

# 中国科学技术大学

# 博士学位论文



## 利用 ATLAS 探测器上 ZZ 玻色子到全 轻子通道的衰变事例进行电弱对称性破 缺的研究

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： 二〇二〇年一月二十二日

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



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

Author: Heling Zhu

Speciality: Particle and Nuclear Physics

Supervisors: Zhengguo Zhao, Hong Ma

Finished time: January 22, 2020

## 中国科学技术大学学位论文原创性声明

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

作者签名：\_\_\_\_\_

签字日期：\_\_\_\_\_

## 中国科学技术大学学位论文授权使用声明

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

保密的学位论文在解密后也遵守此规定。

公开  保密（\_\_\_\_ 年）

作者签名：\_\_\_\_\_

导师签名：\_\_\_\_\_

签字日期：\_\_\_\_\_

签字日期：\_\_\_\_\_

## 摘要

中文摘要

关键词：

## **ABSTRACT**

English abstract.

**Key Words:**

## **Acknowledgments**

## Contents

Acknowledgments . . . . .	III
Chapter 1 Introduction . . . . .	1
Chapter 2 Theory . . . . .	3
2.1 The Standard Model of Particle Physics . . . . .	3
2.1.1 Elementary particles in the Standard Model . . . . .	3
2.1.2 Electroweak theory . . . . .	5
2.1.3 Higgs mechanism and Electroweak symmetry breaking . . . . .	6
2.2 Phenomenology of Large Hadron Collider . . . . .	8
2.2.1 Physics at hadronic collision . . . . .	9
2.2.2 Higgs physics at LHC . . . . .	11
2.2.3 Diboson physics . . . . .	12
Chapter 3 The Large Hadron Collider and the ATLAS Detector . . . . .	18
3.1 The Large Hadron Collider . . . . .	18
3.1.1 Operation history and machine layout . . . . .	18
3.1.2 Luminosity and pile-up . . . . .	20
3.2 ATLAS detector . . . . .	23
3.2.1 Detector overview . . . . .	23
3.2.2 Physics requirement . . . . .	25
3.2.3 Magnet system . . . . .	26
3.2.4 Inner detector . . . . .	27
3.2.5 Calorimeters . . . . .	28
3.2.6 Muon spectrometer . . . . .	31
3.2.7 Trigger system . . . . .	32
Chapter 4 Simulation and Event Reconstruction for the ATLAS Experiment . . . . .	36
4.1 Event simulation . . . . .	36
4.2 Event reconstruction . . . . .	39
4.2.1 Track . . . . .	39
4.2.2 Primary vertex . . . . .	41
4.2.3 Electron . . . . .	43

4.2.4	Muon	47
4.2.5	Jets	52
4.2.6	Missing transverse energy	55
Chapter 5	Observation of electroweak ZZ production and measurement of SM ZZ cross section	58
5.1	Introduction	58
5.2	Data and MC samples	59
5.2.1	Data samples	59
5.2.2	MC simulation	60
5.3	Objects and Event selection	61
5.3.1	Objects selection	61
5.3.2	Event selection in reconstruction level	62
5.4	Background estimation	63
5.4.1	QCD backgrounds	64
5.4.2	Fake backgrounds	64
5.5	Systematics	68
5.5.1	Theoretical systematics	69
5.5.2	Experimental systematics	70
5.6	Measurement of fiducial cross section	72
5.6.1	Calculation of C-factor	73
5.6.2	Result of fiducial cross section	74
5.7	Search for EW-ZZjj	74
5.7.1	MD discriminant	76
5.7.2	Profile likelihood ratio method	76
5.7.3	Fitting procedure	77
5.7.4	Result of statistical fit	78
5.8	Prospect study of EW-ZZjj production in HL-LHC	80
5.8.1	Introduction	80
5.8.2	The ATLAS detector at HL-LHC	80
5.8.3	Simulation	80
5.8.4	Event selection	81
5.8.5	Systematics	82
5.8.6	Results	84
5.9	Conclusion	87

---

Contents

Bibliography . . . . .	88
------------------------	----

---

## Contents

*test*

## Chapter 1 Introduction

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

### Elementary particles

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

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

The well-known Standard model (SM) was developed in stages throughout the latter half of the 20th century. Since then, confirmation of the top quark (1995), the tau neutrino (2000), and the Higgs boson (2012) have added further credence to the Standard Model. Now, the quarks, leptons and gauge bosons are the elementary constituents in a framework of Standard Model of particle physics, which theoret-

cally describes three of the four known fundamental forces (the electromagnetic, weak, and strong interactions, and not including the gravitational force) in the universe, as well as classifies all known elementary particles.

### Higgs mechanics and electroweak symmetry breaking

In 1961, Sheldon Glashow brought forward a unified electroweak theory to combine the electromagnetic and weak interactions. In the standard model, at energy high enough that electroweak symmetry is unbroken, all elementary particles are massless. But measurements show the fact that the W and Z bosons actually have masses. Later on, the Higgs mechanics resolves this conundrum. The simplest description of the mechanism adds a Higgs field that permeates all space to the Standard Model. Below some extremely high energy, the field causes spontaneous symmetry breaking during interactions. All massive particles in the Standard Model, including the W and Z bosons, interact with Higgs boson to acquire their mass.

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

### Contents of this thesis

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

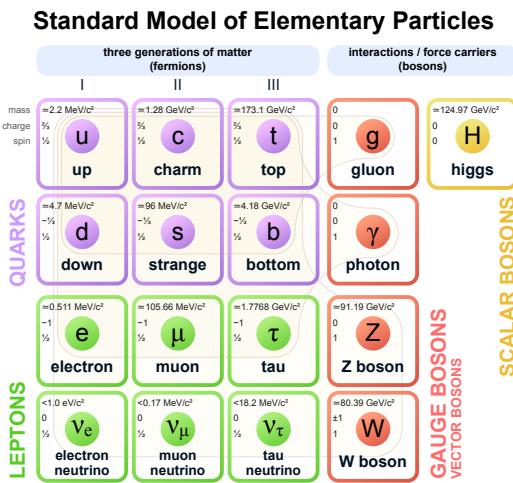
# Chapter 2 Theory

## 2.1 The Standard Model of Particle Physics

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

### 2.1.1 Elementary particles in the Standard Model

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.**

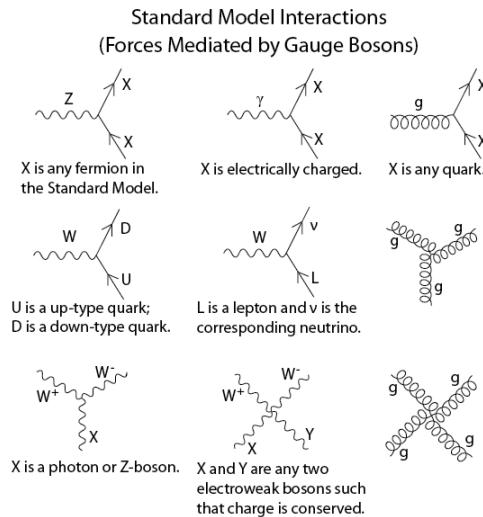
**Fermions** The Standard Model includes 12 elementary particles of spin- $\frac{1}{2}$  obeying the Fermi-Dirac statistics, known as fermions. They are classified into two types: *leptons* and *quarks* according to their interactions. The *leptons* include three generations: electron ( $e$ ) and electron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

from one to another, forming color-neutral composite particles (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 (the 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.  $Z$  boson is electrically neutral.

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



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

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

### 2.1.2 Electroweak theory

The electroweak interaction is the unified description of two of the four known fundamental interactions of nature: electromagnetism and the weak interaction. It is based on the gauge group of  $SU(2)_L \times SU(1)_Y$ , in which  $L$  is the left-handed fields and  $Y$  is the weak hypercharge<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  $B_\mu$  respectively present the  $SU(2)_L$  and  $SU(1)_Y$  gauge fields, with the corresponding field strength tensors of

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

### 2.1.3 Higgs mechanism and Electroweak symmetry breaking

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

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

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

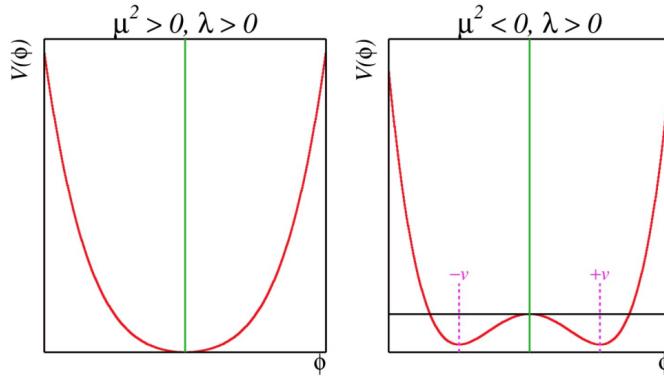
where  $\phi_i = \phi_i^+$  stand for four Hermitian field. In this new basis, the Higgs potential in Eq. 2.5 can be expressed as:

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

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

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

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



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

doublet can be expressed by

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

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

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

Where  $H$  is a Hermitian field for physical Higgs scalar. In this gauge, the Lagrangian  $L_{Higgs}$  in Eq. 2.4 takes a simple form

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

where the  $W$  and  $Z$  fields are

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

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

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

where  $\theta_W$  is the weak angle defined as

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

Then another gauge boson photon remains massless with the field of

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

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

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

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

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

Due to the unknown of quartic Higgs coupling  $\lambda$ , the Higgs mass is not predicted. The third and fourth terms in Higgs potential  $V$  denote the induced cubic and quartic interactions of the Higgs scalar.

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

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

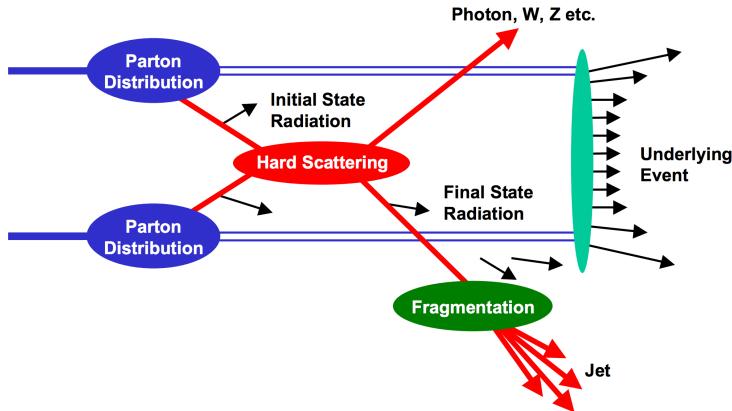
in which  $m_f = \frac{y_f v}{\sqrt{2}}$  for  $f = d, u, l$ .

## 2.2 Phenomenology of Large Hadron Collider

The Large Hadron Collider (LHC) was built as a bridge between the theories and the experiment. Physicists hope that the LHC can help to answer some of the fundamental open questions in physics, concerning the basic laws of interactions and forces among the elementary particles, the deep structure of space and time, and in particular the interrelation between quantum mechanics and general relativity. This section will talk about firstly the general introduction of Physics inside hadronic collision, then followed by two importants LHC phenomenology of Higgs physics and Diboson physics which is related closely to this thesis.

### 2.2.1 Physics at hadronic collision

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



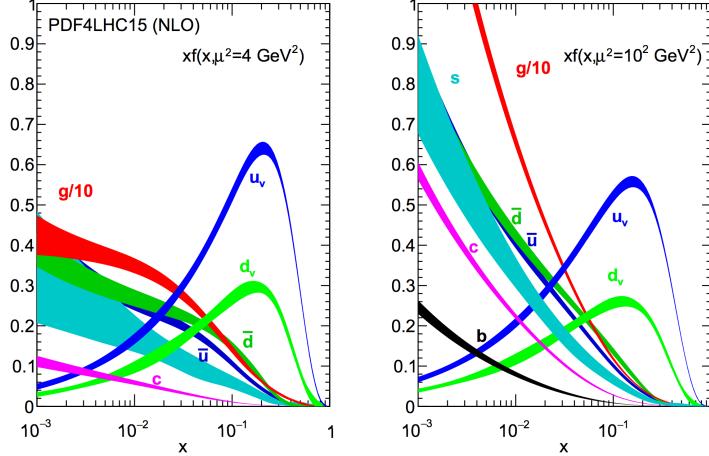
**Fig. 2.4** Schematic view of a hadron-hadron collision<sup>[4]</sup>.

typical features are summarized as below:

- **Parton Distribution Function (PDF):**  $f_i(x, Q^2)$  gives the probability of a parton with flavor  $i$  (quark or gluon), carrying amomentum fraction of  $x$  and at the energy of  $Q$  in a proton. Parton distribution function cannot be fully calculated by perturbative QCD because of the inherent non-perturbative nature of partons. There are many different sets of PDFs that are determined by a fit to data from experimental observables in various processes. As an example, figure 2.5 for *PDF4LHC15* which is based on the combination of the *CT14*, *MMHT14* and *NNPDF3.1* NNLO PDF sets<sup>[5]</sup>.
- **Fragmentation and hadronization:** The processes to produce final state particles (or jets) from the partons produced in hard scattering.
- **Initial/Final state radiation:** The incoming and outgoing partons that carry color

charge can emit QCD radiation, which gives rise to additional jets. Also the charged incoming and outgoing particles can emit QED radiations with photons.

- **Underlying events:** Products from soft processes (not come from the primary hard 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  $x$ .**

### Cross section of hard scattering

According to *QCD factorization theorems*<sup>[6]</sup>, the perturbative calculations can be applied to many important hard processes involving hadrons. The basic problem addressed by factorization theorems is how to calculate high energy cross sections. Consider the process of scattering between two hardons A and B to produce a final state X, the cross section  $\sigma$  can be obtained by summing over all the subprocess cross section  $\hat{\sigma}$ <sup>[7]</sup>

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

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

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

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

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

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

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

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

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

### 2.2.2 Higgs physics at LHC

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

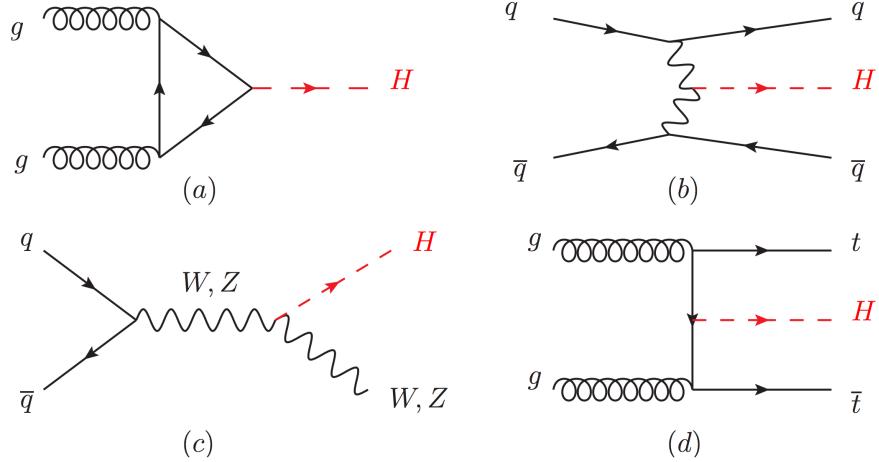
#### Higgs productions

Higgs boson can be produced through several processes. There are 4 main production modes at LHC: gluon-gluon fusion ( $ggF$ ), vector boson fusion ( $VBF$ ), associated production with vector-bosons (also called Higgs strahlung) ( $VH$ ) and associated production with a pair of top/antitop quarks ( $ttH$ )<sup>[8]</sup>. Figure 2.6 shows the corresponding Feynman diagrams of each process (at LO). For pp collision, the cross section of productions of Higgs boson is as a function of center-of-mass-energy  $\sqrt{s}$ . Figure 2.7 summarizes the cross section for SM Higgs with mass of 125 GeV.

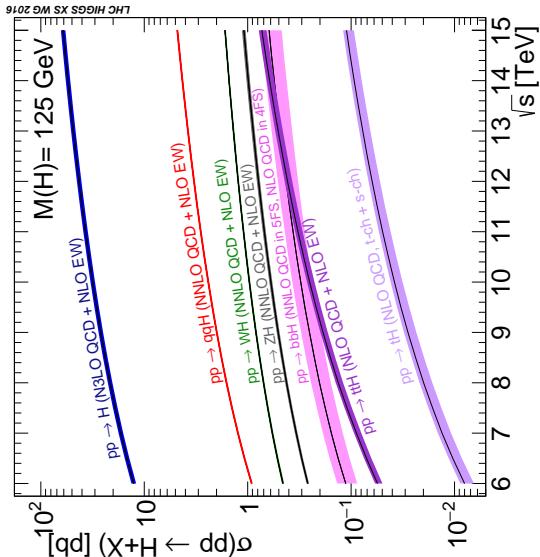
Figure 2.8 summarizes the prospect of different Higgs boson production cross sections as a function of Higgs mass for pp collision center-of-mass-energy at 13 TeV and 14 TeV<sup>[9]</sup>.

#### Higgs decays

Higgs boson can interact with gauge bosons and fermions through gauge coupling and Yukawa coupling as introduced in section 2.1.3. Figure 2.9 depicts Feynman dia-



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



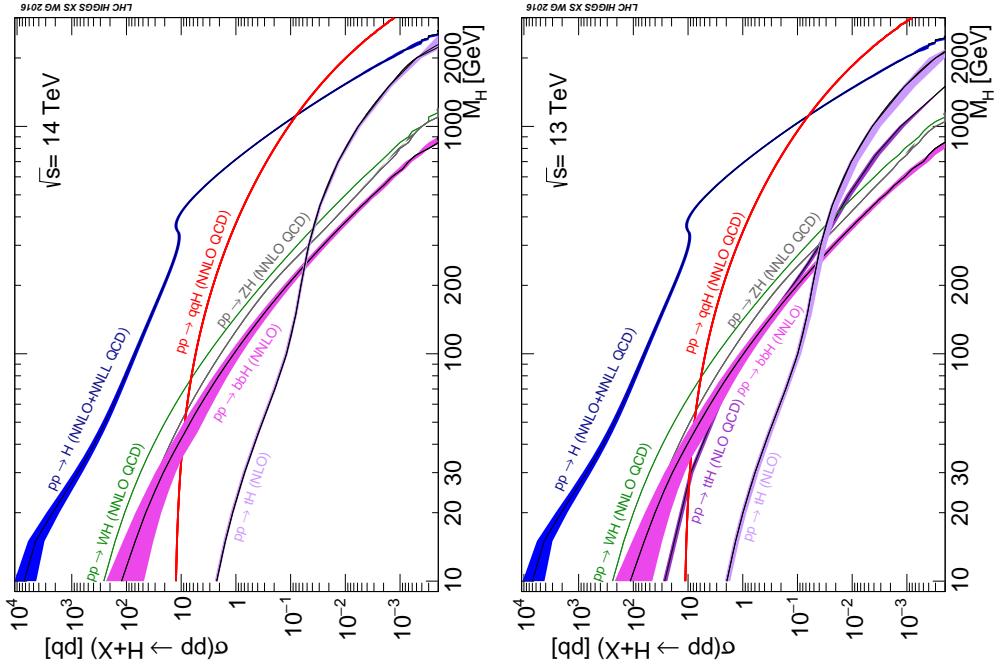
**Fig. 2.7** The SM Higgs boson production cross sections as a function of the center-of-mass energy for pp collision.

grams of possible Higgs decay channels. The branching ratio of Higgs boson decaying into different final states as a function of Higgs mass is shown in figure 2.10.

### (BSM Higgs models)

#### 2.2.3 Diboson physics

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



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

### Diboson productions

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

### Vector boson scattering

The  $SU(2)_L \times U(1)_Y$  structure in SM predicts self-interactions between electroweak gauge bosons. Those self-couplings can involve either three or four gauge bosons at a single vertex, known as triple gauge coupling (*TGC*) and quartic gauge couplings (*QGC*), respectively. Vector boson scattering or fusion (*VBS* or *VBF*) 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 Higgs boson described in figure 2.14.

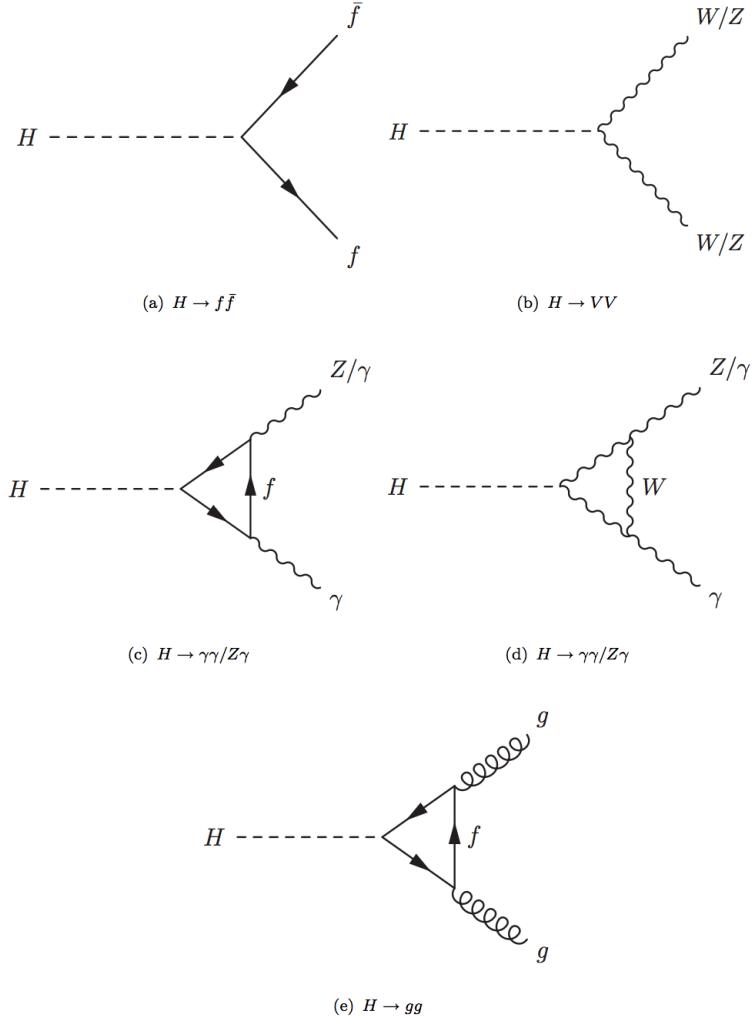
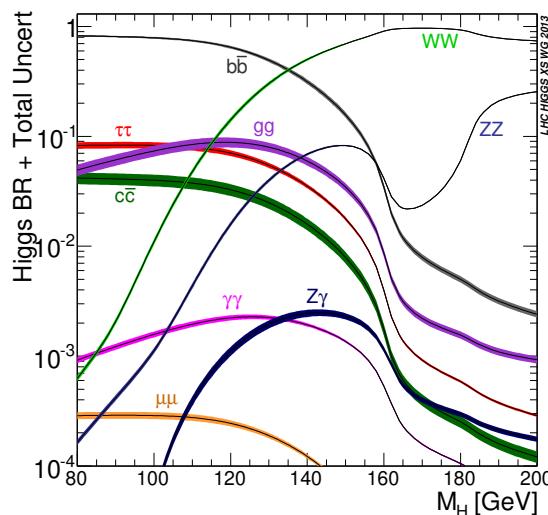
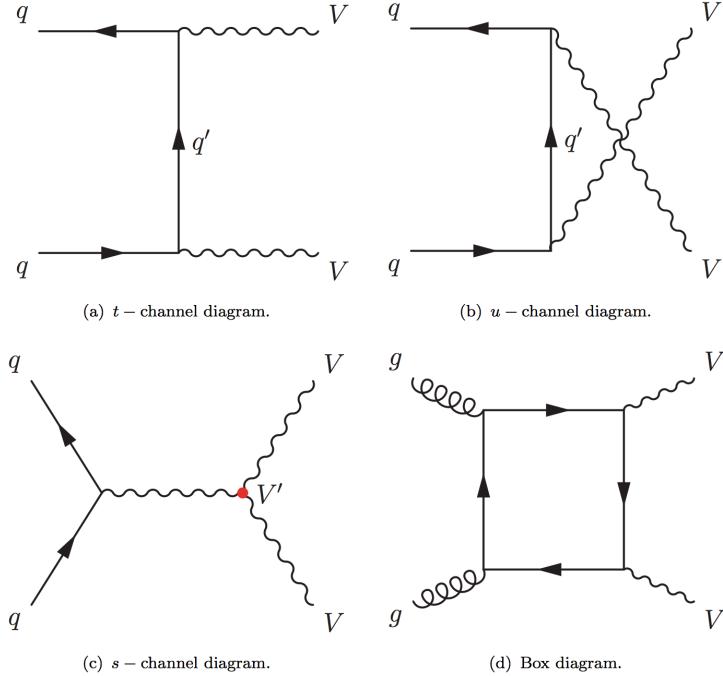


Fig. 2.9 SM Higgs decay channels.


