

Search for Heavy Boson Resonances in Final States with a W
or Z Boson and a Photon Using 139 fb^{-1} of Proton-Proton
Collision Data at $\sqrt{s} = 13 \text{ TeV}$ Collected with the ATLAS
Detector

by

Minyu Feng

Department of Physics
Duke University

Date: _____
Approved: _____

Alfred Goshaw, Advisor

Ayana Arce

Ashutosh Kotwal

Thomas Mehen

Stephen Teitsworth

Dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Department of Physics
in the Graduate School of
Duke University

2020

ABSTRACT

Search for Heavy Boson Resonances in Final States with a W or Z Boson and a Photon Using 139 fb^{-1} of Proton-Proton Collision Data at $\sqrt{s} = 13 \text{ TeV}$ Collected with the ATLAS Detector

by

Minyu Feng

Department of Physics
Duke University

Date: _____
Approved: _____

Alfred Goshaw, Advisor

Ayana Arce

Ashutosh Kotwal

Thomas Mehen

Stephen Teitsworth

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy
in the Department of Physics
in the Graduate School of
Duke University

2020

Copyright © 2020 by Minyu Feng
All rights reserved

Abstract

This thesis presents a search for boson resonances that decay into an energetic photon and the Standard Model’s W or Z boson: $X \rightarrow W\gamma$ and $X \rightarrow Z\gamma$. The dataset used for the measurement was collected at the CERN Large Hadron Collider during a run period from 2015 to 2018. The proton-proton collision center-of-mass energy was 13 TeV with a useful integrated luminosity recorded using the ATLAS detector of 139 fb^{-1} . New particle identification techniques are employed, and event selection criteria are optimized to improve the search sensitivity. No significant deviations from the Standard Model predictions are observed over a $W\gamma$ and $Z\gamma$ mass range from 1.0 to 6.8 TeV. The data are used to set limits with 95% confidence level on the production cross section of massive bosons that are predicted by various extensions of the Standard Model’s theory of elementary particles.

Contents

Abstract	iv
List of Figures	ix
List of Tables	xv
Acknowledgements	xvi
1 Introduction	1
2 Theory Background	4
2.1 The Standard Model	4
2.1.1 Quantum Electrodynamics (QED)	5
2.1.2 Yang-Mills Theory	7
2.1.3 Symmetry Breaking and the Higgs Mechanism	11
2.1.4 Electro-Weak Unification	15
2.1.5 Quantum Chromodynamics (QCD)	17
2.1.6 Success and Limitation of Standard Model	19
2.2 Beyond the Standard Model	22
3 The Large Hadron Collider and the ATLAS Experiment	24
3.1 The Large Hadron Collider	24
3.2 The ATLAS Experiments	27
3.2.1 The Inner tracking Detectors	32
3.2.2 Calorimeters	34
3.2.3 ATLAS Trigger and Data Acquisition System	39
4 Analysis Methods	43

4.1	The measurement of photons and W/Z bosons	43
4.1.1	Photon identification and reconstruction	43
4.1.2	Reconstruction of boosted jets	45
4.1.3	Identification of boosted W/Z boson jets	47
4.1.4	Identification of jets from beauty hadron decays	48
4.2	Statistical Methods	53
4.2.1	Statistical treatment in High Energy Physics	53
4.2.2	Optimization of Sensitivity	58
4.2.3	Setting Exclusive Limits	60
4.2.4	Shape Modeling and Spurious Signal	60
5	Search for $X^\pm \rightarrow W^\pm\gamma$ and $X^0 \rightarrow Z^0\gamma$	63
5.1	Introduction	63
5.2	Data and Simulation Samples	64
5.2.1	Data Samples	64
5.2.2	Signal Simulation Samples	64
5.2.3	Background Simulation	65
5.3	Selection and Categorization of Data	66
5.3.1	Baseline Selection	67
5.3.2	Categorization of the Baseline-selected Events	73
5.4	Signal Shape Modeling	80
5.4.1	Signal Shape Fitting	80
5.4.2	Interpolation	83
5.5	Background composition and modeling	85
5.5.1	Spurious signal test on functional form	86

5.5.2	Parametrization and Estimation of the Background	90
5.6	Systematic Uncertainty	94
5.6.1	Background systematic uncertainties	94
5.6.2	Signal systematic uncertainties	94
5.7	Statistical analysis implementation	97
5.8	Expected limits and the impact of systematic uncertainties	98
6	Results of the $W/Z + \gamma$ heavy resonance search	104
6.1	Data distributions of $m_{J\gamma}$	105
6.2	Significance scan for $W\gamma$ and $Z\gamma$ resonances	105
6.3	Measured limits on the cross-section for production of BSM $X^\pm \rightarrow W^\pm\gamma$ and $X \rightarrow Z\gamma$ resonances	106
7	Conclusions	116
A	Appendices	119
A.1	Control Region	119
A.2	Signal Cutflow	122
A.3	Data/MC Comparison	122
A.4	Optimization of p_T cuts	126
A.4.1	Optimization Methodology	127
A.4.2	Extrapolation vs. constant cut at high mass region	128
A.5	Tests of Background Fit	129
A.5.1	More spurious signal test results	129
A.5.2	Spurious signal with data sideband	134
A.5.3	Spurious signal test with fractional MC samples	134
A.6	Checks for the maximum significance mass points	141

A.6.1	Distributions and Fits	141
A.6.2	Pull and Ranking	141
Bibliography		152

List of Figures

2.1	A table of the elementary particles in the Standard Model.	4
2.2	A summary of SM experimental measurements compared to corresponding theoretical predictions.	20
3.1	A landscape of LHC ring and its surrounding, and the locations of 4 experiments on LHC.	25
3.2	The integrated luminosity delivered by LHC machine, and collected by ATLAS detector, versus years.	28
3.3	The distribution and average value of pileup for each year from 2015 to 2018.	28
3.4	An overview of ATLAS detector structure.	29
3.5	An illustration of different particle signatures in ATLAS detector. .	31
3.6	The structure of the inner detector	32
3.7	The geometry of the Inner Detector. For different pseudo-rapidity, particles will go through different layers of the Inner Detector.	33
3.8	An overview of ATLAS calorimeter system.	35
3.9	A simple demonstration of both electron/photon and hadron showering cascade happening in calorimeters.	37
3.10	ATLAS ECAL barrel Calorimeter geometry.	39
3.11	ATLAS Hadron Tile Calorimeter geometry.	40
3.12	Sliding window algorithm for e/γ L1 triggers.	41
3.13	Ringer algorithm for e/γ HLT identification.	42
4.1	The performance for ATLAS photon triggers.	44

4.2	A simulated example for the anti- k_t algorithm.	45
4.3	The contour region of D2 variable and its discriminating power between boosted Z boson jets and QCD jets.	48
4.4	The cut values for TCC jets tagging for W and Z bosons.	49
4.5	The comparison of performance between TCC jets and LCTopo jets. .	50
4.6	Performance of the MV2c20 algorithm used to discriminate beauty hadron decay products.	52
4.7	Relationship between the p-value and the significance.	54
5.1	Comparison of signal baseline selection efficiencies.	69
5.2	Differences of photon η distribution among different signal channels. .	70
5.3	Differences of photon p_T distribution among different signal channels. .	71
5.4	Differences of jet p_T distribution among different signal channels. . .	72
5.5	The flowchart for the criteria of categorization for $Z + \gamma$ (left) and $W + \gamma$ (right) signals.	74
5.6	Mass dependent photon p_T cut funtions for different signal channels and categories.	75
5.7	Signal categorization efficiencies for different signal channels and categories.	76
5.8	Overall signal efficiencies for different signal channels and categories. .	77
5.9	SM Background categorization efficiencies for spin-0 $Z\gamma$ channel. . .	78
5.10	SM Background categorization efficiencies for spin-1 $W\gamma$ channel. .	79
5.11	Plots of fit parameter σ dependence of $M_{\gamma J}$	81
5.12	Some fit results for the spin-0 $Z\gamma$ signal model.	82

5.13	Some fit results for the spin-1 $W\gamma$ signal model.	83
5.14	Interpolated signal shapes for the spin-0 $Z\gamma$ model.	84
5.15	Interpolated signal shapes for the spin-1 $W\gamma$ model	84
5.16	Spurious signal test on MC samples with different number of parameters for the spin-0 $Z\gamma$ channel.	88
5.17	The absolute value of raw number of spurious signal events $ N_{SS} $ for the spin-0 $Z\gamma$ channel.	89
5.18	Background fit to the $m_{J\gamma}$ mass spectrum from $\gamma + Jet$ MC in the signal region of spin-0 $Z\gamma$ channels.	92
5.19	Background fit to the $m_{J\gamma}$ mass spectrum from $\gamma + Jet$ MC in the signal region of spin-1 $W\gamma$ channels.	93
5.20	The expected 95% CL cross section limits for spin-0 $Z\gamma$ channel.	100
5.21	The expected 95% CL cross section limits for spin-2 $Z\gamma$ channel. The production mode is $q\bar{q} \rightarrow X$	101
5.22	The expected 95% CL cross section limits for spin-2 $Z\gamma$ channel. The production mode is $gg \rightarrow X$	102
5.23	The expected 95% CL cross section limits for spin-1 $W\gamma$ channel.	103
6.1	Fit to the observed $m_{J\gamma}$ distributions in spin-0 $Z\gamma$ channel.	108
6.2	Fit to the observed $m_{J\gamma}$ distributions in spin-2 gg $Z\gamma$ channel.	109
6.3	Fit to the observed $m_{J\gamma}$ distributions in spin-2 qq $Z\gamma$ channel.	110
6.4	Fit to the observed $m_{J\gamma}$ distributions in spin-1 $W\gamma$ channel.	111
6.5	Local p -value scans for different signal channels.	111
6.6	95% CL cross section limits for spin-0 $Z\gamma$ signal channel.	112
6.7	95% CL cross section limits for spin-2 gg $Z\gamma$ signal channel.	113

6.8	95% CL cross section limits for spin-2 qq $Z\gamma$ signal channel.	114
6.9	95% CL cross section limits for spin-1 $W\gamma$ signal channel.	115
A.1	The $m_{J\gamma}$ distribution for the CR dataset.	120
A.2	The signal leakage rate for the CR.	120
A.3	Comparison of spin-0 $Z\gamma$ SR categories with CR dataset, and the function fitting to the ratio of them.	121
A.4	Comparison of spin-1 $W\gamma$ SR categories with CR dataset, and the function fitting to the ratio of them.	121
A.5	Data/MC comparison for $P_T(\gamma)$	124
A.6	Data/MC comparison for $P_T(jet)$ and $\eta(jet)$	125
A.7	Data/MC comparison for M_{jet}	125
A.8	Data/MC comparison for $\Delta\eta_{\gamma J}$ and $\Delta\phi_{\gamma J}$	126
A.9	An example of p_T cut scan for optimization.	128
A.10	Optimized p_T cut values and corresponding fitting functions.	129
A.11	Spurious signal test on MC samples with different number of parameters for the spin-1 $W\gamma$ channel.	130
A.12	Spurious signal test on MC samples with different number of parameters for the spin-2 gg $Z\gamma$ channel.	130
A.13	Spurious signal test on MC samples with different number of parameters for the spin-2 qq $Z\gamma$ channel.	131
A.14	The absolute value of raw number of spurious signal events $ N_{SS} $ for the spin-1 $W\gamma$ channel.	131
A.15	The absolute value of raw number of spurious signal events $ N_{SS} $ for the spin-2 gg $Z\gamma$ channel.	132

A.16 The absolute value of raw number of spurious signal events $ N_{SS} $ for the spin-2 qq $Z\gamma$ channel.	133
A.17 Comparison of N_{ss} results with data control region (black line) and MC samples (blue line) for spin-0 $Z\gamma$ d2, vmass categories.	134
A.18 Comparison of N_{ss} results with data control region (black line) and MC samples (blue line) for spin-2 qq $Z\gamma$ d2, vmass categories.	135
A.19 Comparison of N_{ss} results with data control region (black line) and MC samples (blue line) for spin-2 gg $Z\gamma$ d2, vmass categories.	135
A.20 Comparison of N_{ss} results with data control region (black line) and MC samples (blue line) for spin-1 $W\gamma$ d2, vmass categories.	135
A.21 Coverage of $2\times$ fitted function to the spurious signal results for spin-0 $Z\gamma$ channel.	136
A.22 Coverage of $2\times$ fitted function to the spurious signal results for spin-2 gg $Z\gamma$ channel.	136
A.23 Coverage of $2\times$ fitted function to the spurious signal results for spin-2 qq $Z\gamma$ channel.	137
A.24 Coverage of $2\times$ fitted function to the spurious signal results for spin-1 $W\gamma$ channel.	137
A.25 Spurious signal test on half of MC samples for the spin-0 $Z\gamma$ channel.	138
A.26 Spurious signal test on half of MC samples for the spin-1 $W\gamma$ channel.	139
A.27 Spurious signal test on half of MC samples for the spin-2 qq $Z\gamma$ channel.	139
A.28 Spurious signal test on half of MC samples for the spin-2 gg $Z\gamma$ channel.	140
A.29 Signal+Background fit with $m_X = 3640$ GeV signal for spin-0 $Z\gamma$ channel.	142
A.30 Signal+Background fit with $m_X = 3580$ GeV signal for spin-2 gg $Z\gamma$ channel.	142

A.31 Signal+Background fit with $m_X = 3560$ GeV signal for spin-2 qq $Z\gamma$ channel.	143
A.32 Signal+Background fit with $m_X = 2820$ GeV signal for spin-1 $W\gamma$ channel.	143
A.33 Pull and ranking distributions for spin=0 $Z\gamma$ $m_X = 3640$ TeV signal.	144
A.34 Pull and ranking distributions for spin=1 $W\gamma$ $m_X = 2820$ TeV signal.	145
A.35 Pull and ranking distributions for spin=2 gg $Z\gamma$ $m_X = 3580$ TeV signal.	146
A.36 Pull and ranking distributions for spin=2 qq $Z\gamma$ $m_X = 3560$ TeV signal.	147

List of Tables

4.1	Performance of B-tagging for different efficiency operation points. . .	53
5.1	The simulated samples used in this search is listed in the table. The corresponding theory models and simulation generators are listed too.	65
5.2	Summary of NLO photon+jet QCD background samples generated with SHERPA.	66
5.3	Observed data and expected compositions of backgrounds in the $Z\gamma$ Spin-0 channel.	85
5.4	Observed data and expected composition of backgrounds in the $W\gamma$ spin-1 channel.	85
5.5	A summary of the number of parameters used for the background fit functions in different channels and categories.	91
6.1	Observed number of events in each signal channel and category. . . .	104
A.1	Cutflow for signal samples at Mass = 1 TeV.	122
A.2	Cutflow for signal samples at Mass = 2 TeV.	123
A.3	Cutflow for signal samples at Mass = 4 TeV.	123

Acknowledgements

There are so many people supported and helped me during the seven years of my graduate study at Duke. Firstly, thanks to my advisor, Al Goshaw. Thank you for guiding me through my graduate student years. Not only your knowledge in physics expanded my perspective, your attitude and passion about research also keep encouraging me. Many thanks to Ashutosh, Ayana, and Mark, who provided lots of suggestions during my graduate study at Duke. You are all very good teachers, and also very nice people to work with. Thanks to Andrea, Shu, Zhijun, Chen and Mia. You gave me so much help during my life and study at CERN. I can never imagine myself enjoy the life and study in Europe without you. Also thanks to Bo, Boping, Vincent and Wei. The experience of conducting this search is memorable. It is a pleasure to exchange ideas with you and inspire each other. Finally, thanks to my loving parents, who supported me financially and mentally during the whole program. Especially in this COVID-19 pandemic, your support is indispensable for me to finish my dissertation.

Chapter 1

Introduction

One of the main philosophies behind science is reductionism, which believes the universe can be described by some fundamental building blocks, and all higher level phenomena can be derived from the knowledge of these blocks. Although further researches imply that reductionism does not tell the full story, it will definitely not be the complete story without the knowledge of fundamental blocks of universe. Particle physics is at the uttermost frontier guided by this spirit of reductionism. The most fundamental building blocks known by physicists are called the elementary particles, including quarks, leptons and force carriers (bosons). Most of the knowledge of these elementary particles are unified under a theory called the Standard Model (SM).

The SM is one of the most significant achievements of all fields of science in the last half century. Based on SM, physicists can predict with high precision and explain in details most of the phenomena observed. It successfully unified three out of four interactions discovered by physicists: electromagnetic, strong and weak interactions. Only the gravity is left outside of the SM. The interactions in SM are described by symmetries – $SU(3)_C \times SU(2)_L \times U(1)_Y$. The $SU(3)_C$ describes the strong interaction, and the $SU(2)_L \times U(1)_Y$ unified the electromagnetic and weak interaction into a so-called electroweak sector. This electroweak sector originally has only the unified electroweak interaction, and then the symmetry is broken into two parts with the Higgs mechanism. In standard model, all these interactions are propagated by its corresponding force carriers: photons for electromagnetic interaction, W^\pm and Z bosons for weak interaction and gluons for strong interaction. However, half a century ago, the only force carrier discovered is the photon. All other three bosons, and also

the Higgs boson, are discovered and validated on colliders since then. Actually, the particle physics has been substantially developed in last several decades, because of this improvement of the observation tools, large particle colliders and detectors, for producing particles and probing them with high precision.

Although SM has achieved brilliant success, there are still some observations deviate from SM predictions. Just like the two clouds obscuring the classical dynamics, these unsolved problems like the composition of dark matter and source of neutrino mass, can potentially lead to some new physics beyond standard model (BSM). For example, one of the hypothesis for dark matter, is that dark matter can weakly interact with normal particles. This interaction could be originally unified with electroweak sector, and was separated after a symmetry breaking process, which means the original unified electroweak interaction has more symmetry dimensions than we expected, and the Higgs mechanism is more complicated than our current understanding. If this hypothesis is true, the real electroweak sector will have more gauge bosons. The new gauge boson might not be able to interact with other fermions, but at least should be able to interact with other bosons like photons and W/Z bosons. This new gauge boson decaying into a photon and a W/Z boson is the target process of the search in this thesis.

In this thesis, the search is conducted on the largest, and highest energy dataset ever collected for high energy particle collisions. This $139\ fb^{-1}$ dataset is collected during Run 2 of the Large Hadron Collider (LHC) by the ATLAS experiment. The protons collide with each other in LHC at 13 TeV. At this energy, some new processes might be unlocked, among them is the hypothetical new gauge boson in electroweak sector. Even though the chance of producing this boson can be extremely low, with such a large dataset, it could be produced many times, and there is still a good chance to find them.

The content of this thesis are organized as follows. Chapter 2 provides the background theory of SM, together with the theoretical motivation for this search. Chapter 3 demonstrates the details of the Large Hadron Collider and the ATLAS experiment at CERN. Chapter 4 explains the experimental and statistical methods used in this search. The final results including the corresponding statistical and systematical uncertainties are shown in Chapter 5. In Chapter 6, the results of this search is shown, and in Chapter 7 conclusions are made and future plans to extend this search are listed.

Chapter 2

Theory Background

2.1 The Standard Model

mass \rightarrow	$\approx 2.3 \text{ MeV}/c^2$	charge \rightarrow	$2/3$	spin \rightarrow
			$1/2$	
	u	c	t	
	up	charm	top	
charge \rightarrow	$\approx 4.8 \text{ MeV}/c^2$	charge \rightarrow	$\approx 95 \text{ MeV}/c^2$	charge \rightarrow
	$-1/3$	$-1/3$	$-1/3$	$1/2$
	d	s	b	
	down	strange	bottom	
mass \rightarrow	$\approx 0.511 \text{ MeV}/c^2$	mass \rightarrow	$\approx 105.7 \text{ MeV}/c^2$	mass \rightarrow
	-1	-1	-1	-1
	e	μ	τ	
	electron	muon	tau	
mass \rightarrow	$<2.2 \text{ eV}/c^2$	mass \rightarrow	$<0.17 \text{ MeV}/c^2$	mass \rightarrow
	0	0	0	$1/2$
	ν_e	ν_μ	ν_τ	
	electron neutrino	muon neutrino	tau neutrino	
charge \rightarrow	$\approx 91.2 \text{ GeV}/c^2$	charge \rightarrow	$\approx 80.4 \text{ GeV}/c^2$	charge \rightarrow
	0	0	± 1	1
	Z		W	
	Z boson		W boson	

Figure 2.1: A table of the elementary particles in the Standard Model.

The Standard Model (SM) is a gauge theory describing features of all known particles and interactions (except gravity). It unifies the strong, weak and electromagnetic interactions with a common framework. This is the best attempt so far to build a theory of everything. There are three generations of quarks and leptons (spin 1/2), four gauge bosons (spin 1) and a Higgs boson (spin 0) contained in the SM. These particles are shown in Fig. 2.1.[1] Four gauge bosons are the mediators for interactions: gluons mediate the strong interaction, W^\pm and Z^0 bosons mediate the weak interaction, and photons mediate the electromagnetic interaction. The Higgs boson generates mass for the W^\pm and Z^0 bosons, and for the quarks and leptons. This successful theory is built upon some simple symmetry assumptions, and de-

veloped within the framework of non-abelian gauge field theories. In this section, I will demonstrate how these assumptions can be developed into a theory of almost everything. The natural units ($\hbar = c = 1$) and Einstein summation convention are used in this chapter.

2.1.1 Quantum Electrodynamics (QED)

In classical electrodynamics, the electromagnetic field is not unique. Any transformation in the form of Eqn. 2.1 will not change the dynamic equations, and therefore is effectively identical. Such a transformation is called a gauge transformation, and the invariance (of physics) under this transformation is called gauge invariance.

$$\vec{A} \rightarrow \vec{A}' = \vec{A} + \nabla\lambda, V \rightarrow V' = V - \frac{\partial\lambda}{\partial t} \quad (2.1)$$

In 1918, Hermann Weyl tried to unify the only two forces known at that time – the electromagnetic force and gravity – with the help of this gauge transformation.[2] This is the first attempt to develop gauge theories. Although this attempt failed, together with the Noether's theorem also published the same year, a new perspective emerges by connecting invariance with conservation laws that has deeply changed the philosophy of modern physics.

The first light of gauge field theory appears when Weyl pointed out the Lagrangian of the Dirac equation for ψ (Eqn. 2.2) and the Maxwell equation for four potentials A_μ and its tensor $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ (Eqn. 2.3) are invariant under the gauge transformation of Eqn. 2.4.

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi \quad (2.2)$$

$$\mathcal{L} = -\frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} \quad (2.3)$$

$$\psi \rightarrow e^{ig_e \lambda(x)} \psi, A_\mu \rightarrow A_\mu + \partial_\mu \lambda \quad (2.4)$$

In Eqn. 2.4, the g_e is a constant and its meaning will be revealed later. The λ is dependent on spacetime coordinates, which implies the phase term of the wavefunction is localized. Such a gauge invariance is called “local gauge invariance” compared to the “global gauge invariance” with λ constant over all space and time.

The Lagrangian describing the electromagnetic interaction can be derived from the Dirac Lagrangian based on this gauge invariance. [3] [4] If the Dirac Lagrangian is transformed by Eqn. 2.4, there will be an extra term:

$$\mathcal{L}' = i\bar{\psi} \gamma^\mu \partial_\mu \psi - m\bar{\psi} \psi - (g_e \bar{\psi} \gamma^\mu \psi) \partial_\mu \lambda.$$

The coefficient g_e is related to both the fractional charge of fermion (f), and electromagnetic fine structure constant (α_{em}):

$$g_e = f \sqrt{4\pi \alpha_{\text{em}}}.$$

Since the other half of the gauge transformation is also related to $\partial_\mu \lambda$:

$$A_\mu \rightarrow A_\mu + \partial_\mu \lambda.$$

By adding an extra term into the Dirac Lagrangian, the invariance can be retained:

$$\mathcal{L} = i\bar{\psi} \gamma^\mu \partial_\mu \psi - m\bar{\psi} \psi - (g_e \bar{\psi} \gamma^\mu \psi) A_\mu.$$

On the other hand, the Lagrangian terms for a massive spin-1 particle field is the Proca Lagrangian:

$$\mathcal{L} = -\frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} + \frac{m_A^2}{8\pi} A^\nu A_\nu. \quad (2.5)$$

The first term is invariant under Eqn. 2.4, but the second term is not. Fortunately, since there is a mass factor m_A in this term, a massless spin-1 field like photon can eliminate this term and fully avoid such a problem. Therefore, the full Lagrangian derived from Dirac Lagrangian based on local gauge invariance becomes:

$$\mathcal{L}_{QED} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - \frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} - (g_e\bar{\psi}\gamma^\mu\psi)A_\mu. \quad (2.6)$$

The first and second term of Eqn. 2.6 describe a free spin- $\frac{1}{2}$ field (fermion) with mass m , third term is for a massless free spin-1 field (gauge boson), and the last term for the interaction between these two fields. This is exactly how the electromagnetic interaction works.

Since Eqn. 2.4 describes a rotation of ψ in the complex plane, which means the gauge transformation is isomorphic to the $U(1)$ group, this gauge invariance is called “ $U(1)$ gauge invariance”. Electric charge conservation corresponds to this $U(1)$ gauge invariance, just like energy and momentum conservation corresponds to the space and time translations.

2.1.2 Yang-Mills Theory

The $U(1)$ gauge invariance used to derive the theory of the electromagnetic interaction was noticed after the theory had been established. It seems to be more like an alternative formalization of QED, which undermined the significance of this gauge invariance. The situation persists until Chen-Ning Yang and Robert Mills applied the same local gauge invariance using the group $SU(2)$. They built a framework for the foundation used to describe the weak interaction called the “Yang-Mills Theory”,

and also inspired further study of other non-abelian gauge field theories.

In this section, an overview of the development of the Yang-Mills Theory based on the section 10.4 of [3] will be presented. From just the assumption of $SU(2)$ symmetry, a model for the weak interaction can be built with three massless vector bosons – one neutral, two oppositely charged.

