. The new  
 778 arrangement helps to reduce the complexity and duplication of algorithm, which leads to a  
 779 more flexible high level trigger system. During the long-shutdown between the LHC run-  
 780 1 and run-2, lots of reoptimizations have been done for trigger reconstruction algorithms  
 781 as well as the offline analysis selections, which can improve the efficiency by more than  
 782 a factor of two in some cases like hadronic tau triggers. For some triggers, the HLT  
 783 processing performed within RoIs also allows to aggregate from RoIs to single objects.  
 784 This improvement reduces the CPU processing for events with overlapping RoIs, and the  
 785 average output rate has been increased from 400 Hz to 1 kHz.

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

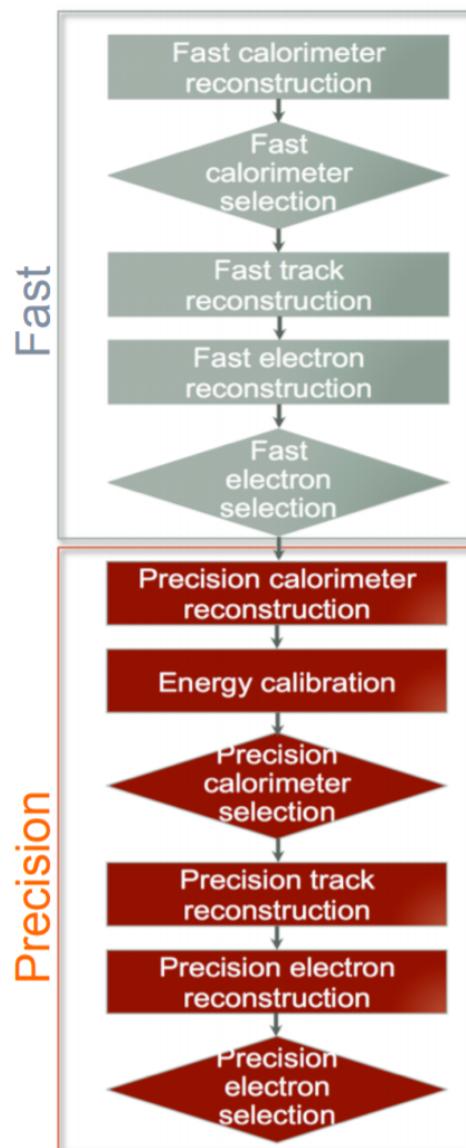


Fig. 3.17 The HLT trigger algorithm sequence<sup>[37]</sup>.

## 792      Chapter 4 Simulation and Event Reconstruction for 793      the ATLAS Experiment

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

### 799    4.1 Event simulation

800      The ATLAS simulation program is integrated into the ATLAS software framework  
801      called *Athena*<sup>[38]</sup>, which uses Python as an object-oriented scripting and interpreter lan-  
802      guage to configure and load C++ algorithms and objects. Figure 4.1 shows the overview  
803      of ATLAS simulation data flow<sup>[39]</sup>. In the diagrams, the square-cornered boxes represents  
algorithms and applications to be run and round-cornered boxes denote data objects.

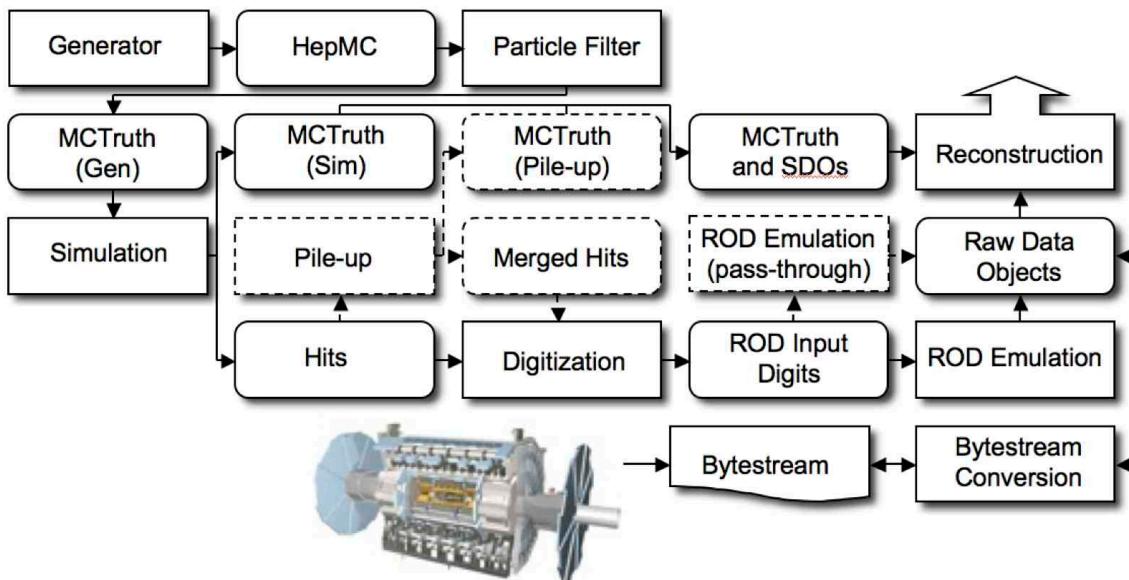


Fig. 4.1 The flow of the ATLAS simulation software.

804  
805      First of all, events are produced by MC generators in standard HepMC format and  
806      then read into the simulation. During the simulation, particles are propagated through the  
807      full ATLAS detector whose configurations can be set by users via GEANT4 toolkit. The  
808      energies deposited in the sensitive regions of the detector are recorded as *hits* that contains  
809      the total energy deposition, position and time, and are written to a simulation hit file. In

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

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

### Event generation

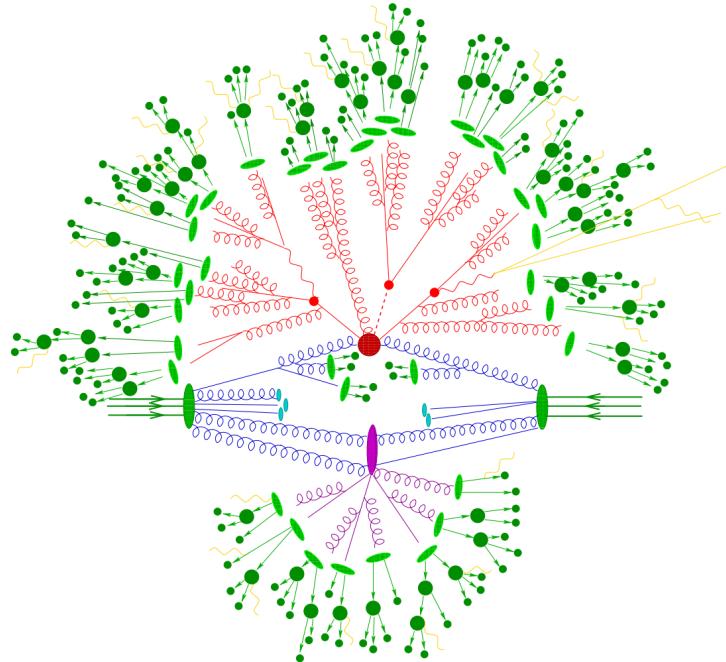
As shown in figure 4.2<sup>[40]</sup>, at hardon colliders, multiple scattering and rescattering effects arise, which needs to be simulated by Monte Carlo (MC) event generators to reflect the full complexity of those event structures. Several MC event generators can be used to generate events in HepMC format. The events can be filtered at generation time with some certain requirements (eg. decay channel or missing energy above a certain threshold). The generator is responsible for any prompt decays (e.g. W or Z bosons) and stores any “stable” particle expected to propagate through a part of the detector. During the generation steps, any interactions with detector are ignored and only immediate decays are considered.

There are several MC generators that have been widely used with general purpose, including Sherpa<sup>[41]</sup>, Herwig++<sup>[42]</sup>, PowhegBox<sup>[43]</sup>, MC@NLO<sup>[44]</sup> and Pythia8<sup>[45]</sup>.

### Simulation

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

Since the standard ATLAS detector simulation cost very large computing resources to accurately model the complex detector geometry and physics descriptions, some fast simulation programss are developed according to different user purpose. Some popular fast-sim toolkits include *Fast G4 Simulation*<sup>[46]</sup>, *ATLFAST-I*<sup>[47]</sup> and *ATLFAST-II*<sup>[48]</sup>.



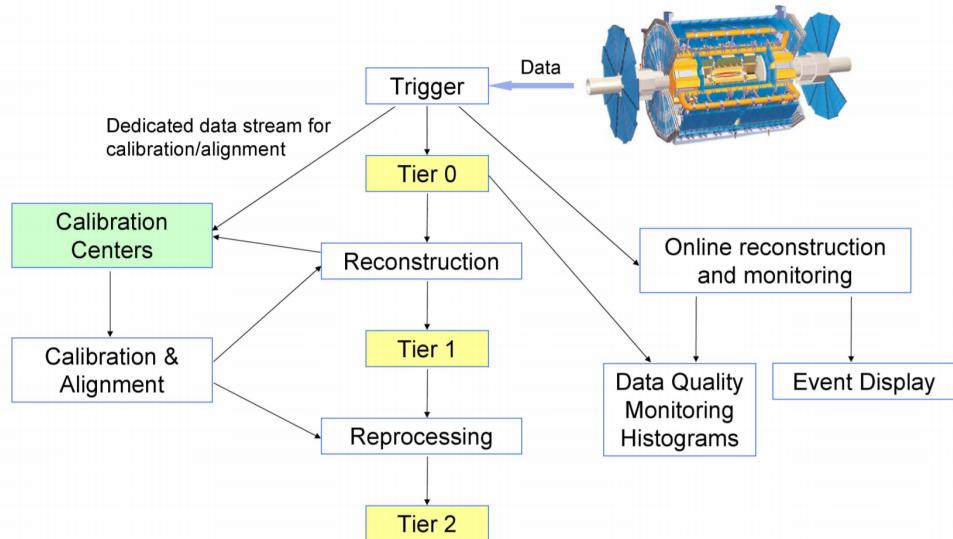
**Fig. 4.2 Sketch of a hardon-hardon collision simulated by MC event generator. The red blob in center denotes the hard collision, surrounded by tree-like structures representing Bremsstrahlung which is simulated by Parton Showers. The purple blob stands for a secondary hard scattering event. The light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

### 845      **Digitization**

846      The hit outputs from simulated events, including hard scattering signal, minimum  
 847 bias, beam halo, beam gas and cavern background events, are then sent into digitization  
 848 procedure, converted into detector response called “digits”. Before converted into detec-  
 849 tor signal as “digits” formart, each type of event can be overlaid at a user-specified rate.  
 850 Those overlay, called “pile-up”, can be done during digitization to save the CPU time. At  
 851 this stage, the detector noise and the first level trigger that implemented with hardware on  
 852 the real detector are added into events. The digitization firstly constructs “digits” inputs to  
 853 the readout drivers (RODs) in the detector electronics. Then the ROD functionality is em-  
 854 ulated, and the output digits are written out as Raw Data Object (RDO) file. In addition,  
 855 the digitization algorithms can also produce Simulated Data Objects (SDOs), which con-  
 856 tain information about all the particles, noise and the amount of energy that contributed  
 857 to the signal. Then all information are sent into reconstruction level described in next  
 858 subsection.

## 4.2 Event reconstruction

The data flow of ATLAS data processing is sketched in figure 4.3<sup>[49]</sup>. Data from detector is firstly filtered by online trigger system before sending to the *Tier-0* (*T0*) for initial processing by offline reconstruction software based on Athena. A small amount of data named “express stream” is processed in almost real time in *T0* for online data quality monitoring. In addition, some other dedicated data streams are sent out at trigger level for detector alignment and calibration. These calibration and alignment information are then used for bulk reconstruction in *T0*. At the end of the reconstruction chain, the data are delivered into *Tier-1* (*T1*) and *Tier-2* (*T2*) centers 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.**

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

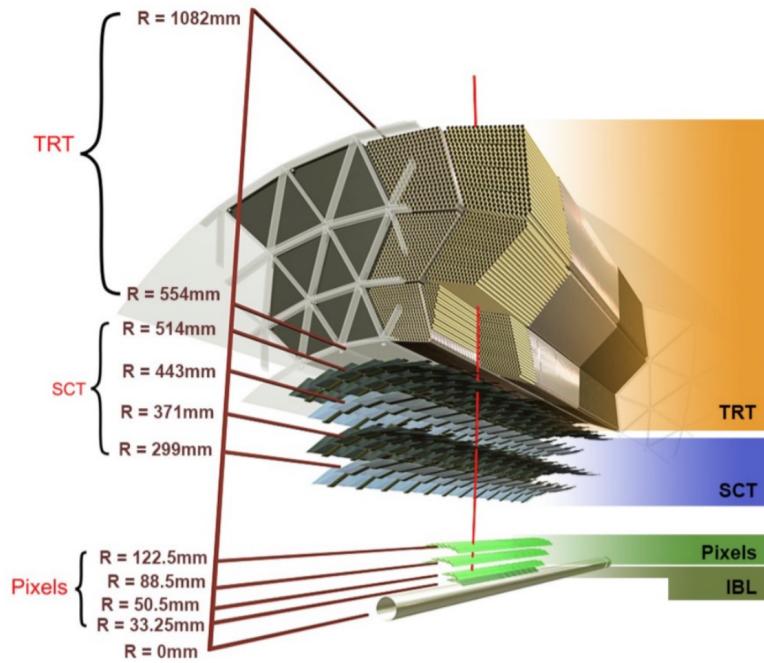
### 4.2.1 Track

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

880 amount of inert material, the large background and the highly inhomogeneous magnetic  
 881 field<sup>[50]</sup>. More details of these two types of track are given as below:

882     **Inner detector track**

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



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

885  
 886     For inside-out tracking, it exploits the high granularity of the pixel and SCT detectors  
 887 to discover prompt tracks originating from the interaction point. In first step, the track  
 888 seeds are formed by combining the information of space-points in the three pixel layers  
 889 and the first SCT layer. Then, these seeds are extended throughout the SCT to build track  
 890 candidates. After that, these candidates are fitted with some quality cuts applied to remove  
 891 the outlier clusters, reject the fake tracks and resolve ambiguities in the cluster-to-track  
 892 association. The selected tracks are then further extended to TRT, and refitted with the  
 893 full information from pixel, SCT and TRT detectors.

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

898     **Muon spectrometer track**

899     The MS track reconstruction<sup>[51]</sup> starts from searching hit patterns inside each muon  
900    chamber to form segments. In each MDT chamber and nearby trigger chamber, a Hough  
901    transform<sup>[52]</sup> is used to search the hits lie on a certain trajectory in the bending plane of the  
902    detector. The MDT segments are reconstructed by performing a linear fit to the hits found  
903    in each layer. The RPC or TGC hits can be built by measuring the coordinate orthogonal to  
904    the bending plane. And the segments of CSC can be built using a separate combinatorial  
905    search in the  $\eta$  and  $\phi$  detector planes.

906     Then muon track candidates are built by fitting hits from segments in different layers  
907    together. This task makes use of the algorithm by performing a segment-seeded combi-  
908    natorial search, which starts by using the segments generated in the middle layers of the  
909    detector where more trigger hits are available as seeds. The search is then extended to use  
910    the segments as seeds from the inner and outer layers. The segments are selected based  
911    on criteria of hit multiplicity and fit quality, and are matched using their relative positions  
912    and angles. To build a track, at least two matching segments are required, except in the  
913    barrel-endcap transition region where a single high-quality segment with  $\eta$  and  $\phi$  infor-  
914    mation can be used to build a track. At beginning, the same segment can be used to build  
915    more than one track candidates. Later on, an overlap removal algorithm is performed to  
916    select the best assignment to a single track, or decide whether allows the certain segment  
917    to be shared between two tracks.

918     The hits associated with each track candidate are then fitted using a global  $\chi^2$  fit. The  
919    algorithm accepts the track candidate if its fitting  $\chi^2$  passes the selection criteria. Hits  
920    with large contribution to  $\chi^2$  are removed and the track fit is repeated. In addition, the  
921    algorithm performs a hit recovery procedure looking for additional hits consistent with  
922    the candidate trajectory, and the track candidate is refit if additional hits are found.

## 923       4.2.2 Primary vertex

924     The reconstruction of primary vertex (PV) uses the reconstructed tracks introduced in  
925    previous section as inputs. The tracks must satisfy the following criteria<sup>[53]</sup>:

- 926
  - $p_T > 400$  MeV
  - $|\eta| < 2.5$
- 927
  - Number of silicon hits  $\geq \begin{cases} 9 & \text{if } |\eta| \leq 1.65 \\ 11 & \text{if } |\eta| > 1.65 \end{cases}$
- 928
  - IBL hits + B-layer hits  $\geq 1$
- 929
  - A maximum of 1 shared module (1 shared pixel hit or 2 shared SCT hits)
- 930
  - Pixel holes = 0

- 932     • SCT holes  $\leq 1$

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

934     The reconstruction of PV can be divided into two steps<sup>[54]</sup>: vertex finding and vertex  
935 fitting. The first step is the pattern recognition process, namely the association of recon-  
936 structed tracks to vertex candidates. The latter one works on the reconstruction of the  
937 actual vertex position and its covariance matrix. More details are described as below:

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

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

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

961     

### 4.2.3 Electron

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

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

968 In ATLAS, in central region, the electrons leave tracks in inner detector (ID) and  
969 deposit the energies in the electromagnetic (EM) calorimeter. Firstly the signals from  
970 calorimeter are used for L1 trigger system, and then combined with the information from  
971 ID tracks to reconstruct electron candidates that will be used for the high level trigger  
972 (HLT) decision algorithms<sup>[55]</sup>. The backgrounds mentioned above can then be further  
973 suppressed by using several identification criteria. In addition, electrons are required to  
974 be isolated from other activities to be further distinguished from background.

975 More details of electron *reconstruction*, *identification* and *isolation* are described as  
976 below.

### 977 Electron reconstruction

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

- 980 1. **Seed-cluster reconstruction:** A sliding window with size of  $3 \times 5$  in unit of  
981  $\Delta\eta^{tower} \times \Delta\phi^{tower} = 0.025 \times 0.025$  in  $\eta \times \phi$  space is utilized to search for elec-  
982 tron cluster seeds with total cluster transverse energy greater than 2.5 GeV. Then  
983 a clustering algorithm<sup>[56]</sup> is applied to form the clusters around the seeds, which  
984 can take advantage of removing the duplications. The kinematics of clusters are  
985 then reconstructed by using an extended window depending on the cluster posi-  
986 tion. The efficiency of cluster search is from about 95% at  $E_T = 7GeV$  to 99% for  
987  $E_T \geq 15GeV$ .
- 988 2. **Track reconstruction:** The track reconstruction can be divided into two steps: pat-  
989 tern recognition and track fit. The standard pattern recognition in ATLAS uses pion  
990 hypothesis for energy loss caused by interactions with detector material. If a track  
991 seed with  $p_T > 1$  GeV cannot be successfully extended to a full track required at  
992 least seven hits using this pion hypothesis, but still falls inside one of the EM clus-  
993 ter region of interest, as a second attempt, the pattern recognition using electron  
994 hypothesis is then used to allow larger energy loss. Depending on the pattern used  
995 in previous stage, the track candidates are then fitted with either the pion hypothesis  
996 or the electron hypothesis by using ATLAS Global  $\chi^2$  Track Fitter<sup>[57]</sup>. If a track  
997 candidate fails the fit by using pion hypothesis, it can be refit with the electron hy-  
998 pothesis again. In this method, a specific electron-oriented algorithm is integrated  
999 into the ATLAS standard track reconstruction, which improves the performance for  
1000 electron and as well as maintain minimal interference with the main track recon-

1001 struction.

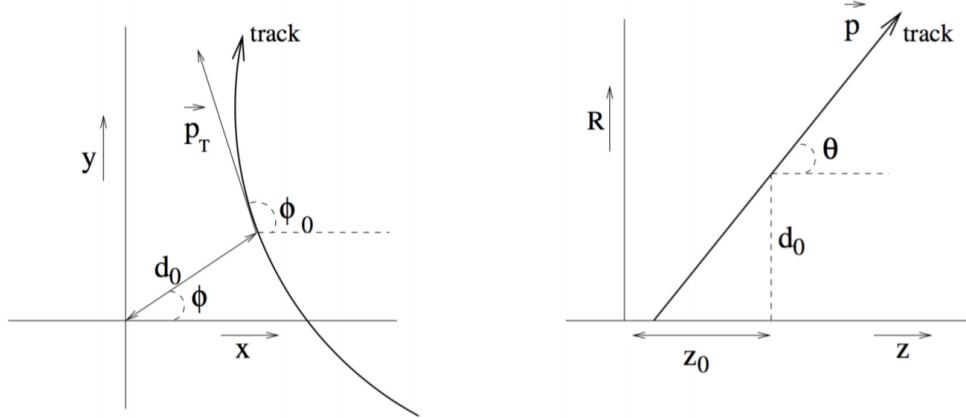
1002 **3. Electron specific track fit:** Once the tracks are obtained, they are loosely matched  
 1003 to EM cluster using the distance in  $\eta$  and  $\phi$  between the position of track (after ex-  
 1004 trapolation) in calorimeter's middle layer and the cluster barycentre. The matching  
 1005 conditions take into account the energy loss of bremsstrahlung and the number of  
 1006 precise hits in silicon detector.

1007 **4. Electron candidate reconstruction:** The electron candidate is reconstructed by  
 1008 matching the track candidate to EM cluster seed to eventually completes the electron  
 1009 reconstruction procedure. If more than one track satisfy the matching condition,  
 1010 one track is chosen as primary track based on the information of the cluster-track  
 1011 distance R, the number of pixel hits and the presence of a hit in the first silicon  
 1012 layer<sup>[58]</sup>. In addition, the electron candidates are removed from electron pool if it's  
 1013 without any associated precise hit tracks, and moved into photon candidates pool.  
 1014 Then we reformed the electron clusters by using  $3 \times 7$  ( $5 \times 5$ ) longitudinal towers  
 1015 of cells in barrel (end-caps) in EM calorimeter. The measured energy is calibrated  
 1016 to original electron energy based on MC simulated samples by using multivariate  
 1017 techniques (MVA).

1018 In addition, in physics analysis, to reduce the background from photon conversions  
 1019 and secondary particles, the track associated with electron is required to be compatible  
 1020 with the primary vertex of the hard collision. Practically, the impact parameters cuts such  
 1021 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  
 1022 the track to the measured beam-line,  $z_0$  is the distance along the beam-line between the  
 1023 point where  $d_0$  is measured and the beam-spot position, and  $\theta$  is polar angle of the track,  
 1024  $\sigma_{d_0}$  denotes the estimated uncertainty of  $d_0$  parameter. Figure 4.5 depicts the definition of  
 1025 each track impact parameter.

## 1026      **Electron identification**

1027 The electron identifications are applied to determine whether the reconstructed elec-  
 1028 tron candidate is more signal-like or background-like object. The identification algorithms  
 1029 make use of quantities of related variables from electron cluster and track measurements  
 1030 including calorimeter shower shapes, track properties, as well as variables measuring  
 1031 bremsstrahlung effects for distinguishing signal from background. Taking the advantage  
 1032 of new IBL in run-2, the number of hits in this innermost pixel layer is utilized for discrim-  
 1033 inating between electrons and converted photons. In addition, a likelihood method based  
 1034 on the TRT high-threshold hits is adopted to compensate the lower transition radiation  
 1035 absorption probability of the argon.



**Fig. 4.5 Schematic of the impact parameters of a track in the transverse plane (left) and RZ-plane (right), as defined in the global ATLAS tracking frame<sup>[59]</sup>.**

1036 The baseline identification algorithm introduced for ATLAS run-2 data analysis is the  
 1037 likelihood-based (LH) method, which uses a MVA technique to simultaneously evaluate  
 1038 several properties of electron candidates when making a decision. The LH method utilizes  
 1039 the probability density functions (PDFs) of signal and background as the input discrimi-  
 1040 nating variables. Based on these PDFs, it can calculate overall probabilities of the object  
 1041 to be signal or background. Then the probabilities of signal and background are combined  
 1042 together into a discriminant  $d_{\mathcal{L}}$ :

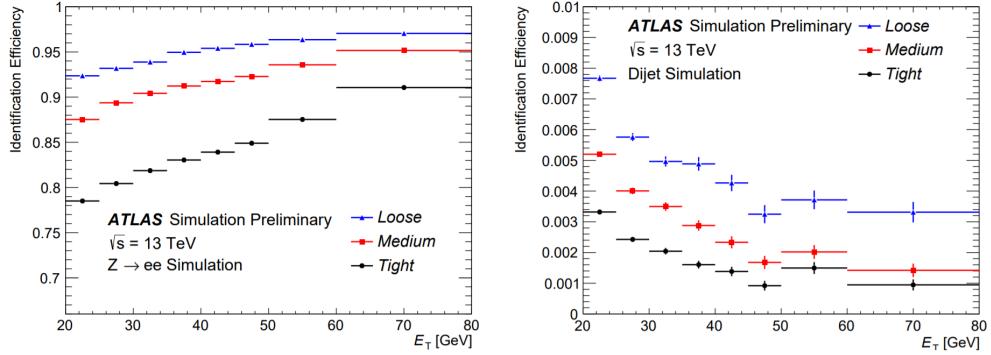
$$d_{\mathcal{L}} = \frac{\mathcal{L}_S}{\mathcal{L}_S + \mathcal{L}_B}, \quad \mathcal{L}_{S(B)}(\mathbf{x}) = \prod_{i=1}^n P_{s(b),i}(x_i) \quad (4.1)$$

1043 where  $\mathbf{x}$  denotes the vector of discriminating variables and  $P_{s(b),i}(x_i)$  represents the value  
 1044 of signal (background) PDF of the  $i^{th}$  variable as  $x_i$ .

1045 Three levels of working points (WPs) for electron identification are provided: *Loose*,  
 1046 *Medium* and *Tight*, in order of increasing background rejection. Samples selected by a  
 1047 looser WP are subsets of a tighter one, for example, the electrons passing Medium can  
 1048 all be selected by Loose. The identification efficiency varies as function of transverse  
 1049 energy ( $E_T$ ) as shown in figure 4.6. For evaluations, the electron candidates from MC  
 1050 simulation of  $Z \rightarrow ee$  decays (di-jet) are used as signal (background). Depending on the  
 1051 working point, the signal (background) efficiencies for reconstructed electron candidates  
 1052 at  $E_T = 25 GeV$  are in the range of 78 to 90% (0.3 to 0.8%), and increase (decrease) with  
 1053  $E_T$ .

### 1054 **Electron isolation**

1055 In addition to the identification criteria, most analyses have electron isolation require-  
 1056 ment to further distinguish signal from background. To quantify the energy of particles



**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).

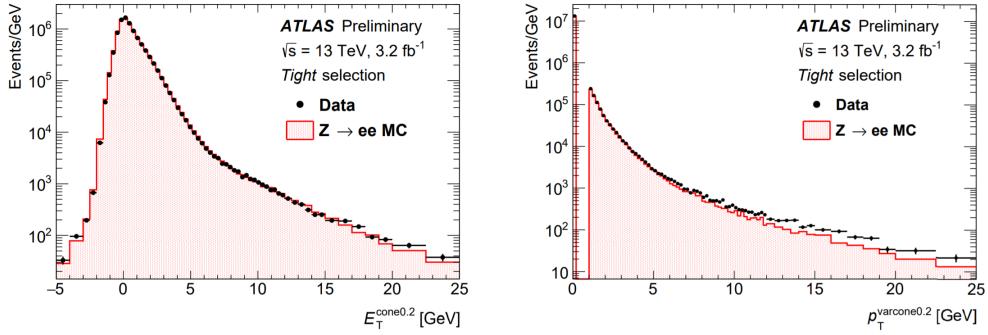
1057 around the electron candidate, the isolation variables can help to separate the prompt elec-  
 1058 tron from other non-isolated electrons, like the electrons from converted photons or from  
 1059 heavy flavour hadron decays. There are two kinds of discriminating variables that have  
 1060 been designed:

1061 • **Calorimeter-based variable:**  $E_T^{topocone20}$ . It's defined as the sum of transverse en-  
 1062 ergies of topological clusters<sup>[60]</sup>, calibrated at EM scale within a cone of  $\Delta R = 0.2$   
 1063 around the candidate electron cluster. It only considers the clusters with positive  
 1064 reconstructed energy. In addition, a correction as a function of  $(E_T, \eta)$  values is  
 1065 applied to account for the electron energy leakage outside the cluster.

1066 • **Track-based variable:**  $p_T^{varcone20}$ . It's calculated as the sum of transverse momen-  
 1067 tum of all satisfied tracks within a cone of  $\Delta R = \min(0.2, 10\text{GeV}/E_T)$  around the  
 1068 candidate electron track. To calculate the sum, it requires the tracks are originating  
 1069 from the reconstruction PV of hard collision, and exclude the associated tracks of  
 1070 electron itself.

1071 Based on the values of  $E_T^{topocone20}/p_T$  and  $p_T^{varcone20}/p_T$ , a series of working points with dif-  
 1072 ferent selection requirements are defined. The resulting WPs are divided into two kinds:  
 1073 • Efficiency targeted working points: varying requirements to obtain a certain isola-  
 1074 tion efficiency, which can either be a constant or as a function of  $E_T$ .  
 1075 • Fixed requirement working points: set the constant upper thresholds on isolation  
 1076 variables.

1077 The distribution of two discriminating variables are shown in figure 4.7 for  $ZZ \rightarrow ee$   
 1078 events with  $E_T > 27\text{GeV}$  and satisfying *Tight* requirement.



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

#### 4.2.4 Muon

Muons are distinctive signatures in final states of many physics analyses at the LHC including the Higgs analyses, SM measurements, BSM searches and so on. High performance of muon reconstruction and identifications are crucial. This section briefly describes some more details of the *reconstruction, identification and isolation* of muon.

##### Muon reconstruction

Muon reconstruction is firstly performed in inner detector (ID) and muon spectrometer (MS) independently as given in section 4.2.1. The information from each individual detector is then combined together to form the muon tracks for physics analyses. The combined ID-MS reconstruction is developed according to several algorithm based on the information from ID, MS and calorimeters. Four different muon types are defined<sup>[51]</sup>:

- **Combined (CB) muons:** a combined track is formed by using the reconstructed tracks performed independently in ID and MS with a global refit. To improve the fit quality, the hits from MS may be added to or removed from the track. The outside-in pattern recognition is utilized for the reconstruction of most muons, in which the muons are first reconstructed in MS and then extrapolated inward to match the ID track. In the meantime, the inside-out pattern is also used as a complementary method.
- **Segment-tagged (ST) muons:** a reconstructed track in ID is defined as muon, if it can be associated with at least one track segment in MDT or CSC chambers. These ST muons are used when they can only pass across one layer of MS chambers due to their low  $p_T$  or falling into regions with less MS acceptance.
- **Calorimeter-tagged (CT) muons:** a reconstructed track in ID is categorized as muon if it's matched to the energy deposit in calorimeter which is recognized with

a minimum-ionizing particle. This CT muons have lowest purity amount all types of muons, but it covers the region where ATLAS muon spectrometer is only partially constructed. For the region of  $|\eta| < 0.1$  and  $15GeV < p_T < 100GeV$ , the identification of CT muons are optimal.

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

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

### Muon identification

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

- *q/p significance*, the absolute difference between q/g (charge over momentum) of muons measured in ID and MS divided quadratic sum of their corresponding uncertainties;
- $\rho'$ , the absolute value of difference between the  $p_T$  (transverse momentum) measured in ID and MS, divided by the  $p_T$  of combined track;
- *Nomalized  $\chi^2$*  of the combined track fit;
- *Number of hits in ID and MS*

In addition, some new variables used for *LowPt* muon working point what will be described later<sup>[61]</sup>:

- *Momentum balance significance (MBS)* is computed as momentum difference between the ID and MS standalone measurements with respect to the uncertainty  $\sigma$  on energy lost in the calorimeter system.
- *Scattering neighbor significance (SNS)* is defined to estimated the significance of a change in trajectory along the track, expected in the presence of a hadron decaying

1138 to a muon.

- 1139 • *Scattering curvature significance (SCS)* is defined as the normalized integral of the  
1140 scattering angle significances, corrected for large kinks along the trajectory.

1141 Five selection working points are developed to satisfy the different needs for different  
1142 physics goals: *LowPt*, *Loose*, *Medium*, *Tight* and *HighPt*. The *Tight*, *Medium*, *Loose* are  
1143 subsets from the tighter one to looser one. More detailed definition of each working point  
1144 is given as follow:

- 1145 • *Loose*: this working point is designed to maximize the reconstruction efficiency  
1146 while keeping good-quality of muon tracks. And they are specifically developed  
1147 for reconstructing the Higgs boson candidates from four-lepton final states. All  
1148 four muon types are used for this selection level. The CB and ME muons passing  
1149 Medium WP that will mentioned below are all included into Loose category. In  
1150 addition, the CT and ST muons are restricted to  $|\eta| < 0.1$  region. In the range of  
1151  $|\eta| < 2.5$ , around 97.5% Loose muons are CB muons, and about 1.5% are CT while  
1152 remaining 1% are ST muons.

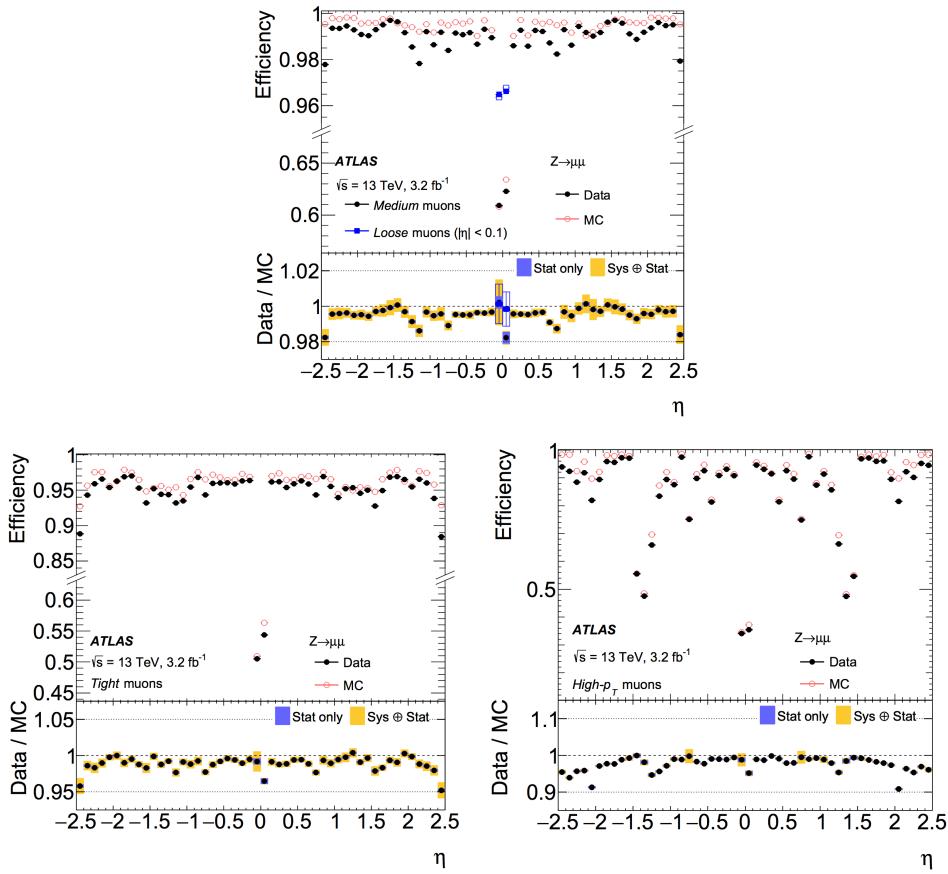
- 1153 • *Medium*: this working point is the default criteria of muon identification in ATLAS.  
1154 This selection minimizes the systematic uncertainties of muon reconstruction and  
1155 calibration. In this category, we only use CB and ME muons. For CB muons, at  
1156 least 3 hits in at least two layers of MDT are required, except  $|\eta| < 0.1$  region,  
1157 in which tracks with  $\geq 1$  MDT layer but  $\leq 1$  MDT hole layer are allowed. For  
1158 ME muons, at least 3 MDT/CSC layers are required. Furthermore, a loose cut on  
1159 the compatibility between measured momentum in ID and MS is applied to reduce  
1160 the fake muons from hadrons misidentification. Besides, the q/p-significance is  
1161 required to be less than 7.