 Fig. 2.10 Branching ratio of Higgs decays<sup>[10]</sup>.

The amplitudes of leading-order (LO) VBS can be expressed as<sup>[11]</sup>:

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



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

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

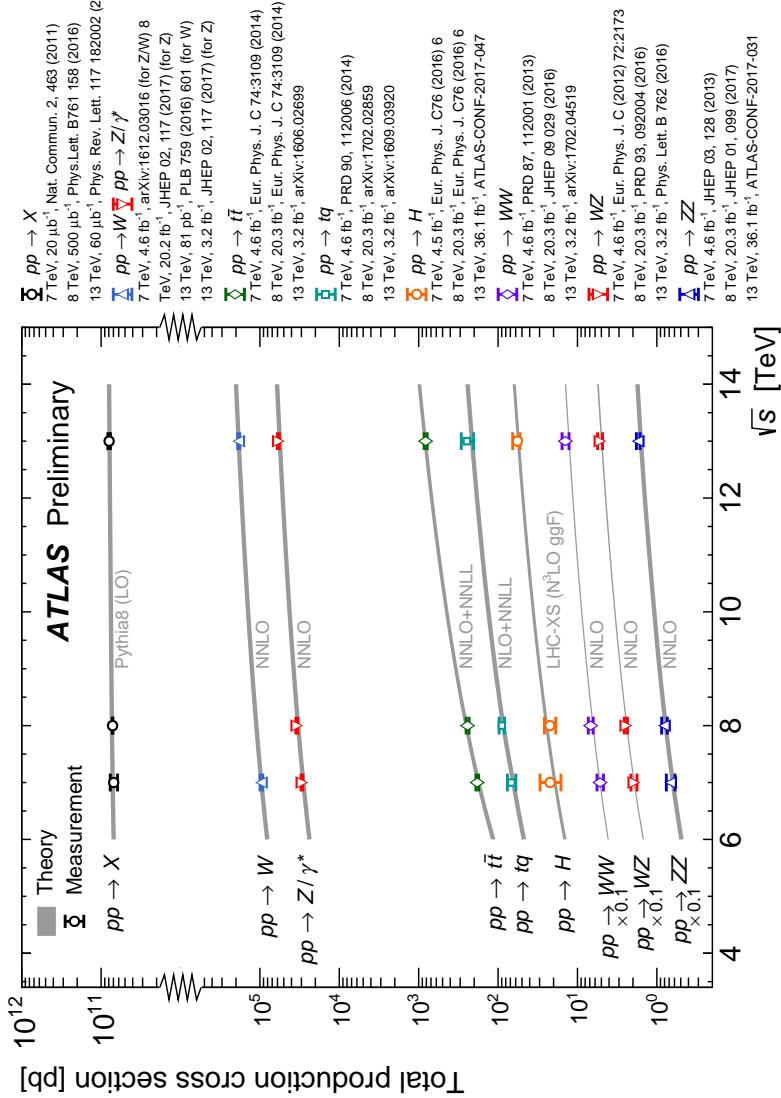
$$\begin{aligned} iM_{Higgs} &= -i \frac{C_v^2 g_1^2}{4m_W^2} \left[ \frac{(s-2m_W^2)^2}{s-m_H^2} + \frac{(t-2m_W^2)^2}{t-m_H^2} \right] \\ &\simeq -i \frac{C_v^2 g_1^2}{4m_W^2} (s+t) \end{aligned} \quad (2.29)$$

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

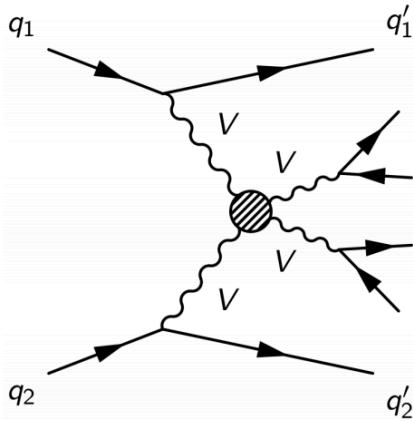
$$iM_{TGC} + iM_{QGC} = i \frac{g_1^2}{4m_W^2} (s+t) + O((s/m_W^2)^0) \quad (2.30)$$

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

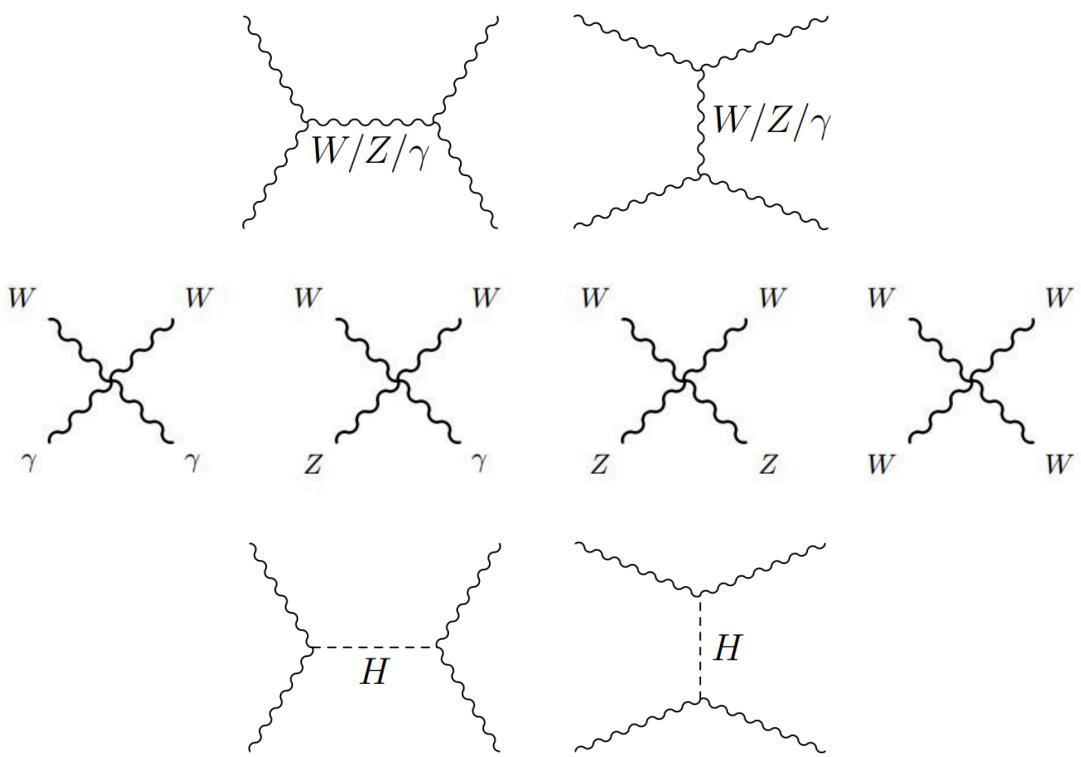
In conclusion, Higgs boson acts as "moderator" to unitarize high-energy longitudinal vector boson scattering by restoring unitarity of total amplitude in high energy region.



**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.**



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



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

# Chapter 3 The Large Hadron Collider and the ATLAS Detector

## 3.1 The Large Hadron Collider

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

### 3.1.1 Operation history and machine layout

#### Operation history

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

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

The second operation period with higher center-of-mass energy at 13 TeV started from 2015 called "run-2". And it continued to the end of 2018 with total integrated luminosity reaching about  $147 \text{ fb}^{-1}$  for ATLAS. Figure 3.1 shows the cumulative luminosity versus month delivered to ATLAS during stable beams at each years from 2011 to 2018.

#### Machine layout

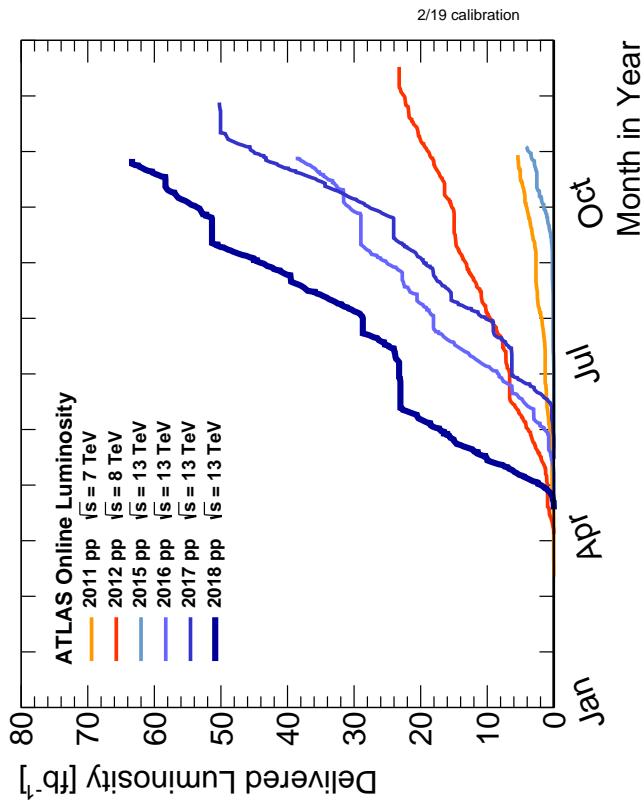
**Table 3.1 Summary of design parameters of LHC for pp collisions.**

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

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

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 the largest column detector at LHC. It is designed to search for Higgs boson, test standard 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



**Fig. 3.1 Cumulative luminosity versus time in ATLAS.**

(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

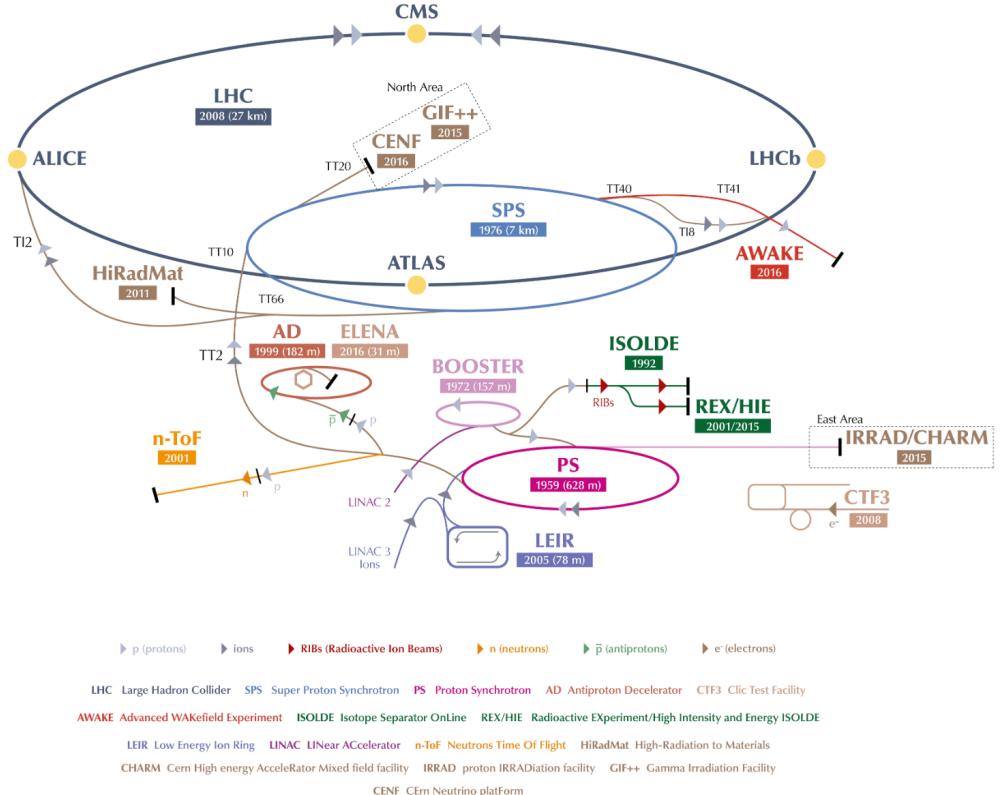
In beam-beam collisions, the event rate for a given process is given by<sup>[15]</sup>:

$$N = L\sigma \quad (3.1)$$

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

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

where  $N_b$  denotes the number of particles per bunch,  $n$  is the number of bunches per



**Fig. 3.2 CERN accelerator complex**<sup>[16]</sup>.

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 RMS (root mean square) 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 imported, which is the integral over time:

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

The unit of integrated luminosity we commonly use is  $b^{-1}$  ( $1b^{-1} = 10^{24} cm^{-2}$ ). Figure 3.3 shows integrated luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during run-2 pp collisions.

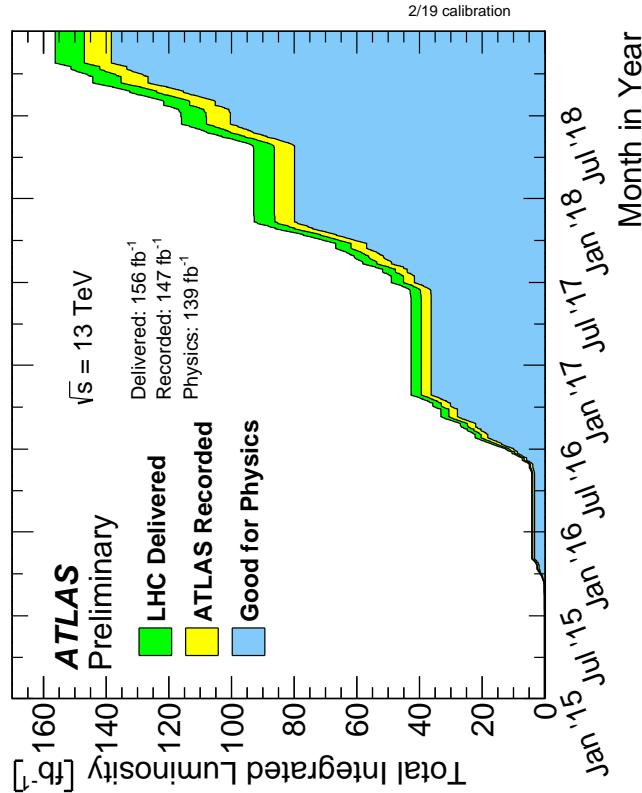


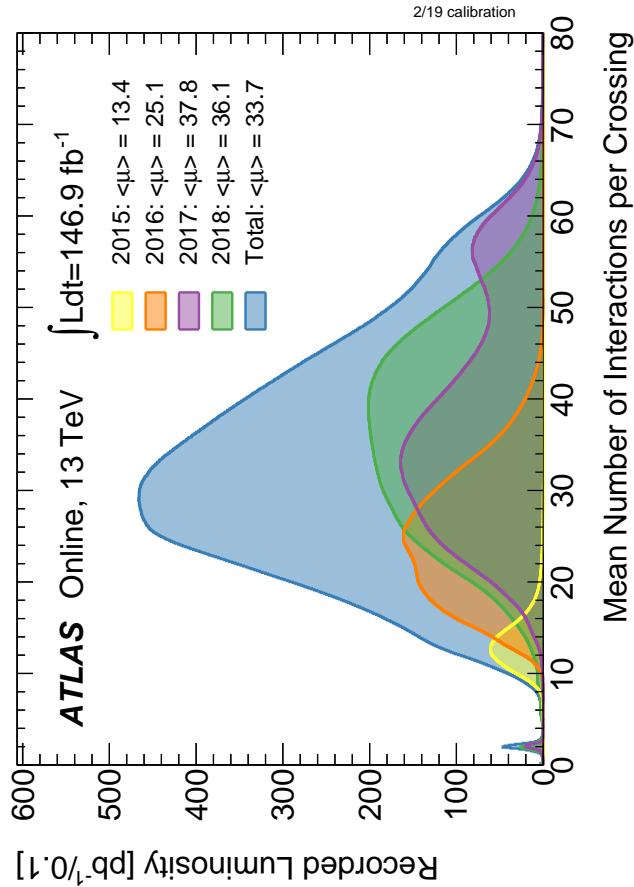
Fig. 3.3 Integrated luminosity in ATLAS.

### Pile-up

In collisions, multiple interactions can happen in one single bunch crossing, which is called "pile-up". The variable  $\mu$ , which represents the average number of interactions per bunch crossing, is defined to describe pile-up effect:

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

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



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

## 3.2 ATLAS detector

### 3.2.1 Detector overview

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

This paragraph briefly summarizes the coordinate system and nomenclature used to describe the ATLAS detector<sup>[17]</sup>. As shown in figure 3.5, we define the nominal interaction point as the origin of the coordinate system, the beam direction as the  $z$ -axis and the  $x$ - $y$  plane is transverse to the beam direction. The positive  $x$ -axis is defined to be the direction pointing to the center of LHC ring, while the positive  $y$ -axis is pointing upwards. There are two sides of detector A and C, in which A(C)-side is defined as with positive (negative)  $z$ . The azimuthal angle  $\phi$  is measured as usual around the beam axis, while the polar angle  $\theta$  is the angle from the beam axis In physics analysis, we usually use the pseudorapidity instead of  $\theta$  angle, which is designed as  $\eta = -\ln[\tan(\frac{\theta}{2})]$ .

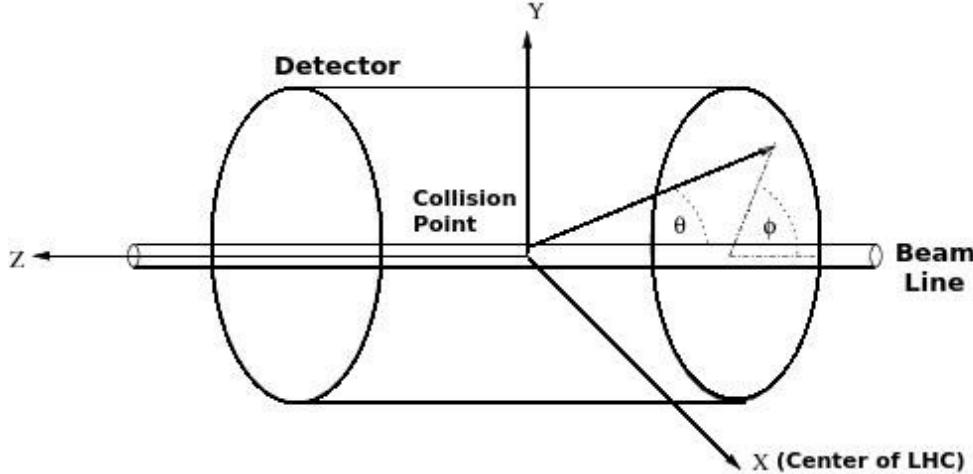


Fig. 3.5 Coordinate system used by the ATLAS experiment at the LHC<sup>[18]</sup>.

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

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

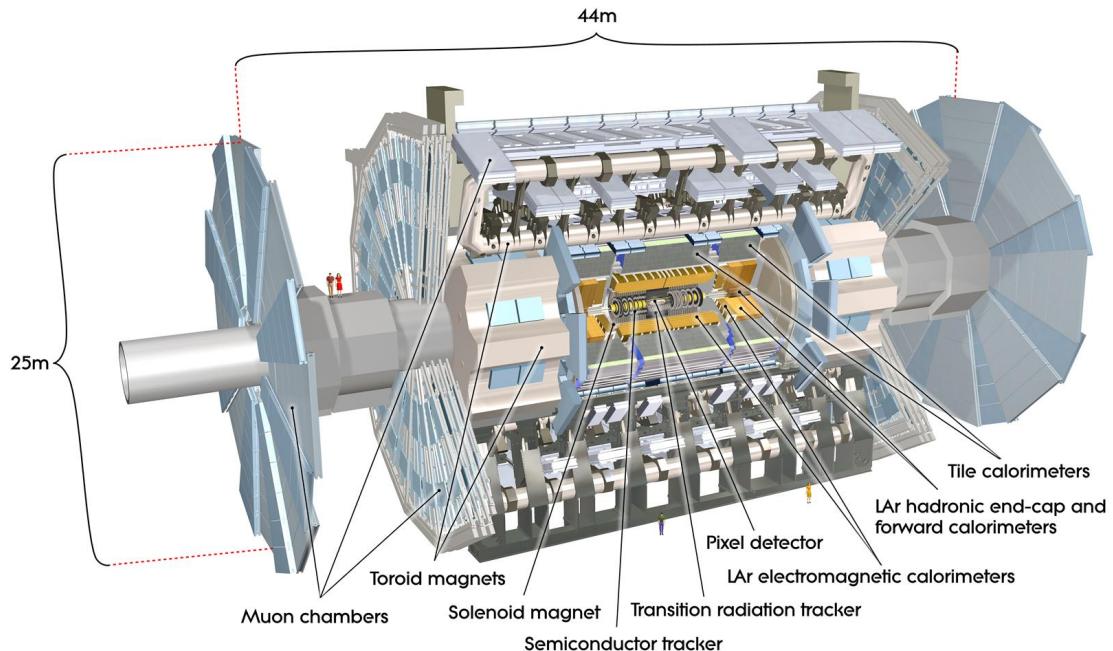


Fig. 3.6 Cut-away view of the ATLAS detector<sup>[19]</sup>.

superconducting solenoid surrounding the inner-detector cavity, and three large superconducting toroids (one barrel and two end-caps) arranged with an eight-fold azimuthal symmetry around the calorimeters.

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

**The calorimeter** is outside the inner detector, for electromagnetic and hadronic energy measurements. High granularity liquid-argon (LAr) electromagnetic sampling calorimeters is used to measure energy and position resolution with range up to  $|\eta| < 3.2$  for electrons and photons. For hadronic calorimetry, a scintillator-tile calorimeter is used in the range of  $|\eta| < 1.7$ . The LAr forward calorimeters provide both electromagnetic and hadronic energy measurements with the coverage up to  $|\eta| = 4.9$ .

**The muon spectrometer** is in the outermost side. The air-core toroid system, with a long barrel and two inserted end-cap magnets, provides strong bending power in a large volume within a light and open structure. Multiple-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 LHC. It's designed to take advantage of the unprecedented energy at LHC. The Higgs boson was discovered as one of its benchmark, and lots of precise tests and measurements of SM is on going. In the meantime, ATLAS is also designed to observe the phenomena that involve highly massive particles, such as heavy beyond standard model (BSM) gauge bosons  $Z'$  and  $W'$ . It 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 speed-fast and radiation-hard electronics are required due to the experimental conditions at 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 importantly required, since these measurements form the basis of

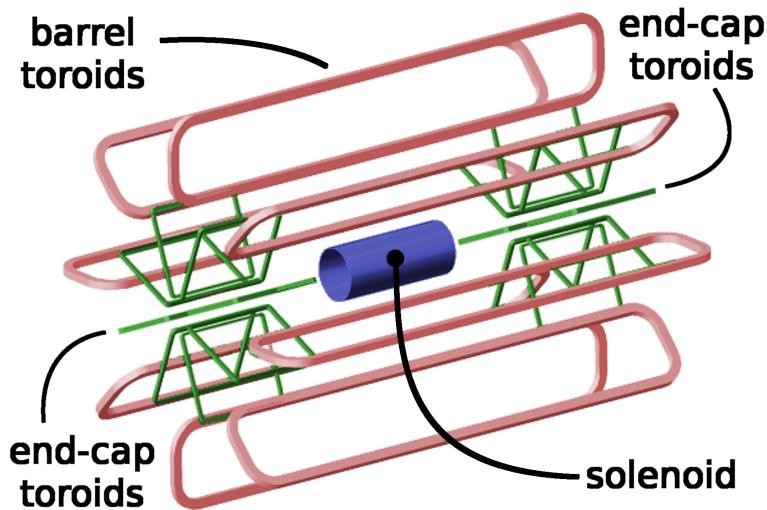
many studies.

