# Search for ZZ Resonances in the Final State of Leptons and Neutrinos in Proton-Proton Collisions at 13 TeV

Yanchu Wang Charlottesville, VA

B.S., University of Science and Technology of China, 2013

A Thesis presented to the Graduate Faculty of the University of Virginia in Candidacy for the Degree of Doctor of Philosophy

**Department of Physics** 

University of Virginia May, 2019 We approve the dissertation of Yanchu Wang.

Date of Signature

Supervisor: Dr. Bob Hirosky

Buelly Ge

4 9 0

4/29/18

Reviewer: Dr. Bradley Cox

29 /19 a Reviewer: Dr. P.Q. Hung 19 Independent Reviewer: Dr. Nitya Kallivayalil

# ACKNOWLEDGEMENT

First and foremost, I would like to express my immerse gratitude to my advisor Prof. Bob Hirosky, for his continual guidance and support throughout the last 4 years for me. He guided me into the world of experimental particle physics and fascinated me with his physics insight and technical skills.

I would also like to express my gratitude to Dr. Hengne Li, the former post-doctor in our group who now works as a professor at the South China Normal University. His dedication to research has always been an invaluable source of inspiration to me.

Moreover, I appreciate the efforts from the members of my defense committee, including Dr. Brad Cox, Dr. P.Q. Hung and Dr. Nitya Kallivayalil, for their time reviewing my research, providing precious comments for my thesis, and approving my dissertation.

In addition, I am grateful to the physicists I worked with and learnt from during my 2 years' stay at CERN, such as Tote Hughes and Andrew Whitbeck in the HCAL Phase I upgrading team, Mario Galanti and Josh Kunkle in the HCAL Operation team, Thomas Peiffer and Vieri Candelise in the B2G MC Production team, and many more amazing researchers in the collaboration.

Finally, I could not have achieved this without the encouragement and love from my parents. I am grateful to my mother Cailing Hu and father Zhiping Wang for supporting me along the way.

Yanchu Wang Charlottesville, VA April, 2019

# Preface

# — CMS and me

In August 2014, I joined the UVA CMS group and worked on calibrating a prototype Shashlik calorimeter using the muon data as my first project here. Afterwards, in the middle of 2015, I left for CERN, Switzerland and stayed there for 2 years. During my stay at CERN I was involved in several projects and activities both on the detector and physics analysis.

My most detector-related work focused on the Hadron Calorimeter (HCAL). Starting with the Hadron Forward (HF) detector frontend electronics Phase I upgrade in 2015, I joined the HCAL upgrade team, helping design and carry out a series of electronics tests. In 2016 I helped the HCAL Endcap (HE) upgrade with system monitoring and the HCAL Data Quality Management (DQM) group with their online DQM system for the HF/HE upgrades. I was also taking HCAL detector on call expert shifts continually for 1 year and a half since 2016.

In terms of physics analysis, I started working on this di-boson analysis [1] in the December of 2015, helping to design and implement the analysis framework from scratch, studying the pileup reweighting, trigger efficiency, lepton identification algorithms and their efficiencies, as well as determining the non-resonant background using data-driven modeling. The muon tracker High  $p_T$  efficiency calculated by me has been widely used in related CMS analyses and I am also contributing to the high  $p_T$ muon paper carried out by the muon Physics Object Group (POG). Apart from this di-boson analysis, I also worked as the Monte Carlo/Generator contact person for physics simulations in the Beyond Two Generations (B2G) Physics Analysis Group (PAG) for the whole year of 2017.

# Contents

1	Intr	oductio	on	9
	1.1	The Standard Model		9
		1.1.1	Fermions	9
		1.1.2	Bosons and the interactions	11
	1.2	The L	imitations of SM and the Hierarchy Problem	13
	1.3	Bulk RS Graviton Model		14
		1.3.1	Introduction of Extra Dimension Models	14
		1.3.2	Randall–Sundrum Model	15
		1.3.3	The Bulk RS Graviton Model	16
		1.3.4	Phenomenology of the Bulk Graviton Model	17
	1.4	4 Status of Searches for the Bulk Graviton		19
		1.4.1	Previous Searches	19
		1.4.2	$2\ell 2 u$ channel and Search Strategy	19
2	The	LHC a	and CMS detector	23
	2.1	The Large Hadron Collider		
	2.2	The Compact Muon Solenoid (CMS) Experiment		26
		2.2.1	Inner Tracking System	29
		2.2.2	Electromagnetic Calorimeter	31
		2.2.3	Hadronic Calorimeter	32
		2.2.4	Muon System	34
		2.2.5	Data Acquisition and Trigger System	36
3	Phy	sics Ob	jects Reconstruction	39
	3.1 Particle Flow Algorithm		e Flow Algorithm	39
		3.1.1	Charged-particle tracks and vertices	39

		3.1.2	Calorimeter clusters	40
		3.1.3	Tracks for electrons	41
		3.1.4	Tracks for muons	42
		3.1.5	Link algorithm	42
	3.2	Electro	on Reconstruction	43
		3.2.1	Electron Identification and Isolation	43
	3.3	Photor	n Reconstruction	45
		3.3.1	Photon Identification and Isolation	46
	3.4	Muon	Reconstruction	46
		3.4.1	Muon Identification and Isolation	47
	3.5	Missir	ng Transverse Momentum ( $p_T^{\text{miss}}$ )	49
4	Data	asets, E	vent Selections and Event Reconstruction	51
	4.1	Data a	nd Triggers	51
		4.1.1	Data	51
		4.1.2	High Level Triggers	52
	4.2	Monte	Carlo Simulation Samples	52
		4.2.1	Standard Model Samples	53
		4.2.2	Signal Samples	54
	4.3	Event	Preselection and Reconstruction	55
		4.3.1	Lepton Selection	55
		4.3.2	Leptonic Z Boson Reconstruction	58
		4.3.3	MET Filters	58
	4.4	Signal	Region	59
5	Bac	kgroun	d Modeling and Plots	61
	5.1	Pileup	re-weighting	62
	5.2	Monte	Carlo Efficiencies	62
		5.2.1	Muon ID/Iso Efficiency	63
		5.2.2	Electron ID/Iso Efficiency	67
		5.2.3	Trigger Efficiency	68
	5.3	Z+iet	ts Background Modeling	75
		5.3.1	$\gamma$ +jets Event Selection	77
		5.3.2	$\gamma$ +jets HLT Prescale Reweighting	78

		5.3.3	Physical $p_T^{\text{miss}}$ Subtraction	79
		5.3.4	$p_T^{\gamma}$ to $p_T^{Z}$ Reweighting	82
		5.3.5	Photon Mass Generation	82
		5.3.6	$p_T^{\text{miss}}$ Hadronic Recoil Tuning	83
	5.4	Non-re	esonant Background Modeling	89
		5.4.1	$e\mu$ Pair Selection	89
		5.4.2	$e\mu$ Pair Event-based Reweighting $\ldots \ldots \ldots$	89
		5.4.3	HLT Efficiency Reweighting	92
		5.4.4	$e\mu$ Events Rescaling $\ldots \ldots \ldots \ldots \ldots \ldots$	93
		5.4.5	Comparing to MC Modeling	94
		5.4.6	Uncertainty Table	97
	5.5	Resona	ant Background Modeling	100
	5.6	Presele	ection Plots	102
	5.7	Plots in	n the Signal Region	108
6	Syst	ematic	Uncertainties and Interpretation of Results	115
6	<b>Syst</b> 6.1	ematic Systen	Uncertainties and Interpretation of Results	<b>115</b> 115
6	Syst 6.1	ematic Systen 6.1.1	Uncertainties and Interpretation of Results natic Uncertainty	<b>115</b> 115 115
6	<b>Syst</b> 6.1	ematic Systen 6.1.1 6.1.2	Uncertainties and Interpretation of Resultsnatic UncertaintyIntegrated Luminosity UncertaintyPDF and QCD Scale Uncertainty	<ul><li>115</li><li>115</li><li>115</li><li>115</li></ul>
6	Syst 6.1	ematic Systen 6.1.1 6.1.2 6.1.3	Uncertainties and Interpretation of Resultsnatic UncertaintyIntegrated Luminosity UncertaintyPDF and QCD Scale UncertaintyTrigger/Lepton Uncertainty	<b>115</b> 115 115 115 115
6	Syst 6.1	ematic Systen 6.1.1 6.1.2 6.1.3 6.1.4	Uncertainties and Interpretation of Resultsnatic UncertaintyIntegrated Luminosity UncertaintyPDF and QCD Scale UncertaintyTrigger/Lepton UncertaintyMissing Transverse Energy Uncertainty	<ul> <li>115</li> <li>115</li> <li>115</li> <li>115</li> <li>116</li> <li>116</li> </ul>
6	Syst 6.1	ematic Systen 6.1.1 6.1.2 6.1.3 6.1.4 6.1.5	Uncertainties and Interpretation of Resultsnatic UncertaintyIntegrated Luminosity UncertaintyPDF and QCD Scale UncertaintyTrigger/Lepton UncertaintyMissing Transverse Energy UncertaintyUncertainties in the Background Modeling methods	<ul> <li>115</li> <li>115</li> <li>115</li> <li>116</li> <li>116</li> <li>117</li> </ul>
6	Syst 6.1	ematic Systen 6.1.1 6.1.2 6.1.3 6.1.4 6.1.5 6.1.6	Uncertainties and Interpretation of Resultsnatic UncertaintyIntegrated Luminosity UncertaintyPDF and QCD Scale UncertaintyTrigger/Lepton UncertaintyMissing Transverse Energy UncertaintyUncertainties in the Background Modeling methodsSystematic Uncertainty Summary	<ul> <li>115</li> <li>115</li> <li>115</li> <li>116</li> <li>116</li> <li>117</li> <li>117</li> </ul>
6	<b>Syst</b> 6.1 6.2	ematic Systen 6.1.1 6.1.2 6.1.3 6.1.4 6.1.5 6.1.6 Result	Uncertainties and Interpretation of Results         natic Uncertainty         Integrated Luminosity Uncertainty         PDF and QCD Scale Uncertainty         Trigger/Lepton Uncertainty         Missing Transverse Energy Uncertainty         Uncertainties in the Background Modeling methods         Systematic Uncertainty Summary         s and Interpretation	<ul> <li>115</li> <li>115</li> <li>115</li> <li>116</li> <li>116</li> <li>117</li> <li>117</li> <li>120</li> </ul>
6	<b>Syst</b> 6.1 6.2 <b>Sum</b>	ematic Systen 6.1.1 6.1.2 6.1.3 6.1.4 6.1.5 6.1.6 Result	Uncertainties and Interpretation of Results natic Uncertainty	<ul> <li>115</li> <li>115</li> <li>115</li> <li>116</li> <li>116</li> <li>117</li> <li>117</li> <li>120</li> <li>127</li> </ul>
<b>6</b> 7	<b>Syst</b> 6.1 6.2 <b>Sum</b> 7.1	ematic Systen 6.1.1 6.1.2 6.1.3 6.1.4 6.1.5 6.1.6 Result mary a Analys	Uncertainties and Interpretation of Results natic Uncertainty	<ul> <li>115</li> <li>115</li> <li>115</li> <li>116</li> <li>116</li> <li>117</li> <li>120</li> <li>127</li> <li>127</li> </ul>
6	Syst 6.1 6.2 Sum 7.1 7.2	ematic Systen 6.1.1 6.1.2 6.1.3 6.1.4 6.1.5 6.1.6 Result mary a Analys Compa	Uncertainties and Interpretation of Results         natic Uncertainty	<ul> <li>115</li> <li>115</li> <li>115</li> <li>116</li> <li>116</li> <li>117</li> <li>120</li> <li>127</li> <li>127</li> <li>127</li> </ul>

# **Chapter 1**

# Introduction

# **1.1 The Standard Model**

The Standard Model (SM)[2, 3, 4] is the fundamental theory of elementary particle physics. After half a century's development, the Standard Model framework has been confirmed by numerous experiments and can be used to explain most of the data collected about interactions of the fundamental particles. In the Standard Model there are generally 2 categories of particles: fermions and bosons. Fermions, which always have half-integer spins, make up all the matter in the universe. On the other hand, bosons with integer spins mediate the fundamental interactions among the fermions, and the interactions here include the electro-magnetic interaction, the weak interaction and the strong interaction. Figure 1.1 shows all the particles that have been discovered and included in the Standard Model.

### 1.1.1 Fermions

Fermions are particles that follow Fermi–Dirac statistics and obey the Pauli exclusion principle. Every fermion has its anti particle, which is a fermion with opposite charge but same mass and spin. The elementary fermions in the Standard Model include leptons and quarks, both of them consist of 3 generations and each generation consists of 2 flavors of fermions. Every elementary fermion in the Standard Model has a spin of one half.



#### **Standard Model of Elementary Particles**

Figure 1.1: The Standard Model contains 3 generations of leptons and quarks, 4 kinds of vector bosons and the Higgs boson.

#### Lepton

The leptons do not participate in the strong interaction. There are 3 generations of leptons, and each generation consists of two flavors of leptons: one charged lepton, and one neutral lepton also known as a "neutrino". The charged lepton always carries 1 unit of elementary electric charge, negative for a lepton and positive for its anti-lepton. The masses of the charged leptons have been precisely measured as shown in Figure 1.1. In the SM, neutrinos are regarded to be massless.

The first generation of leptons includes the electron  $(e^-)$  and electronneutrino  $(\nu_e)$ . Both electron and electron-neutrino have an electron number  $L_e = 1$ , while their anti leptons have  $L_e = -1$ . The second generation leptons, the muon  $(\mu)$  and muon-neutrino  $(\nu_{\mu})$ , have a muon number  $L_{\mu} =$ 1, while  $L_{\mu} = -1$  for the anti-muon  $(\bar{\mu})$  and anti-muon-neutrino  $(\bar{\nu}_{\mu})$ . Similarly, the third generation consists of the tau  $(\tau)$  and tau-neutrino  $(\nu_{\tau})$ , with a tau number  $L_{\tau}$  accordingly.

In the SM, under the assumption that neutrinos are massless, the lepton numbers are strictly conserved in any kind of interaction. However, recent experiments [5, 6] indicate that the neutrinos have small masses, which implies the lepton numbers can be mixed among different generations.

## Quark

A quark can participate in any of the 3 interactions in the Standard Model. Like leptons, quarks fall into 3 generations, and each generation contains 2 flavors of quarks having 2/3 and -1/3 elementary electric charge correspondingly. Besides electric charge, quarks also have another intrinsic property known as color charge. The color charge a quark carries is typically defined to be red, blue or green, while the anti quark carries a corresponding anti-color. Quarks are never observed directly in an isolated state due to the phenomenon of color confinement, which confines quarks to only exist in composite, colorless particles known as hadrons.

The first generation includes up and down quarks, the second generation includes charm and strange quarks, and the third generation consists of top and bottom quarks. Every quark has its anti quark, with opposite charge. The mass for each quark is shown in Figure 1.1. It is possible for heavier quarks to decay into lighter quarks through the weak interaction, especially for quarks within the same generation.

## **1.1.2** Bosons and the interactions

In contrast to fermions, bosons are particles that follow Bose–Einstein statistics allowing multiple particles in the same state. In the SM there are 4 kinds of vector bosons each with spin 1 and one spin-0 scalar boson, which is the recently discovered Higgs boson. The vector bosons work as force carriers of the 3 interactions, while the masses of the elementary particles are generated by their interaction with the Higgs field.

#### **Vector Bosons**

In the SM, the vector bosons (gluon, photon, Z boson and W boson) work as mediators of the 3 fundamental interactions among fermions. Each of the vector bosons has spin 1.

- the **gluon** is the force mediator of the strong interaction, described in Quantum Chromodynamics (QCD) theory, a gauge theory based on SU(3). Gluons are massless and have no electric charge.
- the **photon** is the force mediator of the electromagnetic interaction, described in Quantum Electrodynamics (QED) theory, a  $U(1)_{EM}$  theory. Photons are massless and have no electric charge.
- the **Z boson** is the mediator of the weak interaction with no electric charge flow. The Z boson has no charge, while it has mass of 91.2 GeV, which makes it possible to decay into a fermion–antifermion pair.
- the W boson is the mediator of the weak interaction with electric charge flow. The W carries either 1 or -1 elementary electric charge, denoted as W<sup>+</sup> and W<sup>-</sup>, both having masses of 80.4 GeV. Like the Z boson, W bosons can decay into fermion–antifermion pairs.

In the SM, the unification of the electromagnetic and weak interactions is realized by an SU(2)  $\times$  U(1) gauge group. The photon, Z boson and W boson are generated in the SU(2)  $\times$  U(1) group due to the process of spontaneous symmetry breaking described by the Higgs mechanism.

#### The Higgs Boson

In the 1960s the Higgs boson and Higgs mechanism was proposed in order to explain the source of the masses of the gauge bosons[7, 8, 9], which are expected to be massless according to the gauge theory. That assumption clearly conflicts with the experiment facts. The Higgs mechanism suggests that an SU(2) doublet of complex scalar fields breaks the SU(2) symmetry as its potential in the form of  $\mu^2 \phi^{\dagger} \phi + \lambda^2 (\phi^{\dagger} \phi)^2$  with  $\mu^2 < 0$  and  $\lambda^2 > 0$  leads to a non-zero vacuum expectation value which does not follow the SU(2) symmetry. This phenomenon is referred to as "spontaneous symmetry breaking" and the field here is the Higgs field. Due to the spontaneous symmetry breaking, three gauge bosons gain masses when interacting with the scalar field, and the Higgs Boson comes from one degree of freedom of the field while the other 3 degrees are no longer observable.

On July 4 of 2012, the discovery of the Higgs boson was announced by both the CMS and ATLAS Collaborations at the CERN, LHC[10, 11]. Hence the Higgs boson officially became one member of the standard model elementary particle family. The Higgs boson has no spin or electric charge, and its mass is measured to be 125 GeV. It is very unstable and mainly decays into a  $b\bar{b}$  quark pair,  $\tau\bar{\tau}$  pair or off shell gauge boson pair.

# 1.2 The Limitations of SM and the Hierarchy Problem

Although the SM has been tested and demonstrated as a great success among numerous particle physics experiments and provides reliable physics predictions for most of the sceneries, it is not yet believed to be a complete theory. The SM does not provide theoretical support for either dark energy or dark matter particles which are believed to exist according to cosmological observations. Moreover, within the SM particles, neutrino oscillation proved by several experiments conflicts with the SM's assumption of massless neutrinos. Above all, as mentioned above, the SM incorporates only 3 of the 4 fundamental interactions, leaving gravitation completely unexplained in the scope of particle physics.

The Hierarchy Problem [12] in particle physics refers to the huge discrepancy between the electroweak scale and the gravitational scale, as the weak force is about  $10^{24}$  times stronger than gravity. The Fermi's constant denoting the scale of the weak interaction is expected to be smaller and closer to the Newton's constant for gravity, based on the calculation of SM. In other words, we expect the large quantum contributions to the square of the Higgs boson mass would make the Higgs boson much heavier than the measured 125 GeV. It could be that the measured Higgs boson mass is the result of incredibly fine-tuned constants within the SM. Alternatively, some new theoretical mechanism is expected. The Bulk RS Graviton Model is one of the possible solutions.

# **1.3 Bulk RS Graviton Model**

The Bulk RS Graviton model offers an efficient solution to the Hierarchy problem. It also provides theoretical support for the production of heavy di-boson resonances, which can result from interactions involving an extra spatial dimension. The development and features of this model are summarized below.

### **1.3.1** Introduction of Extra Dimension Models

In the 1920s, the Kaluza–Klein theory was proposed as a means to unify gravitation and electromagnetism. It assumed a 5th dimension beyond our usual four dimension of space and time and started purely from the 5 dimensional extension of General Relativity. The quantum mechanical interpretation predicts that gravitons or other particles in the extra dimension will acquire quantized excited modes, which are referred to as Kaluza-Klein (K-K) modes. The K-K theory is considered an important precursor to subsequent theories that introduce extra spatial dimensions as a solution to the Hierarchy Problem.