- 1162 • *Tight*: this working point is used to maximize the purity of muons but with sacrifice  
1163 of some selection efficiency. Only CB muons with hits in  $\geq 2$  stations of MS and  
1164 passing Medium criteria are selected. In addition, the normalized  $\chi^2$  of combined  
1165 track fit should be smaller than 8. Then, a two-dimensional cut of q/p-significance  
1166 and  $\rho'$  is adopted as a function of muon  $p_T$  to ensure tighter background rejection  
1167 for momentum below 20 GeV, in which the fake rate is usually higher.

- 1168 • *High- $p_T$* : this set of selections aims to maximize the momentum resolution for  
1169 tracks with  $p_T > 100\text{GeV}$  region. The selection is especially optimized for search-  
1170 ing high-mass  $Z'$  and  $W'$  resonances. CB muons satisfying Medium selection and  
1171 with  $\geq 3$  hits in 3 MS stations are chosen. The specific region in MS where align-  
1172 ment is suboptimal are removed as a precaution.

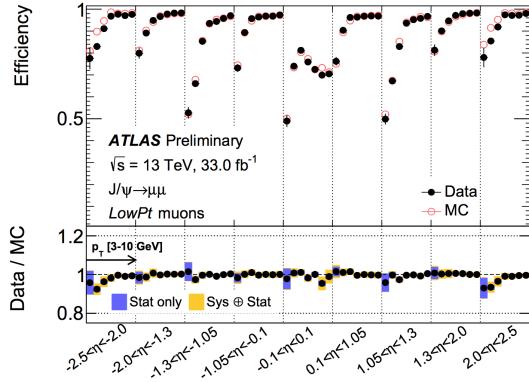
- *Low- $p_T$* : this type of muon is newly designed for physics analyses with ATLAS software release version 21. It's designed to obtain a optimal muon identification with very low transverse momentum of  $3GeV < p_T < 5GeV$ , which is crucial for B-physics measurement in ATLAS. In this muon requirement, 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 working points. For *Medium (Loose)*, *Tight* and *High- $p_T$* :  $Z \rightarrow \mu\mu$  events with  $p_T > 10GeV$  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  $3GeV < p_T < 10GeV$  are used for measurement.



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

## Muon isolation



**Fig. 4.9 Muon reconstruction efficiency for Low- $p_T$  working point as a function of  $\eta$ .**

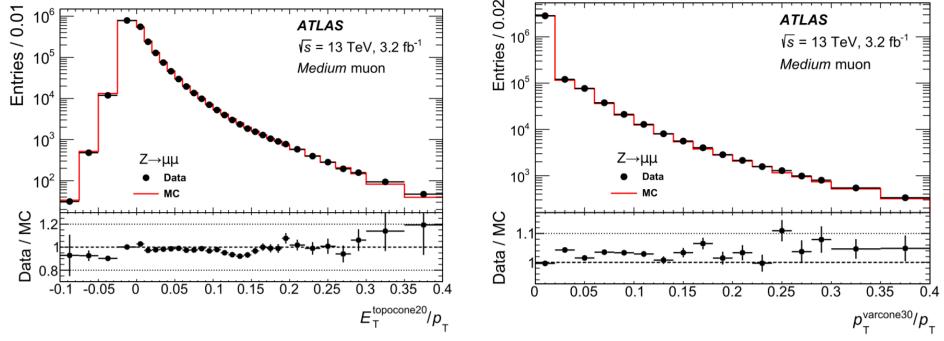
Similar to electron, the muon isolation is used to further distinguish the prompt muon from non-prompt backgrounds. There are also two types of isolation variables for muon:

- **Calorimeter-based variable:**  $E_T^{topocone20}$ . It's defined as the sum of the transverse energy of topological clusters within a cone of size  $\Delta R = 0.2$  around the candidate muon, after subtracting the contribution from the energy deposit of the muon itself and correcting for pile-up effects. The contributions from pile-up and underlying events are computed using the ambient energy-density technique<sup>[62]</sup> and are corrected on an event-by-event basis.
- **Tracked-based variable:**  $p_T^{varcone30}$ . It's computed as the scalar sum of the transverse momenta of the tracks with  $p_T > 1\text{GeV}$  in a cone size of  $\Delta R = \min(10\text{GeV}/p_T^\mu, 0.3)$  around the candidate muon whose transverse momenta is  $p_T^\mu$  after excluding the muon track itself. This  $p_T$ -dependent cone size can help to improve the performance for muons produced in the decay of particles with a large transverse momentum.

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

#### 4.2.5 Jets

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



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

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

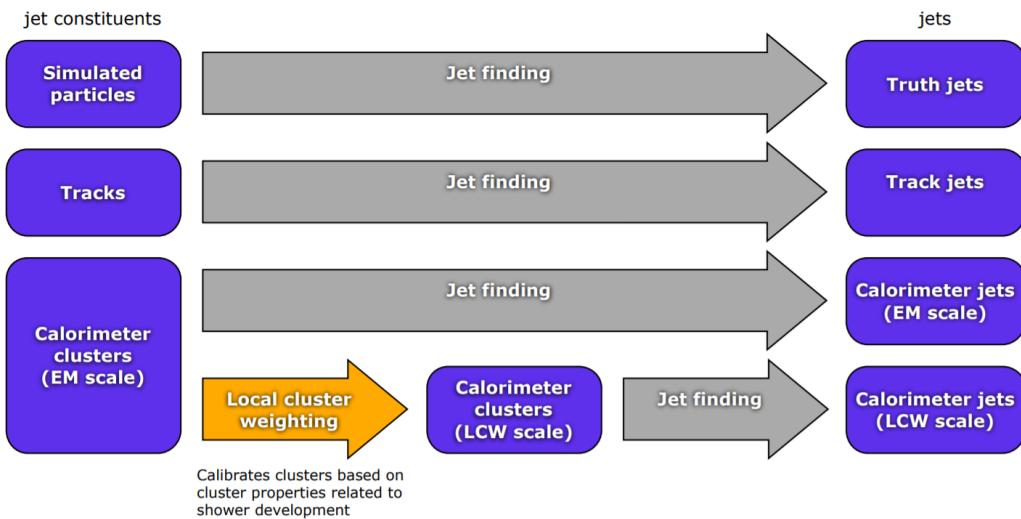
#### Jet reconstruction

Jets are reconstructed using anti- $k_t$  algorithm<sup>[63]</sup> and with radius parameter of  $R = 0.4$  in most cases. The FastJet software package<sup>[64]</sup> is utilized for jet finding and reconstruction. A collection of four-vectors are used as inputs at each combination step in jet clustering, the total four-momentum is therefore computed as the sum of four-vector of all its constituents. There are three types of jets in ATLAS:

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

Figure 4.11 shows the schematic of ATLAS jet reconstruction.

#### Jet reconstruction



**Fig. 4.11 A overview schematic of ATLAS jet reconstruction<sup>[65]</sup>.**

1223

1224 The *truth jets* are reconstructed using anti- $k_t$  algorithm with  $R = 0.4$  by using final-

state, stable particles from MC simulation as inputs. It requires the candidate particles with lifetime  $c_\tau > 10$  mm and excludes the particles from pile-up. Truth jets with  $p_T > 7\text{GeV}$  and  $|\eta| < 4.5$  are then used for jets calibration described later.

The *track jets* are reconstructed from charged particles within the full acceptance of inner detector ( $|\eta| < 2.5$ ). The track reconstruction has been introduced in section 4.2.1. Reconstructed jets with  $p_T > 500\text{MeV}$  and associated with primary vertex are then selected. Tracks are assigned to jets using ghost association<sup>[62]</sup>, 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<sup>[66]</sup>. Topo-clusters are built from near-by calorimeter cells that contains a significant energy above a noise threshold, which is estimated from measurements of calorimeter electronic noise and simulated pile-up noise. Those calorimeter cell energies are measured at electromagnetic energy scale (EM scale) corresponding to the energy deposited by electromagnetically interacting particles. And jets passing a  $p_T$  threshold of 7 GeV are reconstructed with the anti- $k_t$  algorithm.

### Jet energy scale calibration

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

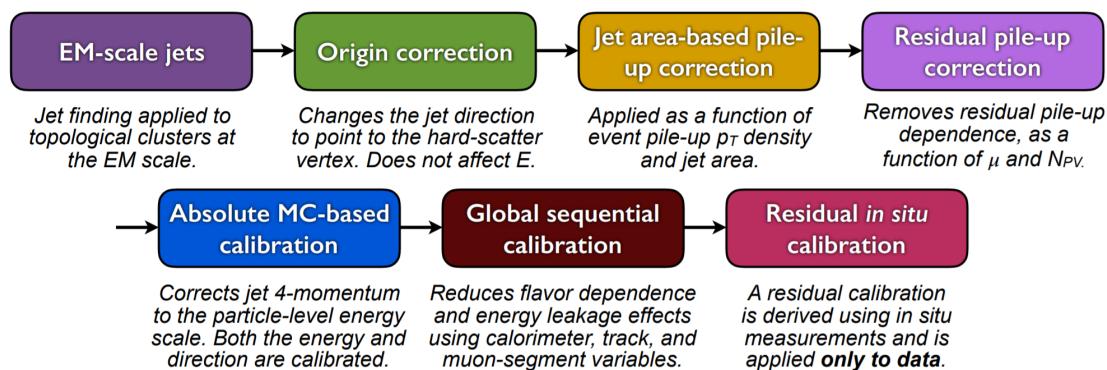


Fig. 4.12 A overview schematic of ATLAS jet calibration<sup>[66]</sup>.

1250 First of all, the origin correction recompute the four-momentum of jets to point them to  
1251 the hard-scatter primary vertex instead of the centre of detector, and in the meantime keep  
1252 the jet energy unchanged. This correction improves the  $\eta$  resolution of jets by roughly  
1253 25% at a jet  $p_T$  of 20 GeV and > 5 times improvement for jet with  $p_T$  above 200 GeV,  
1254 as measured from the difference between reconstructed jets and truth jets in MC simu-  
1255 lation. Secondly, the pile-up correction is adopted to remove the excess energy due to  
1256 in-time and out-of-time pile-up, which consists of two processes: an area-based  $p_T$  den-  
1257 sity subtraction applied on the top of each event; and a residual correction derived from  
1258 the simulation. Thirdly, the absolute JES calibration corrects the jet four-momentum to  
1259 the particle-level energy scale, using truth jets in di-jet MC events. Furthermore, the step  
1260 of global sequential calibration uses calorimeter, track and MS-based variables to reduce  
1261 the flavor dependence and energy leakage effects. Finally, the residual in situ calibration  
1262 is adopted to correct jets in data by using well-measured objects eg. photons, Z bosons  
1263 and calibrated jets.

1264 **B-jet tagging**

1265 Tagging of b-jets plays a important role in many physics analyses involving b- or t-  
1266 quark. In the meantime, lots of analyses need to apply b-jet veto to suppress  $t\bar{t}$  process.  
1267 There are three major types of algorithms that have been developed to distinguish b-quark  
1268 jets from light-quark (u,d,s) jets<sup>[67]</sup>:

- 1269 • **Impact parameter based algorithms (IP2D and IP3D):** b-hadrons usually have  
1270 long lifetime ( $\sim 1.5$  ps,  $c_\tau \sim 450$   $\mu\text{m}$ ), which leads to large impact parameter for  
1271 tracks produced from b-hadron decay. The impact parameter taggers are devel-  
1272 oped based on these variables. The IP2D tagger makes use of the transverse im-  
1273 pact parameter significance  $d_0/\sigma(d_0)$  as discriminant, while IP3D tagger uses two-  
1274 dimensional discriminant of both transverse and longitudinal impact parameter sig-  
1275 nificances:  $d_0/\sigma(d_0)$  and  $z_0 \sin\theta/\sigma(z_0)$ .
- 1276 • **Secondary vertex finding algorithm (SV1)** makes use of the secondary vertex  
1277 formed by decay products of b-hadron within the jet. All track pairs within a jet are  
1278 tested for a two-track vertex hypothesis, and removed if they are likely to originate  
1279 from a long-live particle decay (eg.  $K_s$  or  $\Lambda$ ), hadronic interactions or photon con-  
1280 versions. After that, a new vertex is fitted with all tracks from remaining two-track  
1281 vertices, and the outliers are removed from this set of tracks.
- 1282 • **Decay chain multi-vertex algorithm (JetFitter)**<sup>[68]</sup> exploits the topological struc-  
1283 ture of weak b- and c- hadron decays inside the jet and tries to reconstruct the full

1284 b-hadron decay chain. A Kalman filter is adopted to find a common line between  
 1285 primary vertex and b/c- vertices, as well as their position in this line, which gives  
 1286 a approximated flight path for the b-hadron. In this approach, the b- and c-hadron  
 1287 vertices, whenever resolution allows, can be resolved, even when there is only a  
 1288 single track associated to them.

1289 The final discrimination commonly used in many physics analyses is called **Multivariate**  
 1290 **Algorithm (MV2)**, which is based on Boosted Decision Tree (BDT) implemented in the  
 1291 TMVA package<sup>[69]</sup> by combining the outputs from underlying taggers mentioned above.  
 1292 The MV2 was trained using jets in  $t\bar{t}$  sample, where the b-jets are treated as signal while  
 1293 the c- and light-flavor jets are treated as backgrounds. There are three kinds of MV2  
 1294 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.

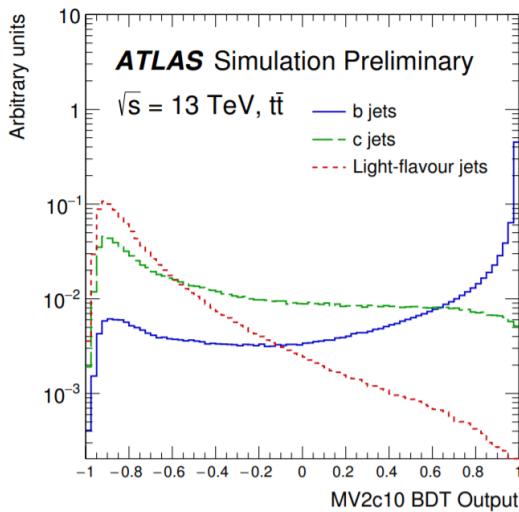


Fig. 4.13 MV2c10 BDT output for b- (solid blue), c- (dashed green) and light-flavour (dotted red) jets in  $t\bar{t}$  events<sup>[67]</sup>.

1295

#### 1296 4.2.6 Missing transverse energy

1297 Many interesting physics processes are with the involvement of neutrinos. Since they  
 1298 do not interact with any materials in the detector, neutrinos cannot be detected directly;  
 1299 but instead, they can result in imbalance in the plane transverse to the beam axis, where  
 1300 momentum conservation is assumed. It is known as the missing transverse momentum  
 1301 denoted as  $E_T^{miss}$ , which is obtained from the negative vector sum of the momenta of all  
 1302 particles detected in a proton-proton collision event.

1303 The  $E_T^{miss}$  is measured using selected, reconstructed and calibrated hard objects in an

1304 event. Its x- and y- components can be calculated as follow<sup>[70]</sup>:

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

1305 where each object term is given by the negative vectorial sum of the momenta of the  
 1306 respective calibrated objects. The calorimeter signals are associated with the recon-  
 1307 structed objects in the following order: electrons, photons, hadronically decaying taus,  
 1308 jets, muons. The soft term is reconstructed from detected objects not match any hard  
 1309 object passing the selections, but associated with the primary vertex. Details of applied  
 1310 selections for each term are summarized in table 4.1.

1311 Based on  $E_{x(y)}^{miss}$ , the magnitude of  $E_T^{miss}$  and the azimuthal angle  $\phi^{miss}$  are computed:

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

1312 In equation 4.2, each objects are required to pass certain reconstruction and calibrated  
 1313 criteria and selections mentioned above before taken as inputs.

1314 In figure 4.14, left plot shows the observed  $E_T^{miss}$  distribution for data and MC of  $Z \rightarrow$   
 1315  $\mu\mu$  events without genuine missing transverse momentum; and right plot shows the  $E_T^{miss}$   
 1316 distribution for  $W \rightarrow e\nu$  events that has genuine (true) missing transverse momentum  
 due to real neutrino.

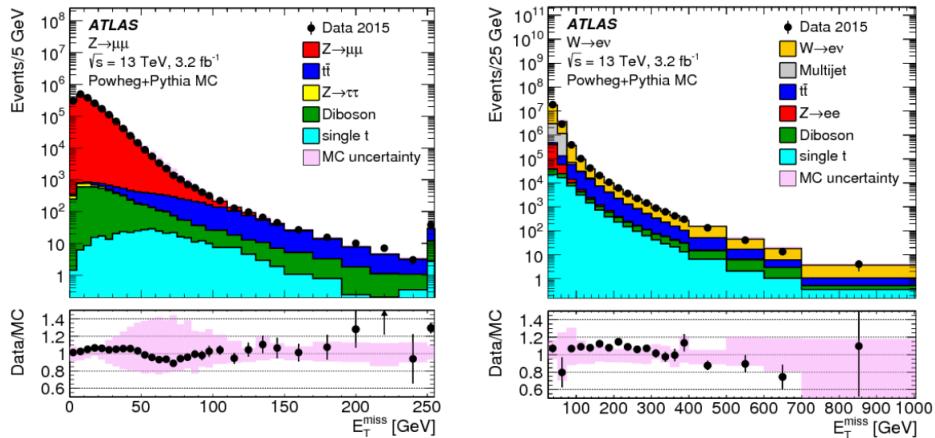


Fig. 4.14 Measured  $E_T^{miss}$  distribution for  $Z \rightarrow \mu\mu$  events (left) and  $W \rightarrow e\nu$  events (right).

1317

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

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

## 1318    **Chapter 5 Statistical treatment of searching for new 1319    particles or processes**

1320    In the experiments of particle physics, one often searches for particles or processes  
1321    that have been predicted but not yet observed, such as the two analysis presented in this  
1322    dissertation: searching for the vector boson scattering process and searching for the heavy  
1323    resonance(s). Usually two hypotheses are defined:

- 1324    •  $H_0$ : null hypothesis, in most cases are designated as background-only hypothesis.
- 1325    •  $H_1$ : signal plus background hypothesis, where signal is a new model one would  
1326    like to search for.

1327    For the purpose of discovering a new signal process, the  $H_0$  hypothesis is tested against  
1328    the alternative  $H_1$ . When setting limits, the  $H_1$  hypotheses with different signal strengths  
1329    are tested against the  $H_0$ .

1330    The level of agreement between observed data and a given hypothesis can be quantified  
1331    by computing the  $p$ -value, the probability under this hypothesis assumption, or  
1332    its equivalent Gaussian significance. This section describes the statistical treatment for  
1333    searches related to this dissertation.

### 1334    5.1 The likelihood function

1335    The likelihood function is defined as the product of a set of the probability density  
1336    functions (pdfs) of variables  $x$ , that used to evaluate the probability of the observed  
1337    dataset:

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

1338    where  $\theta_1, \dots, \theta_M$  are the nuisance parameters that can be written as  $\boldsymbol{\theta}$ , and  $x_1, \dots, x_N$   
1339    denote the observables of dataset. Usually one measures the variable  $x$  by constructing a  
1340    histogram  $\mathbf{n} = (n_1, \dots, n_N)^T$ . The expectation value of the  $i$ th bin  $n_i$  can be written as:

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

1341    where  $\mu$  is the signal strength,  $s_i$  and  $b_i$  are the number of signal and background events  
1342    in that bin. In addition to the histogram  $\mathbf{n}$ , in some cases, one would like to use subsidiary  
1343    measurements to help further constrain the nuisance parameters. For instance, due to the  
1344    lack of background simulation or the mismodelling issue of one MC sample, one can

choose a control region and construct another histogram  $\mathbf{m} = (m_1, \dots, m_M)$  to constrain the contribution of one certain background in data. For this measurement, the expectation value of the  $i$ th bin  $m_i$  can be written as:

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

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

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_i^N \frac{(\mu s_i + b_i)^{n_j}}{n_j!} e^{-(\mu s_i + b_i)} \prod_i^M \frac{u_k^{m_k}}{m_k!} e^{-u_k} \quad (5.4)$$

Then to test the hypothesized value of  $\mu$ , the profile likelihood ratio is defined as:

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

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

## 5.2 Test statistic

To test the level of agreement between the data and the hypothesized value  $\mu$ , a test statistic  $t_\mu$  can be defined:

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

From the definition of  $\lambda(\mu)$  in equation 5.5, one can see that  $0 \leq \lambda \leq 1$ , and a  $\lambda$  with value close to 1 implies good agreement between data and  $\mu$ . Thus, larger value of  $t_\mu$  means the increase of incompatibility between data and  $\mu$ . To quantify the level of disagreement, one can calculate the  $p$ -value as:

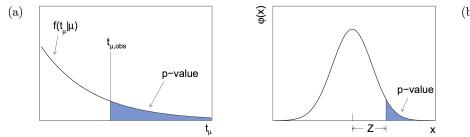
$$p_\mu = \int_{t_{\mu,obs}}^{\infty} f(t_\mu | \mu) dt_\mu \quad (5.7)$$

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

1364 corresponding observed significance,  $Z$ , can be defined as:

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

1365 The relationship between the  $t_\mu$ ,  $p$ -value and significance  $Z$  are depicted in figure 5.1.  
 1366 When searching for a signal process, such as Higgs boson, the particle physics commu-  
 1367 nity tends to claim a discovery when the rejection of background-only hypothesis has a  
 1368 significance of at least  $Z = 5$ .



**Fig. 5.1 (a) Illustration of the relationship between the observed  $t_\mu$  and its  $p$ -value. (b) The relationship between  $p$ -value and the observed significance  $Z$ , where  $\phi(x)$  is a standard normal distribution.**

1369 In most cases, one assumes that the presence of a new signal can only increase the  
 1370 event rate comparing to the background only model, then the signal strength  $\mu \geq 0$ . And  
 1371 for the case of discovery, the hypothesis of a positive signal strength should be tested  
 1372 against to the background-only (null) hypothesis by using the test statistic called  $p_0$ :

$$p_0 = \begin{cases} -2\ln(\lambda(0)) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \quad (5.9)$$

1373 which corresponds to the  $p$ -value called  $p_0$ :

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

1374 to quantify the level of disagreement between the data and the null hypothesis ( $\mu = 0$ ).

### 1375 5.3 The CLs upper limit

1376 For a signal hypothesized value  $\mu$ , one can compute the probability that this hypothesis  
 1377 (called S+B hypothesis) gives a **greater** test statistic value than the observed one  $q_{obs}$ :

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

1378 In the meantime, the probability that the background-only hypothesis gives a **smaller** test  
 1379 statistic than observed data can also be calculated as:

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

1380 Then we define the CLs<sup>[71]</sup> of a hypothesized value  $\mu$  as:

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

1381 For purpose of excluding a signal hypothesis, a threshold CLs of 0.05 (namely 95% con-  
 1382 fidence level) is often used. For this reason, usually under the circumstance that no signif-  
 1383 icance derivation between data and background-only hypothesis is found, one would like  
 1384 to find the value of hypothesized signal strength  $\mu$  by requiring its  $CLs = 95\%$  (called  
 1385 95% CLs upper limit) for exclusion.

1386 The sensitivity of an experiment to exclude a new signal process is quantified by *me-  
 1387 dian upper limit*, which is obtained using "Asimov dataset". The Asimov dataset is de-  
 1388 fined such that when one uses it to evaluate the estimators for all parameters, one obtains  
 1389 the true parameter values. Moreover, it is useful to use Asimov dataset to compute how  
 1390 much the sensitivity is expected to vary, given the expected fluctuations in the data. The  $\hat{\mu}$   
 1391 is assumed to follow a Gaussian distribution with a mean value of  $\mu'$  and the standard de-  
 1392 viation of  $\sigma$ . First of all, the test statistic from profile likelihood ratio can be approximated  
 1393 as<sup>[1]</sup>:

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

1394 Given that the Asimov dataset corresponding to a signal strength  $\mu'$ , one finds:

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

1395 where  $q_{\mu,A} = -2\ln\lambda_A(\mu)$  is the observed test statistic of Asimov dataset. Then the stan-  
 1396 dard derivation can be computed as:

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

1397 In a special situation where one wants to find the median exclusion significance for the  
 1398 hypothesis  $\mu$  assuming that there is no signal ( $\mu' = 0$ ), one gets:

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



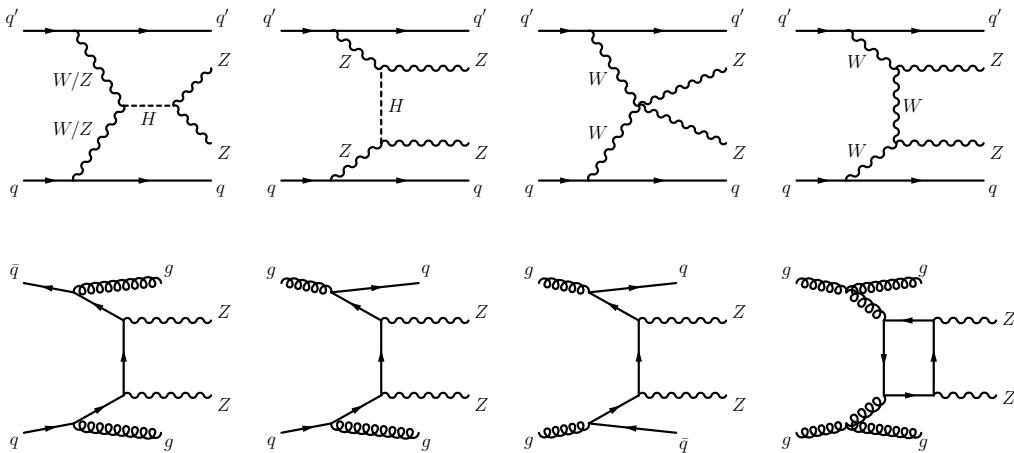
## 1400   **Chapter 6 Studies of SM $ZZ$ production in $\ell\ell\ell'\ell'$ final** 1401   **state using pp collision data collected by ATLAS** 1402   **detector from 2015 to 2018**

### 1403   6.1 Introduction

1404   After the discovery of Higgs boson<sup>[3,4]</sup>, the examination of electroweak symmetry  
1405   breaking (EWSB) becomes a main focus at the LHC. In addition to measuring the prop-  
1406   erties of Higgs boson directly, the vector boson scattering (VBS) process is another key  
1407   avenue to probe EWSB<sup>[72-74]</sup>. As introduced in section 2.1.3, in Standard Model (SM),  
1408   the Higgs boson acts as “moderator” to unitarize high-energy longitudinal VBS ampli-  
1409   tudes at the TeV scale. Therefore, studying high-energy behaviours of VBS is crucial to  
1410   understand the mechanism of EWSB.

1411   Since no VBS process was observed prior to the LHC era, LHC provides an exception-  
1412   able opportunity to study them due to its unprecedented high energy and luminosity. At  
1413   the LHC, the VBS process is typically studied through the measurements of electroweak  
1414   (EW) production of two vector bosons radiated from quark-quark initial state, plus a pair  
1415   of hadronic jets with high energy in the back and forward regions (denoted as EW- $VVjj$ ).  
1416   The quantum chromodynamics (QCD) production of  $VVjj$  containing two QCD vertices  
1417   at the lowest order (denoted as QCD- $VVjj$ ) is an irreducible background to the search  
1418   of EW- $VVjj$  production. The features of EW- $VVjj$  production including a large invari-  
1419   ant mass of jet pair ( $m_{jj}$ ) and a significant separation of rapidity between two jets ( $\Delta y_{jj}$ ).  
1420   Figure 6.1 presents some typical Feynman diagrams of EW- and QCD-  $ZZjj$  processes.  
1421

1422   The first evidence of the EW- $VVjj$  process was seen in same-sign  $WW$  channel  
1423   (EW- $W^\pm W^\pm jj$ ) by ATLAS collaboration with  $20.3 \text{ fb}^{-1}$  8 TeV data<sup>[75]</sup>, in which a  $3.6\sigma$   
1424   excess was observed in data over the background-only prediction. In the LHC run-2, the  
1425   observation (with  $> 5 \sigma$  statistical significance) of EW- $W^\pm W^\pm jj$  process has been re-  
1426   ported in both ATLAS and CMS collaboration with  $36 \text{ fb}^{-1}$  13 TeV data<sup>[76-77]</sup>. In  $WZ$   
1427   channel (EW- $WZjj$ ), an observation with  $5.3 \sigma$  excess was also reported by the AT-  
1428   LAS collaboration recently<sup>[78]</sup>. As for the EW- $ZZjj$  production, it was searched by  
1429   CMS using  $35.9 \text{ fb}^{-1}$  13 TeV data but no evidence was found<sup>[79]</sup>. The EW production in  
1430    $ZZ$  final state (EW- $ZZjj$ ) is typically rare, whose fiducial cross section has an order of  
1431    $O(0.1) \text{ fb}^{-1}$  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  $139 \text{ fb}^{-1}$  luminosity. It is a new milestone in the study of EWSB at the LHC, and completes the last missing part of observation of weak boson scattering for massive bosons. In the meantime, the measurement of fiducial cross-sections for SM  $ZZ$  production including both EW and QCD processes is also reported. The  $ZZjj$  production involving intermediate  $\tau$ -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  $139.0 \pm 2.4 \text{ fb}^{-1}$ .

### 1452    6.2.2 MC simulations

1453    The EW- $Z Z jj$  production is modelled using MadGraph5\_aMC@NLO 2.6.1<sup>[80]</sup> with  
1454    the matrix elements (ME) calculated in the leading-order (LO) approximation in pertur-  
1455    bative QCD (pQCD) and with the NNPDF2.3LO<sup>[81]</sup> parton distribution functions (PDF).  
1456    The VBF Higgs process is also included.

1457    The QCD- $Z Z jj$  production is modelled using Sherpa 2.2.2<sup>[82]</sup> with the  
1458    NNPDF3.0NNLO<sup>[83]</sup> PDF, where events with up to one (three) outgoing partons  
1459    are generated at NLO (LO) in pQCD. The production of  $Z Z jj$  from the gluon-gluon  
1460    initial state with a four-fermion loop or with an exchange of the Higgs boson has an  
1461    order of  $\alpha_S^4$  in QCD, and is not included in the Sherpa simulation. A separate  $gg$  induced  
1462     $Z Z + 2\text{jets}$  sample is modelled using Sherpa 2.2.2 with the NNPDF3.0NNLO PDF and  
1463    with an additional 1.7 k-factor<sup>[84]</sup> being applied. Then the interference between EW- and  
1464    QCD- $Z Z jj$  is modelled with MadGraph5\_aMC@NLO 2.6.1 calculated at LO.

1465    The diboson productions from QCD  $WW \rightarrow \ell\nu qq$  as well as QCD and EW  $WZ \rightarrow$   
1466     $\ell\ell qq$  are modelled using Sherpa 2.2.2 with the NNPDF3.0NNLO PDF. The productions  
1467    of semileptonic decays ( $WW \rightarrow \ell\nu qq$  and  $WZ \rightarrow qq\ell\ell$ ) are modelled using Powheg-  
1468    Box v2<sup>[85]</sup> with the CT10 PDF<sup>[86]</sup>. The triboson production is modelled using Sherpa  
1469    2.2.2 with the NNPDF3.0NNLO PDF.

1470    For top-quark pair ( $t\bar{t}$ ) production, the Powheg-Box v2 is used with the CT10 PDF.  
1471    The single top-quark production in  $t$ -channel,  $s$ -channel and  $Wt$ -channel are simulated  
1472    using the Powheg-Box v1 event generator<sup>[87-89]</sup>. The productions of  $t\bar{t}$  in association with  
1473    vector boson(s) ( $t\bar{t}V$ ) are modelled with MadGraph5\_aMC@NLO 2.3.3 for  $t\bar{t}W$  and  $t\bar{t}Z$   
1474    with  $Z \rightarrow \nu\nu/qq$  decays, with Sherpa 2.2.1 for  $t\bar{t}Z$  where the  $Z$  decays to dilepton, and  
1475    with MadGraph5\_aMC@NLO 2.2.2 for  $t\bar{t}WW$  respectively.

1476    The  $Z+jets$  processes are modelled using Sherpa 2.2.1 with the NNPDF3.0NNLO  
1477    PDF, in which the ME is calculated for up to two partons with next-to-leading-order  
1478    (NLO) accuracy in pQCD and up to four partons with LO accuracy.

1479    For all the samples except those from Sherpa, the parton showering is modelled with  
1480    Pythia8<sup>[45]</sup> using the NNPDF2.3<sup>[81]</sup> PDF set, and the A14 set of tuned parameters<sup>[90]</sup>.  
1481    While for Sherpa samples, the parton showering is simulated within the programme.

1482    All simulated events are processed with detector response simulation based on Geant4  
1483    described in section 4.1. In addition, simulated inelastic pp collisions are overlaid to  
1484    model additional pp collision in the same and neighbouring bunch crossings (pile-up),  
1485    and reweighted to match the pile-up conditions in data. Moreover, all simulated events  
1486    are processed using the same reconstruction algorithms as data. And the leptons and

1487 jets reconstruction, energy scale and resolution, and the leptons identification, isolation,  
 1488 trigger efficiencies for simulated events, as described in section 4.2, are all corrected to  
 1489 match the data measurements.

## 1490 6.3 Objects and Event selection

### 1491 6.3.1 Objects selection

1492 The selection of analysis relies on the definition of multiple objects: *electrons*, *Muons*,  
 1493 and *jets*. Details of definition for each object are described as below:

1494 **Muon:** To increase the acceptance range in reconstruction (reco) -level for  $\ell\ell\ell'\ell'$   
 1495 channel, all four types of muons (CB, ST, CT, ME muons, described in section 4.2.4)  
 1496 are used. The identified muons are then required to pass  $p_T > 7$  GeV and  $|\eta| < 2.7$ ,  
 1497 and satisfy the *Loose* identification criterion (see definition in sec 4.2.4). The impact  
 1498 parameter cuts are further applied to suppress the contribution from cosmic muons and  
 1499 non-prompt muons, with the value of:  $|d_0/\sigma(d_0)| < 3.0$  and  $|z_0 \sin\theta| < 0.5$  mm, where  $d_0$   
 1500 is the transverse impact parameter relative to the beam line,  $\sigma(d_0)$  is its uncertainty, and  
 1501  $z_0$  is the longitudinal impact parameter relative to the primary vertex. In order to avoid  
 1502 muons associated with jets, all muons are required to be isolated and pass *FixedCutLoose*  
 1503 isolation criteria of  $E_T^{\text{topocone}20}/p_T < 0.3$  and  $p_T^{\text{varcone}30}/p_T < 0.15$ .

1504 **Electron:** As described in section 4.2.3, electrons are reconstructed from energy de-  
 1505 posits in the EM calorimeter matched to a track in the inner detector. The electron candi-  
 1506 dates must satisfy the *Loose* criterion defined by the likelihood-based (LH) method. And  
 1507 electrons are required to have  $p_T > 7$  GeV and  $|\eta| < 2.47$ . Moreover, the impact param-  
 1508 eter requirements of  $|d_0/\sigma(d_0)| < 5.0$  and  $|z_0 \sin\theta| < 0.5$  mm are applied. Same as muon,  
 1509 all electrons are required to satisfy *FixedCutLoose* isolation criteria, which, for electrons,  
 1510 is  $E_T^{\text{topocone}20}/p_T < 0.2$  and  $p_T^{\text{varcone}20}/p_T < 0.15$ .