- Good muon spectrometer is also required for muon identification and momentum resolution 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>[20]</sup>. The total magnet system is 22m in diameter and 26m in length as shown in figure 3.7.



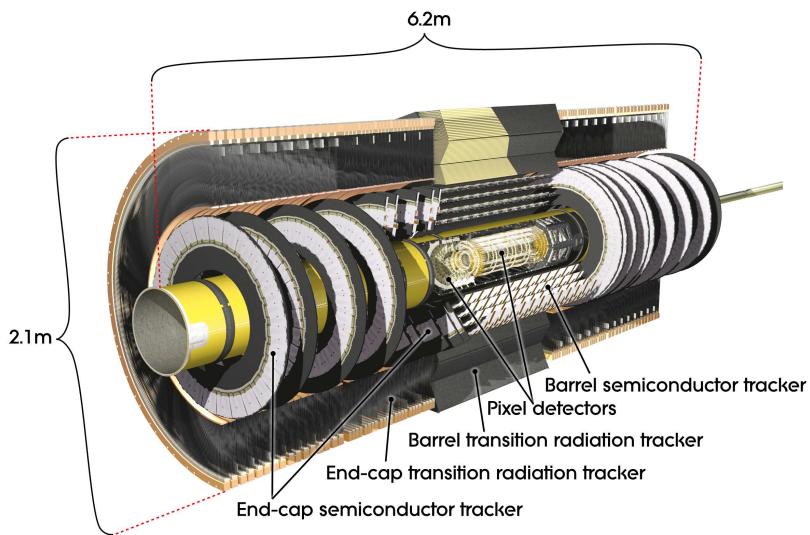
**Fig. 3.7 Schematic diagram of the ATLAS magnet system.**

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

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

### 3.2.4 Inner detector

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



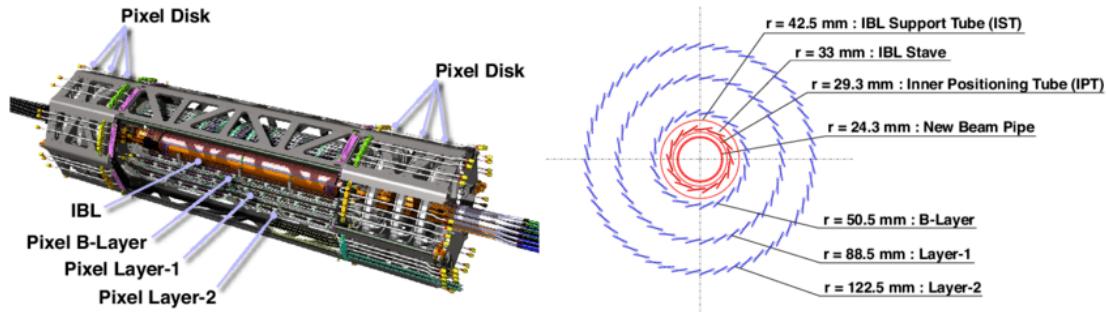
**Fig. 3.8 Schematic diagram of the ATLAS inner detector<sup>[21]</sup>.**

#### Pixel detector

The pixel detector is the innermost part of ATLAS tracking system. With finest granularity of materials, it has the best spatial resolution and 3-dimensional space-point measurement in inner detector. ATLAS Pixel Detector for LHC Run-2 is composed of 4 layers of barrel pixel detector and two end-caps with three pixel disks each, as shown in figure 3.9. There are three outer layers that originally installed for run-1 and one additional layer called Insertable B-Layer (IBL) that newly constructed in run-2<sup>[22]</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.

#### Semiconductor Tracker

The Semiconductor Tracker (SCT) is the middle component of the inner detector that outside the pixel 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, it contains four concentric layers



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

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, it have accuracies of  $17 \mu\text{m}$  in the  $z$ - $\phi$  coordinate and  $580 \mu\text{m}$  in  $r$  coordinate.

### Transition radiation tracker

The transition radiation tracker (TRT)<sup>[24]</sup> is the outermost part of inner detector. It has a very different design with the two previously sub-detectors. It's composed of thin-walled drift tubes called straw, also in three parts: a barrel and two end-cap regions. There are 73 barrel layers and 224 end-cap layers (112 in each) with 372000 straws in total, and about 351000 readout channels for TRT. The TRT procides better  $z$  resolution but much worse  $r$ - $\phi$  resolution (about  $130 \mu\text{m}$ ) compared to the pixel detector and SCT per straw. But the straw hits still make significant contributions to momentum measurement, since its lower precision per point (compared to silicon) can be 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. As mentioned in overview section, there are two basic calorimeter systems: an inner electromagnetic (EM) calorimeter and an outer hadronic calorimeter. The EM calorimeter is designed for precise measurements for electrons and photons, so that with fine granularity; while the hardronic one with relative coarser granularity but satisfied the physics requirements for jets reconstructions and

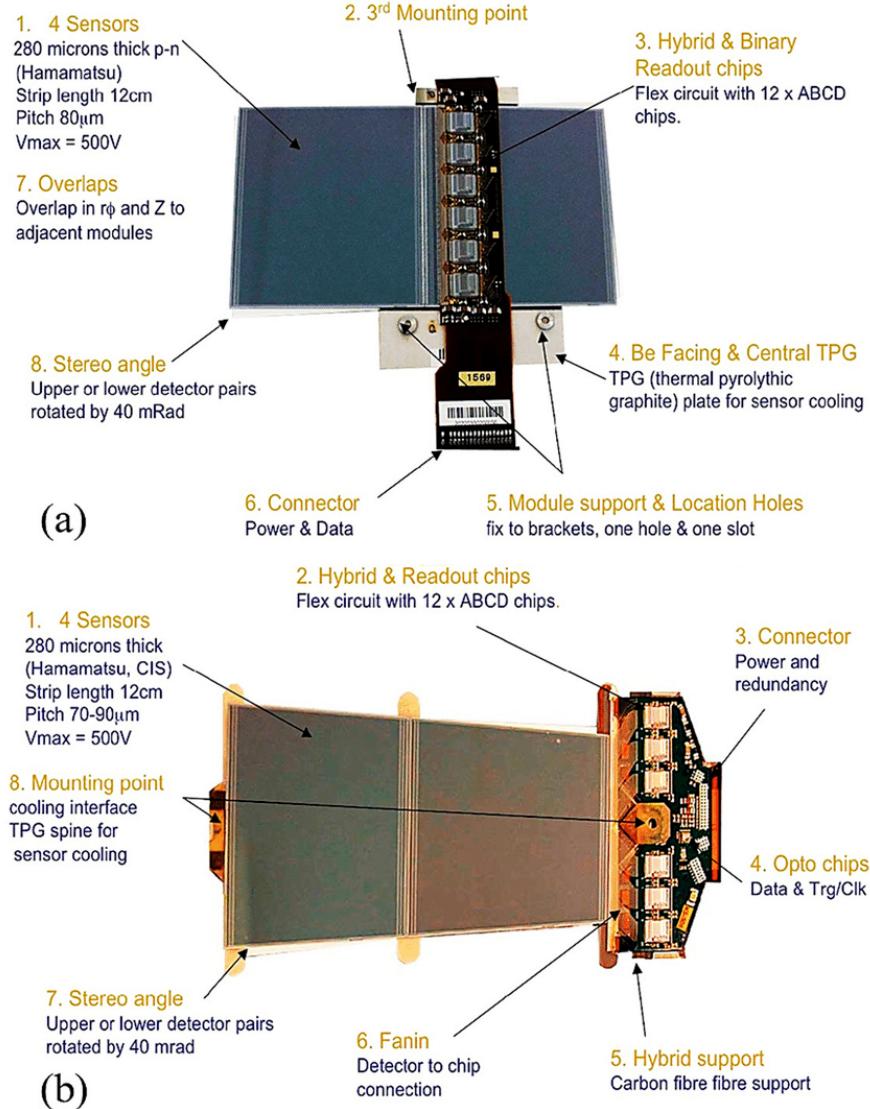
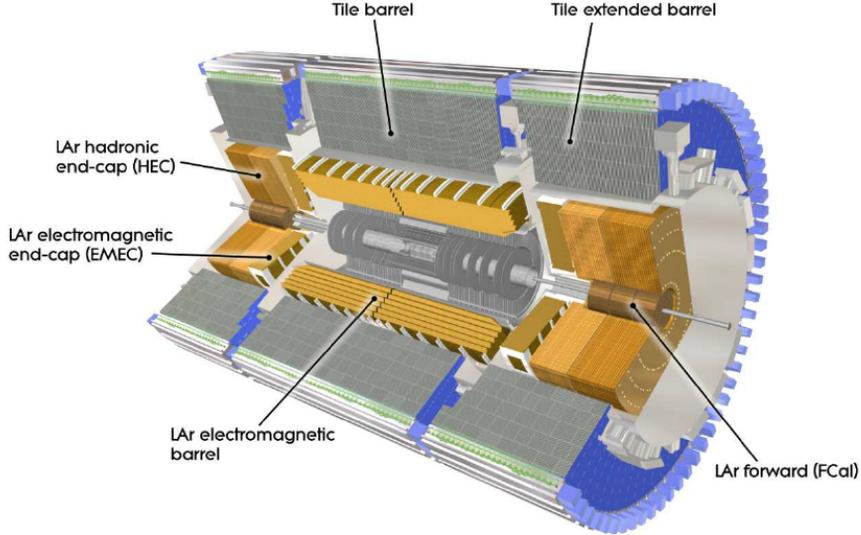


Fig. 3.10 SCT (a) barrel module and (b) end-cap<sup>[23]</sup>.

$E_T^{miss}$  measurements. Two different sampling techniques are used, the EM calorimeter is purely based on liquid-argon (LAr) technology, hadronic calorimeter use both LAr and scintillating tiles calorimeters. More details are described as belows.

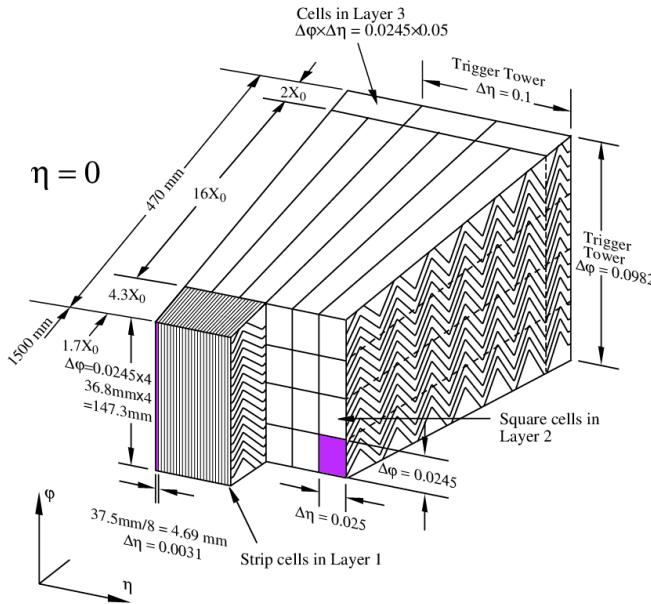
### Liquid Argon calorimeter

The LAr calorimeter is the one uses liquid-argon as active medium. The Liquid Argon 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 from  $|\eta| = 1.4$  up to the acceptance limit  $|\eta| = 4.9$ <sup>[26]</sup>. Figure 3.12 shows the shape of a barrel module as accordion geometry. For barrel EM calorimeter, the absorbing material is lead-liquid argon, while the hadronic end-cap calorimeter use copper plates as the absorbing material. In addition, the forward calorimeter is splitted



**Fig. 3.11** Cut-away view of the ATLAS calorimeters. The LAr calorimeters are seen inside the scintillator- based Tile hadronic calorimeters<sup>[25]</sup>.

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

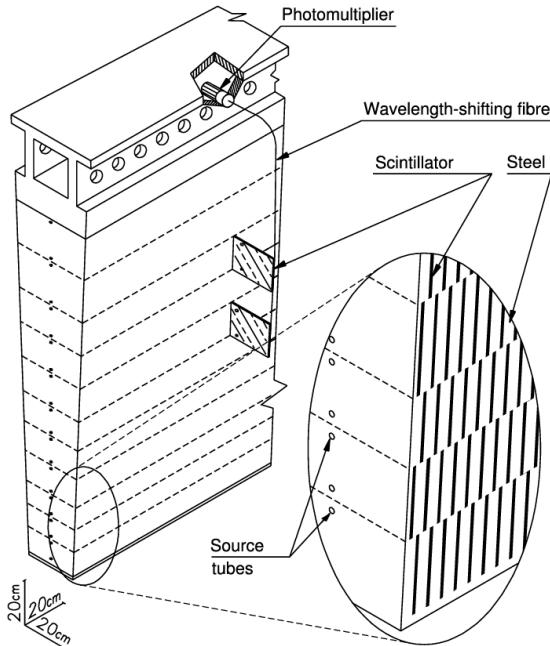


**Fig. 3.12** Schematic diagram of a LAr EM calorimeter barrel module.

### Tile calorimeter

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

hadronic calorimeter.



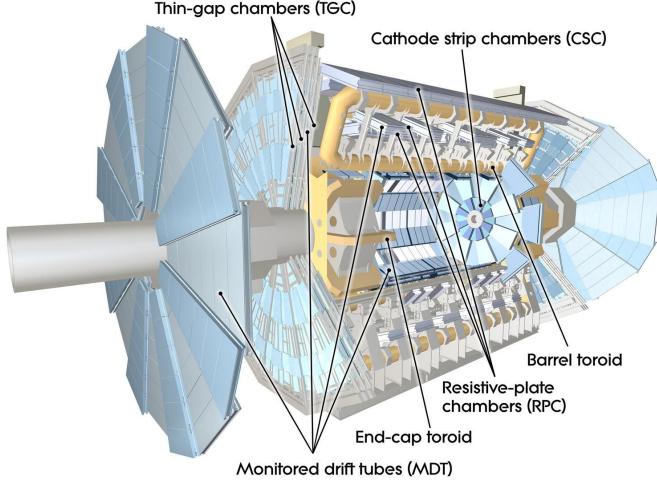
**Fig. 3.13 Schematic diagram of tile calorimeter module<sup>[27]</sup>.**

### 3.2.6 Muon spectrometer

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

The details of four-type chambers are given as belows:

- **Monitored Drift Tubes (MDT).** MDTs provide the precise momentum measurement with the  $|\eta|$  range up to 2.7, except in the innermost end-cap layer where the coverage is limited to  $|\eta| < 2.0$ . The chambers are comprised of three or eight layers of drift tubes, with a diameter of 29.970 mm, operated with Ar/CO<sub>2</sub> gas (93/7) at 3 bar. The average resolution can reach 80  $\mu\text{m}$  per tube and 30  $\mu\text{m}$  per chamber.
- **Cathode strip chambers (CSC).** CSCs are used in the forward region of  $2 < |\eta| < 2.7$  in the innermost tracking layers, due to their good time resolution and high rate capability. The CSCs are multi-wire proportional chambers (MWPC)



**Fig. 3.14 Cut-away view of the ATLAS muon spectrometer<sup>[28]</sup>.**

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.

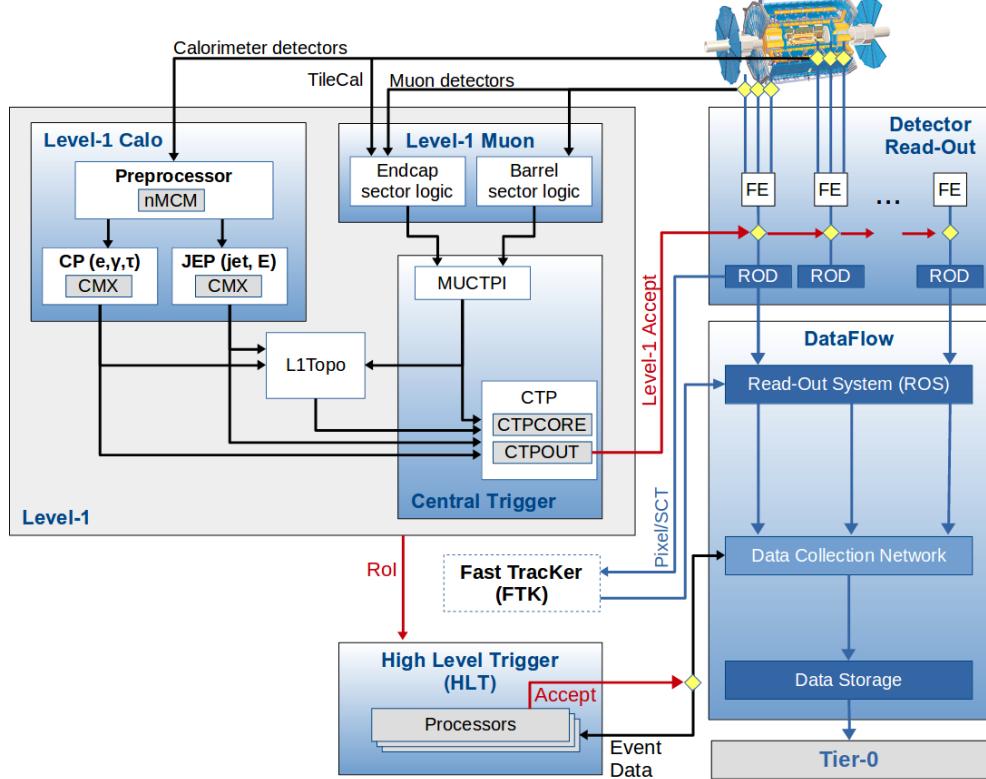
- **Resistive plate chambers (RPC).** The RPCs serve as fast triggers in the barrel region of  $|\eta| < 1.05$  due to the high rate capability and good spatial and time resolution. It is a gaseous parallel electrode-plate detector without any wires. There are three concentric cylindrical layers around the beam axis, referred to as the 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 operated 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 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 cause to a even larger challenge and more strict requirement of trigger system.

The trigger system in run-2 is comprised of a hardware-based first level trigger (Level-1) and a software-based high level trigger (HLT)<sup>[29]</sup>. As depicted in figure 3.15, in Level-1, the inputs from coarse granularity calorimeter and muon detector informa-

tion together with some other subsystems are sent to the Central Trigger Processor to determine Regions-of-Interest (RoIs) in the detector. The events rate can be reduced by



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

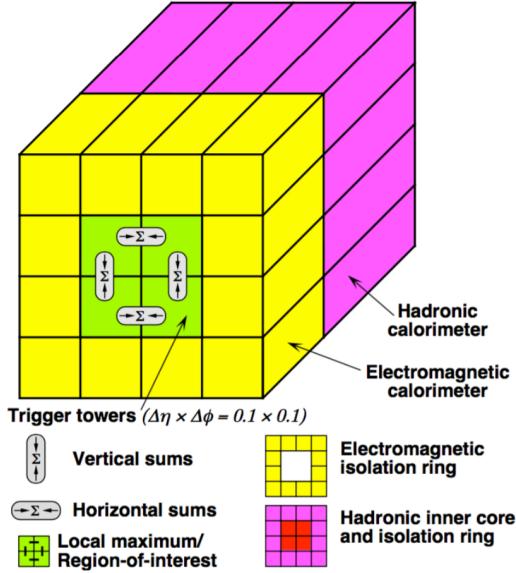
Level-1 triggers from 30 MHz to 100 kHz. After that, the RoI information from Level-1 is sent to HLT, in which more sophisticated selection algorithms are run for regional reconstruction. The HLT reduces the rate from Level-1 of 100 kHz to about 1 kHz on average. At the end, the events that accepted by HLT are transferred to local storage at experimental site for offline reconstruction. Details about Level-1 and HLT trigger systems will be described as belows.

### 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, allows the trigger rate increasing from 70 kHz (run-1) to 100 kHz (run-2). As mentioned above, there are two major parts of Level-1 triggers, which include 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 is that, the new FPGA-based



**Fig. 3.16** An examples of L1 calorimeter trigger tower for electron and photon triggers<sup>[30]</sup>.

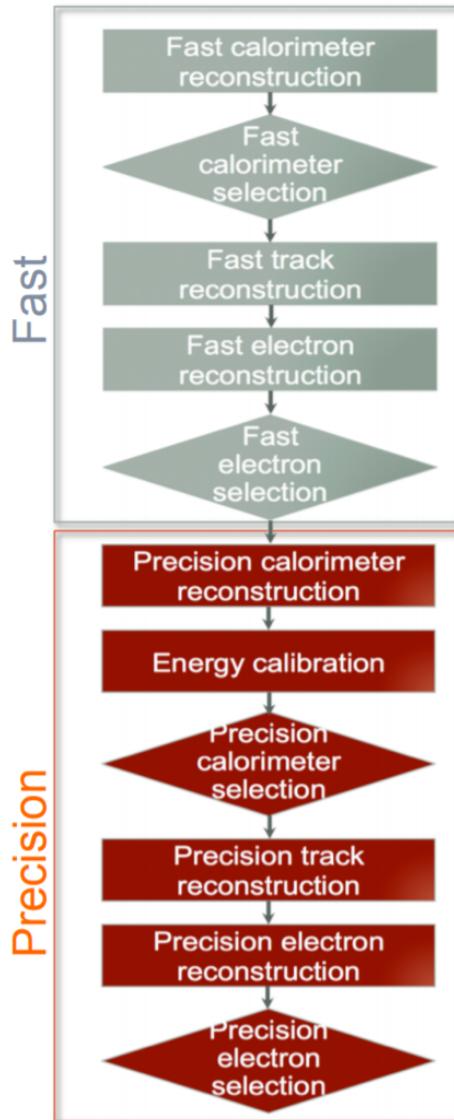
(field-programmable gate array) Multi-Chip Modules are used to replace the ASICs (application-specific integrated circuits) included in the modules used in run-1, which allows the use of auto-correlation filters to suppress pile-up.

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  $1.3 < |\eta| < 1.9$  while only loss around 2% signal efficiency. In addition, the coverage is extended by around 4% due to installing new chambers in the feet region of the muon detector.

### High Level Trigger

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

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



**Fig. 3.17** The HLT trigger algorithm sequence<sup>[30]</sup>.

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

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

### 4.1 Event simulation

The ATLAS simulation program is integrated into the ATLAS software framework called *Athena*<sup>[31]</sup>, which uses Python as an object-oriented scripting and interpreter language to configure and load C++ algorithms and objects. Figure 4.1 shows the overview of ATLAS simulation data flow<sup>[32]</sup>. In the diagrams, the square-cornered boxes represent algorithms and applications to be run and round-cornered boxes denote data objects.

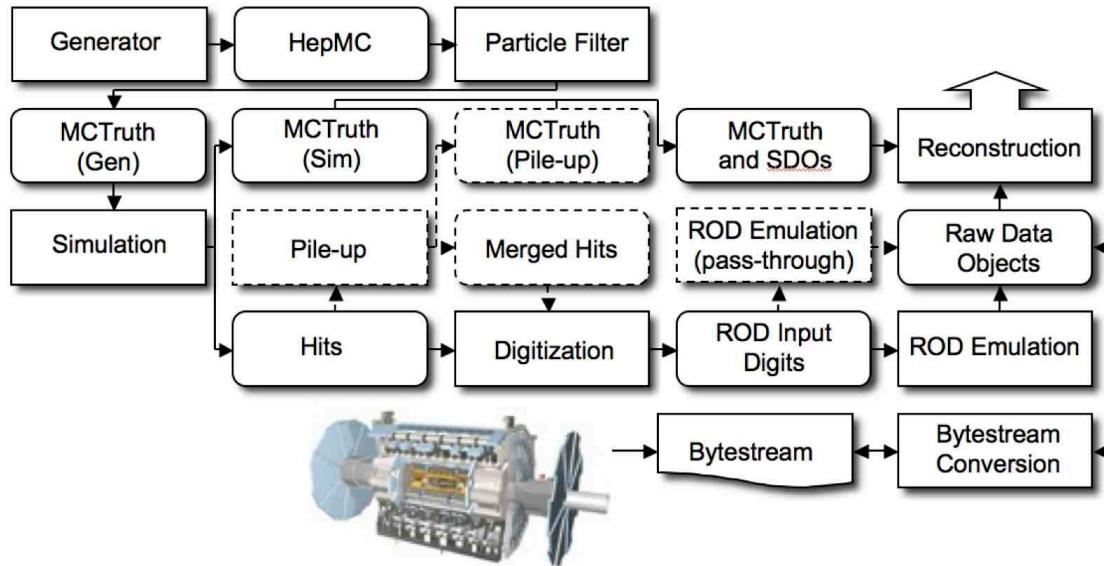


Fig. 4.1 The flow of the ATLAS simulation software.