Consider a composite wave-function constructed from a doublet of particles like (u, d) or (e^-, ν_e) . We cannot describe all possible unitary transformations for this doublet with just a $U(1)$ group. For example, define the new wave-function for the particle doublet as:

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}.$$

These two components (ψ_1 and ψ_2) are both 4-component Dirac wave-functions. Assume the mass of the two particles is the same. The Lagrangian for such a particle doublet becomes:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi.$$

The corresponding gauge transformation is again a unitary matrix, but now belongs to the $SU(2)$ symmetry group:

$$\psi \rightarrow U\psi.$$

The U matrix is a 2×2 unitary matrix, and since all unitary matrices can be expressed by exponential of Hermitian matrices, we can have:

$$U = e^{iH},$$

$$H = \theta \mathbf{1} + \boldsymbol{\tau} \cdot \mathbf{a}.$$

\mathbf{a} is a three dimensional arbitrary vector, and the $\boldsymbol{\tau}$ is constructed from the Pauli matrices as:

$$\boldsymbol{\tau} = \begin{pmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{pmatrix},$$

$$\tau_1 = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tau_2 = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tau_3 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Therefore, the gauge transformation for SU(2) becomes:

$$\psi \rightarrow e^{i\theta} e^{i\boldsymbol{\tau} \cdot \mathbf{a}} \psi. \quad (2.7)$$

The first part is exactly the U(1) gauge transformation, and the second factor is for the SU(2) gauge transformation. For future convenience, an SU(2) local gauge transformation can be defined by letting $\mathbf{a}(x) = -g_w \boldsymbol{\lambda}(x)$.

$$\begin{aligned} \psi &\rightarrow e^{-ig_w \boldsymbol{\tau} \cdot \boldsymbol{\lambda}(x)} \psi \\ \mathbf{A}_\mu &\rightarrow \mathbf{A}_\mu + \partial_\mu \boldsymbol{\lambda} + 2g_w (\boldsymbol{\lambda} \times \mathbf{A}_\mu) \end{aligned} \quad (2.8)$$

In Yang-Mills theory used to describe weak interaction, the 3-components of the weak field \mathbf{A}_μ can be defined as:

$$\mathbf{A}_\mu = (W_{1\mu}, W_{2\mu}, W_{3\mu}).$$

The extra term needed for the free fermion Lagrangian to be invariant under such a transformation (Eqn. 2.8) is:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - (g_w\bar{\psi}\gamma^\mu\boldsymbol{\tau}\psi)\cdot\mathbf{A}_\mu.$$

For a multi dimensional field \mathbf{A}_μ , the field tensor also need to be modified because there is a new asymmetric term $\mathbf{A}_\mu \times \mathbf{A}_\nu$ available. In this case, the field tensor $\mathbf{F}^{\mu\nu}$ is defined as:

$$\mathbf{F}^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu - 2g_w(A^\mu \times A^\nu).$$

This coefficient choice makes the term $\mathbf{F}^{\mu\nu}\mathbf{F}_{\mu\nu}$ invariant under Eqn. 2.8 when the boson is massless, which makes the full Yang-Mills Lagrangian as:

$$\mathcal{L}_{YM} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - (g_w\bar{\psi}\gamma^\mu\boldsymbol{\tau}\psi)\cdot\mathbf{A}_\mu - \frac{1}{16\pi}\mathbf{F}^{\mu\nu}\mathbf{F}_{\mu\nu}. \quad (2.9)$$

The second last term describe the interaction between fermions and bosons. Since $\mathbf{A}_\mu = (W_{1\mu}, W_{2\mu}, W_{3\mu})$, it can be expanded as:

$$(g_w\bar{\psi}\gamma^\mu\boldsymbol{\tau}\psi)\cdot\mathbf{A}_\mu = \frac{g_w}{2}[\bar{\psi}_2\gamma^\mu\psi_1(W_1^\mu + iW_2^\mu) + \bar{\psi}_1\gamma^\mu\psi_2(W_1^\mu - iW_2^\mu) + (\bar{\psi}_1\gamma^\mu\psi_1 - \bar{\psi}_2\gamma^\mu\psi_2)W_3^\mu] \quad (2.10)$$

Each of the interaction terms represent a coupling in the weak interaction, between either same or opposite part of the weak doublet. This means we can redefine each of the component of weak field as:

$$\begin{aligned} W_\mu^- &= \frac{1}{\sqrt{2}}(W_1^\mu + iW_2^\mu), \\ W_\mu^+ &= \frac{1}{\sqrt{2}}(W_1^\mu - iW_2^\mu), \\ W_\mu^0 &= W_3^\mu. \end{aligned}$$

This shows how a theory framework for weak interaction can be built by solely requiring the local gauge invariance for $SU(2)$. The main characteristic for weak interaction is included in this simplified model of Yang-Mills theory: there are two charged bosons and a neutral boson, intermediating either flavor changing or flavor neutral current of weak interaction. However, several problems in this theory have to be addressed. The most critical one is that, mass term itself is not invariant under gauge transformation. It seems the Yang-Mills theory prohibit vector bosons to acquire any mass.

2.1.3 Symmetry Breaking and the Higgs Mechanism

Although the Yang-Mills theory looks beautiful in mathematics, physicists, including Wolfgang Pauli, questioned the theory due to its lack of boson mass, and therefore making the weak interaction effective at long distance. This seems to be a problem for all field theory derived from gauge invariance, because the mass term of spin-1 Proca Lagrangian (Eqn. 2.5) is not invariant under local gauge transformation. This problem is finally solved by introducing symmetry breaking into the gauge theory. A simplified derivation of the Higgs mechanism will be presented in this section based on section 10.7 to 10.9 of [3].

In order to find a way to generate mass for spin-1 bosons, we need to identify the mass term in Lagrangian first. For example, in the Klein-Gordon Lagrangian (Eqn. 2.11), the mass term can be easily identified as $-\frac{1}{2}m^2\phi^2$.

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2 \quad (2.11)$$

However, when replacing the mass term $-\frac{1}{2}m^2\phi^2$ with the Higgs potential $+\frac{1}{2}\mu^2\phi^2 - \frac{1}{4}\lambda^2\phi^4$,¹ the new mass term becomes less obvious.

¹This λ is a constant for Higgs potential, not the same as defined in section 2.1.1 and 2.1.2.

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi + \frac{1}{2}\mu^2\phi^2 - \frac{1}{4}\lambda^2\phi^4 \quad (2.12)$$

Since the Feynmann calculus is perturbative, we can consider wave-functions as excitations from ground state (vacuum). This ground state is determined by the wave-function with minimum potential. The Klein-Gordon Lagrangian is just a coincident with ground state being $\phi = 0$, and the equation itself is automatically a expansion around it. In more general cases, the mass term is the second order term when the Lagrangian is expanded around the ground state of the wave-function. Take Eqn. 2.12 as an example. By considering the Lagrangian as $\mathcal{L} = \mathcal{T} - \mathcal{U}$, the kinetic and potential parts are:

$$\begin{aligned}\mathcal{T} &= \frac{1}{2}\partial_\mu\phi\partial^\mu\phi, \\ \mathcal{U} &= -\frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda^2\phi^4.\end{aligned}$$

This potential reaches its minimum when $\phi = \pm\mu/\lambda$. With a wise choice of gauge transformation, the positive sign can be taken. By defining the perturbative wave-function as η , the full Lagrangian can be expanded around $\phi = \mu/\lambda + \eta$.

$$\begin{aligned}\mathcal{L} &= -\frac{1}{2}\partial_\mu(\eta + \frac{\mu}{\lambda})\partial^\mu(\eta + \frac{\mu}{\lambda}) + \frac{1}{2}\mu^2(\eta + \frac{\mu}{\lambda})^2 - \frac{1}{4}\lambda^2(\eta + \frac{\mu}{\lambda})^4 \\ &= -\frac{1}{2}\partial_\mu\eta\partial^\mu\eta - \mu^2\eta^2 - \mu\lambda\eta^3 - \frac{1}{4}\lambda^2\eta^4 + \text{constant}\end{aligned}$$

The mass term is now clearly $-\mu^2\eta^2$, and corresponding mass of boson is $m = \sqrt{2}\mu$.

After clarifying the mass term in Lagrangian, the mechanism of generating mass for spin-1 boson can be explored. This can be done by introducing a scalar field

similar to Eqn. 2.12 and a spin-1 field \mathbf{A}_μ . The scalar field ϕ is not real but complex ($\phi = \phi_1 + i\phi_2$), and the derivative is replaced by covariant derivative ($\mathcal{D}_\mu = \partial_\mu + iq\mathbf{A}_\mu$) to fulfill local gauge invariance. Then the full Lagrangian becomes:

$$\mathcal{L} = \frac{1}{2}(\mathcal{D}_\mu\phi)^*\mathcal{D}^\mu\phi + \frac{1}{2}\mu^2\phi^*\phi - \frac{1}{4}\lambda^2(\phi^*\phi)^2 - \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu}. \quad (2.13)$$

Since the potential in Eqn 2.13 is the same as Eqn. 2.12, the ground state is at $\phi_1^2 + \phi_2^2 = \mu^2/\lambda^2$. Due to the global gauge invariance, the ground state can be taken as Eqn 2.14 without losing generality.

$$\begin{aligned} \phi_1 &= \eta + \frac{\mu}{\lambda} \\ \phi_2 &= \xi \end{aligned} \quad (2.14)$$

Then the perturbative expansion of Lagrangian is:

$$\begin{aligned} \mathcal{L} &= \frac{1}{2}[(\partial_\mu - iq\mathbf{A}_\mu)\phi^*][(\partial^\mu + iq\mathbf{A}^\mu)\phi] + \frac{1}{2}\mu^2\phi^*\phi - \frac{1}{4}\lambda^2(\phi^*\phi)^2 - \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} \\ &= \frac{1}{2}[(\partial_\mu - iq\mathbf{A}_\mu)(\eta + \frac{\mu}{\lambda} - i\xi)][(\partial^\mu + iq\mathbf{A}^\mu)(\eta + \frac{\mu}{\lambda} + i\xi)] \\ &\quad + \frac{1}{2}\mu^2(\eta^2 + \frac{2\mu\eta}{\lambda} + \frac{\mu^2}{\lambda^2} + \xi^2) - \frac{1}{4}\lambda^2(\eta^2 + \frac{2\mu\eta}{\lambda} + \frac{\mu^2}{\lambda^2} + \xi^2)^2 - \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu}. \end{aligned}$$

There is a term $\frac{q\mu}{\lambda}\partial_\mu\xi\mathbf{A}^\mu$ left from the kinetic term implying direct interchange between scalar boson ξ and vector boson \mathbf{A} , which has never been observed. This can be solved by applying a proper local gauge transformation making ϕ always a real value. In other words, ξ has to be fixed at 0. This can greatly simplify the calculation:

$$\begin{aligned}
\mathcal{L} &= \frac{1}{2}[(\partial_\mu - iq\mathbf{A}_\mu)(\eta + \frac{\mu}{\lambda})][(\partial^\mu + iq\mathbf{A}^\mu)(\eta + \frac{\mu}{\lambda})] \\
&\quad + \frac{1}{2}\mu^2(\eta^2 + \frac{2\mu\eta}{\lambda} + \frac{\mu^2}{\lambda^2}) - \frac{1}{4}\lambda^2(\eta^2 + \frac{2\mu\eta}{\lambda} + \frac{\mu^2}{\lambda^2})^2 - \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} \\
&= \frac{1}{2}[\partial_\mu\eta\partial^\mu\eta + (\eta + \frac{\mu}{\lambda})^2q^2\mathbf{A}_\mu\mathbf{A}^\mu] \\
&\quad - \mu^2\eta^2 - \mu\lambda\eta^3 - \frac{1}{4}\lambda^2\eta^4 - \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \text{constant} \\
&= \frac{1}{2}\partial_\mu\eta\partial^\mu\eta + \frac{q^2}{2}\eta^2\mathbf{A}_\mu\mathbf{A}^\mu + \frac{q^2\mu}{\lambda}\eta\mathbf{A}_\mu\mathbf{A}^\mu + \frac{q^2\mu^2}{2\lambda^2}\mathbf{A}_\mu\mathbf{A}^\mu \\
&\quad - \mu^2\eta^2 - \mu\lambda\eta^3 - \frac{1}{4}\lambda^2\eta^4 - \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \text{constant}
\end{aligned}$$

By properly re-grouping all terms:

$$\begin{aligned}
\mathcal{L} &= \frac{1}{2}\partial_\mu\eta\partial^\mu\eta - \mu^2\eta^2 \rightarrow (\text{kinetic and mass terms of spin-0 boson}) \\
&\quad - \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \frac{q^2\mu^2}{2\lambda^2}\mathbf{A}_\mu\mathbf{A}^\mu \rightarrow (\text{kinetic and mass terms of spin-1 boson}) \\
&\quad + \frac{q^2\mu}{\lambda}\eta\mathbf{A}_\mu\mathbf{A}^\mu - \mu\lambda\eta^3 \rightarrow (\text{triple boson coupling}) \\
&\quad + \frac{q^2}{2}\eta^2\mathbf{A}_\mu\mathbf{A}^\mu - \frac{1}{4}\lambda^2\eta^4 \rightarrow (\text{quadruple boson coupling})
\end{aligned}$$

Comparing the kinetic and mass terms of spin-1 boson with Eqn. 2.5, the mass of boson can be found:

$$\begin{aligned}
\frac{m_A^2}{8\pi} &= \frac{q^2\mu^2}{2\lambda^2}, \\
m_A &= \frac{2\sqrt{\pi}q\mu}{\lambda}.
\end{aligned}$$

Therefore, by introducing a new scalar boson, and applying both global and local gauge invariance, the mass term for spin-1 boson can be generated. This mechanism is called the Higgs mechanism, and the scalar boson introduced is called the Higgs boson. With this mechanism, the weak bosons (W^\pm and Z^0) can be massive, which makes the weak interaction to be effective only at short distance. The idea of the Higgs mechanism was originally developed by Peter Higgs and other physicists in 1964 [5] [6] [7]. The Higgs boson has been searched for half a century until the discovery of the Higgs boson at a mass of 125 GeV in 2012 by the ATLAS and CMS experiments at the LHC [8]. The Higgs mechanism was finally confirmed and the gauge theory for elementary particles – the Standard Model – got its strongest evidence.

2.1.4 Electro-Weak Unification

Since the Higgs mechanism provided a solution to the generation of weak boson mass, the Yang-Mills theory became more convincing. Further efforts were made to unify both the electromagnetic and weak interactions. These two interactions are unified into a single framework by extending the symmetry group so that it contains subgroups of both $U(1)$ for the electromagnetic interaction and $SU(2)$ for the weak interaction. This is called the Glashow-Weinberg-Salam (GWS) theory.[3] However, there are two corrections needed to build this framework. Firstly, a change is required for a practical theory of the weak interaction to account for parity violation. It turns out neutrinos are all left-handed, and so are the leptons interacting with them through the weak interaction. This means a parity operator for left-handed particles (P_L) is needed for the weak interaction terms in the Lagrangian. Secondly, in order to deal with the components in weak isospin doublets with different eigenstates of electric and weak charges, a hyper-charge operator Y is introduced. This operator is defined by the Gell-Mann Nishijima relation:

$$Q = \tau_3 + \frac{Y}{2}. \text{ (only for left-handed particles)}$$

Under the symmetry of this $U(1)_Y \times SU(2)_L$ group, the unified electro-weak Lagrangian can be built. The flavor changing terms of W^\pm in Eqn. 2.10 will be preserved, and these W^\pm fields correspond to the W^\pm bosons. The flavor neutral terms from $SU(2)_L$ and $U(1)_Y$ will come together:

$$\mathcal{L}_{neutral} = -g'(\bar{\psi}\gamma^\mu \frac{Y}{2}\psi)B_\mu - g_w(\bar{\psi}\gamma^\mu \tau_3 P_L \psi)W_\mu^0.$$

These two fields can be mixed with each other:

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_w & -\sin \theta_w \\ \sin \theta_w & \cos \theta_w \end{pmatrix} \begin{pmatrix} W_\mu^0 \\ B_\mu \end{pmatrix}.$$

With proper choice of the Weinberg angle θ_w , the QED interaction term can be restored, which requires:

$$\begin{aligned} \mathcal{L}_{QED} &= (\text{kinetic terms}) - \bar{\psi}\gamma^\mu(g' \cos \theta_w \frac{Y}{2} + g_w \sin \theta_w \tau_3 P_L)\psi A_\mu \\ &= (\text{kinetic terms}) - g_e(\bar{\psi}\gamma^\mu Q\psi)A_\mu. \end{aligned}$$

Compare this equation with Gell-Mann Nishijima relation, the coupling constants can be determined based on the known g_e :

$$\begin{aligned} g' &= \frac{g_e}{\cos \theta_w}, \\ g_w &= \frac{g_e}{\sin \theta_w}. \end{aligned}$$

And the remaining term for Z_μ is:

$$\begin{aligned}
\mathcal{L}_Z &= -\bar{\psi}\gamma^\mu(g_w \cos \theta_w \tau_3 P_L - g' \sin \theta_w \frac{Y}{2})\psi Z_\mu \\
&= -\bar{\psi}\gamma^\mu[g_w \cos \theta_w \tau_3 P_L - g' \sin \theta_w (Q - \tau_3 P_L)]\psi Z_\mu \\
&= -\bar{\psi}\gamma^\mu[(g_w \cos \theta_w + g' \sin \theta_w)\tau_3 P_L - g' \sin \theta_w Q]\psi Z_\mu \\
&= -\bar{\psi}\gamma^\mu(\frac{g_e}{\sin \theta_w \cos \theta_w}\tau_3 P_L - \frac{g_e \sin \theta_w}{\cos \theta_w}Q)\psi Z_\mu \\
&= -\frac{g_e}{\sin \theta_w \cos \theta_w}\bar{\psi}\gamma^\mu(\tau_3 P_L - \sin^2 \theta_w Q)\psi Z_\mu \\
&= -g_z \bar{\psi}\gamma^\mu(\tau_3 P_L - x_w Q)\psi Z_\mu
\end{aligned}$$

where

$$g_z = \frac{g_e}{\sin \theta_w \cos \theta_w}$$

$$x_w = \sin^2 \theta_w.$$

The electromagnetic interaction and weak interaction are unified under the same framework built by non-abelian gauge theory as shown in this section. The predictions of this theory have been tested for decades, and results are all consistent. After the discovery of Higgs boson in 2012, the electroweak unification theory acquired the last missing piece. Even if this theory is finally proved to be inaccurate in some regime, it has already provided a sound ground for future exploration.

2.1.5 Quantum Chromodynamics (QCD)

Similar to the weak interaction, the strong interaction can also be interpreted using a non-abelian gauge theory. Unlike the electro-weak case, the only kind of fermions affected by strong interaction are the quarks. They each form a triplet with 3 different color charges, usually call red, green and blue. This is why the theory of strong interaction is called Quantum Chromodynamics (QCD).

The wave-function of a color triplet can be defined as:

$$\psi = \begin{pmatrix} \psi_r \\ \psi_g \\ \psi_b \end{pmatrix}.$$

Although each flavor of quarks comes with three different colors, the mass of them are supposed to be identical. The Lagrangian of a free quark is:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi.$$

The corresponding local gauge transformation can be applied by a U(1) and an SU(3) symmetry group. The transformation U of such a group can be written as:

$$U = e^{i\theta} e^{i\lambda \cdot a}.$$

Here λ are the Gell-Mann matrices:

$$\begin{aligned} \lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda_2 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ \lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} & \lambda_5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} & \lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\ \lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} & \lambda_8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \end{aligned}$$

A modification from Yang-Mills theory is required due to the additional dimension of wave-function and symmetry group. The cross product is now defined as:

$$(\mathbf{B} \times \mathbf{C})_i = \sum_{j,k=1}^8 f_{ijk} B_j C_k.$$

With the f_{ijk} to be the structure constants of SU(3) as:

$$[\lambda_i, \lambda_j] = 2i f_{ijk} \lambda_k.$$

The QCD Lagrangian is therefore very similar to Eqn. 2.9, except for the interaction term:

$$\mathcal{L}_{QCD} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - (g\bar{\psi}\gamma^\mu\lambda\psi)\cdot\mathbf{A}_\mu - \frac{1}{16\pi}\mathbf{F}^{\mu\nu}\mathbf{F}_{\mu\nu}. \quad (2.15)$$

The eight Gell-Mann matrices correspond to eight different gluons. They are massless spin-1 vector bosons.

2.1.6 Success and Limitation of Standard Model

The predictions of the Standard Model have achieved great success in the past several decades. As a theory covering almost every aspect of the microscopic world, its predictions are in extraordinary agreement with experimental observations. Figure 2.2 provided a brief summary of experimental measurements done to test the SM. It shows fantastic agreement between SM theory predictions and the experimental measurements. Furthermore, some new particles are discovered with the guidance of SM. Based on calculations using the GWS theory [10] and measurement of θ_w , the W^\pm and Z^0 boson mass were predicted to be:

$$m_W = 82 \pm 2 \text{ GeV},$$

$$m_Z = 92 \pm 2 \text{ GeV}.$$

These two bosons were discovered in 1983 [11] with masses measured to be:

$$m_W = 81.0 \pm 2.5 \pm 1.3 \text{ GeV},$$

$$m_Z = 91.9 \pm 1.3 \pm 1.4 \text{ GeV}.$$

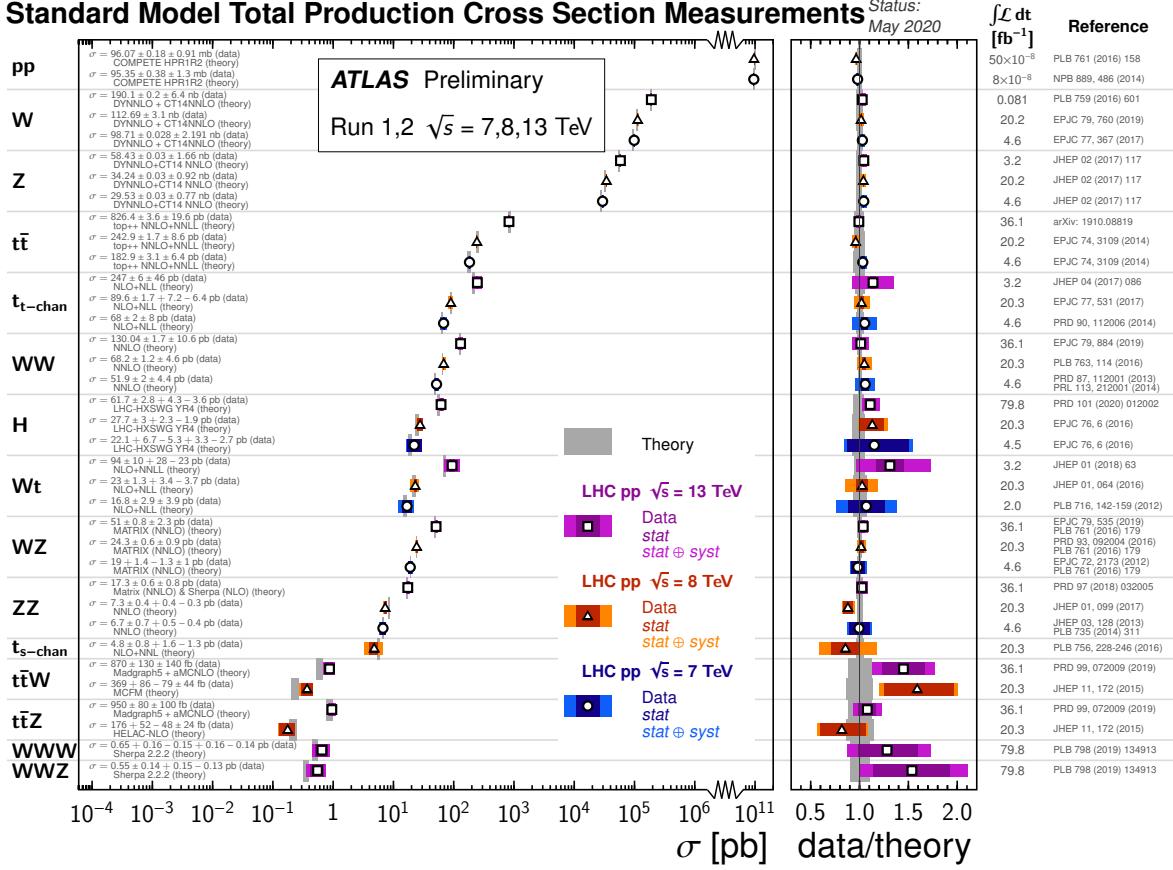


Figure 2.2: Comparison of the measurements and theoretical predictions of the Standard Model process cross-sections.[9] The measurements are all done by experiments at the LHC with $\sqrt{s} = 7, 8$ and 13 TeV. The theoretical predictions are calculated with next-to-leading order (NLO) or higher precisions. From the rare processes ($\sigma \sim 10^{-1}$ pb) to common processes ($\sigma \sim 10^{11}$ pb), data and theory agree fantastically.

After several decades, the measurements of these two bosons have become more precise. In 2020, the latest measurements are [1]:

$$m_W = 80.379 \pm 0.012 \text{ GeV},$$

$$m_Z = 91.1876 \pm 0.0021 \text{ GeV}.$$

In 2012, the last particle predicted by the SM – the Higgs boson – was discovered at the LHC.[8] This means all the mechanisms needed to build the SM are proven to be true. This is certainly one of the most successful theories in physics history.

Although there is no doubt about the success of the SM, some fundamental shortcomings exist. Firstly, it fails to include gravity, which is one of the four fundamental interactions. Since general relativity has also achieved great success, it is believed that the ultimate theory of everything must incorporate general relativity in some way. Secondly, the dark matter candidate is not included in the basic SM. There are many modifications and extensions of the SM with good dark matter candidates, like the axions and some Supersymmetry (SUSY) particles, but none of them has been found. Our benchmark theory models discussed in the second half of this chapter are motivated by this problem.

Except for the gravity and dark matter problems, some limitations are also emerging when we explore further in particle physics. One of the problem is concerning matter anti-matter asymmetry. The SM does not predict strong enough Charge-Parity (CP) symmetry violation, which is required for a source of matter anti-matter asymmetry. One of the main contributions of CP violation predicted by the SM comes from the weak interaction. An extension of the electroweak sector of the SM might be a solution of this problem. This thesis will explore some possible extensions of the SM in the electroweak sector.

2.2 Beyond the Standard Model

In order to search for new particles, a study of theories beyond the standard model (BSM) is necessary. These BSM theories should preserve features and most of conclusions of the SM, while providing possibilities to cover the shortcomings. Such studies can provide both guidance to the most probable new particles, and understanding kinematic and substructure signatures for such events if they are detected. In this thesis, three BSM theories are used as benchmark models to generate simulation of events for target processes. However, the search itself is generic. The main purpose of this search is not to test some specific models, but to look for any anomalies in the $W\gamma$ and $Z\gamma$ mass spectra.

The first theory [12] introduces a new singlet scalar boson similar to the Higgs boson. It was proposed before the SM's 125 GeV boson was observed. It provides an alternative theory model if we observe some resonance similar to the Higgs boson, but with non-standard branching ratios into one or more final states. In this specific model, it can enhance the decay of a singlet scalar into $\gamma\gamma$ and $Z\gamma$ final states by an order of magnitude even if the resonance mass is above the WW kinematic threshold. In this thesis, the model is used to produce a process with a spin 0 resonance decaying into a $Z\gamma$ final state.