1511 **Jets:** Jet are key signatures for VBS processes. This analysis use the jets clustered  
 1512 using the anti- $k_t$  algorithm with radius parameter  $R = 0.4$ , more details of jets' reconstruc-  
 1513 tion can be found in section 4.2.5. The jets are required to satisfy  $p_T > 30$  (40) GeV in  
 1514 the  $|\eta| < 2.4$  ( $2.4 < |\eta| < 4.5$ ) region. To further reduce the effects of pile-up jets, a jet  
 1515 vertex tagger (JVT) is applied to jets with  $p_T < 60$  GeV and  $|\eta| < 2.4$  to select jets from  
 1516 hard-scattering vertex<sup>[91]</sup>.

1517 **Overlap removal:** An overlap-removal procedure is applied to selected leptons and  
 1518 jets. To enhance the selection efficiency, leptons are given higher priority to be kept when  
 1519 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 <b>AND</b> ( $N_{Trk}(jet) < 3$ 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.

### 6.3.2 Event selection

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

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. At most one muon is allowed to be ME or CT muon. The  $p_T$  threshold of first three leading leptons are 20, 20 and 10 GeV. If more than one quadruplets are found, the one with minimum sum of difference between two dilepton pair masses and Z boson mass ( $|m_{l_1^+ l_1^-} - m_Z| + |m_{l_2^+ l_2^-} - m_Z|$ ) is selected. Both two dilepton pair masses are required to be between 66 to 116 GeV. In addition, the invariant masses of all possible OSSF pairs are required to be greater than 10 GeV to reject events from  $J/\phi$  or  $Y$  decay.

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

Electrons	$p_T > 7 \text{ GeV}$ , $ \eta  < 2.47$ $ d_0/\sigma_{d_0}  < 5$ 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) $\text{GeV}$ 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  $139 \text{ fb}^{-1}$ . 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- $Z Z jj$	$20.6 \pm 2.5$
QCD- $q\bar{q} \rightarrow ZZ$	$77 \pm 25$
QCD- $gg \rightarrow ZZ$	$13.1 \pm 4.4$
Others	$3.2 \pm 2.1$
Total	$114 \pm 26$
Data	127

**Table 6.3 Observed data and expected signal and background yields in  $139 \text{ fb}^{-1}$  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

1547 by reverting either the  $m_{jj}$  or  $\Delta y_{jj}$  requirements as:

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

1548 Then this normalization factor is included into statistical fit as a free parameter to properly  
 1549 treat the uncertainty correlations between SR and CR, while the shapes are taken from MC  
 1550 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	$3.9 \pm 0$
QCD-ZZjj	$136.9 \pm 0.6$
QCD-ggZZjj	$16.8 \pm 0.1$
Diboson	$0.3 \pm 0.1$
Triboson	$1.6 \pm 0.1$
Z+jets	0
t̄t	0
Total	$159.5 \pm 0.62$
Data	152

Table 6.4 Observed data and expected signal and background yields in  $139 \text{ fb}^{-1}$  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.

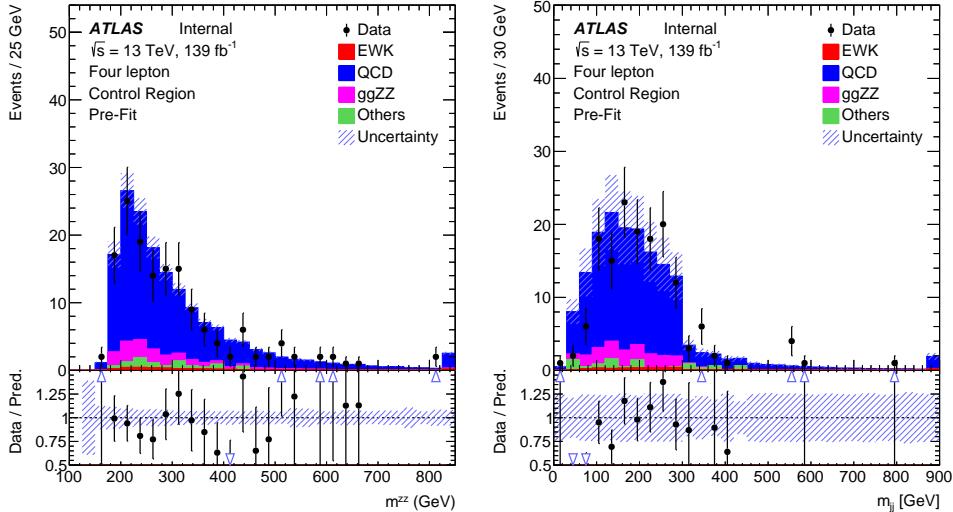
1551

1552 dijet in QCD CR are shown in figure 6.2.

#### 1553 6.4.2 Reducible backgrounds

1554 Backgrounds from Z+jets, top-quark and WZ processes called reducible back-  
 1555 grounds can be estimated by data-driven method. These events usually contain two or  
 1556 three leptons from Z/W decays, together with heavy-flavor jets or misidentified compo-  
 1557 nents of jets reconstructed as leptons called “fake leptons”. A *fake factor* method is used  
 1558 to estimate these backgrounds, where the lepton misidentification is measured in data with  
 1559 the region enhances contributions from Z+jets and top-quark processes. The method is  
 1560 described as below:

1561 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:

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

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:

$$\mathcal{N}_{fake} = (N_{gggp} - N_{ggp}) \times \mathcal{F} - (N_{gppg} - N_{gpp}) \times \mathcal{F}^2 \quad (6.3)$$

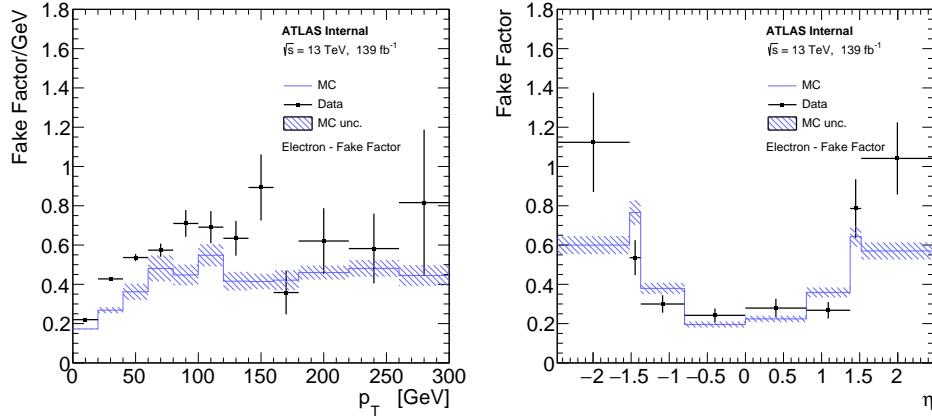
with the subtraction of  $ZZ$  contribution, and the double counting between  $N_{gggp}$  and  $N_{gppg}$ .

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.

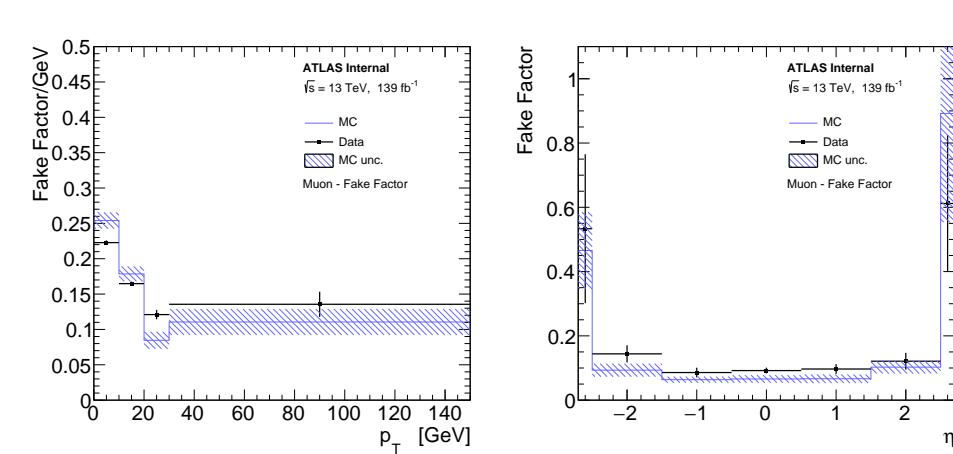
### 1. Fake factor for $Z+jets$

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  $\eta$  as shown in

figure 6.3 for electrons and figure 6.4 for muons. During calculation, the contributions 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  $\eta$  (right).



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

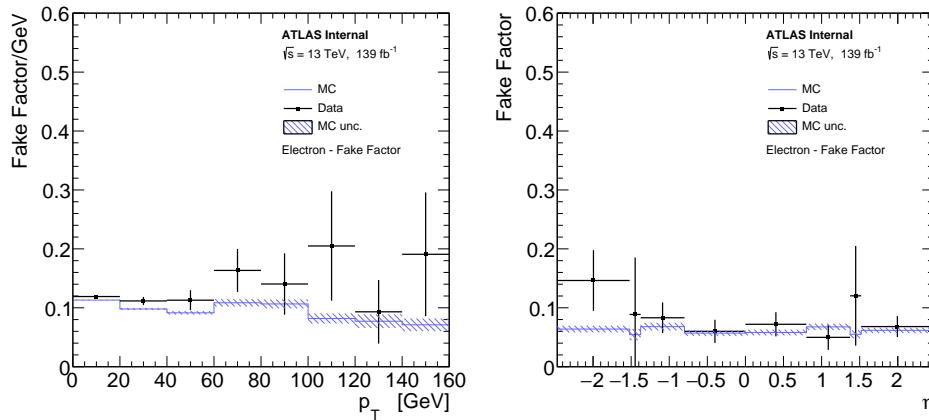
## 2. Fake factor for $t\bar{t}$

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

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

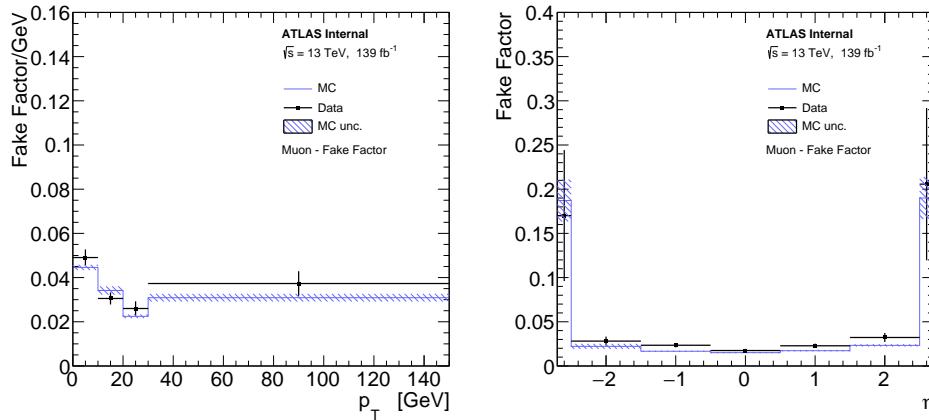
In addition, at least one b-jet is required to enhance the top component. The fake factors of  $t\bar{t}$  calculated from data as the function of  $p_T$  and  $\eta$  are shown in figure 6.5 for electrons and 6.6 for muons. The non- $t\bar{t}$  contributions including  $Z+jets$ ,  $ZZ$  and  $WZ$ ,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

1591



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

1592

### 3. Systematics of fake estimation and results

1593

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.

1596

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

1598

- 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.

1599

1600

1601

1602

- The difference of fake factors when changing to one bin measurement (instead of  $p_T$  or  $\eta$  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 $\mu$	4 $\mu$	inclusive
Nominal estimate	$0.678 \pm 0.652$	$1.023 \pm 0.740$	$0.566 \pm 0.240$	$2.268 \pm 1.015$
$F$ stat. uncertainty varied down	$0.698 \pm 0.622$	$0.872 \pm 0.652$	$0.509 \pm 0.214$	$2.079 \pm 0.926$
$F$ stat. uncertainty varied up	$0.657 \pm 0.685$	$1.173 \pm 0.840$	$0.622 \pm 0.267$	$2.452 \pm 1.116$
One bin $F$	$0.653 \pm 0.590$	$0.594 \pm 0.558$	$0.646 \pm 0.313$	$1.892 \pm 0.870$
MC $F$	$0.534 \pm 0.471$	$1.415 \pm 0.993$	$0.439 \pm 0.184$	$2.389 \pm 1.114$
Isolation varied down	$0.938 \pm 0.686$	$0.552 \pm 0.466$	$0.215 \pm 0.107$	$1.704 \pm 0.837$
Isolation varied up	$0.723 \pm 0.646$	$1.104 \pm 0.739$	$0.559 \pm 0.237$	$2.386 \pm 1.010$
MC corr. varied down	$0.697 \pm 0.695$	$1.048 \pm 0.811$	$0.832 \pm 0.385$	$2.577 \pm 1.136$
MC corr. varied up	$0.660 \pm 0.614$	$0.984 \pm 0.687$	$0.316 \pm 0.159$	$1.961 \pm 0.935$

**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.

1608

## 1609 6.5 Systematics

1610 The analysis includes both the statistical fit to MD distribution to search the EW-  
 1611  $ZZjj$  process, as well as the cross section measurement of inclusive EW and QCD  $ZZjj$   
 1612 process in fiducial volume. Therefore, theoretical and experimental uncertainties may  
 1613 affect the predicted background yields and shapes, the correction factors from detector-  
 1614 level to particle-level measurement, as well as the  $ZZjj$  MD shapes and so on. Moreover,  
 1615 the statistical uncertainties of simulated samples are also taken into account. Due to the  
 1616 extremely low cross section of  $\ell\ell\ell'\ell'$  channel, the analysis is still data statistic dominant.  
 1617 This section describes the measurement of both theoretical and experimental systematics  
 1618 for  $ZZjj$  productions. The systematics for fake backgrounds have been elaborated in  
 1619 section 3.

### 1620 6.5.1 Theoretical systematics

1621 The theoretical systematics on EW- and QCD- $ZZjj$  processes including the uncer-  
 1622 tainties from PDF, QCD scale,  $\alpha_S$  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 ( $\mu_R$ ) and factorisation scale ( $\mu_F$ ) 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<sup>[92-93]</sup> algorithm. The  $\alpha_S$  uncertainty is estimated by varying the value of  $\alpha_S$  within  $\pm 0.001$ . 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
$\alpha_S$	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 uncertainties of each theoretical components in fiducial volume of SR, while table 6.8 shows the numbers in QCD-enriched CR region. For QCD process, the uncertainty is QCD scale dominant. Both of them are taken as inputs for statistical fit.

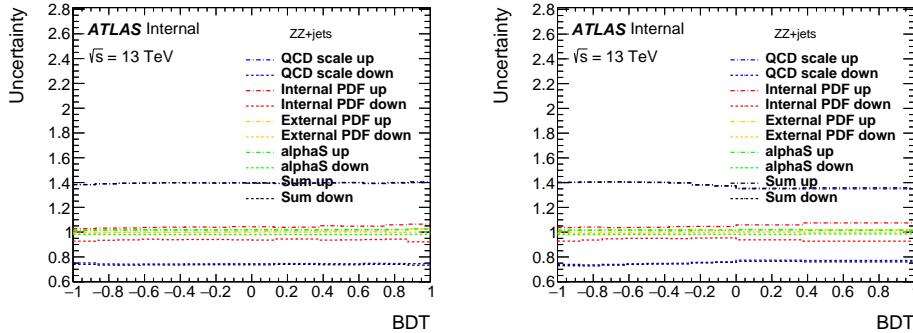
Process	PDF (%)	$\alpha_S$ (%)	QCD scale (%)	Parton shower (%)
EW	+5.9 -5.9		+6.1 -5.6	+3.3 -3.3
qqQCD	+2.0 -1.0	+2.6 -2.6	+34.2 -22.8	

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

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

**Table 6.8 Summary of theoretical uncertainties for the control region for EW and qqQCD processes.**

1638 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).

1639

### 1640 6.5.2 Experimental systematics

1641 The dominant experimental uncertainties are from the luminosity uncertainty, the mo-  
1642 mentum scale and resolution of leptons and jets, as well as the lepton reconstruction and  
1643 selection efficiency. Some smaller uncertainties, such as trigger efficiency and pile-up  
1644 correction, are also considered. Table 6.9 lists the major systematic components from  
1645 leptons and jets for signal and major background processes in  $\ell\ell\ell'\ell'$  channel. The to-  
1646 tal uncertainties for sources from electron, muon and jet respectively, as well as the sum  
1647 (quadratic sum) of them are also summarized in this table.

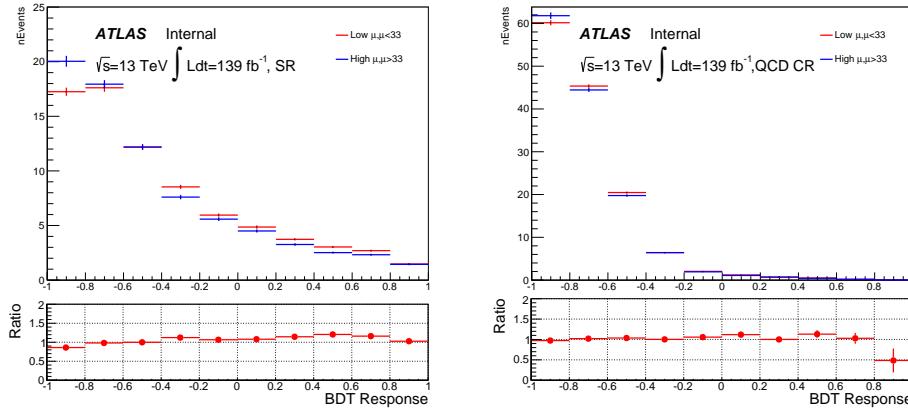
name	EW- $ZZjj$	QCD $qq$ -initial	QCD $gg$
nominal yield	20.61	76.69	13.10
EG_RESOLUTION_ALL	$\pm^{0.00\%}_{0.03\%}$	$\pm^{0.02\%}_{0.04\%}$	$\pm^{0.01\%}_{1.41\%}$
EG_SCALE_ALL	$\pm^{0.03\%}_{0.05\%}$	-0.04%	$\pm^{0.01\%}_{0.06\%}$
EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	$\pm^{2.66\%}_{2.58\%}$	$\pm^{2.60\%}_{2.53\%}$	$\pm^{2.65\%}_{2.57\%}$
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 6.9 Experimental systematic uncertainties in  $\ell\ell'\ell'\ell'$  channel with the luminosity of  $139 \text{ fb}^{-1}$ . The “Electron Exp.”, “Muon Exp.” and “Jet Exp.” represent the quadrature of the respective sources from electron, muon, and jets.**

1648 In addition, the uncertainty of the combined 2015 to 2018 integrated luminosity is  
 1649 1.7%<sup>[94]</sup> in ATLAS experiment, obtained using the LUCID-2 detector<sup>[95]</sup> for the primary  
 1650 luminosity measurements.

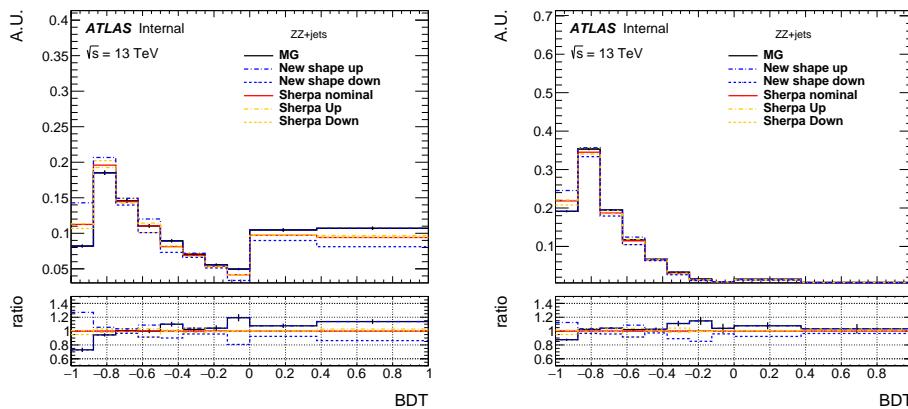
1651 On top of them, a systematic uncertainty for MD distribution with different pile-up  
 1652 ( $\langle\mu\rangle$ ) is also considered for QCD- $ZZjj$  background by comparing the distributions be-

1653 tween events with low and high pile-up conditions. A boundary of  $\langle\mu\rangle = 33$  is used to  
 1654 defined low/high pile-up according to the average  $\langle\mu\rangle$  for signal (about 34.5) and QCD  
 1655 background (about 33). Figure 6.8 shows the MD distribution in SR (left) and QCD CR  
 1656 (right) in two different PU conditions, the difference as function of MD is then taken into  
 1657 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).**

1658 Moreover, a conservative uncertainty is assigned to QCD- $Z Z jj$  process by comparing  
 1659 the sample modelled by Sherpa generator (nominal) with MadGraph5\_aMC@NLO.  
 1660 The MD shape difference for both SR (left) and QCD CR (right) are shown in figure 6.9.  
 1661 The modelling uncertainty is then calculated from the envelop between nominal and al-  
 ternative 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).**

1662

## 6.6 Measurement of fiducial cross section

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

For electrons and muons, QED final-state radiation is for the most part recovered by adding the four-momenta of surrounding photons that are not originating from hadrons and within an angular distance  $\Delta R < 0.1$  to the lepton four-momentum, called lepton “dressing” at truth level. Particle-level jets are built with anti- $k_T$  algorithm with radius parameter  $R = 0.4$  using all final-state particles except leptons and neutrinos as inputs. Comparing to the events selection at detector-level in section 6.3, at particle-level, the selected dilepton pair mass required is relaxed to be within 60 to 120 GeV for the reasons of reducing the migration effect, as well as being more compatibility with CMS publication<sup>[79]</sup>. All other kinematic selections are the same as the definition at detector-level.

### 6.6.1 Calculation of C-factor

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

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

The value of C-factor for each  $ZZjj$  process are calculated from each individual simulation samples as listed in table 6.10 together with their systematics.

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

**Table 6.10** C Factor of different  $ZZjj$  processes.

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

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

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

1686 and exp. denote the theoretical and experimental uncertainties described in section 6.5.

1687 **6.6.2 Result of fiducial cross section**

1688 The cross section in fiducial volume is computed as:

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

1689 where  $N_{data}$  and  $N_{bkg}$  denote the number of events selected from detector-level selection  
1690 from data and sum of backgrounds, and  $C$  is the C-factor calculated above, Lumi repre-  
1691 sents the integrated luminosity of data from 2015 to 2018 of  $139 \text{ fb}^{-1}$ . Table 6.11 shows  
1692 the fiducial cross section for  $\ell\ell\ell'\ell'$  final state measured from equation 6.7, as well as  
1693 the predicted cross section measured from MC simulation directly.

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

**Table 6.11 Measured and predicted fiducial cross-sections in  $\ell\ell\ell'\ell' jj$  final-state. Uncertainties due to different sources are presented.**

1694 The measured cross section has a total uncertainty of 11%, and is found to be compat-  
1695 ible with SM prediction. This measurement is still dominant by data statistic.

1696 **6.7 Search for EW- $ZZjj$**

1697 **6.7.1 MD discriminant**

1698 To further separate the EW- $ZZjj$  component from QCD- $ZZjj$ , a MD based on  
1699 Gradient Boosted Decision Tree (BDT) algorithm<sup>[96]</sup> is trained with simulated events  
1700 via TMVA framework<sup>[69]</sup>. Training is performed between EW (signal) and QCD (back-  
1701 ground) processes. Twelve event kinematic variables sensitive to the characteristics of  
1702 the EW signal are used as input features in training. Table 6.12 lists those input variables  
1703 with the order of their importance in BDT response provided by TMVA tool. One can see  
1704 the jet-related information provides larger sensitivity. Then the MD distributions in both  
1705 SR and QCD CR region are used for statistical fit.

1706 **6.7.2 Fitting procedure**

1707 A profile likelihood fit is performed on MD discriminant to extract the EW- $ZZjj$  sig-  
1708 nal from backgrounds. The binning of MD distributions in SR is optimized to maximize

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	$\Delta y_{jj}$	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 6.12 Input features for the training of MD.**

1709 the sensitivity for detecting EW signal. The normalization of QCD- $Z Z jj$  production  
 1710 ( $\mu_{QCD}^{III}$ ) in  $\ell\ell\ell'\ell'$  channel is determined by data from simultaneously fit in SR and QCD  
 1711 CR as described in section 6.4. The signal strength of EW- $Z Z jj$  production ( $\mu_{EW}$ ) is  
 1712 taken as parameter of interest and floated in the fit. The effects of the uncertainties related  
 1713 to normalizations and shapes described previously in section 6.5 of background processes  
 1714 in the MD distribution are all taken into account.

1715 In most case, a common nuisance parameter is used for each source of systematic in  
 1716 all bins and all categories. The statistical uncertainties for simulated samples are uncorre-  
 1717 lated among all bins, and the background uncertainties only applied to their corresponding  
 1718 backgrounds. Furthermore, to be more conservative, the generator modelling uncertainty  
 1719 for QCD- $Z Z jj$  production mentioned in section 6.5 is separated to be two nuisance pa-

1720 parameters in low and high MD region.

### 1721 6.7.3 Result of fit

1722 The results of fitting to  $\ell\ell\ell'\ell'$  final state are presented in table 6.13. To drive expected  
 1723 results, the observed data is used for QCD CR to extract normalization factor of QCD  
 1724 component ( $\mu_{QCD}^{III}$ ), while in SR, asimov data built from background prediction and signal  
 1725 model with SM assumed cross section is used.

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

Table 6.13 Observed  $\mu_{EW}$  and  $\mu_{QCD}^{\ell\ell\ell'jj}$ , as well as the observed and expected significance from the individual  $\ell\ell\ell'\ell'$  channel. The full set of systematic uncertainties are included.

1726 As a conclusion, the background-only hypothesis is rejected at  $5.5\sigma$  ( $3.9\sigma$ ) for ob-  
 1727 served (expected) data, which leads to the observation of EW- $ZZjj$  production.

1728 Figure 6.10 shows the post-fit MD distributions for  $\ell\ell\ell'\ell'$  channel in SR (left) and  
 1729 QCD CR (right). The EW- $ZZjj$  cross section measured in  $\ell\ell\ell'\ell'$  channel is extracted  
 to be  $0.94 \pm 0.26$  fb.

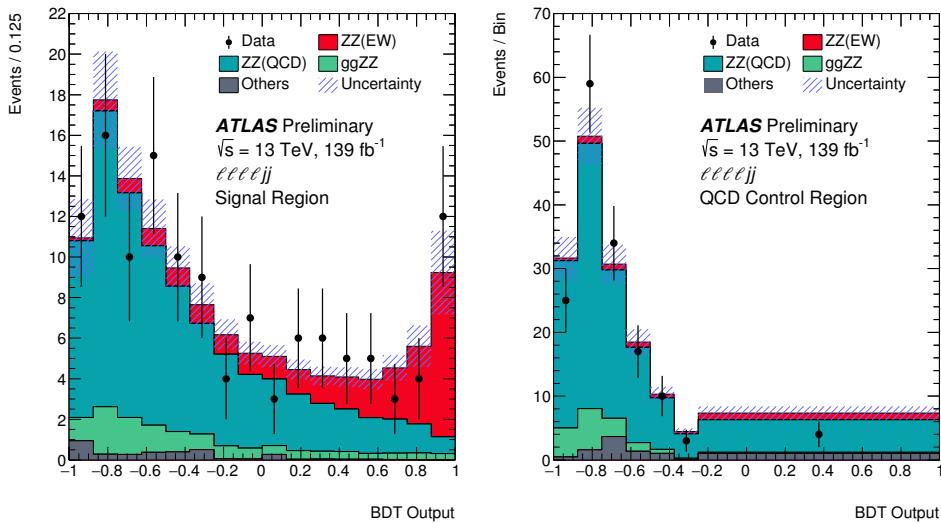
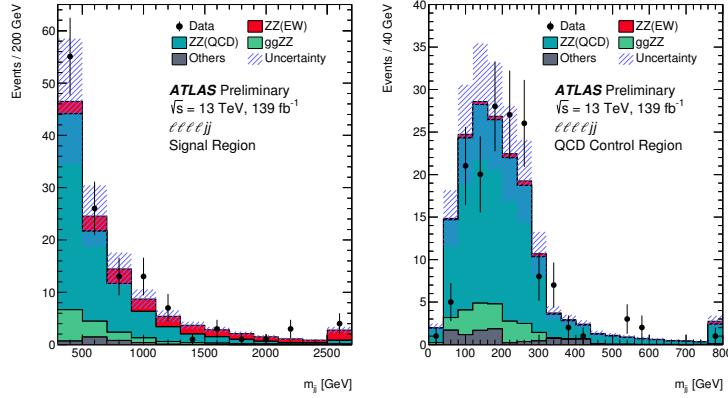


Fig. 6.10 Observed and post-fit expected multivariate discriminant distributions after the statistical fit in the  $\ell\ell\ell'\ell'$  SR (left) and QCD CR (right). The error bands include the experimental and theoretical uncertainties, as well as the uncertainties in  $\mu_{EW}$  and  $\mu_{QCD}^{\ell\ell\ell'jj}$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

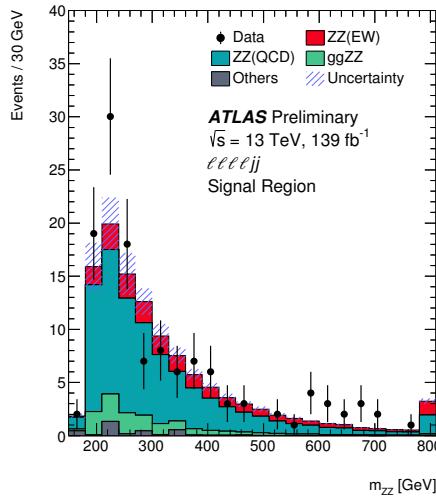
1730

1731 Figure 6.11 shows the  $m_{jj}$  distribution in SR (left) and QCD CR (right), where the  
 1732 normalization of EW and QCD processes are scaled according to their observed value in  
 1733 table 6.13. High  $m_{jj}$  region is more sensitive for EW- $ZZjj$  events detection from this

1734 figure. Figure 6.12 shows the spectrum of invariant mass of  $\ell\ell\ell'\ell'$  system ( $m_{ZZ}$ ) in SR  
 1735 also with the normalization of EW and QCD processes scaled.

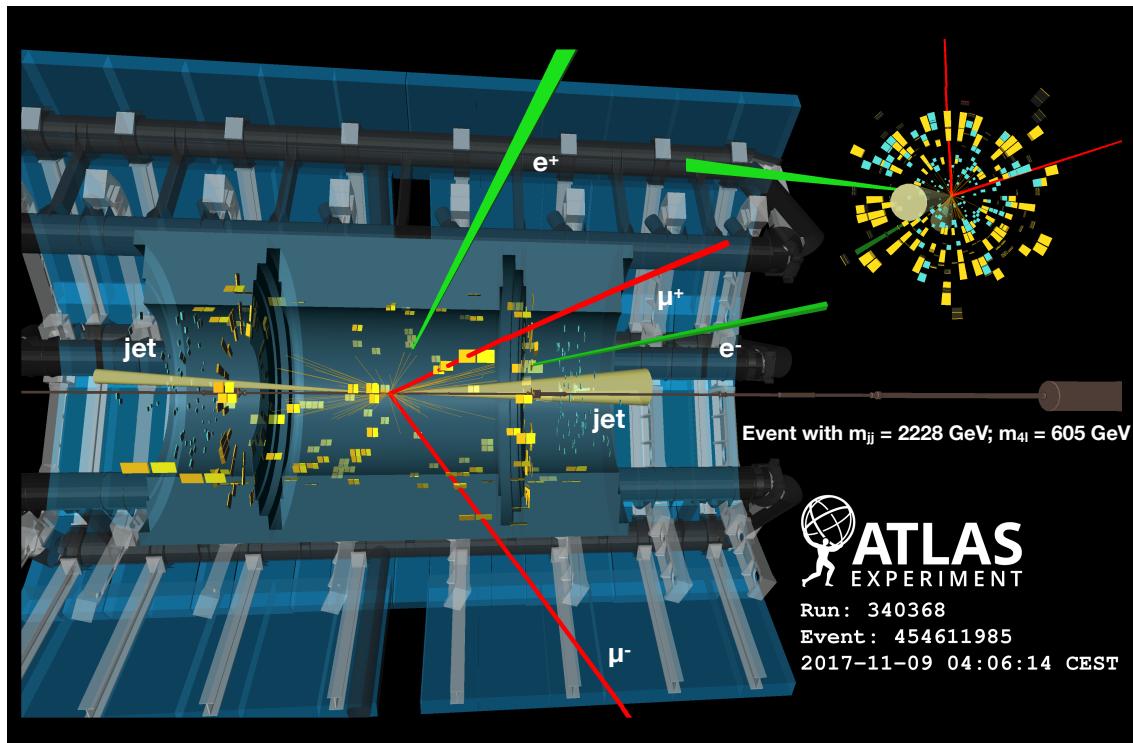


**Fig. 6.11** Observed and post-fit expected  $m_{jj}$  distributions in SR (left) and QCD CR (right). The error bands include the expected experimental and theoretical uncertainties. The error bars on the data points show the statistical uncertainty. The contributions from the QCD and EW production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

1736 Figure 6.13 is the display of one event candidate of EW- $Z Z jj$  production in  $2e2\mu$   
 1737 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.**

## 1738 6.8 Prospect study of EW- $Z Z jj$ production in HL-LHC

1739 The High-Luminosity Large Hadron Collider (HL-LHC) project aims to increase the  
 1740 luminosity by a factor of 10 beyond the LHC's design value to increase the potential for  
 1741 discoveries after 2025. The expected luminosity will reach  $3000 \text{ fb}^{-1}$  with the centre-of-  
 1742 mass energy of 14 TeV.

1743 As introduced in previous sections, with full run-2 data of  $139 \text{ fb}^{-1}$  collected by AT-  
 1744 LAS detector at the LHC, the EW- $Z Z jj$  production is the last channel of observation for  
 1745 VBS processes with massive bosons due to its very low cross section in  $Z Z$  channel. So  
 1746 we expect that this channel will benefit significantly from the increased luminosity at the  
 1747 HL-LHC, and can be studied in great details for this known mechanism.

1748 In this section, a prospective study is performed for EW- $Z Z jj$  production at the HL-  
 1749 LHC in the  $\ell\ell\ell'\ell'$  channel. The study uses  $3000 \text{ fb}^{-1}$  of simulated pp collision data at a  
 1750 centre-of-mass energy of 14 TeV as expected to be recorded by the ATLAS detector at the  
 1751 HL-LHC. All simulated events are produced at particle-level, and the detector effects of  
 1752 leptons and jets reconstruction and identification are estimated by corrections assuming  
 1753 the mean number of interactions per bunch crossing ( $\langle\mu\rangle$ ) of 200.

### 6.8.1 The ATLAS detector at HL-LHC

As the expectation of HL-LHC, the new Inner Tracker (ITk)<sup>[97]</sup> will extend the tracking acceptance capability of ATLAS detector to pseudorapidity ( $|\eta|$ ) up to 4.0. By including a forward muon trigger, the upgraded Muon Spectrometer<sup>[98]</sup> is also expected to provide muon identification capabilities to  $|\eta|$  up to 4.0. In addition, the new high granularity timing detector (HGTD)<sup>[99]</sup> designed to mitigate the pile-up (PU) effects is also expected to be installed in the forward region of  $2.4 < |\eta| < 4.0$ . More details of expected performance of the upgraded ATLAS detector at the HL-LHC has been reported in Ref.<sup>[100]</sup>

### 6.8.2 Simulation

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

### 6.8.3 Event selection

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

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

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

Table 6.14 summarized the number of selected signal and background events normalized to  $3000 \text{ fb}^{-1}$ . In addition to the *baseline* selection listed above, to compare the different detector scenarios at the HL-LHC, two alternative selections are also studied:

1. Reduce the lepton  $\eta$  region to 2.7, to understand the effect due to forward lepton reconstruction and identification with the upgraded ATLAS detector.
2. Only apply the PU jet suppression with region  $|\eta| < 2.4$ , to measure the improvement of *baseline* by extending the rejection range of PU jets at the HL-LHC with the installation of HGTD.