First of all, events are produced by MC generators in standard HepMC format and then read into the simulation. During the simulation, particles are propagated through the full ATLAS detector whose configurations can be set by users via GEANT4 toolkit. The energies deposited in the sensitive regions of the detector are recorded as *hits*, which

contains 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 below.

### Event generation

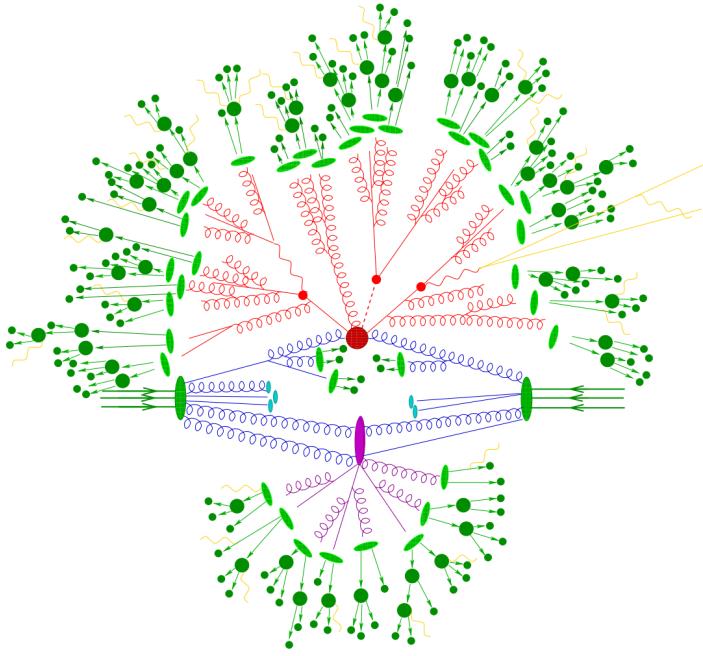
As shown in figure 4.2<sup>[33]</sup>, at hardon colliders, multiple scattering and rescattering effects arise, which must be simulated by Monte Carlo (MC) event generators to reflect the full complexity of those event structures. Several MC event generators can be used to generate events originally 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) but 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, which include Sherpa<sup>[34]</sup>, Herwig++<sup>[35]</sup>, PowhegBox<sup>[36]</sup>, MC@NLO<sup>[37]</sup> and Pythia8<sup>[38]</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 cuts 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 and also outgoing particles. The output of GEANT4 is called *hit file*, which contains metadata describing the configuration of the simulation during the run, all truth information requested 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



**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 indicate the parton-to-hardon transitions and the dark green blobs represents hadron decays. The yellow lines are soft photon radiations.

simulation programss are developed according to different user purpose. Some widely-used fast-sim toolkits include *Fast G4 Simulation*<sup>[39]</sup>, *ATLFAST-I*<sup>[40]</sup> and *ATLFAST-II*<sup>[41]</sup>.

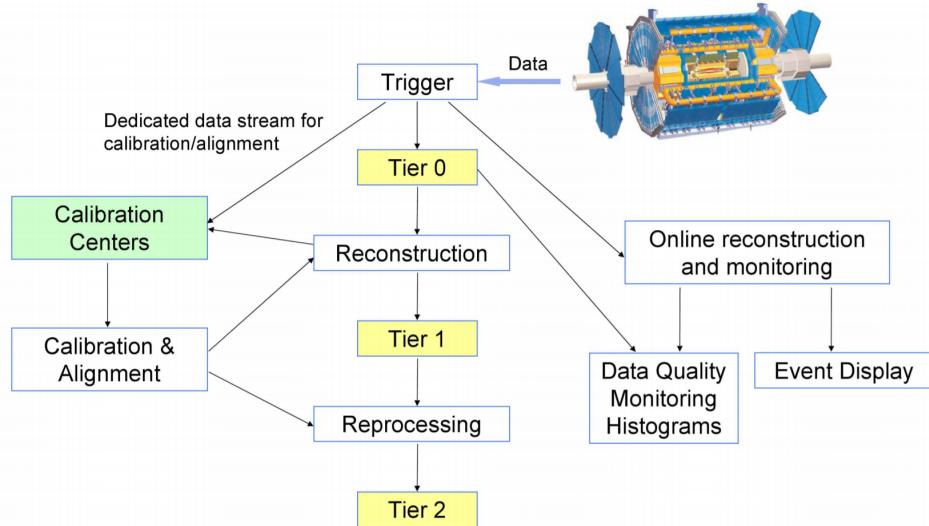
### Digitization

The hit outputs from simulated events, including hard scattering signal, minimum bias, beam halo, beam gas and cavern background events, are then send into digitization to convert into detector response called "digits". Before converting into detector signal with 'digits' formart, each type of events can be overlaid at a user-specified rate. Those overlay, called "pile-up", can be done during digitization to save the CPU time in simulation level. At this stage, the detector noise and the first level trigger that implemented with hardware on the real detector are added into events. The digitization firstly constructs "digits" inputs to the read out drivers (RODs) in the detector electronics. The ROD functionality is then emulated, and the output digits are written out as Raw Data Object (RDO) file. In addition, the digitization algorithms can also produce Simulated Data Objects (SDOs), which contain information about all the particles, noise and the amount of energy that contributed to the signal. Then all information are sent

into reconstruction that will be described in next subsection.

## 4.2 Event reconstruction

The data flow of ATLAS data processing is sketched in figure 4.3<sup>[42]</sup>. Data from detector is firstly filtered by online trigger system and then send to the *Tier-0* (*T0*) for initial processing by offline reconstruction software also based on Athena. A small amount of data named "express stream" is processed in almost real time in *T0* for online data quality monitoring. In addition, some other dedicated data streams are sent out at trigger level for detector alignment and calibration. These calibration and alignment information are then used for bulk reconstruction in *T0*. At the end of the reconstruction chain, the data are delivered into *Tier-1* (*T1*) and *Tier-2* (*T2*) centers for further analysis and production of simulated data. Besides, *T1* centers are also responsible for data reprocessing by re-running data reconstruction with improved calibration and alignment constants and with improved reconstruction algorithms. This section describes the reconstruction of some important physics objects in ATLAS experiment, i.e. tracks, vertices, electrons, muons, jets, and missing energies.



**Fig. 4.3** The flowchart of the ATLAS data processing.

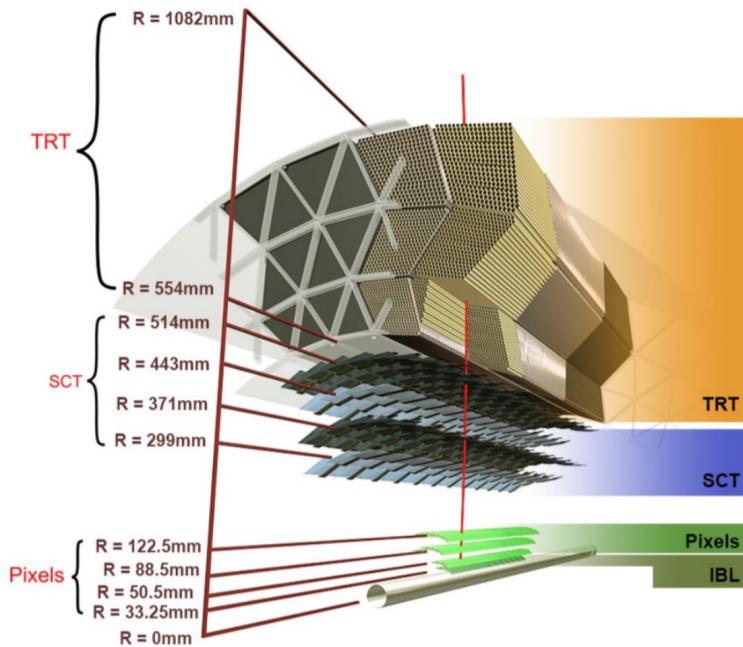
### 4.2.1 Track

The ATLAS detector is composed of two independent tracking systems: the Inner Detector (ID) close to the interaction point, and the Muon Spectrometer (MS). The reconstructed charged-particle trajectories in the ID and MS are referred to as ID tracks and MS tracks respectively. The ID reconstruction needs to handle high track density

that imposes a large number of combinatorial track candidates, the MS reconstruction is however largely limited by the huge amount of inert material, the large background and the highly inhomogeneous magnetic field<sup>[43]</sup>. More details of two types of track reconstructions are given below.

### Inner detector track

Figure 4.4 shows the ID system used for detect charge-particle tracks. The ID 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.**

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

Another complementary approach, outside-in, searches for unused track segments start from TRT instead. These segments are then extended into the SCT and pixel detectors to improve the tracking efficiency for secondary tracks from conversions or decays

of long-lived particles.

### Muon spectrometer track

The MS track reconstruction<sup>[44]</sup> starts from searching hit patterns inside each muon chamber to form segments. In each MDT chamber and nearby trigger chamber, a Hough transform<sup>[45]</sup> is used to search the hits lies on a certain trajectory in the bending plane of the detector. The MDT segments are reconstructed by performing a linear fit to the hits found in each layer. The RPC or TGC hits can be built by measuring the coordinate orthogonal to the bending plane. And the segments of CSC can be built using a separate combinatorial search in the  $\eta$  and  $\phi$  detector planes.

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

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

### 4.2.2 Primary vertex

The primary vertex (PV) is reconstructed by using the reconstructed tracks introduced in previous section as inputs. The tracks to be considered for vertex reconstruction must satisfy the following criterias<sup>[46]</sup>:

- $p_T > 400 \text{ MeV}$
- $|\eta| < 2.5$
- Number of silicon hits  $\geq \begin{cases} 9 & \text{if } |\eta| \leq 1.65 \\ 11 & \text{if } |\eta| > 1.65 \end{cases}$

- IBL hits + B-layer hits  $\geq 1$
- A maximum of 1 shared module (1 shared pixel hit or 2 shared SCT hits)
- Pixel holes = 0
- SCT holes  $\leq 1$

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

The reconstruction of PV can be divided into two steps<sup>[47]</sup>: vertex finding and vertex fitting. The first step represents the pattern recognition process, namely the association of reconstructed tracks to vertex candidates. The latter one works on the reconstruction of the actual vertex position and its covariance matrix. More details are described as below:

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

After determining the seed position, the iterative primary vertex finding procedure starts. An adaptive vertex fitting algorithm is used to find the optimal vertex position by using an iterative  $\chi^2$  minimization. The seed position is used as start point and the reconstruction tracks are used as input measurements. The input tracks are assigned weights to reflect their compatibility with the vertex estimation, and the vertex position is calculated based on the weighted tracks. It's the iterative procedure, in each iteration, the less compatible tracks are down-weighted and then the vertex position is recomputed based on the reweighted tracks.

After the last iteration, the final weight of each track used in vertex fit is estimated. And based on their final weights, the incompatible tracks are then rejected from this vertex candidate and moved back to the unused pool for next determination of another vertex. Then the procedures describes above are repeated again, until no unassociated tracks are left or no additional vertex could be found in remaining tracks.

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

### 4.2.3 Electron

Many of the interesting physical processes with the involvement of one or more electrons (or positrons) at LHC. But these electrons can be subjected to large amount of backgrounds such as hardrons, non-prompt electrons from photon conversions and non-isolated electrons from heavy flavor hadron decays. It is therefore essential to efficiently reconstruct and identify electrons and in the meantime to keep high background rejection.

In ATLAS, in central region, the electrons leave tracks in inner detector (ID) and the energy deposite in the electromagnetic (EM) calorimeter. Firstly the signals from calorimeter are used for L1 trigger system, and them combined with the information from ID tracks to reconstruct electron candidates that will be used for the high level trigger (HLT) decision algorithms<sup>[48]</sup>. The backgrounds mentioned above can then be further suppressed by using several identification criterias. In addition, electrons are required to be isolated from other activities to be further distinguished from background.

More details of electron *reconstruction*, *identification* and *isolation* will be described as below.

#### Electron reconstruction

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

1. **Seed-cluster reconstruction:** A sliding window with size of  $3 \times 5$  in unit of  $\Delta\eta^{tower} \times \Delta\phi^{tower} = 0.025 \times 0.025$  in  $\eta \times \phi$  space is utilized to saerch for electron cluster seeds with total cluster transverse energy greater than 2.5 GeV. Then a clustering algorithm<sup>[49]</sup> is applied to form the clusters around the seeds, which can take advantage of removing the depllications. The kinematics of clusters are then reconstructed by using an extended window depending on the cluster position. The efficiency of cluster search is from about 95% at  $E_T = 7GeV$  to 99% for  $E_T \geq 15GeV$ .
2. **Track reconstruction:** The track reconstruction can be divided into two steps: pattern recognition and track fit. The standard pattern recognition in ATLAS uses pion hypothesis for energy loss caused by interactions with detector material. If a track seed with  $p_T > 1GeV$  cannot be successfully extended to a full track required at least seven hits using this pion hypothesis, but still falls inside one of the EM cluster region of interest, as a second attempt, the pattern recognition using electron hypothesis is then used to allow larger energy loss. Depending on

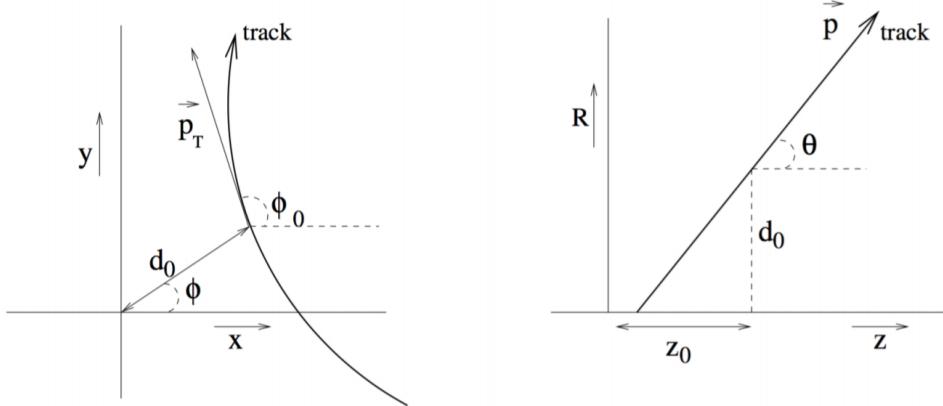
the pattern that has been used in previous stage, the track candidates are then fitted with either the pion hypothesis or the electron hypothesis by using ATLAS Global  $\chi^2$  Track Fitter<sup>[50]</sup>. If a track candidate fails the fit by using pion hypothesis, it can be refit with the electron hypothesis again. In this method, a specific electron-oriented algorithm is integrated into the ATLAS standard track reconstruction, which improves the performance for electron and as well as maintain minimal interference with the main track reconstruction.

3. **Electron specific track fit:** Once the tracks are obtained, they are loosely matched to EM cluster using the distance in  $\eta$  and  $\phi$  between the position of track (after extrapolation) in calorimeter's middle layer and the cluster barycentre. The matching conditions take into account the energy loss of bremsstrahlung and the number of precise hits in silicon detector.
4. **Electron candidate reconstruction:** The electron candidate is reconstructed by matching the track candidate to EM cluster seed to eventually completes the electron reconstruction procedure. If more than one track satisfy the matching condition, one track is chosen as primary track based on the information of the cluster-track distance  $R$ , the number of pixel hits and the presence of a hit in the first silicon layer<sup>[51]</sup>. In addition, we remove the electron candidates mentioned above but without any associated precise hit tracks from electron pool and move them into photon candidates. Then we re-formed the electron clusters by using  $3 \times 7$  ( $5 \times 5$ ) longitudinal towers of cells in barrel (endcaps) in EM calorimeter. The measured energy is calibrated to original electron energy based on MC simulated samples by using multivariate techniques (MVA).

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

### Electron identification

The electron identifications (ID) are applied to determine the reconstructed electron candidates are more signal-like or background-like objects. The ID algorithms make use of quantities of related variables from electron cluster and track measurements in-



**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>[52]</sup>.

cluding calorimeter shower shapes, track properties, as well as variables measuring bremsstrahlung effects for distinguishing signal from background. Taking the advantage of new IBL in run-2, the number of hits in this innermost pixel layer is utilized for discriminating between electrons and converted photons. In addition, a likelihood method based on the TRT high-threshold hits is adopted to compensate the lower transition radiation absorption probability of the argon.

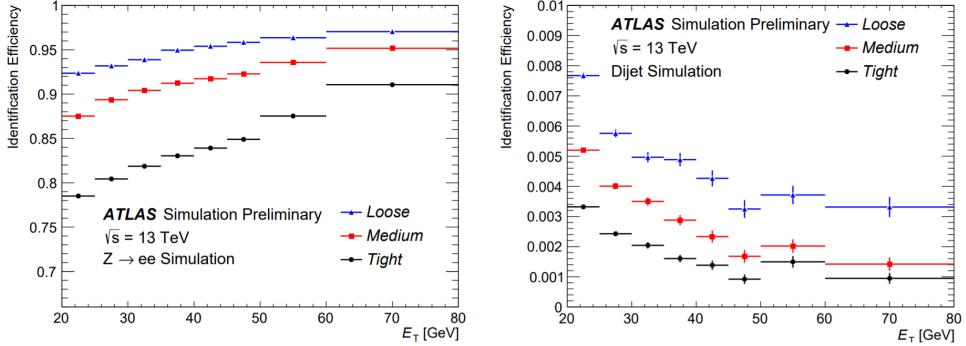
The baseline ID algorithm introduced for ATLAS run-2 data analysis is the likelihood-based (LH) method, which use a MVA technique to simultaneously evaluate several properties of electron candidates when making a decision. The LH method utilizes the probability density functions (PDFs) of signal and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

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

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

Three levels of working points (WPs) for electron ID are provided: *Loose*, *Medium* and *Tight*, in order of increasing background rejection. Samples selected by a looser WP are subsets of a tighter one, eg. the electrons passing Medium can all be selected by Loose. The ID efficiency varies as function of electron energy ( $E_T$ ) as shown in figure 4.6. For evaluations, the electron candidates from MC simulation of  $Z \rightarrow ee$  decays (dijet) is used as signal (background). Depending on the working point, the signal (background) efficiencies for reconstructed electron candidates at  $E_T = 25 GeV$

are in the range from 78 to 90% (0.3 to 0.8%), and increase (decrease) with  $E_T$ .



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

### Electron isolation

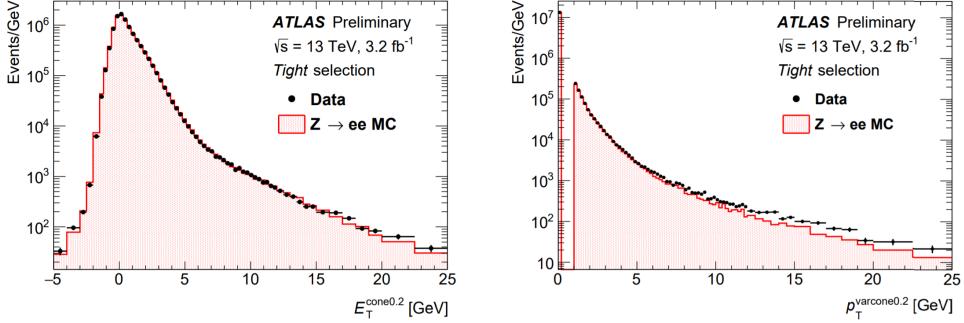
In addition to the ID criteria, most analyses have electron isolation requirement to further distinguish signal from background. To quantify the energy of particles around the electron candidate, the isolation variables can help to separate the prompt electron from other, non-isolated electron candidates,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

- **Calorimeter-based variable:**  $E_T^{cone20}$ . It's defined as the sum of transverse energies of topological clusters<sup>[53]</sup>, calibrated at EM scale within a cone of  $\Delta R = 0.2$  around the candidate electron cluster. It only consider the clusters with positive reconstructed energy. Besides, a correction as a function of  $(E_T, \eta)$  values is then applied to account for the electron energy leakage outside the cluster.
- **Track-based variable:**  $p_T^{varcone20}$ . It's calculated as the sum of all transverse momentum of all satisfied tracks within a cone of  $\Delta R = \min(0.2, 10\text{GeV}/E_T)$  around the candidate electron track. For the sum calculation, it requires the tracks are originating from the reconstruction PV of hard collision, and exclude the associated tracks of electron itself.

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

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

The distribution of two discriminating variables are shown in figure 4.7 for  $Z Z \rightarrow ee$  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 LHC which include the Higgs analyses, SM measurements and 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 detectors are 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>[44]</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, if 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;
- Normalized  $\chi^2$  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 below include<sup>[54]</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 estimate the significance of a change in trajectory along the track, expected in the presence of a hadron decaying to a muon.
- *Scattering curvature significance (SCS)* is defined as the normalized integral of the scattering angle significances, corrected for large kinks along the trajectory.

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

- *Loose*: this working point is designed to maximize the reconstruction efficiency while keeping good-quality of muon tracks. And they are specifically developed for reconstructing Higgs boson candidates from four-lepton final states. All four muon types are used for this selection level. The CB and ME muons passing Medium WP that will be mentioned below are all included into Loose category. In addition, the CT and ST muons are restricted to  $|\eta| < 0.1$  region. In the range of  $|\eta| < 2.5$ , around 97.5% Loose muons are CB muons, and about 1.5% are CT while remaining 1% are ST muons.
- *Medium*: this working point is the default criteria of muon identification in ATLAS. This selection minimizes the systematic uncertainties of muon reconstruction and calibration. In this category, we only use CB and ME muons. For CB muons, at least 3 hits in at least two layers of MDT is required, except  $|\eta| < 0.1$  region, in which tracks with  $\geq 1$  MDT layer but  $\leq 1$  MDT hole layer are allowed. For ME muons, at least 3 MDT/CSC layers is required. Furthermore, a loose cut on the compatibility between measured momentum in ID and MS is applied to reduce the fake muons from hadrons misidentification. Besides, the q/p-significance is required to be less than 7.
- *Tight*: this working point is used to maximize the purity of muons but with sacrifice of some selection efficiency. Only CB muons with hits in  $\geq 2$  stations of MS and passing Medium criteria are selected. In addition, the normalized  $\chi^2$  of combined track fit should be smaller than 8. Then, a two-dimensional cut of q/p-significance and  $\rho'$  is adopted as a function of muon  $p_T$  to ensure tighter background rejection for momentum below 20 GeV, in which the fake rate is usually higher.