The second model [13] used for generating Monte-Carlo events is a heavy vector triplets model. By introducing a new three dimensional real vector corresponds to an adjoint representation of $SU(2)_L$ with vanishing hypercharge, the model predicts a new charged boson and a new neutral boson both with spin one. The charged boson decay into $W\gamma$ is suppressed, but not completely forbidden. Therefore, we use this model to generate our simulated events for a charged boson decay into a $W\gamma$ final state.

The third model [14] is motivated by the fine-tuning issue for the Higgs mass

by constructing a theory that can automatically make the radiative corrections for the Higgs mass to be finite. This theory is built by an $SU(N)$ symmetry in higher space dimensions. Such a model predicts a spin 2 neutral boson which can decay into the $Z\gamma$ final states. We use it to produce two different simulations with either a quark-quark collision or gluon fusion production modes.

Chapter 3

The Large Hadron Collider and the ATLAS Experiment

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the largest proton-proton collider ever constructed in human history. As shown in Fig. 3.1, the LHC is located near Geneva, Switzerland at the European Organization for Nuclear Research (CERN). The tunnel of the LHC is 26.7 km in length, and lies 45 to 170 meters underground. This powerful tool for searching for SM and BSM particles was finally approved in December 1994 by the CERN Council, and built between 1998 and 2008.[15] It has been the energy and luminosity frontier of high energy physics ever since. Currently in the LHC, protons collide with a center-of-mass energy of 13 TeV, with a total integrated luminosity of 156 fb^{-1} delivered by the end of 2018. There are 4 interaction points on the ring. Each of the 4 major experiments, ATLAS, ALICE, CMS and LHCb, are built on one of the interaction points. ATLAS [16] and CMS [17] are two general purposed experiments, searching for all types of phenomena including but not limited to the Higgs search and measurements, electroweak interaction studies, and dark matter searches; LHCb is an experiment focusing on particles and processes containing b and anti-b quarks; and ALICE is designed to study heavy-ion (lead nuclei and more) collisions to understand strong interaction more profoundly.

The tunnel of the LHC was originally built for the Large Electron-Positron Collider (LEP) running from 1989 to 2000, which is famous for its precise measurements of the Z boson. LEP failed to observe evidence of the Higgs boson, a particle pro-

posed by Peter Higgs in 1964 to generate mass to W and Z bosons, and instead set a lower mass bound of 114.4 GeV, barely missing the discovery at LEP. This goal was finally achieved by the LHC in 2012, with the ATLAS and CMS experiments combined to claim the discovery of the 125 GeV Higgs boson. After that, the LHC has become a factory of Higgs boson, and more precise measurements of Higgs features are being conducted. There are also many searches being conducted at the LHC for BSM particles and phenomena before and after the Higgs discovery.

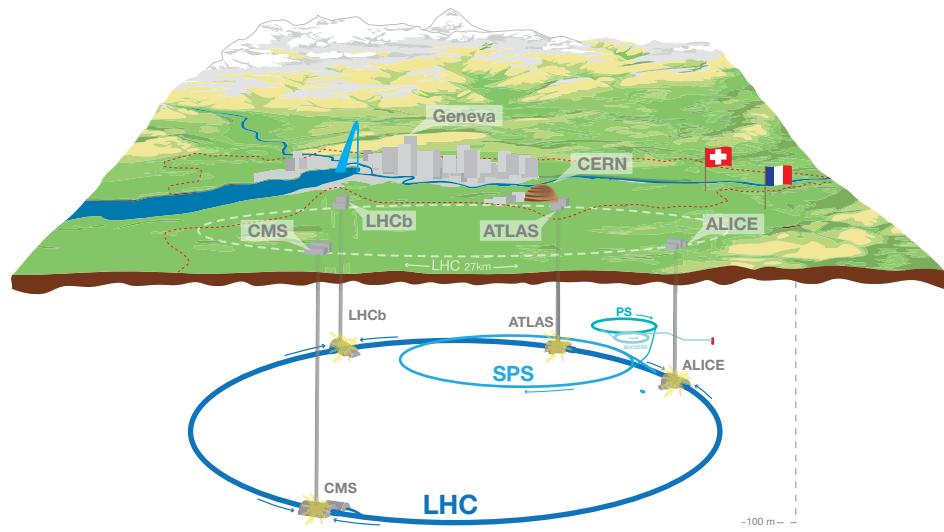


Figure 3.1: A landscape of LHC ring and its surrounding, and the locations of 4 experiments on LHC.

The high center-of-mass energy achieved by the LHC is strongly benefited by innovative technologies in superconductivity and cryogenics. Since the collider geometry is determined by the size of the LEP tunnel, the bending force provided by superconducting magnets will directly determine the collision energy limit the LHC

can reach. The magnetic field must be strong enough to bend the trajectory of protons inside the LHC, making them follow the curvature of the machine vacuum pipe. After careful consideration and validation, with effort both from the physics community and industry, the parameters of 1232 dipole and 392 quadrupole magnets were decided. Unlike older generations of colliders, the LHC used the well-proved NbTi superconductor, but cooled using the latest technology with superfluid helium to a temperature of 2 K, compared to supercritical helium with temperature of slightly above 4.2 K. This allowed the magnet system of the LHC to run at 8 T, compared to other colliders (Tevatron-FNAL, HERA-DESY and RHIC-BNL) that were usually below or around 5 T, and therefore pushed the proton-proton collision energy limit of the LHC to 14 TeV. The design of twin-bore magnets was also adopted by the LHC to enable protons traveling in opposite directions in the ring to collide head to head. [18]

The LHC not only collides particles with the highest center-of-mass energy ever achieved, but also has reached the highest luminosity for proton-proton collisions. The luminosity mentioned here is a number measuring the intensity of interaction collisions. The instantaneous luminosity is defined to be the event rate per cross-section for a process, as defined in Eqn. 3.1, and the integrated luminosity is defined as the time integral of instantaneous luminosity as shown in Eqn. 3.2. The total number of events for a specific process can be predicted by multiplying the integrated luminosity by the cross-section of the corresponding process, $N_{event} = \sigma_{process} L$. For the LHC, the instantaneous luminosity is determined by parameters of colliding beams as shown in Eqn. 3.3. n_b is the number of proton bunches; N_b is the number of protons in each bunch; f_{rev} is the frequency of revolution and γ_r is the relativistic gamma factor. These four parameters combine to estimate the rate of protons running across each other. They are fully determined by the beam of the LHC. There are about

2800 bunches filled into the LHC, with about 1.1×10^{11} protons in each bunch. These bunches are separated by a time interval of 25 ns, which result in a collision rate of 40 MHz. $\epsilon_n = \gamma_r \epsilon$ is the normalized beam emittance; β^* is the beta function at the interaction point and F is a geometric factor related to the beam crossing angle. These three factors are determined by the geometry of the interaction points.[19]

$$\mathcal{L} = \frac{1}{\sigma} \frac{dN}{dt} \quad (3.1)$$

$$L = \int \mathcal{L} dt \quad (3.2)$$

$$\mathcal{L} = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} \times F \quad (3.3)$$

Since modern methodology in high energy physics strongly relies on statistics, evaluating and constraining all results with rigorous statistical treatments, the high luminosity of the LHC improves both the sensitivity of searches, and accuracy of measurements. LHC has reached a peak instantaneous luminosity of approximately $2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, and has delivered a total integrated luminosity of 156 fb^{-1} . (See Fig. 3.2) This high luminosity is achieved with a price of multiple pp collisions happening simultaneously in each proton-proton bunch crossing at the LHC. This is called pileup, and monitored for each of the runs as shown in Fig. 3.3. This pileup information is useful in both monitoring the running status of the collider, and providing collision background information for measurements.

3.2 The ATLAS Experiments

The ATLAS experiment uses a general purposed detector (Fig. 3.4) operating at one of the LHC collision points.[16] The detector started construction in the early 1990s,

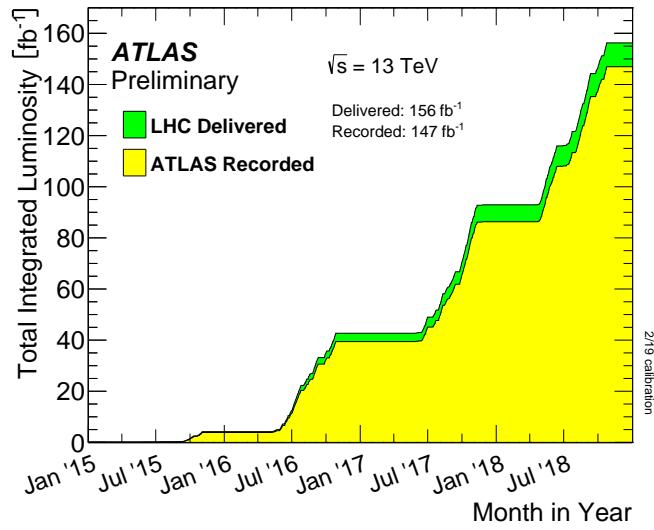


Figure 3.2: The integrated luminosity delivered by LHC machine, and collected by ATLAS detector, versus years.

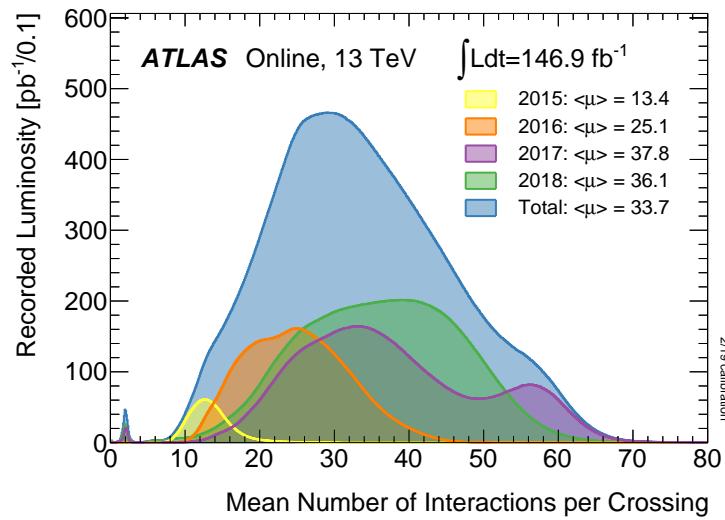


Figure 3.3: The distribution and average value of pileup for each year from 2015 to 2018.

and started taking data in 2009. The first period of operation, called Run 1, started on 20 November 2009, and stabilized in early 2010 to be running at 7 TeV pp collision center-of-mass energy until 2011. Run 1 continued with an energy increase to 8 TeV, and ended in early 2013. The second period of operation, Run 2, running from 2015 to the end of 2018, recorded 147 fb^{-1} of 13 TeV collision data with 139 fb^{-1} good for physics studies. The research described here uses this dataset.

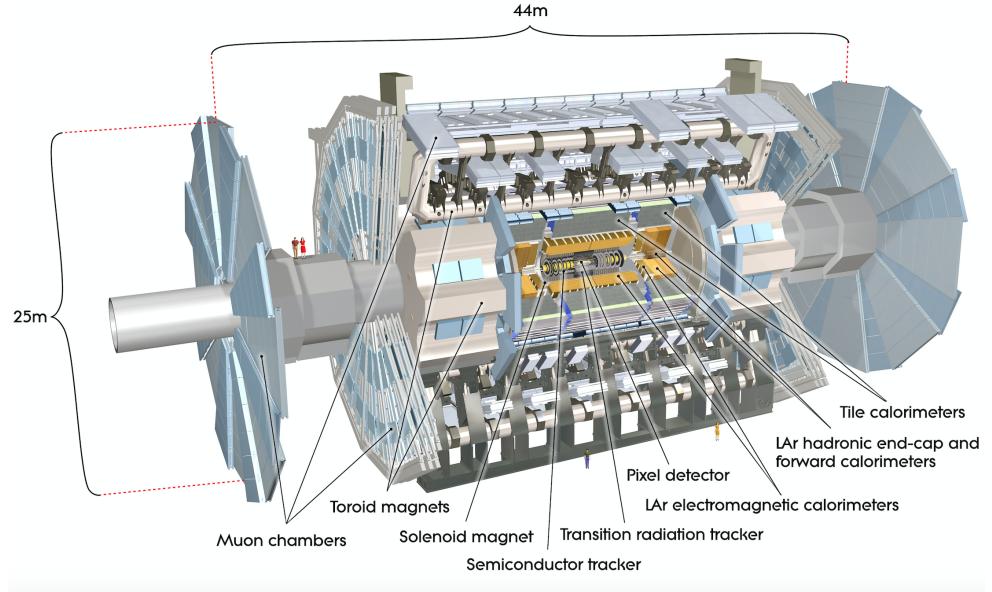


Figure 3.4: An overview of ATLAS detector structure.

The coordinate system used by the ATLAS detector is defined as follows. The x-axis is in the plane of the LHC ring, pointing to the center, the y-axis is defined going up vertically, and the z-axis is along the beam according to the right hand rule. The angle ϕ is defined in the x-y plane, and θ is defined relative to the z-axis. A more frequently used parameter, pseudorapidity η , is defined based on θ as shown in Eqn. 3.4.

$$\eta = -\ln \tan \frac{\theta}{2} \quad (3.4)$$

As a general purposed detector, ATLAS is designed to capture signatures of all possible collision products. Different layers of the ATLAS detector are each sensitive to specific features of outgoing particles. Starting from the point of pp collisions, particles will first go through the Inner Detector, which consist of a silicon pixel layer, a silicon microstrip layer (SCT) and a Transition Radiation Tracker (TRT). These sub-detectors provide information about the tracks and vertices, which is useful for tagging the particle type and measuring the momenta of charged particles. A solenoid magnet producing a magnetic field of 2 T surrounds this Inner Detector. The next layer just outside this solenoid is the electromagnetic calorimeter measuring energy of the outgoing particles. Electrons and photons are expected to be fully absorbed in this layer. The next layers are hadronic calorimeters, where the remaining baryons and mesons showered from quarks and gluons completely deposit their energy. The particles not fully absorbed by these calorimeters, most likely to be muons, will get into the Muon Spectrometers. There are superconducting toroid magnets providing 0.2-3.5 T magnetic field for measuring the momenta of muons. [16]

As mentioned before, different sections of the detector are sensitive to different features of particles, which means each type of particle will leave a unique signature in the detector as shown in Fig. 3.5. In this analysis, we expect to produce photons and W/Z bosons in the final state, with the bosons decaying to $q\bar{q}$ pairs. These $q\bar{q}$ pairs subsequently produce jets of hadrons. The information provided by the Inner Detector tracking system, electromagnetic calorimeters and hadronic calorimeters is utilized to reconstruct these particles. The following sections of this thesis will explain in detail the operation and performance of these detector systems.

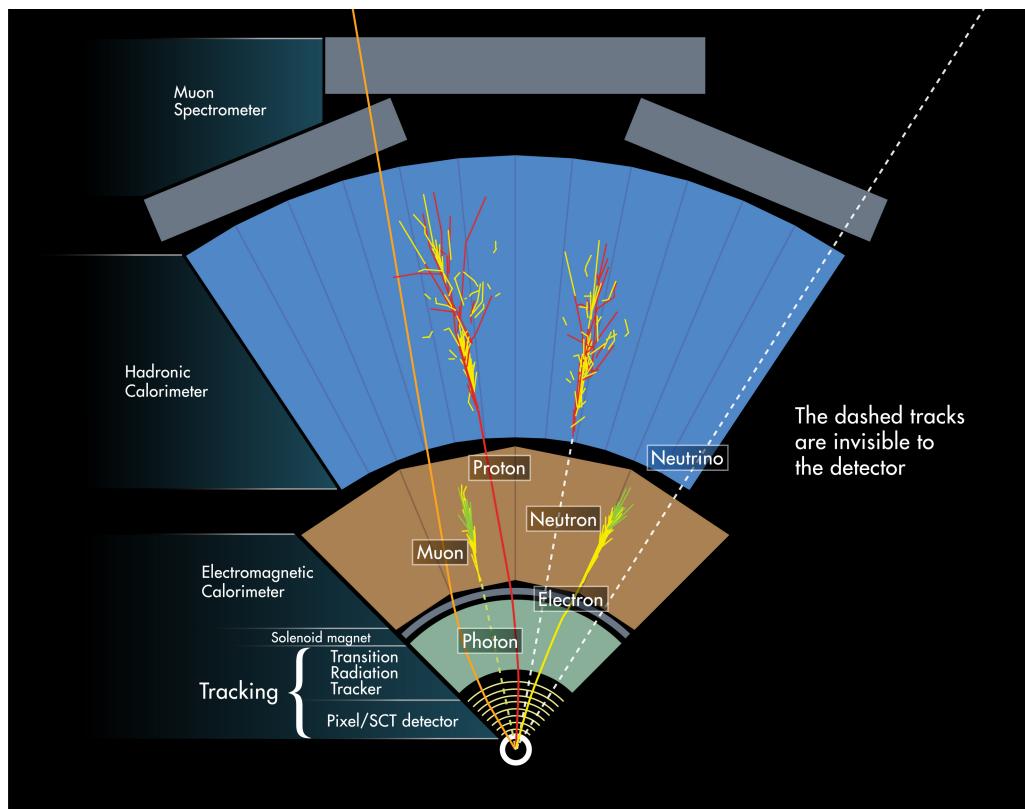


Figure 3.5: An illustration of different particle signatures in ATLAS detector.

3.2.1 The Inner tracking Detectors

The inner detector (ID) contains a silicon pixel layer, a silicon microstrip layer (SCT) and a Transition Radiation Tracker (TRT) as shown in Fig. 3.6. It is the innermost component of the ATLAS detector. The charged particles can be precisely tracked by the ID with the help of a solenoid magnet that produces a 2 T magnetic field to measure their momenta. The reconstructed tracks of charged particles provide crucial information to determine their vertex positions, magnitude of momenta and signs of charge. The TRT also provides particle identification (PID) information for electrons. More details of the layers and geometry of the inner detector are shown in Fig. 3.7, and will be described in following sections.

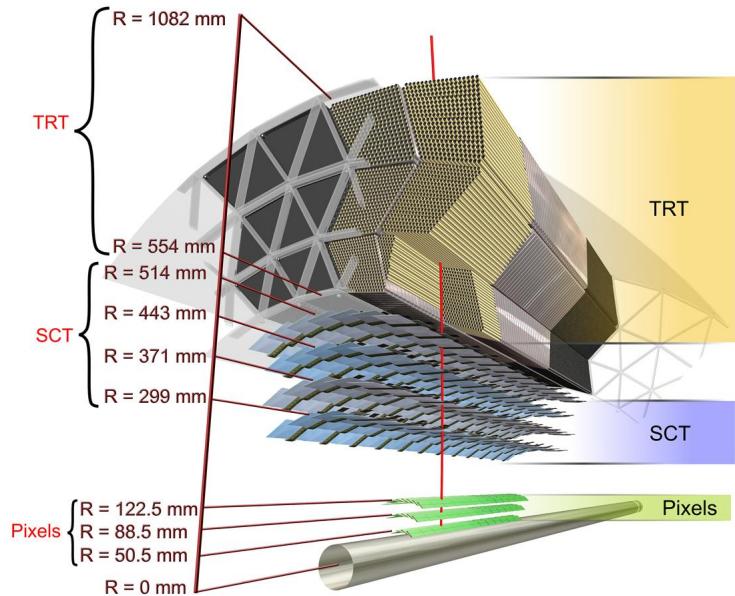


Figure 3.6: The structure of the inner detector

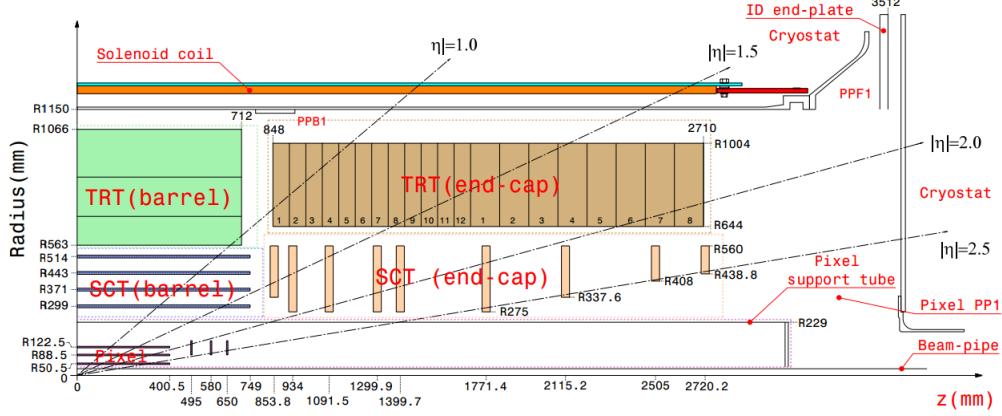


Figure 3.7: The geometry of the Inner Detector. For different pseudo-rapidity, particles will go through different layers of the Inner Detector.

3.2.1.1 Pixel Detector and Semiconductor Tracker (SCT)

The pixels [20] and SCT [21] are both semiconductor detectors. When a charged particle passes through a semiconductor, it leaves ionization. Since there is a biased voltage applied to the semiconductor, the charge carriers (electrons or positive holes) are driven to produce ionization current. This signal is picked up by readout electronics. There are three barrel layers and three endcap disks for the pixels. An extra layer, called the insertable B-layer (IBL), was added in 2014 to improve tracking and b-tagging performance for the increased luminosity of Run 2 and beyond. With the IBL and pixel readouts combined, ATLAS reached an intrinsic accuracy of $10 \mu\text{m} \times 115 \mu\text{m}$ ($r\phi \times z$), and a vertex reconstruction resolution of $11 \mu\text{m}$ in x and y, and $24 \mu\text{m}$ in z.

The SCT provides another four semiconductor layers outside the pixels with double sided strip detector modules. The strips on either side of a module are set at a 40 mrad stereo angle, allowing for 2D reconstruction in the modules. This saves money and readout bandwidth at the cost of accuracy which become $17 \mu\text{m} \times 580 \mu\text{m}$ ($r\phi \times z$). The pixels and SCT endcaps provide tracking coverage up to $|\eta| < 2.5$.

3.2.1.2 Transition Radiation Tracker (TRT)

The TRT [22] is a gas drift tube detector. The basic elements of TRT are polyimide drift (straw) tubes of 4 mm diameter. When a charged particle travels through a straw, it ionizes gas along its track. The ionization clusters drift to a central wire under high voltage with an electromagnetic cascade amplifying the signal. The electric signals are collected on the central wire. The distance from the track to the wire is estimated by the drift time of collected charge. With both barrel and endcap sections, the TRT provides coverage of $|\eta| < 2$, and an intrinsic accuracy of $130 \mu\text{m}$.

In addition to measuring charge particle hits, the TRT also provides particle identification (PID) for electrons. Relativistic charged particles will induce transition radiation (TR) photons when they pass across boundaries with different indices of refraction. In the TRT, polypropylene fibers filled between TRT straws provide many transition boundaries. The energy of TR photons is proportional to the Lorentz factor of the particle $\gamma = E/m$. Electrons, due to their low mass, can induce TR photons with high enough energy to eject electrons from the gas filled in the TRT straws. This produces a much stronger signal than direct charged particle ionization. Therefore, the TRT can provide a very effective electron PID.

3.2.2 Calorimeters

Calorimeters, as implied by their name, measure the energy of particles. There are several different types and designs of calorimeters. The ATLAS calorimeters [23] are built in a classical design, with alternating layers of absorbing material to initiate radiation and active layers to measure the radiation energy. The absorbers make the incoming particles shower into multiple objects with lower energy, and the active layers will be scintillated or ionized by these objects and provide electronic signals related to their energy deposit, which provide enough information to reconstruct the

incoming particles. Multiple sub-detectors either for detecting photons or electrons, which are called electromagnetic (EM) calorimeters, or if more sensitive to hadrons called hadronic calorimeters, are assembled together to build the ATLAS calorimeter system. [24, 25] As shown in Fig. 3.8, different sub-detectors cover different angles for outgoing particles. For example, the EM barrel calorimeter covers the low η region, and the high η region is covered by the EM endcap calorimeter.

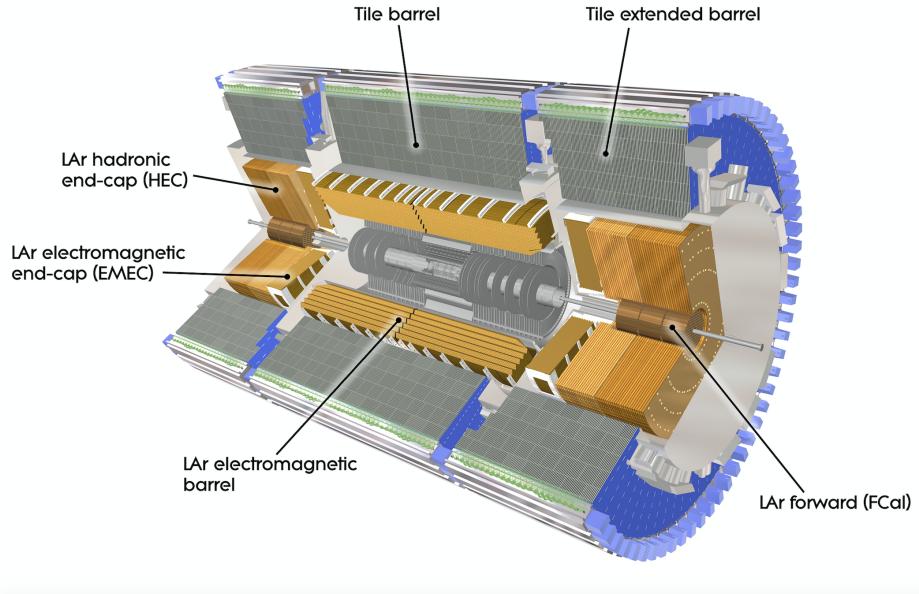


Figure 3.8: An overview of ATLAS calorimeter system.

The electrons, photons and hadrons, all interact in the calorimeters, but have different features and produce different interaction products (see Fig. 3.9). When high-energy electrons enter the calorimeter, they radiate bremsstrahlung photons, which happens whenever charged particles accelerate. On the other hand, photons will induce some electron positron pair production in the mediator. This means high-energy photons or electrons will produce cascades as lower energy electrons and photons, and finally can be fully absorbed by the calorimeter. This process is called an electromagnetic cascade. For a specific kind of material, the radiation length X_0

is defined as the distance traveled by an electron which reduces its energy by a factor of $1/e$. This radiation length is used to evaluate the interaction strength between electrons and the corresponding material, and therefore determines the depth of each layer for the ATLAS EM calorimeter.