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

**Table 6.14 Comparison of event yields for signal ( $N_{\text{EW-ZZjj}}$ ) and background ( $N_{\text{QCD-ZZjj}}$ ) processes, and expected significance of EW- $ZZjj$  processes, normalized to  $3000 \text{ fb}^{-1}$  data at 14 TeV, with baseline and alternative selections. Uncertainties in the table refer to expected data statistical uncertainty at 14 TeV with  $3000 \text{ fb}^{-1}$ .**

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

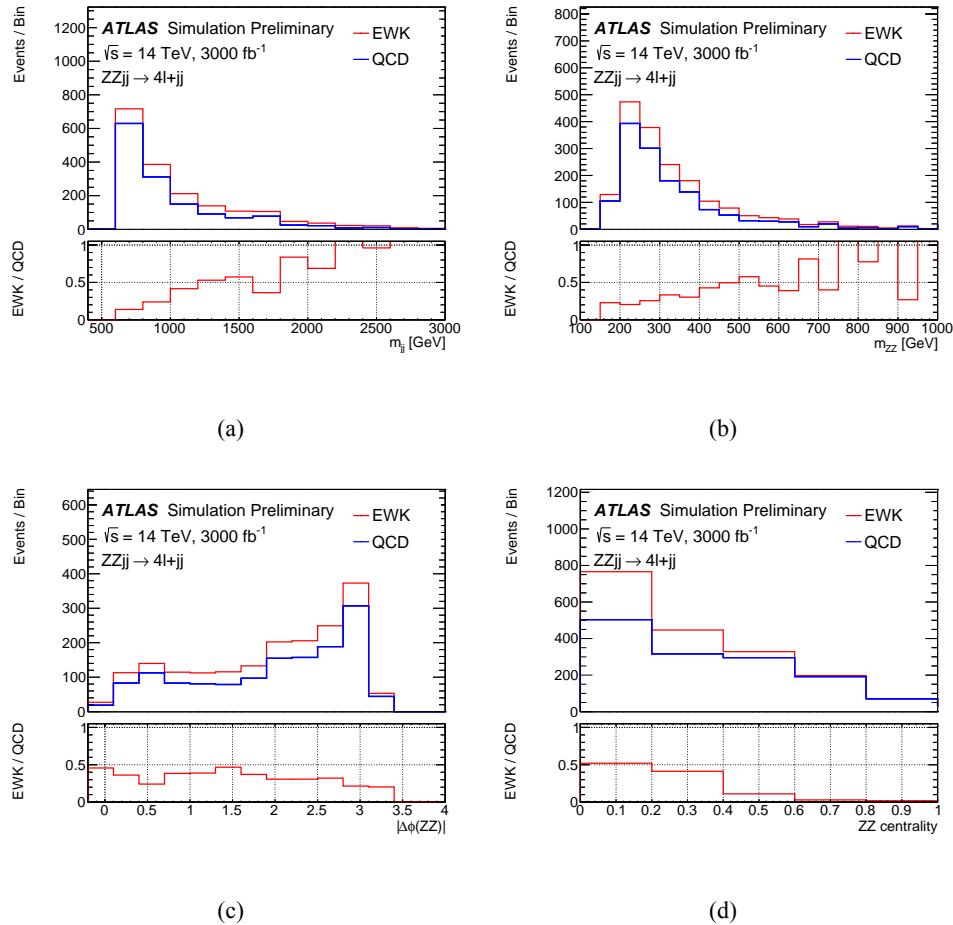
Figure 6.14 shows the kinematic distributions of di-jet invariant mass ( $m_{jj}$ ), the  $ZZ$  invariant mass ( $m_{ZZ}$ ) and the  $\phi$  separation of two Z bosons ( $|\Delta\phi(ZZ)|$ ) as well as the centrality of the  $ZZ$  system. The  $ZZ$  centrality is defined as:

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

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

#### 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- $ZZjj$  background process. Different sizes of systematics

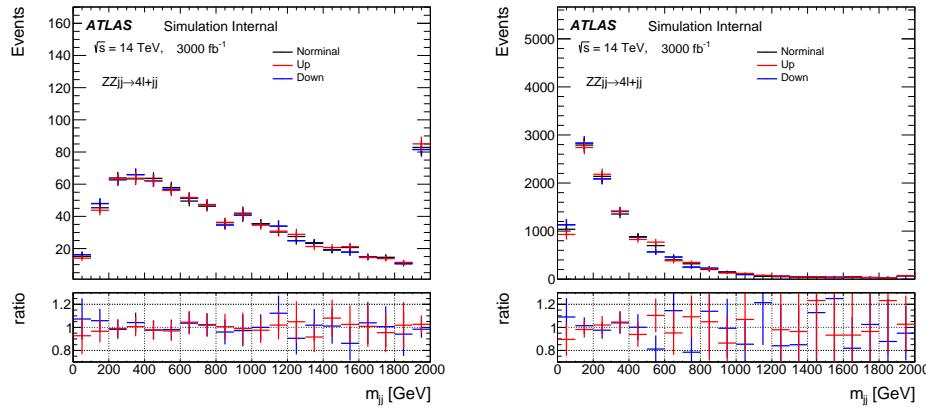


**Fig. 6.14** Detector-level distributions of EW- and QCD- $Z Z j j$  processes with selected events in defined phase space at 14 TeV of (a)  $m_{jj}$ , (b)  $m_{ZZ}$ , (c)  $|\Delta\phi(ZZ)|$ , (d) ZZ centrality, normalized to  $3000 \text{ fb}^{-1}$ .

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.<sup>[100]</sup>, 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.

Since the final result relies greatly on the uncertainties, especially the theoretical un-



**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  $3000 \text{ fb}^{-1}$  at 14 TeV. Upgrade Performance Function is used to extract the uncertainties with *baseline* setting.**

certainties 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

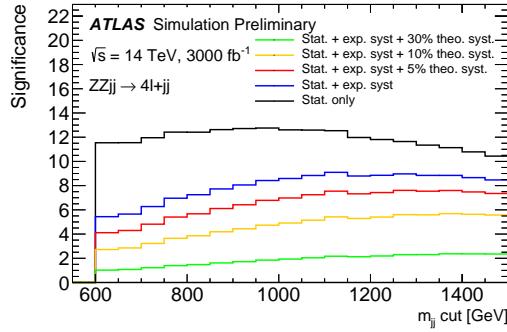
In this analysis, instead of a statistical fit, the expected significance of EW- $Z Z jj$  production is calculated as:

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

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  $3000 \text{ fb}^{-1}$ .

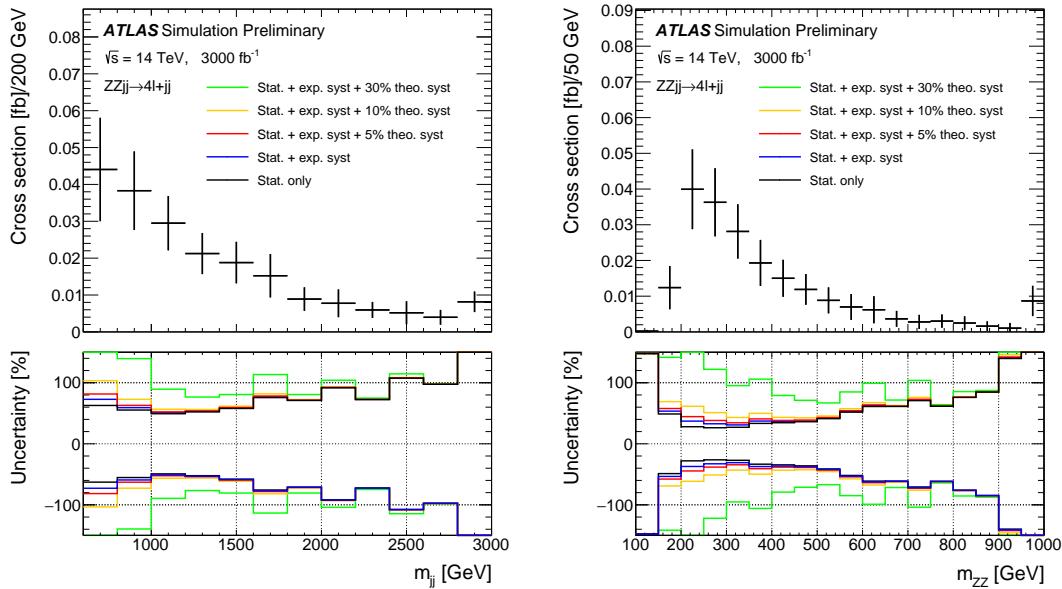
Based on baseline selection of  $m_{jj} > 600 \text{ GeV}$ , an additional scan over different  $m_{jj}$  cuts is also performed with a step of 50 GeV under different systematic conditions, as shown in figure 6.16.

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



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

The expected differential cross sections are calculated as:



**Fig. 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  $3000 \text{ fb}^{-1}$ . The bottom panel shows impact of different sizes of systematic uncertainties.

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

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

1845 at 14 TeV, and  $N_{QCD-ZZjj}$  and  $N_{EW-ZZjj}$  are the number of predicted events of QCD-  
 1846  $ZZjj$  and EW- $ZZjj$  processes in particle-level. The  $C_{EW-ZZjj}$  factor represents the  
 1847 detector efficiency for EW- $ZZjj$  processes introduced in section 6.6.1. The interference  
 1848 between EW- and QCD-  $ZZjj$  processes is ignored due to its minor contribution.

1849 The value of expected integrated cross section as well as its uncertainty under different  
 1850 systematic conditions are shown in table 6.15 with  $3000 \text{ fb}^{-1}$  luminosity at 14 TeV. The  
 1851 statistical uncertainty is at 10% level when with such large luminosity. The result is dom-  
 1852 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	$\pm 0.02$	$\pm 0.04$	$\pm 0.05$	$\pm 0.08$	$\pm 0.21$

**Table 6.15 Summary of expected cross-section measured with different theoretical uncer-  
 tainties. The statistical uncertainty is computed from expected data yield with  
 $3000 \text{ fb}^{-1}$  at 14 TeV. Different uncertainties are treated as uncorrelated and  
 summed quadratically.**

1853

## 1854 6.9 Conclusion

1855 The fiducial cross section for inclusive  $ZZjj$  production is measured in this section,  
1856 with a total relative uncertainty of 11% for the  $\ell\ell\ell'\ell'$  final state, and found to be com-  
1857 patible with the SM prediction. The observation of electroweak production of two jets in  
1858 association with a  $Z$ -boson pair decay to  $\ell\ell\ell'\ell'$  final state using  $139 \text{ fb}^{-1}$  of 13 TeV pp  
1859 collision data collected by ATLAS experiment at the LHC is presented in this section.  
1860 The search for electroweak production of two jets in association with a  $Z$ -boson pair is  
1861 based on multivariate discriminants (MD) to enhance the separation between the signal  
1862 and backgrounds. In  $\ell\ell\ell'\ell'$  final state, the background-only hypothesis is rejected with  
1863 an observed (expected) significance of  $5.5$  ( $3.9$ )  $\sigma$ , which gives the first observation of  
1864 electroweak production in  $ZZjj$  channel.

1865 In addition, the prospective study for the EW- $ZZjj$  production at the HL-LHC in  
1866 the  $\ell\ell\ell'\ell'$  channel, using  $3000 \text{ fb}^{-1}$  simulated pp collision data at a centre-of-mass  
1867 energy of 14 TeV has been presented. The precision of the expected measurements of the  
1868 integrated and differential cross sections as a function of dijet or  $\ell\ell\ell'\ell'$  invariant mass are  
1869 shown. Under the assumption of theoretical uncertainty for the QCD- $ZZjj$  processes and  
1870 experimental uncertainty for jets being constraint at 5% level respectively, with statistical  
1871 uncertainty in  $3000 \text{ fb}^{-1}$  being considered, the observation of the EW- $ZZjj$  process can  
1872 reach a significance of  $7 \sigma$ .

1873 **Chapter 7 Search for heavy resonances decaying into**  
1874 **a pair of  $Z$  bosons in  $\ell\ell\ell'\ell'$  final state using pp**  
1875 **collision data collected by ATLAS detector from 2015 to**  
1876 **2018**

1877 **7.1 Introduction**

1878 A new particle was discovered by the ATLAS and CMS Collaborations at the LHC<sup>[3-4]</sup>  
1879 in 2012. Both two experiments have confirmed that the properties including spin, cou-  
1880 plings and parity of this new particle are consistent with Higgs boson predicted in the  
1881 Standard Model (SM), which is an important milestone in understanding of the mecha-  
1882 nism of EWSB. Nevertheless, the possibility that this newly discovered particle is just a part  
1883 of the extended Higgs sector as predicted by various extensions in the SM cannot be ruled  
1884 out. There are many models predicted the existence of new heavy resonances decaying  
1885 into dibosons, such as a heavy spin-0 neutral Higgs boson<sup>[101]</sup> and the two-Higgs-doublet  
1886 models (2HDM)<sup>[6]</sup>, as well as the spin-2 Kaluza–Klein (KK) excitations of the graviton  
1887 ( $G_{KK}$ )<sup>[102]</sup>.

1888 Though with smaller branching ratio comparing to semileptonic or fully hadronic de-  
1889 cay channels, the  $\ell\ell\ell'\ell'$  final state has its unique sensitivity in mass range smaller than  
1890 1 TeV region due to its good mass resolution and relative smaller experimental and the-  
1891 oretical systematics. This section presents the search for heavy resonance decaying into  
1892 a pair of  $Z$  bosons to the  $\ell\ell\ell'\ell'$  final state, in which  $\ell$  denotes to either an electron  
1893 or a muon. Several signal hypotheses are considered. The first hypothesis is a heavy  
1894 Higgs boson (spin-0 resonance) under the narrow-width approximation (NWA). Then as  
1895 several theoretical models prefer non-negligible natural widths, the models under large-  
1896 width approximation (LWA), assuming widths of 1%, 5%, 10% and 15% of the resonance  
1897 mass, are also studied. In addition, the graviton excitations (spin-2 resonance) under the  
1898 Randall–Sundrum model are also searched. It is assumed that the heavy resonance is  
1899 produced predominantly via the gluon-gluon Fusion (ggF) and the Vector Boson Fusion  
1900 (VBF) productions, but with the unknown ratio of two production rates. So the results are  
1901 separated for ggF and VBF production modes. To gain more sensitivity, the  $\ell\ell\ell'\ell'$  events  
1902 are classified into ggF- and VBF-enriched categories. Moreover, for the NWA model, the  
1903 categorizations are studied under both cut-based and multivariate (MVA) -based methods,  
1904 the details of categorization are shown in following sections.

1905 The search uses the four-lepton invariant mass in the range of 200 GeV to 2000 GeV for  
 1906 signal hypothesis of spin-0 resonance under the NWA model, and from 400 GeV to  
 1907 2000 GeV for the one under the LWA models. And the spin-2 graviton signals are searched  
 1908 in the mass range from 600 GeV to 2000 GeV. The data collected by ATLAS detector at  
 1909 the LHC from 2015 to 2018 at the centre-of-mass energy of 13 TeV is used. In case of no  
 1910 excess, upper limits on the production rate of different signal hypotheses are computed  
 1911 from statistical fits to  $m_{4l}$  distribution.

1912 **7.2 Data and MC samples**

1913 **7.2.1 Data samples**

1914 The data used in this analysis are collected by ATLAS detector at the centre-of-mass  
 1915 energy of 13 TeV during the years of 2015 to 2018. Only events passing the latest Good  
 1916 Run List (GRL) released by the Data Quality group from ATLAS experiment as listed in  
 1917 section 6.2.1 corresponding to an integrated luminosity of  $139.0 \pm 2.4 \text{ fb}^{-1}$  are used.  
 1918 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	$3.86 \text{ fb}^{-1}$	$36.2 \text{ fb}^{-1}$	13.4	28.1
2016	$35.6 \text{ fb}^{-1}$		25.1	52.2
2017	$46.9 \text{ fb}^{-1}$	$44.3 \text{ fb}^{-1}$	37.8	79.8
2018	$60.6 \text{ fb}^{-1}$	$58.5 \text{ fb}^{-1}$	36.1	88.6

1920 **7.2.2 Background MC simulations**

1921 Background processes considered in this analysis include  $ZZ$  ( $q\bar{q} \rightarrow ZZ$ ,  $gg \rightarrow$   
 1922  $ZZ$ ), triboson ( $WWZ$ ,  $WZZ$ ,  $ZZZ$ ),  $Z+jets$  and top-quark ( $t\bar{t}$ ,  $ttV$ ) processes.

1923 The QCD  $q\bar{q} \rightarrow ZZ$  process is modelled using Sherpa 2.2.2<sup>[82]</sup> with the  
 1924 NNPDF3.0NNLO<sup>[83]</sup> PDF, where events with up to one (three) outgoing partons are gen-  
 1925 erated at NLO (LO) in pQCD. The production of  $ZZ$  from the gluon-gluon initial state  
 1926 with a four-fermion loop or with an exchange of the Higgs boson, which has an order  
 1927 of  $\alpha_S^4$  in QCD, is not included in this Sherpa simulation. So a separate  $gg$  induced  $ZZ$   
 1928 sample including the continuum background, the SM Higgs boson, and the interference  
 1929 contribution is modelled using Sherpa 2.2.2 with the NNPDF3.0NNLO PDF set, and with

1930 an additional 1.7 k-factor<sup>[84]</sup> being applied. The EW- $ZZjj$  production is simulated using  
1931 Sherpa 2.2.2 with the NNPDF3.0NNLO PDF, and the  $ZZZ \rightarrow \ell\ell\ell'\ell'qq$  process is also  
1932 taken into account in this sample.

1933 The  $Z+jets$  events are generated using Sherpa 2.2.2 with the NNPDF3.0NNLO PDF,  
1934 in which the ME is calculated for up to two partons with next-to-leading-order (NLO)  
1935 accuracy in pQCD and up to four partons with LO accuracy. The  $Z+jets$  events are  
1936 normalized using the next-to-next-to-leading-order (NNLO) cross section. The triboson  
1937 processes with full leptonic decays and at least four prompt charged leptons are generated  
1938 using Sherpa 2.1.1. For top-quark pair ( $t\bar{t}$ ) production and the single top-quark produc-  
1939 tions in  $t$ -channel,  $s$ -channel and  $Wt$ -channel, the Powheg-Box v2 is used with the CT10  
1940 PDF. The productions of  $t\bar{t}$  in association with  $Z$  boson(s) ( $ttZ$ ) is modelled with Mad-  
1941 Graph5\_aMC@NLO.

### 1942 7.2.3 Signal MC simulations

1943 One model considered in this analysis is heavy spin-0 resonance under the Narrow  
1944 Width Approximation (NWA) simulated using Powheg-Box v2 MC event generator with  
1945 the CT10 PDF. The gluon-gluon fusion (ggF) production mode and vector-boson fusion  
1946 (VBF) production mode are calculated separately with matrix elements up to NLO in  
1947 QCD. The Powheg-Box is interfaced to Pythia8 for parton showering, and for decaying  
1948 the Higgs boson into the  $H \rightarrow ZZ \rightarrow \ell\ell\ell'\ell'$  final states. Events of NWA signal are  
1949 generated at mass points between 200 GeV to 2000 GeV using the step of 100 (200) GeV up  
1950 to (above) 1 TeV in both ggF and VBF production modes.

1951 In addition, heavy Higgs boson events under the Large Width Approximation (LWA)  
1952 with widths of 1%, 5%, 10% and 15% of the boson mass are generated using Mad-  
1953 Graph5\_aMC@NLO 2.3.2 interfaced to Pythia8. Only ggF production is consid-  
1954 ered. Mass points between 400 GeV to 2000 GeV are simulated with the step of 100  
1955 (200) GeV up to (above) 1 TeV. To describe jet multiplicity, MadGraph5\_aMC@NLO is  
1956 used to simulated process of  $pp \rightarrow H + \geq 2\text{jets}$  at NLO in QCD with the FxFx merging  
1957 scheme<sup>[103]</sup>.

1958 Spin-2 Kaluza–Klein (KK) gravitons ( $G_{KK}$ ) from the Bulk Randall–Sundrum  
1959 model<sup>[104]</sup> are also studies in this analysis. Events are generated by Mad-  
1960 Graph5\_aMC@NLO at LO in QCD, which is then interfaced to Pythia8 for parton show-  
1961 ering. The  $G_{KK}$ -gluon coupling  $k/\overline{M}_{\text{Planck}}$ , where  $k$  is the curvature scale of the extra  
1962 dimension and  $\overline{M}_{\text{Planck}}$  is the reduced Planck mass, is set to 1. The width of the resonance  
1963 is correlated with the coupling  $k/\overline{M}_{\text{Planck}}$  and in this configuration is around  $\square 6\%$  of its

mass. The mass of the  $G_{KK}$  is the only free parameter in this simplified model. Mass points between 600 GeV to 2 TeV with 200 GeV spacing were generated.

## 7.3 Analysis selections

### 7.3.1 Objects selection

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

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

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

**Jets:** Jets are clustered using the anti- $k_t$  algorithm with radius parameter  $R = 0.4$  implemented in the FastJet package as described in section 4.2.5. The ‘particle flow’ (PFlow) objects<sup>[105]</sup>, which combines measurements from both the tracker and the calorimeter, are used as inputs to the FastJet package. The energy deposited in the calorimeter by all charged particles is removed, and the jet reconstruction is performed on an ensemble of PFlow objects consisting of the remaining calorimeter energy and tracks which are matched to the hard interaction. This improves the accuracy of the charged-hadron measurement, while retaining the calorimeter measurements of neutral-particle energies. Compared to only using topological clusters, jets reconstructed with the particle flow algorithm with  $p_T > 30$  GeV have approximately 10% better transverse momentum resolution. The jets used in this analysis are then required to have  $p_T > 30$  GeV and  $|\eta| < 4.5$ . To further reduce the effects of pile-up jets, a jet vertex tagger (JVT) is applied to jets with

<sup>1997</sup>  $p_T < 60 \text{ GeV}$  and  $|\eta| < 2.4$ .

<sup>1998</sup> **Overlap removal:** As the selected jet and lepton candidates can be reconstructed 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 <sup>2000</sup> calorimeter-tagged and does not have a MS track, or is a segment-tagged muon, otherwise <sup>2001</sup> the muon is selected. The jet overlapping with electron (muon) within a cone of size of <sup>2002</sup>  $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.2(0.1)$  are removed. <sup>2003</sup>

### <sup>2004</sup> 7.3.2 Event selection

<sup>2005</sup> First of all, the four-lepton events are required to pass single or multi-lepton triggers. <sup>2006</sup> Due to the increasing of peak luminosity and pile-up, the  $p_T$  and  $E_T$  thresholds of triggers <sup>2007</sup> increase slightly during the data-taking periods from 2015 to 2018. Table 7.2 summarizes <sup>2008</sup> the triggers used for  $\ell\ell\ell'\ell'$  channel. The overall trigger efficiency for selected signal <sup>2009</sup> 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	$\mu 20; \mu 50; \mu 60$	$\mu 24; \mu 26; \mu 40; \mu 50$	$\mu 26; \mu 50; \mu 60$	$\mu 26; \mu 50; \mu 60$
single electron	$e24; e60; e120$	$e26; e60; e140; e300$	$e26; e60; e140; e300$	$e26; e60; e140; e300$
dimuon	$2\mu 10; \mu 18_{-\mu 8}$	$2\mu 10; 2\mu 14; \mu 22_{-\mu 8}$	$2\mu 14; \mu 22_{-\mu 8}$	$2\mu 14; \mu 22_{-\mu 8}$
dielectron	$2e12$	$2e15; 2e17$	$2e17; 2e24$	$2e17; 2e24$
electron-muon	$e24_{-\mu 8}$	$e24_{-\mu 8}; e26_{-\mu 8}$ $e17_{-\mu 14}; e7_{-\mu 24}; 2e12_{-\mu 10}; e12_{-\mu 10}$	$e26_{-\mu 8}$	$e26_{-\mu 8}$
trimuon	$\mu 18_{-2\mu 4}$	$\mu 11_{-2\mu 4}; \mu 6_{-2\mu 4}; \mu 20_{-2\mu 4}; 3\mu 4$ $3\mu 6$	$4\mu 4; \mu 20_{-2\mu 4}; 3\mu 4$	$\mu 20_{-2\mu 4}$
trilepton	$e17_{-2e9}$	$e17_{-2e9}; e17_{-2e10}$	$e24_{-2e12}$	$e24_{-2e12}$

<sup>2010</sup> 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 <sup>2011</sup> and 10 GeV. If there are more than one combination of lepton pairing in quadruplet, the <sup>2012</sup> pairing is selected by keeping it with the mass of lepton pairs closest (leading pair, refers <sup>2013</sup> as  $m_{12}$ ) and second closest (sub-leading pair, refers as  $m_{34}$ ) to  $Z$  boson mass. The mass <sup>2014</sup> of leading pair is required to satisfy  $50 < m_{12} < 106$  GeV, while the sub-leading pair is <sup>2015</sup> required to be less than 115 GeV and larger than 50 GeV. <sup>2016</sup>

<sup>2017</sup> The two lepton pairs in quadruplet are required to have angular separation with  $\Delta R >$  <sup>2018</sup> 0.1. To suppress the contribution from  $J/\psi \rightarrow \ell\ell$  decays, for 4 $\mu$  and 4 $e$  quadruplets, <sup>2019</sup> the events are rejected if any opposite-sign same-flavour lepton pair is found with mass

2020 below 5 GeV. If there are more than one quadruplets from different channels in event at  
 2021 this point, the one with highest expected signal rate is selected in the order of  $4\mu$ ,  $2e2\mu$ ,  
 2022  $4e$ . The transverse impact-parameter significance ( $|d_0|/\sigma_{d_0}$ ) for muons (electrons) is than  
 2023 required to be smaller than 3 (5) to suppress the backgrounds from heavy-flavour hadrons.

2024 In addition, the track- and calorimeter- based isolation criteria is required for all elec-  
 2025 trons and muons to further suppress the reducible backgrounds of  $Z+jets$  and  $t\bar{t}$ . For  
 2026 lepton isolation selection, the two track- and calorimeter- based variables,  $E_T^{topocone}$  and  
 2027  $p_T^{varcone}$  as described in section 4.2.4 (section 4.2.3) for muons (electrons), are vulnerable  
 2028 to pileup. For track-based variable, this is because of additional tracks in the event. The  
 2029 definition of  $p_T^{varcone}$  attempts to limit the tracks used in the calculation to those from the  
 2030 vertex via a loose cut of  $|z_0 \sin(\theta)| < 3$ , which proved to be too loose in new pile-up  
 2031 regime 2017 and 2018 datasets. So new track-based variable is used, by adding a require-  
 2032 ment that the track be used in determining the vertex, or that, if not, it both pass the cut on  
 2033  $|z_0 \sin(\theta)|$  and not be used in determining any other vertex, which makes the track-based  
 2034 variable to be more isolation-robust in the high pile-up regime. The new variable is named  
 2035 as  $ptvarcone[cone]_TightTTVA_pt[p_T \text{ cut}]$ , where [cone] is the cone size and [ $p_T$  cut] is  
 2036 the cutoff for including tracks in the calculation.

2037 For calorimeter-based variable, the calculation of  $E_T^{topocone}$  corrects the pile-up effects  
 2038 by subtracting an average pileup contribution computed over the whole detector. But with  
 2039 the increasing of energy density of pile-up events, the root mean square (RMS) of  $E_T^{topocone}$   
 2040 variable increases, which leads to the increment of possibility that the pile-up fluctuations  
 2041 are not be accounted for correctly. One possible solution is that use particle-flow (PFlow)  
 2042 method to calculate the calorimeter isolation. As part of PFlow reconstruction process, it  
 2043 assigns the clusters to tracks which improves the track-cluster association for better de-  
 2044 termination of the raw value of the  $E_T$  in the cone. And using PFlow jets to calculate the  
 2045 pileup correction provides a further improvement. So a resulting variable named neflow-  
 2046 isol[cone] is used. Finally, a requirement of isolation, called *FixedCutPFlowLoose*, which  
 2047 gives better performance in high piup-up condition is applied to electrons and muons as:  
 2048  $(\max(ptcone20_TightTTVA_pt500, ptvarcone30_TightTTVA_pt500) + 0.4 \times neflow-$   
 2049  $isol20) / p_T < 0.16$

2050 On the top of impact parameter cut and lepton isolation cut, the four-lepton candidates  
 2051 are also required to originate from a common vertex to reduce  $Z+jets$  and  $t\bar{t}$  backgrounds.  
 2052 This is ensured by applying a vertex fit  $\chi^2$  cut of 4 ID tracks of lepton candidates satisfying  
 2053  $\chi^2/N_{dof} < 6$  (9) for events in  $4\mu$  ( $4e$  and  $2e2\mu$ ) channel(s).

2054 To improve the mass resolution, the QED process of final state radiation (FSR) pho-

2055 tons in  $Z$  boson decays are taken into account in the reconstruction of  $Z$  bosons. The  
2056 four-momentum of any reconstructed photon that is consistent with having been radiated  
2057 from lepton(s) in leading pair are added into final state. Moreover, the four-momenta of  
2058 leptons in both (leading and sub-leading) pairs are recomputed by performing a  $Z$ -mass-  
2059 constrained kinematic fit, which uses a Breit–Wigner  $Z$  boson line-shape and Gaussian  
2060 function with width set to the expected lepton resolution per lepton to model the momen-  
2061 tum response function. The  $Z$ -mass-constrained mass improves the  $m_{4\ell}$  resolution by up  
2062 to 15% depending on  $m_H$ .

2063 In summary, table 7.3 lists a comprehensive object and event level selection as de-  
2064 scribed above. Table 7.4 to 7.7 shows the cutflow of NWA ggF and VBF signal at the  
2065 mass points of 600 and 1000 GeV as examples.

2066 **7.3.3 Event categorizations**

2067 To improve the sensitivity of search in both VBF and ggF production mode in NWA  
2068 model, events are classified into the VBF- and ggF- enriched categories. With the statistic  
2069 increasing in full run-2 data, a multivariate (MVA) based classifier has been studied for  
2070 NWA signal, while in the meantime the traditional cut-based classifier is also used as a  
2071 model-independent result for all three (NWA, LWA, graviton) models.

2072 **1. Cut-based categorization**

2073 There are four categories in total: one VBF-enriched category and three ggF-enriched  
2074 categories. The categorization is defined based on kinematic cuts:

- 2075 • VBF-CBA-enriched category: Events have at least two selected jets as defined in  
2076 section 7.3.1, with the two leading jets being separated by  $|\Delta\eta_{jj}| > 3.3$  and invariant  
2077 mass satisfying  $m_{jj} > 400$  GeV;
- 2078 • ggF-CBA-enriched categories: The remaining events that are not classified into  
2079 VBF-enriched category. Then events are categorized into three channels based on  
2080 lepton-flavor, namely ggF\_2e2μ, ggF\_4e and ggF\_4μ.

2081 **2. MVA-based categorization**

2082 In order to target different production modes, two types of classifiers, one dedicate to  
2083 VBF production while the other one for ggF, have been trained using deep neural network  
2084 technique. Details of two classifiers are described as below:

2085 **DNN models**

2086 Figure 7.1 shows the architecture of VBF (left) and ggF (right) network. The VBF  
2087 network includes three parts: two recurrent neural networks (RNNs) and one multilayer  
2088 perceptron (MLP). The ggF network consists of one RNN and one MLP.

**Table 7.3 Summary of the object and event selection requirements.**

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

**Table 7.4 Cutflow table for a narrow-width ggF signal sample at  $m_H = 600$  GeV.**  $N_{\text{event}}$  denotes the number of events selected after each cut is applied, normalized to  $139 \text{ fb}^{-1}$ , according to the expected upper limit on the cross section. The acceptances (the proportion of events selected relative to the initial number of events) are also included.

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

**Table 7.5 Cutflow table for a narrow-width ggF signal sample at  $m_H = 1000$  GeV.**  $N_{\text{event}}$  denotes the number of events selected after each cut is applied, normalized to  $139 \text{ fb}^{-1}$ , according to the expected upper limit on the cross section. The acceptances (the proportion of events selected relative to the initial number of events) are also included.

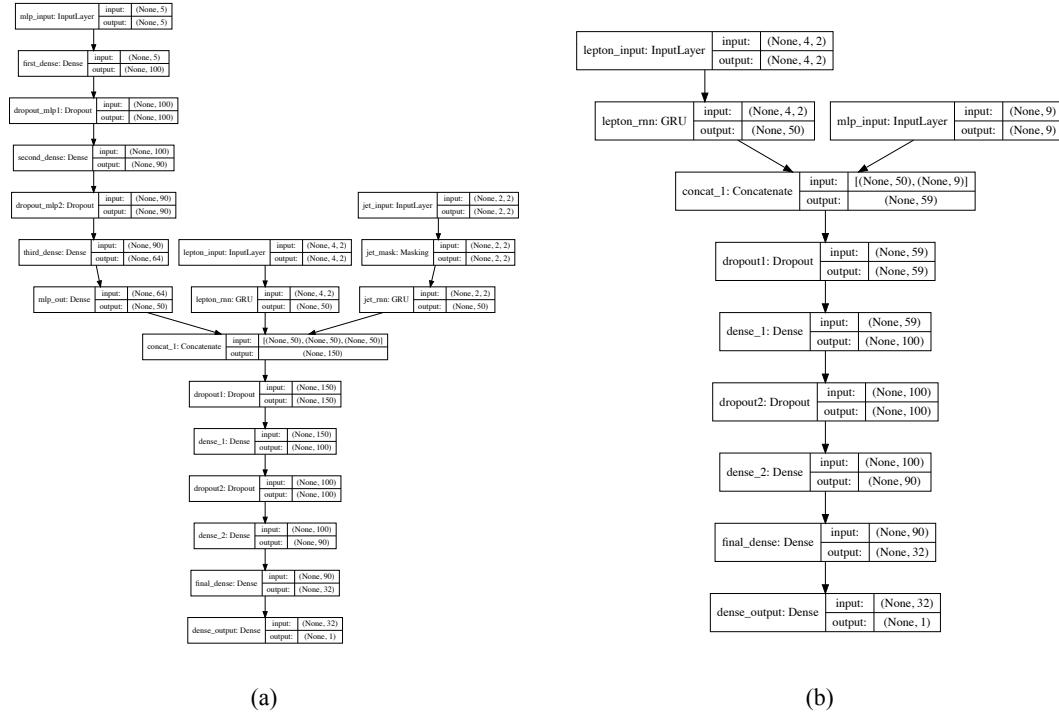
	$N_{\text{event}}$	$N_{\text{event}}/\text{BR}(ZZ \rightarrow \ell\ell\ell'\ell')$	Acc. [%]	Acc. · $\text{BR}(ZZ \rightarrow \ell\ell\ell'\ell') \cdot 1000$
Initial	5.603	1 240.8	100.00	4.516
Lepton selection	2.141	474.1	38.21	1.725
SFOS	1.944	430.5	34.70	1.567
Kinematic cuts	1.943	430.3	34.68	1.566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

**Table 7.6 Cutflow table for a narrow-width VBF signal sample at  $m_H = 600$  GeV.**  $N_{\text{event}}$  denotes the number of events selected after each cut is applied, normalized to  $139 \text{ fb}^{-1}$ , according to the expected upper limit on the cross section. The acceptances (the proportion of events selected relative to the initial number of events) are also included.

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

**Table 7.7 Cutflow table for a narrow-width VBF signal sample at  $m_H = 1000$  GeV.**  $N_{\text{event}}$  denotes the number of events selected after each cut is applied, normalized to  $139 \text{ fb}^{-1}$ , according to the expected upper limit on the cross section. The acceptances (the proportion of events selected relative to the initial number of events) are also included.