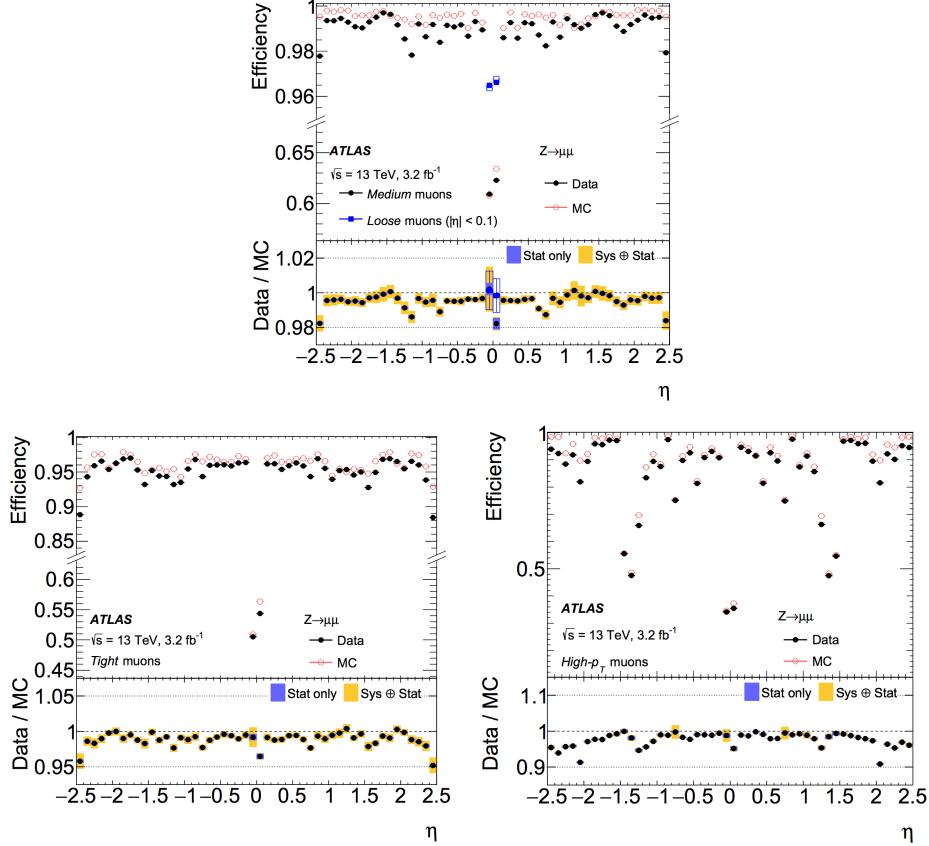
- *High- $p_T$* : this set of selections aims to maximize the momentum resolution for tracks with  $p_T > 100\text{GeV}$  region. The selection is especially optimized for searching high-mass  $Z'$  and  $W'$  resonances. CB muons satisfying Medium selection and with  $\geq 3$  hits in 3 MS stations are chosen. The specific region in MS where alignment is suboptimal are removed as a precaution.
- *Low- $p_T$* : this type of muon is newly designed for physics analyses with ATLAS software release version 21. It's designed to obtain a optimal muon identification with very low transverse momentum of  $3\text{GeV} < p_T < 5\text{GeV}$ , which is crucial for B-physics measurement in ATLAS. In this muon requirement, only CB muons are used. In the range of  $|\eta| < 1.3$ , it requests muons hit at least one MS station; in  $1.3 < |\eta| < 1.55$ , a least two MS stations are required; while in region of  $|\eta| > 1.55$ , *Medium WP* is required. In addtion, cuts are applied to suppress fakes:  $|\text{MBS}| < 3.0$ ,  $|\text{SNS}| < 3.0$  and  $|\text{SCS}| < 3.0$ .

Figure 4.8 and 4.9 show the selection efficiency of different muon identification working points. For *Medium*, *Tight* and *High- $p_T$* :  $Z \rightarrow \mu\mu$  events with  $p_T > 10\text{GeV}$  are used for measurement. In addition, the top plot also shows the efficiency of the *Loose* selection (squares), in which the Loose and Medium selections differ significantly in region of  $|\eta| < 0.1$ . For *LowPt*,  $J/\Psi \rightarrow \mu\mu$  events with  $3\text{GeV} < p_T < 10\text{GeV}$  are used for measurement.

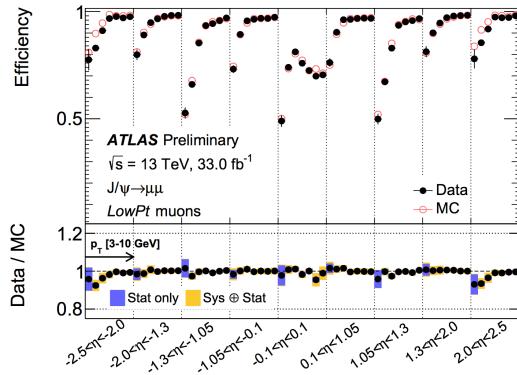
### Muon isolation

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

- **Calorimeter-based variable:**  $E_T^{\text{topocone}20}$ . It's defined as the sum of the transverse energy of topological clusters within a cone of 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>[55]</sup> and are corrected on an event-by-event basis.
- **Tracked-based variable:**  $p_T^{\text{varcone}30}$ . 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.

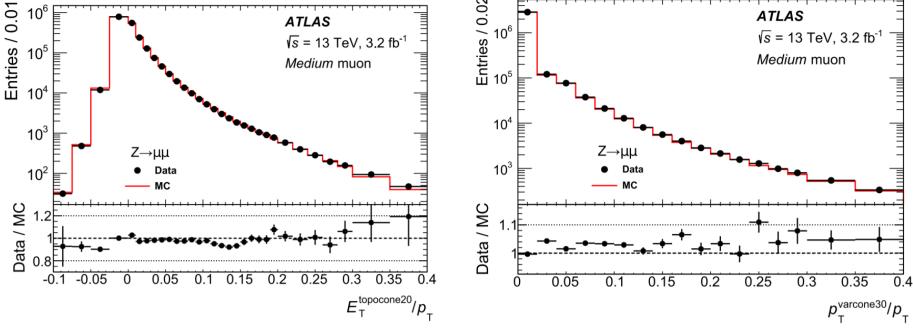


**Fig. 4.8 Muon reconstruction efficiency as a function of  $\eta$  for: Medium (and Loose), Tight and High- $p_T$  working points.**



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

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



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

#### 4.2.5 Jets

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

##### Jet reconstruction

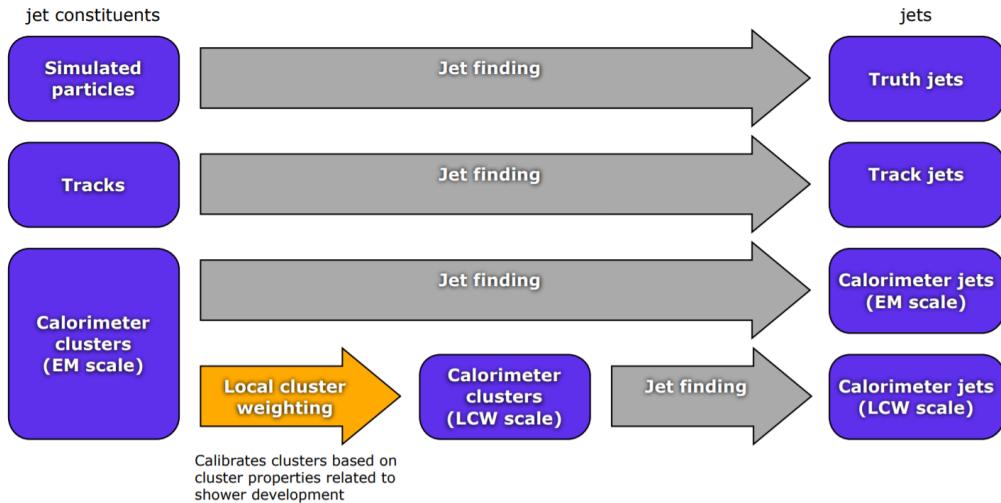
Jets are reconstructed using anti- $k_t$  algorithm<sup>[56]</sup> and with radius parameter of  $R = 0.4$  in most cases. The FASTJET software package<sup>[57]</sup> is utilized for jet finding and reconstruction. A collection of four-vectors are used as inputs at each combination step in jet clustering, 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.

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>[59]</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 pT threshold of 7 GeV are reconstructed with the anti- $k_t$  algorithm.

## Jet reconstruction



**Fig. 4.11** A overview schematic of ATLAS jet reconstruction<sup>[58]</sup>.

The *truth jets* are reconstructed also 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  $c_\tau > 10\text{mm}$  and muons, neutrinos, and excludes particles from pile-up. Truth jets with  $p_T > 7\text{GeV}$  and  $|\eta| < 4.5$  are then used for jets calibration that will be mentioned 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>[55]</sup>, a procedure that treats selected tracks as four-vectors of infinitesimal magnitude during the jet reconstruction and assigns them to the jet with which they are clustered. In addition, muon track segments are used as a compensation for those uncaptured jet energy from energetic particles passing through the calorimeters without fully being absorbed. The segments are tracks reconstructed from hits in MS and assigned to jets using the method of ghost association mentioned above as well.

### Jet energy scale calibration

Figure 4.12 depicts an overview of ATLAS jet calibration scheme for EM-scale calorimeter jets. This procedure restores the jet energy scale to that of 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 pT, energy, and mass.

First of all, the origin correction recompute the four-momentum of jets to point them to the hard-scatter primary vertex instead of the centre of detector, and in the meantime

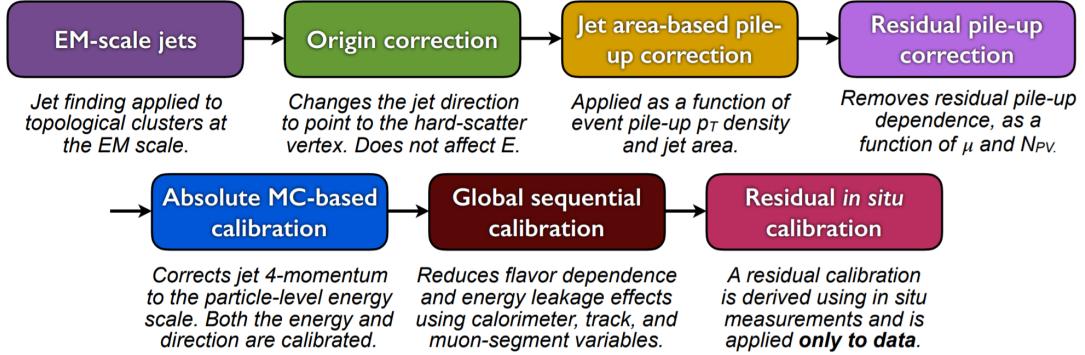


Fig. 4.12 A overview schematic of ATLAS jet calibration<sup>[59]</sup>.

keep the jet energy unchanged. This correction improves the  $\eta$  resolution of jets for roughly 25% at a jet  $p_T$  of 20 GeV and > 5 times improvement for jet with  $p_T$  above 200 GeV, as measured from the difference between reconstructed jets and truth jets in MC simulation. Secondly, the pile-up correction is adopted to remove the excess energy due to in-time and out-of-time pile-up, which consists two processes: an area-based  $p_T$  density subtraction applied on the top of each event; and a residual correction derived from the simulation. Thirdly, the absolute JES calibration corrects the jet four-momentum to the particle-level energy scale, using truth jets in dijet MC events. Furthermore, the step of global sequential calibration use calorimeter, track and MS-based variables to reduce the flavor dependence and energy leakage effects. Finally, the residual in situ calibration is adopted to correct jets in data by using well-measured objects eg. photons, Z bosons and calibrated jets.

### B-jet tagging

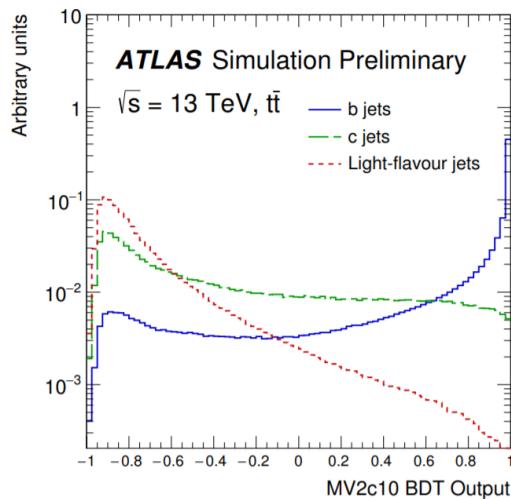
Tagging of b-jets plays a important role in many physics analyses involving b- or t- quark. On the other hand, lots of analyses need to apply b-jet veto to suppress top-antitop process. There are three major types of algrothms that have been developed to distinguish b-jet from light-quark (u,d,s) jets<sup>[60]</sup>:

- **Impact parameter based algorithms (IP2D and IP3D):** b-hardons usually have long lifetime ( $\sim 1.5$  ps,  $c_\tau \sim 450 \mu\text{m}$ ), which leads to large impact parameter for tracks produced from b-hardon decay. The impact parameter taggers are developed based on these variables. The IP2D tagger makes use of the transverse impact parameter significance  $d_0/\sigma(d_0)$  as descriminant, while IP3D tagger uses two-dimensional descriminant of both transverse and longitudinal impact parameter significances:  $d_0/\sigma(d_0)$  and  $z_0 \sin\theta/\sigma(z_0)$ .
- **Secondary vertex finding algorithm (SV1)** makes use of the secodary vertex formed by decay products of b-hardon within the jet. All track pairs within a jet are

tested for a two-track vertex hypothesis, and removed if they are likely to originate from a long-live particle decay (eg.  $K_s$  or  $\Lambda$ ), hadronic interactions or photon conversions. After that, a new vertex is fitted with all tracks from remaining two-track vertices, and the outliers are removed from this set of tracks.

- **Decay chain multi-vertex algorithm (JetFitter)**<sup>[61]</sup> exploits the topological structure of weak b- and c- hadron decays inside the jet and tries to reconstruct the full b-hadron decay chain. A Kalman filter is adopted to find a common line between primary vertex and b-/c- vertices, as well as their position in this line, which gives an approximated flight path for the b-hadron. In this approach, the b- and c-hadron vertices, whenever resolution allows, can be resolved, even when there is only a single track associated to them.

The final discrimination commonly used in many physics analyses is called **Multivariate Algorithm (MV2)**, which is based on Boosted Decision Tree (BDT) implemented in the TMVA package<sup>[62]</sup> by combining the outputs from underlaying taggers mentioned above. The MV2 was trained using jets in  $t\bar{t}$  sample, where b-jets are treated as signal and c- and light-flavor jets are treated as backgrounds. There are three kinds of MV2 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>[60]</sup>.

#### 4.2.6 Missing transverse energy

Many interesting physics processes are with the involvement of neutrinos. Since they do not interact with any materials in the detector, neutrinos cannot be detected directly;

but instead, they can result in imbalance in the plane transverse to the beam axis, in which momentum conservation is assumed. It is known as the missing transverse momentum denoted as  $E_T^{miss}$ , which is obtained from the negative vector sum of the momenta of all particles detected in a proton-proton collision event.

The  $E_T^{miss}$  is measured using selected, reconstructed and calibrated hard objects in an event. Its x- and y- components can be calculated as follow<sup>[63]</sup>:

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

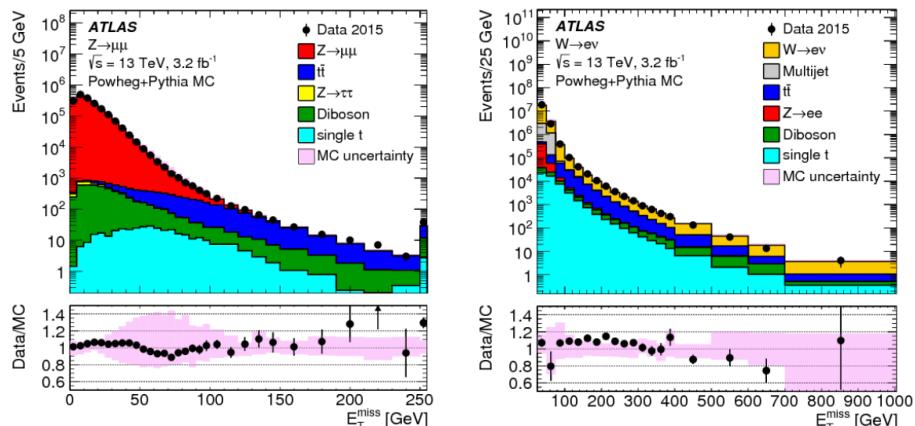
where each object term is given by the negative vectorial sum of the momenta of the respective calibrated objects. The calorimeter signals are associated with the reconstructed objects in the following order: electrons, photons, hadronically decaying taus, jets, muons. The soft term is reconstructed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

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

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

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

In figure 4.14, left plot shows the  $E_T^{miss}$  distribution for data and MC of  $Z \rightarrow \mu\mu$  events, in which there is no genuine missing transverse momentum; and right plot shows the  $E_T^{miss}$  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).

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

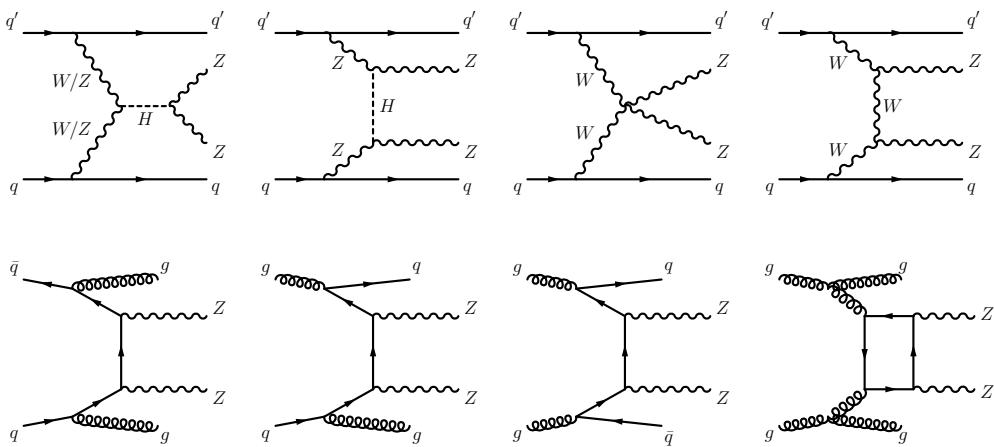
Objects contributing to $E_T^{miss}$				
Priority	Type	Selections	Variables	Comments
(1)	$e$	$ \eta  < 1.37 \text{ or } 1.52 <  \eta  < 1.47$ $p_T > 10GeV$	$E_T^{miss,e}$	all $e^\pm$ passing medium reconstruction quality and kinematic selections
(2)	$\gamma$	$ \eta  < 1.37 \text{ or } 1.52 <  \eta  < 1.47$ $p_T > 25GeV$	$E_T^{miss,\gamma}$	all $\gamma$ passing tight quality and kinematic selections in reconstruction, and without signal overlap with (1)
(3)	$\tau_{had}$	$ \eta  < 1.37 \text{ or } 1.52 <  \eta  < 1.47$ $p_T > 20GeV$	$E_T^{miss,\tau}$	all $\tau_{had}$ passing medium reconstruction quality and kinematic selections, and without signal overlap with (1) and (2)
(4)	$\mu$	$ \eta  < 2.7$ $p_T > 10GeV$	$E_T^{miss,\mu}$	all $\mu$ passing medium quality and kinematic selections in reconstruction
(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 reconstruction quality (jet cleaning) and kinematic selections, and without signal 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-scatter vertex passing reconstruction quality and kinematic selections, and not associated with any particle from (1), (3) or (4), or ghostassociated with a jet from (5)

## Chapter 5 Observation of electroweak ZZ production and measurement of SM ZZ cross section

### 5.1 Introduction

After discovering Higgs boson<sup>[64-65]</sup>, the examine of electroweak symmetry breaking (EWSB) becomes a main focus at the LHC. In addition to measuring the properties of Higgs boson directly, the vector boson scattering (VBS) process is another key avenue to probe EWSB<sup>[66-68]</sup>. As introduced in section 2.1.3, in Standard Model (SM), the Higgs boson acts as "moderator" to unitarize high-energy longitudinal VBS amplitudes at the TeV scale. Therefore, studying high-energy behaviours of VBS is crucial to understand the mechanism of EWSB.

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



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

The first evidence of the EW-VVjj process was seen in same-sign WW channel (EW- $W^\pm W^\pm$ jj) by ATLAS collaboration with  $20.3\text{ fb}^{-1}$  8 TeV data<sup>[69]</sup>, in which a  $3.6\sigma$  excess was observed in data over the background-only prediction. In LHC run-2, the observation of EW- $W^\pm W^\pm$ jj process has been reported in both ATLAS and CMS collaboration with  $36\text{ fb}^{-1}$  13 TeV data<sup>[70-71]</sup>. In WZ channel (EW-WZjj), an observation with  $5.3\sigma$  excess was also reported by the ATLAS collaboration recently<sup>[72]</sup>. The EW production in ZZ final state (EW-ZZjj) is typically rare, whose fiducial cross section has an order of  $O(0.1)\text{ fb}^{-1}$  in the final state where both Z bosons decay leptonically. The EW-ZZjj production was searched by CMS using  $35.9\text{ fb}^{-1}$  13 TeV data, no evidence was found<sup>[73]</sup>. But in the meantime,  $ZZ \rightarrow 4l$  process offers a extremely clean channel than all the others, with more data collected in LHC, the observation of EW-ZZjj becomes possible.

This section will present the first observation of EW-ZZjj production by ATLAS collaboration using the complete set of LHC run-2 data with  $139\text{ fb}^{-1}$  luminosity. It is a new milestone in the study of EWSB at LHC, and completes the last missing part of observation of weak boson scattering for *massive bosons*. The thesis will focus on the final state of Z bosons pair decay to four charged leptons with two jets ( $\ell\ell\ell\ell$  jj), includes both search of EW production and the fiducial cross-sections measurement for the inclusive production of the EW and QCD processes. The ZZjj production involving intermediate  $\tau$ -leptons from Z decays is considered as signal but has a negligible contribution to the selected event sample. Reducible backgrounds give minor contributions in the  $\ell\ell\ell\ell$  jj channel. To further separate the EW signal and the QCD background, multivariate discriminant (MD) is trained using event kinematic information from simulated samples. The MD distribution is then used as discriminant in statistical fit to evaluate the signal strength of EW process.

## 5.2 Data and MC samples

### 5.2.1 Data samples

The data sets for this analysis are the full run-2 pp collision data collected by the ATLAS experiment during the years from 2015 to 2018. Data event is only used if it passed the latest Good Run List (GRL) released by the Data Quality group from ATLAS experiment as listed below:

```
data15_13TeV.periodAllYear_DetStatus-v89-pro21-02_Unknown_PHYS_StandardGRL_All_Good_25ns.xml  
data16_13TeV.periodAllYear_DetStatus-v89-pro21-01_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns.xml  
data17_13TeV.periodAllYear_DetStatus-v99-pro22-01_Unknown_PHYS_StandardGRL_All_Good_25ns_Triggerno17e33prim.xml  
data18_13TeV.periodAllYear_DetStatus-v102-pro22-04_Unknown_PHYS_StandardGRL_All_Good_25ns_Triggerno17e33prim.xml
```

The events are required to additionally recorded by single and 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%.

### 5.2.2 MC simulation

The EW-ZZjj production (signal) is modelled using **MADGRAPH5\_aMC@NLO 2.6.1**<sup>[74]</sup> matrix elements (ME) calculated in the leading-order (LO) approximation in perturbative QCD (pQCD) and with NNPDF2.3LO<sup>[75]</sup> parton distribution functions (PDF). VBF Higgs process is also included.

The QCD-ZZjj production is modelled using **SHERPA 2.2.2**<sup>[76]</sup> with the NNPDF3.0NNLO<sup>[77]</sup> PDF, in which events with up to one (three) outgoing partons are generated at NLO (LO) in pQCD. The production of ZZjj from the gluon-gluon initial state with a four-fermion loop or with an exchange of the Higgs boson has an order of  $\alpha_S^4$  in QCD, and is not included in the **SHERPA** simulation. A separate gg induced  $Z Z + 2\text{jets}$  sample is modelled using **SHERPA 2.2.2** with the NNPDF3.0NNLO PDF, and with an additional 1.7 k-factor<sup>[78]</sup> being applied.

Then the interference between EW- and QCD-ZZjj is modelled with **MADGRAPH5\_aMC@NLO 2.6.1** calculated at LO.

The diboson productions from QCD  $WW \rightarrow l\nu qq$  as well as QCD and EW  $WZ \rightarrow llqq$  are modelled using **SHERPA 2.2.2** with NNPDF3.0NNLO PDF. The productions of semileptonic decays ( $WW \rightarrow l\nu qq$  and  $WZ \rightarrow qql l$ ) are modelled using **POWHEG-Box v2**<sup>[79]</sup> with the CT10 PDF<sup>[80]</sup>. Other diboson processes are not included due to negligible contributions. The triboson production is modelled using **SHERPA 2.2.2** with NNPDF3.0NNLO PDF.

For top-quark pair ( $t\bar{t}$ ) production, the **POWHEG-Box v2** is used with the CT10 PDF. The single top-quark production in  $t$ -channel,  $s$ -channel and  $Wt$ -channel were simulated using the **POWHEG-Box v1** event generator<sup>[81-83]</sup>. The productions of  $t\bar{t}$  in association with vector boson(s) ( $t\bar{t}V$ ) is modelled with **MADGRAPH5\_aMC@NLO 2.3.3** for  $t\bar{t}W$  and  $t\bar{t}Z$  with  $Z \rightarrow \nu\nu/qq$  decays, with **SHERPA 2.2.1** for  $t\bar{t}Z$  with the  $Z$  to dilepton decays, and with **MADGRAPH5\_aMC@NLO 2.2.2** for  $t\bar{t}WW$  respectively.