The hadronic cascade is similar to the electromagnetic cascade in many aspects. The largest difference is that, hadronic cascade is based on the strong interaction. Therefore, although hadronic showers contain some EM shower components, they also have non-EM components. The electrons, photons and neutral pions present in a hadronic cascade process are absorbed as EM showers. The other non-EM components like protons, neutrons, charged pions and kaons. will inelastically collide with nucleus in absorber material, excite these nucleus into unstable state, making them radiate more mesons, baryons and photons for further interactions. Both of the components combined make a full hadronic cascade, and its ratio strongly varies event by event based on the decay products and energy distribution of the initial hadron. Unlike EM components, non-EM components contain some invisible energy. There is energy stored in the nuclear binding energy, the production of neutrinos and high energy muons, and the kinetic energy left in absorber material. All these energy components cannot provide signals in the active layers, and therefore become effectively invisible to the hadronic calorimeter. A large portion of energy from hadrons, sometimes as high as 30% to 40%, are of this type. Although a heavy effort is made to compensate and calibrate the hadronic calorimeter performance, the precision of the hadron energy measurements is lower than that of the EM calorimeter.

[26]

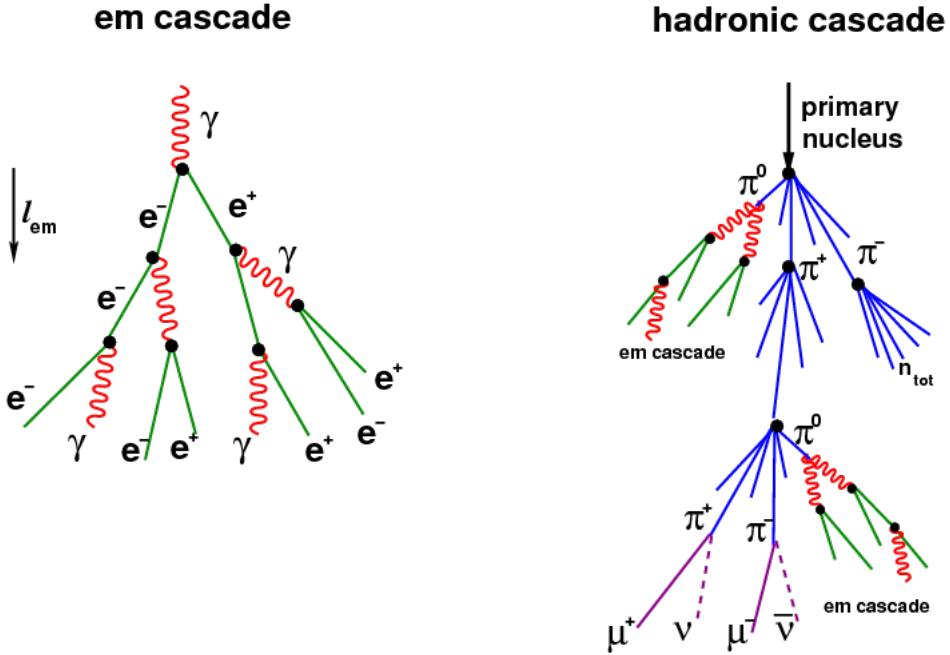


Figure 3.9: A simple demonstration of both electron/photon and hadron showering cascade happening in calorimeters.

3.2.2.1 The ATLAS Electromagnetic Calorimeter

The EM calorimeter (ECAL) [24] in ATLAS detector is built in three parts, the barrel calorimeter covering η less than 1.475, and the two endcap calorimeters covering η between 1.375 and 3.2. All of them are liquid Argon (LAr) sampling calorimeters, with lead as the absorber and LAr in the active layers. When traveling in the lead layer, electrons and photons shower, losing their energy, and when they enter the LAr layer, they will ionize LAr atoms. The electrons are drifted to electrode plates, and then recorded as electric signals.

The alternating layers of ECAL are in accordion-like shape to provide both phi symmetry and complete azimuthal coverage without cracks (see Fig. 3.10). They are divided into two or three different layers depending on the location of the detector. In each layer, calorimeters consist of tower-like cells with different $\Delta\phi \times \Delta\eta$ granularities. For example, in the barrel calorimeter, the first sampling layer as il-

lustrated in Fig. 3.10, has a extremely fine resolution in η to distinguish showers induced by different kinds of particles. This layer is constructed with strips in size of $\Delta\eta \times \Delta\phi = 0.0031 \times 0.098$ with the depth of $4.3 X_0$. The second sampling layer is $16 X_0$ in depth, which collects most of energy from electromagnetic showers. The angular size of this layer is increased to $\Delta\eta \times \Delta\phi = 0.025 \times 0.0245$ because of the cone size spread during cascade process. At the third sampling layer, where only the highest energy photons and electrons could reach, the cell size is doubled in η direction without any cost of resolution. The depth of this third layer is only $2 X_0$. The thickness of the ECAL in total is about $22 X_0$, so that showers of electrons and photons are usually fully absorbed and measured. The resolution of the ECAL, shown in Eqn. 3.5, is energy dependent, and is about 1% for 100 GeV electrons.

$$\frac{\sigma(E)}{E} = \frac{10\%}{\sqrt{E(\text{GeV})}} \oplus 0.17\% \quad (3.5)$$

3.2.2.2 ATLAS Hadron Calorimeter

Similar to the ECAL, the ATLAS hadron calorimeter (HCAL) is divided into several parts. There is a scintillating tile barrel hadronic calorimeter [25], with two extended pieces, covering eta less than 1.7, and two end-cap hadronic calorimeters [24] covering eta from 1.5 to 3.2. The end-cap hadronic calorimeters use LAr like the ECAL, but copper as absorber layers instead. In the central part of the HCAL system, plastic scintillating tiles are used as the active material and steel as the absorber (see Fig. 3.11). The hadrons traveling through this calorimeter will scintillate the plastic tile and produce ultra-violet (UV) radiation. The produced UV light is shifted to visible wavelength with organic fluorescence embedded in the plastic tiles, collected with fiber optic cables, and measured by photomultiplier tubes on the outer edge of HCAL. There are three layers in the barrel, and four in end-cap. Due to the

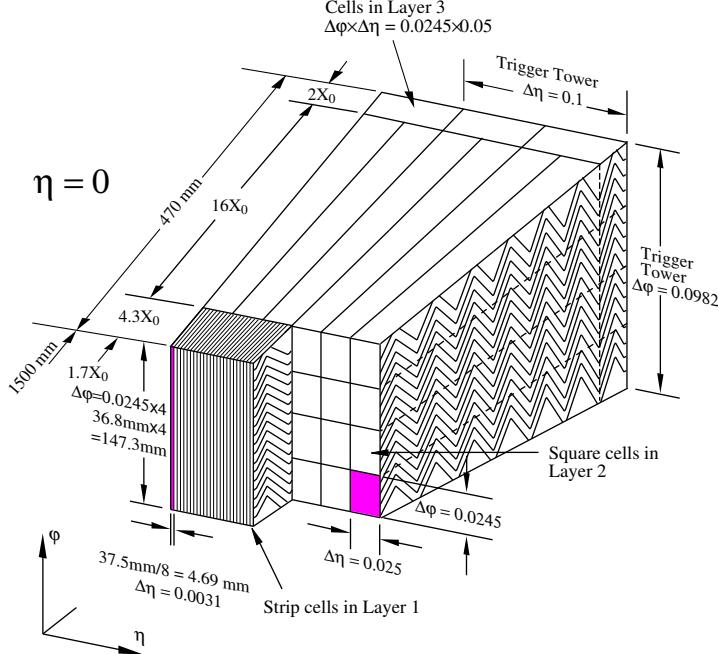


Figure 3.10: ATLAS ECAL barrel Calorimeter geometry.

missing energy in hadronic cascade process, the resolution of this HCAL, as shown in Eqn. 3.6, is worse than ECAL. For a 350 GeV pion in central region, the resolution is about 6.6%.

$$\frac{\sigma(E)}{E} = \frac{56.4\%}{\sqrt{E(\text{GeV})}} \oplus 5.5\% \quad (3.6)$$

3.2.3 ATLAS Trigger and Data Acquisition System

The ATLAS detector collects the proton-proton bunch collision information at the rate of 40 MHz. This provides high luminosity with the cost of high pile-up, and also heavy load on the data reading bandwidth. Each of the events can produce many physical objects, such as electrons, photons, muons, jets and missing transverse energy (MET). As a result, the data recorded is 1-2 MB for a single event, and the

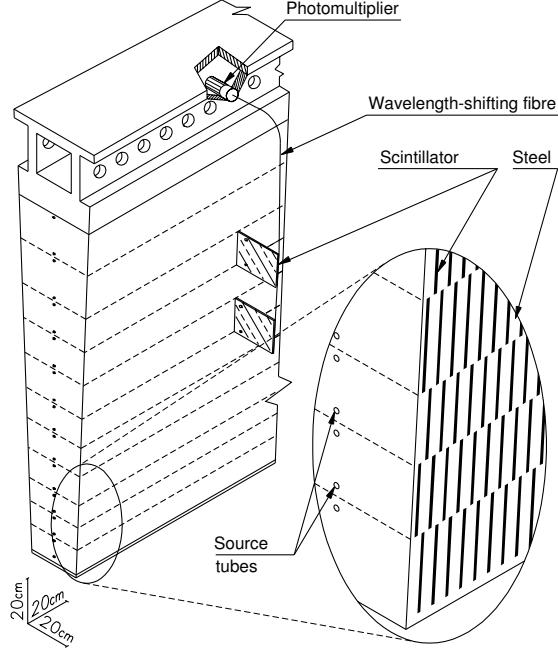


Figure 3.11: ATLAS Hadron Tile Calorimeter geometry.

total processing rate will be 40-80 TB/s without any selection. Therefore, the collision information must be selected in real-time. The trigger and data acquisition system (TDAQ) is designed to do this job. There are three layers of the trigger system called Level-1 (L1), Level-2 (L2) and an event filter. L2 and the event filter together are called High-Level Trigger (HLT).

The L1 trigger select events with high- p_T objects (photons, electrons, muons and jets), large MET or large total transverse energy. The L1 trigger uses reduced-granularity information from a subset of detectors to decide the regions of interest (RoI). The event rate is reduced from 40 MHz to 100 kHz after L1 trigger selection. The L1 decision reaches the front-end electronics in 2.5 μ s. In order to improve the performance of the L1 trigger, the identification and reconstruction must be done in clever ways. For example, the electron/photon (e/γ) reconstruction algorithm is called a sliding window algorithm. As shown in Fig. 3.12, some fixed window (2×2) is

used to scan through EM calorimeter towers looking for two neighboring towers that have a total energy deposit over some pre-defined threshold. Isolation-veto thresholds are set for the 12-tower around the ring in the EM calorimeter, as well as for the 2×2 hadronic-tower core sum behind the cluster and the surrounding 12-tower hadronic ring.

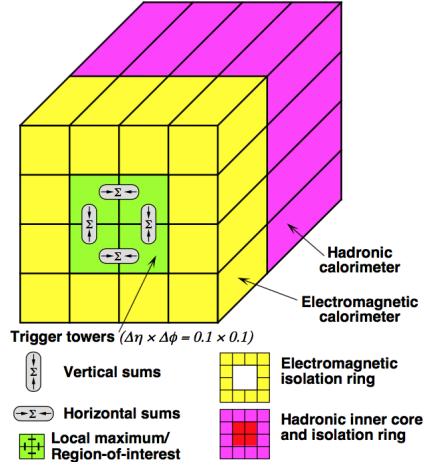


Figure 3.12: Sliding window algorithm for e/γ L1 triggers.

The L2 trigger uses RoI information decided by L1 to obtain additional information from the detector readout, and makes further selection. This reduces the event rate further to below 3.5 kHz, with an average event processing time of approximately 40 ms. The event filter further selects events to a rate which can be recorded with the fully built event information. It reduces the event rate to about 1 kHz, with an average event processing time in order of seconds.

Most of the complex algorithms are applied at the HLT level due to its longer processing time. For e/γ identification, a ringer algorithm is introduced in Run 2 to replace an older cut-based method. As shown in Fig. 3.13, this algorithm makes decision based not only on the information from each calorimeter cell, but the sum of energy deposit in each of the rings. The total energy deposits in each of the rings are

sent into a Neural Network to make particle identification. This new algorithm provides basically the same efficiency but 2-3 times greater rejection of the background. The time spent for processing is almost halved due to the background removal at an early stage of the HLT. [27]

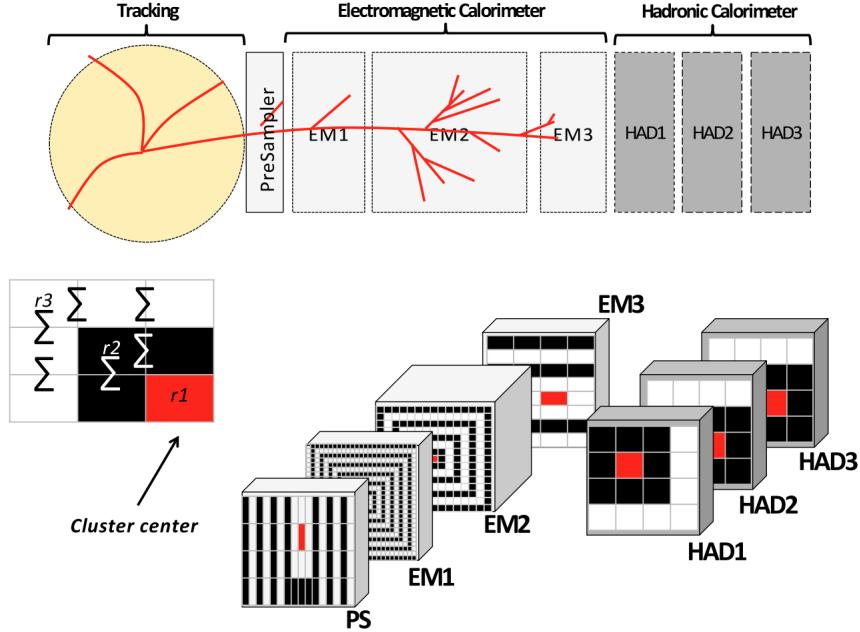


Figure 3.13: Ringer algorithm for e/γ HLT identification.

In summary, the data acquisition system (DAQ) receives and buffers the event data at the L1 trigger rate. Then this DAQ transmits any data requested by the L2 trigger. For those events passed the L2 trigger, event-building starts. The reconstructed events are then sent to the event filter by the DAQ, and the events passed the event filter can finally be recorded for physics studies at a rate of 1 kHz.

Chapter 4

Analysis Methods

4.1 The measurement of photons and W/Z bosons

The search presented in this thesis looks for new heavy boson decaying into a final state with a high energy photon and a strongly boosted W/Z boson which decays into $q\bar{q}$ pairs. In the ATLAS detector, high energy photons are directly identified, and the products of hadronic W/Z boson decays are reconstructed as a shower of particles called jets. The reconstruction of these objects utilizes information collected by both the inner detector and calorimeter systems described in Chapter 3. The methods of reconstruction are described in this section.

4.1.1 Photon identification and reconstruction

The reconstruction of photons utilizes both calorimeter and tracking information from the inner detector. The full reconstruction has several steps. An algorithm called the “ringer” algorithm described in section 3.2.3 uses calorimeter information to provide rectangular area cells containing the local maximum energy cells in the calorimeter. The positions of these clusters are matched to tracks from the inner detector, and determine the particle hypothesis among electrons, photons and converted photons. After the particle hypothesis is determined, the cluster size is rebuilt based on the type of particle, and cluster calibrations are applied accordingly. If there is a track matched to this cluster, the direction is determined by this track, otherwise the position of the cluster will be used to determine the photon direction. Finally, some discriminant variables are used to reject backgrounds. About half of the photons are

converted to electron-positron pairs before reaching the calorimeter, so the converted vertex reconstruction is critical. A good efficiency for conversion radius below 80 cm is achieved by considering not only the double-track conversions, but also single-track conversions, where one of the electron or positron is not reconstructed in the inner detector tracking system.

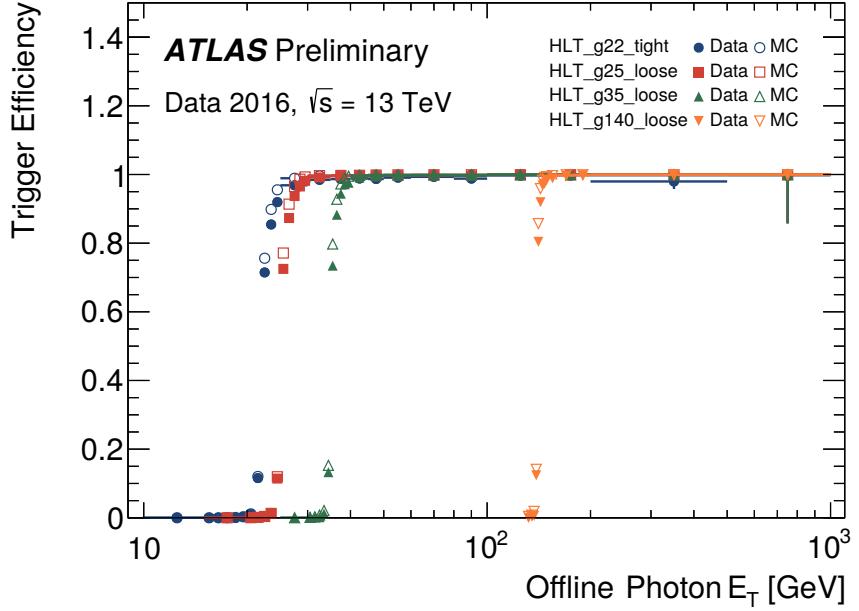


Figure 4.1: The orange inverted triangle dots shows the HLT_g140_loose photon trigger efficiency. This photon trigger is applied in pre-selection for our measurement. There is good agreement between data (solid symbols) and the simulation (open symbols) for the trigger performance.[28]

The ATLAS trigger system is described in section 3.2.3. We use the HLT_g140_loose trigger for event pre-selection. The transverse energy threshold for the photon is 140 GeV, and the photon identification criteria here is relatively loose, which requires less stringent application of discriminant variables. The performance of this trigger for selecting photons is shown in Fig. 4.1.

4.1.2 Reconstruction of boosted jets

Unlike leptons and photons, which leave relatively clean signals in the detector, gluons and quarks have more complicated signatures, and need special treatment to reconstruct. These colored particles will shower into many softer colorless hadrons. In order to reconstruct the original quark or gluon, an algorithm clustering the shower produced by the quark/gluon hadronic products is needed.

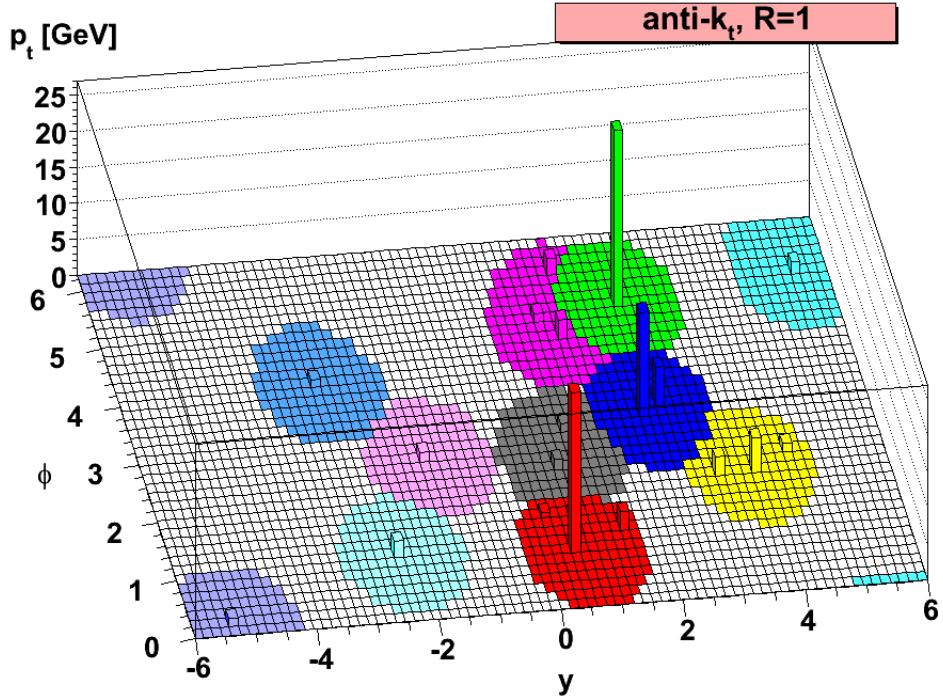


Figure 4.2: A simulated parton-level example together with some random soft objects called “ghost”, clustered by anti- k_t algorithm. As shown in figure, high energy objects are reconstructed with a good shape, not affected by the choice of ghosts, which implies the stability of this anti- k_t algorithm. More discussion and comparison with other algorithms are made in [29].

The clustering algorithm for jet reconstruction used in ATLAS is called the anti- k_t algorithm.[29] This is named after the function used to measure the distance between candidate objects produced in the hadronic shower. This definition of the distance for

different jet clustering algorithms can be written in a general form shown in Eqn. 4.1. A measure of distance to beam is also defined as Eqn. 4.2. i and j denotes two different cells or subjet candidates. k_t means transverse momentum of a calorimeter cell. ΔR_{ij}^2 is defined as $\phi_{ij}^2 + \eta_{ij}^2$. R_{jet} is a parameter used to define the cone size of final jets. The so-called anti- k_t algorithm takes the parameter p to be -1.

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta R_{ij}^2}{R_{jet}^2} \quad (4.1)$$

$$d_{iB} = k_{ti}^{2p} \quad (4.2)$$

When the clustering algorithm starts, all the d_{ij} and d_{iB} values are computed. If the minimum one of these values is some d_{ij} , i-th and j-th subjet will be clustered into one, if it is d_{iB} , the i-th subjet is identified as a different jet. This process will be repeated until all candidate objects are assigned to jets. Note in this process, if $\Delta R_{ij} > R_{jet}$, d_{ij} must be larger than d_{iB} or d_{jB} depending on which object has smaller transverse momentum. This guarantees a well defined maximum jet cone size R_{jet} . $R_{jet} = 0.4$ is usually used for reconstructing quark and gluon jets, but sometimes two collinear partons from a boson decay become interesting, and a larger cone jet with $R_{jet} = 1.0$ is useful. $R_{jet} = 0.2$ is used to reconstruct subjets within these large cone jets to further reveal substructure, and therefore the original source of these large cone jets.

In this search, a very broad mass range of $W\gamma$ and $Z\gamma$ is explored, which means the energy of the boosted W and Z jets can be extremely large. For these strongly boosted jets, the angular resolution is not enough to study the jet substructure. Therefore, a new technique called Track-Calorimeter Clusters (TCC) is introduced. [30]

The Track-Calorimeter Clusters jets are reconstructed by including the track information. Since the angular resolution of the tracker is much better than the calorimeter, by

matching tracks to clusters of calorimeter cells, the angular information of objects in jet clustering algorithm is more accurate. This improves the performance of reconstruction for both the jet mass and the jet substructure used to identify W/Z boson hadronic decay products.

4.1.3 Identification of boosted W/Z boson jets

In order to discriminate large cone jets originating from W/Z boson with other similar jets, some jet substructure variables are defined. When a vector boson is very energetic, it decays hadronically into two quarks traveling collinearly, making a so-called two-prong jet. The relative positioning of the small-cone jets become a very strong discriminating feature, and n-point energy correlation functions provide a well-defined measure of such feature. The 1-point and 2-point functions are defined as shown in Eqn. 4.3 and Eqn. 4.4. The variables p_i in these functions are the momentum four-vector and E_i is the total energy of the i-th small-cone jets. E_J is the total energy of all the objects in the large-cone jet.

$$\epsilon_2^{(\alpha)} = \frac{1}{E_J^2} \sum_{i < j}^N E_i E_j \left(\frac{2p_i \cdot p_j}{E_i E_j} \right)^{\alpha/2} \quad (4.3)$$

$$\epsilon_3^{(\alpha)} = \frac{1}{E_J^3} \sum_{i < j < k}^N E_i E_j E_k \left(\frac{2p_i \cdot p_j}{E_i E_j} \frac{2p_i \cdot p_k}{E_i E_k} \frac{2p_j \cdot p_k}{E_j E_k} \right)^{\alpha/2} \quad (4.4)$$

According to the work of Andrew J. Larkoski, Ian Moult, and Duff Neill [32] [33], a variable $D_2^{(\alpha)}$ is defined as Eqn. 4.5. In ATLAS, the value of α is usually chosen to be 1, and the variable $D_2^{(\alpha=1)} = D_2$ is used as convention. This variable takes small values for a two-prong jet and large values for a one-prong jet. The performance of this variable is shown in Fig. 4.3.