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



**Fig. 7.1 (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.

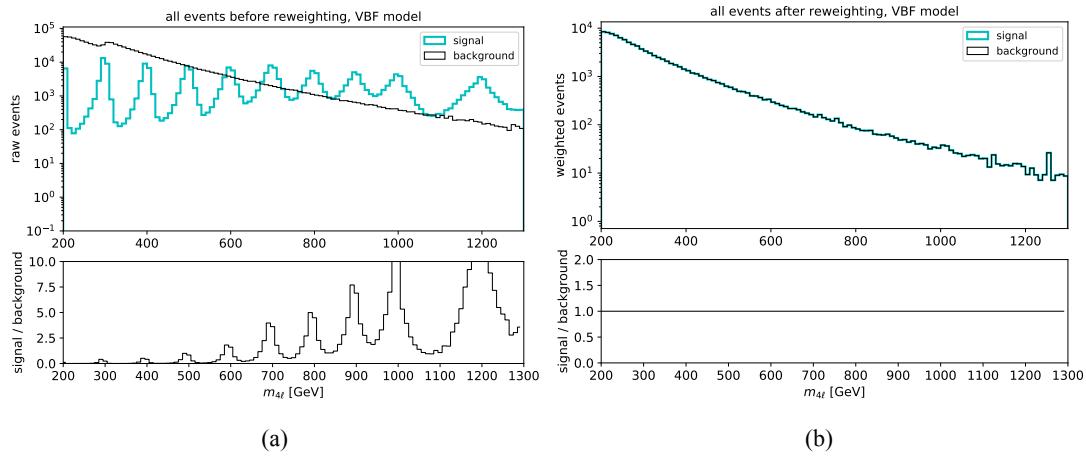
In order to assign equivalent importance to signals with different mass assumptions, during the training, signal events are reweighted to follow the  $m_{4\ell}$  distribution from background, as shown in figure 7.2 (figure 7.3) before (left) and after(right) reweighting for VBF (ggF) samples.

## Input features

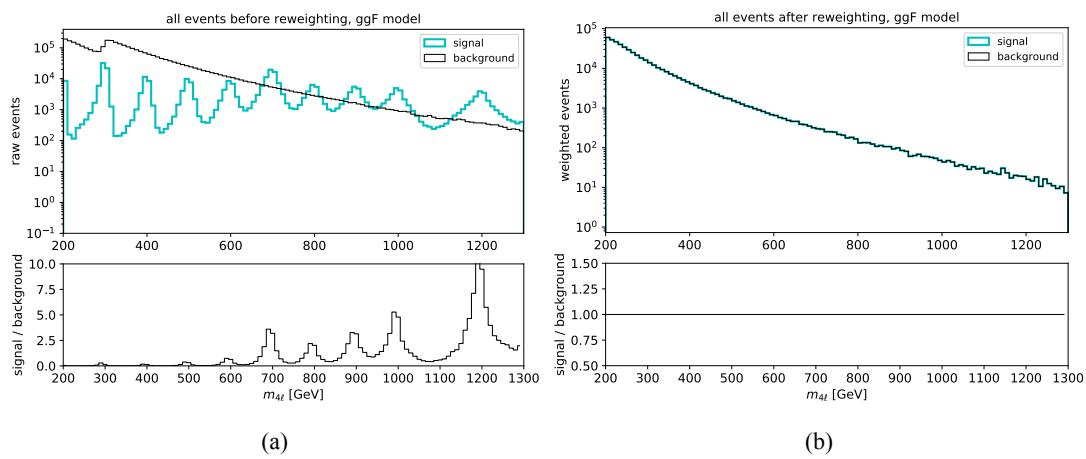
Table 7.8 (table 7.9) lists the input features used for VBF (ggF) network during the training. For VBF network, one RNN (the other one) takes the  $p_T$  and  $\eta$  of  $p_T$ -ordered four leptons (two leading jets) as input features, which intends to study the time relationship from particle decay between leptons (jets). For ggF network, the only one RNN model takes the  $p_T$  and  $\eta$  of  $p_T$ -ordered four leptons as inputs.

## Evaluation of models

Figure 7.4 shows the output of “ggF-classifier” and “VBF-classifier” for data, SM



**Fig. 7.2** (a)  $m_{4\ell}$  distribution of raw (unweighted) training events for VBF signal (blue) and background (black); (b)  $m_{4\ell}$  distribution of weighted VBF signal (blue) and background (black) used at training time.



**Fig. 7.3** (a)  $m_{4\ell}$  distribution of raw (unweighted) training events for ggF signal (blue) and background (black); (b)  $m_{4\ell}$  distribution of weighted ggF signal (blue) and background (black) used at training time.

**Table 7.8 Input features used in the “VBF-classifier” for the  $\ell\ell\ell'\ell'$  analysis. The RNN stands for the recurrent neural network and MLP for the multilayer perceptron.**

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

backgrounds and an example signal at 600 GeV. The ggF and VBF signals cross section are set to be one hundred times of their observed upper limit described in section 7.7.3 for ggF output and fifty times of the observed upper limit for VBF output for best visibility.

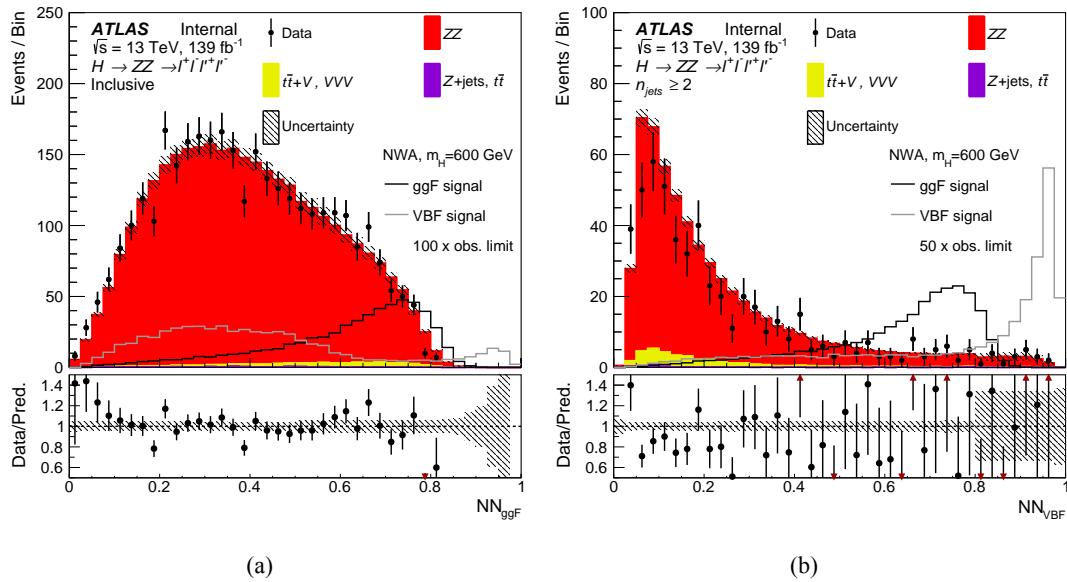
Then the optimal cut at output score from each classifier is chosen based on an overall good performance of classifier to have a large significance improvement while retaining a high signal efficiency. Figure 7.5 shows the significance improvements of MVA-based cuts when comparing with cut-based one at different VBF (left) and ggF (right) mass samples, where the significance is calculated as:

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

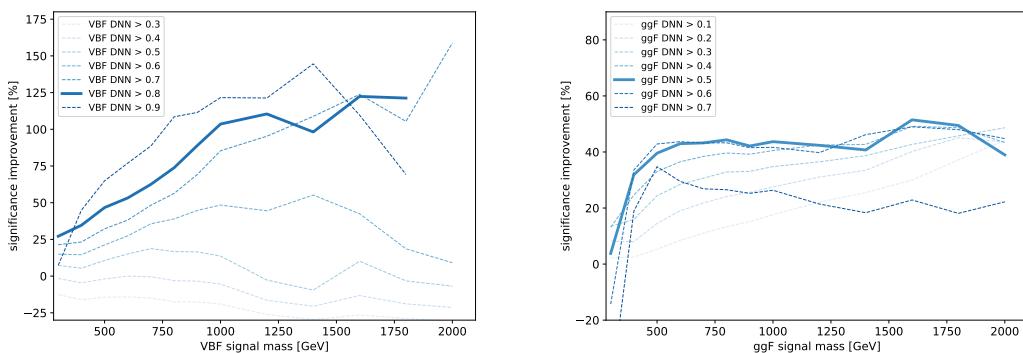
Cut at 0.5 (0.8) for VBF (ggF) classifier is chosen as shown in solid lines.

Then the events passing VBF classifier are categorized into VBF-MVA-enriched category. Otherwise, the events failing VBF classifier but passing ggF classifier are categorized into ggF-MVA-high category, which is further split into 3 channels. All remaining events are sorted into one additional ggF-MVA-low category. Thus there are five categories defined in MVA-based categorization. In summary, cuts applied in categorization are defined as follow, and these different phase spaces are also illustrated in figure 7.6.

- VBF-MVA-enriched category: Events have at least two selected jets ( $N_{\text{jets}} \geq 2$ ), and with  $DNN_{\text{VBF}} > 0.8$ ;



**Fig. 7.4** The output score of “ggF-classifier” (a) and “VBF-classifier” (b) with the events passing the common event selections for the data, the SM backgrounds and an example of a NWA signal with a mass of 600 GeV. For the “VBF-classifier”, an additional requirement of at least two jets in the event is applied. The signals cross section are set to one hundred times of the observed limit for the “ggF-classifier” and fifty times of the observed limit for the “VBF -classifier”. The  $ZZ$  backgrounds are scaled by the normalisation factors shown in Table 7.15. The lower panels show the ratio of data to prediction. Only statistical and experimental systematic uncertainties are included.



**Fig. 7.5** Significance improvements of the MVA-based over the cut-based categorization of the VBF (ggF) category for VBF (ggF) signal samples from 300 to 2000 GeV for seven different cuts on the VBF (ggF) output score. The optimal cut of 0.8 (0.5) for VBF (ggF) score is chosen as the solid line, while other alternative cuts are plotted with dashed lines. For VBF category, results at 2000 GeV for cuts of 0.8 and 0.9 are missing due to a lack of background events passing this tight selection.

**Table 7.9 Input features used in the “ggF-classifier” for the  $\ell\ell\ell'\ell'$  analysis. The RNN stands for the recurrent neural network and MLP for the multilayer perceptron.**

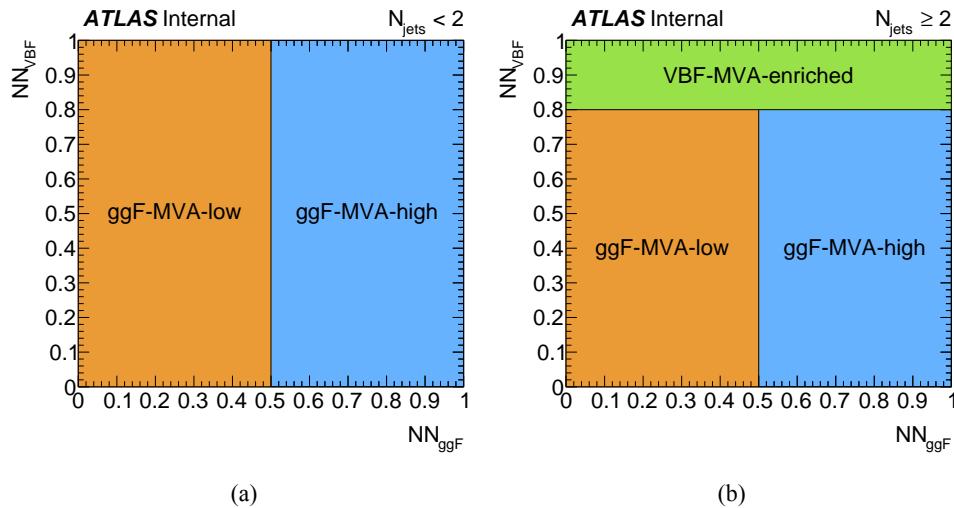
Model	Inputs	Description
RNN	$p_T^{\ell0,\ell1,\ell2,\ell3}$	transverse momenta of the four leptons
	$\eta^{\ell0,\ell1,\ell2,\ell3}$	pseudorapidity of the four leptons
MLP	$m_{4\ell}$	invariant mass of the four lepton system
	$p_T^{4\ell}$	transverse momentum of the four lepton system
	$\eta^{4\ell}$	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
	$\Phi$	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
	$\eta^{j0}$	pseudorapidity of the leading jet

- 2125 • ggF-MVA-high categories: ( $N_{\text{jets}} \geq 2 \ \&\& \ DNN_{\text{VBF}} \leq 0.8 \ \&\& \ DNN_{\text{ggF}} >$   
 2126       $0.5) || (N_{\text{jets}} < 2 \ \&\& \ DNN_{\text{ggF}} > 0.5);$   
 2127 • ggF-MVA-low category: All remaining events that fail VBF and ggF cuts men-  
 2128      tioned above.

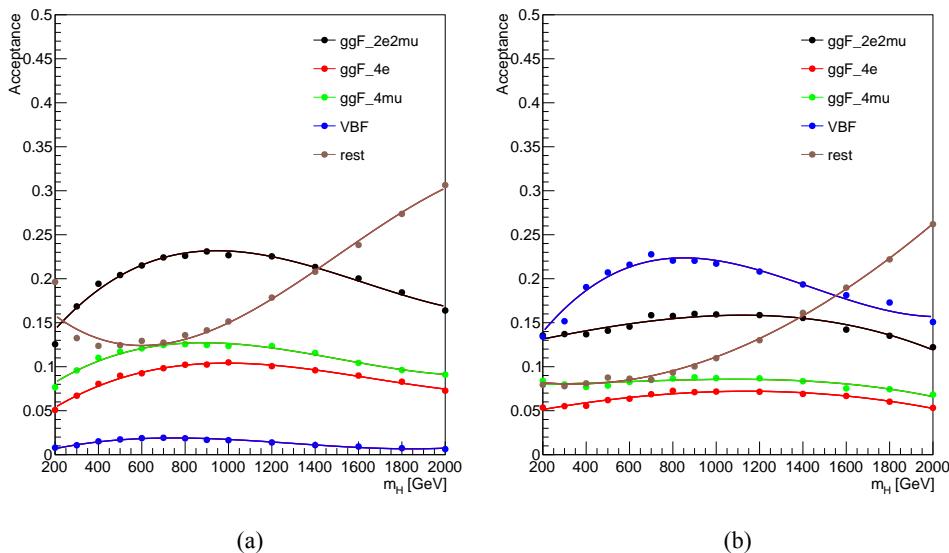
### 2129      7.3.4 Signal acceptance

2130      The signal acceptance is defined as the ratio of events passing all analysis selection in  
 2131      each category to the total number of simulated events in whole phase space. In denom-  
 2132      inator, the events with  $\tau$  final states are not taken into account. And the contribution of  
 2133       $\tau$ -lepton decay to electrons and muons final states is found to be negligible.

2134      Figure 7.7 and 7.8 show the acceptance of NWA signals in DNN- and Cut- based  
 2135      categorization, estimated by merging the three signal MC campaigns, mc16a, mc16d and  
 2136      mc16e. A 3-rd order polynomial fit is applied for each category.



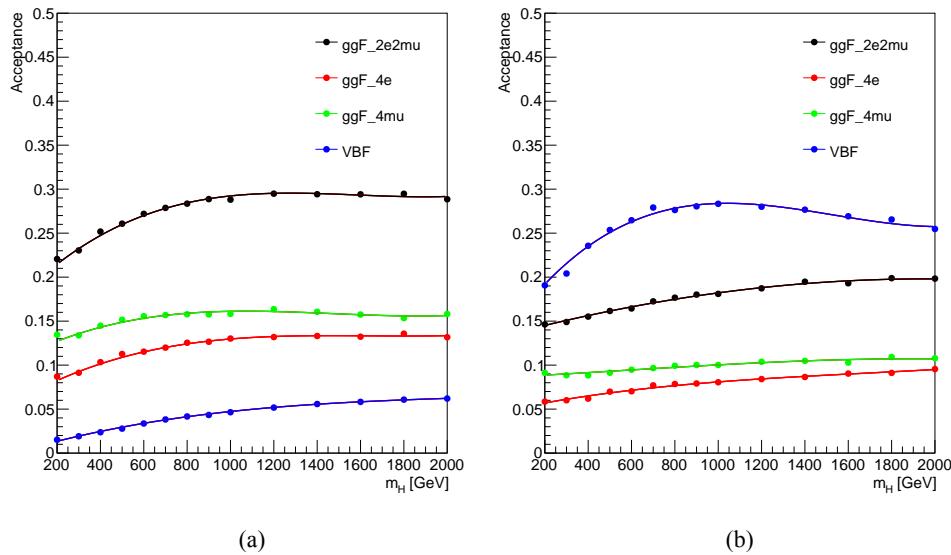
**Fig. 7.6** Illustration of the MVA-based VBF and ggF event classification for events with (a)  $N_{\text{jets}} < 2$  and (b)  $N_{\text{jets}} \geq 2$ .



**Fig. 7.7** NWA acceptance as a function of  $m_H$  for the MVA-based categorization for the samples of (a) ggF production; (b) VBF production.

## 7.4 Background estimation

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



**Fig. 7.8 NWA acceptance as a function of  $m_H$  for the Cut-based categorization for the samples of (a) ggF production mode; (b) VBF production mode.**

2144 In addition to irreducible backgrounds, events from  $Z+jets$  and  $t\bar{t}$  processes, represent  
 2145 as reducible backgrounds, contribute at a few percent level and can be measured using data  
 2146 driven method that will be described briefly later. Additional background called ‘Others’,  
 2147 including  $ttV$  and triple-V (VVV) processes, has tiny contribution and is estimated from  
 2148 MC simulation directly.

#### 2149 7.4.1 Irreducible backgrounds

2150 The Irreducible backgrounds have events with four prompt leptons. The normalization  
 2151 of two dominant backgrounds  $q\bar{q} \rightarrow ZZ$  and  $gg \rightarrow ZZ$  are taken from data by statistical  
 2152 fit, and the normalization of small  $q\bar{q} \rightarrow ZZ$  EW background is measured directly from  
 2153 MC simulation.

2154 The  $m_{4\ell}$  shapes of all three background components are taken from MC samples and  
 2155 then parameterized by an empirical function for each of them in each category respec-  
 2156 tively. Details of background modellings are illustrated as below:

2157 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,

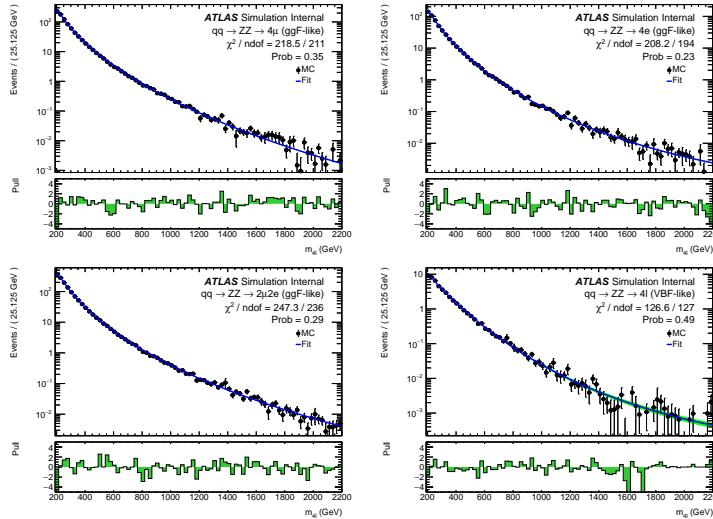
$$f_1(x) = \left( \frac{x - a_4}{a_3} \right)^{a_1-1} \left( 1 + \frac{x - a_4}{a_3} \right)^{-a_1-a_2},$$

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

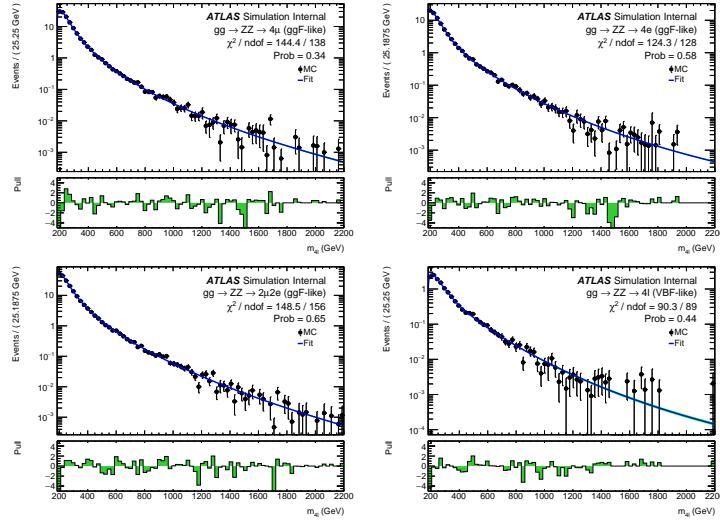
$$C_0 = \frac{f_2(m_0)}{f_1(m_0)}.$$

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.

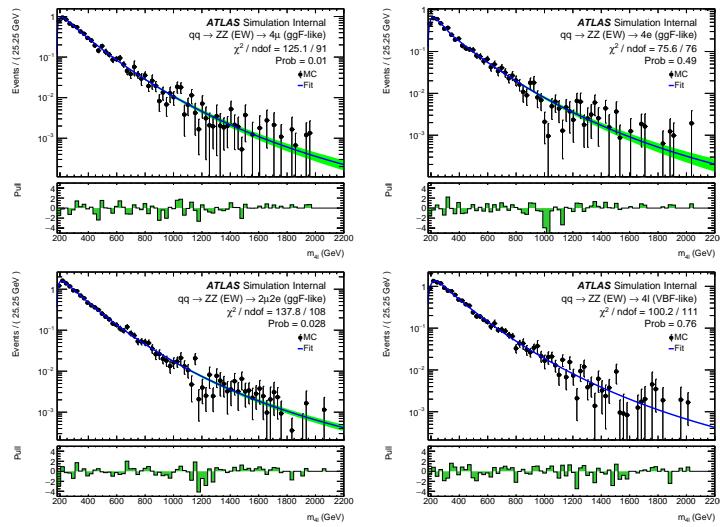
Figure 7.9 to 7.11 shows the fitting results of  $q\bar{q} \rightarrow ZZ$ ,  $gg \rightarrow ZZ$ ,  $q\bar{q} \rightarrow ZZ$  EW backgrounds in four cut-based categories (ggF-CBA-enriched- $2e2\mu$ , ggF-CBA-enriched- $4e$ , ggF-CBA-enriched- $4\mu$  and VBF-CBA-enriched). Figure 7.12 to 7.14 shows the fitting results of those backgrounds in five MVA-based categories (ggF-MVA-high- $2e2\mu$ , ggF-MVA-high- $4e$ , ggF-MVA-high- $4\mu$ , ggF-MVA-low and VBF-MVA-enriched).



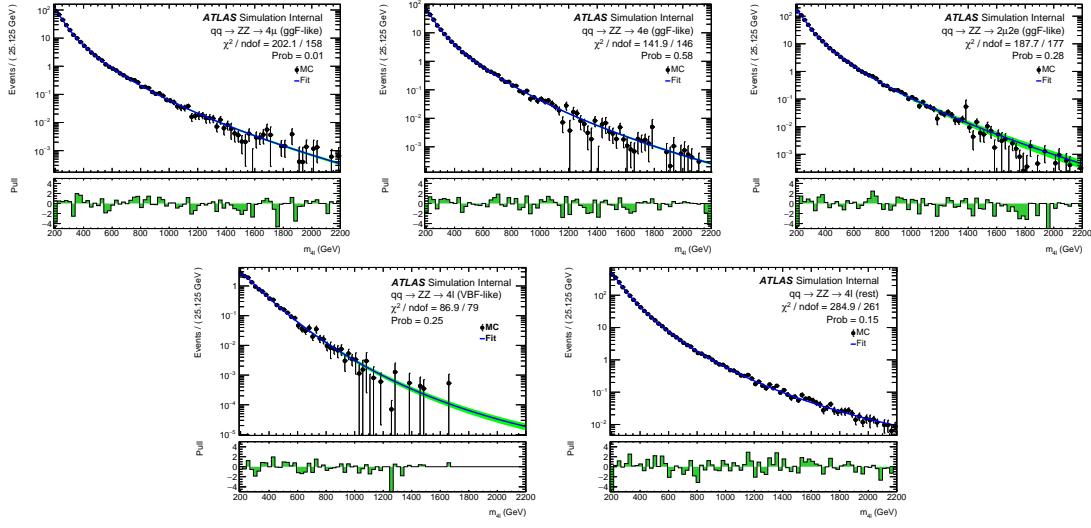
**Fig. 7.9 Distributions of the  $m_{4\ell}$  invariant mass fit projections of the  $q\bar{q} \rightarrow ZZ$  background samples for the  $4\mu$ ,  $4e$  and  $2\mu2e$  final states in the ggF-CBA-enriched category, and the  $4\ell$  inclusive VBF-CBA-enriched category. Cut-based categorization is used.**



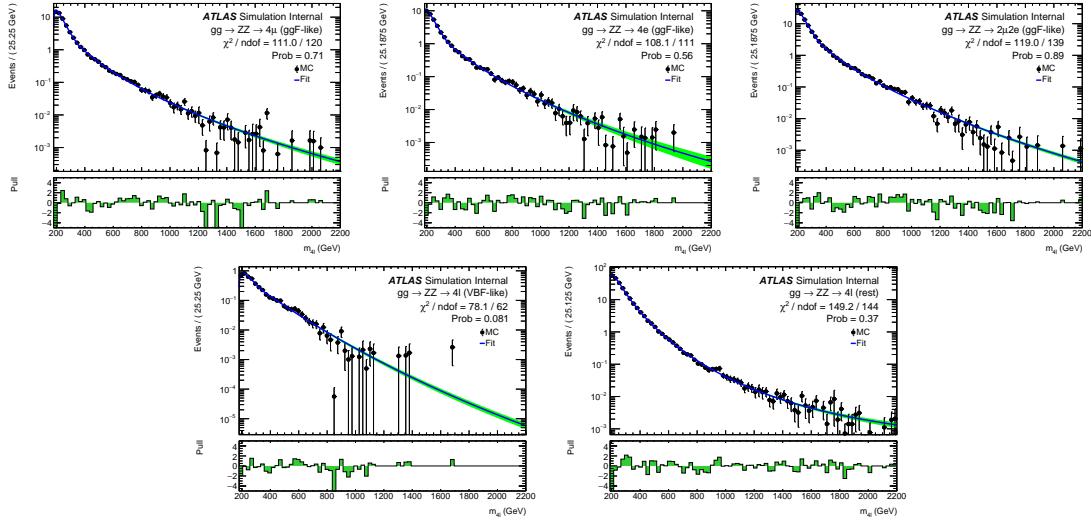
**Fig. 7.10** Distributions of the  $m_{4\ell}$  invariant mass fit projections of the  $gg \rightarrow ZZ$  background samples for the  $4\mu$ ,  $4e$  and  $2\mu2e$  final states in the ggF-CBA-enriched category, and the  $4\ell$  inclusive VBF-CBA-enriched category. Cut-based categorization is used.



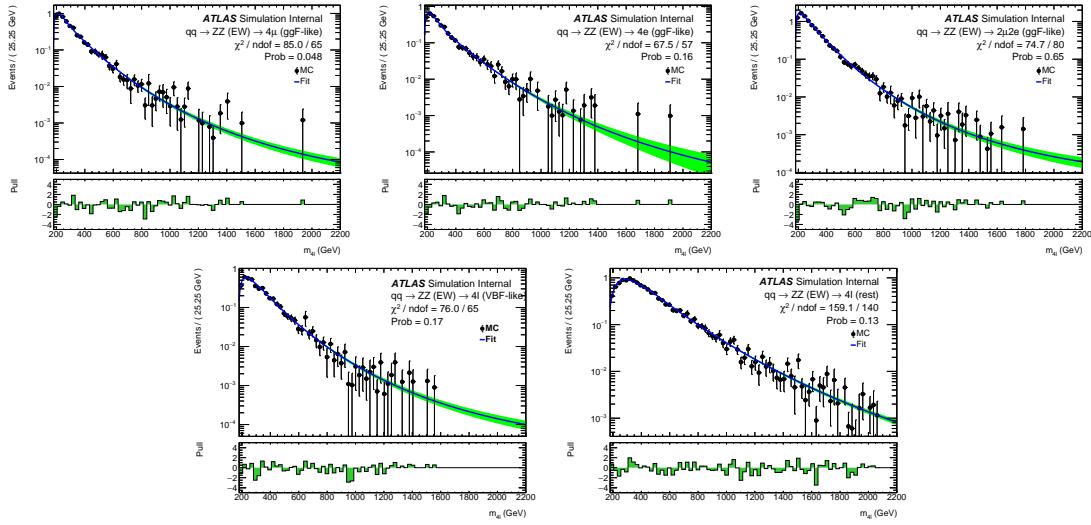
**Fig. 7.11** Distributions of the  $m_{4\ell}$  invariant mass fit projections of the  $q\bar{q} \rightarrow ZZ$  (EW) background samples for the  $4\mu$ ,  $4e$  and  $2\mu2e$  final states in the ggF-CBA-enriched category, and the  $4\ell$  inclusive VBF-CBA-enriched category. Cut-based categorization is used.



**Fig. 7.12** Distributions of the  $m_{4\ell}$  invariant mass fit projections of the  $q\bar{q} \rightarrow ZZ$  background samples for the  $4\mu$ ,  $4e$  and  $2\mu2e$  final states in the ggF-MVA-high category, the  $4\ell$  inclusive ggF-MVA-low category and VBF-MVA-enriched category. DNN-based categorization is used.



**Fig. 7.13** Distributions of the  $m_{4\ell}$  invariant mass fit projections of the  $gg \rightarrow ZZ$  background samples for the  $4\mu$ ,  $4e$  and  $2\mu2e$  final states in the ggF-MVA-high category, the  $4\ell$  inclusive ggF-MVA-low category and VBF-MVA-enriched category. DNN-based categorization is used.



**Fig. 7.14 Distributions of the  $m_{4\ell}$  invariant mass fit projections of the  $q\bar{q} \rightarrow ZZ$  (EW) background samples for the  $4\mu$ ,  $4e$  and  $2\mu2e$  final states in the ggF-MVA-high category, the  $4\ell$  inclusive ggF-MVA-low category and VBF-MVA-enriched category. DNN-based categorization is used.**

#### 2172      7.4.2 Reducible backgrounds

2173      Similar as section 6.4, the reducible backgrounds include  $Z+jets$  (consists of both  
 2174      heavy- and light-flavour jets), top quark pair, and  $WZ$  production, which contain fake  
 2175      and non-isolated leptons. The simulations are not very robust in terms of the selection  
 2176      efficiencies. Thus, the data-driven method is applied to estimate the normalization of  
 2177      those processes in different control regions (CRs). The estimations in this analysis are  
 2178      performed separately for  $\ell\ell + \mu\mu$  and  $\ell\ell + ee$  final states, with slightly different ap-  
 2179      proaches for “muon” and “electron” backgrounds.

2180      The “electron” backgrounds mostly come from process of a  $Z$  boson with light-  
 2181      flavour jets ( $Z+LF$ ) misidentified as electrons. The large contribution of “muon” back-  
 2182      grounds come from heavy-flavour jets produced in association with a  $Z$  boson ( $Z+HF$ )  
 2183      or in the decays of top quark. The estimations are done following the common H4l stud-  
 2184      ies without a specific  $m_{4\ell}$  range requirement<sup>[106]</sup>, and then the corresponding fraction of  
 2185      event yield in  $m_{4\ell} > 200$  GeV is calculated from MC simulation.

##### 2186      $\ell\ell + \mu\mu$ final states

2187      The normalizations of “muon” backgrounds are extracted from simultaneous fits of  
 2188      the leading lepton pair’s invariant mass ( $m_{12}$ ) in four orthogonal CRs:

- 2189      • **Inverted  $d_0$  CR:** this CR is formed by inverting the  $d_0$  selection for at least one  
 2190      lepton in subleading lepton pair while the leptons in leading pair are required to  
 2191      pass all standard selection. This CR enhances  $Z+HF$  and  $t\bar{t}$  as leptons from heavy-

flavour hadronic decays are characterised by large  $d_0$ .

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

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

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

### $\ell\ell + ee$ final states

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

The  $m_{4\ell}$  shapes of reducible backgrounds are obtained from MC simulation in signal region, and then smoothed by an one-dimensional kernel estimation, which models the input data as a superposition of Gaussian kernels, one for each data point with contributing  $1/N$  to total integral  $N$ <sup>[107]</sup>. The difference from using different smoothing strength ( $\rho$ ) 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

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

### 7.5.1 Modelling of narrow-width signal

For narrow-width (NWA) signal, the  $m_{4\ell}$  width is totally determined by detector resolution, which is modelled by the sum of a Crystal Ball ( $C$ ) function<sup>[108-109]</sup> and a Gaussian ( $G$ ) function:

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

The two functions share the same central value  $\mu$ , while the resolution parameters,  $\sigma_C$  and  $\sigma_G$ , are different. In the Crystal Ball function, the parameters  $\alpha_C$  and  $n_C$  model the shape of non-Gaussian tail, and the fraction parameter  $f_C$  is used to ensure the relative normalization between two functions.

The parameters are obtained by fitting to signal MC simulations combining the mc16a, mc16d and mc16e campaigns for each category at each mass points from 200 GeV to 2000 GeV respectively, and the shape of ggF and VBF signals are found to be similar. Figure 7.15 shows the  $m_{4\ell}$  distribution and fitted curves for ggF production at mass from 200 GeV to 2000 GeV in  $2e2\mu$  channel as examples.

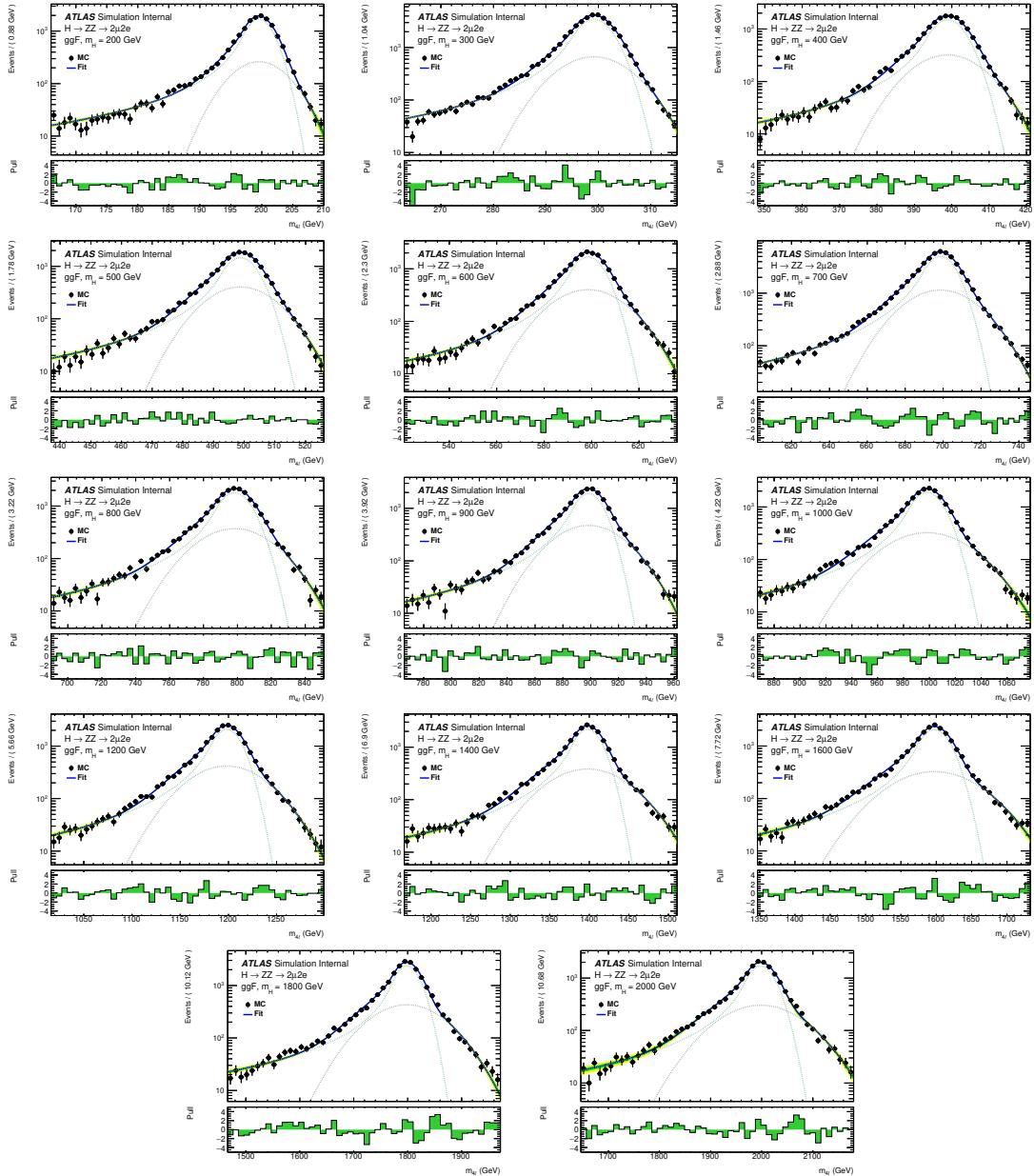
Then the  $C + G$  parameters are fitted with a polynomial function as the function of generated mass points ( $m_H$ ), as an example shown in figure 7.16 for  $2e2\mu$  channel. The fitting quality can be measured by the Pearson's  $\chi^2$ , which is within 3 (2) for  $2e2\mu$  ( $4e$  and  $4\mu$ ) channel.