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

For all the samples except those from **SHERPA**, the parton showering is modelled with **PYTHIA8**<sup>[38]</sup> using NNPDF2.3<sup>[75]</sup> PDF set, and the A14 set of tuned parameters<sup>[84]</sup>. For

SHERPA samples, the parton showering is simulated within the programme.

All simulated events were processed with detector response simulated based on GEANT4 described in section 4.1. In addition, simulated inelastic pp collisions were overlaid to model additional pp collisions in the same and neighbouring bunch crossings (pile-up), and reweighted to match the pile-up conditions in data. Moreover, all simulated events were processed using the same reconstruction algorithms as data. And the leptons' and jets' reconstruction, energy scale and resolution, and the leptons' identification, isolation, trigger efficiencies for simulated events, as described in section 4.2, were all corrected to match the data measurements.

## 5.3 Objects and Event selection

### 5.3.1 Objects selection

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

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

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

**Jets:** Jet are key signatures for VBS processes. This analysis use the jets clustered using the anti-kt algorithm with radius parameter  $R = 0.4$ , more details of jets' reconstruction can be found in section 4.2.5. The jets are required to have  $p_T > 30\text{(40)}\text{ GeV}$  in

the  $|\eta| < 2.4$  ( $2.4 < |\eta| < 4.5$ ) region. To further reduce the effects of pile-up jets, a jet vertex tagger (JVT) is applied to jets with  $p_T < 60$  GeV and  $|\eta| < 2.4$  to select jets from hard-scattering vertex.

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

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

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

### 5.3.2 Event selection in reconstruction level

The  $\ell\ell\ell\ell$  quadruplets are formed by two opposite-sign, same-flavour (OSSF) lepton pairs ( $l^+l^-$ ), in which leptons are required to be separated by  $\Delta R > 0.2$  in table 5.1. At most one muon is allowed to be ME or CT muon. The  $p_T$  threshold of first three leading muons are 20, 20 and 10 GeV. If more than one quadruplets are found, the one with minimum sum of difference between two muon 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  $J/\psi$  or  $\Upsilon$  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 5.2 summarizes the above

selection requirements, which is defined as signal region (SR).

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

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

## 5.4 Background estimation

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

Process	$\ell\ell\ell\ell jj$
EW- $ZZjj$	$20.6 \pm 2.5$
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 5.3** 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.

### 5.4.1 QCD backgrounds

The QCD-ZZjj production, which include both qq and gg induced processes, is an irreducible background in the search of EW-ZZjj production. A QCD-enriched control region (CR) is defined to constrain the contribution by reverting either the  $m_{jj}$  or  $\Delta y_{jj}$  requirements:

$$m_{jj} < 300 \text{ GeV} \text{ or } \Delta y_{jj} < 2$$

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

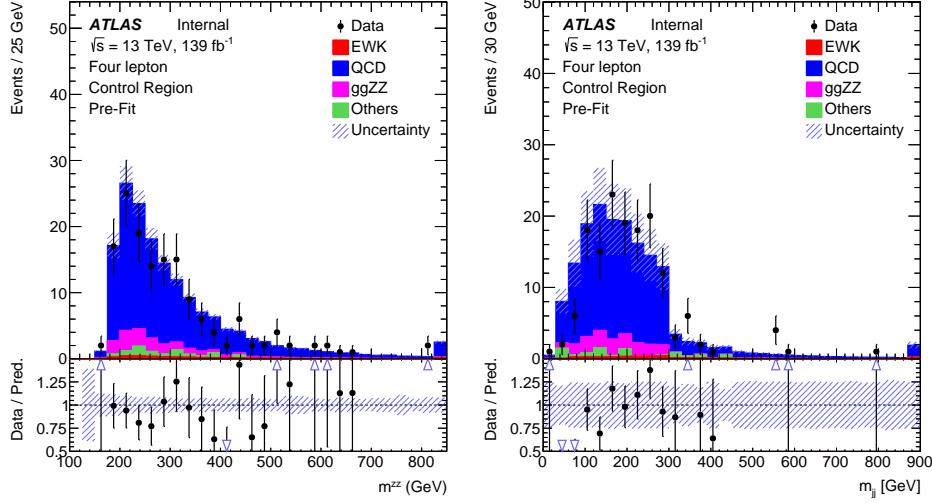
Process	$\ell\ell\ell\ell jj$
EW ZZjj	$3.9 \pm 0$
QCD ZZjj	$136.9 \pm 0.6$
QCD ggZZjj	$16.8 \pm 0.1$
Diboson	$0.3 \pm 0.1$
Triboson	$1.6 \pm 0.1$
Z+jets	0
t <bar>t</bar>	0
Total	$159.5 \pm 0.62$
Data	152

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

distributions of 4l and dijet invariant mass in QCD CR are shown in figure 5.2.

### 5.4.2 Fake backgrounds

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



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

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

$$\mathcal{F} = \mathcal{N}_{good}/\mathcal{N}_{pool} \quad (5.1)$$

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  $\ell\ell\ell\ell jj$  fake control region, where one or two leptons pass *poor* requirement while all the other leptons are required to have SR selection.
3. The number of fake events are calculated as:

$$\mathcal{N}_{fake} = (N_{gggp} - N_{ggpp}^{ZZ}) \times \mathcal{F} - (N_{ggpp} - N_{ggpp}^{ZZ}) \times \mathcal{F}^2 \quad (5.2)$$

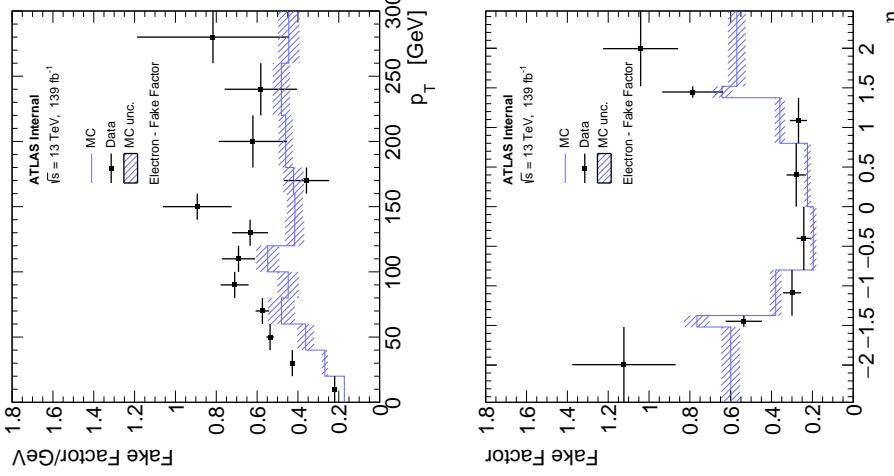
with the subtraction of ZZ contribution, and the double counting between ( $N_{gggp}$  and  $N_{ggpp}$ ).

For the definition of *poor* leptons: The poor electrons are defined as failing "FixedCutLoose" isolation requirement or "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.

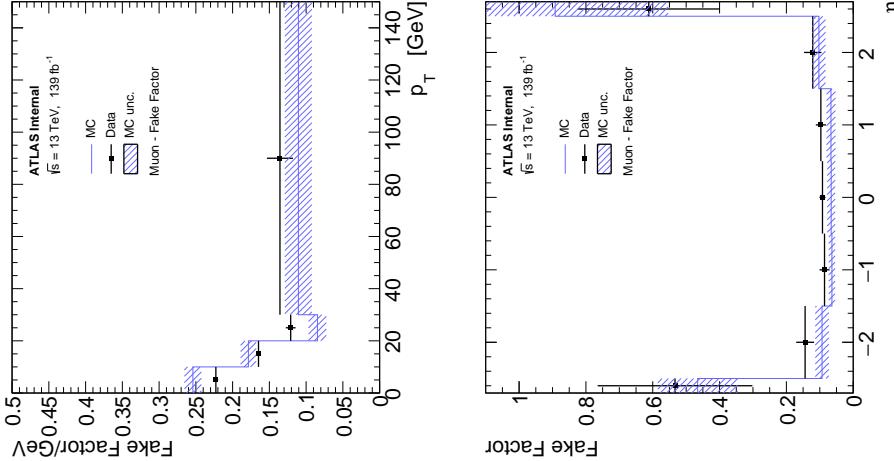
### 1. Fake factor for Z+jets

Fake factor for Z+jets background is calculated in Z+jets-enriched region, where events with one SFOS lepton pair around Z mass associated with two jets are selected.

The value of fake factor is driven from data, and as a function of  $p_T$  and  $\eta$  as shown in figure 5.3 for electrons and figure 5.4 for muons. During calculation, the contributions from non-Z+jets backgrounds ( $t\bar{t}$ , ZZ, WZ) have been subtracted from data. The values calculated from Z+jets MC directly are also shown in plots for comparison.



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



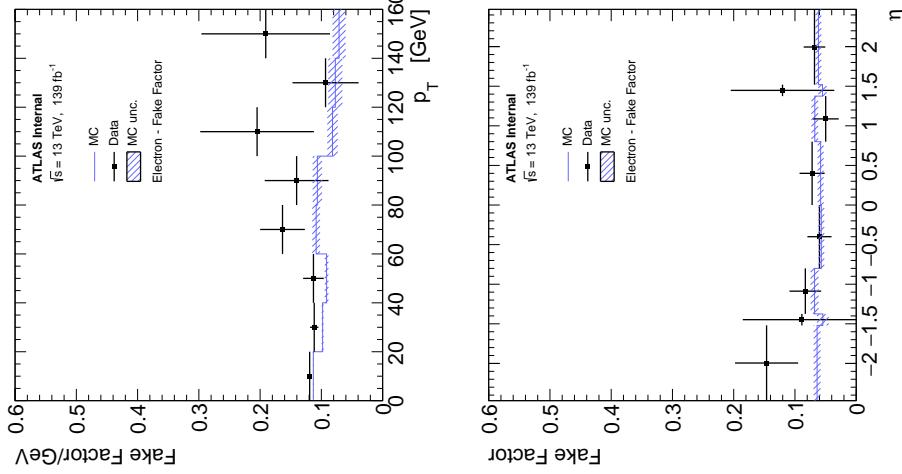
**Fig. 5.4** Fake factor for Z+jets background, constructed with additional muon, as a function of  $p_T$  (left) and  $\eta$  (right).

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

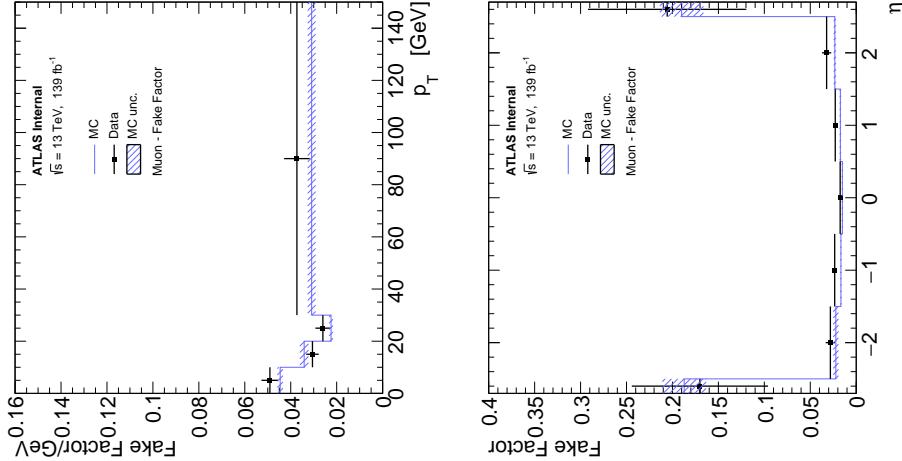
The fake factor for  $t\bar{t}$  are calculated in  $t\bar{t}$  dominated region by selecting one  $e\mu$ -pair with additional two jets. For events with three leptons,  $m_T^W < 60$  GeVcut is applied to reject the contribution from  $t\bar{t} + W$  events. The  $m_T^W$  is defined as below:

$$m_T^W = \sqrt{2p_T^{l_3}E_T^{miss} \left[ 1 - \cos(\Delta\phi(p_T^{l_3}, E_T^{miss})) \right]} \quad (5.3)$$

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 5.5 for electrons and 5.6 for muons. The non- $t\bar{t}$  contributions, which include  $Z+jets$ ,  $ZZ$  and  $WZ$ , are subtracted from data.



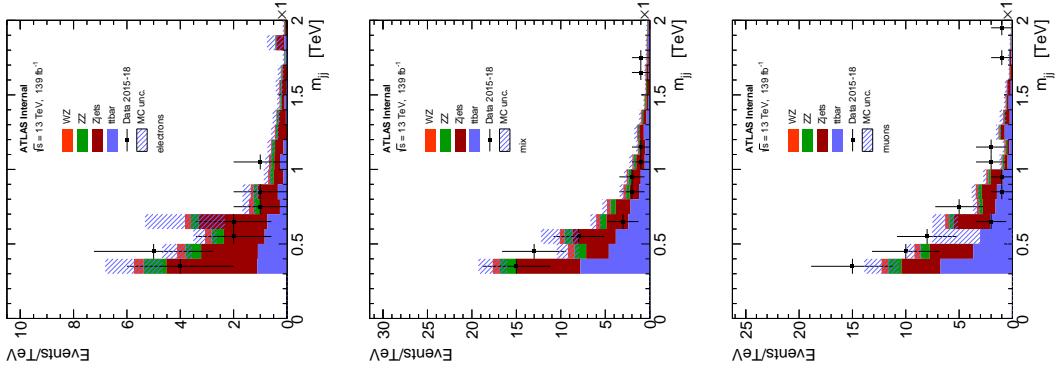
**Fig. 5.5** Fake factor for  $t\bar{t}$  background, constructed with additional electron, as a function of  $p_T$  (left) and  $\eta$  (right).



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

### 3. Combination

The fake factors calculated from each dedicated region are then combined together according to their contributions in fake control region described previously. Figure 5.7 shows the  $m_{jj}$  distribution with data and major fake backgrounds in three different 4l channels.



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

#### 4. Systematics of fake estimation and results

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

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

Table 5.5 summarizes the contribution of fake backgrounds in signal region under different systematic conditions mentioned above as well as the nominal one. Uncertainties of each value in table are statistical one.

## 5.5 Systematics

The analysis performances both the statistical fit to MD distribution to extract the EW-ZZjj contributions and the cross section measurements in fiducial volume. Therefore, theoretical and experimental uncertainties may affect the predictions background yields and shapes, correction factors from detector-level to particle-level measurement, as well as the ZZjj MD shapes and so on. Moreover, the statistical uncertainties of

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

simulated samples are also taken into account. And due to the extremely low cross section of  $\ell\ell\ell\ell$  channel, the analysis is still data statistic dominated. This section will described the measurement of both theoretical and experimental systematics for  $ZZjj$  productions. The systematics for fake backgrounds have been elaborated in section 4.

### 5.5.1 Theoretical systematics

The theoretical systematics on EW- and QCD-ZZjj processes include the uncertainties from PDF, QCD scale,  $\alpha_S$  and parton showering variations. The PDF uncertainty is estimated from 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 renormalisation 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 choosen 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>[85-86]</sup> algorithm. The  $\alpha_S$  uncertainty is estimated by varying the value of  $\alpha_S$  within  $\pm 0.001$ . Details of those variation components are summarized in table 5.6. Due to the lack of simulation sample for alternative parton showering on QCD-ZZjj process, the value of parton showering component is taken from the measurement of EW process.

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

Process	EW-ZZjj	QCD-ZZjj
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 5.6** Summary of different variations for EW- and QCD-ZZjj theoretical uncertainties measurement.

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 5.7** Summary of theoretical uncertainties for the fiducial volum (SR) for both EW and QCD qq-initial processes.

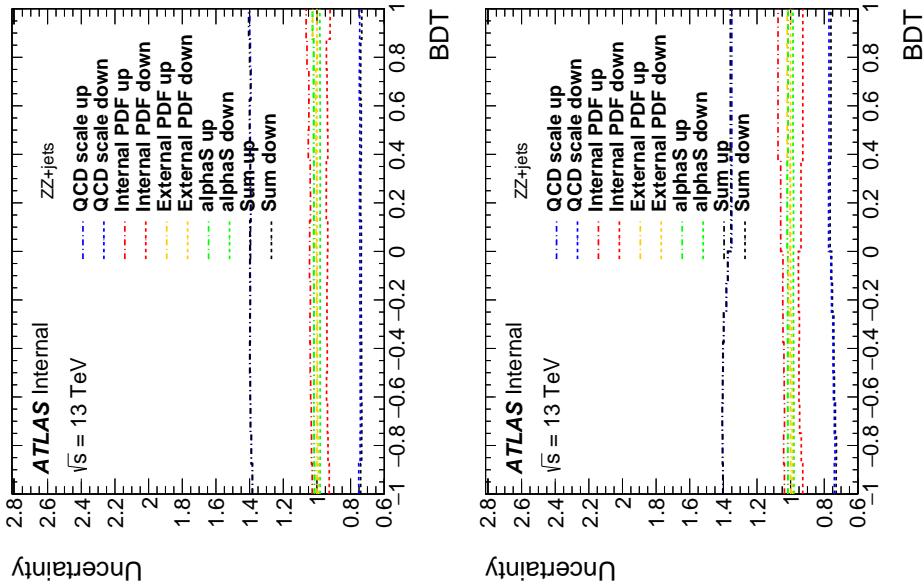
Process	PDF (%)	$\alpha_S$ (%)	QCD scale (%)	Parton shower (%)
EW $\ell\ell\ell\ell$	+6.1 -6.1		+0.8 -1.1	+10.1 -10.1
qqQCD $\ell\ell\ell\ell$	+2.0 -1.0	+2.6 -2.6	+31.5 -22.0	

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

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

### 5.5.2 Experimental systematics

The major experimental uncertainties are from the luminosity uncertainty, the momentum scale and resolution of leptons and jets, as well as the lepton reconstruction and selection efficiency. Some smaller uncertainties, such as trigger efficiency and pile-up correction, are also considered. Overall, most large systematics are from leptons and jets. Table 5.9 lists the major systematic components from leptons and jets for major processes in  $\ell\ell\ell\ell$  channel. The total uncertainties for sources from electron, muon and jet respectively, as well as the sum (quadratic sum) of them are also summarized in this table.

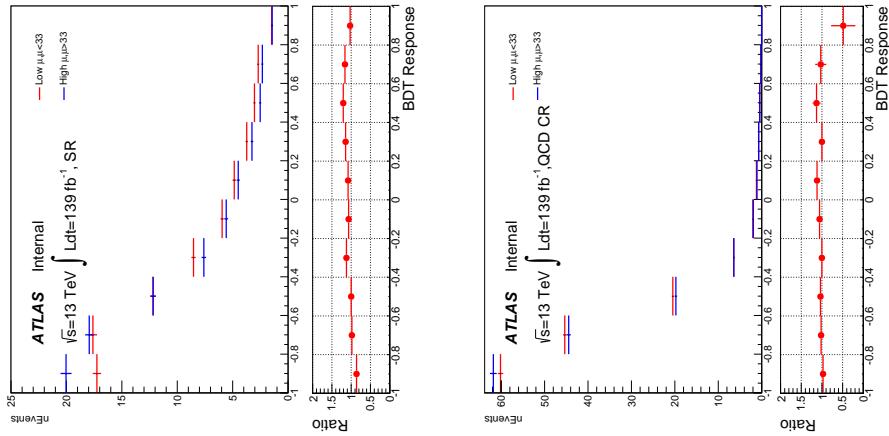


**Fig. 5.8** The theoretical uncertainties for  $gg \rightarrow ZZ$  background in particle-level SR (left) and CR (right).

name	EW-ZZjj	QCD qq-initial	QCD gg
nominal yield	20.61	76.69	13.10
EG_RESOLUTION_ALL	$\pm 0.00\%$ $\pm 0.03\%$	$\pm 0.02\%$ $\pm 0.04\%$	$\pm 0.01\%$ $\pm 1.41\%$
EG_SCALE_ALL	$\pm 0.03\%$ $\pm 0.05\%$	-0.04%	$\pm 0.01\%$ $\pm 0.06\%$
EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	$\pm 2.66\%$ $\pm 2.58\%$	$\pm 2.60\%$ $\pm 2.53\%$	$\pm 2.65\%$ $\pm 2.57\%$
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	$\pm 1.97\%$	$\pm 11.82\%$ $\pm 10.14\%$	$\pm 16.21\%$ $\pm 12.92\%$
JET_GroupedNP_2	$\pm 0.23\%$	$\pm 1.26\%$	+5.3%
JET_GroupedNP_3	$\pm 0.55\%$	$\pm 2.94\%$	$\pm 3.14\%$ $\pm 0.12\%$
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%	$\pm 0.04\%$ $\pm 0.01\%$	<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	71	$\pm 0.02\%$ $\pm 0.03\%$ $\pm 0.02\%$	<0.001%
PRW_DATASF		$\pm 0.5\%$ $\pm 0.42\%$ $\pm 1.02\%$	$\pm 2.17\%$ $\pm 1.46\%$
Electron Exp		$\pm 2.8\%$ $\pm 2.7\%$	$\pm 2.75\%$ $\pm 2.64\%$
Muon Exp		-1.2%	-1.2%
Photon Exp		-1.04%	-1.04%

In addition, the uncertainty of the combined 2015–2018 integrated luminosity is 1.7%<sup>[87]</sup>, obtained using the LUCID-2 detector<sup>[88]</sup> for the primary luminosity measurements.

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



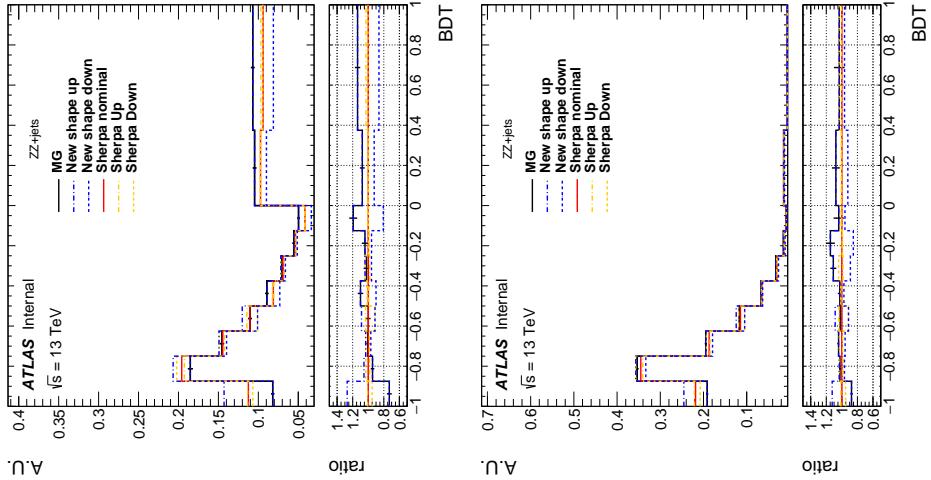
**Fig. 5.9** MD distribution between low and high pile-up events for SR (left) and CR (right).

Moreover, a conservative uncertainty is signed to QCD-ZZjj process by comparing the sample modeled by SHERPA generator (nominal) to MADGRAPH5\_aMC@NLO. The MD shape differences for both SR (left) and QCD CR (right) are shown in figure 5.10. The modeling uncertainty is calculated from the envelop of MD shape difference between nominal and alternative samples as function of MD.

## 5.6 Measurement of fiducial cross section

The fiducial cross section for the production of inclusive  $Z Z jj$ , which includes both EW and QCD components, is then measured.