Following the K-K theory, a number of extra dimension theories were proposed attempting to address the Hierarchy problem, such as the ADD model. The ADD model [13, 14] was proposed in 1998 by Nima Arkani-Hamed, Savas Dimopoulos, and Gia Dvali, suggesting the existence of additional large extra spatial dimensions. It proposed that the Planck scale  $M_{Pl}$  is not a fundamental scale but is instead simply a consequence of the large size of the new dimensions. While gravitons can freely propagate in the new dimensions, at sub-weak energies the SM fields must be localized to a 4-dimensional manifold of weak scale thickness in the extra dimensions.

#### **1.3.2 Randall–Sundrum Model**

The Randall-Sundrum (RS) model [15] was originally proposed in 1999 by Lisa Randall and Raman Sundrum, because they were not satisfied with those large extra dimension models that involved fine tunings of the bulk cosmological constant and brane tensions. This theory is often referred to as the RS1 model. The framework is based on a slice of 5-dimensional antide Sitter space ( $AdS_5$ ) [16], with two flat four-dimension branes on each boundary. According to the  $AdS_5$  theory, if the flat 4D branes carry energy, the geometry of the additional dimension has to be warped to be consistent with the General Relativity. The line segment (ds) which determines the distance scale is given by equation 1.1,

$$ds^{2} = e^{-2ky} \eta_{\mu\nu} dx^{\mu} dx^{\nu} + dy^{2}, \qquad (1.1)$$

where  $\eta_{\mu\nu}$  is the Minkowski metric for flat four-dimension space;  $dx^{\mu}$  is the regular time-space four-vector; y represents the 5th dimension coordinate, bounded by  $0 \le y \le \pi R$  and R is the extra dimension length; k is the curvature constant.

Fluctuations on 2 of the variables are possible, k and R, each corresponding to a particle field: the graviton and radion. In the RS1 model, the SM particles are all trapped on one of the branes, referred to as the TeV Brane, located at  $y = \pi R$ . The other brane in this model, named the Plank Brane and located at y = 0, has a Planckian fundamental scale and is the brane where the 4D (or zero-mode) graviton concentrates. Given this assumption, it is possible to unify the fundamental gravitational and weak mass scales, since the 4D physical masses on the TeV brane acquire an exponential rescaling of  $e^{-2ky}$ . Figure 1.2 gives the idea of this extra dimension model and the metric behavior.

The key feature of the RS1 model is that, though the zero-mode spin-2 graviton is localized in the Plank Brane, its K-K excited mode (K-K graviton) is localized near the TeV brane and has mass around a TeV, so that K-K graviton coupling to the entire SM is only  $\sim$  TeV suppressed.



Figure 1.2: In RS theory, the Gravity Brane (Planck Brane) and the Weak Brane (TeV Brane) are the 4 dimensional boundaries of the extra dimension. The metric behavior along the extra dimension is also shown.

## **1.3.3** The Bulk RS Graviton Model

However, in the RS1 model, the higher-dimensional operators in the 5D effective field theory are suppressed only by the warped scale  $\sim$  TeV, giving contributions that are too large to flavor-changing neutral current (FCNC) processes [17, 18] which are strongly suppressed in the SM.

A solution to this issue is to allow the SM fields to propagate in the extra dimension, which leads to the scenario of the Bulk RS Graviton model [19, 20, 21]. In this model, the SM particles are identified with the zero-modes of the 5D fields and the profile of a SM fermion in the extra dimension depends on its 5D mass parameter. We can then choose to localize 1st and 2nd generation fermions near the Planck brane, so that their interactions to the K-K gauge bosons are suppressed, as the K-K gauge bosons are near the TeV brane and have little overlapping with the light fermions. Therefore, the FCNCs from higher-dimensional operators are suppressed by scales far beyond TeV scale.

Like the RS1 model, the K-K gravitons are localized near the TeV brane. Figure 1.3 shows the zero-modes of the SM matter fields and the K-K graviton along the 5th dimension, for the Bulk RS model and RS1 model.



Figure 1.3: Zero-modes of the SM matter fields and the K-K graviton along the extra dimension for the Bulk RS model scenario (left), and RS1 (right).

From Figure 1.3 one can see that in the Bulk RS scenario the light fermions are localized near the Plank brane while the K-K gravitons are near the TeV brane, as a result, the couplings of the K-K gravitons to light fermions are highly suppressed compared to the RS1 model. Additionally, the SM massless gauge bosons have flat distribution across the extra dimension, which also leads to the suppression of their couplings to the K-K gravitons by roughly a factor of  $k\pi R$ .

### **1.3.4** Phenomenology of the Bulk Graviton Model

As in the RS1 model, there are 2 free parameters in the RS Bulk theory, the curvature constant k and the extra dimension length R. Equivalently they can be expressed as  $\tilde{k}$ , which is defined as the ratio of k and the reduced Plank mass ( $\overline{M_{Pl}} \equiv M_{Pl}/\sqrt{8\pi}$ ), and the K-K graviton mass ( $m_G$ ). The K-K Graviton in the Bulk Graviton model can be produced in several ways, while the dominant process is QCD gluon fusion. Figure 1.4 shows the cross sections for the production of a Bulk Graviton production for different processes [22], with  $\tilde{k} = 0.1$  and proton-proton center of mass energy of 13TeV. Here the  $pp \rightarrow Gjj$  mode is dominated by the Vector Boson Fusion (VBF) process, and the  $pp \rightarrow GV$  denotes the associated production with a massive vector boson. For  $\tilde{k} < 1$  and  $m_G < 2$  TeV the width of the K-K graviton is less than 6% of the graviton mass, therefore a narrow resonance is expected. In this case the production cross section scales with



Figure 1.4: Production cross section for the K-K Graviton in the Bulk RS model,  $\tilde{k} = 0.1$ 

 $\tilde{k}$  as described in Equation 1.2.

$$\sigma(m_G, \tilde{k}) = (\tilde{k}/0.1)^2 \sigma(m_G, \tilde{k} = 0.1)$$
(1.2)

In terms of the decay modes, the K-K gravitons are most likely to decay to SM particles localized near the TeV brane, as they have the strongest coupling to the graviton [22]. Therefore the dominant decay modes are to top quarks and Higgs, as well as Z and W bosons, while the decay modes into massless gauge bosons and light quarks are suppressed and negligible. Figure 1.5 shows the branching ratios for the decay of a Bulk Graviton. The decay branching fractions are independent of  $\tilde{k}$ . Although the top



Figure 1.5: Branch fraction of Bulk Graviton decay modes.

channel and Higgs channel have the highest branch fractions, neither of

these channels can be easily detected or reconstructed. In contrast, despite of the relatively lower branching ratio, the W and Z boson channels are more preferred for the experimental search.

# **1.4 Status of Searches for the Bulk Graviton**

As described above, the Bulk Graviton can be produced by gluon fusion as well as other processes occurring in proton-proton collisions produced at the Large Hadron Collider. Because of the large gluon luminosity at the LHC, these data are ideally suited to search of evidence of a Bulk Graviton.

## 1.4.1 Previous Searches

Previous searches [23, 24, 25, 26, 27, 28] for the Bulk Graviton have been performed based on data collected from both ATLAS and CMS detectors with the proton-proton center of mass energy at 7, 8 and 13 GeV. Limits on the cross section for the production of a Bulk Graviton have been set as a function of  $m_G$ . The existence of the Bulk Graviton was excluded with mass below 610 GeV [26] for  $\tilde{k} = 0.5$  and mass below 1100 GeV [28] for  $\tilde{k} = 1.0$  at 95% confidence level. In these searches, semi-leptonic final states including leptons and jets, comprise the most popular channels. There is no previous search studying the channel ZZ to  $2\ell 2\nu$ .

## 1.4.2 $2\ell 2\nu$ channel and Search Strategy

In this analysis we present a search for the Bulk Graviton or similar resonances decaying into a pair of Z bosons, in which one of the Z boson decays into a pair of charged leptons, either an electron pair or a muon pair (denoted by " $\ell$ "), while the other decays into two neutrinos (denoted by " $\nu$ "). Figure 1.6 shows the Feynman diagram of this process. This analysis is based on the data from proton-proton collisions at a center-of-mass energy of 13 TeV collected by the CMS detector in 2016 and corresponding to an integrated luminosity of 35.9 fb<sup>-1</sup>. The signature of the  $2\ell 2\nu$  channel is a pair of adjacent leptons with high transverse momentum ( $p_T$ ) from the



Figure 1.6: Leading order Feynman diagram for the production of resonance X via gluon–gluon fusion decaying to the  $ZZ \rightarrow 2\ell 2\nu$  final state

decay of a boosted Z boson and large missing transverse momentum from the other Z boson decaying to neutrinos, which is denoted as  $p_T^{\text{miss}}$ .

Compared to the semi-leptonic channels, the  $2\ell 2\nu$  channel has reduced background. Although Z+jets production is a main source of background for both of these channels, it is almost impossible to distinguish the ZZ to  $2\ell 2q$  process from Z+jets background, as they have very similar kinematics – a Z boson and corresponding hadronic recoil. Unlike the Z+jets background, events from the ZZ to  $2\ell 2\nu$  process always includes large  $p_T^{\text{miss}}$ , which makes it possible to strongly suppress the Z+jets background. Although the  $2\ell 2q$  channel has a larger branching fraction, the branching fraction of the  $2\ell 2\nu$  is still considerable, about 1/3 of that of the  $2\ell 2q$  channel and 6 times as large as that for the four charged-lepton final state. The combination of lower statistics and the effects on resolution of a signal due to the undetected neutrinos explains why no previous search in  $2\ell 2\nu$ channel was carried out. But now in RunII data with larger cross section from the higher collision energy and the larger and growing statistics, the  $2\ell 2\nu$  channel has become more advantageous.

Because of the invisible neutrinos from the second Z boson decay, we will not have full momentum information for that Z boson, instead, only  $p_T^{\text{miss}}$  is available, which can be observed as the projection of the invisible Z boson's momentum on the x-y detector plane. Therefore it is not possible

to reconstruct the invariant mass of the  $2\ell 2\nu$  system. In this analysis the transverse mass  $(m_T)$  is designated as the discriminating variable to separate signal from background. The transverse mass variable is defined as:

$$m_T^2 = \left[\sqrt{(p_T^{\ell\ell})^2 + m_{\ell\ell}^2} + \sqrt{(p_T^{miss})^2 + m_{\ell\ell}^2}\right]^2 - \left[\vec{p}_T^{\ell\ell} + \vec{p}_T^{miss}\right]^2, \quad (1.3)$$

Here  $p_T^{\ell\ell}$  and  $m_{\ell\ell}$  each represent the  $p_T$  and mass of the Z boson constructed from the charged lepton pair system. The  $m_{\ell\ell}$  in the middle term provides an estimator of the mass of the invisibly decaying Z boson. This choice has negligible impact on the expected signal at large  $m_T$ , but is found to preferentially suppress backgrounds from  $t\bar{t}$  and WW decays.

Although the  $m_T$  variable only equals the invariant mass confined to the transverse plane, a kinematic edge is still expected from the putative heavy resonance, and the position of the edge strongly depends on the mass of the resonance. Figure 1.7 gives an example of what the transverse mass spectrums look like from the decay of resonances with invariant masses each at 800 GeV, 1000 GeV, 1200 GeV and 1400 GeV. The kinematic edges on the right side of each spectrum are very close to the graviton mass values.



Figure 1.7: The transverse  $mass(m_T)$  spectrums of resonances with masses at 800 GeV, 1000 GeV, 1200 GeV and 1400 GeV.

The transverse mass can also be written in an alternate form in which no vector is involved:

$$m_T^2 = \left[\sqrt{(p_T^{\ell\ell})^2 + m_{\ell\ell}^2} + \sqrt{(p_T^{miss}_{\perp})^2 + (p_T^{miss}_{\parallel})^2 + m_{\ell\ell}^2}\right]^2 - \left(p_T^{\ell\ell} + p_T^{miss}_{\parallel}\right)^2 - (p_T^{miss}_{\perp})^2,$$
(1.4)

where  $p_T_{\parallel}^{miss}$  refers to the projection of  $p_T^{miss}$  in the direction of  $p_T^{\ell\ell}$  and  $p_T_{\perp}^{miss}$  is the fraction perpendicular to  $p_T^{\ell\ell}$ .

# **Chapter 2**

# The LHC and CMS detector

This analysis is based on the 2016 data collected by the Compact Muon Solenoid (CMS) experiment. The CMS detector was constructed at the Large Hadron Collider (LHC) located in Geneva, Switzerland. This chapter gives an overview of the LHC and the CMS experiment.

## 2.1 The Large Hadron Collider

The Large Hadron Collider is the world's most powerful particle accelerator and collider [29, 30]. It was built by the European Organization for Nuclear Research (CERN) between 1998 and 2008, in the tunnel of its predecessor, the Large Electron Positron Collider (LEP), with a circumference of 27 km and as deep as 175 meters (574 ft) beneath the France–Switzerland border near Geneva.

The LHC is designed mainly for proton-proton collisions, but can also be used to produce ion collisions. In this thesis only proton collisions are discussed. The first long proton run was performed in 2010, starting with the center of mass energy of 7 TeV, or 3.5 TeV per beam. After a short break at the end of 2011 the beam energy was increased from 3.5 TeV to 4 TeV (8 TeV center of mass energy). The period of these 7 TeV/8 TeV runs are often referred to as LHC Run I. Run I ended early 2013 as the LHC began a 2-year shutdown for the LHC upgrades, to enable operations at higher collision energy. The Run II data taking started on 3 June 2015, with a center of mass energy of 13 TeV. The process of particle acceleration begins from a simple tank of hydrogen as the source of protons, which are progressively accelerated to higher energies in sequential machines ending at the LHC. A diagram of the CERN accelerator complex is shown in Figure 2.1. The hydrogen atoms are first accelerated by a linear accelerator, Linac2, after being stripped of their electrons. In Linac2 the proton beam is accelerated by electric fields generated by alternately charged cylindrical conductors, and concentrated by magnetic fields generated by superconducting quadrupole magnets. The protons exit Linac2 with an energy of 50 MeV and enter the Proton Synchrotron Booster (PSB), where the particles are further accelerated to 1.4 GeV. The protons are then transferred to the Proton Synchrotron (PS) with a circumference of 628m where the energy of the protons is increased to 25 GeV. Following is the 7km long Super Proton Synchrotron (SPS) as the final step of the acceleration chain before the beams are injected into the LHC ring. There the protons are accelerated to an energy of 450 GeV.





The beams are injected into the LHC rings for the final acceleration before collisions. There the protons reach the energy of 6.5 TeV after about 20 minutes of acceleration. Nearly 10,000 magnets are used on the LHC. Among them over 1000 superconducting dipole magnets are used to produce a magnetic field of 8.3T and bend the proton beams onto the circular trajectory. Additional, superconducting quadrupole magnets are used to focus the beam, and sextupoles and higher order magnets are used to correct the beam chromaticity.

Inside the LHC particle collisions can happen at 4 interaction points of the tunnel, which correspond to the locations of 4 experiments: CMS, ALICE, ATLAS and LHCb. The ALICE experiment is designed to study quark-gluon plasma using data collected during heavy ion operations of the LHC. These measurements are designed to draw conclusions about the initial state of the universe. LHCb focuses on precisely measuring B-meson decays and CP-violating processes. CMS and ATLAS are the two general purpose experiments at the LHC built for studying a broad range of physics processes. These studies include precision measurements of Standard Model processes and parameters, thereby deepening our knowledge and understanding of the Standard Model. Study of the production and properties of the Higgs boson and searches for physics beyond the Standard Model are major fields of study.

The number of collisions generated at the LHC is proportional to the cross section for proton-proton interactions and the integrated luminosity, and can be written as  $N = \sigma \times L$ . Here L is the integrated luminosity and is defined as the integral of instantaneous luminosity (denoted by  $\mathcal{L}$ ) over time, as shown in Equation 2.1.

$$L = \int \mathcal{L}dt \tag{2.1}$$

The maximal instantaneous luminosity of LHC in 2016 reached 15.30 Hz/nb, and the total integrated luminosity delivered by LHC in 2016 was 40.82 fb<sup>-1</sup>. A total of 37.76 fb<sup>-1</sup> was recorded by CMS and 35.9 fb<sup>-1</sup> satisfied the data quality requirements used in this analysis. Figure 2.2 shows the development of instantaneous and integrated luminosity during year of 2016.



Figure 2.2: Daily peak instantaneous luminosity (left) and integrated luminosity (right) of 2016 LHC proton-proton collisions.

# 2.2 The Compact Muon Solenoid (CMS) Experiment

The Compact Muon Solenoid (CMS) detector [31] is one of the two general purpose detectors at the LHC. It sits 100 meters underground at the LHC interaction point opposite the CERN site, in the French village of Cessy, and measures 15 meters in height, 22 meters in length. And the weight of 14,000 tons makes the CMS detector the heaviest particle physics detector in the world. The CMS detector is designed as a barrel-like detector around the interaction point of the proton beams delivered by the LHC, consisting of multiple layers of subdetectors. The detector can be separated into parts: the central part, ofter referred to as "barrel", and the two sections closing out the barrel on the ends of the detector called "endcaps".

Figure 2.3 shows the structure and components of the CMS detector.

One characteristic of the CMS detector is its strong axial magnetic field at a uniform strength of 3.8 T, generated by a superconducting solenoid magnet surrounding the inner detector subsystems. Figure 2.4 shows the magnetic field distribution measured using cosmic ray tracks [32]. The magnetic field is mostly confined within the steel yoke around the solenoid. Most particles from the collisions except muons would either deposit all their energy in the subdetectors before the yoke or be stopped by the yoke.



Figure 2.3: Illustration of CMS detector structure and subdetectors as components in a cut-out quadrant view (upper), and a cross sectional slice view as well as the trajectories of various particles from p-p collisions in the detector (lower).



Figure 2.4: CMS magnetic field strength (left) and field lines (right)

Two sets of coordinate systems are widely used in CMS. One is an orthogonal right handed coordinate system, with the origin sitting at the nominal collision point inside the experiment. The x-axis points radially toward the center of the LHC ring, the y-axis points upward perpendicular to the LHC ring plane, and the z-axis correspondingly points counterclockwise along the beam pipe. Apart from the orthogonal coordinate system, a spherical coordinate system is used due to the cylindrically symmetric design of the detector. Differing from the regular spherical coordinate with parameter  $\phi$  denoting the azimuthal angle to the x axis in the x - y plane and parameter  $\theta$  denoting the polar angle measured from the z axis, this coordinate in CMS keeps the  $\phi$  parameter but uses the pseudorapidity  $\eta$  defined in Equation 2.2 in the place of  $\theta$ .

$$\eta = -\ln[\tan(\frac{\theta}{2})] \tag{2.2}$$

Based on this spherical coordinate, parameter  $\Delta R$  is introduced as a measurement of the distance between two particles, and defined as  $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}$ .

## 2.2.1 Inner Tracking System

The inner tracking system [33] is the innermost component of the CMS detector. The main functionality of the inner tracking system is to precisely measure the trajectory of charged particles such as charged leptons and hadrons. It consists of an inner pixel detector and a strip tracker. Figure 2.5 shows the structure of the tracking system

### **Inner pixel detector**

The inner pixel detector, consisting of 3 cylindrical layers with minimal radius of 4 cm and disks on each end as endcaps, is the closest detector to the interaction point of the experiment, which makes it vital in reconstructing the tracks of very short-lived particles. But being close to the interactions also means an enormous particle flux. Therefore the pixel detector's design was driven by the goal of getting the best track position resolution possible while also being very radiation tolerant. The inner pixel detector contains 65 million silicon pixel sensors each with a dimension of  $100\mu m \times 150\mu m$ . These pixel sensors were built using high dose n-implants on a high resistance n-substrate, which ensures high signal collection efficiency with only moderate bias voltages even after high doses of radiation.