In addition, possible difference on the signal yield extracted from parameterization and MC simulation is studied. Figure 7.17 shows this difference by computing  $\frac{N_{\text{reco}} - N_{\text{fit}}}{N_{\text{fit}}}$ , where  $N_{\text{reco}}$  denotes the total number of reconstructed events observed from MC simulation at that mass point and  $N_{\text{fit}}$  depicts the number of events obtained from the fitted PDF. The differences are treated as an additional systematic uncertainty with the value of 2% (1%) for  $2e2\mu$  ( $4e$  and  $4\mu$ ) channel in statistical fit.

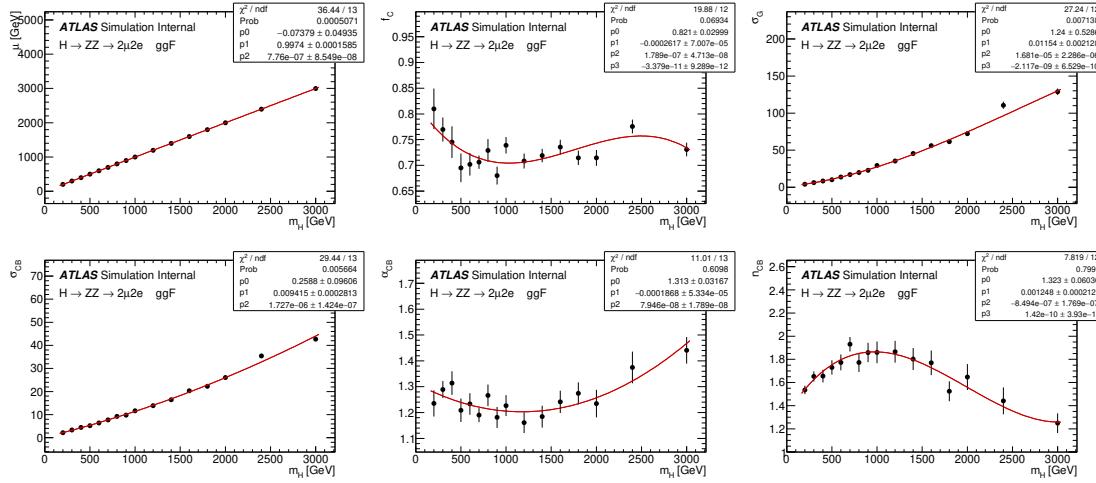
In summary, the final interpolated signal shapes for the ggF production mode are shown together in figure 7.18 for mass points with step of 100 GeV from 200 GeV to 3000 GeV.

### 7.5.2 Modelling of large-width signal

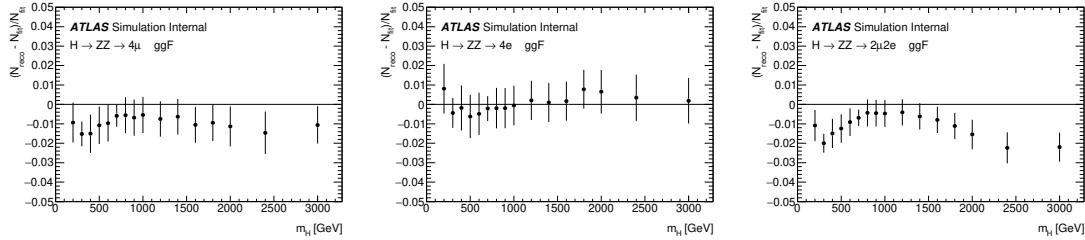
The  $m_{4\ell}$  shape of heavy Higgs model in large-width (LWA) hypothesis can be described by a convolution of a truth distribution and a resolution from detector effect. The



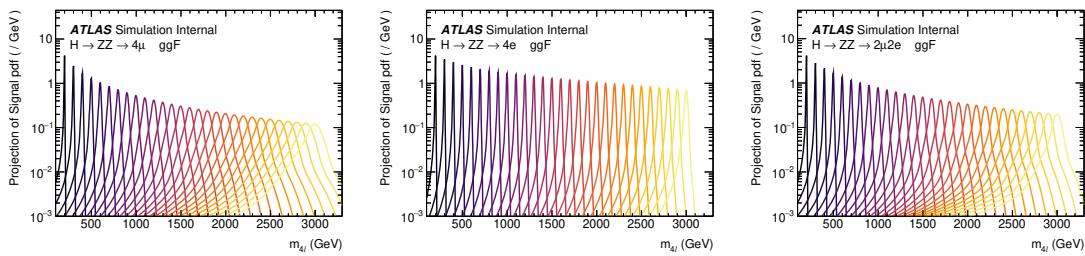
**Fig. 7.15** Distributions of the  $m_{2\mu 2e}$  and fit projection for signal samples between 200 to 3000 GeV for ggF production mode. Three MC campaigns, mc16a, mc16d and mc16e, are combined. The lower panel in each plot shows the pull distribution.



**Fig. 7.16** Polynomial fits of the parameters  $\mu$ ,  $f_C$ ,  $\sigma_G$ ,  $\sigma_C$ ,  $n_c$  and  $\alpha_c$  for the signal  $\mathcal{C} + \mathcal{G}$  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. 7.17** The difference between MC simulation and parameterization of  $4\mu$  (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. 7.18** The final signal shapes for the ggF production mode, interpolated from the polynomial fit parameters.

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

The differential parton cross section for the heavy Higgs model can be written as<sup>[110]</sup>:

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

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

Using the definition of a partial width,

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

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<sup>[110-111]</sup>:

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

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

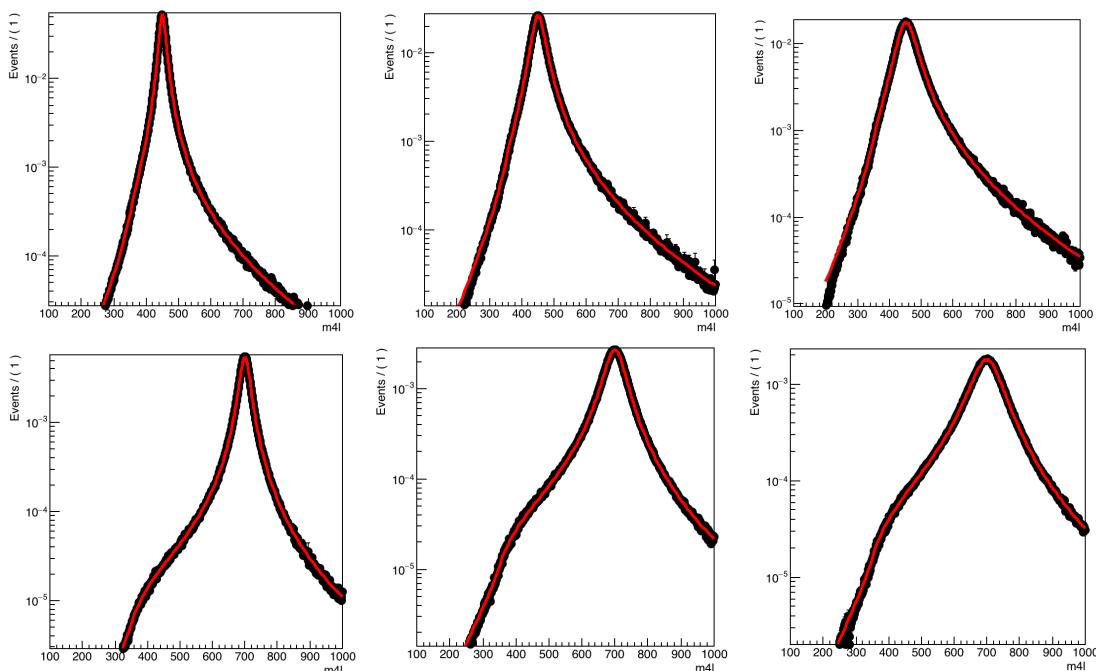
$$\begin{aligned} \Gamma_{H \rightarrow gg}(s) &= C \cdot s^{\frac{3}{2}} \cdot |A_t(\tau_t)|^2 \\ A_t(\tau) &= 2 \frac{\tau + (\tau - 1)f(\tau)}{\tau^2} \\ \tau_t &= \frac{s}{4m_t^2} \\ f(\tau) &= \begin{cases} \arcsin^2(\sqrt{\tau}), & \tau \leq 1 \\ -\frac{1}{4} \left[ \log \frac{1+\sqrt{1-\tau^{-1}}}{1-\sqrt{1-\tau^{-1}}}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

where  $m_f$  stands for the mass of a fermion  $f$ , and  $\Gamma_H$  denotes an assumed total width 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<sup>[112]</sup>. 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 7.19.



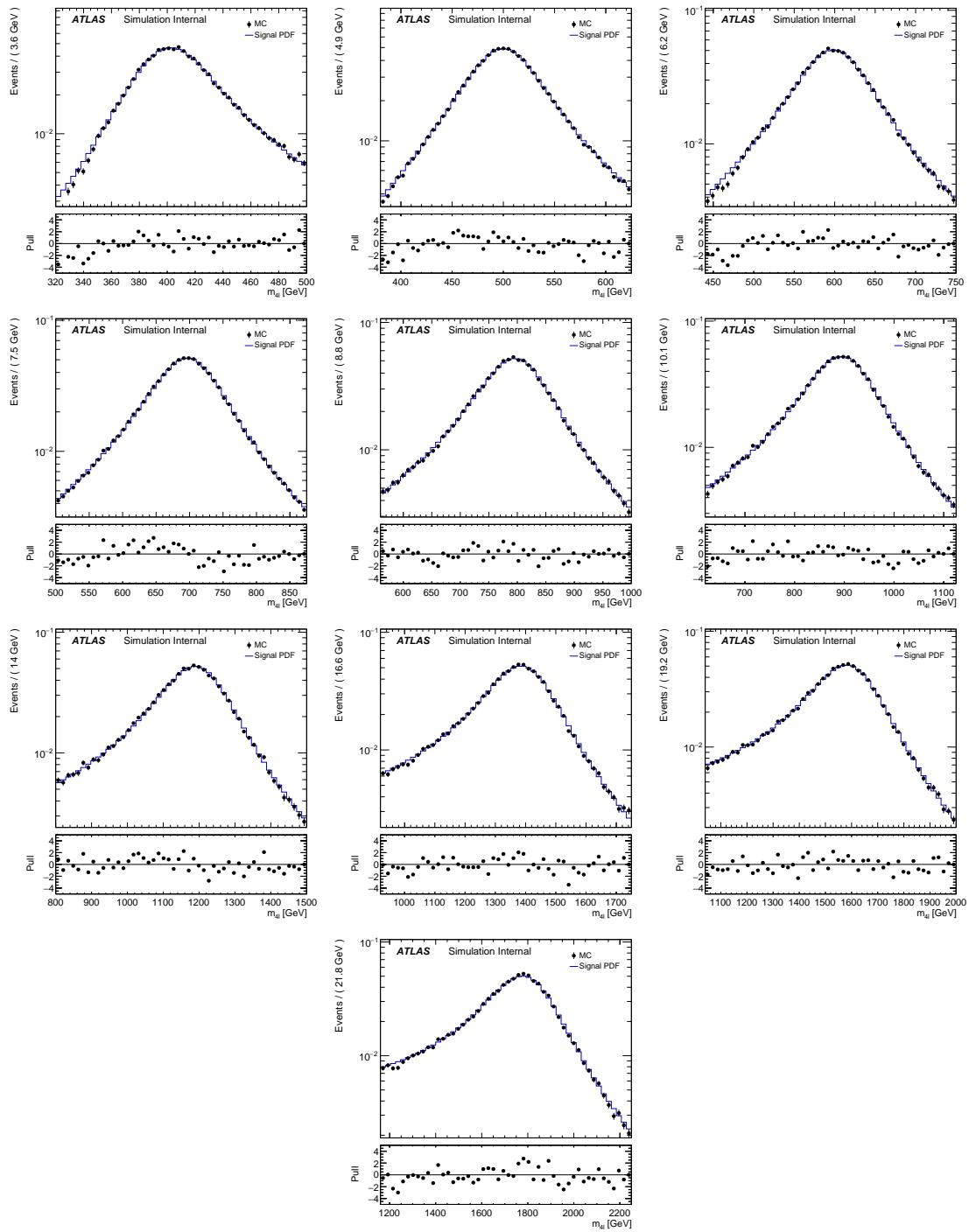
**Fig. 7.19 Comparison of the analytical shape to a truth  $m_{4\ell}$  distribution of gg2VV MC samples for  $m_H = 450$  GeV (top), 700 GeV (bottom) and width equal to 5% (left), 10% (middle), 15% (right) of the mass.**

The reconstruction level signal shape can then be modelled by the analytical truth shape convoluted with detector effects modelled in section 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 7.20, the shapes are well compatible between each other. This modelling is not valid for lower masses due to the rapid change of detector resolution.

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

- 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:

$$\begin{aligned}
 \sigma_{gg \rightarrow (X) \rightarrow ZZ}(s) &= \frac{1}{2s} \int d\Omega |A_h(s, \Omega) + A_H(s, \Omega) + A_B(s, \Omega)|^2 \\
 &= \frac{1}{2s} \int d\Omega \left( |A_h(s, \Omega)|^2 + |A_H(s, \Omega)|^2 + |A_B(s, \Omega)|^2 \right) + \\
 &\quad + \frac{1}{s} \int d\Omega \left( \text{Re}[A_h(s, \Omega) \cdot A_B^*(s, \Omega)] \right. \\
 &\quad \left. + \text{Re}[A_H(s, \Omega) \cdot A_B^*(s, \Omega)] + \text{Re}[A_H(s, \Omega) \cdot A_h^*(s, \Omega)] \right) \\
 &\quad + \frac{1}{s} \text{Re} \left[ \frac{1}{s - s_H} \int d\Omega \cdot A_H^P(s, \Omega) \cdot A_H^D(s, \Omega) \cdot A_B^*(s, \Omega) \right] \\
 &\quad + \frac{1}{s} \int d\Omega \cdot \text{Re} \left[ A_H^P(s, \Omega) \cdot \frac{1}{s - s_H} \cdot A_H^D(s, \Omega) \cdot A_h^{P*}(s, \Omega) \cdot \frac{1}{(s - s_h)^*} \cdot A_h^{D*}(s, \Omega) \right]
 \end{aligned} \tag{7.11}$$

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 has been written down in equation 7.11. By assuming that this function has a smooth behaviour, it can be replaced with complex polynomial:

$$\int d\Omega \cdot A_H^P(s, \Omega) \cdot A_H^D(s, \Omega) \cdot A_B^*(s, \Omega) \approx (a_0 + a_1 \cdot \sqrt{s} + \dots) + i \cdot (b_0 + b_1 \cdot \sqrt{s} + \dots) \tag{7.12}$$

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

2302 into a hadron cross section as a function of  $m_{4\ell}$ :

$$\sigma_{pp}(m_{4\ell}) = \mathcal{L}_{gg} \cdot \frac{1}{m_{4\ell}} \cdot \text{Re} \left[ \frac{1}{s - s_H} \cdot ((a_0 + a_1 \cdot m_{4\ell} + \dots) + i \cdot (b_0 + b_1 \cdot m_{4\ell} + \dots)) \right] \quad (7.13)$$

2303 where the propagators are shown in equation 7.7.

2304 Figure 7.21 shows the distributions of interference function obtained by simultaneous  
2305 fitting to  $m_{4\ell}$  shape from truth level H-B interference simulation at different mass in  $2e2\mu$   
2306 channel as an example.

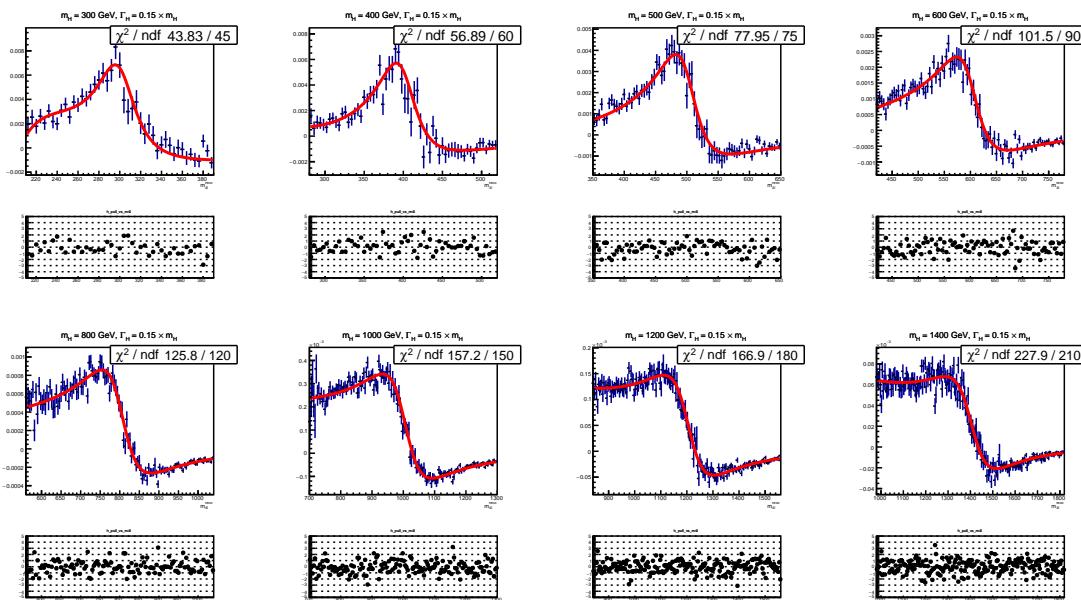


Fig. 7.21 The interference (H-B) model fitted to the truth  $m_{4\ell}$  MC distribution after signal region selection for  $2\mu2e$  channel.

## 2307 2. Interference between heavy Higgs and SM Higgs

2308 The parton cross section of this interference term has been written down in equa-  
2309 tion 7.11. By assuming the production and decay amplitudes are the same for heavy  
2310 Higgs boson and SM Higgs boson, the cross section function can be simplified to:

$$\sigma_{gg}(s) = \frac{1}{s} \int d\Omega \cdot \text{Re} \left[ \frac{1}{s - s_H} \cdot \frac{1}{(s - s_h)^*} \right] \cdot |A_{gg \rightarrow H}(s, \Omega)|^2 |A_{H \rightarrow ZZ}(s, \Omega)|^2 \quad (7.14)$$

2311 Taking into account Equation 7.5:

$$\sigma_{gg}(s) = 4 \cdot \text{Re} \left[ \frac{1}{s - s_H} \cdot \frac{1}{(s - s_h)^*} \right] \cdot \Gamma_{H \rightarrow gg}(s) \cdot \Gamma_{H \rightarrow ZZ}(s) \quad (7.15)$$

2312 where the propagators are described in equation 7.7, and the partial widths are de-  
2313 scribed in equations 7.8 and 7.9.

2314     Same as previous procedure, the parton cross section can be transformed to a hadron  
 2315     cross section as a function of  $m_{4\ell}$ :

$$\sigma_{pp}(m_{4\ell}) = 4 \cdot m_{4\ell} \cdot \mathcal{L}_{gg} \cdot \text{Re} \left[ \frac{1}{s - s_H} \cdot \frac{1}{(s - s_h)^*} \right] \cdot \Gamma_{H \rightarrow gg}(m_{4\ell}) \cdot \Gamma_{H \rightarrow ZZ}(m_{4\ell}) \quad (7.16)$$

2316     The modelling procedure of interference is the same as the way for large-width signal  
 2317     described in section 7.5.2. The truth line shape is measured as analytical function from  
 2318     equation 7.16, and then convolute with detector effect from NWA parameterization to get  
 2319     the reconstruction level shape.

2320     For LWA signal model, these two interferences are carefully token into account, and  
 2321     the integration of the pure LWA signal with the interferences is used for further studies.  
 2322     Figure 7.22 shows the signal model for large-width scenario at mass points of 400 GeV,  
 2323     600 GeV, 800 GeV, for three different signal widths: 5%, 10%, 15%, with and without  
 2324     interference. Additionally, the contribution of the interference between heavy Higgs and  
 2325     SM Higgs (H-h) is shown together with the one between heavy Higgs and SM  $gg \rightarrow$   
 2326      $ZZ$  background (H-b). One can see the interference effect on signal shape becomes less  
 2327     important when going to higher mass.

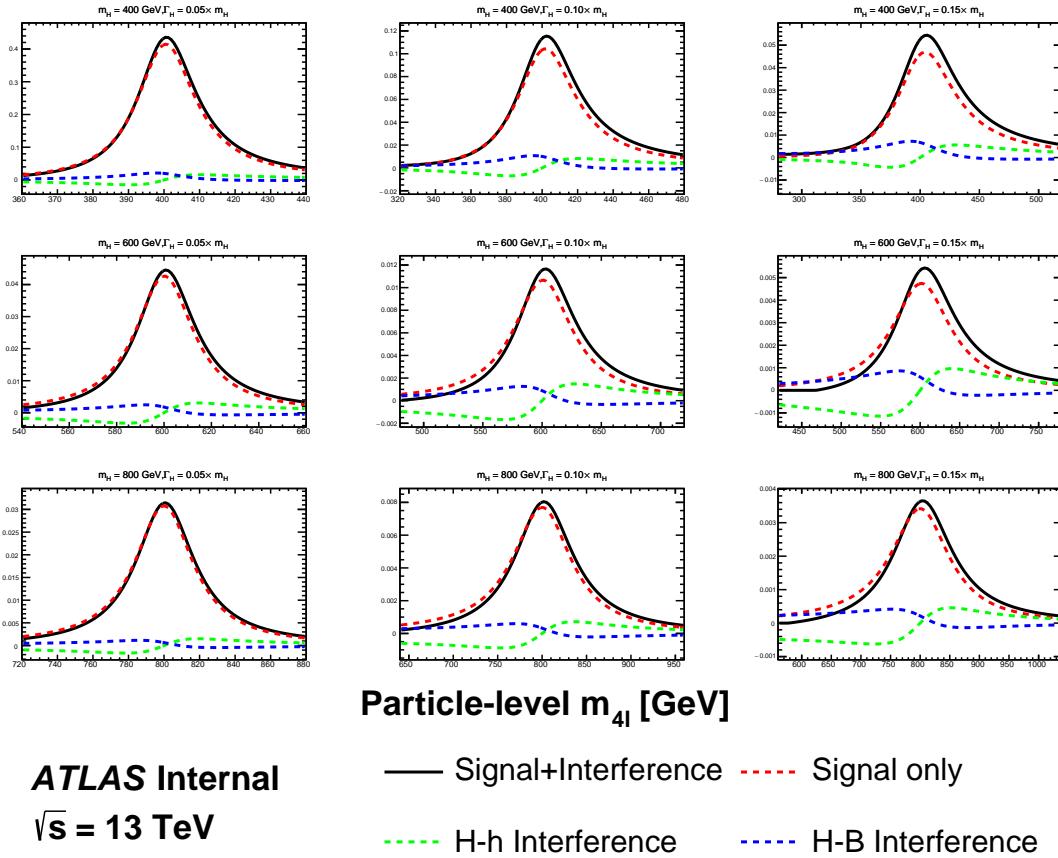
#### 2328     7.5.4 Modelling of spin-2 RS Graviton signal

2329     The search for Randall-Sundrum (RS) graviton is performed in mass region between  
 2330     600 to 2000 GeV. The width of resonance is determined by the  $k/\overline{M}_{\text{Planck}}$ , which, as  
 2331     mentioned in section 7.2.3, is set to be 1. In this configuration, the width of signal is  
 2332     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 for truth-level shape, for graviton, it's modelled as the product of a relativistic Breit-Wigner (RBW) term, a term corresponding to the squared matrix element of the production process and a parton luminosity term  $\mathcal{L}$  as given in [113]. So the truth lineshape of  $m_{4\ell}$  is token from:

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

2333     The truth-level signal model is extracted by fitting to MC simulation at truth-level with  
 2334     the mass  $m_G$  and width  $\Gamma$  parameters floating at each mass points respectively. And then



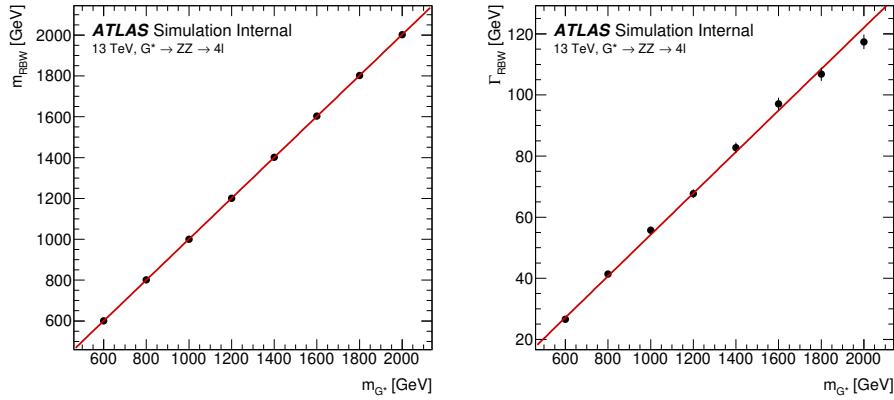
**Fig. 7.22** The signal modelling for the large-width scenario at  $m_H$  of 400 GeV (top), 600 GeV (middle) and 800 GeV (bottom), as well as three different signal width: 5% (left), 10% (middle) and 15% (right). The contribution of the interference between heavy Higgs and SM Higgs (H-h) is shown together with the one between heavy Higgs and SM  $gg \rightarrow ZZ$  background (H-b).

2335 the two parameters are parameterized as the function of  $m_H$  by a linear fit as shown in  
 2336 figure 7.23.

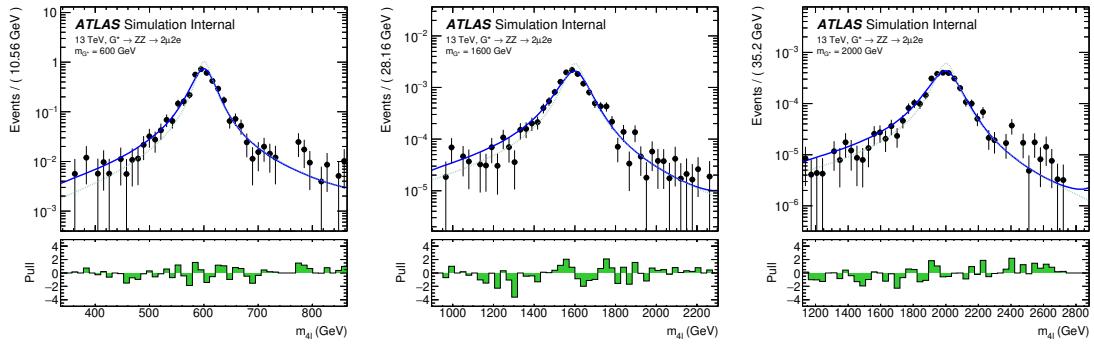
2337 The final signal model is obtained by convolving the truth-level lineshape with the  
 2338 detector resolution function. To verify the result, figure 7.24 compares the  $m_{4\ell}$  lineshape  
 2339 from parameterization with the one observed from reconstructed-level MC simulation in  
 2340  $2e2\mu$  channel at masses of 600 GeV, 1600 GeV and 2000 GeV as examples.

## 2341 7.6 Systematic uncertainties

2342 This section describes the sources and value of theoretical and experimental system-  
 2343 atic uncertainties considered in this analysis. In addition, as mentioned in previous sec-  
 2344 tions, the uncertainties of irreducible background modelling, reducible background shape  
 2345 smoothing procedure and signal yield difference between simulation and parameterization



**Fig. 7.23 Fitted parameters of the graviton RBW,  $m_{\text{RBW}}$  and  $\Gamma_{\text{RBW}}$ , as a function of the graviton resonance mass,  $m_G$ .**



**Fig. 7.24 Reconstructed  $m_{4\ell}$  distributions in the  $2\mu 2e$  channel with the final signal model superimposed for each RS graviton signal sample at masses of 600 GeV, 1600 GeV and 2000 GeV. The lower panel in each plot shows the pull distribution. The dashed green lines show the truth-level graviton signal models for reference.**

2346 are all taken into account.

### 2347 7.6.1 Theoretical uncertainties

2348 The theoretical modelling uncertainties include the PDF variations, missing QCD  
2349 higher-order corrections via the variations of factorisation and renormalization scales, and  
2350 the parton showering uncertainties.

#### 2351 1. Theoretical uncertainties for signal

2352 The PDF, QCD scale and parton showering uncertainties affecting the acceptance dif-  
2353 ference originating from analysis selection for signal are taken into account in different  
2354 categories. The acceptance uncertainties are calculated on the acceptance factor which  
2355 extrapolates from the fiducial space to the full phase space by a simple ratio:

$$A = \frac{N_{\text{fiducial}}}{N_{\text{total}}} \quad (7.17)$$

2356 For PDF uncertainties, the standard derivations of 100 PDF replicas of NNPDF3.0  
 2357 NNLO, as well as comparison to two external PDF sets: MMHT2014 NNLO, CT14  
 2358 NNLO are considered. For missing QCD higher-order corrections, the effects are studied  
 2359 with truth events by comparing weights corresponding to variations of the renormalization  
 2360 and factorization scale factors, up and down by a factor of two, and the envelop of dif-  
 2361 ferent variations is used. The parton showering uncertainties are estimated by comparing  
 2362 events with different setting via Pythia8.

2363 Systematic uncertainties are studied for both cut- and MVA- based event catego-  
 2364 rizations, in two different categories: the inclusive ggF-CBA-enriched and VBF-CBA-  
 2365 enriched category for cut-based analysis and in three different categories: inclusive ggF-  
 2366 MVA-high, ggF-MVA-low and VBF-MVA-enriched category for MVA-based one. This  
 2367 section shows the MVA-based results as an example.

2368 Table 7.10 and 7.11 show the theoretical uncertainties mentioned above for ggF and  
 2369 VBF signal respectively in MVA-based categorization.

**Table 7.10 Summary of acceptance uncertainties of PDF, QCD scale and parton shower vari-  
 ations for ggF production. The MVA-based categorization is used.**

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

**Table 7.11 Summary of acceptance uncertainties of PDF, QCD scale and parton shower vari-  
 ations for VBF production. The MVA-based categorization is used.**

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

2370 **2. Theoretical uncertainties for SM background processes**

2371 The theoretical uncertainties of irreducible  $ZZ$  backgrounds are considered in terms  
 2372 of both the variations of shape of  $m_{4\ell}$  distributions and the acceptance originating from  
 2373 the event selection.

2374 The PDF and QCD scale uncertainties are considered by using the same method as  
 2375 described for signal. The parton showering uncertainties for those Sherpa samples are  
 2376 evaluated by varying the resummation scale by a factor of 2, changing the CKKW setting  
 2377 and using different showering option, following the PMG recommendation in ref.<sup>[114]</sup>,  
 2378 and the quadratic sum between the uncertainties in different kinds of showering option

2379 is taken as final result of uncertainties. Moreover, the shape uncertainty associated with  
 2380 electroweak higher-order correction for  $q\bar{q} \rightarrow ZZ$  process is also taken into account.

2381 Same as for signals, these theoretical uncertainties for irreducible backgrounds are  
 2382 studied for both cut- and MVA- based event categorizations. The value of shape uncer-  
 2383 tainties vary from less than 1% at low mass region to 50% at high mass tail due to large  
 2384 statistic fluctuation. As for the acceptance uncertainties, the values vary from about 1%  
 2385 for PDF variations to 40% for parton showering variations. The VBF category has relative  
 2386 larger uncertainties.

2387 Table 7.12 summarizes the acceptance uncertainties of PDF, QCD scale, and parton  
 2388 showering variations for the dominant background:  $q\bar{q} \rightarrow ZZ$ .

**Table 7.12 Summary of acceptance uncertainties of PDF, scale, and parton showering varia-  
 tions for QCD  $q\bar{q} \rightarrow ZZ$  background. The MVA-based categorization is used.**

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

## 2389 7.6.2 Experimental systematics

2390 The signal and background predictions used in this analysis are also affected by  
 2391 various sources of experimental systematic uncertainties. Similar as described in sec-  
 2392 tion 6.5.2, the dominant experimental uncertainties in this analysis come from the en-  
 2393 ergy/momentum scales and reconstruction and identification efficiencies of the leptons  
 2394 and jets, as well as the luminosity uncertainty. The systematic uncertainties are calculated  
 2395 using the recommendations from the Combined Performance (CP) groups of ATLAS ex-  
 2396 periment. Table 7.13 summarizes the experimental systematics considered in this analysis  
 2397 that affect either the normalization of total event yield or the shape of  $m_{4\ell}$  distribution.  
 2398 The impact of those systematics in statistical fit are studied in section 7.7.

## 2399 7.7 Results in $\ell\ell\ell'\ell'$ channel

2400 The statistical treatment in searching for heavy resonances in  $ZZ \rightarrow \ell\ell\ell'\ell'$  final  
 2401 state is described in this section. Results are presented in both cut- and MVA- based  
 2402 analysis.

### 2403 7.7.1 Statistical procedure

2404 The upper limits on heavy resonances are obtained using the unbinned profile likelihood fits.  $m_{4\ell}$  is the discriminant. The likelihood function is a product of a Poisson term  
 2405 representing the probability of observing  $n$  events and a weighted sum of both signal and  
 2406 background probability distribution functions (PDFs) evaluated at all observed events.  
 2407

$$L(x_1..x_n|\sigma_{ggF}, \sigma_{VBF}) = \text{Pois}(n|S_{ggF} + S_{VBF} + B) \left[ \prod_{i=1}^n \frac{S_{ggF}f_{ggF}(x_i) + S_{VBF}f_{VBF}(x_i) + Bf_B(x_i)}{S_{ggF} + S_{VBF} + B} \right] \quad (7.18)$$

2408 where  $f_X$ s are the probability distribution functions of signal and backgrounds modelled  
 2409 in section 7.5 and 7.4,  $S_X$  and  $B$  are the normalizations of signal and sum of backgrounds.