The defination of fiducial volume, which is used for cross section measurement, follows closely to the detector-level selection but use physics objects in "particle-level", which are reconstructed in simulation from stable final-state particles, prior to their interactions with the detector. For electrons and muons, QED final-state radiation is for the most part recovered by adding the four-momenta of surrounding photons that are



**Fig. 5.10** MD shape difference for QCD  $q\bar{q} \rightarrow ZZ$  background between different SHERPA theoretical uncertainties and sample from MADGRAPH5\_aMC@NLO on SR (left) and CR (right).

not originating from hadrons and within an angular distance  $\Delta R < 0.1$  to the lepton four-momentum, called lepton "dressing". Particle-level jets are built with anti- $k_T$  algorithm with radius parameter  $R = 0.4$  using all final-state particles except leptons and neutrinos as inputs. Comparing to the events selection in detector-level in section 5.3, in particle-level, the selected dilepton pair mass required is relaxed to be within 60 to 120 GeV to reduce the migration effect as well as be more compatibility with previous CMS publication<sup>[73]</sup>. All the other kinematics selection requirements are the same as the definition in detector-level.

### 5.6.1 Calculation of C-factor

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

$$C = \frac{N_{\text{detector-level}}}{N_{\text{FV}}} \quad (5.4)$$

The C-factor value of each  $ZZjj$  processes calculated from each individual simulation samples are listed in table 5.10 as well as their systematics.

Process	$C$	$\Delta C(\text{stats})$	$\Delta C(\text{sys})$	$\Delta C(\text{theo})$
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 5.10** C Factor of different  $ZZjj$  processes.

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

$$C = \sum_i \frac{N_{FV.}^i}{\sum_j N_{FV.}^j} \times C_i = 0.699 \pm 0.003(\text{stats.}) \pm^{0.011}_{0.013} (\text{theo.}) \pm 0.028(\text{exp.}) \quad (5.5)$$

The stats. refers to the statistical uncertainty from MC simulation statistics. The theo. and exp. denote the theoretical and experimental uncertainties described in section 5.5.

### 5.6.2 Result of fiducial cross section

The cross section in fiducial volume is computed as:

$$\sigma^{FV.} = \frac{N_{data} - N_{bkg}}{C \times Lumi} \quad (5.6)$$

where  $N_{data}$  and  $N_{bkg}$  denote the number of events selected from detector-level selection from data and sum of backgrounds, and  $C$  is the C-factor calculated above, Lumi represents the integrated luminosity of data 2015–2018 of  $139 \text{ fb}^{-1}$ .

As shown in table 5.3, in inclusive measurement, only "Others" represents background, processes of EW-ZZjj, QCD- $q\bar{q} \rightarrow ZZ$  and QCD-gg  $\rightarrow ZZ$  are signals. Table 5.11 shows the fiducial cross section for  $\ell\ell\ell\ell$  channel measured from equation 5.6, as well as the predicted cross section measured from signals MC directly. The mea-

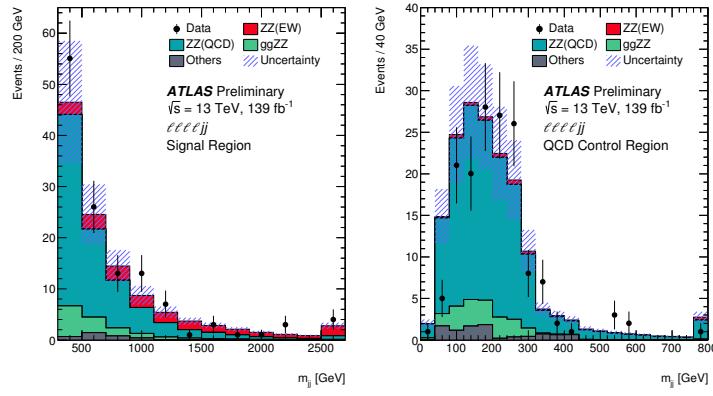
Measured fiducial $\sigma$ [fb]	Predicted fiducial $\sigma$ [fb]
$1.27 \pm 0.12(\text{stat}) \pm 0.02(\text{theo}) \pm 0.07(\text{exp}) \pm 0.01(\text{bkg}) \pm 0.03(\text{lumi})$	$1.14 \pm 0.04(\text{stat}) \pm 0.20(\text{theo})$

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

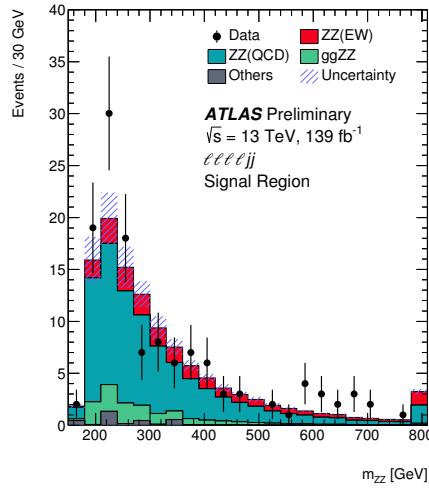
sured cross section has a total uncertainty of 11%, and is found to be compatible with SM prediction. The data statistic is still dominant for the measurement.

## 5.7 Search for EW-ZZjj

Figure 5.11 represents the  $m_{jj}$  distribution in SR (left) and QCD CR (right), where the normalization of EW and QCD processes are scaled according to their observed value explained later in this section. High  $m_{jj}$  region is more sensitive for EW-ZZjj events detection from this figure. Figure 5.12 shows the spectrum of invariant mass of  $\ell\ell\ell\ell$  system ( $m_{ZZ}$ ) in SR.



**Fig. 5.11** Observed and expected  $m_{jj}$  distributions in SR (left) and QCD CR (right). The error bands include the expected experimental and theoretical uncertainties. The error bars on the data points show the statistical uncertainty on data. The contributions from the QCD and EW production of  $ZZjj$  events are scaled by 0.96 and 1.35, respectively, which correspond to the observed normalization factors in the statistical fit to the combined channel. The last bin includes the overflow events.



**Fig. 5.12** Observed and expected  $m_{ZZ}$  spectrum in SR. The error bands include the expected experimental and theoretical uncertainties. The error bars on the data points show the statistical uncertainty on data. The contributions from the QCD and EW production of  $ZZjj$  events are scaled by 0.96 and 1.35, respectively, which correspond to the observed normalization factors in the statistical fit to the combined channel. The last bin includes the overflow events.

### 5.7.1 MD discriminant

To further separate the EW-ZZjj component from QCD-ZZjj, a MD based on *Gradient Boosted Decision Tree (BDTG)* algorithm<sup>[89]</sup> is trained with simulated events via TMVA framework<sup>[62]</sup>. For  $\ell\ell\ell\ell$  channel, training is performed between EW (signal) and QCD (background) processes. Twelve event kinematic variables sensitive to the characteristics of the EW signal is used as input features in training. Table 5.12 listed those input variables with the order of their ranking provide by TMVA tool. The jet-related information provides larger sensitive in  $\ell\ell\ell\ell$  final-state. Then the MD distributions in both SR and QCD CR region are used for statistical fit.

### 5.7.2 Profile likelihood ratio method

To examine the compatibility between data and the signal-plus-background hypothesis, a test statistic is driven by using the profile likelihood ratio method. The binned likelihood function is given as”

$$\mathcal{L}(\mu, \sigma) = \prod_i^{\text{bins}} \mathcal{L}_{\text{poiss}}(N_{\text{data}} | \mu s(\theta) + b(\theta))_i \times \mathcal{L}_{\text{gauss}}(\theta)_i \quad (5.7)$$

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

The test statistic  $q_\mu$  is defined as:

$$q_\mu = -2 \ln \left( \frac{\mathcal{L}(\mu, \hat{\theta}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \right) \quad (5.8)$$

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

Rank	Variables	Description
1	$m_{jj}$	Dijet invariant mass
2	$p_T^{j1}$	$p_T$ of the leading jet
3	$p_T^{j2}$	$p_T$ of the sub-leading jet
4	$\frac{p_T(ZZjj)}{H_T(ZZjj)}$	$p_T$ of the $ZZjj$ system divided by the scalar $p_T$ sum of Z bosons and two jets
5	$y_{j1} \times y_{j2}$	Product of jet rapidities
6	$\Delta y_{jj}$	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	$p_T^{\ell 3}$	$p_T$ of the third lepton

**Table 5.12 Input features for the training of MD.**

### 5.7.3 Fitting procedure

A profile likelihood fit is performed on MD discriminant to extract the EW-ZZjj signal from backgrounds. The binning of MD distributions in SR is optimized to maximize the sensitive for detecting EW signal. The normalization of QCD-ZZjj production ( $\mu_{QCD}^{III}$ ) in  $\ell\ell\ell\ell$  channel is varied simultaneously in the fit in SR and QCD CR as

described in section 5.4. The signal strength of EW-ZZjj production ( $\mu_{EW}$ ) is taken as parameter of interest and floated in the fit. The effects of the uncertainties related to normalizations and shapes described previously in section 5.5 of background processes in the MD distribution are all taken into account.

In most case, a common nuisance parameter is used for each source of systematic in all bins and all channels. The statistical uncertainties for simulated samples are uncorrelated among all bins, and the background uncertainties only applied to their corresponding backgrounds. For combination between two channels, the theoretical uncertainties between  $\ell\ell\ell\ell$  and  $\ell\ell\nu\nu$  are uncorrelated due to different fiducial volumes defination. Furthermore, to be more conservative, the generator modelling uncertainty for QCD-ZZjj production mentioned in section 5.5 is separated to be two nuisance parameters in low and high MD region.

#### 5.7.4 Result of statistical fit

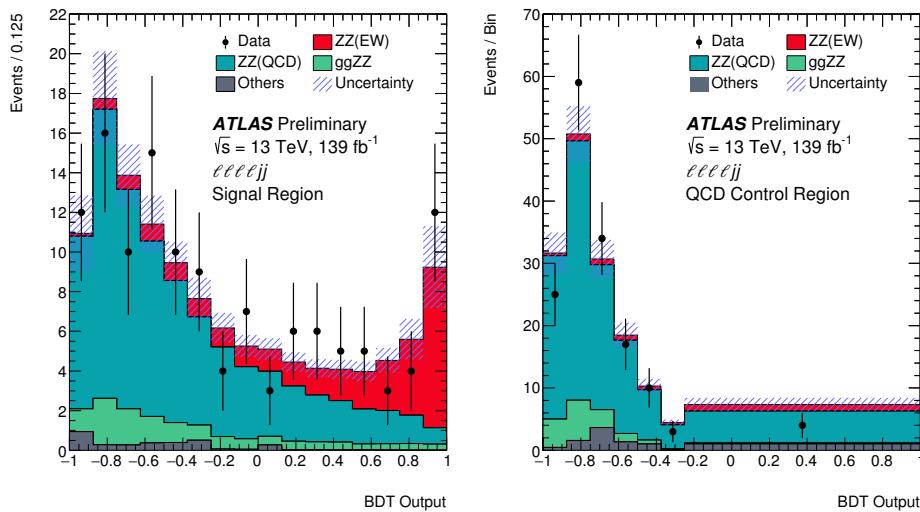
The statistical fit is performed both in individual  $\ell\ell\ell\ell$  channel, as well as the combination between  $\ell\ell\ell\ell$  and  $\ell\ell\nu\nu$  channel to gain more statistic. The results of statistical fit for  $\ell\ell\ell\ell$  channel and the combined channel are presented in table 5.13. The  $\ell\ell\nu\nu$  analysis will not be talked about in this thesis, but more details can refer to<sup>[90]</sup>. To drive expected results, the oberved data is used for QCD CR to extract normalization factor of QCD component ( $\mu_{QCD}^{III}$ ), while in SR, asimov data built from background prediction and signal model with SM assumed cross section is used. For  $\ell\ell\ell\ell$  channel,

	$\mu_{EW}$	$\mu_{QCD}^{\ell\ell\ell\ell jj}$	Significance Obs. (Exp.)
$\ell\ell\ell\ell jj$	$1.54 \pm 0.42$	$0.95 \pm 0.22$	$5.48 (3.90) \sigma$
Combined	$1.35 \pm 0.34$	$0.96 \pm 0.22$	$5.52 (4.30) \sigma$

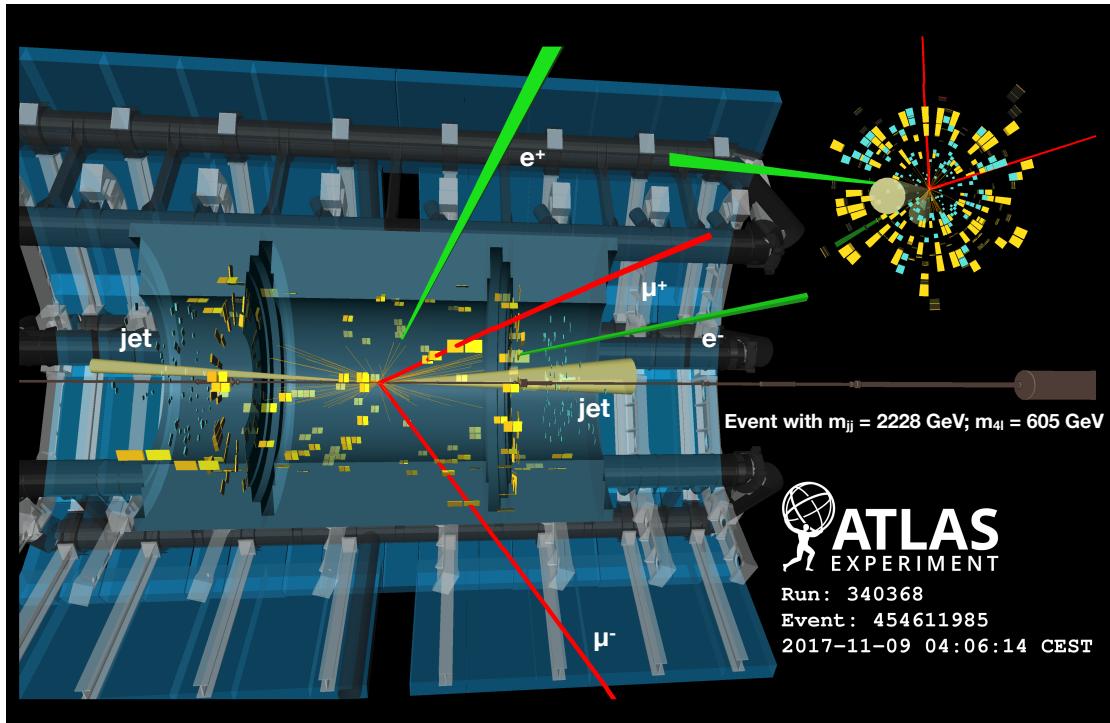
**Table 5.13** Observed  $\mu_{EW}$  and  $\mu_{QCD}^{\ell\ell\ell\ell jj}$ , as well as the observed and expected significance from the individual  $\ell\ell\ell\ell jj$  channel, and the combined fits. The full set of systematic uncertainties is included.

the background-only hypothesis is rejected at  $5.5\sigma$  ( $3.9\sigma$ ) for data (expectation), which leads to the observation of EW-ZZjj production. Figure 5.13 shows the post-fit MD distributions for  $\ell\ell\ell\ell$  channel in SR (left) and QCD CR (right). The EW-ZZjj cross section measured in  $\ell\ell\ell\ell$  channel is extracted to be  $0.94 \pm 0.26$  fb.

Figure 5.14 depicts the display of an event candidate of EW-ZZjj production in  $2e2\mu$  final state with two jets in forward and backward region, and with a MD value from 0.875 to 1.0.



**Fig. 5.13** Observed and expected multivariate discriminant distributions after the statistical fit in the  $\ell\ell\ell\ell$  SR (left) and QCD CR (right). The error bands include the experimental and theoretical uncertainties, as well as the uncertainties in  $\mu_{\text{EW}}$  and  $\mu_{\text{QCD}}^{\ell\ell\ell\ell jj}$ . The error bars on the data points show the statistical uncertainty on data.



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

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

### 5.8.1 Introduction

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

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

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

### 5.8.2 The ATLAS detector at HL-LHC

As the expectation of HL-LHC, the new Inner Tracker (ITk)<sup>[91]</sup> will extend the acceptance capability of ATLAS detector to pseudorapidity ( $|\eta|$ ) up to 4.0. By including a forward muon trigger, the upgraded Muon Spectrometer<sup>[92]</sup> is also expected to provide muon identification capabilities to  $|\eta|$  up to 4.0. The new high granularity timing detector (HGTD)<sup>[93]</sup> designed to mitigate the pile-up (PU) effects is also foreseen in the forward region of  $2.4 < |\eta| < 4.0$ . The expected performance of the upgraded ATLAS detector at HL-LHC has been studied as reported in Ref.<sup>[94]</sup>.

### 5.8.3 Simulation

The analysis is performed using particle-level events for simulated samples. The samples are generated at  $\sqrt{s} = 14 \text{ TeV}$  and with a fast simulation based on setting for ATLAS detector at HL-LHC. The signal in this analysis is EW-ZZjj process, while only the dominated irreducible background of QCD-ZZjj is considered. Both signal

and background are generated using **SHERPA** with NNPDF3.0NNLO PDF set. The signal sample is generated with two jets at Matrix Element (ME) level. The background is qq-initial process, in which events with up to one (three) outgoing partons are generated at NLO (LO) in perturbative QCD. As a quick study, other irreducible backgrounds like fake backgrounds from Z+jets and top-quark processes, as well as Diboson without 4l final-state and Triboson processes are not included into this analysis. Furthermore, for hard scattering events, the pile-up collisions are set with a mean value of 200 interactions per bunch crossing. In studies, signal and background yields are then scaled to 3000 fb<sup>-1</sup> for HL-LHC.

#### 5.8.4 Event selection

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

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

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

Table 5.14 summarized the number of selected signal and background events normalized to 3000 fb<sup>-1</sup>. In addition to the *baseline* selection listed above, two alternative selections are also studied:

- Reduce the lepton  $\eta$  region to 2.7, to understand the effect due to forward lepton reconstruction and identification with the upgraded ATLAS detector.
- 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.

From this table, one can see the extended track coverage increases the  $\ell\ell\ell\ell jj$  events by 15–30%, by 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

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

events.

Figure 5.15 shows the kinematic distributions of dijet 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:

$$\text{ZZ centrality} = \frac{|y_{ZZ} - (y_{j1} + y_{j2})/2|}{|y_{j1} - y_{j2}|} \quad (5.9)$$

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

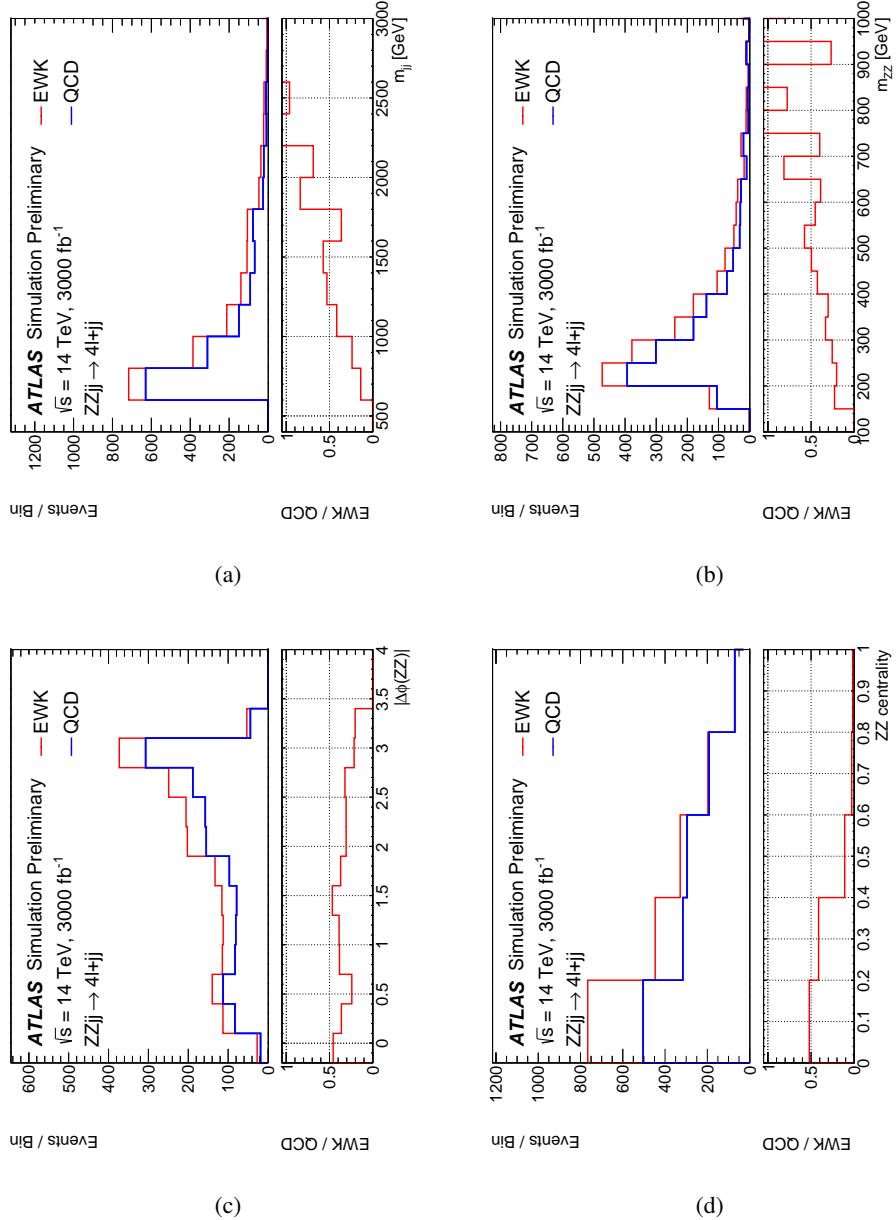
### 5.8.5 Systematics

According to studies in section 5.5, the dominated systematic in  $\ell\ell\ell\ell$  channel is from theoretical systematic for QCD-ZZjj background process. Different sizes of systematics have been tested, at 5, 10 and 30% level on background modeling. The 5% uncertainty is an optimal estimation when there is enough data events from QCD-enrich control region at HL-LHC to constrain the theoretical modeling on QCD-ZZjj process. The 30% one is a conservative estimation, in which the uncertainties are calculated from different PDF sets and QCD renormalization and factorization scales, following recommendation from PDF4LHC mentioned in section 5.5.

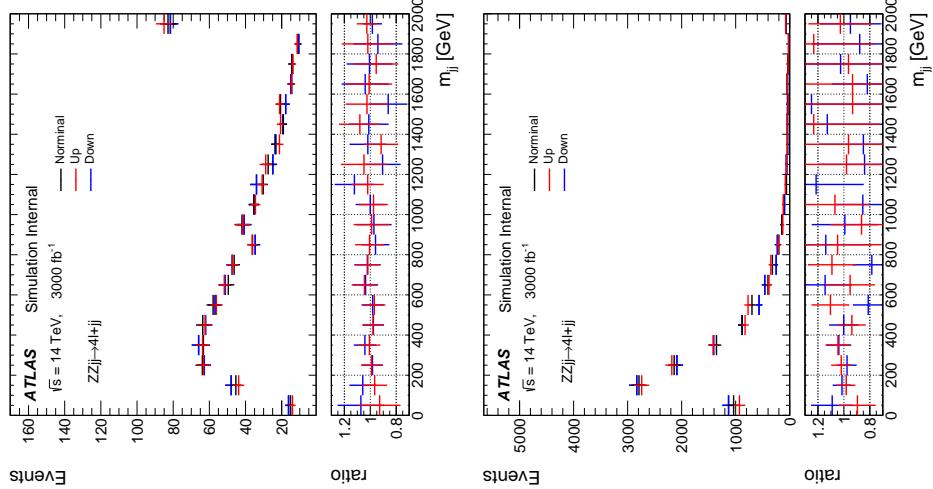
For experimental sources, the jet systematic has been checked following the setting provided by HL-LHC in Ref.<sup>[94]</sup>, and the uncertainties are within 5% level, which is smaller than run-2 measurement at 10%. Figure 5.16 depicts the up and down variations for jet uncertainty provided by HL-LHC performance tool as function of dijet invariant mass. Therefore, a conservative 5% uncertainty is used as experimental uncertainty.

The final results rely largely on the uncertainties, especially the theoretical uncertainties on QCD-ZZjj production. So results with different uncertainty conditions will be shown:

- The case with statistical uncertainty of simulated samples only.