### Strip tracker

After passing through the pixel detector, particles reach the outer silicon strip tracker at radius of 130 cm. The strip detector consists of the Tracker Inner Barrel (TIB), Tracker Inner Disks (TID), Tracker Outer Barrel (TOB), and the Tracker End Caps (TEC). The TIB contains 4 layers of silicon sensors, among which the inner 2 layers are built with double sided sensors while the other 2 layers are single sided. The TID, placed on each end of TIB as the endcaps, is composed of 2 sets of disks of sensors, with 3 disks in each set. Surrounding the inner tracker is the TOB, consisting of 6 layers. The inner 2 layers carry double sided sensors. Finally the endcaps (TEC) close off the tracking system, with 9 disks of sensors on each side.





Figure 2.5: Structure of CMS tracking system in the plane parallel (upper) and perpendicular (lower) to the z axis.

The CMS tracking system covers the detector region up to  $|\eta| = 2.5$ and has a resolution of up to 10  $\mu m$  in the x - y direction and 20  $\mu m$  in the z direction.

#### 2.2.2 Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) [34] is designed mainly for the measurement of the energy of photons and electrons. It sits between the tracker and the Hadron Calorimeter, covering the  $\eta$  range from -3.0 to 3.0. The ECAL is composed of the ECAL Barrel (EB), ECAL Endcaps (EE) and ECAL Preshower (ES). Figure 2.6 shows the structure of ECAL.



Figure 2.6: Structure of the Electromagnetic Calorimeter (ECAL) in the y - z plane.

#### **ECAL Barrel and Endcap**

ECAL Barrel and Endcap contains near 80,000 lead tungstate (PbWO4) crystals each with face dimension of approximately  $3cm \times 3cm$ . These crystals are primarily composed of metal with high density, and have good radiation tolerance and short radiation length. This material produces scintillation light with fast photon showers. The lead tungstate crystals

provide precise energy resolution and allow the calorimeter to be compact enough for the CMS design. However, a drawback of the crystal is that its light yield strongly depends on temperature. The nominal operating temperature of the ECAL is maintained at  $18^{\circ}C$ , with variation within  $0.1^{\circ}C$ , for the desired energy measurement resolution.

The EB consists of over 60,000 crystals and covers  $|\eta|$  up to 1.479. Avalanche photodiodes (APD) are mounted on the rear face of these crystals to collect the scintillation light and amplify the signal. Each of the two endcaps contains over 7,000 crystals. The crystals are grouped in  $5 \times 5$ arrays referred to as "super crystals". Vacuum phototriodes (VPT) are glued to these crystals for signal collection and amplification, similar to the APDs in the EB. The EE covers the  $|\eta|$  range between 1.479 and 3.0.

#### **ECAL Preshower**

The ECAL Preshower (ES) detectors sit in front of the EEs. They are composed of two planes of lead followed by silicon sensor strips. When photons pass through, electromagnetic showers are produced in the lead layer, and then detected by the silicon sensors. The silicon sensor strips on the ES with width of 2mm give much better position resolution, compared to the  $3cm \times 3cm$  ECAL crystals, and can distinguish individual photons in the adjacent photon pairs from  $\pi^0$  decays.

#### 2.2.3 Hadronic Calorimeter

The Hadron Calorimeter (HCAL) [35] is designed for the measurement of energy of the hadrons produced from the p-p collisions, and plays a crucial role in the indirect measurement of particles having no interaction with the detector, such as neutrinos. The HCAL can be divided into 4 components, which are the HCAL Barrel (HB), HCAL Endcaps (HE), HCAL Forward calorimeter (HF) and HCAL Outer calorimeter (HO). Figure 2.7 shows the structure of the HCAL detector.



Figure 2.7: Structure of the Hadron Calorimeter (HCAL) in the y - z plane. Muon Chambers are also shown to illustrate the position of HF.

#### **HCAL Barrel and Endcap**

The HB and HE are located inside the solenoid magnet, surrounding the ECAL. Unlike ECAL which is built completely by crystals, HB and HE are sampling calorimeters made of alternating layers of metal as absorbers and tiles of Kuraray SCSN81 plastic scintillator. When a particle passes through the absorbers, numerous secondary particles are produced. The particle shower produced in the absorbers causes the scintillator layers to emit violet light. This violet light is then shifted into green light in wavelength-shifting fibers and carried to the readout boxes by clear fibers, where the light is converted into fast electronic signals by Hybrid Photodiodes (HPDs) photosensors.

The absorber material for HB is steel for the innermost and outermost layers and brass for the layers in between. The HB consists of 2 halves, referred to as HB Plus (HBP) and HB Minus (HBM). They each contains 18  $\phi$  sections called wedges. Each wedge weighs 26 tonnes and is further subdivided in 4  $\phi$  sectors and 16  $\eta$  sectors. The HB covers the  $|\eta|$  up to 1.3.

Like the HB, the endcaps are made of alternating brass layers and

Kuraray SCSN81 plastic scintillator layers. Each HE is divided into 36  $\phi$  sectors and 13  $\eta$  towers. The HE covers  $|\eta|$  between 1.3 and 3.0.

### **HCAL Outer Calorimeter**

The HO sits outside the magnet coil and covers the region of  $|\eta| < 1.3$ . It consists of a single 10 mm Bicron BC408 scintillator layer at a radial distance of 4.07 m. Like the HB and HE, the scintillation light is collected by wavelength-shifting fibers and sent to HPDs by clear fibers. The HO expands the radial sampling depth of the HCAL system and ensures no energy leakage undetected by HB.

### **HCAL Forward Calorimeter**

The HFs are placed outside the muon system, 11.2m away from the interaction point of the experiment along the z-axis on each side. The HF receives a great amount of radiation from the collision due to its position and thus has to be more resistant to radiation than the rest of HCAL components. Each HF is made of steel, shaped as a cylinder. The steel in this case works as absorbers while radiation-hard quartz fibers are placed along grooves inside, 5.0 mm apart, and working as Cherenkov radiators. The cylinder has height of 165 cm, outer radius of 130.0 cm, and inner radius of 12.5 cm for the beams to pass through. The HF extends the  $|\eta|$  coverage of HCAL to 5.0.

## 2.2.4 Muon System

Unlike other particles, muons can penetrate the calorimeters and yoke without being stopped. The CMS muon system aims at the identification and measurement of muons [36]. The muon system is the outermost layer of the CMS detector system, and is the largest subdetector in CMS. It includes 1400 muon chambers, which can be categorized into 3 groups: Drift Tubes (DT), Cathode Strip Chambers (CSC) and Resistive Plate Chambers (RPC). The muon system covers the region  $|\eta| < 2.4$ . Figure 2.8 shows the structure of the muon system.



Figure 2.8: Structure of CMS muon system in the y - z plane.

#### **Drift Tubes**

The Drift Tubes are used in the barrel of the muon system, and cover the region of  $|\eta| < 1.2$  with four layers called stations. The inner 3 stations each contain three super layers (SL) with four chambers of DTs filled with a gas mixture of 85% Ar and 15%  $CO_2$ . The outer two SLs provide precise  $\phi$  measurements while the middle one measures  $\eta$ . The outermost station has only two SLs each measuring either  $\phi$  or  $\eta$ .

#### **Cathode Strip Chambers**

The Cathode Strip Chambers make up the endcap region of the muon system covering  $|\eta|$  between 0.9 and 2.4. A CSC covers  $10-20^{\circ}$  in  $\phi$ . It consists of 7 cathode strip panels alternating with 6 layers of anode wire planes, filled with a gas mixture of 40% Ar,  $50\% CO_2$  and  $10\% CF_4$ . The cathode strips are arranged in the radial direction to give the measurement of  $\phi$ , while the anode wires are wound in the  $\phi$  direction for radial position measurements.
#### **Resistive Plate Chambers**

In addition to DTs and CSCs, RPCs are embedded in both the barrel and endcap for  $|\eta|$  up to 2.1. Despite of their coarse spatial resolution, RPCs provide very precise timing measurement of muons and are used to supplement the muon triggering system. An RPC is composed of an anode and a cathode as two parallel plates, filled with a gas mixture of 95.2% Freon, 4.5% isobutane, as well as 0.3% hexafluoride and water vapor.

## 2.2.5 Data Acquisition and Trigger System

Proton-proton collisions take place every 25 ns with an average of 20 protonproton interactions per bunch crossing, leading to an input data rate of nearly 1 GHz seen by the CMS detector. However, it is impossible to record and store all the information as the computing system can only handle events up to 400 Hz. Therefore a trigger system is needed for fast processing to select the most interesting events. In CMS a two-stage trigger system is used [37]: the *Level* 1 trigger (L1T) and *High Level trigger* (HLT).

### Level 1 Trigger

The *Level* 1 *trigger* is the first stage of CMS trigger system. It makes decisions completely based on the coarse information from the calorimeters and muon detectors, allowing for fast response. Local trigger information and decisions are first calculated by each subsystem and then sent to the Global Trigger (GT) where a final decision is made for L1T based on the subsystem information.

The L1T reduces the event rate to 100kHz before sending data on to the *High Level trigger*.

#### **High Level Trigger**

The *High Level trigger* is entirely software based and relies much more on the reconstruction of physics objects than L1T. This is similar to the offline object reconstructions, but is implemented in a simpler and faster way. This allows the HLT to define events based on physics objects and to select those that are most interesting for data analysis. A dedicated computing farm is used for the HLT processing.

The HLT brings the event rate down to the order of 100 Hz. Events selected by the HLT are kept and sent to the data storage system to await offline processing.

# **Chapter 3**

# **Physics Objects Reconstruction**

The information received from the detector consists of digitized information from sensors measuring energy deposition in active regions of the detector. This is not directly suitable for physics analyses. In this chapter we introduce the reconstruction and identification of physics objects used in this analysis, based on the detector-level information.

In the CMS Experiment the reconstruction of the event to interpret the digital information in terms of particle production is accomplished using the Particle-Flow (PF) algorithm [38].

# **3.1** Particle Flow Algorithm

The PF algorithm starts from reconstructing basic elements based on the information collected from each subsystem of the detector, including the trajectories of charged particles, calorimeter energy clusters and electron/muon tracks. Then a link algorithm is used to form a block of these PF elements possibly related to a single physics object. Then particle reconstruction and identification is applied based on these blocks. A description of the reconstructed objects returned by the PF algorithm is given below.

#### **3.1.1** Charged-particle tracks and vertices

The construction of charged-particle tracks provides a basis for particle reconstruction and identification, including the measurement of the momentum of energetic and isolated muons, identification of energetic and isolated hadronic  $\tau$  decays, and tagging b quark jets. A combinatorial track finder based on the Kalman Filtering (KF) is used for track construction [39, 40]. It starts with a few hits compatible with a charged-particle trajectory in the tracking system; then finds hits from all the tracker layers along this charged-particle trajectory; and finally performs a fit to determine the trajectory and the particle properties including origin, transverse momentum and direction. To increase the reconstruction efficiency and suppress the fake rate, the combinatorial track finder is applied in several successive iterations.

The vertex reconstruction [41] determines the locations of all the protonproton interactions associated with each beam crossing. The vertex reconstruction uses information from the reconstructed charged-particle tracks. Candidate vertices are identified using a deterministic annealing (DA) algorithm [42] and fitting them with at least two matched tracks using an adaptive vertex fitter. Then vertices are selected by applying quality criteria including being consistent with the collision region (referred to as beam spots) and being matched to a minimum of four tracks. The primary vertex of an event is considered to be the vertex with the largest sum of the squared track momenta. The other vertices are regarded as pile-up vertices from additional unrelated proton-proton interactions. Figure 3.1 shows the vertex multiplicity distribution for the data collected in 2016.

### 3.1.2 Calorimeter clusters

The calorimeter clustering serves four purposes: to detect and measure the energy and direction of stable neutral particles; to separate these neutral particles from charged hadron energy deposits; to reconstruct and identify electrons and all accompanying bremsstrahlung photons; and to improve the energy measurement of charged hadrons for which the track parameters were not determined accurately. The clustering process is performed separately for EB, EE, ES, HB and HE. Cluster seeds are first selected from calorimeter cells with deposited energy larger than a threshold and also any neighbouring cells. The topological clusters then spread from the seed to



Figure 3.1: Number of interactions per bunch crossing for data collected in 2016.

the nearby cells with energy above defined thresholds to suppress noise. Finally a fit based on the expectation-maximization algorithm is applied for each topological cluster to evaluate the position and amplitude of the resulting particle clusters [38].

#### **3.1.3** Tracks for electrons

Because of the significant thickness of the tracker, most electrons emit a considerable fraction of their energy in the form of bremsstrahlung photons before reaching the ECAL. To reconstruct the properties of an electron, the energy of the bremsstrahlung photons in ECAL must be taken into account in addition to the energy deposited directly by the electron. A tracker-based electron seeding method was developed. When the radiated energy is small, the electron track can be accurately reconstructed across the whole tracking system with a  $\chi^2$  fit, and the reconstructed momentum should match the energy deposited in the corresponding ECAL cluster. However when energetic photons are radiated, the momentum change of the electron will lead to a large  $\chi^2$  and missing hits in the tracker. In these cases, a modified selection of tracker hits is performed based on the  $\chi^2$ 

value and number of hits in the previous KF fit, and these selected hits are fit again with a Gaussian-sum filter (GSF) [43] which is more adapted to the electrons as energy losses along the trajectory are considered.

#### **3.1.4** Tracks for muons

The muons leave hits in the tracking system where their momenta and trajectories can be precisely measured. They are identified over the full detector using additional information from the muon chambers. Hits within the DT and CSC detectors are clustered to form track segments. These track segments are then used as seeds to reconstruct the muon trajectory using DT, CSC, and RPC hits. The result of the final fitting is referred to as a standalone-muon track. With the information of the reconstructed tracks in the tracking system, two collections of high-level muons physics objects can be obtained: the global muon and the tracker muon (See Section 3.4).

#### 3.1.5 Link algorithm

To reconstruct a physics object, a link algorithm is applied to connect related PF elements from different subdetectors. The PF elements that can be considered by the algorithm can only be the nearest neighbours in the  $(\phi, \eta)$  plane to any of the elements in the linked block starting from a seed element. The element will be added to the block after the distance of the link is examined. This process continues until the PF block is formed and the reconstruction of the corresponding particle starts. Links between GSF electron tracks and the ECAL clusters can be established for electrons; links among calorimeter clusters can be established for various particle reconstruction processes, specifically trivial links among ECAL clusters with close  $\eta$  but spread  $\phi$  form link blocks called superclusters; links between tracks and muon segments can be established for muon reconstruction.

# 3.2 Electron Reconstruction

Electron reconstruction and identification are performed mainly based on the information from electron tracks in the tracking system and the ECAL clusters. [44]

The electron reconstruction starts from a GSF electron track. In the link algorithm, apart from the electron cluster in ECAL, clusters nearby the extrapolated tangents of the GSF track will also be linked to the PF block as potential bremsstrahlung photons emitted from the electron. A GSF track will be considered a electron candidate if in the PF block the ECAL cluster corresponding to the electron shower is not linked to more than 2 tracks.

The ECAL clusters in the PF block that can be linked to either the GSF track tangent extrapolations or the supercluster will be associated with the candidate electron and will be used for the energy calculation of the electron. The GSF track is also attached to the candidate electron. The position information  $(\eta, \phi)$  assigned to the electron is obtained from the GSF tracks. Once an element is assigned to a reconstructed particle, it will be masked against further processing in other object constructions. The calculated energy from the calorimeter clusters will be corrected in terms of energy and  $\eta$ , to compensate for energy loss in the process. The final energy assigned to the electron will be obtained from a combination of the corrected energy from the calorimeter clusters and the momentum of the GSF track.

Furthermore, electrons used in physics analyses must meet additional identification and isolation requirements.

### 3.2.1 Electron Identification and Isolation

The electron candidates used in this analysis are required to pass the loose cut-based identification (ID) and isolation (Iso) recommended by the CMS EGamma Physics Object Group (POG) for 2016 data.

The ID and Iso criteria set cuts on the following variables, and the cut

values are given in Table 3.1, these include:

- the  $\eta_{\rm SC}$  denoting the  $\eta$  value of the corresponding ECAL super cluster;
- the  $\sigma_{i\eta,i\eta}$  describing the shape of the supercluster;
- the geometric distance,  $|\Delta \eta_{in}|$  and  $|\Delta \phi_{in}|$ , between the supercluster and the matched track;
- the ratio of the energy deposits in HCAL and ECAL, hOverE;
- the relative combined PF isolation following correction for pile-up contamination in the Effective Area (EA), relisoWithEA;
- the difference between the tracker momentum and ECAL energy, |1/E 1/p|;
- the maximal expected missing hits in the inner track;
- a veto on electrons that are likely to be produced by photon conversions.

Table 3.1: The cuts used in the POG 1	oose electron identification.
---------------------------------------	-------------------------------

Variable	Barrel	Endcap
$ \eta_{\rm SC} $ acceptance	(0, 1.479)	(1.479, 2.5)
$\sigma_{i\eta,i\eta} <$	0.011	0.0314
$ \Delta\eta_{in}  <$	0.00477	0.00868
$\Delta \phi_{in} <$	0.222	0.213
hOverE <	0.298	0.101
rellsoWithEA <	0.0994	0.107
1/E - 1/p  <	0.241	0.14
expectedMissingInnerHits $\leq$	1	1
conversion veto	yes	yes

#### **More About PF Isolation**

The PF isolation requirement is introduced in the electron ID, in order to suppress the electron-like signals caused by jets. The calculation of the PF Iso is based on the PF algorithm: for a given electron or photon, the sum of transverse momenta of all the PF elements with the type of charged hadron, neutral hadron or photon will be calculated if the PF element falls in the isolation cone around the electron/photon. The cone is usually defined as the region of  $\Delta R < 0.3$ . The separately calculated isolations for the charged hadrons, neutral hadrons and photons can be noted as  $Iso_{ch}$ ,  $Iso_{nh}$ ,  $Iso_{photon}$ .

Corrections are applied to the calculated PF Iso to compensate the effect caused by contamination from pile-up. In the case of electron Iso,  $\rho$ -effective area corrections are applied. The effect from the pile-up is considered to be  $PU = \rho \times effective\_area$ , where  $\rho$  is the event-specific average pile-up energy density per unit area in the  $\phi - \eta$  plane, and the effective area suggests the effective area affected by pile-up for each type of Iso.

Because tracks are associated with specific interaction vertices, the  $Iso_{ch}$  term is nearly independent of the  $\rho$ -effective area correction, and the combined PF isolation is defined as

$$Iso = Iso_{ch} + max(0, Iso_{nh} + Iso_{photon} - PU)$$
(3.1)

The relative isolation shown in Table 3.1 is defined as  $relIso = Iso/p_T$ .

### **3.3** Photon Reconstruction

The photon reconstruction has many similarities with the electron reconstruction. Photons interact electromagnetically with material in the detector and can convert into  $e^-e^+$  pairs and emit bremsstrahlung photons like the electrons. The reconstruction of photons relies mostly on ECAL. The photon reconstruction starts from an ECAL supercluster which has no link to a GSF track. The energy deposited in the supercluster is calculated and ECAL energy correction is applied to the supercluster. The result is assigned to the photon as its energy. Due to the missing information in the track, a photon cannot be matched to a vertex.

### 3.3.1 Photon Identification and Isolation

The photon objects used in this analysis are also required to pass the loose cut-based ID and Iso following the recommendation of the EGamma POG. The criteria are listed in Table 3.2.

Variable	Barrel	Endcap
$ \eta_{\rm SC} $ acceptance	(0, 1.479)	(1.479, 2.5)
$\sigma_{i\eta,i\eta} <$	0.0103	0.0301
hOverE <	0.597	0.481
$Iso_{ch} <$	1.295	1.011
$Iso_{nh} <$	$10.910 + 0.0148 \times p_T + 0.000017 \times p_T^2$	$5.931 + 0.0163 * pt + 0.000014 * pt^2$
$Iso_{photon} <$	$3.630 + 0.0047 \times p_T$	$6.641 + 0.0034 \times p_T$

Table 3.2: The cuts used in the POG loose photon identification.