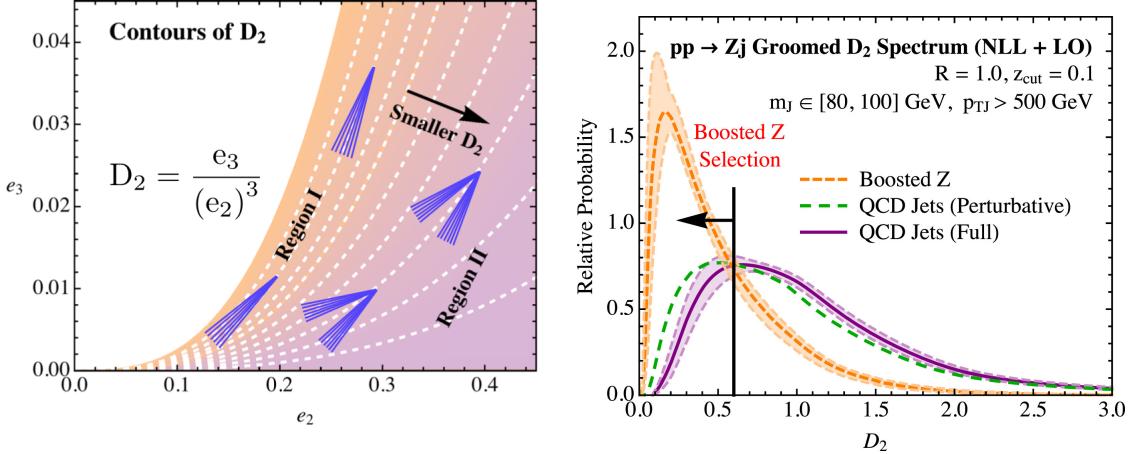


Figure 4.3: The left figure shows the D_2 variable contours in (e_2, e_3) phase space for 1-prong jets (Region I) and 2-prong jets (Region II). The right figure shows the separation power of D_2 variable for boosted Z boson jets from the background jets produced by QCD processes. The estimation for QCD jets in dashed line shows the prediction made by perturbation calculation, and the solid line shows prediction including non-perturbative corrections. [31]

$$D_2^{(\alpha)} = \frac{e_3^{(\alpha)}}{(e_2^{(\alpha)})^3} \quad (4.5)$$

The boosted boson tagging in this search utilizes the jet mass and D_2 variables. This tagging criteria is optimized on boosted W and Z boson induced jets produced from a simulated sample of $W' \rightarrow WZ$. The cut values for jet mass and D_2 as functions of jet p_T are shown in Fig. 4.4. The improvement with a new TCC algorithm, compared to the old jet clustering algorithm using locally calibrated topological clusters called LCTopo, is shown in Fig. 4.5. [30]

4.1.4 Identification of jets from beauty hadron decays

The jets from beauty hadron decays have special features that differ from gluon jets or other light-quark jets. Almost 70% of Z bosons decay hadronically, and about 15% out of all possible decay modes, the final products are a pair of $b\bar{b}$ quarks. For other

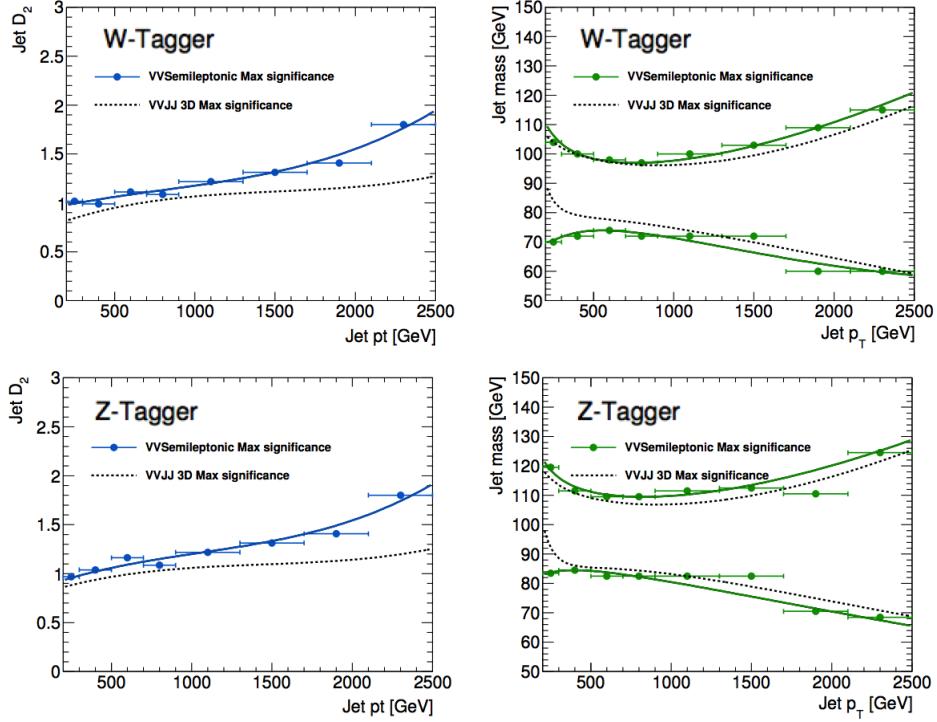


Figure 4.4: These plots show the cut functions for TCC jets tagging for W and Z bosons. The solid lines are the selections used in this thesis. The lines in the left two plots show the threshold for D_2 variables for W/Z boson tagging, with low values accepted and high values rejected. The lines in right two plots shows the jet mass window cut for W/Z boson tagging in different jet p_T regime. The solid lines in these plots are optimized for the significance of a heavy diboson resonance search in semi-leptonic final states (one of the W/Z boson decays hadronically and the other decays leptonic), and the dashed lines (together with a cut on number of tracks) for a heavy resonance search from two boosted jets both produced by W/Z bosons. Plots are made by the ATLAS VV semi-leptonic search group using ATLAS simulated $W' \rightarrow WZ$ samples in [34].

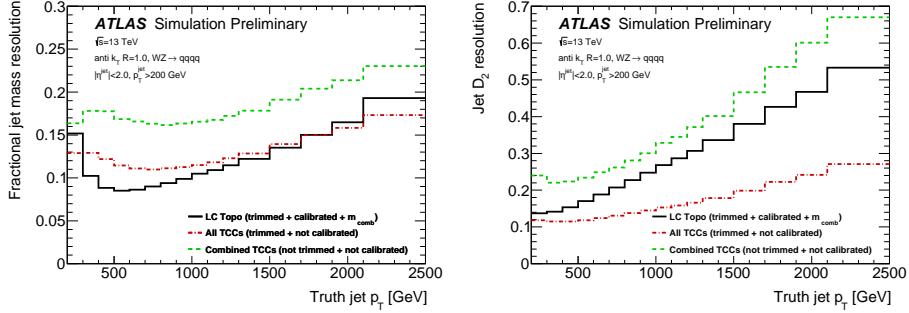


Figure 4.5: These plots are the comparison of performance between TCC jets in dashed red lines and LCTopo jets in solid black lines (using only calorimeter information for jet reconstruction). The left one is for the jet mass resolution, and the right one is for jet substructure variable D_2 resolution. [30]

sources of boosted jets, mostly from QCD processes, the fraction of $b\bar{b}$ final states are much less than from Z bosons. Due to the power of background rejection from beauty hadron identification, this can significantly improve selection of boosted jets from Z boson decays.

In the ATLAS experiment, the b-jets are identified by combining three different algorithms: an impact parameter-based algorithm, an inclusive secondary vertex reconstruction algorithm, and a decay chain multi-vertex reconstruction algorithm.[35] The first two of these algorithms are based on the fact that beauty hadrons have a long lifetime, and the last one attempts to reconstruct the full decay chain of the beauty hadron.

The impact parameter is defined as the distance of the charged track extrapolated to the primary vertex. For each track, there is a transverse impact parameter (d_0) defined in the $r - \phi$ plane and a longitudinal impact parameter (z_0) in the z direction. Due to the long lifetime of b-hadrons, large impact parameters of tracks usually increase the chance of finding b-jets. By combining all tracks' contributions with a log-likelihood ratio function $\sum_{i=1}^N \log \frac{p_b}{p_u}$, a score of this impact parameter-based b-tagger can be used for the next step. N is the number of tracks associated to the

jet, p_b and p_u are the template probability density functions (PDF) for b-flavour and light-flavour jet hypotheses. The assumption of no correlations among the various tracks in the jet is made.

The secondary vertex finding algorithm explicitly reconstructs potential secondary vertices by looking at every pair of tracks. All these potential secondary vertices are checked carefully based on their silicon hits (sum of pixel and SCT hits), kinematic variables and impact parameters. They will be rejected if they are likely to be a product of long-lived particles, photon conversions or hadron interacting with detector material. The highest p_T candidate vertex will be reconstructed as the secondary vertex.

The decay chain multi-vertex algorithm uses a Kalman filter to find a common line through the primary vertex and the bottom/charm vertices, and assuming it to be the fight path and position of the b-hadron. This provides extra information so that with a single track attached to the vertex, it can still be found.

Both the secondary vertex finding algorithm and the decay chain multi-vertex algorithm provide information about displaced vertices, and the corresponding kinematics information for tracks associated to them. Combining this information with the log-likelihood ratio from the impact parameter algorithm, a Boosted Decision Tree (BDT) is trained to optimize b-tagging performance. The final output of this b-tagger is the BDT score. This full b-tagging algorithm is named as MV2c20, which is upgraded from MV1 algorithm used in Run-1. Not only the improvement on performance is significant, MV2c20 algorithm also simplified the multi-variable training steps by directly taking input from each of the algorithms. Fig. 4.6 shows the discriminating power of this MV2c20 b-tagger. There are four operating points for the MV2c20 b-tagging algorithm with different efficiency provided as shown in Table. 4.1, and in this thesis, the 60% efficiency operating point is applied. More details of the

full b-tagging algorithms in ATLAS are described in [35].

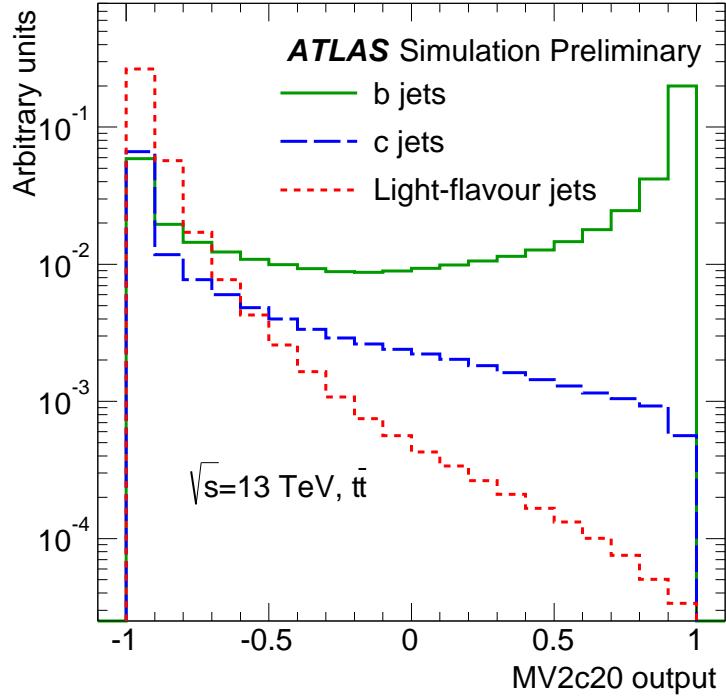


Figure 4.6: This plot shows the performance of the MV2c20 algorithm used to discriminate beauty hadron decay products. The training of this BDT is performed on a set of approximately 5 million $t\bar{t}$ events, which contains beauty hadron showered jets (b jets), charm hadron showered jets (c jets), and jets showered from other light-flavour hadrons (light-flavour jets). The b-jets are assigned as signal, and a mixture of 80% light-flavour jets and 20% c jets as background. As we can see in the plot, this MV2c20 b-tagger provide a very strong rejection power for c jets and light-flavour jets, while keeping most of the b jets in the high score area.[35]

Table 4.1: Operating points provided for the MV2c20 b-tagging algorithm, with corresponding efficiency and rejection (reciprocal of efficiency) rates shown in this table. The statistical uncertainties for these efficiency and rejection rates are negligible. [35]

Efficiency	Cut Value	c-jet Rejection	τ -jet Rejection	Light-jet Rejection
60%	0.4496	21	93	1900
70%	-0.0436	8.1	26	449
77%	-0.4434	4.5	10	140
85%	-0.7887	2.6	3.8	28

4.2 Statistical Methods

4.2.1 Statistical treatment in High Energy Physics

In high energy physics experiments, the physical results of interest are usually buried within huge amount of experimental data collected, and can only be extracted by some statistical treatment. Furthermore, the level of belief we can have on a result is sometimes as important as the result itself. This section will provide the details of the statistical treatment of the data used in this thesis.

In general, a search for certain physical processes or objects can be formalized as a frequentist statistical test. Theories provide hypothesis, data provides observation, and with proper statistical treatment, results with physical meanings can be extracted. Depending on whether the search targets a discovery or an exclusion, the final result can be either a significance or a limit. The significances or limits are essential criteria for a discovery or an exclusion. However, they cannot solely decide the validity of a theory or hypothesis. The merit of signal hypothesis and the compatibility between data and theory will also be taken into consideration. In this thesis, both the significance and the upper limit on production cross-section are calculated. The observation is tested between a background-only hypothesis considering all known SM physics processes, and a signal plus background hypothesis including also the BSM process we are searching for. For limit setting, the signal plus background hypothesis is treated as the null hypothesis H_0 , and the background-only hypothesis

is treated as the alternative hypothesis H_1 . For calculation of a signal significance, the background only hypothesis is treated as null hypothesis H_0 , and the signal plus background hypothesis is used as the alternative hypothesis H_1 . [36]

The direct result from a statistical test is a probability called p-value. This p-value is the probability of finding an observation with equal or greater incompatibility with the predictions made by the hypothesis. The hypothesis is regarded as excluded if the corresponding p-value below a certain threshold. Another value mentioned frequently call significance Z is defined solely based on the p-value as shown in Eqn. 4.6 and Figure. 4.7.

$$Z = \Phi^{-1}(1 - p) \quad (4.6)$$

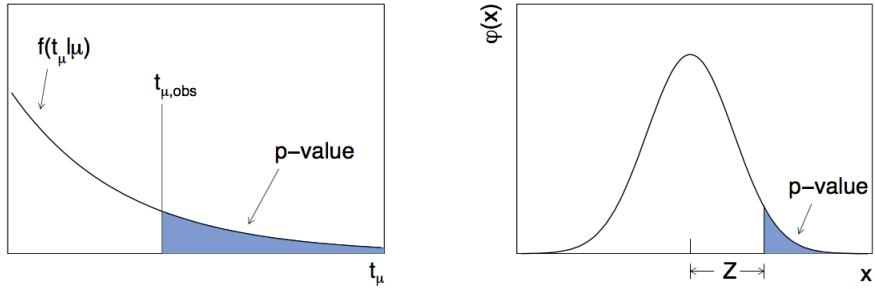


Figure 4.7: The left figure shows some random distribution $f(t_\mu | \mu)$, and corresponding p-value for the specific observation $t_{\mu, \text{obs}}$. The right figure shows how this p-value can be translated into significance by a standard normal distribution function $\phi(x) = (1/\sqrt{2\pi})\exp(-x^2/2)$. The signal strength μ is defined later in Eqn. 4.7, and test statistics t_μ in Eqn. 4.12.

The function Φ^{-1} here is the quantile (inverse of the cumulative function) of a standard Gaussian distribution. Although the probability distribution is not always Gaussian, this function translates the p-value directly into a significance Z for a Gaussian distribution, therefore providing an intuition of how much the observation deviates from the hypothesis. For example, a significance Z larger than 5 is usually

regarded as the criteria for discovery with a corresponding p-value of 2.87×10^{-7} . In order to set limits, a threshold of p-value of 0.05 (95% confidence level) is chosen, which corresponds to a significance $Z = 1.64$. An important feature of equation 4.6 is that, the significance Z decreases monotonically when the p-value increases. Although the relationship between p-value and Z is non-linear, the median of Z is always the corresponding significance of the median of the p-value. Since the median of the p-value can also be determined from the median of observation, the median of the expected significance can be calculated directly from the median of the observation. This median of expected significance is usually reported as the sensitivity of the experiment.

The frequentist significance test is started by choosing functions to describe signal and background processes. These functions contain not only the parameters of interest like the rate (cross section) of the signal and background, but also other nuisance parameters which are not fixed a priori. Some of these nuisance parameters are constrained by the experimental systematic uncertainties, the others are determined by fitting the signal and background functions to the simulation or the data. The functions need to be sufficiently flexible so that the true signal and background exist somewhere within the phase space allowed for the nuisance parameters. The extra flexibility introduced by these nuisance parameters results in a loss of sensitivity, while if the flexibility is not enough, an additional systematic uncertainty will be added to the measurement. In this thesis, this extra uncertainty is determined by a method widely used in the ATLAS experiment called a spurious signal test which will be discussed later in this section.

To compare the compatibility of the observation to different hypotheses, a well-defined measure of compatibility is necessary. In high energy physics, this compatibility can be defined with a likelihood ratio function. The first step is to build a

likelihood function for a given hypothesis. Assume in a binned analysis, which divides events into bins based on some kinematic variable, the expected number of events falling into the i -th bin can be expressed as Eqn. 4.7.

$$\langle n_i \rangle = \mu s_i + b_i \quad (4.7)$$

The variable μ here means the strength of the signal process. $\mu = 0$ implies background-only hypothesis, and $\mu = 1$ is the nominal signal plus background hypothesis. s_i and b_i are expected number of events from signal and background, each determined by a set of nuisance parameters as shown in Eqn. 4.8.

$$\begin{aligned} s_i &= s_{tot} \int_{\text{bin } i} f_s(x; \theta_s) dx \\ b_i &= b_{tot} \int_{\text{bin } i} f_b(x; \theta_b) dx \end{aligned} \quad (4.8)$$

The s_{tot} and b_{tot} are total amount of events from signal and background. The distribution of signal and background are each determined by a set of nuisance parameters θ_s and θ_b . Here we define $\theta = (b_{tot}, \theta_s, \theta_b)$ to denote all the nuisance parameters. The s_{tot} is related to the cross-section of the signal process, which is the parameter we are measuring, so it is not included in the list of nuisance parameters.

The likelihood of actually observing n_i events in that i -th bin follows a Poisson probability as Eqn. 4.9. The θ here represent all nuisance parameters in this equation, including those in signal and background functions used to estimate s_i and b_i .

$$L_i(\mu, \theta) = \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)} \quad (4.9)$$

The full likelihood function can be built by directly multiplying all bins together as Eqn. 4.10.

$$L(\mu, \theta) = \prod_{i=1}^N L_i(\mu, \theta) P(\theta) \quad (4.10)$$

The extra term $P(\theta)$ here is the probability distribution for different nuisance parameters, such as gaussian distributions for some of the systematic uncertainties. The corresponding profile likelihood ratio used to measure incompatibility of a hypothesis is defined as Eqn: 4.11.

$$\lambda(\mu) = \frac{L(\mu, \hat{\theta}(\mu))}{L(\hat{\mu}, \hat{\theta})} \quad (4.11)$$

The $\hat{\mu}$ and $\hat{\theta}$ in the denominator are the variables that maximize this likelihood function, called the (unconditional) maximum-likelihood estimators of μ and θ . The $\hat{\theta}(\mu)$ in numerator is the conditional maximum-likelihood estimator of θ if μ is already fixed. This makes $0 \leq \lambda \leq 1$, and larger λ means better agreement. Therefore a test statistics called t_μ can be defined as Eqn. 4.12 to be equivalent and more convenient since it can take all real values. Depending on different assumptions, this test statistics can be defined differently.

$$t_\mu = -2 \ln(\lambda(\mu)) \quad (4.12)$$

As the measure of incompatibility is defined by this profile likelihood ratio, the p-value can be calculated from Eqn. 4.13 to evaluate the probability of the hypothesis with observation. This is also shown in the left figure of Fig. 4.7.

$$p_\mu = \int_{t_{\mu,obs}}^{\infty} f(t_\mu | \mu) dt_\mu \quad (4.13)$$

The function $f(t_\mu|\mu)$ here is the probability density function (p.d.f.) of the test statistics t_μ . One of the important test statistics used to make a discovery is defined by assuming only physically possible cases with $\mu \geq 0$. This assumption is common for searching an excess made by some new physics process above a known background. And for discovery, we are effectively trying to reject the $\mu = 0$ hypothesis. We then use a special notation q_0 , and define it as Eqn. 4.14.

$$q_0 = \begin{cases} -2 \ln(\lambda(0)) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \quad (4.14)$$

4.2.2 Optimization of Sensitivity

Before applying the full statistical treatment, the first step of analysis is to define the signal region of events selected to do the search. This requires a quantitative definition of sensitivity so that the selection criteria can be optimized based on it. In this search, a counting experiment is used to estimate the search sensitivity. This counting experiment is done by defining a mass window around the signal as a one-bin analysis, and then counting the number of events from signal and background in the bin. The likelihood function follows a Poisson distribution. Assuming the number of background events b is already known (or can be estimated very precisely), and we finally observe n events in data, the likelihood function can be written as:

$$L(\mu) = \frac{(\mu s + b)^n}{n!} e^{-(\mu s + b)}. \quad (4.15)$$

According to Eqn: 4.11 and Eqn. 4.14, the test statistics for discovery is:

$$q_0 = \begin{cases} -2 \ln\left(\frac{L(0)}{L(\hat{\mu})}\right) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \quad (4.16)$$

Since the background b is known, and assuming the signal strength s is normalized to unity, $\hat{\mu} = n - b$. According to [36], by taking the approximation with sufficiently large b , the significance can be written as: [36]

$$Z_0 = \sqrt{q_0} = \begin{cases} \sqrt{2(n \ln(n/b) + b - n)} & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \quad (4.17)$$

The measure of the sensitivity can be defined as the median significance with the nominal signal hypothesis ($\mu = 1$). This can be predicted by using the most representative distribution as the ‘Asimov’ dataset.[37] The expected observation $n = s + b$ in this Asimov dataset. The median of the significance then becomes:

$$\text{med}[Z_0|0] = \sqrt{q_0} = \sqrt{2((s + b) \ln(1 + s/b) - s)} = \frac{s}{\sqrt{b}}(1 + \mathcal{O}(s/b)), s \geq 0. \quad (4.18)$$

In this search, we use the approximation $\text{med}[Z_0|0] \approx s/\sqrt{b}$ to optimize our search strategy, but will go back to the exact Poisson likelihood function Eqn. 4.15 for sensitivity estimation, especially in the region with very little background events.

In many analyses, the signal region is also divided into different parts based on the expected purity of the signal events. By separating high purity and low purity signal regions into different subsets called categories, the sensitivity of the search will be improved. Each category can be effectively an independent study if the signal regions for each of them are orthogonal (no overlap between any two of them). The final result requires combination of the different categories. Just as in doing multi-binned analysis, more categories can be treated as more bins being observed in a different (but orthogonal) phase space. Therefore, the likelihood function of each of the categories will be multiplied together to make the full likelihood function as in Eqn. 4.10. Then the following steps to build the profile likelihood ratio and test statistics will be just the same as described above.

4.2.3 Setting Exclusive Limits

In this search, the significance of the data's deviation from the background only hypothesis is reported. In addition, exclusive limits on cross-sections are set for potential heavy resonance decay into $W\gamma$ or $Z\gamma$ final states. This is done by floating the signal strength parameter μ defined in Eqn. 4.7. A modified confidence level is defined as:[38]

$$CL_{s+b} = P(t_\mu < t_{obs}) = \int_{-\infty}^{t_{obs}} f(t_\mu | \mu),$$

$$CL_b = P(t_{\mu=0} < t_{obs}) = \int_{-\infty}^{t_{obs}} f(t_\mu | \mu = 0), \quad (4.19)$$

$$CL_s = CL_{s+b}/CL_b.$$

The functions $f(t_\mu | \mu)$ here is defined similar to the p-value calculation (Eqn. 4.13), which is the p.d.f. of the test statistics t_μ . CL_{s+b} represents the possibility of finding the observation excluding the signal plus background hypothesis, and the CL_b represents the possibility of excluding the background only hypothesis. By looking for a proper value of μ which makes the $CL_s = 0.95$, the corresponding cross-section of the signal process can be regarded as a limit with confidence level of 95%. Scanning through the full mass spectrum, exclusive limits with 95% confidence level can be produced. Any cross section above this limit represents a hypothesis that is excluded by more than 95% probability.

4.2.4 Shape Modeling and Spurious Signal

In this thesis, we also need to study the shape of the signal and the background to constrain the nuisance parameters. This is done by doing a maximum-likelihood fit with chosen functions or families of function candidates. The likelihood function

can be defined either in a binned perspective, by dividing events into different small windows on a full mass spectrum, or in an unbinned perspective, with each of the events treated independently, and the likelihood for each of the bins or events will be multiplied together. The function that maximizes this likelihood among all possible candidates will be taken as the reasonable model for the real shape of the signal or background.

The signal shape is relatively easy to model since it has a resonance in the mass spectrum, but the background is usually more complicated. We can rarely guarantee the real background shape function can be included in the phase space covered by our background function candidates. In order to proceed from this dilemma, we make a compromise by adding a new uncertainty to cover the systematic uncertainty induced by the background mis-modeling. This uncertainty is called a ‘spurious signal’, and the procedure for getting this uncertainty is called a ‘spurious signal test’.

The idea of this spurious signal test is to test the null hypothesis on a simulated background only sample. We require the sample to be much larger than real data, so that statistical fluctuation will be negligible. Ideally, the simulated background sample should match perfectly with the background only fit. This means any mis-modeling will be represented by a mismatch in fit. We test the alternative hypothesis by doing a signal plus background fit to the simulated background sample. A strength of signal (called spurious signal N_{SS}) is extracted even though we do not expect any. Then we divide this spurious signal by the corresponding statistical uncertainty, to estimate the relative impact from this mis-modeling. Firstly, if this ratio is too large, more than 20-50%, we need to change the choice of function candidates because this implies the mis-modeling is one of the dominating effect on this analysis, and our modeling is not good enough. If this ratio is low enough, it implies the modeling uncertainty is almost negligible compare to the statistical uncertainty. This means the

chosen function can model the background. However, since there are still some spurious signal generated by our background functions, an extra uncertainty is needed to cover the mis-modeling. Based on the the ratio between spurious signal and statistical uncertainty, the systematic uncertainty from mis-modeling can be converted to an additional fraction of statistical uncertainty.

There are some other ways to test the validity of functional forms chosen to model background. For example, an injection test, adding a reasonable strength of signal into simulated background events, and extracting the same signal strength through the full procedure of analysis. However, the advantage of a spurious signal test is that it not only provides validation, but also estimates possible mis-modeling, and therefore makes the final result conservative.

Chapter 5

Search for $X^\pm \rightarrow W^\pm\gamma$ and $X^0 \rightarrow Z^0\gamma$

5.1 Introduction

In this search for new massive resonances, pp collision events are simulated for both the new physics signals and the SM backgrounds. The signal simulations are used to study kinematic features of the hypothetical processes as they would appear in the ATLAS detector, and SM backgrounds are generated to increase the statistics in corners of phase space where the signals are dominant. By comparing the features of the signal and backgrounds, a search strategy can be developed. We define a phase space region called the signal region (SR) for the search, and split the SR into subsets called categories to improve detection sensitivity. After event selection procedures and categorizations are determined, we treat each category separately and model the production of $X^\pm \rightarrow W^\pm\gamma$ and $X^0 \rightarrow Z\gamma$ signals in each region. Parameterized background distribution functions are compared to simulated SM background samples to validate their flexibility, and to test for the creation of spurious signals. The data are analyzed following the methodology described in Section 4.2 to measure the significance of any possible new resonance signals, and if none are found to set limits on their production cross sections.

5.2 Data and Simulation Samples

5.2.1 Data Samples

The data used in this search includes all data collected by the ATLAS detector from 2015 to 2018 with LHC pp collisions at 13 TeV center of mass energy (called Run 2). The HLT_g140_loose trigger is used for the collection of events having a photon with transverse energy (E_T) higher than 140 GeV and passing a loose photon identification criteria. The performance of this trigger is shown in Fig. 4.1. After demanding good performance of all ATLAS detector components the total pp collision integrated luminosity of this dataset is 139 fb^{-1} .

5.2.2 Signal Simulation Samples

Simulated signal events are generated using the benchmark theory models described in Section 2.2. However, this study is a generic search for massive resonances of the type $X^\pm \rightarrow W^\pm \gamma$ and $X^0 \rightarrow Z \gamma$. It uses these models to provide general information about signal shapes and production kinematic features for the determination of the detection efficiencies needed for the calculation of production cross sections.