2410 The parameters of interest (POI) in the search is  $\sigma_{ggF}$  (and  $\sigma_{VBF}$  only for NWA signal),  
 2411 which is the cross section of signal model in ggF (and VBF) production mode. In the case  
 2412 of there are two POIs, when testing one POI, the other one is profiled along with other  
 2413 nuisance parameters (except left unconstrained) during the minimization. These POIs  
 2414 enter the likelihood inside the expected signal yields  $S_{ggF}$  and  $S_{VBF}$  as:

$$S_{ggF(VBF)} = \sigma_{ggF(VBF)} \times BR(S \rightarrow ZZ) \times A \times C \times \int \mathcal{L} \quad (7.19)$$

2415 where  $A \times C$  is the signal acceptance as parameterized in 7.3.4, and  $\int \mathcal{L} = 139 \text{ fb}^{-1}$  is the  
 2416 integrated luminosity of the dataset.

2417 The dependence of the expected number of signal and background events (normaliza-  
 2418 tions) and the shape of the PDFs on the systematic uncertainties measured in section 7.6  
 2419 is described by a set of nuisance parameters (NPs)  $\theta_i$ . The Gaussian constraints are ap-  
 2420 plied to those NPs. The constraints are implemented as additional ‘penalty’ terms added  
 2421 to the likelihood which increase the negative log-likelihood when any nuisance parameter  
 2422 is shifted from its nominal value. The final likelihood function  $L(\sigma_{ggF}, \sigma_{VBF}, m_H, \theta_i)$  is  
 2423 therefore a function of  $\sigma_{ggF}$ ,  $\sigma_{VBF}$ ,  $m_H$ , and  $\theta_i$ .

2424 Furthermore, the normalization of SM background  $pp \rightarrow ZZ$ , including both  $q\bar{q} \rightarrow$   
 2425  $ZZ$  and  $gg \rightarrow ZZ$ , is a free parameter ( $\mu_{ZZ}$ ) and profiled during the minimization.  
 2426 Floating  $ZZ$  normalization in fit takes the advantage of reducing the dependence on the-  
 2427 ory predictions and their associated uncertainties, especially given that the increased data  
 2428 luminosity would provide precise determination of the SM  $ZZ$  background rate.

2429 At the end, the upper limit on production cross-section  $\sigma_{ggF(VBF)}$  at a given heavy  
 2430 resonance model is obtained by setting the mass of signal  $m_H$  parameter as constant at the  
 2431 desired value, and maximising the likelihood function with respect to nuisance parame-  
 2432 ters. The  $CL_s^{[115]}$  method is used to obtain exclusion limits.

2433     7.7.2 For to likelihood function under background-only hypothesis for  
2434                   MVA-based analysis

2435     Both MVA- and cut-based analysis are studied by performing likelihood fit to the  
2436     (pseudo-) data under the background-only hypothesis and under different signal models.  
2437     Due to the same background estimation and modelling procedures, as well as the same  
2438     method of systematic measurements, this section only shows the results of background-  
2439     only fits for MVA-based analysis under the model of heavy Higgs resonance with narrow-  
2440     width as an example. The final results of interpretation in both MVA- and cut- based  
2441     analysis in all signal models described in section 7.5 will be measured in next section.

2442     First of all, table 7.14 summarized the expected and observed number of events for  
2443     region of  $m_{4\ell} > 200$  GeV together with their systematic uncertainties after background-  
2444     only fit. The post-fit  $m_{4\ell}$  spectrum in each category is shown in figure 7.25.

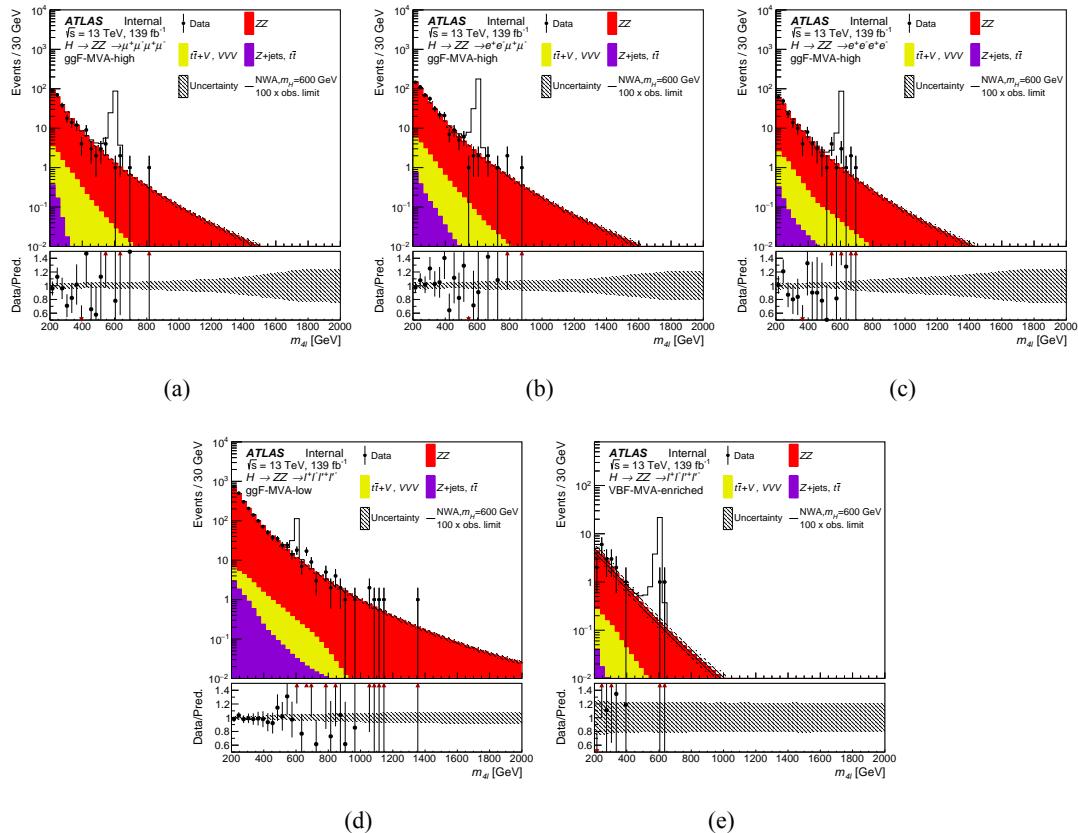
2445     To inspect the likelihood model, pulls and constraints as well as the correlation matrix  
2446     of NPs are studied by performing a background only fit. Figure 7.26 shows the pulls and  
2447     constraints when fitting to pseudo-data (top) and observed data (bottom). Figure 7.27  
2448     shows the correlation matrix, only for NPs with correlation between each others greater  
2449     than 0.1 when fitting to pseudo-data. The normalization of  $ZZ$  background is taken from  
2450     data for one category each, as shown in table 7.15.

2451     The impact of a systematic uncertainty on the result depends on the production mode  
2452     and the mass hypothesis. To check the impact of systematic uncertainties on expected sig-  
2453     nal sensitivity, a NP ranking study is performed using signal injected Asimov data with  
2454     the injected cross section close to 95% CLs upper limit at the masses of 400 GeV and  
2455     1000 GeV. The results are shown in table 7.16. For ggF production, at lower masses, the  
2456     systematic uncertainties of parton showering variation for signal, the luminosity uncer-  
2457     tainty, and the parametrization of signal acceptance dominate, while at higher masses, the  
2458     shape uncertainties from PDF variation for  $ZZ$  ( $q\bar{q} \rightarrow ZZ$  and  $gg \rightarrow ZZ$ ) background  
2459     become important, as also seen in VBF production mode. In addition for VBF, jet related  
2460     uncertainties become more important comparing to ggF production. Moreover, the domi-  
2461     nate uncertainties include the acceptance uncertainty from QCD scale variation for signal  
2462     and the luminosity uncertainty.

2463     7.7.3 Interpretations

2464       1. Spin-0 resonance with NWA

2465     In the absence of a specific model, the ratio of ggF and VBF production mode is un-  
2466     known for this additional heavy scalar. For this reason, the fits for ggF and VBF processes

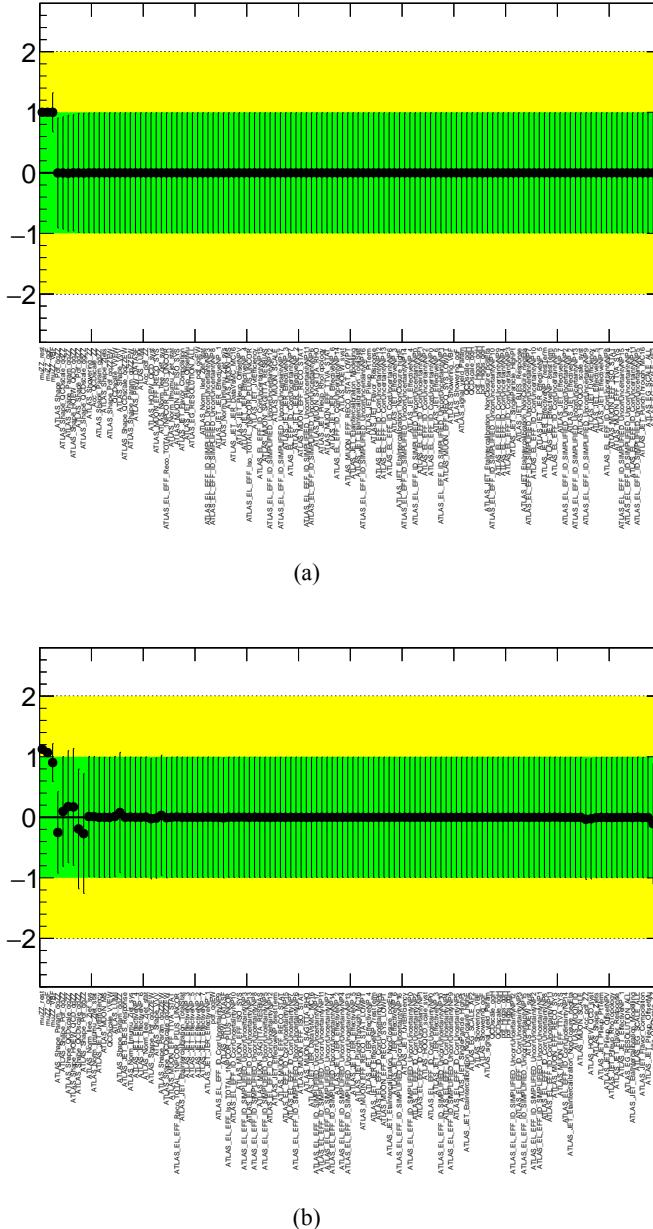


**Fig. 7.25 Distribution of the four-lepton invariant mass  $m_{4\ell}$  in the  $\ell\ell\ell'\ell'$  search for (a), (b), (c) the ggF-MVA-high categories, (d) the ggF-MVA-low category and (e) the VBF-MVA-enriched category. The backgrounds are determined from a combined likelihood fit to the data under the background-only hypothesis. The simulated signal at 600 GeV is normalized to a cross section corresponding to one hundred times the observed upper limit given in section 7.7.3. The error bars on the data points indicate the statistical uncertainty, while the systematic uncertainty in the prediction is shown by the hatched band. The lower panels show the ratio of data to prediction.**

are done separately, and in each case the cross section of the untested process is allowed to be a free parameter in the statistical fit. The observed and expected upper limit at 95% confidence level (CL) on the  $\sigma \times BR(H \rightarrow ZZ)$  of a narrow scalar resonance for both ggF (left) and VBF (right) production mode with the integrated luminosity of  $139 \text{ fb}^{-1}$  is shown in figure 7.28 (7.29) for MVA- (cut-) based analysis. No excess over  $2\sigma$  is found.

## 2. Spin-0 resonance with LWA

In the case of LWA model, only ggF production mode is studied. The interference between the heavy scalar and SM Higgs boson ( $H-h$ ), as well as the heavy scalar and SM  $gg \rightarrow ZZ$  continuum background ( $H-B$ ) as modelled in section 7.5.3 are taken into account. The upper limit at 95% confidence level (CL) on ggF cross section times branch ratio ( $\sigma_{ggF} \times BR(H \rightarrow ZZ)$ ) is shown in figure 7.30 for a width of 1, 5, 10 and 15% of

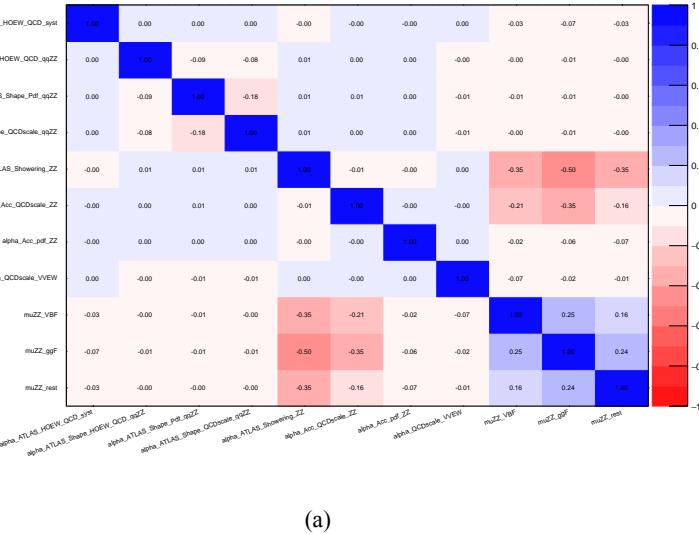


**Fig. 7.26** Pulls and constraints of nuisance parameters after a background only fit to (a) Asimov data and (b) observed data in the  $\ell\ell\ell'\ell'$  channel. The Asimov data is generated with background data only, and the observed data includes datasets from 2015 to 2018.

2478  $m_H$ .

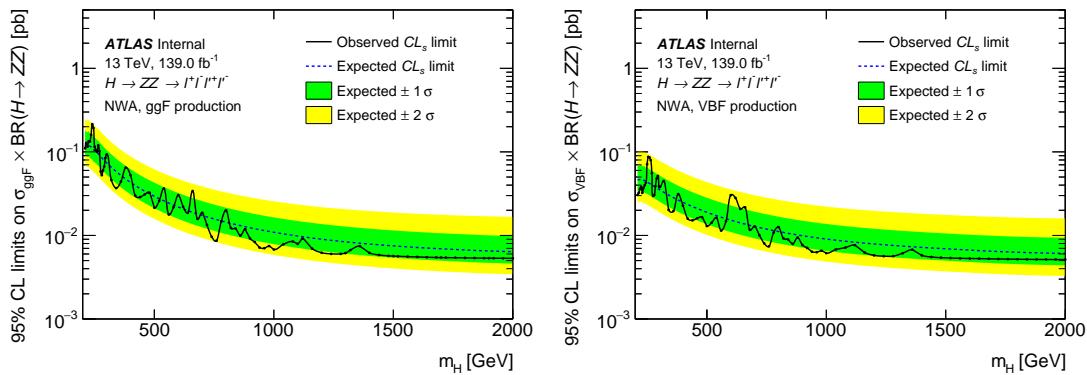
2479     3. Spin-2 RS Graviton resonance

2480     The observed and expected 95% upper limit on the cross section times branching ratio  
 2481     for RS Graviton (RSG) scenario is shown in figure 7.31. Same as LWA case, only  $4e$ ,  
 2482      $4\mu$  and  $2e2\mu$  channel of ggF production mode are used. On top of the expected and  
 2483     observed upper limits in this model, a predicted cross section as function of mass provided  
 2484     by theorist is also shown in the figure. Comparing with the observed result provided by  
 2485      $ZZ \rightarrow \ell\ell\ell'\ell'$  decay, this spin-2 graviton is excluded up to a mass of 1500 GeV.



(a)

**Fig. 7.27 Correlation of nuisance parameters after a background only fit to Asimov data in the  $\ell\ell\ell'\ell'$  channel. The Asimov data is generated with background data only.**



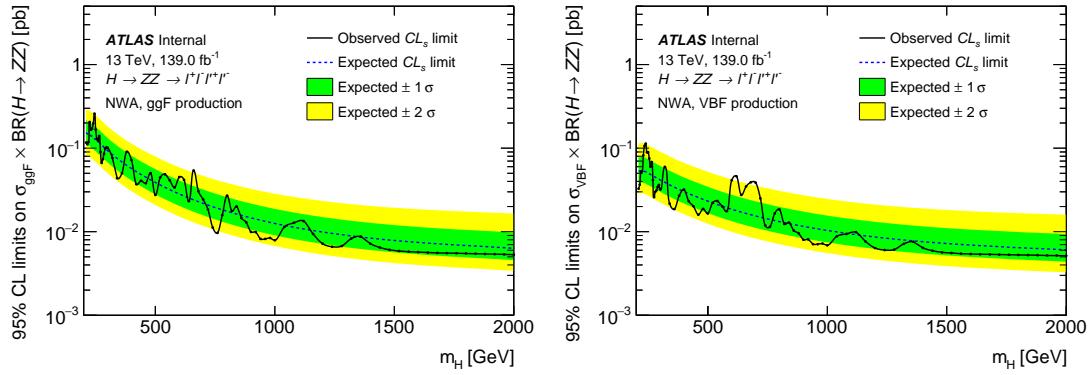
**Fig. 7.28 The expected and observed upper limits at 95% CL on  $\sigma \times BR(H \rightarrow ZZ)$  using the MVA-based analysis for ggF (left) and VBF (right) production. The green and yellow bands represent the  $\pm 1\sigma$  and  $\pm 2\sigma$  uncertainties in the expected limits.**

#### 4. Summary of interpretation

As a summary, figure 7.32 shows the comparison of expected and observed 95% CL upper limits between different models described above.

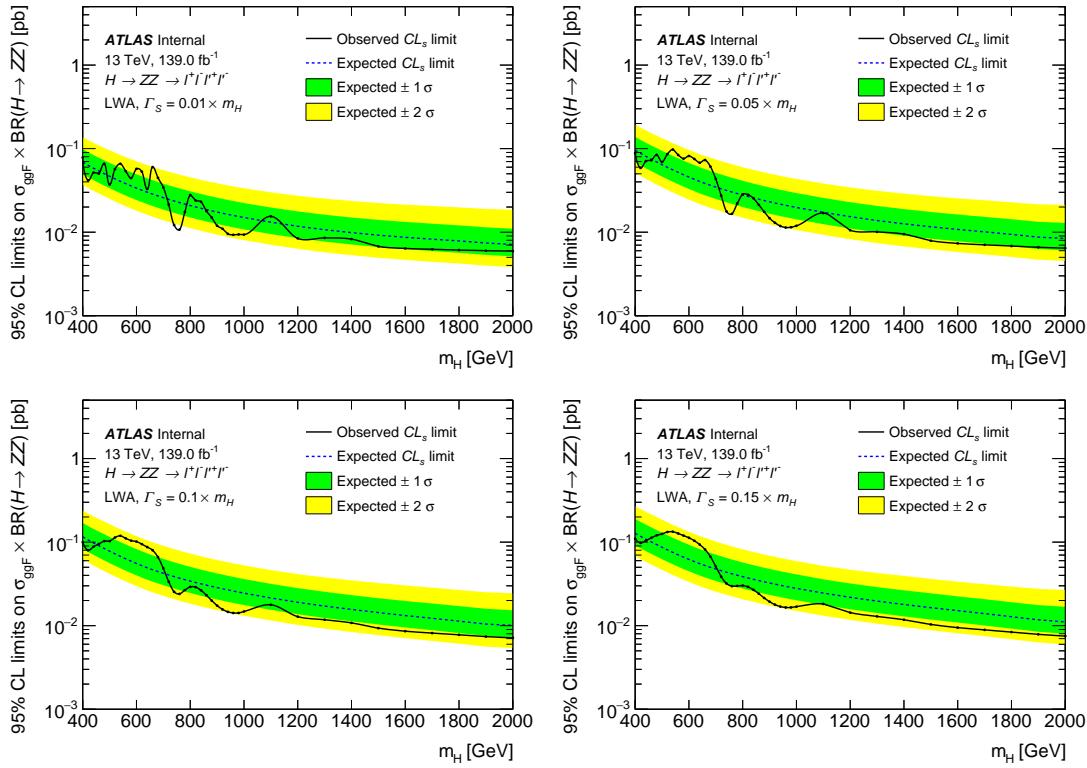
Figure 7.33 compares the expected 95% CL upper limits as a function of the NWA resonance mass in this analysis with full run-2 data and the one in previous publication<sup>[116]</sup> with the integrated luminosity of  $36.1 \text{ fb}^{-1}$ . With a significant increase of integrated luminosity and an improved analysis strategy, comparing to the previous publication, the expected sensitivities of searching for narrow-width heavy resonance reduce by up to 70% in MVA-based analysis, where 50% of reduction is due to luminosity increase while other improvement mainly comes from inviting multivariate method.

Figure 7.34 shows the display of one candidate event passing analysis selection in

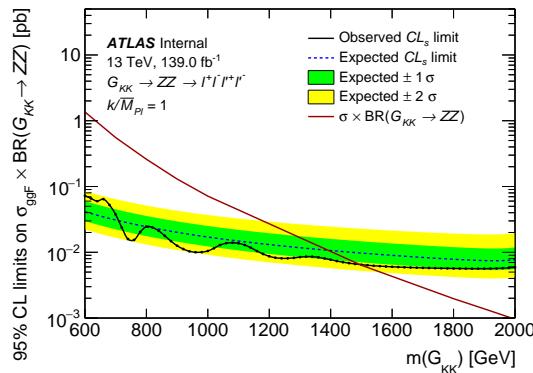


**Fig. 7.29** The expected and observed upper limits at 95% CL on  $\sigma \times \text{BR}(H \rightarrow ZZ)$  using the cut-based analysis for ggF (left) and VBF (right) production. The green and yellow bands represent the  $\pm 1\sigma$  and  $\pm 2\sigma$  uncertainties in the expected limits.

2497 four-muon final state with four-muon invariant mass of 1.34 TeV.



**Fig. 7.30** The upper limits at 95% confidence level on  $\sigma_{ggF} \times BR(H \rightarrow ZZ)$  as a function of the heavy resonance mass  $m_H$  for the ggF production mode with an intrinsic width of 1% (top left), 5% (top right), 10% (bottom left) and 15% (bottom right) for both the case where interference with Standard Model processes is considered. The green and yellow bands represent the  $\pm 1\sigma$  and  $\pm 2\sigma$  uncertainties in the expected limits.



**Fig. 7.31** The upper limits at 95% confidence level on  $\sigma_{ggF} \times BR(G_{KK} \rightarrow ZZ)$  as a function of the heavy resonance mass  $m(G_{KK})$  for the ggF production mode in RS Graviton model. The green and yellow bands represent the  $\pm 1\sigma$  and  $\pm 2\sigma$  uncertainties in the expected limits.

**Table 7.13** A list of the experimental systematics considered in this analysis. The NPs have been separated by whether they only affect the normalisation (left column) or if they affect the shape (right column) of the  $m_{4\ell}$  distribution. They are further subdivided into the primary objects that they affect.

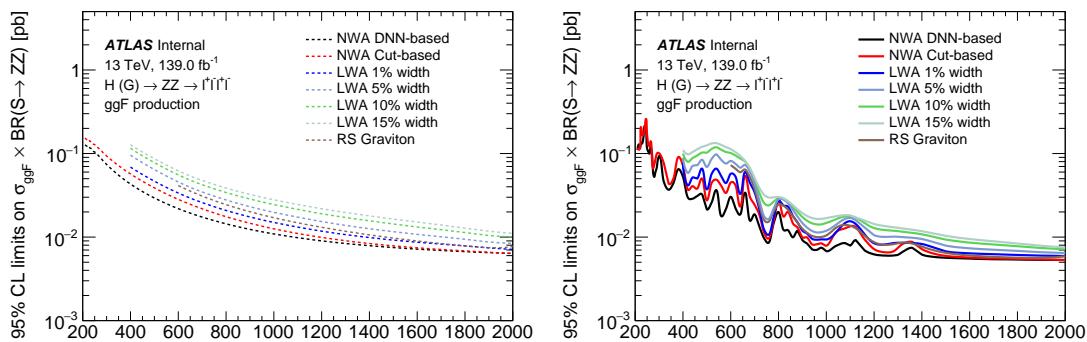
Normalisation NPs	Shape NPs
<b>Electrons</b>	
EL_EFF_ID_CorrUncertaintyNP[0-15]	EG_RESOLUTION_ALL
EL_EFF_ID_SIMPLIFIED_UncorrUncertaintyNP[0-17]	EG_SCALE_ALLCORR
EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	EG_SCALE_E4SCINTILLATOR
EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	EG_SCALE_LARCALIB_EXTRA2015PRE EG_SCALE_LARTEMPERATURE_EXTRA2015PRE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

**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 7.15, are also taken into account.

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

**Table 7.15**  $ZZ$  normalization factor in each category, obtained from a likelihood fit to the data under the background-only hypothesis.

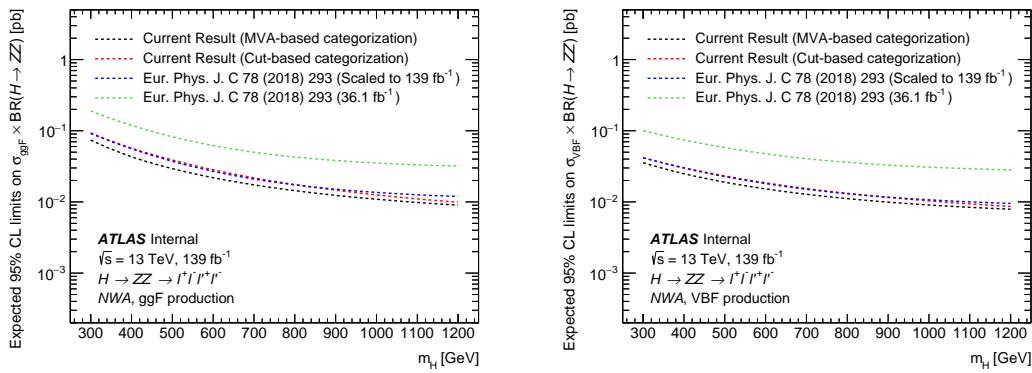
Normalization factor	Fitted value
$\mu_{ZZ}^{ggF-MVA-high}$	$1.07 \pm 0.047$
$\mu_{ZZ}^{ggF-MVA-low}$	$1.12 \pm 0.026$
$\mu_{ZZ}^{VBF-MVA-enriched}$	$0.91 \pm 0.314$



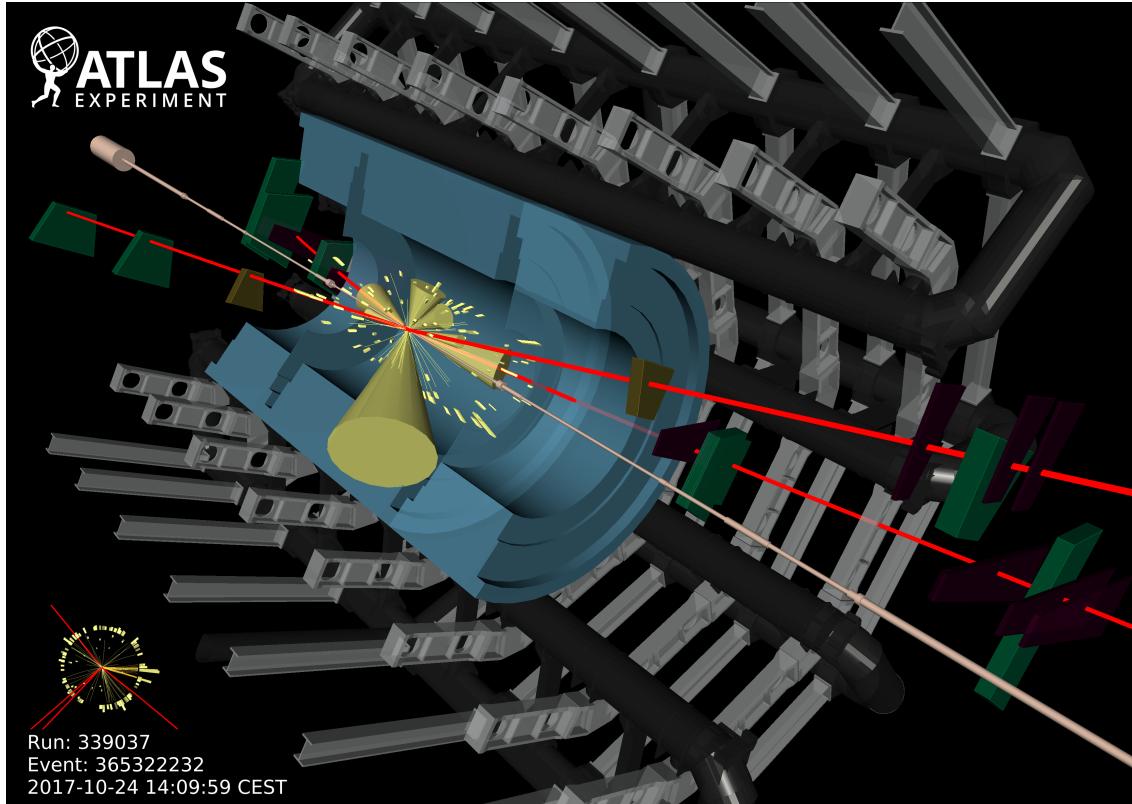
**Fig. 7.32** The expected (left) and observed (right) upper limits at 95% confidence level on  $\sigma \times BR(S \rightarrow ZZ)$  for ggF production mode at different assumptions.

**Table 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



**Fig. 7.33 Comparisons of the expected upper limits at 95% CL on the cross section times branching ratio as a function of the heavy resonance mass  $m_H$  for the ggF production mode (left) and for the VBF production mode (right) in the case of the NWA. The expected limits from the previous publication are shown in the green dashed line and are projected to the  $139 \text{ fb}^{-1}$  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. 7.34 Display of one candidate event in  $4\mu$  final state with the mass of 1.35 TeV.**

## 2498 7.8 Conclusion

2499 Searches of heavy resonances decaying into a pair of  $Z$  boson to  $\ell\ell\ell'\ell'$  final state  
2500 are performed using  $139 \text{ fb}^{-1}$  of 13 TeV pp collision data collected by ATLAS experiment  
2501 at the LHC. The results are interpreted as 95% CL upper limits on the production cross  
2502 section of a spin-0 and spin-2 resonances under different theoretical models. The search  
2503 range of the hypothetical resonances is between 200 GeV to 2000 GeV depending on the  
2504 signal model.

2505 The spin-0 resonance is assumed to be a heavy Higgs like scalar produced predom-  
2506 inantly from gluon–gluon fusion (ggF) and vector-boson fusion (VBF) decays, and it is  
2507 studied under both the narrow-width approximation and with the large-width assumption.  
2508 For narrow-width approximation, limits on cross section of heavy scalar decaying into  
2509 two  $Z$  bosons are set separately for ggF and VBF production modes, under DNN- and  
2510 cut- based analysis. In DNN-based analysis, the 95% CL upper limit range is from  $215 \text{ fb}$   
2511 at  $m_H = 240 \text{ GeV}$  to  $5.3 \text{ fb}$  at  $m_H = 2000 \text{ GeV}$  for ggF production mode, and from  $87 \text{ fb}$   
2512 at  $m_H = 255 \text{ GeV}$  to  $5.1 \text{ fb}$  at  $m_H = 1960 \text{ GeV}$  for VBF production mode. In cut-based  
2513 analysis, the 95% CL upper limit range is from  $259 \text{ fb}$  at  $m_H = 245 \text{ GeV}$  to  $5.3 \text{ fb}$  at  
2514  $m_H = 2000 \text{ GeV}$  for ggF production mode, and from  $113 \text{ fb}$  at  $m_H = 240 \text{ GeV}$  to  $5.1 \text{ fb}$   
2515 at  $m_H = 2000 \text{ GeV}$  for VBF production mode. DNN-based analysis gains about 20% im-  
2516 provement on upper limits at lower mass region comparing to the cut-based analysis, while  
2517 for mass above 1500 GeV, both analyses perform closely. For large-width approximation,  
2518 limits are studied on ggF production rate at four different widths assumptions: 1%, 5%,  
2519 10% and 15% of resonance’s mass, with the interference between the heavy scalar and the  
2520 SM Higgs boson as well as the heavy scalar and the SM  $gg \rightarrow ZZ$  continuum background  
2521 taken into account. The maximum and minimum of upper limits are obtained as  $78 \text{ fb}$  at  
2522  $m_H = 400 \text{ GeV}$  to  $5.9 \text{ fb}$  at  $m_H = 2000 \text{ GeV}$  for 1% width;  $98 \text{ fb}$  at  $m_H = 540 \text{ GeV}$  to  $6.4$   
2523  $\text{fb}$  at  $m_H = 2000 \text{ GeV}$  for 5% width;  $119 \text{ fb}$  at  $m_H = 540 \text{ GeV}$  to  $7.1 \text{ fb}$  at  $m_H = 2000 \text{ GeV}$   
2524 for 10% width;  $133 \text{ fb}$  at  $m_H = 540 \text{ GeV}$  to  $7.5 \text{ fb}$  at  $m_H = 2000 \text{ GeV}$  for 15% width.  
2525 Last but not least, the framework of the Randall–Sundrum model with a graviton excita-  
2526 tion spin-2 resonance with  $m(G_{KK}) < 1500 \text{ GeV}$  is excluded at 95% CL.

2527

## Chapter 8 Summary

2528 On December 3rd, 2018, the LHC finished its second run (run-2) after three fantastic  
2529 years. Thanks to run-2 with largely increased statistic, we now know the masses of the  
2530 Higgs boson, top quark and W boson to considerably greater precision. And also confirm  
2531 the Standard Model as a stable theory.

2532 In this dissertation, various physics processes in  $ZZ \rightarrow \ell\ell\ell'\ell'$  final state are studied,  
2533 taking the advantage of full run-2 pp collision data in the LHC. Using this signature,  
2534 we measured the fiducial cross section of  $ZZ$  production to  $\ell\ell\ell'\ell' jj$  channel in SM,  
2535 which is an important physics process and major background in many analysis with  $ZZ$   
2536 production, eg. Higgs analysis ( $HZZ$ ). In addition, we searched the electroweak  $ZZ$   
2537 production via vector boson scattering in associated with 2-jet process in  $\ell\ell\ell'\ell'$  final  
2538 state. In the meantime, the searches of heavy resonances decaying into a pair of  $Z$  bosons  
2539 to  $\ell\ell\ell'\ell'$  final state for several different hypothetical resonances are conducted in this  
2540 dissertation. The results of several analyses are summarized as below:

2541 **Measurement of fiducial cross section of  $ZZ$  production in  $\ell\ell\ell'\ell' jj$  final state**

2542 The fiducial cross section of inclusive SM  $ZZ \rightarrow \ell\ell\ell'\ell'$  production is measured to  
2543 be:

$$\sigma_{ZZ \rightarrow \ell\ell\ell'\ell'}^{fid} = 1.27 \pm 0.12(stat) \pm 0.02(theo) \pm 0.07(exp) \pm 0.01(bkg) \pm 0.03(lumi) \quad (8.1)$$

2544 which is found to be compatible with the SM prediction. The  $ZZ$  cross section is calcu-  
2545 lated with up to one (three) outgoing partons at NLO (LO) using Sherpa 2.2.2 for QCD  
2546 production, and in LO using MadGraph5\_aMC@NLO 2.6.1 for EW production. The to-  
2547 tal uncertainty is 11%, the analysis is still data static dominant (data statistic uncertainty  
2548 is about 9.5%).