Unlike electrons, the photon Iso has specific requirements on each type of isolations, instead of a single criterion on a combined Iso value.

# 3.4 Muon Reconstruction

The CMS muon reconstruction is mainly based on information collected from the tracker and muon chambers. Generally two collections of muon objects are constructed: the global muon and the tracker muon. [45]

- A Global Muon starts from a stand-alone track in the muon system. If the stand-alone muon track can be matched to a charged-particle track in the tracking system and the properties of the tracks agree, a combined fit is performed over the two tracks to find a global muon track.
- A Tracker Muon starts from a charged-particle track in the tracking system with  $p_T$  over 0.5 GeV and momentum over 2.5 GeV. If one or more muon segments in the muon chamber system match the extrapolation of the charged-particle track, then the track is considered a tracker muon track.

The reconstruction of global muons and tracker muons is performed separately. A muon can be reconstructed both as a global muon and as a tracker muon. The momenta calculated from the inner tracks are assigned to the muons with  $p_T$  less than 200 GeV. For those with  $p_T$  beyond 200 GeV, the Tune-P algorithm [46] is used: the momenta are calculated from the best fit track among the inner track, the global track, the track combining the inner tracker and the first muon station, and the globally fit track after discarding muon chamber stations with high occupancy.

### 3.4.1 Muon Identification and Isolation

Charged hadrons can produce fake muon signatures after reconstruction. To suppress the fake rate of reconstructed muons, additional identification requirements are applied for muons used in CMS analyses. In this analysis, two types of muon ID are used: the muon High  $p_T$  ID and Tracker High  $p_T$  ID.

#### Muon High $p_T$ ID

The High  $p_T$  ID is defined below, based on the CMS Muon POG recommendation for 2016 data analysis.

- The candidate is reconstructed as a Global Muon
- At least one muon-chamber hit is included in the Global Muon track fit;
- Muon segments present in at least two muon stations. This implies that the muon is also a Tracker Muon;
- The relative error in  $p_T$  of the muon best fit track is less than 30%;
- Its tracker track has transverse impact parameter |dxy| < 2mm with respect to the primary vertex;
- The longitudinal distance of the tracker track with respect to the primary vertex is |dz| < 5mm;
- At least one pixel hit is associated with the track;

• The number of tracker layers with hits is required to be more than 5.

The High  $p_T$  ID is designed to enhance the reconstruction of the muon track parameters for high-pT muons. To be consistent with the High  $p_T$  ID, the muon properties obtained from the Tune-P algorithm are used.

#### Tracker High $p_T$ ID

The Tracker High  $p_T$  ID is a customized muon ID for this analysis and is modified from the High  $p_T$  ID, by loosening the global muon requirement to that for a tracker muon, namely removing the following two criteria in the above High  $p_T$  ID:

- The candidate is reconstructed as a Global Muon;
- At least one muon-chamber hit is included in the Global Muon track fit;

and replacing them with the following criterion:

• The candidate is reconstructed as a Tracker Muon

The Tracker High  $p_T$  ID is introduced into this analysis to optimize the muon efficiency. The standard High  $p_T$  ID is designed for the identification of a single isolated high  $p_T$  muon and would result in decreasing muon efficiency when the muon is from the decay of a boosted Z, where the two muons from the Z decay can be too collinear to be reconstructed as two separated stand-alone muon tracks, because of the limited spatial resolution in the muon detector system. The introduction of the Tracker High  $p_T$  ID makes it possible to fully identify the 2 adjacent muons without requiring each of them to contain a stand-alone muon track. The performance of this algorithm is described in Section 4.3.1.

In this analysis the muon momentum information used are always obtained from the Tune-P algorithm.

#### **Muon Isolation**

The muons selected for this analysis are required to be isolated, following the loose tracker isolation recommended by the CMS Muon POG. The sum of the momenta of all charged-particle inner tracks that do not belong to a muon and appears within the cone of  $\Delta R < 0.3$  around the muon track is calculated, to define  $Iso_{tk}$ . Note that only non-muon tracks are taken into account because our events of interest are characterized by adjacent muon pairs.

The isolation requirement is  $relIso_{tk} < 0.1$  for each muon selected, where  $relIso_{tk}$  is the relative tracker Iso and is defined as  $relIso_{tk} = Iso_{tk}/p_T$ .

## **3.5** Missing Transverse Momentum $(p_T^{\text{miss}})$

The Missing Transverse Momentum [47], denoted by  $p_T^{\text{miss}}$  or MET, is calculated as the negative vectorial sum of the momenta of all PF elements,

$$\vec{p_T}^{miss} = -\Sigma_i^{PF} \vec{P}_{T,i}.$$
(3.2)

It is often used to indicates the existence of undetected particles in the final state of an event, such as neutrinos. But the  $p_T^{\text{miss}}$  signature can also result from instrumental effects due to the non linearity of the calorimeter response for hadronic particles, tracker inefficiencies and minimum energy thresholds in the calorimeters. This bias can be reduced by applying jet energy corrections [48] on the particle-level reconstructed jet  $p_T$ . And thus the corrected missing transverse momentum can be obtained by Equation 3.3, where the superscript "corr" indicates the corrected values.

$$\vec{p_T}^{\text{miss,corr}} = \vec{p_T}^{\text{miss}} - \Sigma_j^{jets} (\vec{p_{T,j}}^{\text{corr}} - \vec{p_{T,j}})$$
(3.3)

Here the jets are reconstructed on the basis of all types of PF elements, with the Anti- $k_t$  jet clustering algorithm [49] with the radius parameter R = 0.4.

# **Chapter 4**

# Datasets, Event Selections and Event Reconstruction

# 4.1 Data and Triggers

#### 4.1.1 Data

This analysis uses a data sample recorded by the CMS detector during 2016 with an integrated luminosity of 35.9 fb<sup>-1</sup>. Throughout the year of 2016, the data collection is split into 7 run periods (Run Period B, C, D, E, F, G, H). The data collections and primary datasets (PD) used are listed in Table 4.1. The primary data sets are defined by the presence of identified physics objects.

Data collections	$lumi[fb^{-1}]$	
/PD/Run2016B-03Feb2017_ver2	-v2 5.78	
/PD/Run2016C-03Feb2017-v1	2.57	
/PD/Run2016D-03Feb2017-v1	4.25	
/PD/Run2016E-03Feb2017-v1	4.01	
/PD/Run2016F-03Feb2017-v1	3.10	
/PD/Run2016G-03Feb2017v1	7.54	
/PD/Run2016H-03Feb2017_ver2	-v1 8.39	
/PD/Run2016H-03Feb2017_ver3	-v1 0.22	
Primary datasets (PD):		
SingleElectron, SingleMuon,	SinglePhoton, and MuEG	

Table 4.1: Datasets used in the analysis.

## 4.1.2 High Level Triggers

In this analysis the single lepton (electron or muon)  $High \ Level$  triggers without pre-selections on isolation are required for the signal data selection. The single lepton triggers are preferred over the di-lepton triggers, to maintain signal efficiency: because the two leptons can be very adjacent to each other and therefore difficult to identify at trigger level. Single photon  $High \ Level$  triggers for the  $\gamma$ +jets event selection are used in the data-driven Drell-Yan (Z+jets) background modeling. The detailed HLT paths are listed in Table 4.2.

Table 4.2: Trigger paths used in 2016 collision data.

Muon	
HLT_Mu50 <b>OR</b> HLT_TkMu50	
Electron	
HLT_Ele115_CaloIdVT_GsfTrkIdT	
Photon	
HLT_PhotonPT_R9Id90_HE10_IsoM, for PT = 22,30,36,50,75,90,120,165 GeV	

# 4.2 Monte Carlo Simulation Samples

Monte Carlo (MC) simulation samples are used in this analysis, both to help model the SM background accurately for the data selected from the experiment and to evaluate the likelihood of the existence of the Bulk Graviton signal. Therefore, simulation samples for both SM background and signal are required. The Parton Distribution Functions (PDF), which describe the fraction of momentum a parton carries within a proton, are modeled using the NNPDF3.0 [50] parametrization, based on the recommendation of the PDF4LHC team [51]. The simulation process includes 3 stages: the simulation for the core physics process including the subsequent decays, handled by matrix element generators; the simulation for the hadron showering and hadronization processes, handled by a parton shower generator, which is Pythia8.212 [52] for this analysis; and the simulation for the experimental observation from the detector, based on GEANT4 [53]. A random number of extra pileup interactions are also added for every MC event to better

match the data from the experiment. The MC samples used in this analysis are in the format of the standard CMS Mini-AOD [54], used in the RunIISummer16 analysis campaign.

### 4.2.1 Standard Model Samples

The SM background in this analysis is mostly dominated by the Z+jets process. The other main sources of the background are the non-resonant production of  $\ell$  and  $p_T^{\text{miss}}$  final states, including  $t\bar{t}$  and WW processes, and the resonant background from SM production of diboson events, composed mainly of ZZ and WZ processes.

Background samples used in this analysis are listed in Table 4.3. The cross-section values in the table are noted in parentheses whether they are calculated at LO, NLO or NNLO level.

Backgrounds	MC Dataset	$\sigma[pb]$
Z+jets	DYJetsToLL M-50 TuneCUETP8M1 13TeV-amcatnloFXFX-pythia8	5765.4 (NNLO)
Z Reso.	ZZTo2L2Nu_13TeV_powheg_pythia8	0.564 (NLO)
	ZZTo4L_13TeV_powheg_pythia8	1.212 (NLO)
	ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.22 (NLO)
	GluGluToContinToZZTo2e2nu_13TeV_MCFM701_pythia8	0.00319 (LO)
	GluGluToContinToZZTo2mu2nu_13TeV_MCFM701_pythia8	0.00319 (LO)
	WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.595 (NLO)
	WZTo3LNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	4.42965 (NLO)
Non-Reso.	TTTo2L2Nu_13TeV-powheg	87.31 (NNLO)
	TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.2043 (NLO)
	WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	61526.7 (NLO)
	WWTo2L2Nu_13TeV-powheg	12.178 (NNLO)
	WWToLNuQQ_13TeV-powheg	49.997 (NNLO)
	WZTo1L1Nu2Q_13TeV_amcatnloFXFX_madspin_pythia8	10.71 (NLO)

Table 4.3: Background MC samples and their cross-sections, RunIISummar16 miniAOD.

In the table, the "TuneCUETP8M1" tag stands for the event tune for the Pythia generator [55]; "amcatnloFXFX" in a sample name indicates that the sample is generated by the MadGraph5\_aMC@NLO 2.3.3 [56] framework with next-to-leading order (NLO) matrix elements, merged with Pythia for parton shower matching using the merging scheme of Frederix and Frixione (FxFx) [57]; "powheg" means the Powheg 2.0 [58] generator is used as the

matrix element generator. More photon related SM MC samples are also used for the Z+jets background modeling while applying the photon+jet data driven method, and will be discussed in Section 5.3

### 4.2.2 Signal Samples

Two sets of MC simulation samples are generated for the signal. For the benchmark model, the signal events are generated at leading order for the bulk graviton model by the MadGraph5\_aMC@NLO 2.3.3 matrix element generator and Pythia 8.212. Because the expected width ( $\Gamma$ ) is small compared to detector resolution for reconstructing the signal, we use a zero width approximation for generating signal events. The mass range of the generated signal events is between 600 GeV and 2500 GeV. Table 4.4 shows the mass points generated and used in this analysis, together with their calculated cross-sections with NLO QCD corrections, for  $\tilde{k} = 0.5$ .

Table 4.4: Mass points for narrow width Bulk Graviton signal samples, for the process of  $G \rightarrow ZZ \rightarrow 2\ell 2\nu$ . The corresponding cross-sections are calculated for  $\tilde{k} = 0.5$ .

Mass points [GeV]	$\sigma[pb]$
600	$8.616 \times 10^{-3}$
700	$3.456\times10^{-3}$
800	$1.580\times10^{-3}$
900	$7.893\times10^{-4}$
1000	$4.217\times10^{-4}$
1100	$2.384\times10^{-4}$
1200	$1.399\times10^{-4}$
1300	$8.505\times10^{-5}$
1400	$5.329\times10^{-5}$
1500	$3.437\times10^{-5}$
1600	$2.244\times10^{-5}$
1800	$1.015\times10^{-5}$
2000	$4.860 \times 10^{-6}$
2500	$9.087 \times 10^{-7}$

In addition, a more general model for the bulk graviton decaying to

ZZ is generated using the JHU Generator 7.0.2 [59, 60, 61] and Pythia. This set of the Bulk Graviton samples contains the same mass points as those generated by Madgraph, but a narrow width is not assumed, and graviton samples are generated with  $\Gamma = 0\%$ , 10%, 20% and 30% of the corresponding graviton mass. Production via both gluon fusion and  $q\bar{q}$  annihilation is generated separately for this set of signal samples.

# 4.3 Event Preselection and Reconstruction

To suppress SM backgrounds while keeping the signal events at high efficiency, event selection criteria are made to select interesting events in the datasets. Preselection criteria are used to define a control region to help optimize the background modeling. Afterwards, further selections are made to define the signal region. These will be discussed in Section 4.4.

### 4.3.1 Lepton Selection

Electrons and muons are selected from both single lepton datasets and the MC simulation samples, and paired to be reconstructed as a Z boson candidate. The detailed selections in addition to the single lepton HLT requirements are described below.

#### **Electron Pair Selection**

Both electrons in a selected electron pair are required to pass the Egamma loose cut-based Identification and Isolation as described in Section 3.2.1 to be regarded as a well-identified electron. Additional criteria are required on the  $p_T$  and  $\eta$  values of the leading and subleading electrons in the pair, as listed below.

- 1. Leading Electron:  $p_T > 120 \text{ GeV}, |\eta| < 2.5$
- 2. Subleading Electron:  $p_T > 35 \text{ GeV}, |\eta| < 2.5$

The restriction on their  $|\eta|$  value is due to the design of the ECAL detector. And for the leading electron, only those with  $p_T$  beyond 120 GeV will be kept, because the single electron HLT (HLT\_Ele115\_CaloIdVT\_GsfTrkIdT) we use has a threshold at 115 GeV. More discussion can be found in Section 5.2.3.

#### **Muon Pair Selection**

Two muon identification approaches are involved in the muon pair selection to optimize the muon pair selection efficiency, as described in Section 3.4.1. Various combinations of muon IDs are considered. Figure 4.1 shows the muon pair efficiencies for different muon ID combinations versus  $\Delta R$ between the two muons in the pair and the  $p_T$  of the muon pair.

The efficiency profiles are obtained from the signal MC samples generated by Madgraph, and generator level information is used to judge if a reconstructed muon is a true muon or not. The two efficiency plots are consistent considering the fact that the  $\Delta R$  between the muons is likely to be smaller in a pair with higher  $p_T$ . From the plot one can see huge sacrifice in efficiency by requiring both muons to pass the High  $p_T$  ID, because the resolution of the muon system is not high enough to distinguish the two adjacent muons and reconstruct them as separated tracks. However, the other two ID combinations (two Tracker High  $p_T$  ID combination and Tracker High  $p_T$  ID + High  $p_T$  ID combination) perform with much higher efficiency in the high  $p_T$  region. Considering that the High  $p_T$  ID is a tighter identification compared to the tracker High  $p_T$  ID, the Tracker High  $p_T$  ID + High  $p_T$  ID combination is used as the muon identification in this analysis, to secure high selection efficiency as well as good muon identification quality. In addition to the muon identification, the isolation requirement  $relIso_{tk} < 0.1$  is also applied for each of the muons in the pair, as described in 3.4.1.

Similar to the electron selection, additional criteria on the muons'  $p_T$  and  $\eta$  values are applied:

- 1. Leading Muon:  $p_T > 60 \text{ GeV}, |\eta| < 2.4$
- 2. Subleading Muon:  $p_T > 20$  GeV,  $|\eta| < 2.4$



Figure 4.1: Efficiency comparison among various muon ID combinations, versus  $\Delta R$  between the muons (upper) and  $p_T$  of the muon pair (lower).

The restriction on their  $|\eta|$  value is due to the design of muon detector system. For the leading muon,  $p_T > 60$  GeV is required because the single muon HLTs (HLT\_Mu50 and HLT\_TkMu50) have thresholds at 50 GeV.

#### 4.3.2 Leptonic Z Boson Reconstruction

A selected lepton pair can either be an electron pair or a muon pair, and the two leptons are required to have opposite charge. In one event, more than one such lepton pairs can exist, and in this case, the best lepton pair will be selected. The best pair selection is based on the invariant mass of the lepton pair: only the lepton pair with the invariant mass closest to the true Z boson mass (91.1876 GeV) will be selected among all the lepton pairs in an event. A Z mass window is set between 70 GeV and 110 GeV, so only events having a lepton pair with invariant mass within the Z mass window are kept.

To suppress the low energy backgrounds,  $p_T^Z > 50 \text{ GeV}$  is required for the pre-selection.

#### 4.3.3 MET Filters

Based on the recommendation by the JetMET group, MET filters are applied to ensure the quality of the  $p_T^{\text{miss}}$  reconstruction, for both Monte Carlo samples and data. The MET filters are listed below:

- Flag\_EcalDeadCellTriggerPrimitiveFilter
- Flag\_HBHENoiseIsoFilter
- Flag\_goodVertices
- Flag\_HBHENoiseFilter
- Flag\_globalTightHalo2016Filter
- Flag\_eeBadScFilter
- Flag\_BadPFMuonFilter

- Flag\_BadChargedCandidateFilter
- Flag\_noBadMuons

# 4.4 Signal Region

To further suppress the background and improve the statistical significance of the search, a signal region is defined. Following the pre-selection, further selection on the lepton pair  $p_T$  and  $p_T^{\text{miss}}$  are made as follows:  $p_T^Z >$ 100 GeV and  $p_T^{\text{miss}} > 50$  GeV. These are applied to match the signature of the two boosted Z bosons of the analysis.

Additionally, because the signal model contains the two Z bosons coming from the decay of a heavy resonance, and thus are mostly back to back, the variable  $|\Delta \Phi(p_T^Z, p_T^{\text{miss}})|$ , which describes the angle in the XY plane between  $p_T^Z$  and  $p_T^{\text{miss}}$ , is used to help separate the signal from the background, especially Z+jets events. The signal events tend to peak at a large  $|\Delta \Phi(p_T^Z, p_T^{\text{miss}})|$  value, while for the Z+jets events  $|\Delta \Phi(p_T^Z, p_T^{\text{miss}})|$  is relatively flatter, because the  $p_T^{\text{miss}}$  is due to instrumental effects. Therefore,  $|\Delta \Phi(p_T^Z, p_T^{\text{miss}})| > 0.5$  is applied in the signal region. Figure 4.2 shows the  $|\Delta \Phi(p_T^Z, p_T^{\text{miss}})|$  distributions for the background and signal processes.