Since these benchmark models describe resonances with different spins and final states, they are treated as different signal hypotheses. A total of four different signal simulation are produced, and listed as in Table. 5.1. All samples are generated using POWHEG [39] or MADGRAPH5.2, and interfaced to PYTHIA8 [40] for parton shower and fragmentation processes. The narrow-width approximation (NWA) is taken for all the signal simulations with the resonance width fixed at 4 MeV. The generated events are passed through a detailed ATLAS detector simulation based on GEANT4 [41] and reconstructed with the same software used for data. The resonance mass points are at integer TeV masses from 1 to 7 TeV to cover our search range, and an

Table 5.1: The simulated samples used in this search is listed in the table. The corresponding theory models and simulation generators are listed too.

Signal Process	Spin	Theory Model	MC Generator
$gg \rightarrow X \rightarrow \gamma Z(\rightarrow q\bar{q})$	0	Singlet scalars model	Powheg+Pythia8
$gg \rightarrow X \rightarrow \gamma Z(\rightarrow q\bar{q})$	2	Higher dimensional operator model	MadGraph+Pythia8
$qq \rightarrow X \rightarrow \gamma Z(\rightarrow q\bar{q})$	2	Higher dimensional operator model	MadGraph+Pythia8
$qq \rightarrow X \rightarrow \gamma W(\rightarrow qq')$	1	Heavy vector triplet model	MadGraph+Pythia8

extra mass point at 1.5 TeV to provide better modeling at low mass region.

5.2.3 Background Simulation

SM $\gamma + Jet$ events can be produced through either scattering of a quark and a gluon or through quark-antiquark annihilation plus an initial state radiation (ISR) photon and 2-jets processes with ISR or FSR (final state radiation) photons. Production through these processes, with the photon participating in the hard scatter and back-to-back with the leading jet, is called “prompt” production. Contributions of $\gamma + Jet$ events with photons from fragmentation (not prompt) are relevant at low photon E_T , and only at a few percent level above 200 GeV.

The fact that the $\gamma + Jet$ cross section is very high, and that the photon is real (i.e. efficiently passing the photon selection) makes these processes the largest background contribution for this analysis.

The events corresponding to single photon with associated jet processes are generated by **SHERPA** generator v2.2.2 [42] with up to two additional parton emissions at next leading order (NLO) accuracy and up to four additional partons at leading order (LO) accuracy using **COMIX** [43] and **OPENLOOPS** [44, 45]. The Sherpa predictions have been compared with data showing a good description of the shape of the main kinematic quantities [46].

The samples have been generated in E_T^γ slices to cover the full spectrum relevant for this analysis, and are passed through a full detector simulation. More details

Table 5.2: Summary of NLO photon+jet QCD background samples generated with SHERPA.

E_T^γ range	σ [pb]	Number of events	Effective Lumi. [fb $^{-1}$]
35–70	43976	34911000	0.794
70–140	4526.5	34923500	7.715
140–280	376.0	34929800	92.90
280–500	21.9	10486500	478.8
500–1000	1.5	3483100	2322
1000–inf	0.015	2162720	144181

on the slices, cross sections, filter efficiency and number of generated events can be found in Table 5.2.

The other less dominant SM backgrounds are the production of QCD multi-jets with jets faking a photon, diboson $W/Z(\rightarrow qq) + \gamma$, which can be enhanced due to boson tagging, and $t\bar{t} + \gamma$, which could be enhanced by both boson tagging and btagging. The SM $W/Z\gamma$ events are modeled by SHERPA2.1 at LO (on-shell only) and the $t\bar{t} + \gamma$ MC is modeled by MADGRAPH5.2 (Matrix Element) and PYTHIA8.1 (Parton Showering). It was shown in a previous study [47] that events with a jet faking a photon do not have visible dependence of the invariant mass of the fatjet and photon and since the final background estimation is done by fitting directly to the data, the estimation for the multi-jets backgrounds is not necessary.

5.3 Selection and Categorization of Data

The goal of this analysis is to search for heavy resonances decaying to $W\gamma$ or $Z\gamma$ pairs, when the vector boson decays to hadronic final states. Thus, the analysis requires a high p_T , isolated photon and one high p_T isolated large- R jet from a boosted W or Z boson. This section presents the reconstruction and selection of the objects used in the analysis, together with the full descriptions of the event selections and

categorization.

5.3.1 Baseline Selection

According to the physics object definition in Section 4.1, the following photons and large radius jets are used for event selection:

- Photons: $p_T(\gamma) > 200$ GeV, $|\eta(\gamma)| < 1.37$, passing tight photon ID and tight calorimeter isolation.
- anti- k_t large- R Jets (R=1.0): $p_T(J) > 200$ GeV, $|\eta(J)| < 2.0$.

The η criteria are defined based on the calorimeter geometry described in Section 3.2.2, focusing on the barrel calorimeter region for better performance. The improvement from endcap photon events is 1% or less, thus can be safely excluded. Events with endcap photons are later used as a Control Region (CR), helping us to validate our background fit functions. The details for the CR can be found in Appendix. A.1.

In order to enhance the purity of the signal in the Search Region (SR), the following event selection cuts are applied using the physics objects (fat jets and photons) defined in section 4.1:

- jet and photon overlap removal: remove fat jets with $\Delta R(\text{jet}, \gamma) < 1.0$, reject events with no surviving fat jets.
- The events are kept if more than one large- R jet and more than one photon passes the aforementioned selection criteria. The leading p_T ones are selected to construct the $J\gamma$ system invariant mass m_X .

The photon and jet overlap removal is necessary because photons are reconstructed also in the jet collection. Since the photon and jet we are interested in are supposed

to be back-to-back, those collinear photon and jets rarely come from signals, and therefore can be rejected.

This is the baseline selection, and events passing this selection are categorized based on W/Z boson tagging and beauty hadron identification (b-tagging). Some stronger photon and jet p_T cuts are applied for several of the categories, so not all events that pass the baseline selection are used in our analysis. Figure 5.1 shows the efficiency of this baseline selection for the various signal samples as a function of the resonance mass m_X . The detailed efficiencies for various criteria of the baseline selection are reported in Appendix. A.2. The signal $\eta(\gamma)$, $p_T(\gamma)$ and $p_T(jet)$ distributions are shown in Figure 5.2, 5.3 and 5.4. The spin and polarization of the resonance strongly affect the decay kinematics, leading to the different signal detection efficiencies shown in Figure 5.1.

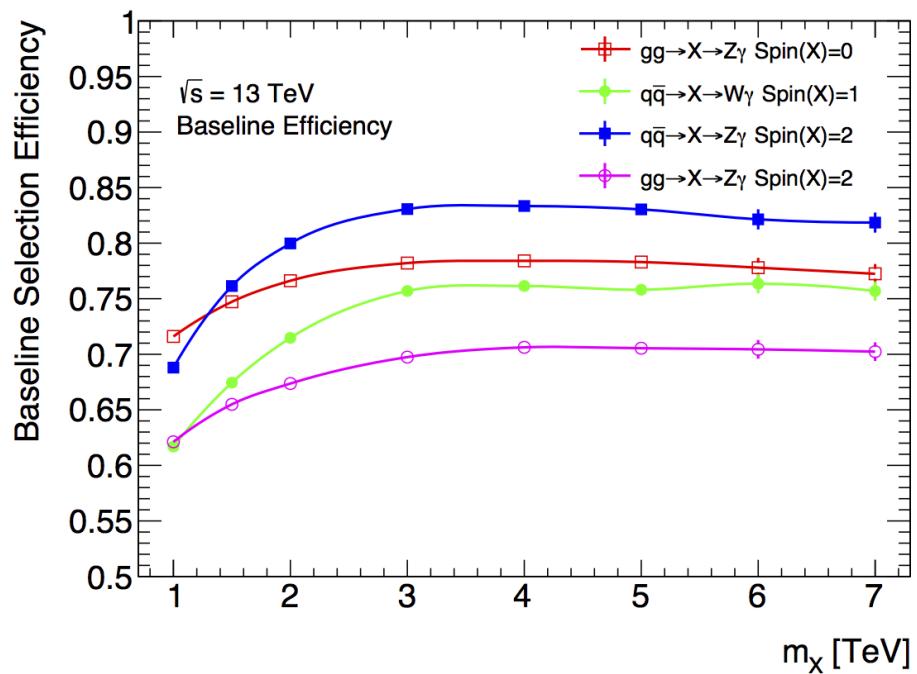


Figure 5.1: Signal Efficiencies of events passing the baseline selection (before categorization). For these efficiencies, the denominator is the total number of events generated from simulation, and the numerator is the events passing the trigger and baseline selections defined in this section.

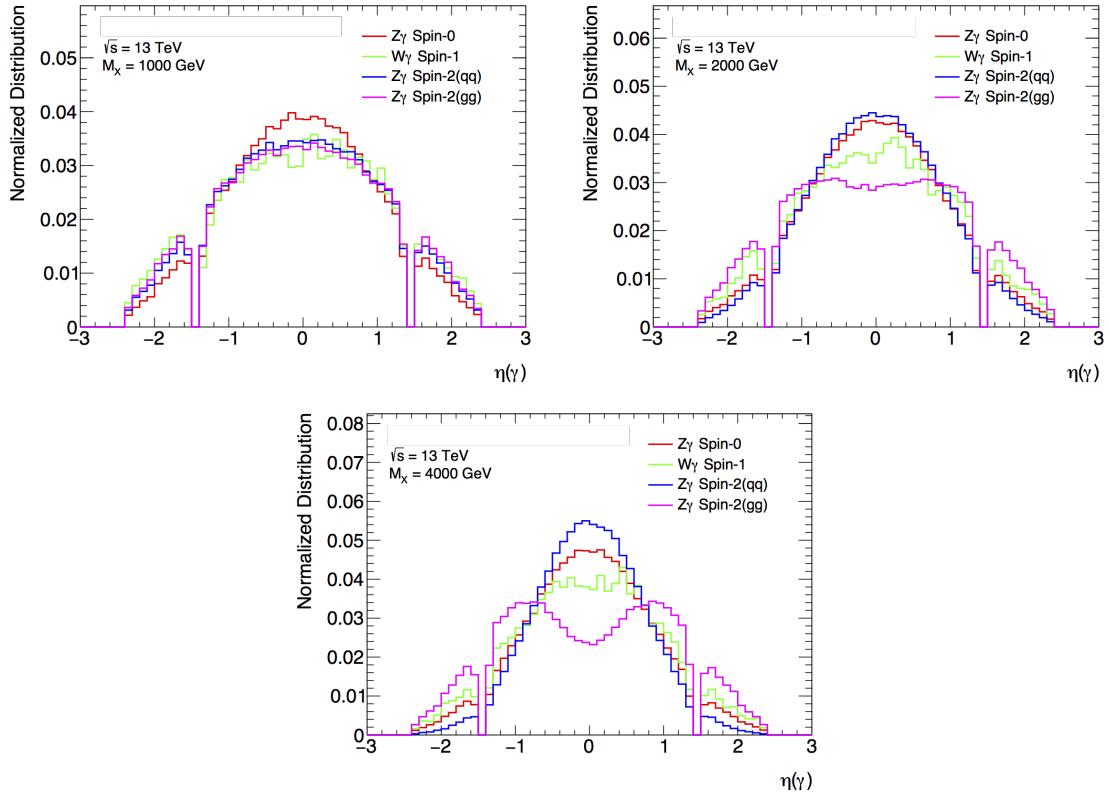


Figure 5.2: These three plots show the differences of photon η distribution among different signal channels, for signal mass at 1 TeV (top left), 2 TeV (top right) and 4 TeV (bottom).

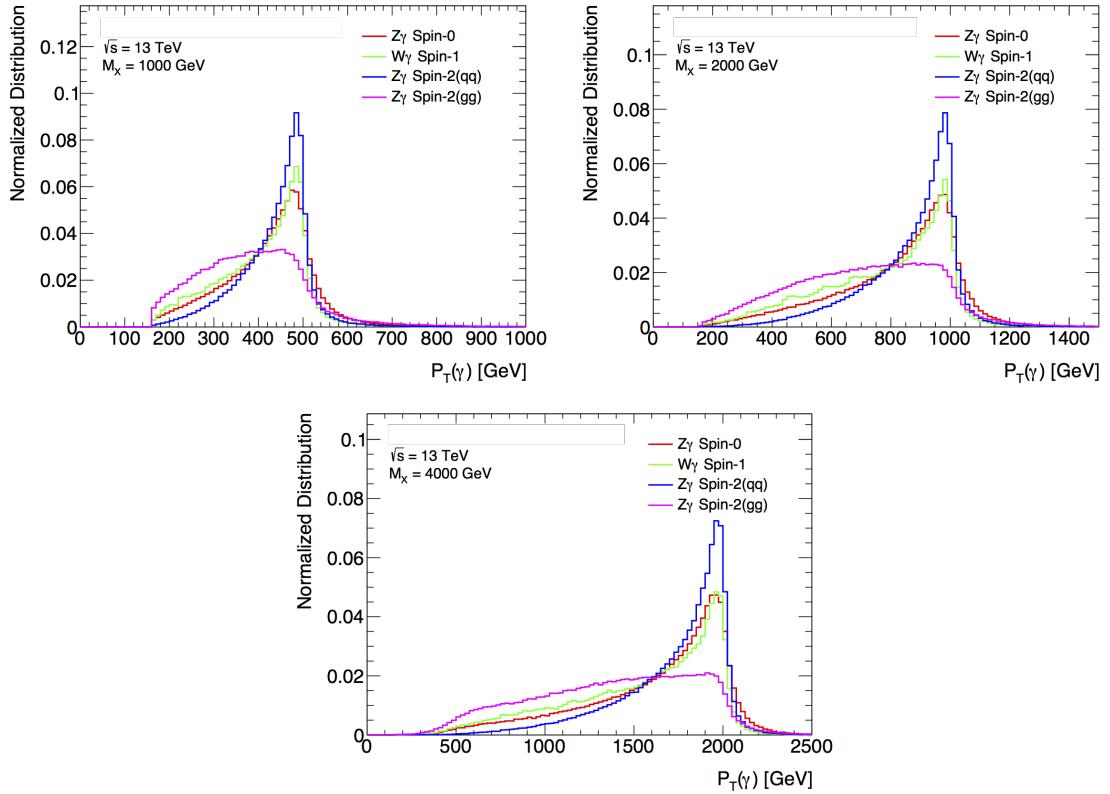


Figure 5.3: These three plots show the differences of photon p_T distribution among different signal channels, for signal mass at 1 TeV (top left), 2 TeV (top right) and 4 TeV (bottom).

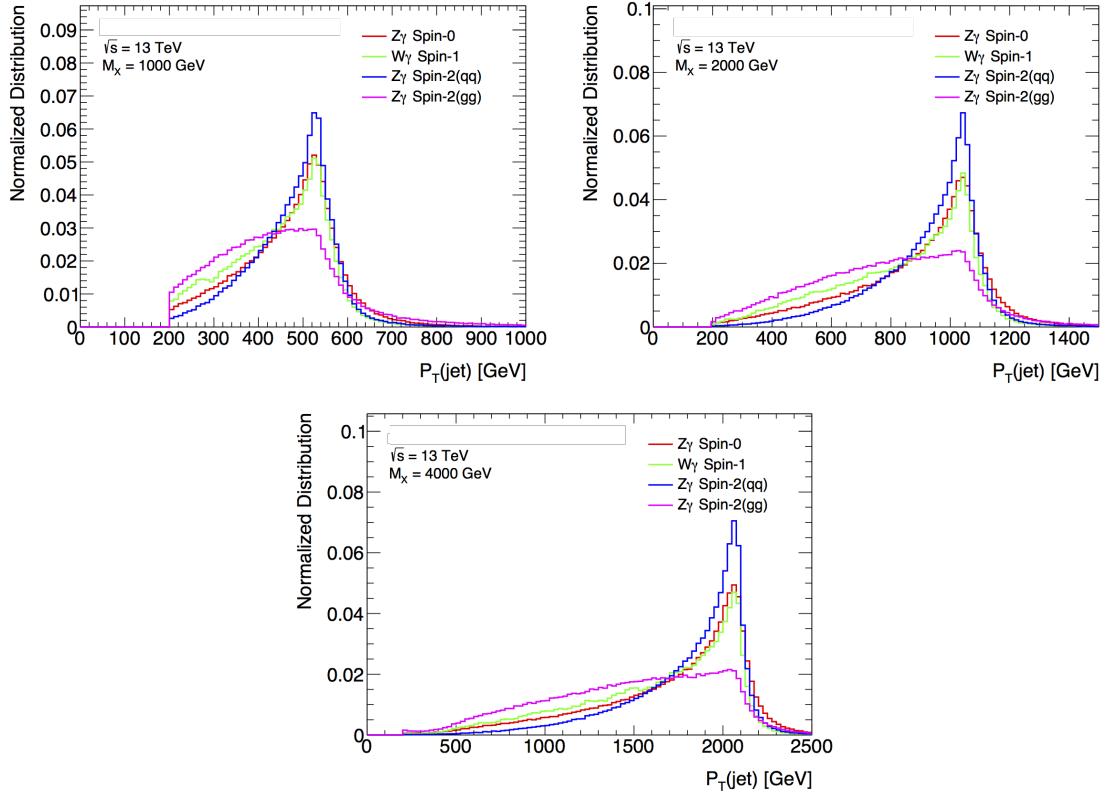


Figure 5.4: These three plots show the differences of jet p_T distribution among different signal channels, for signal mass at 1 TeV (top left), 2 TeV (top right) and 4 TeV (bottom).

5.3.2 Categorization of the Baseline-selected Events

In all events, the photon is the highest P_T photon passing the baseline selection with $P_T(\gamma) > 200$ GeV and $|\eta(\gamma)| < 1.37$, and the jet is the highest P_T large-R jet with $P_T(J) > 200$ GeV and $|\eta(J)| < 2.0$. In the low $m_{J\gamma}$ region, the purity of W/Z boson is improved using boson tagging of the large-R jets as described in section 4.1.3. In addition, the Z boson purity is increased using b-tagging (at the 60% working point, see section 4.1.4). However, this sacrifices some sensitivity for high P_T large-R jets where these taggers become less efficient. This is recovered at high $m_{J\gamma}$ mass by requiring only a W/Z boson mass window cut after the baseline selection. To balance these considerations, we define mutually exclusive categories of events as described below and illustrated in Figure 5.5.

For the $X \rightarrow Z\gamma$ resonance search, three categories are defined:

- **B-tag category:** the selected large-R jet passes boson mass tagging and double b-tagging.
- **D2 category:** the selected large-R jet passes boson mass tagging and D2 jet substructure tagging for Z boson, but fails double b-tagging.
- **Vmass category:** the selected large-R jet passes Z boson mass tagging but fails both double b-tagging and D2 jet substructure boson tagging.

For the $W\gamma$ resonance search the double b-tagging is not useful, so the B-tag category is removed and the events are categorized in only two subsets.

- **D2 category:** the selected large-R jet passes boson mass tagging and D2 jet substructure tagging for W boson.
- **Vmass category:** the selected large-R jet passes W boson mass tagging but fails D2 jet substructure boson tagging.

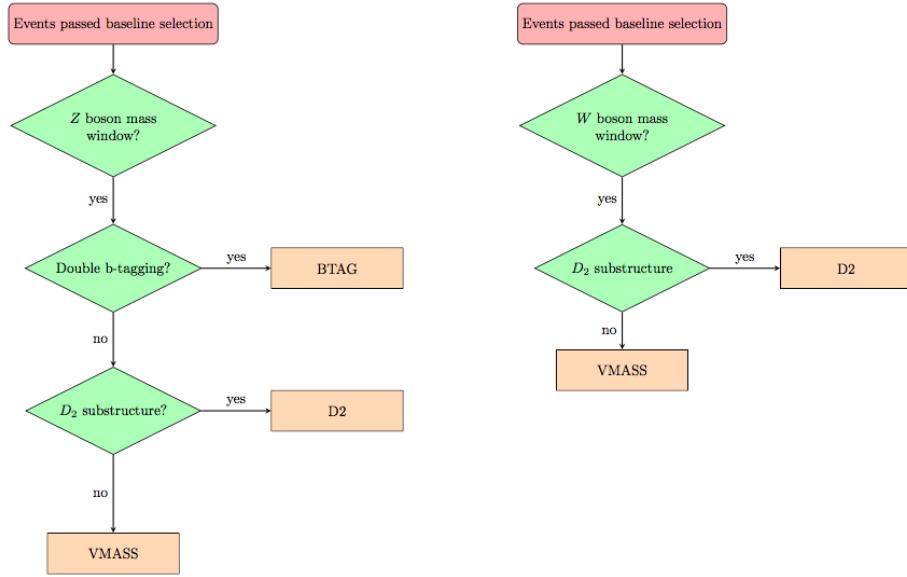


Figure 5.5: The flowchart for the criteria of categorization for $Z + \gamma$ (left) and $W + \gamma$ (right) signals.

In order to further improve the sensitivity of the resonance search, we apply additional mass dependent photon and jet p_T cuts in the D2 and Vmass categories where the signal and background statistics are both relatively high. This cut is optimized independently at each of the generated mass points in different signal channels and categories. The optimized cut values, each set by tolerating 5% loss of sensitivity, are modeled by parabola functions. The final mass dependent p_T cut therefore differs between different signal channels and categories.

The relationship between optimal p_T cut values and the resonance mass points is roughly parabolic. However, if following the parabolic trend, at high mass region, the cut value keep decreasing. This means upon the exponentially decreasing background spectrum, there could be some fake bump produced by this p_T cut. In order to avoid such danger, we decide to use a constant p_T cut after the cut function reaching its maximum. This guarantees the cut to be differentially continuous to produce the background mass distribution as smooth as possible. The functions of this p_T cut is

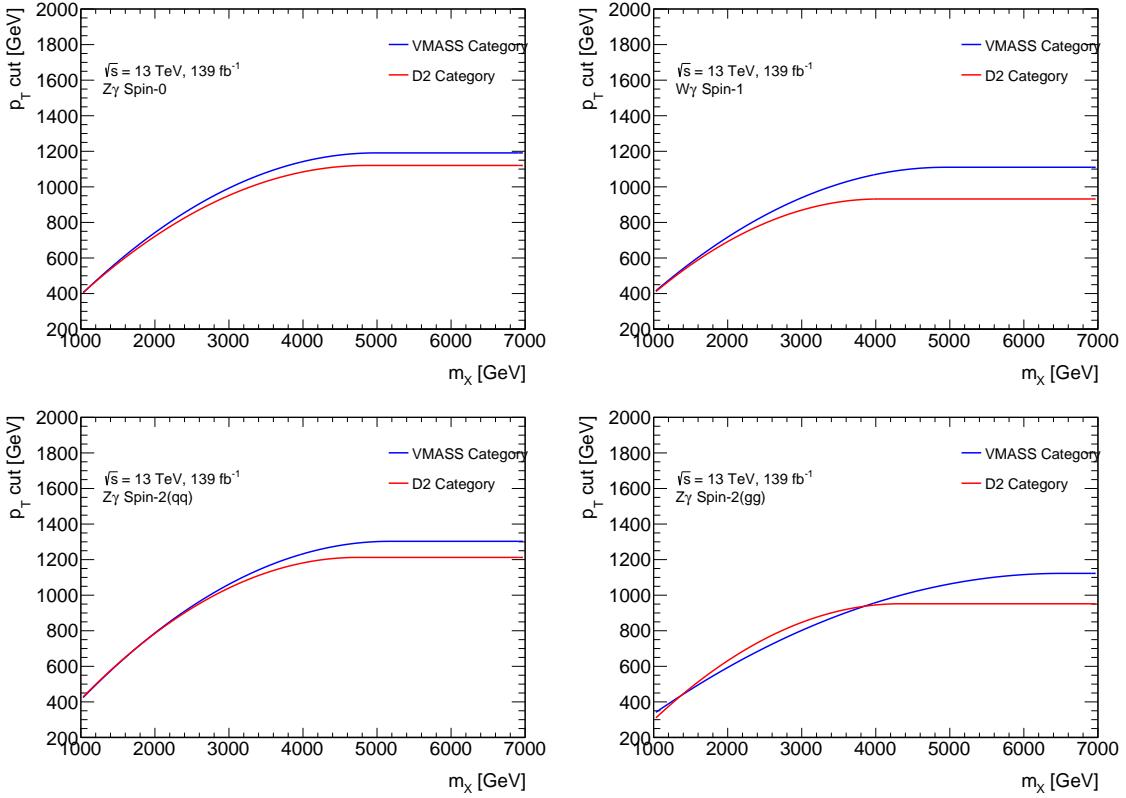


Figure 5.6: The dependence of the photon p_T cut on the resonance mass is shown here. For different signal channels and categories, the cut functions are different. They are all above the baseline 200 GeV threshold, so there will be no discontinuity at the crossing points.

shown in Fig. 5.6.

Details of this cut optimization are described in Appendix A.4.

The fractions of events in each category for the $Z\gamma$ and $W\gamma$ signal samples as a function of the resonance mass m_X are shown in Figure 5.7. By combining Figures 5.1 and 5.7, the overall efficiencies for each category can be found in Figure 5.8.