2549 **Observation of electroweak  $ZZ$  production in  $\ell\ell\ell'\ell' jj$  final state**

2550 Thanks to the largely increased data statistic collected by ATLAS experiment in the  
2551 LHC run-2, the electroweak  $ZZ$  production (EW- $ZZjj$ ) to  $\ell\ell\ell'\ell'$  channel in associa-  
2552 tion with two jets is observed with a significant deviation from the background-only  
2553 hypothesis. The signal strength of EW- $ZZjj$  production, the normalization of QCD-  
2554  $ZZjj$  production, as well as the observed and expected statistical significance measured

2555 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)  $\sigma$

2556 Then in this dissertation, the differential cross section and expected significance of  
2557 EW- $Z Z jj$  production, using 3000  $\text{fb}^{-1}$  simulated pp collision data at a centre-of-mass  
2558 energy of 14 TeV to be recorded by ATLAS experiment at the HL-LHC, are studied via  
2559 simulations. The HL-LHC will for sure give us more opportunity to probe rare process  
2560 like  $Z Z \rightarrow \ell\ell\ell'\ell'$  in the future.

2561 **Searches of heavy  $Z Z$  resonances in  $\ell\ell\ell'\ell'$  final state**

2562 Searches of heavy  $Z Z$  resonances are performed in four-lepton invariant mass  $m_{4\ell}$   
2563 range from 200 GeV to 2000 GeV. Data are found to agree with the background-only  
2564 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).

2566 Search range is from 200 GeV to 2000 GeV.

2567 In DNN-based analysis, the limits are range from 215 fb at  $m_H = 240$  GeV to 5.3  
2568 fb at  $m_H = 2000$  GeV for ggF production mode, and from 87 fb at  $m_H = 255$  GeV  
2569 to 5.1 fb at  $m_H = 1960$  GeV for VBF production mode.

2570 The DNN-based analysis is found to be at most 20% better than cut-based results.

- 2571 • Spin-0 heavy Higgs under large-width approximation (LWA) with the width of 1,  
2572 5, 10, 15% of its mass.

2573 Search range is from 400 GeV to 2000 GeV, and only ggF production is studied.

2574 The maximum and minimum of upper limits are obtained as 78 fb at  $m_H = 400$  GeV  
2575 to 5.9 fb at  $m_H = 2000$  GeV for 1% width; 98 fb at  $m_H = 540$  GeV to 6.4 fb at  $m_H =$   
2576 2000 GeV for 5% width; 119 fb at  $m_H = 540$  GeV to 7.1 fb at  $m_H = 2000$  GeV for  
2577 10% width; 133 fb at  $m_H = 540$  GeV to 7.5 fb at  $m_H = 2000$  GeV for 15% width.

- 2578 • Spin-2 graviton excitation under the Randall–Sundrum model.

2579 Search range is from 600 GeV to 2000 GeV, and only ggF production is studied.

2580 The maximum and minimum of limits are 73 fb at  $m_H = 600$  GeV and 5.6 fb  
2581 at  $m_H = 1880$  GeV for ggF production mode. And the mass of graviton below  
2582 1500 GeV is excluded comparing the observed results with theoretical prediction.

2583

2584 In summary, the  $Z Z \rightarrow \ell\ell\ell'\ell'$  production presented in this dissertation are consis-  
2585 tent with SM prediction. This result completes the observation of weak boson scattering

2586 for massive bosons, which is a new milestone reached in the study of electroweak sym-  
2587 metry breaking. In the meantime, no indication of new physics is observed. We are  
2588 looking forward the HL-LHC, with greatly increased luminosity and higher centre-of-  
2589 mass energy, which should enhance the sensitivity for new physics search and precise  
2590 measurement for rare process like  $\ell\ell\ell'\ell'$  final state.

2591

## Bibliography

- 2592 [1] LANGACKER P. Introduction to the Standard Model and Electroweak Physics[C/OL]//  
2593 Proceedings of Theoretical Advanced Study Institute in Elementary Particle Physics on The  
2594 dawn of the LHC era (TASI 2008): Boulder, USA, June 2-27, 2008. 2010: 3-48. DOI:  
2595 [10.1142/9789812838360\\_0001](https://doi.org/10.1142/9789812838360_0001).
- 2596 [2] PICH A. Electroweak Symmetry Breaking and the Higgs Boson[J/OL]. Acta Phys. Polon.,  
2597 2016, B47:151. DOI: [10.5506/APhysPolB.47.151](https://doi.org/10.5506/APhysPolB.47.151).
- 2598 [3] Observation of a new particle in the search for the standard model higgs boson with the atlas  
2599 detector at the lhc[J/OL]. Physics Letters B, 2012, 716(1):1 - 29. <http://www.sciencedirect.com/science/article/pii/S037026931200857X>. DOI: <https://doi.org/10.1016/j.physletb.2012.08.020>.
- 2600 [4] Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc[J/OL].  
2601 Physics Letters B, 2012, 716(1):30 - 61. <http://www.sciencedirect.com/science/article/pii/S0370269312008581>. DOI: <https://doi.org/10.1016/j.physletb.2012.08.021>.
- 2602 [5] PROFUMO S, RAMSEY-MUSOLF M J, SHAUGHNESSY G. Singlet higgs phenomenology  
2603 and the electroweak phase transition[J/OL]. Journal of High Energy Physics, 2007, 2007(08):  
2604 010-010. <https://doi.org/10.1088%2F1126-6708%2F2007%2F08%2F010>. DOI: [10.1088/1126-6708/2007/08/010](https://doi.org/10.1088/1126-6708/2007/08/010).
- 2605 [6] BRANCO G, FERREIRA P, LAVOURA L, et al. Theory and phenomenology of two-higgs-  
2606 doublet models[J/OL]. Physics Reports, 2012, 516(1):1 - 102. <http://www.sciencedirect.com/science/article/pii/S0370157312000695>. DOI: <https://doi.org/10.1016/j.physrep.2012.02.002>.
- 2607 [7] DREMIN I M. Soft and hard processes in QCD[J/OL]. JETP Lett., 2005, 81:307-310. DOI:  
2608 [10.1134/I.1944068](https://doi.org/10.1134/I.1944068).
- 2609 [8] WOMERSLEY J. QCD at the Tevatron: Status and prospects[C/OL]//Proceedings, 5th Inter-  
2610 national Symposium on Radiative Corrections - RADCOR 2000. 2000. <http://www.slac.stanford.edu/econf/C000911/>.
- 2611 [9] LIN H W, et al. Parton distributions and lattice QCD calculations: a community white paper  
2612 [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).
- 2613 [10] COLLINS J C, SOPER D E, STERMAN G F. Factorization of Hard Processes in QCD[J/OL].  
2614 Adv. Ser. Direct. High Energy Phys., 1989, 5:1-91. DOI: [10.1142/9789814503266\\_0001](https://doi.org/10.1142/9789814503266_0001).
- 2615 [11] STIRLING W J. Perturbative QCD[J/OL]. 2000(CERN-OPEN-2000-296):40 p. <https://cds.cern.ch/record/1194745>. DOI: [10.5170/CERN-2000-007.305](https://doi.org/10.5170/CERN-2000-007.305).
- 2616 [12] GROJEAN C. Higgs Physics[J/OL]. 2017(arXiv:1708.00794):143-158. 12 p. <https://cds.cern.ch/record/2243593>. DOI: [10.5170/CERN-2016-005.143](https://doi.org/10.5170/CERN-2016-005.143).

---

## Bibliography

---

- 2624 [13] Cern yellow reports: Monographs: Handbook of LHC Higgs Cross Sections: 4. Deciphering  
2625 the Nature of the Higgs Sector[M/OL]. 2016. <https://cds.cern.ch/record/2227475>. DOI: [10.23731/CYRM-2017-002](https://doi.org/10.23731/CYRM-2017-002).
- 2627 [14] HEINEMEYER S. Cern yellow reports: Monographs: Handbook of LHC Higgs Cross Sec-  
2628 tions: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group[M/OL].  
2629 2013. <https://cds.cern.ch/record/1559921>. DOI: [10.5170/CERN-2013-004](https://doi.org/10.5170/CERN-2013-004).
- 2630 [15] CHANG J, CHEUNG K, LU C T, et al. *ww* scattering in the era of post-higgs-boson discovery  
2631 [J/OL]. Phys. Rev. D, 2013, 87:093005. <https://link.aps.org/doi/10.1103/PhysRevD.87.093005>.
- 2632 [16] BRUNING O S, COLLIER P, LEBRUN P, et al. LHC Design Report Vol.1: The LHC Main  
2633 Ring[J]. 2004.
- 2634 [17] BUNING O, COLLIER P, LEBRUN P, et al. LHC Design Report. 2. The LHC infrastructure  
2635 and general services[J]. 2004.
- 2636 [18] BENEDIKT M, COLLIER P, MERTENS V, et al. LHC Design Report. 3. The LHC injector  
2637 chain[J]. 2004.
- 2638 [19] EVANS L, BRYANT P. LHC machine[J/OL]. Journal of Instrumentation, 2008, 3(08):S08001-  
2639 S08001. <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08001>. DOI: [10.1088/1748-0221/3/08/s08001](https://doi.org/10.1088/1748-0221/3/08/s08001).
- 2641 [20] MOBS E. The CERN accelerator complex. Complexe des accélérateurs du CERN[J/OL]. 2016.  
2642 <https://cds.cern.ch/record/2197559>.
- 2643 [21] COLLABORATION A. The ATLAS experiment at the CERN large hadron collider[J/OL].  
2644 Journal of Instrumentation, 2008, 3(08):S08003-S08003. <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08003>. DOI: [10.1088/1748-0221/3/08/s08003](https://doi.org/10.1088/1748-0221/3/08/s08003).
- 2646 [22] PEREZ G. Unitarization models for vector boson scattering at the lhc[D/OL]. 2018. DOI:  
2647 [10.5445/IR/1000082199](https://doi.org/10.5445/IR/1000082199).
- 2648 [23] PEQUENAO J. Computer generated image of the whole ATLAS detector[Z/OL]. 2008. <https://cds.cern.ch/record/1095924>.
- 2650 [24] MCFAYDEN J. The lhc and atlas detector. in: Third generation susy and  $t^-t^+ + z$  productin.  
2651 [D/OL]. 2014. DOI: [10.1007/978-3-319-07191-6\\_2](https://doi.org/10.1007/978-3-319-07191-6_2).
- 2652 [25] COLLABRATION A. Operation and performance of the ATLAS semiconductor tracker. Op-  
2653 eration and performance of the ATLAS semiconductor tracker[J/OL]. JINST, 2014, 9(CERN-  
2654 PH-EP-2014-049. CERN-PH-EP-2014-049):P08009. 80 p. <https://cds.cern.ch/record/1698966>.  
2655 DOI: [10.1088/1748-0221/9/08/P08009](https://doi.org/10.1088/1748-0221/9/08/P08009).
- 2656 [26] ATLAS pixel detector electronics and sensors[J/OL]. Journal of Instrumentation, 2008, 3(07):  
2657 P07007-P07007. <https://doi.org/10.1088%2F1748-0221%2F3%2F07%2Fp07007>. DOI: [10.1088/1748-0221/3/07/p07007](https://doi.org/10.1088/1748-0221/3/07/p07007).

---

## Bibliography

---

- 2659 [27] MULLIER G. The upgraded pixel detector of the atlas experiment for run-2 at the large hadron  
2660        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).
- 2662 [28] The silicon microstrip sensors of the atlas semiconductor tracker[J/OL]. Nuclear Instruments  
2663        and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Asso-  
2664        ciated Equipment, 2007, 578(1):98 - 118. <http://www.sciencedirect.com/science/article/pii/S0168900207007644>. DOI: <https://doi.org/10.1016/j.nima.2007.04.157>.
- 2666 [29] SULTAN D M S. Development of small-pitch, thin 3d sensors for pixel detector upgrades at  
2667        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).
- 2668 [30] The ATLAS transition radiation tracker (TRT) proportional drift tube: design and performance  
2669        [J/OL]. Journal of Instrumentation, 2008, 3(02):P02013-P02013. <https://doi.org/10.1088%2F1748-0221%2F3%2F02%2Fp02013>. DOI: [10.1088/1748-0221/3/02/p02013](https://doi.org/10.1088/1748-0221/3/02/p02013).
- 2671 [31] BUCHANAN N, CHEN L, GINGRICH D, et al. Design and implementation of the front end  
2672        board for the readout of the atlas liquid argon calorimeters[J/OL]. Journal of Instrumentation,  
2673        2008, 3:P03004. DOI: [10.1088/1748-0221/3/03/P03004](https://doi.org/10.1088/1748-0221/3/03/P03004).
- 2674 [32] Technical design report atlas: ATLAS liquid-argon calorimeter: Technical Design Report  
2675        [M/OL]. Geneva: CERN, 1996. <https://cds.cern.ch/record/331061>.
- 2676 [33] AAD G, ABBOTT B, ABDALLAH J, et al. Readiness of the atlas tile calorimeter for lhc  
2677        collisions[J]. European Physical Journal C, 2010.
- 2678 [34] Technical design report atlas: ATLAS muon spectrometer: Technical Design Report[M/OL].  
2679        Geneva: CERN, 1997. <https://cds.cern.ch/record/331068>.
- 2680 [35] SLIWA K. "ATLAS Overview and Main Results"[C]//Proceedings, International School on  
2681        High Energy Physics : Workshop on High Energy Physics in the near Future. (LISHEP 2013):  
2682        Rio de Janeiro, Brazil, March 17-24, 2013. 2013.
- 2683 [36] RUIZ-MARTINEZ A, COLLABORATION A. The Run-2 ATLAS Trigger System: ATL-DAQ-  
2684        PROC-2016-003[R/OL]. Geneva: CERN, 2016. <https://cds.cern.ch/record/2133909>. DOI:  
2685        [10.1088/1742-6596/762/1/012003](https://doi.org/10.1088/1742-6596/762/1/012003).
- 2686 [37] PÁSZTOR G. The Upgrade of the ATLAS Electron and Photon Triggers towards LHC Run 2 and  
2687        their Performance: ATL-DAQ-PROC-2015-053. arXiv:1511.00334[R/OL]. Geneva: CERN,  
2688        2015. <https://cds.cern.ch/record/2063746>.
- 2689 [38] COLLABORATION A. Atlas computing: Technical design report[R/OL]. Geneva: CERN.  
2690        <http://atlas-computing.web.cern.ch/atlas-computing/packages/athenaCore/athenaCore.php>.
- 2691 [39] AAD G, et al. The ATLAS Simulation Infrastructure[J/OL]. Eur. Phys. J., 2010, C70:823-874.  
2692        DOI: [10.1140/epjc/s10052-010-1429-9](https://doi.org/10.1140/epjc/s10052-010-1429-9).
- 2693 [40] HOCHÉ S. Introduction to parton-shower event generators[C/OL]//Proceedings, Theoretical

---

## Bibliography

---

- 2694 Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Fron-  
2695 tier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014. 2015: 235-  
2696 295. DOI: [10.1142/9789814678766\\_0005](https://doi.org/10.1142/9789814678766_0005).
- 2697 [41] GLEISBERG T, HöCHE S, KRAUSS F, et al. Event generation with SHERPA 1.1[J/OL]. Journal of High Energy Physics, 2009, 2009(02):007-007. <https://doi.org/10.1088%2F1126-6708%2F2009%2F02%2F007>. DOI: [10.1088/1126-6708/2009/02/007](https://doi.org/10.1088/1126-6708/2009/02/007).
- 2700 [42] BÄHR M, GIESEKE S, GIGG M A, et al. Herwig++ physics and manual[J/OL]. The European Physical Journal C, 2008, 58(4):639-707. <https://doi.org/10.1140/epjc/s10052-008-0798-9>.
- 2702 [43] NASON P. A New method for combining NLO QCD with shower Monte Carlo algorithms [J/OL]. JHEP, 2004, 11:040. DOI: [10.1088/1126-6708/2004/11/040](https://doi.org/10.1088/1126-6708/2004/11/040).
- 2704 [44] FRIXIONE S, WEBBER B R. Matching NLO QCD computations and parton shower simulations[J/OL]. Journal of High Energy Physics, 2002, 2002(06):029-029. <https://doi.org/10.1088%2F1126-6708%2F2002%2F06%2F029>. DOI: [10.1088/1126-6708/2002/06/029](https://doi.org/10.1088/1126-6708/2002/06/029).
- 2707 [45] SJOSTRAND T, MRENNA S, SKANDS P Z. A Brief Introduction to PYTHIA 8.1[J/OL]. 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).
- 2709 [46] BARBERIO E, et al. The Geant4-Based ATLAS Fast Electromagnetic Shower Simulation [C/OL]//Astroparticle, particle and space physics, detectors and medical physics applications. Proceedings, 10th Conference, ICATPP 2007, Como, Italy, October 8-12, 2007. 2008: 802-806. <http://cdsweb.cern.ch/record/1064665/files/soft-conf-2007-002.pdf>. DOI: [10.1142/9789812819093\\_0133](https://doi.org/10.1142/9789812819093_0133).
- 2714 [47] RICHTER-WAS E, FROIDEVAUX D, POGGIOLI L. ATLFAST 2.0 a fast simulation package for ATLAS: ATL-PHYS-98-131[R/OL]. Geneva: CERN, 1998. <https://cds.cern.ch/record/683751>.
- 2717 [48] EDMONDS K, FLEISCHMANN S, LENZ T, et al. The Fast ATLAS Track Simulation (FATRAS): ATL-SOFT-PUB-2008-001. ATL-COM-SOFT-2008-002[R/OL]. Geneva: CERN, 2008. <https://cds.cern.ch/record/1091969>.
- 2720 [49] BOYD J T, COSTA M J, TONOYAN A, et al. Commissioning of the ATLAS reconstruction software with first data[J/OL]. Journal of Physics: Conference Series, 2010, 219(3):032059. <https://doi.org/10.1088%2F1742-6596%2F219%2F3%2F032059>. DOI: [10.1088/1742-6596/219/3/032059](https://doi.org/10.1088/1742-6596/219/3/032059).
- 2724 [50] CORNELISSEN T, ELSING M, FLEISCHMANN S, et al. Concepts, Design and Implementation of the ATLAS New Tracking (NEWT): ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002[R/OL]. Geneva: CERN, 2007. <https://cds.cern.ch/record/1020106>.
- 2727 [51] AAD G, et al. Muon reconstruction performance of the ATLAS detector in proton-proton collision data at  $\sqrt{s}=13$  TeV[J/OL]. Eur. Phys. J., 2016, C76(5):292. DOI: [10.1140/epjc/s10052-016-4080-0](https://doi.org/10.1140/epjc/s10052-016-4080-0).

- 2729        016-4120-y.
- 2730 [52] ILLINGWORTH J, KITTNER J. A survey of the hough transform[J/OL]. Computer Vision,  
2731        Graphics, and Image Processing, 1988, 44(1):87 - 116. <http://www.sciencedirect.com/science/article/pii/S0734189X88800331>. DOI: [https://doi.org/10.1016/S0734-189X\(88\)80033-1](https://doi.org/10.1016/S0734-189X(88)80033-1).
- 2733 [53] Vertex Reconstruction Performance of the ATLAS Detector at 13 TeV: ATL-PHYS-PUB-2015-  
2734        026[R/OL]. Geneva: CERN, 2015. <https://cds.cern.ch/record/2037717>.
- 2735 [54] AABOUD M, et al. Reconstruction of primary vertices at the ATLAS experiment in Run 1  
2736        proton-proton collisions at the LHC[J/OL]. Eur. Phys. J., 2017, C77(5):332. DOI: <10.1140/epjc/s10052-017-4887-5>.
- 2738 [55] Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton  
2739        collision data: ATLAS-CONF-2016-024[R/OL]. Geneva: CERN, 2016. <https://cds.cern.ch/reCORD/2157687>.
- 2741 [56] LAMPL W, LAPLACE S, LELAS D, et al. Calorimeter Clustering Algorithms: Description  
2742        and Performance: ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003[R/OL]. Geneva:  
2743        CERN, 2008. <https://cds.cern.ch/record/1099735>.
- 2744 [57] CORNELISSEN T G, ELSING M, GAVRILENKO I, et al. The global  $\chi^2$ track fitter in ATLAS  
2745        [J/OL]. Journal of Physics: Conference Series, 2008, 119(3):032013. <https://doi.org/10.1088%2F1742-6596%2F119%2F3%2F032013>. DOI: <10.1088/1742-6596/119/3/032013>.
- 2747 [58] Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton  
2748        collision data: ATLAS-CONF-2014-032[R/OL]. Geneva: CERN, 2014. <http://cds.cern.ch/reCORD/1706245>.
- 2750 [59] LIMPER M. Track and vertex reconstruction in the ATLAS inner detector[D/OL]. 2009. <http://cds.cern.ch/record/1202457>.
- 2752 [60] AAD G, et al. Topological cell clustering in the ATLAS calorimeters and its performance in  
2753        LHC Run 1[J/OL]. Eur. Phys. J., 2017, C77:490. DOI: <10.1140/epjc/s10052-017-5004-5>.
- 2754 [61] ZHENG Z. Identification of very-low transverse momentum muons with the ATLAS experi-  
2755        ment: ATL-MUON-PROC-2018-018[R/OL]. Geneva: CERN, 2018. <https://cds.cern.ch/reCORD/2649299>.
- 2757 [62] CACCIARI M, SALAM G P. Pileup subtraction using jet areas[J/OL]. Physics Letters B, 2008,  
2758        659(1):119 - 126. <http://www.sciencedirect.com/science/article/pii/S0370269307011094>. DOI:  
2759        <https://doi.org/10.1016/j.physletb.2007.09.077>.
- 2760 [63] CACCIARI M, SALAM G P, SOYEZ G. The anti-k<sub>T</sub>jet clustering algorithm[J/OL]. Journal of  
2761        High Energy Physics, 2008, 2008(04):063-063. <https://doi.org/10.1088%2F1126-6708%2F2008%2F04%2F063>. DOI: <10.1088/1126-6708/2008/04/063>.
- 2763 [64] CACCIARI M, SALAM G P, SOYEZ G. Fastjet user manual[J/OL]. The European Physical

## Bibliography

---

- 2764      Journal C, 2012, 72(3):1896. <https://doi.org/10.1140/epjc/s10052-012-1896-2>.
- 2765 [65] AAD G, et al. Jet energy measurement and its systematic uncertainty in proton-proton collisions  
2766 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).
- 2768 [66] AABOUD M, et al. Jet energy scale measurements and their systematic uncertainties in proton-  
2769 proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector[J/OL]. Phys. Rev., 2017, D96(7):  
2770 072002. DOI: [10.1103/PhysRevD.96.072002](https://doi.org/10.1103/PhysRevD.96.072002).
- 2771 [67] Optimisation of the ATLAS *b*-tagging performance for the 2016 LHC Run: ATL-PHYS-PUB-  
2772 2016-012[R/OL]. Geneva: CERN, 2016. <https://cds.cern.ch/record/2160731>.
- 2773 [68] PIACQUADIO G, WEISER C. A new inclusive secondary vertex algorithm for b-jet tagging  
2774 in ATLAS[J/OL]. Journal of Physics: Conference Series, 2008, 119(3):032032. <https://doi.org/10.1088%2F1742-6596%2F119%2F3%2F032032>. DOI: [10.1088/1742-6596/119/3/032032](https://doi.org/10.1088/1742-6596/119/3/032032).
- 2776 [69] SPECKMAYER P, HÖCKER A, STELZER J, et al. The toolkit for multivariate data analysis,  
2777 TMVA 4[J/OL]. Journal of Physics: Conference Series, 2010, 219(3):032057. <https://doi.org/10.1088%2F1742-6596%2F219%2F3%2F032057>. DOI: [10.1088/1742-6596/219/3/032057](https://doi.org/10.1088/1742-6596/219/3/032057).
- 2779 [70] Performance of missing transverse momentum reconstruction with the atlas detector using  
2780 proton–proton collisions at  $\sqrt{s}=13$  TeV[J/OL]. The European Physical Journal C, 2018, 78(11):903. <https://doi.org/10.1140/epjc/s10052-018-6288-9>.
- 2782 [71] READ A L. Presentation of search results: theCLstechnique[J/OL]. Journal of Physics G:  
2783 Nuclear and Particle Physics, 2002, 28(10):2693-2704. [https://doi.org/10.1088%2F0954-3899/28/10/313](https://doi.org/10.1088%2F0954-3899%2F28%2F10%2F313). DOI: [10.1088/0954-3899/28/10/313](https://doi.org/10.1088/0954-3899/28/10/313).
- 2785 [72] LEE B W, QUIGG C, THACKER H B. The Strength of Weak Interactions at Very High-  
2786 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).
- 2788 [73] CHANOWITZ M S, GAILLARD M K. The TeV Physics of Strongly Interacting W's and Z's  
2789 [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).
- 2790 [74] SZLEPER M. The Higgs boson and the physics of  $WW$  scattering before and after Higgs  
2791 discovery[J]. 2014.
- 2792 [75] Evidence for electroweak production of  $W^\pm W^\pm jj$  in  $pp$  collisions at  $\sqrt{s} = 8$  TeV with the  
2793 atlas detector[J/OL]. Phys. Rev. Lett., 2014, 113:141803. <https://link.aps.org/doi/10.1103/PhysRevLett.113.141803>.
- 2795 [76] Observation of electroweak production of a same-sign  $w$  boson pair in association with two jets  
2796 in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the atlas detector[J/OL]. Phys. Rev. Lett., 2019, 123:  
2797 161801. <https://link.aps.org/doi/10.1103/PhysRevLett.123.161801>.
- 2798 [77] SIRUNYAN A M, et al. Observation of electroweak production of same-sign W boson pairs

---

## Bibliography

---

- 2799       in the two jet and two same-sign lepton final state in proton-proton collisions at  $\sqrt{s} = 13$  TeV  
2800       [J/OL]. Phys. Rev. Lett., 2018, 120(8):081801. DOI: [10.1103/PhysRevLett.120.081801](https://doi.org/10.1103/PhysRevLett.120.081801).
- 2801 [78] Observation of electroweak  $w \pm z$  boson pair production in association with two jets in pp col-  
2802       lisions at  $s=13$  tev with the atlas detector[J/OL]. Physics Letters B, 2019, 793:469 - 492.  
2803       <http://www.sciencedirect.com/science/article/pii/S0370269319303211>. DOI: <https://doi.org/10.1016/j.physletb.2019.05.012>.
- 2805 [79] Measurement of vector boson scattering and constraints on anomalous quartic couplings from  
2806       events with four leptons and two jets in proton–proton collisions at  $s=13$  tev[J/OL]. Physics  
2807       Letters B, 2017, 774:682 - 705. <http://www.sciencedirect.com/science/article/pii/S0370269317308328>. DOI: <https://doi.org/10.1016/j.physletb.2017.10.020>.
- 2809 [80] ALWALL J, FREDERIX R, FRIXIONE S, et al. The automated computation of tree-level and  
2810       next-to-leading order differential cross sections, and their matching to parton shower simula-  
2811       tions[J/OL]. JHEP, 2014, 07:079. DOI: [10.1007/JHEP07\(2014\)079](https://doi.org/10.1007/JHEP07(2014)079).
- 2812 [81] BALL R D, et al. Parton distributions with LHC data[J/OL]. Nucl. Phys. B, 2013, 867:244-289.  
2813       DOI: [10.1016/j.nuclphysb.2012.10.003](https://doi.org/10.1016/j.nuclphysb.2012.10.003).
- 2814 [82] GLEISBERG T, HÖCHE S, KRAUSS F, et al. Event generation with SHERPA 1.1[J/OL].  
2815       JHEP, 2009, 02:007. DOI: [10.1088/1126-6708/2009/02/007](https://doi.org/10.1088/1126-6708/2009/02/007).
- 2816 [83] BALL R D, et al. Parton distributions for the LHC run II[J/OL]. JHEP, 2015, 04:040. DOI:  
2817       [10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040).
- 2818 [84] CAOLA F, MELNIKOV K, RÖNTSCH R, et al. Qcd corrections to  $zz$  production in gluon  
2819       fusion at the lhc[J/OL]. Phys. Rev. D, 2015, 92:094028. <https://link.aps.org/doi/10.1103/PhysRevD.92.094028>.
- 2821 [85] FRIXIONE S, RIDOLFI G, NASON P. A positive-weight next-to-leading-order Monte Carlo  
2822       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).
- 2824 [86] LAI H L, et al. New parton distributions for collider physics[J/OL]. Phys. Rev. D, 2010, 82:  
2825       074024. DOI: [10.1103/PhysRevD.82.074024](https://doi.org/10.1103/PhysRevD.82.074024).
- 2826 [87] ALIOLI S, NASON P, OLEARI C, et al. NLO single-top production matched with shower in  
2827       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).
- 2829 [88] FREDERIX R, RE E, TORRIELLI P. Single-top t-channel hadroproduction in the four-flavour  
2830       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).
- 2832 [89] RE E. Single-top Wt-channel production matched with parton showers using the POWHEG  
2833       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).

---

## Bibliography

---

- 2834 [90] COLLABORATION A. ATLAS Pythia 8 tunes to 7 TeV data[M/OL]. 2014. <https://cds.cern.ch/record/1966419>.
- 2835
- 2836 [91] ATLAS Collaboration. Performance of pile-up mitigation techniques for jets in  $pp$  collisions at  
2837  $\sqrt{s} = 8$  TeV using the ATLAS detector[J/OL]. Eur. Phys. J. C, 2016, 76:581. DOI: [10.1140/epjc/s10052-016-4395-z](https://doi.org/10.1140/epjc/s10052-016-4395-z).
- 2838
- 2839 [92] BELLM J, et al. Herwig 7.0/Herwig++ 3.0 release note[J/OL]. Eur. Phys. J. C, 2016, 76(4):  
2840 196. DOI: [10.1140/epjc/s10052-016-4018-8](https://doi.org/10.1140/epjc/s10052-016-4018-8).
- 2841 [93] BAHR M, et al. Herwig++ Physics and Manual[J/OL]. Eur. Phys. J. C, 2008, 58:639-707. DOI:  
2842 [10.1140/epjc/s10052-008-0798-9](https://doi.org/10.1140/epjc/s10052-008-0798-9).
- 2843 [94] Luminosity determination in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector at the  
2844 LHC: ATLAS-CONF-2019-021[R/OL]. Geneva: CERN, 2019. <http://cds.cern.ch/record/2677054>.
- 2845
- 2846 [95] The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS[J/OL]. Journal  
2847 of Instrumentation, 2018, 13(07):P07017-P07017. <https://doi.org/10.1088/1748-0221/13/07/p07017>. DOI: [10.1088/1748-0221/13/07/p07017](https://doi.org/10.1088/1748-0221/13/07/p07017).
- 2847
- 2848
- 2849 [96] COADOU Y. Boosted decision trees and applications[J/OL]. EPJ Web of Conferences, 2013,  
2850 55:02004-. DOI: [10.1051/epjconf/20135502004](https://doi.org/10.1051/epjconf/20135502004).
- 2851 [97] COLLABORATION A. Technical Design Report for the ATLAS Inner Tracker Pixel Detector:  
2852 CERN-LHCC-2017-021. ATLAS-TDR-030[R/OL]. Geneva: CERN, 2017. <https://cds.cern.ch/record/2285585>.
- 2853
- 2854 [98] COLLABORATION A. Technical Design Report for the Phase-II Upgrade of the ATLAS Muon  
2855 Spectrometer: CERN-LHCC-2017-017. ATLAS-TDR-026[R/OL]. Geneva: CERN, 2017. [ht  
tps://cds.cern.ch/record/2285580](https://cds.cern.ch/record/2285580).
- 2856
- 2857 [99] COLLABORATION A. Technical Proposal: A High-Granularity Timing Detector for the AT-  
2858 LAS Phase-II Upgrade: CERN-LHCC-2018-023. LHCC-P-012[R/OL]. Geneva: CERN, 2018.  
2859 <https://cds.cern.ch/record/2623663>.
- 2860 [100] Expected performance for an upgraded ATLAS detector at High-Luminosity LHC: ATL-PHYS-  
2861 PUB-2016-026[R/OL]. Geneva: CERN, 2016. <https://cds.cern.ch/record/2223839>.
- 2862 [101] HILL A, VAN DER BIJ J J. Strongly interacting singlet-doublet higgs model[J/OL]. Phys. Rev.  
2863 D, 1987, 36:3463-3473. <https://link.aps.org/doi/10.1103/PhysRevD.36.3463>.
- 2864 [102] DAVOUDIASL H, HEWETT J, RIZZO T. Bulk gauge fields in the randall–sundrum  
2865 model work supported by the department of energy, contract de-ac03-76sf00515.1[J/OL].  
2866 Physics Letters B, 2000, 473(1):43 - 49. <http://www.sciencedirect.com/science/article/pii/S03700269399014306>. DOI: [https://doi.org/10.1016/S0370-2693\(99\)01430-6](https://doi.org/10.1016/S0370-2693(99)01430-6).
- 2867
- 2868 [103] FREDERIX R, FRIXIONE S. Merging meets matching in MC@NLO[J/OL]. JHEP, 2012, 12:

## Bibliography

---

- 2869 061. DOI: [10.1007/JHEP12\(2012\)061](https://doi.org/10.1007/JHEP12(2012)061).
- 2870 [104] AGASHE K, DAVOUDIASL H, PEREZ G, et al. Warped Gravitons at the LHC and Beyond  
2871 [J/OL]. Phys.Rev., 2007, D76:036006. DOI: [10.1103/PhysRevD.76.036006](https://doi.org/10.1103/PhysRevD.76.036006).
- 2872 [105] ATLAS Collaboration. Jet reconstruction and performance using particle flow with the ATLAS  
2873 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).
- 2874 [106] Measurements of higgs boson production and couplings in the four-lepton channel in  $pp$  col-  
2875 lisions at center-of-mass energies of 7 and 8 tev with the atlas detector[J/OL]. Phys. Rev. D,  
2876 2015, 91:012006. <https://link.aps.org/doi/10.1103/PhysRevD.91.012006>.
- 2877 [107] CRANMER K S. Kernel estimation in high-energy physics[J/OL]. Comput. Phys. Commun.,  
2878 2001, 136:198-207. DOI: [10.1016/S0010-4655\(00\)00243-5](https://doi.org/10.1016/S0010-4655(00)00243-5).
- 2879 [108] OREGLIA M. A Study of the Reactions  $\psi' \rightarrow \gamma\gamma\psi$ [M/OL]. SLAC, 1980. <https://www.slac.stanford.edu/cgi-wrap/getdoc/slac-r-236.pdf>.
- 2880 [109] GAISER J. Charmonium Spectroscopy From Radiative Decays of the  $J/\psi$  and  $\psi'$ [M/OL].  
2881 SLAC, 1982. <https://www.slac.stanford.edu/cgi-wrap/getdoc/slac-r-255.pdf>.
- 2882 [110] GORIA S, PASSARINO G, ROSCO D. The Higgs Boson Lineshape[J/OL]. Nucl.Phys., 2012,  
2883 B864:530-579. DOI: [10.1016/j.nuclphysb.2012.07.006](https://doi.org/10.1016/j.nuclphysb.2012.07.006).
- 2884 [111] SPIRA M, DJOUADI A, GRAUDENZ D, et al. Higgs boson production at the LHC[J]. Nucl.  
2885 Phys., 1995, B 453:17-82.
- 2886 [112] BALL R D, et al. Parton Distribution Benchmarking with LHC Data[J/OL]. JHEP, 2013, 04:  
2887 125. DOI: [10.1007/JHEP04\(2013\)125](https://doi.org/10.1007/JHEP04(2013)125).
- 2888 [113] BIJNENS J, EEROLA P, MAUL M, et al. Qcd signatures of narrow graviton resonances in  
2889 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).
- 2890 [114] ATLAS Collaboration. PMG Systematic Uncertainty Recipes[EB/OL]. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PmgSystematicUncertaintyRecipes>.
- 2891 [115] READ A L. Presentation of search results: The CL(s) technique[J]. J. Phys. G, 2002, 28:  
2892 2693-2704.
- 2893 [116] AABOUD M, et al. Search for heavy ZZ resonances in the  $\ell^+\ell^-\ell^+\ell^-$  and  $\ell^+\ell^-\nu\bar{\nu}$  final states  
2894 using proton–proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector[J/OL]. Eur. Phys. J.  
2895 C, 2018, 78(4):293. DOI: [10.1140/epjc/s10052-018-5686-3](https://doi.org/10.1140/epjc/s10052-018-5686-3).