To summarize, the selection criteria in the signal region are listed below:

- Trigger: single muon and single electron triggers as documented in Section 4.1.2
- 2. Lepton ID and ISO:

Muons: Tracker High  $p_T$  ID + High  $p_T$  ID combination in a muon pair, both muons are also required to pass tracker ISO

Electrons: both electrons pass Loose Electron ID and corresponding PF Isolation

3. Lepton acceptance cuts:

Muons: Leading  $p_T > 60$  GeV, subleading  $p_T > 20$  GeV, both muons in  $|\eta| < 2.4$ 



Figure 4.2:  $|\Delta \Phi(p_T^Z, p_T^{\text{miss}})|$  distributions for background and signal

Electrons: Leading  $p_T > 120$  GeV, subleading  $p_T > 35$  GeV, both electrons in  $|\eta| < 2.5$ 

- Dilepton pair selection: A pair of same flavor opposite sign leptons. The pair with closest invariant mass to the true Z boson mass is selected
- 5. Z mass window:  $70 < |M_{ll}| < 110 \,\text{GeV}$
- 6. Z boson:  $p_T > 100 \,\text{GeV}$
- 7. Missing  $p_T$ :  $p_T^{\text{miss}} > 50 \text{ GeV}$
- 8.  $|\Delta \Phi(p_T^Z, p_T^{\text{miss}})| > 0.5$

# **Chapter 5**

# **Background Modeling and Plots**

Background models are developed using the data distributions in the control region. The main sources of SM background in this analysis can be divided into 3 groups:

- 1. Z+jets events: the Z+jets process is the dominant background source for this search. Lepton pairs are present in the final state from the Z decay.  $p_T^{\text{miss}}$  is only instrumental.
- 2. Non-resonant events: Events that have lepton pairs and  $p_T^{\text{miss}}$  in the final state, while the lepton pair does not come from the decay of a resonance.  $t\bar{t}$  and WW processes are the main sources of the non-resonant background
- 3. **Resonant events**: Events that have leptons and  $p_T^{\text{miss}}$  in the final state, where the lepton pair comes from the decay of a Z boson. ZZ and WZ are the main sources of the resonant background

The background modeling starts from tuning the MC simulation samples. MC simulation is widely used in this analysis. The modeling of the resonant background and the signal completely rely on MC simulation. Though the modeling of the Z+jets and non-resonant background are done by datadriven methods, simulation samples are also used in the process. Various weights are assigned to the MC events to compensate for discrepancies between the simulation and the actual experiment, in terms of pileup, HLT and lepton ID/Iso. The details of the background strategies for the three background groups are described below.

## 5.1 Pileup re-weighting

The presence of pileup interactions can affect the quality of the event reconstruction, such as the value of  $p_T^{\text{miss}}$  and lepton isolation. For the MC events, a random number of pileup interactions are added during their production, which does not necessarily agree with the pileup distribution in the data. To align the distribution of pileup interactions between data and simulation, pileup re-weighting is applied to the MC samples.

Because the reconstruction efficiencies of the vertices might differ between MC and data, it is preferable to reweight the MC based on the distribution of the number of actual pileup interactions (true pileup), rather than observed interactions. The value of the weight assigned to the MC events is obtained from the ratio of calculated data pileup profile and the MC pileup profile. The MC pileup profile can be found in the configuration of the RunIISummer16 MC production. The calculated data pileup profile is based on the instantaneous luminosity for each bunch crossing noted in the CMS pileup JSON file, and the recommended Mini-Bias cross-section of 69.2 mb  $\pm$  4.6% evaluated by the CMS luminosity POG.

Figure 5.1 shows the pileup distribution for both data and MC, as well as their ratio. The pileup reweighting value for each event in the MC samples is obtained from the ratio plot based on the pileup number in that event.

## **5.2 Monte Carlo Efficiencies**

Weights are also applied to the MC samples in terms of the discrepancies in HLT and lepton ID/Iso efficiencies between data and MC samples. The probability of a lepton passing HLT, ID and Iso is:

$$P(HLT \cdot ID \cdot Iso) = P(HLT|ID \cdot Iso) \times P(ID \cdot Iso)$$
$$= P(HLT|ID \cdot Iso) \times P(Iso|ID) \times P(ID).$$

where  $P(HLT \cdot ID \cdot Iso)$  denotes the efficiency of a lepton passing HLT, ID and Iso;  $P(HLT|ID \cdot Iso)$  is the efficiency of a lepton passing HLT under the condition of it passing ID and ISO, and is also referred to as trigger



Figure 5.1: Pileup number profiles (left) for RunIISummer16 MC and 2016 data; and the pileup re-weighting function (right).

efficiency;  $P(ID \cdot Iso)$  is the combined ID/Iso efficiency; P(Iso|ID) is the efficiency of a lepton passing Iso under the condition that it passes ID, referred to as the Iso efficiency; P(ID) is the ID efficiency. The efficiencies are all calculated using a Tag-and-Probe method in this analysis, as discussed below.

### 5.2.1 Muon ID/Iso Efficiency

The ID/Iso efficiencies may differ between MC and data due to imprecise modeling of detector performance. Therefore efficiencies are measured and scale factors (SF) are calculated for the MC samples to counter this effect. The Muon High  $p_T$  ID, Tracker High  $p_T$  ID and the Tracker Isolation efficiencies are measured separately using the tag-and-probe method described below. Events are selected from the SingleMuon dataset for data and DYJetsToLL\_M-50\_TuneCUETP8M1\_13TeV-madgraphMLM-pythia8 dataset for MC with pileup reweighting applied.

In the tag-and-probe method, a tag muon and a probe muon are selected from an event. The tag muon is a muon required to be identified with high accuracy, which helps eliminate bias, and the probe muon is the muon candidate from which the efficiency is studied. A event is kept only if two muons are found with one passing the "tag" criteria while the other passing the "probe" criteria. For the muon ID/Iso efficiency measurement, the tag criteria are:

- 1. passing the tight Muon ID and the relIso < 0.2 Iso requirement recommended by CMS Muon POG
- 2. passing HLT\_IsoMu24 and  $p_T > 26 \,\text{GeV}$

The ID and Iso requirements ensure that the tag muon is a well defined muon. The HLT\_IsoMu24 requirement is the un-prescaled HLT with lowest  $p_T$  threshold in the SingleMuon dataset. Requiring the tag muon to pass this HLT assures that the selected event will be kept in the SingleMuon dataset so that the efficiencies to be measured for the probe muon would not be biased due to the dataset selection. The  $p_T > 26$  GeV requirement is added to enforce that the trigger is highly efficient for the tag muon. A muon candidate is considered as a probe if it is reconstructed as either a global muon or a tracker muon, with  $p_T > 20$  GeV.

A known issue with the tracker system during Run Periods B to F caused some tracking inefficiency in the presence of heavily ionizing particles (HIPs). The effect on the muon reconstruction is found to be minor, and accounted for by calculating the efficiency in data for muons separately for Runs B to F and Runs G to H. And an additional 1% uncertainty is assigned to cover the effect of the tracking inefficiency issue on the muon reconstruction. Considering 19.71 fb<sup>-1</sup> in run B to F and 16.15 fb<sup>-1</sup> in Run G to H, a flat random number between [0,1] is generated for each MC event, and if it is larger than 19.71 fb<sup>-1</sup>/(19.71 fb<sup>-1</sup> + 16.15 fb<sup>-1</sup>) = 0.5496 the scale factors calculated from the muon efficiency in Run B to F would be assigned to the event, otherwise the SFs from Run GH are assigned.

#### **Muon ID Efficiency**

In the ID efficiency measurement, any muon candidate from the object reconstruction with  $p_T > 20 \text{ GeV}$  and  $|\eta| < 2.4$  (base muon selection in the analysis) could be considered as a probe muon. The invariant mass spectra of the tag-probe muon pairs are calculated for various  $p_T - \eta$  bins,

with each set including a spectrum for those pairs with probe muon passing the ID, while another spectrum for pairs with probe muon failing the ID. The invariant mass spectrum consists of  $Z \rightarrow \mu\mu$  process and various background processes. The  $Z \rightarrow \mu\mu$  events are considered signal and the probe muon in these events are true muons. In this case, the efficiency is calculated as

$$\epsilon = N_{signal}^{pass} / (N_{signal}^{pass} + N_{signal}^{fail}).$$
(5.1)

The spectra are fit into signal + background models, and the integral of the signal shape is  $N_{signal}^{pass}$  in the passing category and  $N_{signal}^{fail}$  in the failing category for each bin.

The signal function used for the fitting is the sum of 2 Voigtians. The background profile can either be RooCMSShape [62] or a third order Chebychev polynomial. An example of the fitting plots is given in Figure 5.2, for the Tracker High  $p_T$  muon ID efficiency measurement.



Figure 5.2: An example of the  $\mu\mu$  invariant mass spectrum in one  $p_T - \eta$  bin with the *signal* + *background* fitting (solid lines) for the efficiency measurement of the Tracker High  $p_T$  ID.

The dashed lines show the background contribution.

Figures 5.3 to 5.5 show the efficiency results calculated from both Tracker High  $p_T$  ID and High  $p_T$  ID.



Figure 5.3: High  $p_T$  Muon ID efficiency for 2016 ReReco data as a function of muon  $p_T$  and  $|\eta|$ , for 2016 Run Peroids B–F (left) and 2016 G–H (right).



Figure 5.4: Tracker High  $p_T$  Muon ID efficiency for 2016 ReReco data as a function of muon  $p_T$  and  $|\eta|$ , for 2016 Run Peroids B–F (left) and 2016 G–H (right).

Because the Tracker High  $p_T$  ID is a loosened version of the High  $p_T$  ID, if a muon passes the High  $p_T$  ID, it will also pass the Tracker High  $p_T$ 



Figure 5.5: Muon ID efficiency for RunIISummer16 MC as a function of muon  $p_T$  and  $|\eta|$ , for High  $p_T$  Muon ID (left) and Tracker High  $p_T$  Muon ID (right).

ID. The muon ID scale factor for an event is therefore calculated as:

$$SF = \epsilon^{data} / \epsilon^{MC}$$

$$= \frac{(\epsilon_{HighPt}(\mu_1) \times \epsilon_{trkHighPt}(\mu_2) + \epsilon_{trkHighPt}(\mu_1) \times \epsilon_{HighPt}(\mu_2) - \epsilon_{HighPt}(\mu_1) \times \epsilon_{HighPt}(\mu_2))_{data}}{(\epsilon_{HighPt}(\mu_1) \times \epsilon_{trkHighPt}(\mu_2) + \epsilon_{trkHighPt}(\mu_1) \times \epsilon_{HighPt}(\mu_2) - \epsilon_{HighPt}(\mu_1) \times \epsilon_{HighPt}(\mu_2))_{MC}}$$

#### **Muon Iso Efficiency**

The muon tracker isolation efficiency is also measured using the tag-andprobe method, with an additional requirement on the probe muon to pass the tracker High  $p_T$  ID. Because the tracker isolation selection applies to both muons, the ratio between the efficiencies of data and MC is used as the MC scale factor for the isolation. The SF values versus  $p_T$  and  $\eta$  are shown in Figure 5.6

#### 5.2.2 Electron ID/Iso Efficiency

Based on the recommendation from the CMS EGamma POG, the Loose cut-based identification (ID) selection is required for all the electron can-



Figure 5.6: tracker ISO data/MC efficiency scale factors as a function of muon  $p_T$  and  $|\eta|$ , for 2016 Run Peroids B–F (left) and 2016 G–H (right).

didates. Because a PF isolation is already included in the Loose ID, no additional electron isolation is needed.

The electron Loose ID (including PF ISO) efficiency and scale factors are provided by the EGamma POG, measured by the tag-and-probe method. Figure 5.7 shows the electron Loose ID scale factors used in this analysis. The electron reconstruction is also affected by the tracking inefficiency issue, the reconstruction scale factors are also provided by the EGamma POG to counter this effect. Figure 5.8 shows the electron reconstruction scale factors from EGamma POG used in this analysis.

#### 5.2.3 Trigger Efficiency

The HLT is designed for making fast decisions for accepting data, and therefore the reconstruction of objects' properties at the HLT level are not as precise compared to the offline reconstruction. Trigger efficiency studies for offline objects are important to physics analyses in two regards: to suppress the trigger efficiency effect on the data by optimizing data selection; and to compensate the discrepancy in terms of trigger efficiency between data and simulation samples by applying trigger efficiency SFs to



Figure 5.7: EGamma POG electron Loose ID (including pf Iso) efficiency scale factors for 2016 dataset analysis.



Figure 5.8: EGamma POG electron reconstruction scale factors for 2016 dataset analysis.

MC samples.

#### SingleMuon HLT Efficiency

For the muon channel selection, events are required to pass the single muon HLT requirement of either HLT\_Mu50 or HLT\_TkMu50. The combined trigger efficiencies and SFs are centrally derived by the Muon POG, measured with the tag-and-probe method. The summary plots in Figure 5.9 show the trigger efficiencies versus  $p_T$  (left) and  $\eta$  (right). In the  $p_T$  plot, a rising edge of the efficiency can be clearly seen around the  $p_T$  threshold of the HLT. To avoid events falling on the rising edge in the trigger efficiency, leading to greater difficulty in the background modeling, the  $p_T$  of the leading muon is required to be at least 60 GeV.

The efficiency results are delivered in the form of 2D histograms in  $p_T - \eta$ . The calculated ratio of efficiencies between data and MC are applied to the leading muon of the pair in the MC events only as the trigger efficiency scale factor, considering that number of events in either data or



Figure 5.9: Muon trigger efficiency for 2016 data and MC, versus  $p_T$  (left) and  $\eta$  (right)

MC is negligible with the subleading muon passing the HLT while leading muon fails the selection.

#### SingleElectron HLT Efficiency

The SingleElectron HLT (HLT\_Ele115\_CaloIdVT\_GsfTrkIdT) efficiency is measured for data and MC using the tag-and-probe method. Events are selected from the SingleElectron dataset in data and the DYJetsToLL\_M-50\_TuneCUETP8M1\_13TeV-amcatnloFXFX-pythia8 dataset for MC with pileup reweighting. The "tag" criteria are:

- 1. passing the tight cut-based identification (ID) and isolation (Iso) recommended by the EGamma POG
- 2. passing HLT\_Ele27\_WPTight\_Gsf with  $p_T > 30$  and  $|\eta| < 2.1$

The criterion of tight ID/Iso ensures the tag electron to be a well identified electron. The HLT\_Ele27\_WPTight\_Gsf is the un-prescaled HLT with lowest  $p_T$  threshold in the SingleElectron dataset. This requirement ensures that the trigger efficiency to be measured from the probe electron is not biased due to the dataset for selection. A electron is considered as a probe
if it satisfies the condition of the trigger efficiency measurement, which is passing the loose ID/Iso in this analysis, as described in Section 3.2.1.

The electron trigger efficiencies are measured as a function of the reconstructed electron  $p_T$  and  $|\eta|$  for both the 2016 full dataset and RunIISummer16 MC. Figure 5.10 gives an example of invariant mass spectrum of the electron pair. Because only electrons passing ID/Iso criteria are selected, the background fraction is very small for the trigger efficiency measurement.



Figure 5.10: An example of one  $p_T - |\eta|$  bin of the electron pair invariant mass spectrum with the *signal* + *background* fit for the efficiency measurement of HLT\_Ele27\_WPTight\_Gsf. The background component shown by the blue line is negligible due to the ID/Iso criteria on the probe electron.

The measured efficiencies are shown in Figures 5.11 and 5.12 for data and MC respectively. The corresponding Data/MC scale factors are shown in Figure 5.13. Similar to the muon channel, the SingleElectron HLT efficiency SF is applied on the leading electron in the MC samples as reweighting factors to model the electron trigger efficiency, and a  $p_T > 120$  GeV selection is added to the leading electron.



Figure 5.11: Electron trigger efficiency from 2016 ReReco dataset as a function of reconstructed electron  $p_T$  and  $|\eta|$ . Left for  $p_T < 150$  GeV, right for  $p_T > 150$  GeV.



Figure 5.12: Electron trigger efficiency from RunIISummer16 MC as a function of reconstructed electron  $p_T$  and  $|\eta|$ . Left for  $p_T < 150$  GeV, right for  $p_T > 150$  GeV.



Figure 5.13: Electron trigger efficiency Data/MC scale factors as a function of reconstructed electron  $p_T$  and  $|\eta|$ . Left for  $p_T < 150$  GeV, right for  $p_T > 150$  GeV.

# 5.3 Z+jets Background Modeling

In this analysis data-driven background modeling methods are used for the majority of the backgrounds, including the Z+jets background. The data-driven method in this analysis has benefits in that:

- 1. The number of available MC events in the high energy region is limited, while the data-driven method offers more statistics in our region of interest.
- 2. The modeling of  $p_T^{\text{miss}}$  for the signal region is more reliable in the data-driven background methods, considering that detector conditions and pileup modeling may be imperfect in MC.

As Figures 5.14 and 5.15 show, even with the pileup reweightings applied, the vertex multiplicity and  $\rho$  (a parameter describing the pileup energy in the detector) distributions are still not precisely reproduced by MC.



Figure 5.14: Number of reconstructed vertices for electron (left) and muon (right) channels comparing data and MC.



Figure 5.15: The  $\rho$  distributions for electron (left) and muon (right) channels comparing data and MC.

The Z+jets background is characterized by a transversely boosted Z boson and a hadronic recoil balancing the momentum of the Z boson. The observed  $p_T^{\text{miss}}$  in this background is primarily instrumental. The  $\gamma$ +jets events have a similar kinematic signature as Z+jets events, and can be used to model the Z+jets background in a data-driven way. This  $\gamma$ +jets datadriven approach is preferred over the MC background modeling, because the  $\gamma$ +jets sample offers more statistics in our regions of interest and the  $p_T^{\text{miss}}$  reconstruction includes all real world detector effects that are difficult to model precisely in the MC samples. Also the  $\gamma$ +jets process is distinct from the signal process and therefore provides a pure background model.

The ultimate goal is to use the  $\gamma$ +jets events to model the  $m_T$  distribution of the Z+jets background. According to Equation 1.4,  $m_T$  is determined by  $m_{\ell\ell}, p_T^{\ell\ell}, p_T^{miss}$  and  $p_T^{miss}$ , where  $m_{\ell\ell}$  is the invariant mass of the lepton pair;  $p_T^{\ell\ell}$  is the  $\vec{p}_T$  sum of the lepton pair;  $p_T^{miss}$  refers to the projection of  $p_T^{miss}$  in the direction of  $p_T^{\ell\ell}$  and  $p_T^{miss}_{\perp}$  is the portion perpendicular to  $p_T^{\ell\ell}$ . It is crucial to model  $m_{\ell\ell}, p_T^{\ell\ell}, p_T^{miss}$  and  $p_T^{miss}_{\perp}$  well. The workflow of the Z+jets data-driven modeling can be summarized as below and will be discussed in the following sections:

### • photon data cleaning

- $\gamma$ +jets Event Selection in Section 5.3.1
- $\gamma$ +jets HLT Prescale Reweighting in Section 5.3.2
- Physical  $p_T^{\text{miss}}$  Subtraction in Section 5.3.3
- $p_T^{\ell\ell}$  correction for photon data
  - $p_T^{\gamma}$  to  $p_T^Z$  Reweighting in Section 5.3.4
- $m_{\ell\ell}$  correction for photon data
  - Photon Mass Generation in Section 5.3.5
- $p_{T\parallel}^{miss}$  and  $p_{T\perp}^{miss}$  corrections for photon data
  - $p_T^{\text{miss}}$  Hadronic Recoil Tuning in Section 5.3.6

# 5.3.1 $\gamma$ +jets Event Selection

The  $\gamma$ +jets events are selected from the SinglePhoton dataset with HLT HLT\_Photon"PT"\_R9Id90\_HE10\_IsoM, for PT = 22, 30, 36, 50, 75, 90, 120, 165 GeV. The Loose Photon ID defined and recommended by the CMS EGamma POG is applied. Furthermore, MET filters listed in Section 4.3.3 are also required in the photon data selection.

Even with the LOOSE Photon ID applied, the photon samples are still contaminated by many fake photons from sources such as ECAL APD spikes, ECAL noise that has not been flagged, and beam halo particles. Therefore, additional selections are applied as listed below:

- Only one reconstructed photon in the event.
- A cleansing filter applied based on problematic ECAL channel maps.
- sigmaletaleta> 0.001

- sigmaIphiIphi> 0.001
- "Swiss Cross": S = (1 E4/E1) < 0.95, where E1 is the seed crystal energy, E4 is the sum of the energies in up, down, left and right crystals adjacent to the seed crystal.
- ECAL seed crystal timing :  $t_0 1.5ns < time < t_0 + 1.5ns$ , where  $t_0$  is the peak time position.
- Minimum ionizing particle (MIP) total energy < 4.9 GeV to suppress halo induced showers in the ECAL.
- Lepton veto: remove events with one or more reconstructed electrons with p<sub>T</sub> > 10 GeV, also remove events with jets containing more than 10 GeV of lepton energy. This is to filter out processes such as Z → ee events, with one electron mis-identified as the photon.