The relative efficiency for categorization of the dominant SM γ +jets background relative to the baseline selection is shown in Figures 5.9 and 5.10. The D2 category efficiency is significantly higher in the $W\gamma$ signal channel than in the other $Z\gamma$ channels due to the difference in polarization of the W/Z bosons. The tagger used

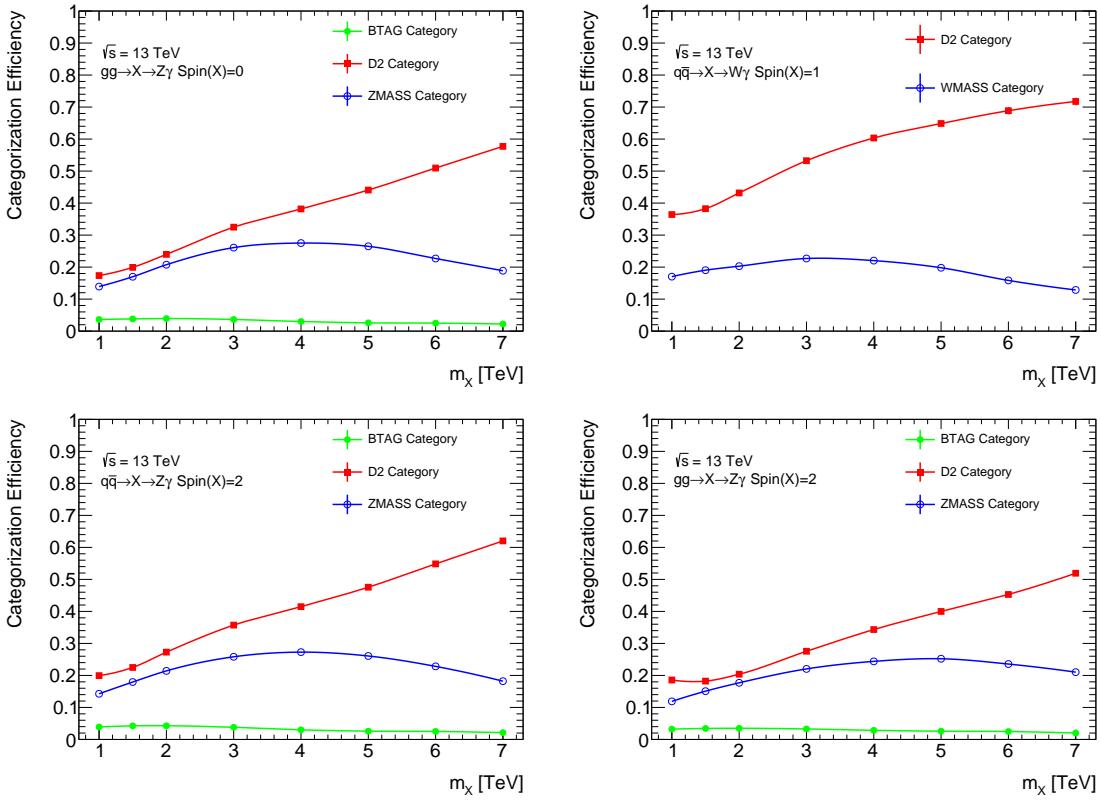


Figure 5.7: Fraction of signal events in each categories for different production modes as function of m_X . The numerator here is the number of events categorized and passing the functional p_T cut, and the denominator is the number of events passing the baseline selection.

here is optimized for longitudinally polarized bosons, as in the case for the $W\gamma$ channel in this search. All other $Z\gamma$ channels have transversely polarized Z bosons, and therefore have much lower D2 category efficiencies.

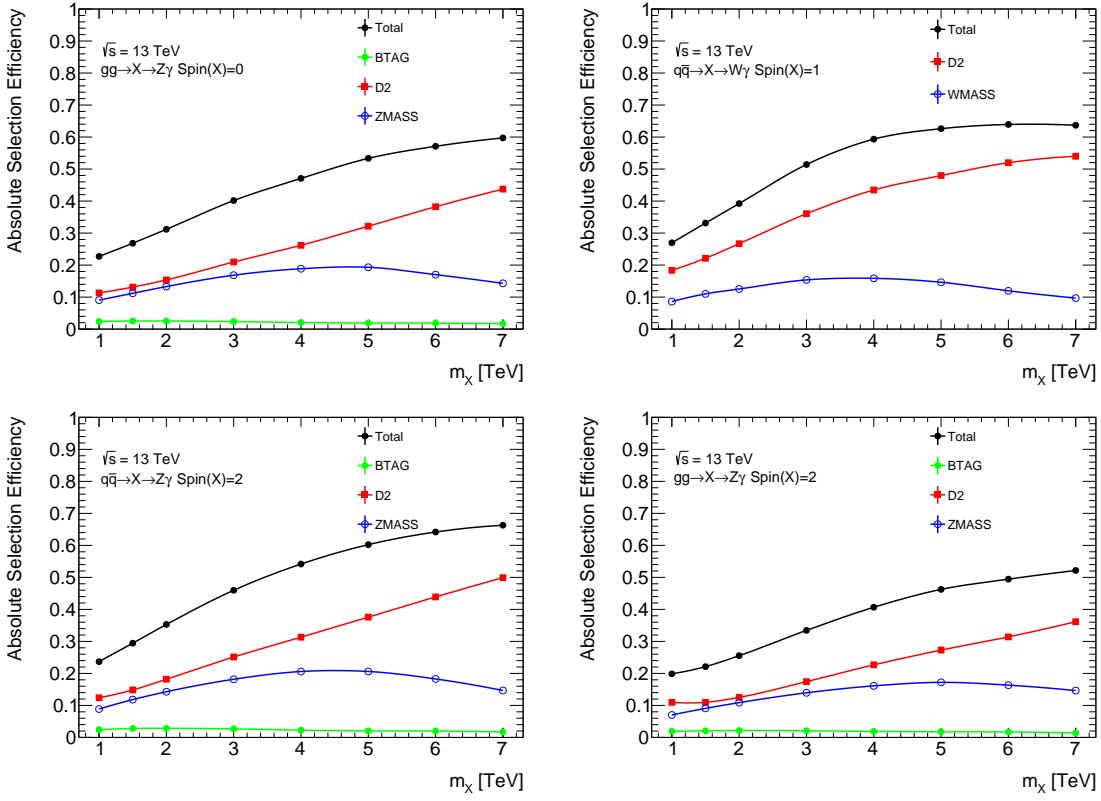


Figure 5.8: Overall efficiencies of signal events in each categories for different production modes as functions of m_X . The numerator here is the number of events categorized and passing the functional p_T cut, and the denominator is the total number of events generated from simulation.

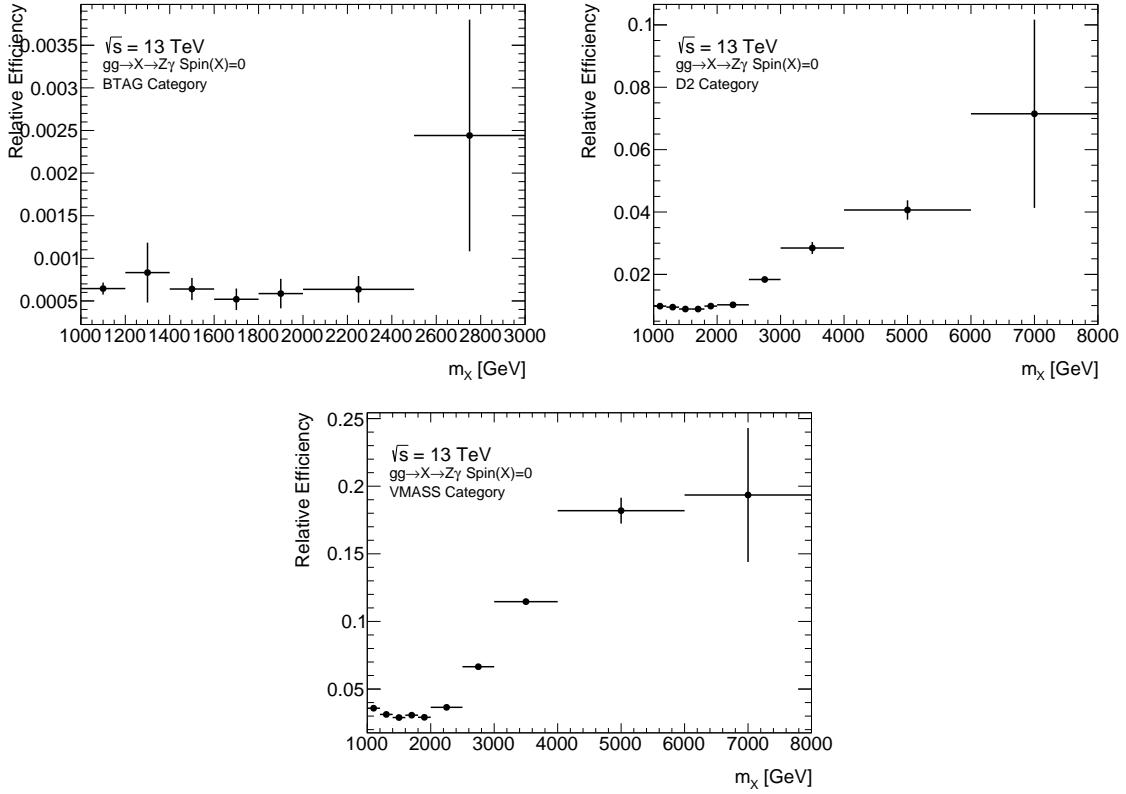


Figure 5.9: The relative background efficiency (vs. baseline selection) of each categories for the SM $\gamma + jets$ events with spin-0 $Z\gamma$ categorization. The numerator here is the number of events categorized and passing the functional p_T cut, and the denominator is the number of events passing the baseline selection. This efficiency is calculated in each of the bins and therefore is strongly biased for the lower edge of each bin due to the general exponentially decreasing trend of background. In the high mass region, the SM background MC runs out of statistics and therefore has relatively large fluctuations.

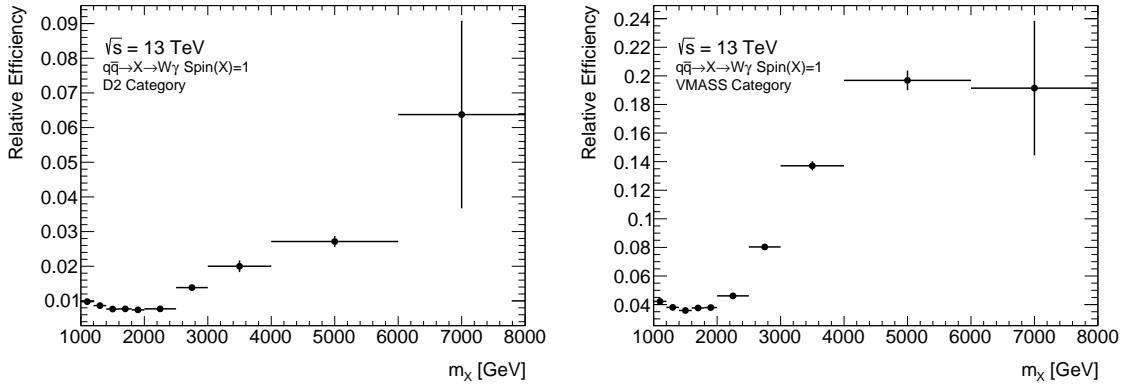


Figure 5.10: The relative background efficiency (vs. baseline selection) of each categories for the SM $\gamma + jets$ events with $W\gamma$ categorization. The numerator here is the number of events categorized and passing the functional p_T cut, and the denominator is the number of events passing the baseline selection. This efficiency is calculated in each of the bins and therefore is strongly biased for the lower edge of each bin due to the general exponentially decreasing trend of background. In the high mass region, the SM background MC runs out of statistics and therefore has relatively large fluctuations.

5.4 Signal Shape Modeling

5.4.1 Signal Shape Fitting

The distribution of the resonance mass $m_{J\gamma}$ is modeled by a probability distribution function (PDF) constructed from a double-sided Crystal Ball function (DSCB) [48]. The center part is a Gaussian function representing the *core* of well-reconstructed events, and modified by a non-Gaussian distribution, modeling both sides of the tails of poorly reconstructed events.

$$DSCB(m; N, \mu, \sigma, \alpha_1, n_1, \alpha_2, n_2) = N \cdot \begin{cases} \left(\frac{n_1}{|\alpha_1|} \right)^{n_1} \exp \left(-\frac{|\alpha_1|^2}{2} \right) \left(\frac{n_1}{|\alpha_1|} - |\alpha_1| - \frac{m-\mu}{\sigma} \right)^{-n_1} & \frac{m-\mu}{\sigma} \leq -\alpha_1 \\ \exp \left(-\frac{(m-\mu)^2}{2\sigma^2} \right) & -\alpha_1 < \frac{m-\mu}{\sigma} \leq \alpha_2 \\ \left(\frac{n_2}{|\alpha_2|} \right)^{n_2} \exp \left(-\frac{|\alpha_2|^2}{2} \right) \left(\frac{n_2}{|\alpha_2|} - |\alpha_2| + \frac{m-\mu}{\sigma} \right)^{-n_2} & \alpha_2 < \frac{m-\mu}{\sigma} \end{cases} \quad (5.1)$$

The parameters μ , σ , α_1 , n_1 , α_2 and n_2 are floated for fitting the signal shape, and N is the normalization factor.

The n_1 and α_1 parameters regulate the shape and position of the non-Gaussian tail on the lower mass side – α_1 is a measure of how far from the peak the distribution becomes non-Gaussian while n_{CB} is related to the slope of the tail. The n_2 and α_2 is similar but for the higher mass side tail. Since these four variables describe the shape of the signal tails, some limits related to the general feature of signals are set for them. All these four variables are set to be positive to stabilize the fit performance without sacrificing any real flexibility of the DSCB function. The n_1 is limited below 200 which is never reached; the α_1 is limited between 0.1 to 4 to avoid the left side non-Gaussian tail too close or too far from the center; the α_2 is limited between 1 to 5 for the same reason; and the n_2 is limited below 10 because the higher mass side tail is always very small. As shown in Figure 5.11, σ represent the peak width of the

signal shape, which is also the signal mass resolution.

The signal shapes of different resonance mass points are investigated with this fit function for all our signal models. Since the same model is used for all different signal channels, the performance of this fit is sometimes not perfect. Figures 5.12 and 5.13 are examples of the signal shape fit results for spin-0 $q\bar{q} \rightarrow X \rightarrow Z\gamma$ and spin-1 $q\bar{q} \rightarrow X^\pm \rightarrow W^\pm\gamma$.

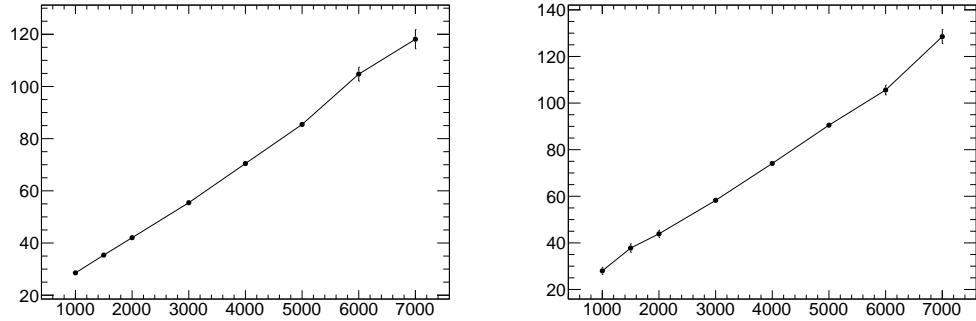


Figure 5.11: These are plots of fit parameter σ dependence of $M_{\gamma J}$. According to Eq.(5.1), these plots represent the mass resolution for different signal mass.

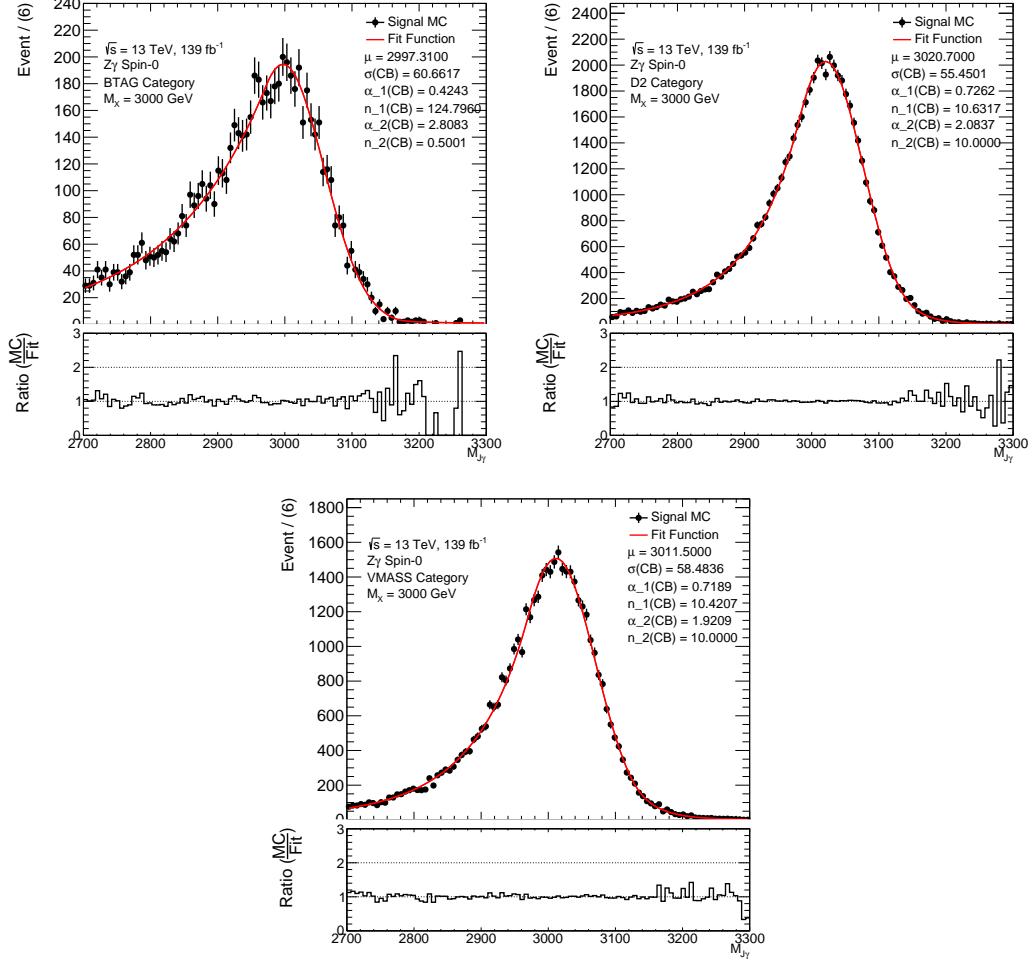


Figure 5.12: $m_{J\gamma}$ mass distributions of the scalar spin-0 $Z\gamma$ model after event selection in each category with fit to the signal shape according to Eq.5.1. The 3 TeV point for BTAG, D2 and Vmass category results are shown here.

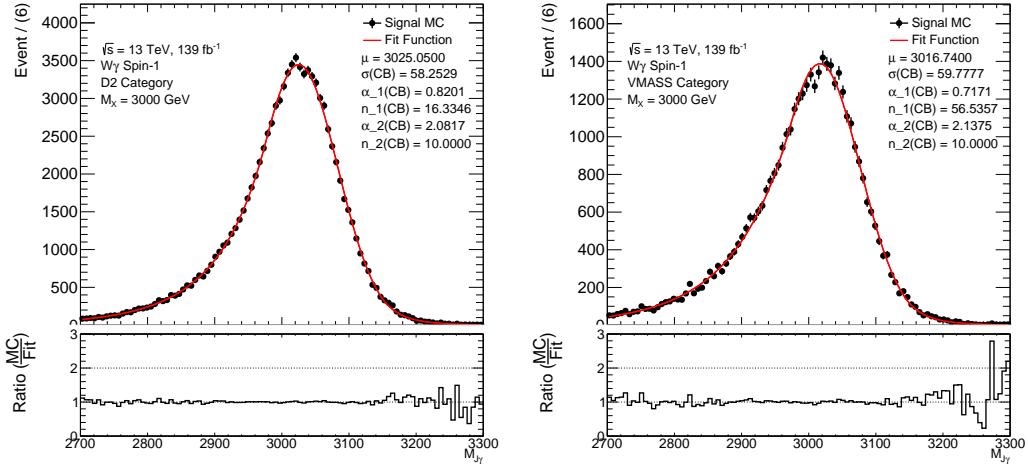


Figure 5.13: $m_{J\gamma}$ mass distributions of the spin-1 $W\gamma$ model after event selection in each category with fit to the signal shape according to Eqn. 5.1. The 3 TeV point for D2 and Vmass category results are shown here.

5.4.2 Interpolation

The signal shapes of the resonance with mass in between the generated mass points are obtained according to signal shape interpolation. All the parameters have clear dependence on m_X and change relatively smoothly. A linear interpolation for each of the fit parameters between two neighbor mass points is a good estimation.

The spectrum of the interpreted $m_{J\gamma}$ mass distributions are shown in Figures 5.14 and 5.15.

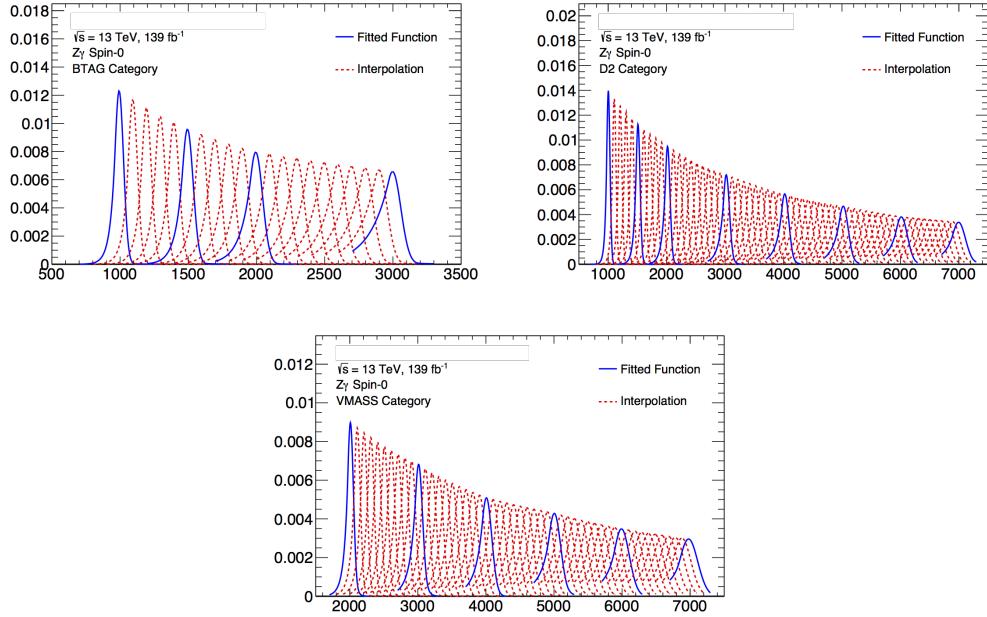


Figure 5.14: These are spectra of the interpolated signal function shape for the scalar spin-0 $Z\gamma$ model. The blue curves are functions directly fit to the full simulation mass points.

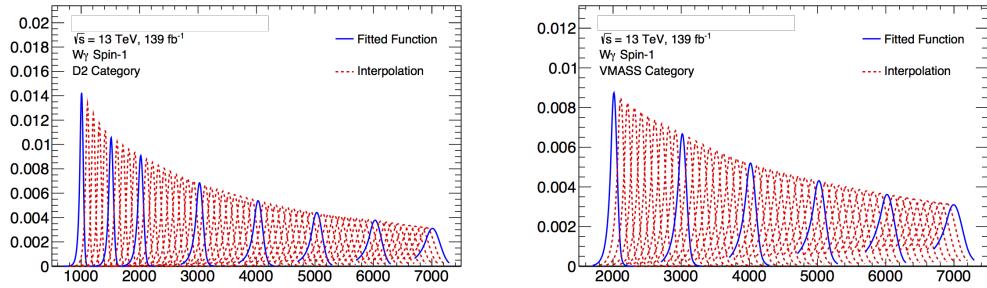


Figure 5.15: These are spectra of the interpolated signal function shape for the spin-1 $W\gamma$ model. The blue curves are functions directly fit to the full simulation mass points.

5.5 Background composition and modeling

In the resonance signal regions, the backgrounds are composed of SM processes with non-resonant smoothly falling distributions as a function of $m_{J\gamma}$. In all the signal regions, the dominant background is the SM $\gamma + \text{jets}$, which contributes more than 90% in the $D2$ and $V\text{mass}$ categories, and approximately 80% in the $B\text{tag}$ signal region. Contributions from other backgrounds such as SM $W/Z + \gamma$ events are small, and are estimated from MC simulated samples.

Tables 5.3 and 5.4 summarize the expected number of background events in the different signal regions.

Table 5.3: Observed data and expected compositions of backgrounds at luminosity = 139 fb^{-1} in the range of resonance mass of 800-7000 GeV in the $Z\gamma$ Spin-0 channel.

SR	btag	d2	zmass	baseline
Observed Data	436	5659	20728	497047
Total SM Background	433	6042	22243	534423
$\gamma + \text{jets}$	382	5564	21438	518531
SM $Z\gamma$	43.8	212	228	3402
SM $W\gamma$	7.3	266	578	12490

Table 5.4: Observed data and expected composition of backgrounds at luminosity = 139 fb^{-1} in the range of resonance mass of 800-7000 GeV in the $W\gamma$ spin-1 channel.

SR	d2	wmass	baseline
Observed Data	6373	25146	497047
Total SM Background	6526	26389	534423
$\gamma + \text{jets}$	5976	25272	518531
SM $Z\gamma$	192	301	3402
SM $W\gamma$	358	816	12490

As shown in tables, the agreement between the SM simulation and observed data is not perfect. The SM background is over-estimated by about 7% after the baseline

selection applied compared to data. The regions with stronger criteria like Btag and D2 categories have better agreement. However, this difference is not going to impact our background modeling, because the SM simulation in this search is only used to determine number of parameters needed for the background function. The final parameter values for background functions are obtained directly from fitting to data.

5.5.1 Spurious signal test on functional form

Before going into the details of background functional fit, studies about the potential bias introduced by mis-modelings are necessary. To understand the bias of the chosen background parametrization with respect to the actual background distribution, a *spurious signal* test is performed following these steps:

- Fit the MC background samples with a background+signal model. The number of fitted signal events is the value of spurious signal N_{SS} . Errors on the MC bins correspond to the expected data statistical error when performing the fit. This will reduce the difference between the background fit on the MC and data, because the number of parameters needed to model a distribution can be totally changed when different uncertainties are assigned to the distribution.
- Count the number of background events (N_B) inside a window of $\pm 2\sigma$ around the signal peak, where σ is the width of double sided crystal-ball function from signal modeling.
- Calculate the value δ_B as the Poisson uncertainty corresponding to 68% confidence level given the expected number of background events N_B . For the N_B large enough, the δ_B is approximately $\sqrt{N_B}$. This δ_B is used to quantify the statistical uncertainty in the search.

- Take the ratio of N_{SS}/δ_B to compare the background modeling uncertainty and the statistical uncertainty from data. Since the two uncertainties are independent, they are added in quadrature. A ratio below 0.5 indicates that the modeling uncertainty is clearly less impactful than the statistical uncertainty. And when the ratio is less than 0.2, the modeling uncertainty will be almost negligible.

The spurious signal test has been performed with various signal mass hypothesis in the corresponding limit-setting range, in a step size of 20 GeV. In the cross-section limit setting, the absolute value of the spurious signal $|N_{SS}|$ is considered as a systematic uncertainty to measure the potential bias related to the choice of the background functional form. To avoid spurious signals due to statistical fluctuation from limited statistics of simulation, the absolute values of spurious signal are parametrized as an exponential function. In the limit setting, the quoted spurious signal used is doubled in order to keep our prediction conservative.

Figure 5.16 shows the results of the test using 2-5 parameters in the function family shown in Eq.5.2 in the spin-0 $Z\gamma$ signal regions. Similar results were obtained in $Btag$ and $D2$ categories, regardless of the number of parameters used.