**Fig. 5.15** 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  $3000 \text{ fb}^{-1}$ .



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

- The case with statistical and experimental uncertainties (5%)
- The case with statistical, experimental and additional theoretical uncertainties at 5%, 10% and 30% respectively.

Three different sources of uncertainties are treated as uncorrelated.

### 5.8.6 Results

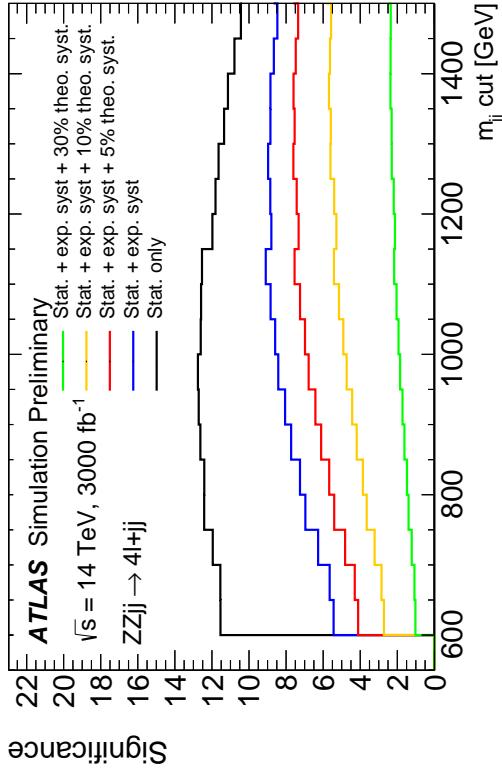
In this analysis, the expected significance of EW-ZZjj production is calculated as:

$$\text{Significance} = \frac{S}{\sqrt{\sigma(B)_{\text{stat.}}^2 + \sigma(B)_{\text{syst.}}^2}}, \quad (5.10)$$

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 is computed from expected yield at  $3000 \text{ fb}^{-1}$ .

Base on baseline selection of  $m_{jj} > 600 \text{ GeV}$ , a additional scan over different  $m_{jj}$  cuts are performed with a step of 50 GeV for luminosity of  $3000 \text{ fb}^{-1}$  under different systematic conditions, as shown in figure 5.17.

In addition, the expected differential cross section of EW-ZZjj process is measured in the defined phase space at 14 TeV, as a function of  $m_{ZZ}$  and  $m_{jj}$ , shown in figure 5.18. The expected differential cross sections are calculated as:



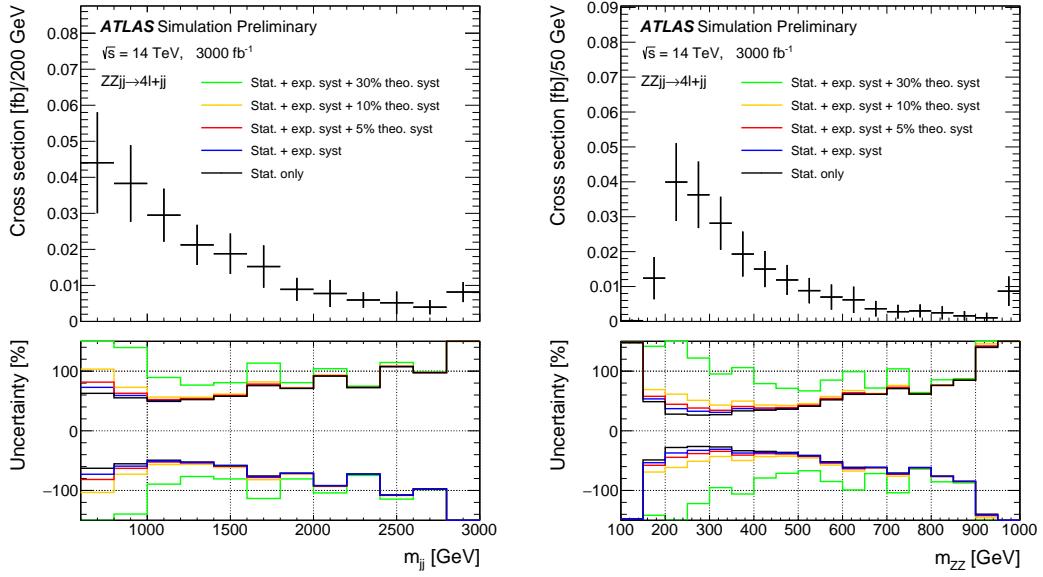
**Fig. 5.17** The expected significance of EW-ZZjj processes as a function of different  $m_{jj}$  cut with  $3000 \text{ fb}^{-1}$ , under conditions of different sizes of theoretical uncertainties on the QCD-ZZjj background modelling. The statistical uncertainty is estimated from expected data yield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

$$\sigma = \frac{N_{\text{pseudo-data}} - N_{QCD-ZZjj}}{L * C_{EW-ZZjj}} \quad (5.11)$$

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

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

The number of expected integrated cross section as well as its uncertainty under different uncertainty conditions are shown in table 5.15 in  $3000 \text{ fb}^{-1}$  luminosity at 14 TeV. The statistical uncertainty is at 10% level when with such large luminosity. The result is dominated by systematics and can reach 100% level when theoretical modeling uncertainty is 30% for QCD-ZZjj processes.



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

	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 5.15** Summary of expected cross-section measurements with different theoretical uncertainties. The statistical uncertainty is estimated from expected data yield at 14 TeV with  $3000 \text{ fb}^{-1}$ . Different uncertainties are summed up quadratically.

## 5.9 Conclusion

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

In addition, the prospect study for the EW- $ZZjj$  production at the HL-LHC in the  $\ell\ell\ell\ell$  channel, using  $3000 \text{ fb}^{-1}$  simulated pp collision at a centre-of-mass energy of 14 TeV has been presented. The precision of the expected measurements of the integrated and differential cross sections as a function of dijet or 4l invariant mass are shown. Under the assumption of theoretical uncertainty being constraint at 5% level for the QCD- $ZZjj$  processes, the observation of the EW- $ZZjj$  process can be reached with a significance of  $7 \sigma$ .

## Bibliography

- [1] LANGACKER P. Introduction to the Standard Model and Electroweak Physics[C/OL]// Proceedings of Theoretical Advanced Study Institute in Elementary Particle Physics on The dawn of the LHC era (TASI 2008): Boulder, USA, June 2-27, 2008. 2010: 3-48. DOI: [10.1142/9789812838360\\_0001](https://doi.org/10.1142/9789812838360_0001).
- [2] PICH A. Electroweak Symmetry Breaking and the Higgs Boson[J/OL]. Acta Phys. Polon., 2016, B47:151. DOI: [10.5506/APhysPolB.47.151](https://doi.org/10.5506/APhysPolB.47.151).
- [3] DREMIN I M. Soft and hard processes in QCD[J/OL]. JETP Lett., 2005, 81:307-310. DOI: [10.1134/I.1944068](https://doi.org/10.1134/I.1944068).
- [4] WOMERSLEY J. QCD at the Tevatron: Status and prospects[C/OL]//Proceedings, 5th International Symposium on Radiative Corrections - RADCOR 2000. 2000. <http://www.slac.stanford.edu/econf/C000911/>.
- [5] LIN H W, et al. Parton distributions and lattice QCD calculations: a community white paper [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).
- [6] COLLINS J C, SOPER D E, STERMAN G F. Factorization of Hard Processes in QCD[J/OL]. Adv. Ser. Direct. High Energy Phys., 1989, 5:1-91. DOI: [10.1142/9789814503266\\_0001](https://doi.org/10.1142/9789814503266_0001).
- [7] STIRLING W J. Perturbative QCD[J/OL]. 2000(CERN-OPEN-2000-296):40 p. <https://cds.cern.ch/record/1194745>. DOI: [10.5170/CERN-2000-007.305](https://doi.org/10.5170/CERN-2000-007.305).
- [8] GROJEAN C. Higgs Physics[J/OL]. 2017(arXiv:1708.00794):143-158. 12 p. <https://cds.cern.ch/record/2243593>. DOI: [10.5170/CERN-2016-005.143](https://doi.org/10.5170/CERN-2016-005.143).
- [9] Cern yellow reports: Monographs: Handbook of LHC Higgs Cross Sections: 4. Deciphering 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).
- [10] HEINEMEYER S. Cern yellow reports: Monographs: Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group[M/OL]. 2013. <https://cds.cern.ch/record/1559921>. DOI: [10.5170/CERN-2013-004](https://doi.org/10.5170/CERN-2013-004).
- [11] CHANG J, CHEUNG K, LU C T, et al.  $w\bar{w}$  scattering in the era of post-higgs-boson discovery [J/OL]. Phys. Rev. D, 2013, 87:093005. <https://link.aps.org/doi/10.1103/PhysRevD.87.093005>.
- [12] BRUNING O S, COLLIER P, LEBRUN P, et al. LHC Design Report Vol.1: The LHC Main Ring[J]. 2004.
- [13] BUNING O, COLLIER P, LEBRUN P, et al. LHC Design Report. 2. The LHC infrastructure and general services[J]. 2004.

## Bibliography

---

- [14] BENEDIKT M, COLLIER P, MERTENS V, et al. LHC Design Report. 3. The LHC injector chain[J]. 2004.
- [15] EVANS L, BRYANT P. LHC machine[J/OL]. Journal of Instrumentation, 2008, 3(08):S08001-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).
- [16] MOBS E. The CERN accelerator complex. Complexe des accélérateurs du CERN[J/OL]. 2016. <https://cds.cern.ch/record/2197559>.
- [17] COLLABORATION A. The ATLAS experiment at the CERN large hadron collider[J/OL]. 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).
- [18] PEREZ G. Unitarization models for vector boson scattering at the lhc[D/OL]. 2018. DOI: [10.5445/IR/1000082199](https://doi.org/10.5445/IR/1000082199).
- [19] PEQUENAO J. Computer generated image of the whole ATLAS detector[Z/OL]. 2008. <https://cds.cern.ch/record/1095924>.
- [20] MCFAYDEN J. The lhc and atlas detector. in: Third generation susy and  $t\bar{t} + z$  productin. [D/OL]. 2014. DOI: [10.1007/978-3-319-07191-6\\_2](https://doi.org/10.1007/978-3-319-07191-6_2).
- [21] COLLABRATION A. Operation and performance of the ATLAS semiconductor tracker. Operation and performance of the ATLAS semiconductor tracker[J/OL]. JINST, 2014, 9(CERN-PH-EP-2014-049. CERN-PH-EP-2014-049):P08009. 80 p. <https://cds.cern.ch/record/1698966>. DOI: [10.1088/1748-0221/9/08/P08009](https://doi.org/10.1088/1748-0221/9/08/P08009).
- [22] MULLIER G. The upgraded pixel detector of the atlas experiment for run-2 at the large hadron 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).
- [23] SULTAN D M S. Development of small-pitch, thin 3d sensors for pixel detector upgrades at 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).
- [24] The ATLAS transition radiation tracker (TRT) proportional drift tube: design and performance [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).
- [25] BUCHANAN N, CHEN L, GINGRICH D, et al. Design and implementation of the front end board for the readout of the atlas liquid argon calorimeters[J/OL]. Journal of Instrumentation, 2008, 3:P03004. DOI: [10.1088/1748-0221/3/03/P03004](https://doi.org/10.1088/1748-0221/3/03/P03004).
- [26] Technical design report atlas: ATLAS liquid-argon calorimeter: Technical Design Report [M/OL]. Geneva: CERN, 1996. <https://cds.cern.ch/record/331061>.
- [27] AAD G, ABBOTT B, ABDALLAH J, et al. Readiness of the atlas tile calorimeter for lhc collisions[J]. European Physical Journal C, 2010.

## Bibliography

---

- [28] SLIWA K. "ATLAS Overview and Main Results"[C]//Proceedings, International School on High Energy Physics : Workshop on High Energy Physics in the near Future. (LISHEP 2013): Rio de Janeiro, Brazil, March 17-24, 2013. 2013.
- [29] RUIZ-MARTINEZ A, COLLABORATION A. The Run-2 ATLAS Trigger System: ATL-DAQ-PROC-2016-003[R/OL]. Geneva: CERN, 2016. <https://cds.cern.ch/record/2133909>. DOI: [10.1088/1742-6596/762/1/012003](https://doi.org/10.1088/1742-6596/762/1/012003).
- [30] PÁSZTOR G. The Upgrade of the ATLAS Electron and Photon Triggers towards LHC Run 2 and their Performance: ATL-DAQ-PROC-2015-053. arXiv:1511.00334[R/OL]. Geneva: CERN, 2015. <https://cds.cern.ch/record/2063746>.
- [31] COLLABORATION A. Atlas computing: Technical design report[R/OL]. Geneva: CERN. <http://atlas-computing.web.cern.ch/atlas-computing/packages/athenaCore/athenaCore.php>.
- [32] AAD G, et al. The ATLAS Simulation Infrastructure[J/OL]. Eur. Phys. J., 2010, C70:823-874. DOI: [10.1140/epjc/s10052-010-1429-9](https://doi.org/10.1140/epjc/s10052-010-1429-9).
- [33] HOCHE S. Introduction to parton-shower event generators[C/OL]//Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014. 2015: 235-295. DOI: [10.1142/9789814678766\\_0005](https://doi.org/10.1142/9789814678766_0005).
- [34] 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/10.1088/1126-6708/2009/02/007>. DOI: [10.1088/1126-6708/2009/02/007](https://doi.org/10.1088/1126-6708/2009/02/007).
- [35] 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>.
- [36] 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).
- [37] 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/10.1088/1126-6708/2002/06/029>. DOI: [10.1088/1126-6708/2002/06/029](https://doi.org/10.1088/1126-6708/2002/06/029).
- [38] 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).
- [39] 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).
- [40] RICHTER-WAS E, FROIDEVAUX D, POGGIOLI L. ATLFAST 2.0 a fast simulation package

## Bibliography

---

- for ATLAS: ATL-PHYS-98-131[R/OL]. Geneva: CERN, 1998. <https://cds.cern.ch/record/683751>.
- [41] 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>.
- [42] BOYD J T, COSTA M J, TONOVAN 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.
- [43] 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>.
- [44] 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/s1052-016-4120-y](https://doi.org/10.1140/epjc/s1052-016-4120-y).
- [45] ILLINGWORTH J, KITTLER J. A survey of the hough transform[J/OL]. Computer Vision, 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).
- [46] Vertex Reconstruction Performance of the ATLAS Detector at 13 TeV: ATL-PHYS-PUB-2015-026[R/OL]. Geneva: CERN, 2015. <https://cds.cern.ch/record/2037717>.
- [47] AABOUD M, et al. Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton-proton collisions at the LHC[J/OL]. Eur. Phys. J., 2017, C77(5):332. DOI: [10.1140/epjc/s10052-017-4887-5](https://doi.org/10.1140/epjc/s10052-017-4887-5).
- [48] Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data: ATLAS-CONF-2016-024[R/OL]. Geneva: CERN, 2016. <https://cds.cern.ch/record/2157687>.
- [49] LAMPL W, LAPLACE S, LELAS D, et al. Calorimeter Clustering Algorithms: Description and Performance: ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003[R/OL]. Geneva: CERN, 2008. <https://cds.cern.ch/record/1099735>.
- [50] CORNELISSEN T G, ELSING M, GAVRILENKO I, et al. The global  $\chi^2$ track fitter in ATLAS [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.
- [51] Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data: ATLAS-CONF-2014-032[R/OL]. Geneva: CERN, 2014. <http://cds.cern.ch/record/1706245>.

## Bibliography

---

- [52] LIMPER M. Track and vertex reconstruction in the ATLAS inner detector[D/OL]. 2009. <http://cds.cern.ch/record/1202457>.
- [53] AAD G, et al. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1[J/OL]. *Eur. Phys. J.*, 2017, C77:490. DOI: [10.1140/epjc/s10052-017-5004-5](https://doi.org/10.1140/epjc/s10052-017-5004-5).
- [54] ZHENG Z. Identification of very-low transverse momentum muons with the ATLAS experiment: ATL-MUON-PROC-2018-018[R/OL]. Geneva: CERN, 2018. <https://cds.cern.ch/record/2649299>.
- [55] CACCIARI M, SALAM G P. Pileup subtraction using jet areas[J/OL]. *Physics Letters B*, 2008, 659(1):119 - 126. <http://www.sciencedirect.com/science/article/pii/S0370269307011094>. DOI: <https://doi.org/10.1016/j.physletb.2007.09.077>.
- [56] CACCIARI M, SALAM G P, SOYEZ G. The anti-ktjet clustering algorithm[J/OL]. *Journal of 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](https://doi.org/10.1088/1126-6708/2008/04/063).
- [57] CACCIARI M, SALAM G P, SOYEZ G. Fastjet user manual[J/OL]. *The European Physical Journal C*, 2012, 72(3):1896. <https://doi.org/10.1140/epjc/s10052-012-1896-2>.
- [58] AAD G, et al. Jet energy measurement and its systematic uncertainty in proton-proton collisions 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).
- [59] AABOUD M, et al. Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector[J/OL]. *Phys. Rev.*, 2017, D96(7): 072002. DOI: [10.1103/PhysRevD.96.072002](https://doi.org/10.1103/PhysRevD.96.072002).
- [60] Optimisation of the ATLAS *b*-tagging performance for the 2016 LHC Run: ATL-PHYS-PUB-2016-012[R/OL]. Geneva: CERN, 2016. <https://cds.cern.ch/record/2160731>.
- [61] PIACQUADIO G, WEISER C. A new inclusive secondary vertex algorithm for *b*-jet tagging 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).
- [62] SPECKMAYER P, HöCKER A, STELZER J, et al. The toolkit for multivariate data analysis, 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).
- [63] Performance of missing transverse momentum reconstruction with the atlas detector using proton–proton collisions at  $\sqrt{s}=13\text{TeV}$ [J/OL]. *The European Physical Journal C*, 2018, 78(11):903. <https://doi.org/10.1140/epjc/s10052-018-6288-9>.
- [64] Observation of a new particle in the search for the standard model higgs boson with the atlas 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.
- [65] Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc[J/OL]. 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>.
- [66] LEE B W, QUIGG C, THACKER H B. The Strength of Weak Interactions at Very High-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).
- [67] CHANOWITZ M S, GAILLARD M K. The TeV Physics of Strongly Interacting W's and Z's [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).
- [68] SZLEPER M. The Higgs boson and the physics of  $WW$  scattering before and after Higgs discovery[J]. 2014.
- [69] Evidence for electroweak production of  $W^\pm W^\pm jj$  in  $pp$  collisions at  $\sqrt{s} = 8$  TeV with the atlas detector[J/OL]. Phys. Rev. Lett., 2014, 113:141803. <https://link.aps.org/doi/10.1103/PhysRevLett.113.141803>.
- [70] Observation of electroweak production of a same-sign  $w$  boson pair in association with two jets in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the atlas detector[J/OL]. Phys. Rev. Lett., 2019, 123:161801. <https://link.aps.org/doi/10.1103/PhysRevLett.123.161801>.
- [71] SIRUNYAN A M, et al. Observation of electroweak production of same-sign W boson pairs in the two jet and two same-sign lepton final state in proton-proton collisions at  $\sqrt{s} = 13$  TeV [J/OL]. Phys. Rev. Lett., 2018, 120(8):081801. DOI: [10.1103/PhysRevLett.120.081801](https://doi.org/10.1103/PhysRevLett.120.081801).
- [72] Observation of electroweak  $w\pm z$  boson pair production in association with two jets in  $pp$  collisions at  $s=13$  tev with the atlas detector[J/OL]. Physics Letters B, 2019, 793:469 - 492. <http://www.sciencedirect.com/science/article/pii/S0370269319303211>. DOI: <https://doi.org/10.1016/j.physletb.2019.05.012>.
- [73] Measurement of vector boson scattering and constraints on anomalous quartic couplings from events with four leptons and two jets in proton–proton collisions at  $s=13$  tev[J/OL]. Physics 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>.
- [74] ALWALL J, FREDERIX R, FRIXIONE S, et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations[J/OL]. JHEP, 2014, 07:079. DOI: [10.1007/JHEP07\(2014\)079](https://doi.org/10.1007/JHEP07(2014)079).
- [75] BALL R D, et al. Parton distributions with LHC data[J/OL]. Nucl. Phys. B, 2013, 867:244-289. DOI: [10.1016/j.nuclphysb.2012.10.003](https://doi.org/10.1016/j.nuclphysb.2012.10.003).
- [76] GLEISBERG T, HÖCHE S, KRAUSS F, et al. Event generation with SHERPA 1.1[J/OL]. JHEP, 2009, 02:007. DOI: [10.1088/1126-6708/2009/02/007](https://doi.org/10.1088/1126-6708/2009/02/007).

## Bibliography

---

- [77] BALL R D, et al. Parton distributions for the LHC run II[J/OL]. *JHEP*, 2015, 04:040. DOI: [10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040).
- [78] CAOLA F, MELNIKOV K, RÖNTSCH R, et al. Qcd corrections to  $zz$  production in gluon fusion at the lhc[J/OL]. *Phys. Rev. D*, 2015, 92:094028. <https://link.aps.org/doi/10.1103/PhysRevD.92.094028>.
- [79] FRIXIONE S, RIDOLFI G, NASON P. A positive-weight next-to-leading-order Monte Carlo 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).
- [80] LAI H L, et al. New parton distributions for collider physics[J/OL]. *Phys. Rev. D*, 2010, 82: 074024. DOI: [10.1103/PhysRevD.82.074024](https://doi.org/10.1103/PhysRevD.82.074024).
- [81] ALIOLI S, NASON P, OLEARI C, et al. NLO single-top production matched with shower in 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).
- [82] FREDERIX R, RE E, TORRIELLI P. Single-top t-channel hadroproduction in the four-flavour 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).
- [83] RE E. Single-top Wt-channel production matched with parton showers using the POWHEG 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).
- [84] COLLABORATION A. ATLAS Pythia 8 tunes to 7 TeV data[M/OL]. 2014. <https://cds.cern.ch/record/1966419>.
- [85] BELLM J, et al. Herwig 7.0/Herwig++ 3.0 release note[J/OL]. *Eur. Phys. J. C*, 2016, 76(4): 196. DOI: [10.1140/epjc/s10052-016-4018-8](https://doi.org/10.1140/epjc/s10052-016-4018-8).
- [86] BAHR M, et al. Herwig++ Physics and Manual[J/OL]. *Eur. Phys. J. C*, 2008, 58:639-707. DOI: [10.1140/epjc/s10052-008-0798-9](https://doi.org/10.1140/epjc/s10052-008-0798-9).
- [87] Luminosity determination in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector at the LHC: ATLAS-CONF-2019-021[R/OL]. Geneva: CERN, 2019. <http://cds.cern.ch/record/2677054>.
- [88] The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS[J/OL]. *Journal of Instrumentation*, 2018, 13(07):P07017-P07017. <https://doi.org/10.1088%2F1748-0221%2F13%2F07%2Fp07017>. DOI: [10.1088/1748-0221/13/07/p07017](https://doi.org/10.1088/1748-0221/13/07/p07017).
- [89] COADOU Y. Boosted decision trees and applications[J/OL]. *EPJ Web of Conferences*, 2013, 55:02004-. DOI: [10.1051/epjconf/20135502004](https://doi.org/10.1051/epjconf/20135502004).
- [90] COLLABORATION T A. Observation of electroweak production of two jets in association with a  $Z$ -boson pair in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector[J]. 2019.
- [91] COLLABORATION A. Technical Design Report for the ATLAS Inner Tracker Pixel Detector:

## Bibliography

---

- CERN-LHCC-2017-021. ATLAS-TDR-030[R/OL]. Geneva: CERN, 2017. <https://cds.cern.ch/record/2285585>.
- [92] COLLABORATION A. Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer: CERN-LHCC-2017-017. ATLAS-TDR-026[R/OL]. Geneva: CERN, 2017. <https://cds.cern.ch/record/2285580>.
- [93] COLLABORATION A. Technical Proposal: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade: CERN-LHCC-2018-023. LHCC-P-012[R/OL]. Geneva: CERN, 2018. <https://cds.cern.ch/record/2623663>.
- [94] Expected performance for an upgraded ATLAS detector at High-Luminosity LHC: ATL-PHYS-PUB-2016-026[R/OL]. Geneva: CERN, 2016. <https://cds.cern.ch/record/2223839>.