In addition, analogous to the Z boson preselection,  $p_T^{\gamma} > 50 \text{ GeV}$  is applied in the photon data selection.

## 5.3.2 $\gamma$ +jets HLT Prescale Reweighting

Both the L1T and HLT can be prescaled in order to suppress the very high rate low energy events. For the HLTs used in the  $\gamma$ +jets event selection: HLT\_Photon< $p_T$ >\_R9Id90\_HE10\_IsoM,

for  $p_T = 22, 30, 36, 50, 75, 90, 120,165 \text{ GeV}$ , prescales are applied as follows:

- $p_T$  threshold = 165 GeV: not pre-scaled,
- $p_T$  threshold = 50, 75, 90, and 120 GeV: pre-scaled at only HLT.
- $p_T$  threshold = 22, 30, 36 GeV: pre-scaled at both L1T and HLT,

For photon triggers with  $p_T$  of 50 GeV and higher, the L1T prescale has no effect on our selected data. The HLT prescale factor is obtained from the trigger conditions information stored for the CMS data, and applied as a weight to correct the effect of the HLT prescale on the  $p_T$  spectrum of the photons. Figure 5.16 shows the photon  $p_T$  spectrum with and without corrections for the HLT prescales.



Figure 5.16: Photon  $p_T$  distributions with and without the HLT prescale correction.

# **5.3.3** Physical $p_T^{\text{miss}}$ Subtraction

Like the Z+jets process, the  $p_T^{\text{miss}}$  in the process of  $\gamma$ +jets is due to instrumental effects. However, processes like  $W\gamma \rightarrow \ell \nu \gamma$  with physical  $p_T^{\text{miss}}$ 

can also contribute to the SinglePhone dataset. To model the kinematics of the  $\gamma$ +jets process better, events with physical  $p_T^{\text{miss}}$  are subtracted from the SinglePhoton dataset using MC. MC samples shown in Table 5.1 are used to describe the single photon dataset. Figures 5.17 and 5.18 show the comparisons of the SinglePhoton data and MC for 2016 full dataset. The same set of SinglePhoton HLTs listed in Section 5.3.2 has been applied on the MC samples, and the HLT prescale corrections are applied too. The discrepancy in the trigger efficiencies between MC and data is evaluated and addressed by the HLT efficiency corrections on the MC samples.

Table 5.1: MC samples and their cross-sections for describing photon data and for physical  $p_T^{\text{miss}}$  subtraction, Summer16 miniAODv2.

MC Dataset	$\sigma[pb]$
Instrumental $p_T^{miss}$	
GJets_HT-*To*_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	32701 (LO)
QCD_Pt-*to*_EMEnriched_TuneCUETP8M1_13TeV_pythia8	$1.86049 \times 10^7$ (LO)
Physical $p_T^{\text{miss}}$	
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	5765.4 (NNLO)
ZJetsToNuNu_HT-*To*_13TeV-madgraph	457.081 (NLO)
WJetsToLNu_HT-*To*_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	2144.75 (NLO)
ZNuNuGJets_MonoPhoton_PtG-130_TuneCUETP8M1_13TeV-madgrap	$0.183 \times 1.43$
ZNuNuGJets_MonoPhoton_PtG-40to130_TuneCUETP8M1_13TeV-madgrap	$2.816 \times 1.43$
WGToLNuG_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	$585.8\times2.51$
TTGJets_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	3.697 (NLO)
ST_t-channel_top_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1	136.02 (NLO)
ST_t-channel_antitop_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1	80.95 (NLO)
ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.6 (NNLO)
ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.6 (NNLO)
TGJets_TuneCUETP8M1_13TeV_amcatnlo_madspin_pythia8	2.967 (NLO)

The MC components with physical  $p_T^{\text{miss}}$  in the final states, such as  $Z \rightarrow \nu\nu$ ,  $Z\gamma \rightarrow \nu\nu\gamma$ ,  $W\gamma \rightarrow \ell\nu\gamma$ ,  $W \rightarrow \ell\nu$ , are subtracted from the SinglePhoton data, by merging these MC events into the photon data sample, with a weight of -1.



Figure 5.17: The photon  $p_T$  (left) and  $\eta$  (right) distributions and the MC sample description for the SinglePhoton dataset.



Figure 5.18: The photon  $p_T^{\text{miss}}$  (left) and  $\Delta \Phi(p_T^{\gamma}, p_T^{\text{miss}})$  (right) distributions and the MC sample description for the SinglePhoton dataset.

# 5.3.4 $p_T^{\gamma}$ to $p_T^Z$ Reweighting

The kinematic signature of  $\gamma$ +jets events is expected to be similar to Z+jets events, especially in the high energy region where the mass of Z bosons can be neglected. However, at lower energies the mass effects will alter the kinematics, and more importantly, the lepton selections applied in the Z boson reconstruction have different efficiencies compared to that for the photon reconstruction. To address this issue, the photon  $p_T$  distribution is reweighted to match that of the Z  $p_T$ .

As there is no easy way to extract a clean Z boson  $p_T$  spectrum from the data with no background processes. The  $p_T$  spectrum of the Z bosons is obtained from the MC sample DYJetsToLL\_M-50\_TuneCUETP8M1\_13TeV-amcatnloFXFX-pythia8, with inclusive cross-section of 5765.4 pb ( $\pm 1.7\%$  PDF uncertainty) calculated at NNLO from FEWZ 3.1 [63]. The differential cross-section with respect to  $p_T^Z$  is reweighted to the Z+jets differential cross-section measured from the 2015 CMS data [64] and corrected to the generator level. The standard preselection is applied to the MC samples, with all efficiency reweightings applied. Figure 5.19 and 5.20 shows the photon  $p_T$  reweighting function for electron and muon channels separately.

#### 5.3.5 Photon Mass Generation

The mass of the leptonic Z boson is used in the transverse mass calculation (Equation 1.3, 1.4), and must be simulated in the  $\gamma$ +jets events to model the  $m_T$  in the Z+jets background. This is done by assigning a random mass to the photon based on the Z boson mass distribution and parameterized as a function of Z boson  $p_T$ .

Figures 5.21 and 5.22, compare the Z mass distributions of Z+jets MC and the simulated Z mass for  $\gamma$ +jets data events, for electron and muon channels separately.



Figure 5.19: Photon  $p_T$  reweighting function for the electron channel. The uncertainty bands includes uncertainties from 2015 CMS Z+jets differential cross-section measurements, the statistical uncertainty from Z+jets MC sample and  $\gamma$ +jets data sample, and the lepton trigger, ID, ISO efficiency scale factors.

## **5.3.6** $p_T^{\text{miss}}$ Hadronic Recoil Tuning

The Z+jets process consists of a Z boson decaying to a lepton pair and a hadronic recoil balancing the Z boson  $p_T$  in the transverse plain. As a result, in theory the  $p_T^{\text{miss}}$  should be 0. However, due to limitations of the detector resolutions for leptons and jets,  $p_T^{\text{miss}}$  is present in Z+jets process due to instrumental effects. The energy resolution of leptons, jets, photons are potentially different between  $\gamma$ +jets data and Z+jets data, and those differences may introduce differences in the resolution and scale of reconstructed  $p_T^{\text{miss}}$ .



Figure 5.20: Photon  $p_T$  reweighting function for the muon channel. The uncertainty bands includes uncertainties from 2015 CMS Z+jets differential cross-section measurements, the statistical uncertainty from Z+jets MC sample and  $\gamma$ +jets data sample, and the lepton trigger, ID, ISO efficiency scale factors.

A single-Gaussian based hadronic recoil fit is developed to tune the  $p_T^{\text{miss}}$  of the  $\gamma$ +jets data to match the  $p_T^{\text{miss}}$  in the Z+jets data. The general idea is to apply a correction to the  $\gamma$ +jets data to better describe the  $p_T_{\parallel}^{\text{miss}}$  and  $p_T_{\perp}^{\text{miss}}$  distributions of the Z+jets process, where  $p_T_{\parallel}^{\text{miss}}$  refers to the projection of  $p_T^{\text{miss}}$  in the direction of  $p_T^{z(\gamma)}$  and  $p_T_{\perp}^{\text{miss}}$  is the fraction perpendicular to  $p_T^{z(\gamma)}$ .

The  $p_T_{\parallel}^{miss}$  and  $p_T_{\perp}^{miss}$  are fit with a Gaussian function in the region of [-50, 50] GeV. The Gaussian mean values and resolutions are parameterized as functions of the  $p_T$  of Z boson and photon. Figure 5.23 gives some



Figure 5.21: Z mass distributions for electron channel, comparing Z+jets MC (left) and the simulated Z mass for  $\gamma$ +jets data events (right).



Figure 5.22: Z mass distributions for muon channel, comparing Z+jets MC (left) and the simulated Z mass for  $\gamma$ +jets data events (right).

example plots showing the Gaussian fits for the  $p_T$  bins 50-60 GeV for Z+jets data, for muon channel and electron channel, for  $p_T_{\parallel}^{miss}$  and  $p_T_{\perp}^{miss}$ , respectively.



Figure 5.23: Example plots for the single Gaussian-based  $p_T^{\text{miss}}$  hadronic recoil fit of a selected Z  $p_T$  bin for Z+jets data, muon channel (upper), electron channel (lower),  $p_T_{\parallel}^{\text{miss}}$  (left),  $p_T_{\perp}^{\text{miss}}$  (right).

The comparison of the recoil fit results between Z+jets data and  $\gamma$ +jets data before correction are shown in Figures 5.24 and 5.25 for muon channel and electron channel, respectively.

The  $\gamma$ +jets data  $p_T_{\parallel}^{miss}$  peak and the resolution of  $p_T_{\parallel}^{miss}$  and  $p_T_{\perp}^{miss}$  are then corrected to match the Z+jets data respectively for muon and electron channels based on Figures 5.24 and 5.25. The correction for the photon data is done in the following way:



Figure 5.24: Comparison of the recoil fitted peak positions and Gaussian resolutions for  $p_T_{\parallel}^{miss}$  and  $p_T_{\perp}^{miss}$  between di-lepton data and  $\gamma$ +jets data for the muon channel. Upper two for  $p_T_{\parallel}^{miss}$ , bottom two for  $p_T_{\perp}^{miss}$ .

- Shift the  $p_{T_{\parallel}}^{miss}$  peak positions of  $\gamma$ +jets data to match that of the Z+jets data;
- Scale the  $p_T_{\parallel}^{miss}$  and  $p_T_{\perp}^{miss}$  resolution of  $\gamma$ +jets data to match that of the Z+jets data.

The peak position correction is only applied to  $p_T_{\parallel}^{miss}$ , as the  $p_T^{miss}$  results from the instrumental imbalance between the Z boson and hadronic recoil, therefore the distribution of  $p_T_{\parallel}^{miss}$  is more imbalanced while the peak position for  $p_T_{\perp}^{miss}$  is expected to be 0.



Figure 5.25: Comparison of the recoil fitted peak positions and Gaussian resolutions for  $p_{T_{\parallel}}^{miss}$  and  $p_{T_{\perp}}^{miss}$  between di-lepton data and  $\gamma$ +jets data for the electron channel. Upper two for  $p_{T_{\parallel}}^{miss}$ , bottom two for  $p_{T_{\perp}}^{miss}$ .

# 5.4 Non-resonant Background Modeling

The non-resonant background contains mostly  $t\bar{t}$  and WW events. A data driven method is used in the analysis to model the non-resonant background. The method is to use the  $e\mu$  pairs to describe the non-resonant background in  $\ell\ell$  (*ee* or  $\mu\mu$ ) events, based on the fact that in  $t\bar{t}$  and WW decays,  $e\mu$  pairs have similar kinematic behavior and cross-section as the  $\ell\ell$  (*ee* or  $\mu\mu$ ) decay states.

## 5.4.1 $e\mu$ Pair Selection

The 35.8 fb<sup>-1</sup> 2016 MuonEG dataset is used for the  $e\mu$  pair selection. Events with one or more  $e\mu$  pair are selected. If more than one  $e\mu$  pair is present in an event, the pair with invariant mass closest to Z boson mass is selected (analogous to the data selection). Electrons are required to pass Loose ID with pfIso, and muons are required to pass High  $p_T$  ID and tracker isolation.

When the  $e\mu$  sample is used as background in the electron channel the selections below are applied to match the requirements applied in data:

- Leading lepton  $p_T > 120 \text{ GeV}, |\eta| < 2.5,$
- Subleading lepton  $p_T > 35 \,\text{GeV}, |\eta| < 2.5$

When the sample is used in the muon channel the selections are:

- Leading lepton  $p_T > 55 \text{ GeV}, |\eta| < 2.4$
- Subleading lepton  $p_T > 20 \text{ GeV}, |\eta| < 2.4.$

### 5.4.2 $e\mu$ Pair Event-based Reweighting

Based on the assumption that electrons and muons have the same behavior in the non-resonant background, the distributions of kinematic variables for the electron and muon in the  $e\mu$  pair are expected to be nearly identical. However, this symmetry is altered by triggers and reconstruction effects. Figure 5.26 shows the  $p_T$  distribution of the electrons and muons, and their ratio in the MuonEG dataset.



Figure 5.26: electron and muon  $p_T$  distribution (left) and ratio (right) in the  $e\mu$  pair sample selected from MuonEG dataset.

The discrepancy between the  $p_T$  distributions of electrons and muons can result from many factors, such as detector effects, triggers, identification criteria and isolation efficiencies. Event based weighting factors are calculated according to the ratio plot in Figure 5.26(right). When modeling the electron channel, the correction factor of the event is set to be the ratio value corresponding to the muon  $p_T$  in the  $e\mu$  event. Conversely, when modeling the muon channel, the correction factor of the event is the inverse of the ratio value corresponding to the electron  $p_T$ .

Figure 5.27 shows the electron and muon  $p_T$  distribution and ratio after the correction for the electron channel, and Figure 5.28 is for the muon channel.

The agreement between the  $p_T$  distributions of electrons and muons are improved with the event based reweighting for both the electron channel and muon channel. In the electron channel, both leptons behave like the electron before weighting, and in the muon channel both leptons behave like the muon before weighting. This reweighting suppresses the systematic uncertainty caused by the performance discrepancy between electrons and muons due to detector and reconstruction effects. With the reweighting



Figure 5.27: electron and muon  $p_T$  distribution (left) and ratio (right) in the  $e\mu$  pair sample (from MuonEG dataset) after the event based reweighting for the electron channel



Figure 5.28: electron and muon  $p_T$  distribution (left) and ratio (right) in the  $e\mu$  pair sample (from MuonEG dataset) after the event based reweighting for the muon channel

there is still a 2% of disagreement between the  $p_T$  distributions of electrons and muons, and this is quoted as the systematic uncertainty of the method.

## 5.4.3 HLT Efficiency Reweighting

The effect of the HLT in our  $\ell\ell$  data selection is calculated and applied in the  $e\mu$  sample to simulate the single lepton HLT. With the event based reweighting described above, the electrons and muons in the  $e\mu$  pair sample behave very similarly and can be treated as  $\ell\ell$  pairs for the purpose of background studies. Both SingleElectron and SingleMuon trigger efficiency for data are calculated for each event in the  $e\mu$  pair data, corresponding to the leading lepton's  $p_T$  and  $\eta$ , regardless of whether the leading lepton is a muon or an electron. The trigger efficiency for data is described in Section 5.2.3. The SingleElectron trigger efficiency is applied when the sample is used in the electron channel, similarly the SingleMuon trigger efficiency is applied when the sample is used in the muon channel. The uncertainty of the trigger efficiency is evaluated and quoted as the standard deviation of the efficiencies among all the selected events for each channel (see Table 5.3).

When applying the SingleLepton trigger efficiency on the MuonEG dataset to emulate the data HLT, the MuonEG dataset is expected to have full acceptance in our preselection region. This explains why no HLT is required in the MuonEG data selection process. By not requiring any MuonEG HLT, events passing any MuonEG HLTs are accepted. Considering the high  $p_T$  threshold in the lepton selection, the acceptance of the dataset is high enough and this effect can be ignored compared to the single lepton trigger efficiency uncertainty. In fact, the acceptance of the MuonEG dataset has negligible effect as long as it is consistent in our preselection region, and the  $e\mu$  sample is rescaled as discussed below. To clarify, the  $p_T$  thresholds of the MuonEG dataset do not affect our event selection. For the muon channel, we require a leading lepton pt > 60 GeV, and subleading lepton pt > 20 GeV. The electron channel has higher lepton  $p_T$  selection. The MuonEG dataset contains non-prescaled triggers:

- HLT\_Mu23\_TrkIsoVVL\_Ele12\_CaloIdL\_TrackIdL\_IsoVL(\_DZ)
- HLT\_Mu12\_TrkIsoVVL\_Ele23\_CaloIdL\_TrackIdL\_IsoVL(\_DZ)

and these  $p_T$  thresholds are much lower than our  $p_T$  selection.

### 5.4.4 $e\mu$ Events Rescaling

The selected  $e\mu$  sample must be scaled to estimate the background in the selected data. To rescale the  $e\mu$  samples to fit the non-resonant background and study the agreement of the background sample and data, a non-resonant control region is defined. The scale factor is calculated by equation 5.2.

$$Scale = \frac{N_{data,ll} - N_{MC,reson}}{N_{data,e\mu}}$$
(5.2)

The  $N_{data,ll}$  is the number of data events in the control region;  $N_{MC,reson}$  is the number of background Z events in the control region, including Z+jets background and resonant background; the  $N_{data,e\mu}$  is the number of  $e\mu$  pair events in the control region. To define the control region, the main idea is to suppress the standard deviation of the calculated scale factor. The standard deviation is calculated based on equation 5.2, and given in equation 5.3.

$$\sigma_{Scale} = \sqrt{\frac{\sigma_{N_{data,ll}}^2 + \sigma_{N_{MC,reson}}^2}{(N_{data,ll} - N_{MC,reson})^2} + \frac{\sigma_{N_{data,e\mu}}^2}{N_{data,e\mu}^2}} \times \frac{N_{data,ll} - N_{MC,reson}}{N_{data,e\mu}}}{N_{data,e\mu}}$$
(5.3)

Here the statistical uncertainty is considered for both MC and data, and the PDF uncertainty is considered for resonant MC (1.5%).

To suppress scale deviation and resonant background in the CR, selections below are applied:

- 1. Z mass veto: invariant mass M < 70 GeV or M > 110 GeV;
- 2.  $p_T^Z > X \text{ GeV};$

To determine the X value ( $p_T^Z$  cut level), a scan over various X values is performed. The calculated scale factor and uncertainty vs X is shown in Table 5.2. X=60 gives relatively small uncertainty, and is used to define the control region. The scale uncertainty corresponding to X=60 is also counted as the systematic uncertainty.

Х	electron channel scale	muon channel scale
0	$0.3442 {\pm} 0.0140$	0.8215±0.0519
10	$0.3442{\pm}0.0140$	$0.7762{\pm}0.0388$
20	$0.3442{\pm}0.0140$	$0.7364{\pm}0.0290$
30	$0.3442{\pm}0.0140$	$0.7046 {\pm} 0.0238$
40	$0.3442{\pm}0.0140$	$0.6962 {\pm} 0.0202$
50	$0.3443 {\pm} 0.0140$	$0.6895 {\pm} 0.0182$
60	$0.3463 {\pm} 0.0140$	$0.6871 {\pm} 0.0171$
70	$0.3490 {\pm} 0.0142$	$0.6907 {\pm} 0.0167$
80	$0.3560 {\pm} 0.0146$	$0.6952{\pm}0.0170$
90	$0.3635 {\pm} 0.0156$	$0.7092{\pm}0.0182$
100	$0.3733 {\pm} 0.0169$	$0.7249{\pm}0.0205$
110	$0.3771 {\pm} 0.0186$	$0.7402{\pm}0.0234$

Table 5.2: scale and deviation vs  $X(p_T^Z \text{ cut value})$  scan result.

Electron channel plots in the control region are shown in Figure 5.29 and those for the muon channel are shown in the Figure 5.30.

The control region plots show that the resonant background is heavily suppressed in the CR and data driven non-resonant background agrees well with the data. Also non-resonant  $e\mu$  events should have similar cross section compared to those in  $\ell\ell$  events, which means that the sum of the scale factors in the electron channel and the muon channel should be close to 1. From Table 5.2 we see that at X=60, the scale factor is 0.346 for the electron channel, and 0.687 for the muon channel, adding up to 1.033, which agrees with the prediction.

### 5.4.5 Comparing to MC Modeling

Figure 5.31 is a comparison of the non-resonant background modeled by the data-driven method and the  $t\bar{t} + WW$  MC samples as a cross-check of the data-driven method, with 70 GeV  $< M_Z < 110$  GeV,  $p_T^{\text{miss}} > 50$  GeV



Figure 5.29: Data-driven non-resonant background in the control region, electron channel plots



Figure 5.30: Data-driven non-resonant background in the control region, muon channel plots

and  $p_T^Z > 100 \text{ GeV}$  (standard signal region selections in this analysis), for electron channel. Similarly Figure 5.32 shows the result for the muon channel. The data-driven modeling method generally gives a very similar non-resonant background distribution as the MC. However, in the electron channel the yield of the data-driven method is slightly higher than the MC modeling.

The yield discrepancy is evaluated and half of the discrepancy value is quoted as a systematic uncertainty (6.7% for electron channel).

### 5.4.6 Uncertainty Table

In addition to the systematic uncertainty of the data-driven method discussed above, differences in acceptance are considered. All muons have  $|\eta| < 2.4$ , while in the electron channel, each lepton from real non-resonant background can also have  $|\eta|$  between 2.4 and 2.5, which differs from the  $e\mu$  pair data-driven background. Fortunately, less than 0.1% events with  $p_T^Z > 100 \text{ GeV}, p_T^{\text{miss}} > 50 \text{ GeV}$  in the electron channel have an electron with  $|\eta|$  between 2.4 and 2.5, which means that the effect can be neglected.

Table 5.3 summarizes the uncertainties that have been evaluated for this data-driven non-resonant background modeling method.

Uncertainty	electron channel	muon channe
$e\mu$ pair reweighting	2%	2%
Trigger efficiency	6.0%	1.3%
Stat. uncert. of Data and MC, PDF/QCD uncert. of subtracted MCs	4.0%	2.4%
Data-driven vs. MC disagreement	6.7%	0.1%
Total	10.0%	3.4%

Table 5.3: Data-driven non-resonant modeling method uncertainties



Figure 5.31: Data-driven non-resonant background and MC non-resonant background comparison, with Z mass selection and  $p_T^Z > 100 \text{ GeV}, p_T^{\text{miss}} > 50 \text{ GeV}$ , electron channel plots



Figure 5.32: Data-driven non-resonant background and MC non-resonant background comparison, with Z mass selection and  $p_T^Z > 100 \text{ GeV}, p_T^{\text{miss}} > 50 \text{ GeV}$ ,muon channel plots

## 5.5 Resonant Background Modeling

The resonant background in this analysis contains mainly SM  $qq \rightarrow ZZ \rightarrow 2\ell 2\nu$  process, as well as WZ processes and ZZ processes with  $\ell\ell qq$  or  $4\ell$  final states. This component of the background is modeled by MC samples. For the  $qq \rightarrow ZZ \rightarrow 2\ell 2\nu$  sample (ZZTo2L2Nu\_13TeV\_powheg\_pythia8), NNLO QCD [65] and NLO EW corrections [66, 67] are applied.

The NNLO/NLO QCD correction is parametrized and applied as a function of  $m_{ZZ}$  at generator level. The correction and the uncertainty band are shown on Figure 5.33. The average NNLO/NLO QCD correction k-factor is 1.11 with an uncertainty of 3%. For generator level  $m_{ZZ} > 500$  GeV, the average k-factor and uncertainty are applied.



Figure 5.33: The NNLO/NLO QCD correction and error band as a function of generator level  $m_{ZZ}$  for SM qqZZ process.

The NLO/LO EW correction is parametrized as a function of initial state quark flavors and event kinematic variables  $\hat{s}$  and  $\hat{t}$  in the center of mass frame at generator level. The variable  $\hat{s}$  is the partonic center of mass

energy, corresponding to  $m_{ZZ}$ , and  $\hat{t}$  is computed as

$$\begin{aligned} \hat{t} &= \left(p_{q_1}^* - p_{Z_1}^*\right)^2 = p_{q_1}^{*2} + p_{Z_1}^{*2} - 2p_{q_1}^* \cdot p_{Z_1}^* \\ &\simeq 0 + m_Z^2 - 2\left(\frac{\hat{s}}{4} - \frac{\sqrt{\hat{s}}}{2}\cos\theta\sqrt{\frac{\hat{s}}{4} - m_Z^2}\right) \\ &= m_Z^2 - \frac{\hat{s}}{2} + \cos\theta\sqrt{\frac{\hat{s}^2}{4} - m_Z^2}\hat{s}, \end{aligned}$$

where  $p_{q_1}^*$  is the four momentum of either one of the quarks initiating the hard process and  $p_{Z_1}^*$  is the four momentum of a Z boson, and quark masses are neglected. The angle  $\theta$  is the angle between a Z boson and the direction of the incident quarks in the center-of-mass frame of the two Z bosons, and it is approximately computed as

$$\cos\theta = \frac{\hat{\vec{p}}_{q_1b} - \hat{\vec{p}}_{q_2b}}{\left| \left( \hat{\vec{p}}_{q_1b} - \hat{\vec{p}}_{q_2b} \right) \right|} \cdot \hat{\vec{p}}_{Z_1b},$$

where  $\hat{\vec{p}}_{q_i/Z_i b}$  represents the unitary direction vector of the *i*th quark/Z boson after the Lorentz boost.

The average k-factor for NLO/LO EW correction is 0.95 with an uncertainty of 3 %. The correction function and the uncertainty band is shown on Figure 5.34 as a function of generator level  $m_{ZZ}$ . The NLO/LO EW correction is only appropriate for on-shell Z bosons with  $m_{ZZ} > 2m_Z$ .



Figure 5.34: The NLO/LO EW correction and error band as a function of generator level  $m_{ZZ}$  for SM qqZZ process.

# 5.6 Preselection Plots

Figures 5.36 to 5.45 show the data vs background distributions in the preselection region, with all the backgrounds modeled using the methods discussed above.



Figure 5.35: Vertex multiplicity distributions for electron (left) and muon (right) channels comparing data and background.



Figure 5.36:  $\rho$  distributions for electron (left) and muon (right) channels comparing data and background.



Figure 5.37:  $m_T$  distributions for electron (left) and muon (right) channels comparing data and background, wide mass window.



Figure 5.38:  $m_T$  distributions for electron (left) and muon (right) channels comparing data and background, narrow mass window.



Figure 5.39: Z mass  $(m_{\ell\ell})$  distributions for electron (left) and muon (right) channels comparing data and background.



Figure 5.40:  $p_T^Z$  distributions for electron (left) and muon (right) channels comparing data and background, wide binning.



Figure 5.41:  $p_T^Z$  distributions for electron (left) and muon (right) channels comparing data and background, narrow binning.



Figure 5.42:  $p_T^{\text{miss}}$  distributions for electron (left) and muon (right) channels comparing data and background.



Figure 5.43:  $p_T^{\text{miss}}$  distributions for electron (left) and muon (right) channels comparing data and background.



Figure 5.44:  $p_T_{\parallel}^{miss}$  distributions for electron (left) and muon (right) channels comparing data and background.



Figure 5.45:  $p_{T_{\perp}}^{miss}$  distributions for electron (left) and muon (right) channels comparing data and background.
# 5.7 Plots in the Signal Region

Figures 5.46 to 5.56 show the data vs background distributions in the signal region, with all the backgrounds modeled using the methods discussed above.



Figure 5.46:  $\rho$  distributions for electron (left) and muon (right) channels comparing data and background, in SR.



Figure 5.47: Vertex multiplicity distributions for electron (left) and muon (right) channels comparing data and background, in SR.



Figure 5.48:  $m_T$  distributions for electron (left) and muon (right) channels comparing data and background, wide mass window, in SR



Figure 5.49:  $m_T$  distributions for electron (left) and muon (right) channels comparing data and background, narrow mass window, in SR



Figure 5.50: Z mass distributions for electron (left) and muon (right) channels comparing data and background, in SR.



Figure 5.51:  $p_T^Z$  distributions for electron (left) and muon (right) channels comparing data and background, wide binning, in SR.



Figure 5.52:  $p_T^Z$  distributions for electron (left) and muon (right) channels comparing data and background, narrow binning, in SR.



Figure 5.53:  $p_T^{\text{miss}}$  distributions for electron (left) and muon (right) channels comparing data and background, in SR.



Figure 5.54:  $p_T^{\text{miss}}$  distributions for electron (left) and muon (right) channels comparing data and background, in SR.



Figure 5.55:  $p_{T_{\parallel}}^{miss}$  distributions for electron (left) and muon (right) channels comparing data and background, in SR.



Figure 5.56:  $p_{T_{\perp}}^{miss}$  distributions for electron (left) and muon (right) channels comparing data and background, in SR.

# **Chapter 6**

# Systematic Uncertainties and Interpretation of Results

In this chapter, systematic uncertainties in this analysis are discussed. The signal cross section limits are evaluated based on the  $m_T$  distributions for data and background, and the corresponding uncertainty propagated from individual sources of systematic uncertainties.

## 6.1 Systematic Uncertainty

The sources of systematic uncertainty considered include those in the background modeling, the signal acceptance, and the integrated luminosity.

### 6.1.1 Integrated Luminosity Uncertainty

The uncertainty on the integrated luminosity measurement is evaluated by the CMS Lumi POG and the value of 2.5% [68] is recommended for all the CMS analyses using 2016 data corresponding to 35.9 fb<sup>-1</sup> of integrated luminosity of proton-proton collisions with a center-of-mass energy of 13 TeV. It is applied to all the signal samples and backgrounds in this analysis, because MC samples are involved in all the background modeling methods.

### 6.1.2 PDF and QCD Scale Uncertainty

Uncertainties arising from the PDF model and renormalization and factorization scales in fixed-order calculations affect MC signals and backgrounds modeled by MC samples, in terms of cross sections and acceptance. The PDF uncertainty effect is estimated by evaluating all the error sets of NNPDF 3.0 PDF (in the form of PDF uncertainty variation weights stored in the MiniAOD files), following the PDF4LHC prescription [51]. This contributes a variation of 1.0–3.4% to the MC background models. As for the simulated signal, the PDF uncertainties on the production cross section can vary between 10-50% depending on the signal mass, but the effect on the signal acceptance is on the scale of 1% and is negligible.

The effect of scale variations is assessed by varying the original factorization and renormalization scales by factors of 0.5 or 2.0. The scale uncertainties are estimated to be about 3–3.5% each in the production cross section and acceptance for the resonant background. For the Z+jets background, the scale choice modifies the normalization by 3.5%. The acceptance varies by 23 (13)% in the electron (muon) channel and the corresponding effect is negligibly small for the signal. An uncertainty of 3.0% is estimated for the (N)NLO correction to the resonant background.

### 6.1.3 Trigger/Lepton Uncertainty

The uncertainties of the HLT efficiency and lepton ID/Iso efficiencies are mainly due to statistics and the *background* + *signal* fitting in the tagand-probe method, as well as a 1% addition to account for muon tracker inefficiency. These affect the signal and background estimates obtained from both simulation and from control samples in data. The combined effect of these uncertainties on the normalizations of the various samples is found to be 0.4-3.6%.

### 6.1.4 Missing Transverse Energy Uncertainty

Given that  $p_T^{\text{miss}}$  is actually a 2-D vector, the uncertainties on both  $|\vec{p_T}^{\text{miss}}|$ and  $\phi(\vec{p_T}^{\text{miss}})$  are evaluated. Considering that all physics objects are involved in the reconstruction of  $p_T^{\text{miss}}$ , the contributions to the  $p_T^{\text{miss}}$  uncertainties includes:

- Jet Energy Scale
- Jet Energy Resolution
- Muon Energy Scale/Resolution
- Electron Energy Scale/Resolution
- Photon Energy Scale/Resolution
- Tau Energy Scale/Resolution
- Unclustered Energy Scale/Resolution

These contributions are evaluated separately for both  $|\vec{p_T}^{miss}|$  and  $\phi(\vec{p_T}^{miss})$ .

# 6.1.5 Uncertainties in the Background Modeling methods

In addition to the general uncertainties discussed above, uncertainties affecting each background modeling method are also evaluated, including:

- $p_T^{\gamma}$  to  $p_T^Z$  reweighting uncertainty shown in Figure 5.19 and 5.20, for Z+jets background modeling;
- Hadronic recoil uncertainty propagated from the uncertainties of the hadronic recoil Gaussian fittings, for Z+jets background modeling;
- Non-resonant background uncertainties summarized in Table 5.3;
- QCD and EW correction uncertainties shown in Figures 5.33 and 5.34.

### 6.1.6 Systematic Uncertainty Summary

The effect of systematic uncertainties on the yields are summarized in Table 6.1 for muon and electron channels respectively. The uncertainties on the acceptance are evaluated in the Signal Region. In the table the uncertainty estimation for the signal is evaluated with the narrow width Bulk Graviton samples with a mass of 1 TeV generated by Madgraph.

Table 6.1: Summary of the uncertainties. "-" denotes uncertainties that of	do
not apply and "(-)" stands for uncertainties with negligible value.	

	Source	Signal	Z+jets	Resonant	Non-Reso
	Luminosity	2.5%	2.5%	2.5%	2.5%
	PDF on cross-section	-	2.3%	1.7%	-
	QCD on cross-section	-	3.5%	3.0%	-
	QCD & EW corrections	-	-	3.0%	-
Electron	PDF on acceptance	1.0%	3.4%	1.0%	-
channel	QCD on acceptance	(-)	22.7%	2.9%	-
	Trigger eff.	1.0%	-	0.1%	-
	Lepton ID eff.	1.9%	-	0.4%	-
	$\mathbf{Z} p_T$ reweighting	-	6.8%	-	-
	Non-reso. scale fact.	-	-	-	10.0%
	$p_T^{\text{miss}}$ lepton/photon $p_T$	(-)	-	4.6%	-
	$p_T^{\text{miss}}$ Jet energy resolution	(-)	-	6.8%	-
	$p_T^{\text{miss}}$ unclustered-energy	(-)	-	5.5%	-
	$p_T^{\text{miss}}$ Hadronic recoil	-	3.4%	-	-
Muon	PDF on acceptance	1.0%	3.4%	1.0%	-
channel	QCD on acceptance	(-)	13.1%	2.9%	-
	Trigger eff.	3.3%	-	0.1%	-
	Lepton ID eff.	1.1%	-	0.2%	-
	Tracking eff.	1.0%	-	1.0%	-
	$\mathbf{Z} p_T$ reweighting	-	3.2%	-	-
	Non-reso. scale fact.	-	-	-	2.4%
	$p_T^{\text{miss}}$ lepton/Photon $p_T$	(-)	-	7.4%	-
	$p_T^{\text{miss}}$ Jet energy resolution	(-)	-	5.6%	-
	$p_T^{\text{miss}}$ Unclustered-Energy	(-)	-	6.3%	-
	$p_T^{\text{miss}}$ Hadronic recoil	-	2.0%	-	-

The distributions of  $p_T^Z$ ,  $p_T^{\text{miss}}$  in the SR with systematic uncertainty are shown in Figs 6.1, 6.2, 6.3, for electron and muon channels, to illustrate the overall systematic and statistic uncertainties in this analysis. The shaded band shows the systematic uncertainties in background, while the statistical uncertainty in the data is shown by the error bars.



Figure 6.1:  $p_T^Z$  for electron (left) and muon (right) channels comparing the data and background. The expected distribution for a zero width bulk graviton resonance with a mass of 1 TeV is also shown for a value of 1 pb for the product of cross section and branching fraction  $\sigma(pp \to X \to ZZ)B(ZZ \to 2\ell 2\mu)$ . The lower panels show the ratio of data to the prediction for the background.



Figure 6.2:  $p_T^{\text{miss}}$  for electron (left) and muon (right) channels comparing the data and background. The expected distribution for a zero width bulk graviton resonance with a mass of 1 TeV is also shown for a value of 1 pb for the product of cross section and branching fraction  $\sigma(pp \to X \to ZZ)B(ZZ \to 2\ell 2\mu)$ . The lower panels show the ratio of data to the prediction for the background.

## 6.2 **Results and Interpretation**

The  $m_T$  distribution is used as the discriminant variable to search for a new resonance decaying to ZZ with the subsequent decay  $ZZ \rightarrow 2\ell 2\nu$ . A binned likelihood fit of the predicted backgrounds to data, combining electron and muon channels, and including the estimated systematic uncertainties is performed and summarized in Table 6.2. Figure 6.3 shows the post-fit  $m_T$ distribution in the signal region using only the background models. The expected distribution for a bulk graviton signal with a mass of 1 TeV and an arbitrary product of cross section and branching fraction  $\sigma(pp \rightarrow X \rightarrow$  $ZZ) \mathcal{B}(ZZ \rightarrow 2\ell 2\nu)$  of 1 pb is also shown. The observed distributions are in agreement with fitted SM background predictions.

	Electron channel	Muon channel
Data	9336	52806
Z+jets Resonant Nonresonant	$8421\pm203$ $637\pm38$ $271\pm28$	$44253 \pm 336$ $2599 \pm 164$ $5961 \pm 211$
Total background	$9329\pm208$	$52813 \pm 439$

Table 6.2: Event yields for different background contributions and those observed in data in the electron and muon channels.

The asymptotic approximation [69] of the semi-frequentist  $CL_s$  technique [70, 71, 72] is applied to determine the expected and observed upper limits for the possible signal strength at 95% confidence level. The same simultaneous combined fit is performed using signal and background distributions, to extract the upper limits for a given signal hypothesis. Statistical uncertainties in the background modeling are taken into account by fluctuating the predicted background histograms within an envelope according to uncertainties in each bin. Systematic uncertainties are treated as nuisance parameters, constrained with Gaussian or log-normal probability density functions in the maximum likelihood fit. For the signal, only uncertainties related to luminosity and acceptance contribute in the limit setting procedure. When the likelihoods for electron and muon channels are combined,



Figure 6.3: The  $m_T$  distributions for electron (left) and muon (right) channels comparing the data and background, after fitting the background-only model to the data. The expected distribution for a zero width bulk graviton resonance with a mass of 1 TeV is also shown for a value of 1 pb for the product of branching fraction and cross section  $\sigma(pp \to X \to ZZ)B(ZZ \to 2\ell 2\mu)$ . The lower panels show the ratio of data to the prediction for the background.

the correlation of systematic effects is taken into account.

The HiggsCombine Tool [73] is used to obtain the limits. A binned shape analysis is applied, and the shape effects on the  $m_T$  distributions from the following sources of systematic uncertainties are evaluated.

- HLT efficiency
- lepton ID/Iso efficiency
- QCD & EW corrections for the resonance background
- Z  $p_T$  reweighting for the Z+jets background
- $p_T^{\text{miss}}$  lepton/photon  $p_T$
- $p_T^{\text{miss}}$  Jet energy resolution
- $p_T^{\text{miss}}$  unclustered-energy

•  $p_T^{\text{miss}}$  Hadronic recoil

Figure 6.4 shows the expected and observed upper limits on the product of the production cross section and the branching fraction for  $X \to ZZ$ determined at the 95% confidence level for the zero width benchmark model, electron and muon channels combined, shown with respect to the Graviton mass. Theoretical expectations of  $\sigma(pp \to X \to ZZ)$  are also shown as a function of the Graviton mass, separately for  $\tilde{k} = 1.0, 0.5, 0.1$ . The observed limits are within 2 standard deviations of expectations from the background-only model, therefore no evidence of new particle is observed. The hypothesis of  $\tilde{k} = 0.5$  can be excluded for masses below 800 GeV at 95% CL, while the current data are not yet sensitive to the hypothesis of  $\tilde{k} = 0.1$ .

Figure 6.5 shows the limit plots for the electron and muon channels separately. The two channels have similar sensitivity, and the combined fit gives better expected results.

The analysis of the upper limits is also performed on the more general wide width version of the bulk graviton model, where the initial state for production is set to be either a gluon–gluon fusion or  $q\bar{q}$  annihilation process. The width of the resonance is set to be 0/10%/20%/30% of the resonance mass. Figure 6.6 shows the limits for the wide width model.

Considering that gluon–gluon fusion is the dominant production for the Bulk Graviton, the zero-width gluon–gluon fusion limit plot looks very similar to Figure 6.4. The  $q\bar{q}$  production limits differ due to the spin and parity effects [74]. Figure 6.7 shows the normalized  $m_T$  distribution of the gluon–gluon production and  $q\bar{q}$  production, as well as the combined production based on the cross-sections of each productions. The  $m_T$  spectrum of the  $q\bar{q}$  production is much less strongly peaked than gluon–gluon fusion, which reduces the sensitivity of the expected and observed cross-section upper limits.



Figure 6.4: Expected and observed limits on the product of cross section and branching fraction of the  $pp \rightarrow G_{bulk} \rightarrow ZZ$  process, with zero-width assumption.



Figure 6.5: Expected and observed limits on the product of cross section and branching fraction of the  $pp \rightarrow G_{bulk} \rightarrow ZZ$  process, with zero-width assumption, for the electron channel (left) and muon channel (right).



Figure 6.6: Expected and observed limits on the product of cross section and branching fraction of the  $pp \rightarrow G_{bulk} \rightarrow ZZ$  process, electron channel and muon channel combined, with signal width being 0/10%/20%/30% of the resonance mass, for gluon–gluon fusion production (left) and  $q\bar{q}$ annihilation production (right).



Figure 6.7: The normalized  $m_T$  distributions for spin-2 resonance signal samples for gluon fusion production and  $q\bar{q}$  production. The combined production has taken the cross-sections of each productions into consideration.

# Chapter 7

# **Summary and Future Prospects**

### 7.1 Analysis Summary

A search for the RS Bulk Graviton has been performed in events with a leptonically decaying Z boson and missing transverse momentum, using data corresponding to an integrated luminosity of 35.9 fb<sup>-1</sup> of protonproton collisions at a center-of-mass energy of 13 TeV, collected by the CMS experiment in 2016. The hypothesis is examined for the range of graviton mass between 600 and 2500 GeV. The observed data are consistent with expectations from standard model processes. At 95% confidence level, the region of  $m_G < 800$  GeV is excluded for the Bulk Graviton model with  $\tilde{k} < 0.5$  and zero-width assumption. The analysis is repeated considering variations of the bulk graviton model to include a large, mass-dependent width. Exclusion limits are provided separately for gluon-gluon fusion and qq annihilation production processes.

### 7.2 Comparison and Future Prospects

The  $G_{bulk} \rightarrow ZZ \rightarrow 2\ell 2\nu$  search has some advantage compared to other diboson final state searches. The background of the  $2\ell 2\nu$  channel is more controllable compared to (semi-)hadronic channels, because the large  $p_T^{\text{miss}}$ can help suppress Z+jets background, which is the dominant background in these searches. The  $2\ell 2\nu$  channel also has an advantage in signal statistics over the  $4\ell$  final state, because its cross-section is about 6 times as high as that of the  $4\ell$  channel.

Similar diboson searches for the RS Bulk Graviton with various final states based on 2016 CMS full dataset are listed below and compared with this analysis:

- $G_{bulk} \rightarrow ZZ \rightarrow 2q2\nu$  [75]
- $G_{bulk} \rightarrow ZZ \rightarrow 4q$  [76]
- $G_{bulk} \rightarrow ZZ \rightarrow 2q2\ell$  [77]
- $G_{bulk} \rightarrow WW \rightarrow 4q$  [76]
- $G_{bulk} \rightarrow WW \rightarrow 2ql\nu$  [78]
- $G_{bulk} \rightarrow HH \rightarrow 4b$  [79]
- $G_{bulk} \rightarrow HH \rightarrow 2b2\tau$  [80]

The expected and observed upper limits on the narrow resonance cross section at the 95% confidence level for these searches are shown in Figure 7.1. The good sensitivity of this  $G_{bulk} \rightarrow ZZ \rightarrow 2\ell 2\nu$  search in the  $m_T < 1000$  GeV region is the result of the controllable backgrounds and in particular the techniques described here to constrain the Z+jets background. In this low mass region the  $G_{bulk} \rightarrow ZZ \rightarrow 2q2\ell$  analysis shows comparable upper limits, but its sensitivity is degraded at higher masses [77].

However the  $G_{bulk} \rightarrow ZZ \rightarrow 2\ell 2\nu$  channel suffers low statistics in the highest accessible mass region. With the integrated luminosity of 35.9 fb<sup>-1</sup>, very few events are observed for either signal region data selection or background modeling for  $m_T > 1000$  GeV. Therefore, this channel is less capable comparing to hadronic channels for exploring the region  $m_G > 2000$  GeV with the 2016 dataset. Nevertheless, considering that for now statistics is the bottleneck of this analysis, with the 2017 and 2018 CMS data becoming available with data corresponding to an integrated luminosity 3 times as high, a significant improvement in the result is expected from the



Figure 7.1: Expected and observed upper limits on the resonance cross section at the 95% confidence level for various diboson searches based on 2016 CMS full dataset.

 $G_{bulk} \rightarrow ZZ \rightarrow 2\ell 2\nu$  channel. With larger data samples being collected, sensitivity to these physics signatures will continue to grow. Each analysis improves our understanding of the standard model processes and advances our measurement techniques. Although solutions to the open questions of the standard model remain elusive, our explorations for new physics will build on present measurements to illuminate ever more rare and challenging processes for our next clues to the path forward.

# Bibliography

- [1] CMS Collaboration. "Search for heavy resonances decaying to a pair of Higgs bosons in the four b quark final state in proton-proton collisions at  $\sqrt{s} = 13$  TeV". Journal of High Energy Physics, Mar 2018.
- [2] K Nakamura and Particle Data Group. "Review of Particle Physics". *Journal of Physics G: Nuclear and Particle Physics*, 37(7A):075021, 2010.
- [3] Steven Weinberg. "A Model of Leptons". *Phys. Rev. Lett.*, 19:1264– 1266, Nov 1967.
- [4] Sheldon L. Glashow. "Partial-symmetries of weak interactions". *Nuclear Physics*, 22(4):579–588, 1961.
- [5] Y. Fukuda et al. "Evidence for Oscillation of Atmospheric Neutrinos". *Phys. Rev. Lett.*, 81:1562–1567, Aug 1998.
- [6] F. P. An et al. "Observation of Electron-Antineutrino Disappearance at Daya Bay". *Phys. Rev. Lett.*, 108:171803, Apr 2012.
- [7] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". *Phys. Rev. Lett.*, 13:321–323, Aug 1964.
- [8] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". *Phys. Rev. Lett.*, 13:508–509, Oct 1964.
- [9] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". *Phys. Rev. Lett.*, 13:585–587, Nov 1964.

- [10] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". *Physics Letters B*, 716(1):30 61, 2012.
- [11] ATLAS. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". *Physics Letters B*, 716(1):1 – 29, 2012.
- [12] Savas Dimopoulos Nima Arkani–Hamed and Gia Dvali. "The hierarchy problem and new dimensions at a millimeter". *Physics Letters B*, 429(3):263 – 272, 1998.
- [13] Nima Arkani–Hamed, Savas Dimopoulos, and Gia Dvali. "The hierarchy problem and new dimensions at a millimeter". *Physics Letters B*, 429(3):263 – 272, 1998.
- [14] Nima Arkani-Hamed, Savas Dimopoulos, and Gia Dvali. "Phenomenology, astrophysics, and cosmology of theories with submillimeter dimensions and TeV scale quantum gravity". *Phys. Rev. D*, 59:086004, Mar 1999.
- [15] Lisa Randall and Raman Sundrum. "Large Mass Hierarchy from a Small Extra Dimension". *Phys. Rev. Lett.*, 83:3370–3373, Oct 1999.
- [16] Lisa Randall and Matthew D Schwartz. "Quantum field theory and unification in AdS5". *Journal of High Energy Physics*, 2001(11):003– 003, nov 2001.
- [17] S. Casagrande, F. Goertz, U. Haisch, M. Neubert, and T. Pfoh. "Flavor physics in the Randall-Sundrum model I. Theoretical setup and electroweak precision tests". *Journal of High Energy Physics*, 2008(10):094, 2008.
- [18] Kaustubh Agashe, Gilad Perez, and Amarjit Soni. "Flavor structure of warped extra dimension models". *Phys. Rev. D*, 71:016002, Jan 2005.
- [19] H. Davoudiasl, J.L. Hewett, and T.G. Rizzo. "Bulk gauge fields in the Randall–Sundrum model". *Physics Letters B*, 473(1):43 – 49, 2000.

- [20] Sanghyeon Chang, Junji Hisano, Hiroaki Nakano, Nobuchika Okada, and Masahiro Yamaguchi. "Bulk standard model in the Randall-Sundrum background". *Phys. Rev. D*, 62:084025, Sep 2000.
- [21] Tony Gherghetta and Alex Pomarol. "Bulk fields and supersymmetry in a slice of AdS". *Nuclear Physics B*, 586(1):141 162, 2000.
- [22] Alexandra Oliveira. "Gravity particles from Warped Extra Dimensions, predictions for LHC". 2014.
- [23] Georges Aad et al. "Search for new phenomena in the WW to  $\ell \nu \ell' \nu'$  final state in pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector". *Phys. Lett. B*, 718:860, 2013.
- [24] Georges Aad et al. "Search for resonant diboson production in the  $\ell \nu j j$  decay channels with the ATLAS detector at 7 TeV". *Phys. Rev. D*, 87:112006, 2013.
- [25] Georges Aad et al. "Search for resonant diboson production in the  $\ell \ell q \bar{q}$ final state in pp collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector". *Eur. Phys. J. C*, 75:69, 2015.
- [26] Serguei Chatrchyan et al. "Search for a narrow spin-2 resonance decaying to a pair of Z vector bosons in the semileptonic final state". *Phys. Lett. B*, 718:1208, 2013.
- [27] Vardan Khachatryan et al. "Search for massive resonances decaying into pairs of boosted bosons in semi-leptonic final states at  $\sqrt{s} = 8$ TeV". *JHEP*, 08:174, 2014.
- [28] Morad Aaboud et al. "Searches for heavy diboson resonances in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector". *JHEP*, 09:173, 2016.
- [29] Lyndon Evans and Philip Bryant. "LHC Machine". *Journal of Instrumentation*, 3(08):S08001, 2008.

- [30] Michael Benedikt, Paul Collier, V Mertens, John Poole, and Karlheinz Schindl. "LHC Design Report". CERN Yellow Reports: Monographs. CERN, Geneva, 2004.
- [31] CMS Collaboration. "The CMS experiment at the CERN LHC". *Journal of Instrumentation*, 3(08):S08004, 2008.
- [32] CMS Collaboration. "Precise mapping of the magnetic field in the CMS barrel yoke using cosmic rays". *Journal of Instrumentation*, 5(03):T03021, 2010.
- [33] CMS Collaboration. "The CMS tracker system project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997.
- [34] CMS Collaboration. "The CMS electromagnetic calorimeter project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997.
- [35] CMS Collaboration. "The CMS hadron calorimeter project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997.
- [36] CMS Collaboration. "The CMS muon project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997.
- [37] CMS Collaboration. "The CMS trigger system". *Journal of Instrumentation*, 12(01):P01020, 2017.
- [38] CMS Collaboration. "Particle-flow reconstruction and global event description with the CMS detector". *Journal of Instrumentation*, 12(10):P10003, 2017.
- [39] Wolfgang Adam, Boris Mangano, Thomas Speer, and Teddy Todorov. "Track Reconstruction in the CMS tracker". Technical Report CMS-NOTE-2006-041, CERN, Geneva, Dec 2006.

- [40] The CMS Collaboration. "Description and performance of track and primary-vertex reconstruction with the CMS tracker". *Journal of Instrumentation*, 9(10):P10009, 2014.
- [41] Thomas Speer, Kirill Prokofiev, R Frühwirth, Wolfgang Waltenberger, and Pascal Vanlaer. "Vertex Fitting in the CMS Tracker". Technical Report CMS-NOTE-2006-032, CERN, Geneva, Feb 2006.
- [42] E Chabanat and N Estre. "Deterministic Annealing for Vertex Finding at CMS". 2005.
- [43] W Adam, R Frühwirth, A Strandlie, and T Todorov. "Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC". *Journal of Physics G: Nuclear and Particle Physics*, 31(9):N9, 2005.
- [44] S. Baffioni et al. "Electron reconstruction in CMS". *The European Physical Journal C*, 49(4):1099–1116, Mar 2007.
- [45] The CMS collaboration. "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at  $\sqrt{s} = 13$  TeV". *Journal of Instrumentation*, 13(06):P06015–P06015, jun 2018.
- [46] Serguei Chatrchyan et al. Performance of CMS muon reconstruction in *pp* collision events at  $\sqrt{s} = 7$  TeV. *JINST*, 7:P10002, 2012.
- [47] "Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector". Technical Report CMS-PAS-JME-16-004, CERN, Geneva, 2016.
- [48] CMS Collaboration. "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV". *Journal of Instrumentation*, 12(02):P02014, 2017.
- [49] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-k<sub>t</sub> jet clustering algorithm". *Journal of High Energy Physics*, 2008(04):063, 2008.

- [50] NNPDF Collaboration. "Parton distributions for the LHC run II". Journal of High Energy Physics, 2015(4):40, Apr 2015.
- [51] Stefano Carrazza et al Jon Butterworth. "PDF4LHC recommendations for LHC Run II". *Journal of Physics G: Nuclear and Particle Physics*, 43(2):023001, 2016.
- [52] Ask Stefan et al Sjöstrand Torbjörn.
- [53] J. Allison et al S. Agostinelli. "Geant4—a simulation toolkit". Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 506(3):250 – 303, 2003.
- [54] A Rizzi G Petrucciani and C Vuosalo. "Mini-AOD: A New Analysis Data Format for CMS". *Journal of Physics: Conference Series*, 664(7):072052, 2015.
- [55] Sirunyan A. M et al Khachatryan V. "Event generator tunes obtained from underlying event and multiparton scattering measurements". *The European Physical Journal C*, 76(3):155, Mar 2016.
- [56] Frederix R et al Alwall J. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". *Journal of High Energy Physics*, 2014(7):79, Jul 2014.
- [57] Rikkert Frederix and Stefano Frixione. "Merging meets matching in MC@NLO". *Journal of High Energy Physics*, 2012(12):61, Dec 2012.
- [58] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with parton shower simulations: the POWHEG method". *Journal of High Energy Physics*, 2007(11):070, 2007.
- [59] Yanyan Gao, Andrei V. Gritsan, Zijin Guo, Kirill Melnikov, Markus Schulze, and Nhan V. Tran. "Spin determination of single-produced resonances at hadron colliders". *Phys. Rev. D*, 81:075022, Apr 2010.

- [60] Sara Bolognesi, Yanyan Gao, Andrei V. Gritsan, Kirill Melnikov, Markus Schulze, Nhan V. Tran, and Andrew Whitbeck. "Spin and parity of a single-produced resonance at the LHC". *Phys. Rev. D*, 86:095031, Nov 2012.
- [61] "Constraining anomalous HVV interactions at proton and lepton colliders, author = Anderson, Ian and Bolognesi, Sara and Caola, Fabrizio and Gao, Yanyan and Gritsan, Andrei V. and Martin, Christopher B. and Melnikov, Kirill and Schulze, Markus and Tran, Nhan V. and Whitbeck, Andrew and Zhou, Yaofu". *Phys. Rev. D*, 89:035007, Feb 2014.
- [62] "RooCMSShape". https://github.com/cms-sw/cmssw/ blob/09c3fce6626f70fd04223e7dacebf0b485f73f54/ PhysicsTools/TagAndProbe/src/RooCMSShape.cc, 2013.
- [63] Ye Li and Frank Petriello. "Combining QCD and electroweak corrections to dilepton production in the framework of the FEWZ simulation code". *Phys. Rev. D*, 86:094034, Nov 2012.
- [64] CMS Collaboration. "Measurement of inclusive W and Z boson production cross sections in pp collisions at sqrt(s)=13 TeV". 2015.
- [65] Massimiliano Grazzini, Stefan Kallweit, and Dirk Rathlev. "ZZ production at the LHC: Fiducial cross sections and distributions in NNLO QCD". *Phys. Lett. B*, 750:407 – 410, 2015.
- [66] Anastasiya Bierweiler, Tobias Kasprzik, and Johann H. Kühn. "Vectorboson pair production at the LHC to  $\mathcal{O}(\alpha^3)$  accuracy". *JHEP*, 1312:071, 2013.
- [67] Stefan Gieseke, Tobias Kasprzik, and Johann H. Kühn. "Vectorboson pair production and electroweakcorrections in HERWIG++". *Eur.Phys.J.*, C74(8):2988, 2014.
- [68] "CMS Luminosity Measurements for the 2016 Data Taking Period". Technical Report CMS-PAS-LUM-17-001, CERN, Geneva, 2017.

- [69] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. "Asymptotic formulae for likelihood-based tests of new physics". *Eur. Phys. J.*, C71:1554, 2011. [Erratum: Eur. Phys. J.C73,2501(2013)].
- [70] Alexander L. Read. "Presentation of search results: The CL(s) technique". J. Phys., G28:2693–2704, 2002. [,11(2002)].
- [71] Thomas Junk. "Confidence level computation for combining searches with small statistics". Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 434(2):435 – 443, 1999.
- [72] Bernhard Mistlberger and Falko Dulat. "Limit setting procedures and theoretical uncertainties in Higgs boson searches". 2012.
- [73] "Procedure for the LHC Higgs boson search combination in Summer 2011". Technical Report CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, Aug 2011.
- [74] Sara Bolognesi, Yanyan Gao, Andrei V. Gritsan, Kirill Melnikov, Markus Schulze, Nhan V. Tran, and Andrew Whitbeck. "On the spin and parity of a single-produced resonance at the LHC". *Phys. Rev.*, D86:095031, 2012.
- [75] "Search for heavy resonances decaying into a Z boson and a vector boson in the  $\nu\nu q\bar{q}$  final state". Technical Report CMS-PAS-B2G-17-005, CERN, Geneva, 2017.
- [76] "Search for massive resonances decaying into WW, WZ, ZZ, qW and qZ in the dijet final state at  $\sqrt{s} = 13$  TeV". Technical Report CMS-PAS-B2G-17-001, CERN, Geneva, 2017.
- [77] "Search for new heavy resonances decaying into a Z boson and a massive vector boson in the  $2\ell 2q$  final state at  $\sqrt{s} = 13$  TeV". Technical Report CMS-PAS-B2G-17-013, CERN, Geneva, 2017.
- [78] "Search for heavy resonances decaying to pairs of vector bosons in the 1 nu q qbar final state with the CMS detector in proton-proton

collisions at sqrt s = 13 TeV". Technical Report CMS-PAS-B2G-16-029, CERN, Geneva, 2017.

- [79] "Search for heavy resonances decaying to a pair of Higgs bosons in the four b quark final state in proton-proton collisions at  $\sqrt{s} = 13$  TeV". Technical Report CMS-PAS-B2G-16-026, CERN, Geneva, 2017.
- [80] "Search for heavy resonances decaying into two Higgs bosons or into a Higgs and a vector boson in proton-proton collisions at 13 TeV". Technical Report CMS-PAS-B2G-17-006, CERN, Geneva, 2017.