Figure 5.17 plots the number of spurious signal events along the mass spectrum for better understanding of the actual fitted spurious signal. The red lines are exponential functions used to fit the absolute value of spurious signals. The spikes in the spurious signal test results are understood as an effect from how smooth the resonance mass spectrum in the simulated background samples is, rather than a limitation of the flexibility of the functional form. The correlation of these spikes are studied by dividing the simulated background sample into fractions, and comparing their behavior as discussed in Appendix. A.5.3.

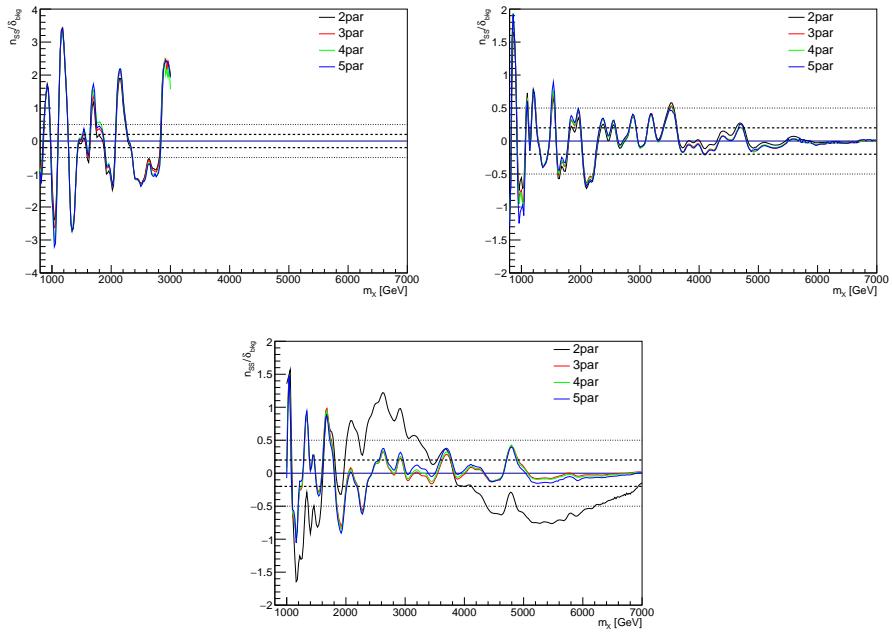


Figure 5.16: Spurious signal test on MC simulation samples with different number of parameters in different event categories (a) $Btag$, (b) $D2$, (c) $Zmass$ of the spin-0 $Z\gamma$ channel.

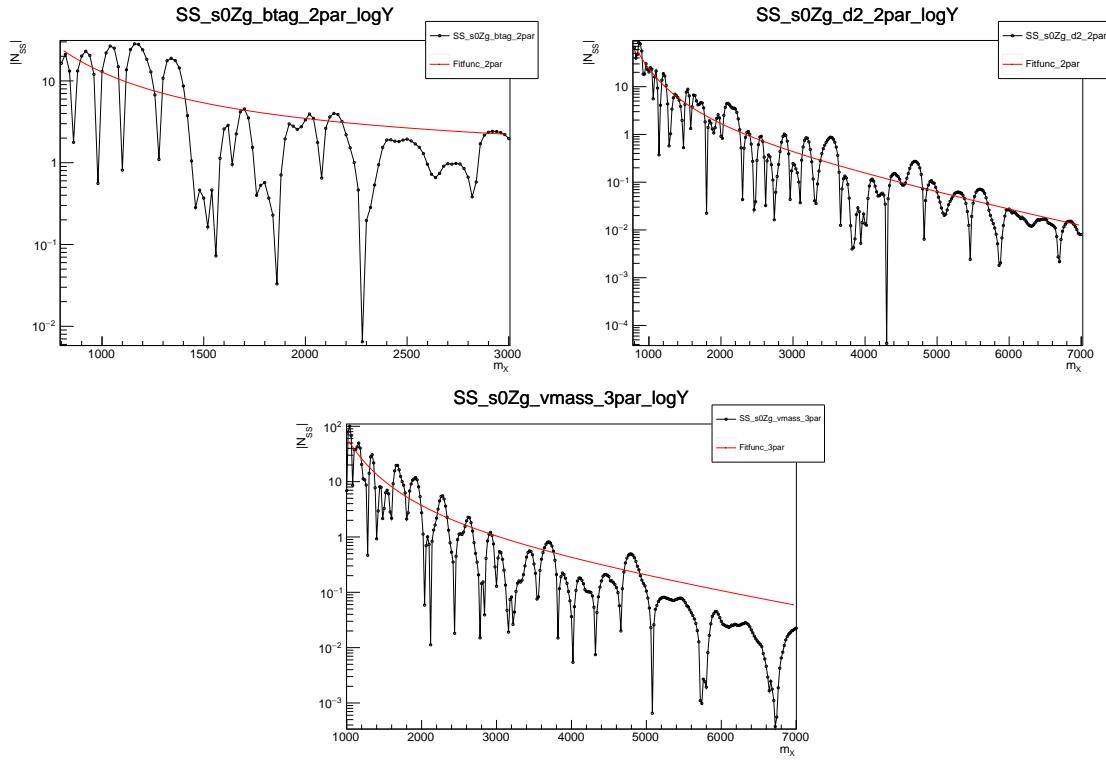


Figure 5.17: The raw number of spurious signal events fitted with MC samples in different event categories (a) $Btag$, (b) $D2$, (c) $Zmass$ of spin-0 $Z\gamma$ channel for chosen number of background function parameters. The $|N_{SS}|$ are then parametrized with an exponential function to minimize the effect due to limited MC statistics on estimation of the spurious signal. In order to have a conservative estimation of the spurious signal, the parametrized results is doubled for the spurious signal systematic uncertainty.

5.5.2 Parametrization and Estimation of the Background

The analysis proceeds by searching for resonant structures over the $m_{J\gamma}$ spectrum. The background is estimated by performing an un-binned maximum-likelihood fit to the data $m_{J\gamma}$ mass distribution with a parametric form based on a family of functions [49].

$$B(m_{J\gamma}; p_i) = (1 - x)^{p_1} x^{p_2 + p_3 \log(x) + p_4 \log^2(x) + p_5 \log^3(x)} \quad (5.2)$$

where $x = m_{J\gamma}/\sqrt{s}$, $p_i, i = 1, 2, \dots$ are dimensionless shape parameters. The *spurious signal test*, which is described in details in Section 5.5.1, is used to test the robustness of the functional form in fitting the SM background. Different numbers of parameters in the functional form are tested to determine the minimum number of parameters that should be used in background modeling. Then an F -test [50] is used to validate the function forms and the number of parameters needed. The F -test, also known as the “joint hypotheses test”, is used to evaluate if some of the parameters in a function are necessary for fitting. Both tests are performed with SM $\gamma + \text{jets}$ MC simulated samples.

In the current analysis, with a small number of events in high mass region, the F -test results are not completely reliable. So results from the spurious signal test are used as the main criteria to determine the number of free parameter for the background fit and the F -test results are treated as a cross check.

Based on the importance in various $m_{J\gamma}$ mass regions, the fit ranges used for different categories are as follows:

- $Btag$: 800 - 3200 GeV
- $D2$: 800 - 7000 GeV
- $Vmass$: 1000 - 7000 GeV

A reduced fit range can also help to avoid the difficult of modeling backgrounds over a large $m_{J\gamma}$ range with analytic functions.

Following determination of the fit ranges, the limit-setting ranges for various signal regions are determined as given below:

- $Btag$: 1000 - 3000 GeV
- $D2$: 1000 - 6800 GeV
- $Vmass$: 1200 - 6800 GeV

After the spurious signal test and F -test, the number of parameters chosen for each signal search category are shown in Table 5.5:

Table 5.5: A summary of the number of parameters used for the background fit functions in different channels and categories.

Channel	btag	d2	vmass
$gg \rightarrow X(spin = 0) \rightarrow \gamma Z(\rightarrow q\bar{q})$	2	2	3
$qq \rightarrow X^\pm(spin = 1) \rightarrow \gamma W^\pm(\rightarrow qq')$	-	2	3
$qq \rightarrow X(spin = 2) \rightarrow \gamma Z(\rightarrow q\bar{q})$	2	3	3
$gg \rightarrow X(spin = 2) \rightarrow \gamma Z(\rightarrow q\bar{q})$	2	2	3

Figures 5.18 and 5.19 show $m_{J\gamma}$ distributions of simulated $\gamma + Jet$ events in different signal regions with fitted curves overlaid. The distributions are determined by performing unbinned maximum likelihood fits with the functions defined in Eq.5.2 to $m_{J\gamma}$ distributions of MC simulated events.

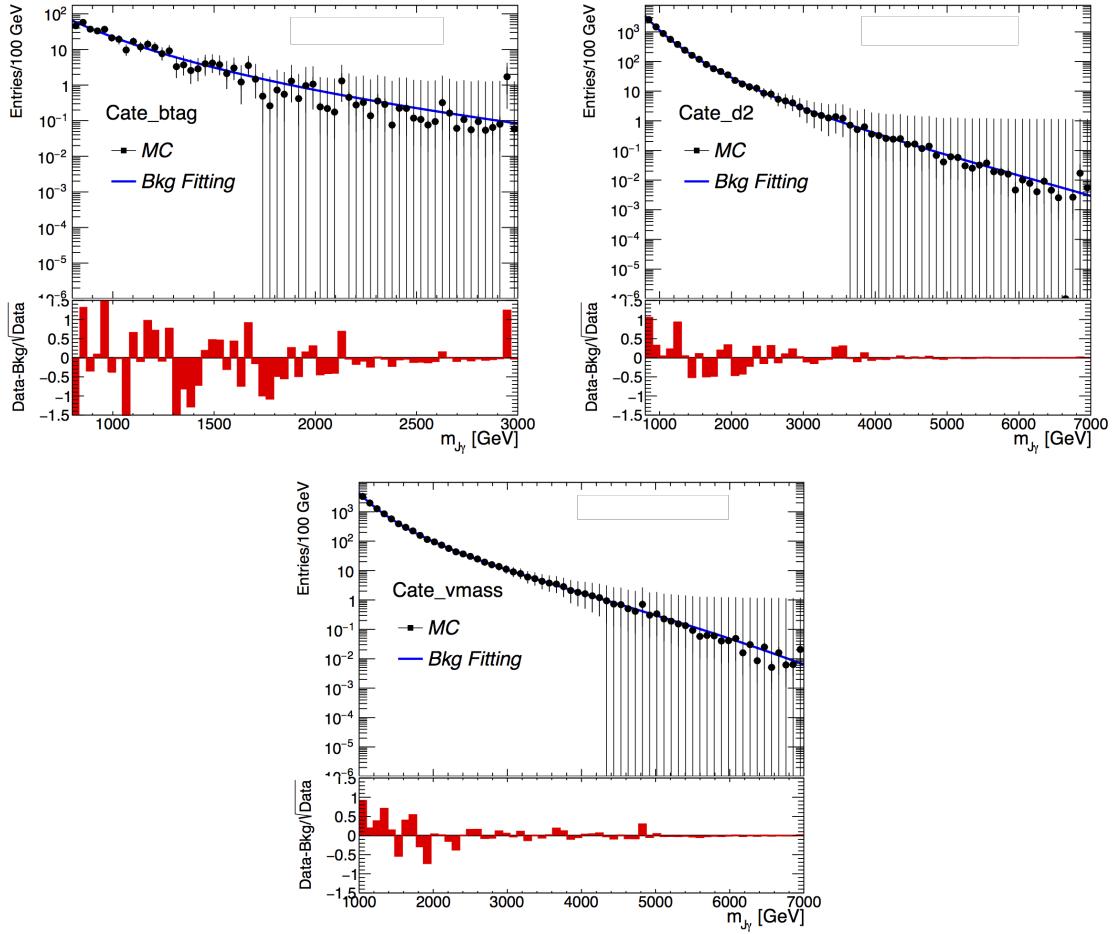


Figure 5.18: Background fit to the $m_{J\gamma}$ mass spectrum from $\gamma + Jet$ MC in the signal region of spin-0 $Z\gamma$ channels. A different fit range is applied to each of the categories. The lower insets of the Figure show the significance, defined as the z -value as described in Ref. [51].

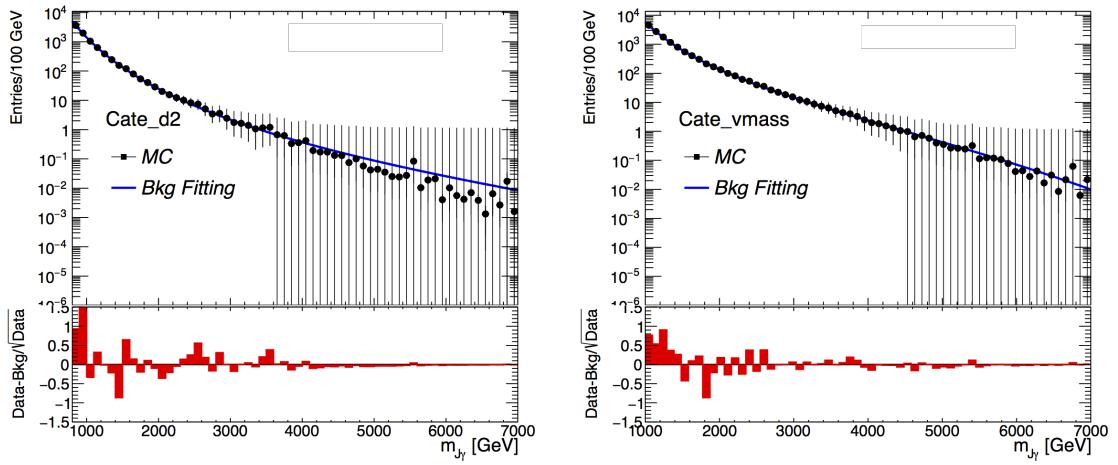


Figure 5.19: Background fit to the $m_{J\gamma}$ mass spectrum from $\gamma + Jet$ MC in the signal region of spin-1 $W\gamma$ channels. A different fit range is applied to each of the categories. The lower insets of the Figure show the significance, defined as the z -value as described in Ref. [51].

5.6 Systematic Uncertainty

The sensitivity of the search for BSM resonance signals is affected by various experimental systematic uncertainties in addition to purely statistical considerations. These systematic uncertainties are included as constrains on nuisance parameters in the final fits to data using a procedure described in Section 5.7. In this search, statistical uncertainties are always the dominating contribution. Spurious signal contribution for the bias in background modeling is more impactful below 2 TeV, and quickly vanishes after 3 TeV. Systematic uncertainties also contribute by several percent, not strongly dependent on resonance mass. A summary of the systematic uncertainties are presented below.

5.6.1 Background systematic uncertainties

The functional form used for modeling the $m_{J\gamma}$ background distribution (see Eq. 5.2) contains parameters determined from fits to the $m_{J\gamma}$ data distribution. The parameters in the background functions are allowed to vary, and are included as nuisance parameters in the likelihood fitting function. An additional systematic uncertainty in the background modeling comes from the inclusion of the spurious signal estimate $|N_{ss}|$ described in Section 5.5.1. To be conservative, $2 \times |N_{ss}|$ is used as a systematic uncertainty in the fit.

5.6.2 Signal systematic uncertainties

The systematic uncertainties are uncertainties generated by reconstructing or calibrating variables for physics objects. The impact of these signal systematic uncertainties are evaluated on signal shape and efficiency, and then introduced into the likelihood function as constrains on signal nuisance parameters as shown in Section 5.7. The included signal systematic uncertainties are listed below:

- *Luminosity*: The amount of data is quantified by the integrated luminosity of the 13 TeV pp collisions. This is measured to be 139 fb^{-1} with an uncertainty of 1.7%. [52] There is no influence from luminosity uncertainty on signal shape, and the efficiencies are varied by about 2% independent of resonance mass.
- *Photon energy*: This uncertainty comes from the photon reconstruction and calibration procedure. The ATLAS photon/electron (e/γ) group evaluated this uncertainty and reported two components. One called the photon energy scale (PES), and the other called the photon energy resolution (PER). Due to the limited range of photon energy calibration up to 1 TeV, the high energy photons are in the extrapolated regime, and therefore use conservative predictions. This greatly increases the influence of PES at high resonance mass region, making it one of the dominating uncertainties. The influence of PER is low over the full search range.
- *Photon identification and isolation*: The identification and isolation tagging of photons can also introduce a source of uncertainty. It can slightly change the baseline selection efficiencies, but is overall insignificant.
- *Jet energy and mass*: The sources of uncertainties on jet energy and mass is mainly from calibration. [53] The following sources contribute to this uncertainty: differences between simulation and data, dependence of MC generators, contribution propagated from the tracks associated to the jet, statistical uncertainty in data used for measurements, and the potential bias from the calibration method. This is our major uncertainty for the full $m_{J\gamma}$ search range especially in the low resonance mass region, where it is almost the only signal systematic uncertainty.
- *W/Z boson and beauty hadron tagging*: These are the uncertainties on W/Z

boson tagging or beauty hadron tagging. The boson tagging uncertainties are provided by Jet Substructure group in ATLAS. The uncertainty of b-tagging is measured from data in $t\bar{t}$ enhanced region compared to simulation prediction. In this search, both of the jet tagging uncertainties are negligible in the full search range.

5.7 Statistical analysis implementation

The data is interpreted with a frequentist analysis as described in section 4.2. In this search, a maximum extended unbinned likelihood fit is performed on the $m_{J\gamma}$ distribution. The likelihood is defined as a global product containing the Poisson model for observed event yield in each category, the product of probabilities for every event according to $m_{J\gamma}$ model based on the sum of signal (\mathcal{S}) and background (\mathcal{B}) shapes and the products of probabilities of auxiliary measurements with their priors distribution (\mathcal{G}). The full likelihood function is written as:

$$\begin{aligned}
\mathcal{L}(\overrightarrow{m_{J\gamma}^{\text{obs}}} | \sigma, \vec{\theta}, \theta^{\text{SS}}, \vec{N^b}, \vec{p}) = \prod_{c \in \mathbb{C}} \left\{ \text{Pois}(N_c^{\text{obs}} | N_c^s(\sigma, \vec{\theta}) + N_c^{SS} + N_c^b) \right. \\
\prod_{i=1}^{N_c^{\text{obs}}} \left[\left(\frac{N_c^s(\sigma, \vec{\theta}) + N_c^{SS}(\theta_c^{SS})}{N_c^s(\sigma, \vec{\theta}) + N_c^{SS}(\theta_c^{SS}) + N_c^b} \right) \mathcal{S}(m_{J\gamma}^{c,i,\text{obs}} | \vec{\theta}) + \right. \\
\left. \left(\frac{N_c^b}{N_c^s(\sigma, \vec{\theta}) + N_c^{SS}(\theta_c^{SS}) + N_c^b} \right) \mathcal{B}(m_{J\gamma}^{c,i,\text{obs}} | \vec{p^c}) \right] \left. \right\} \times \\
\prod_{s \in \mathbb{S}} \mathcal{G}(0 | \theta_s, 1) \prod_{c \in \mathbb{C}} \mathcal{G}(0 | \theta_c, 1).
\end{aligned} \tag{5.3}$$

The $\overrightarrow{m_{J\gamma}^{\text{obs}}} = \{m_{J\gamma}^{1,1,\text{obs}}, \dots, m_{J\gamma}^{c,i,\text{obs}}, \dots\}$ is a set of observed $m_{J\gamma}$ in data, c is the index of category and i the index of events in each category. $\vec{\theta}, \theta^{\text{SS}}$ are nuisance parameters related to signal and spurious signal. N_c^s, N_c^{SS}, N_c^b is the observed number of signal, spurious signal or background events in each category. This cross section σ is an abbreviation for $\sigma(pp \rightarrow X \rightarrow W/Z(\rightarrow \text{hadrons}) + \gamma)$, which is the parameter of interest for this search. It includes the cross section of resonance production $\sigma(pp \rightarrow X)$ and branching fractions of $X \rightarrow W/Z + \gamma$ and $W/Z \rightarrow \text{hadrons}$. The background shape parameters $\vec{p^c} = (p_1^c, p_2^c, \dots)$ are float during the fit to the data

and are uncorrelated among categories.

The first line of Eq. 5.3 is the probability of finding the observed number of events in each category given the prediction from signal and background models determined by nuisance parameters $\vec{\theta}$ and \vec{p}^c . The second and third lines are likelihood for a specific event i appears in observation (either from signal or background). And the last line are the constraints on various nuisance parameters for systematics θ_s and spurious signal contributions for each categories θ_c . These constraints are assumed to be a Gaussian distribution around the estimated central value of the parameters.

The analysis proceeds first with the so-called *search-phase* and second with the so-called *limit-setting* phase. In the search-phase, a global maximum likelihood fit is performed with data distributions according to the model mentioned in Eq. 5.3 to extract both signal and background contributions. Secondly, to test the compatibility of data and background-only hypothesis, the local p -values are calculated for potential signals with m_X over full search range at 20 GeV per step. For the given dataset, the observed p -value is defined as the probability of having equal and higher signal events as estimated in this dataset from the background only model.

In the limit-setting phase, an exclusive cross section limit is derived from observed data and presented as the 95% confident level (CL) upper limit on cross section following the CL_s prescription [38]. The limits are calculated in low mass region with 20 GeV per step while increased to 100 GeV per step in high mass region.

5.8 Expected limits and the impact of systematic uncertainties

Before looking at the data for the presence of BSM $X \rightarrow Z\gamma$ and $X \rightarrow W\gamma$ resonance signals, the $m_{J\gamma}$ distributions in the data were "blinded" in order to avoid any bias in the search procedures. Before "unblinding" the data, the expected limits on the cross

sections for BSM resonances (assuming no signals) are obtained using MC generated events for the SM backgrounds. The events are scaled to the same statistics expected for the data. The predictions for the expected cross section limits are presented in this section, and the actual observed search results and cross section limits are presented in Chapter 6.

Figure 5.20 shows the expected limits on the production cross section for spin=0 $Z\gamma$ signals. The ratio curves shown in the middle quantify the limit changes caused by the spurious signal contribution, and in the bottom panel from the other systematic uncertainties. The curves are produced with a step of 200 GeV on resonance mass. Similar results for spin=1 $W\gamma$ signals are shown in Figure 5.23. For spin=2 $X \rightarrow Z\gamma$ signals, the results are presented in Figures 5.21 and 5.22 for $qq \rightarrow X$ and $gg \rightarrow X$ production modes respectively. By including all systematic uncertainties, the combined limits for spin=0 $Z\gamma$ increase by up to 8%. For spin=1 $W\gamma$, systematic uncertainties worsen the limits by 2%. For both spin=2 $Z\gamma$ signals, the systematic uncertainties worsen the limit by 5%. The spurious signal uncertainties affect the limit up to 30% in low mass region with the effects decreasing at higher mass points. The difference of the spurious signal effects on limits for the $Z\gamma$ signal and the $W\gamma$ signal is due to the btag category.

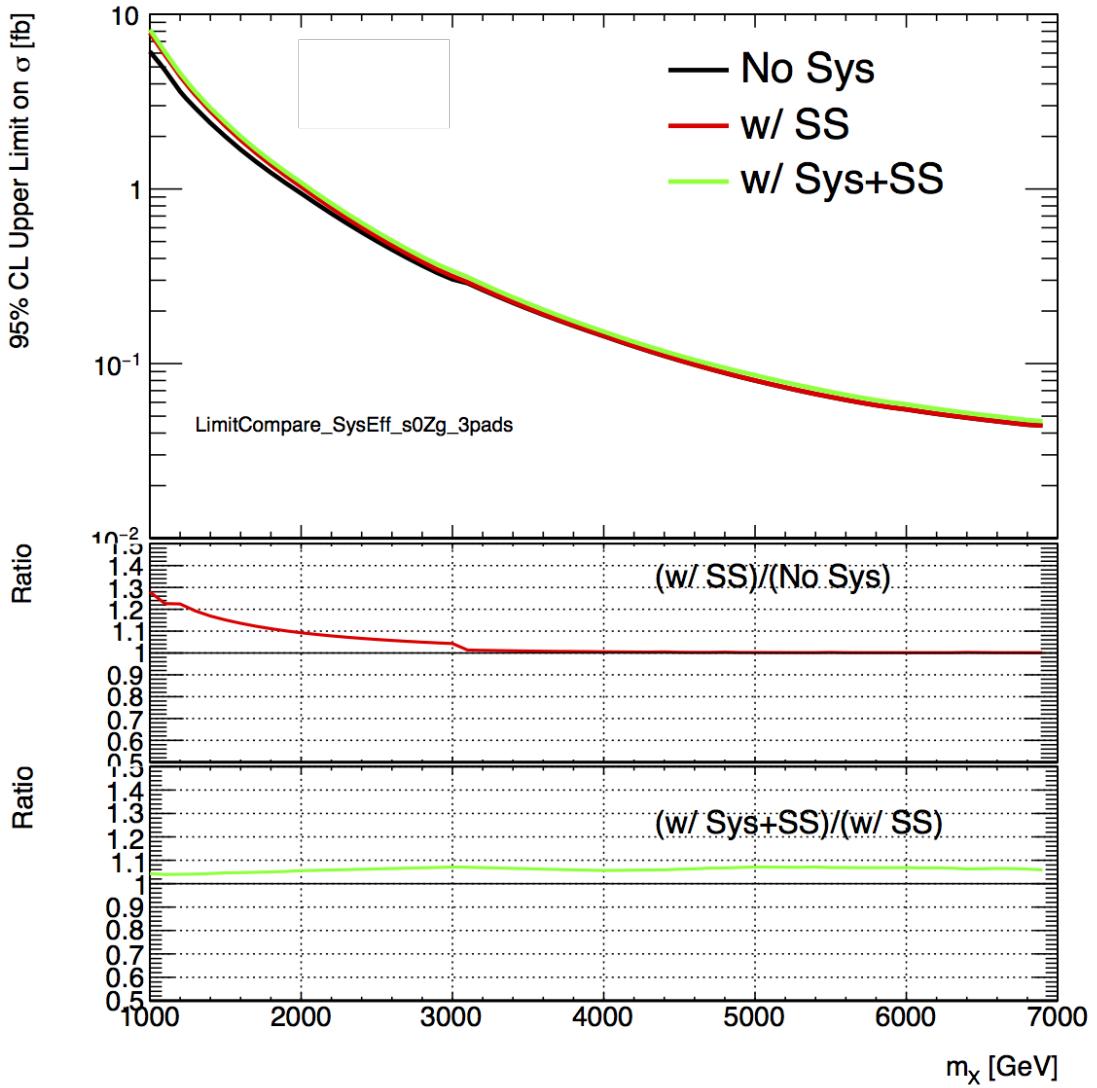


Figure 5.20: This is the expected 95% CL cross section limits for $\sigma(p + p \rightarrow X(\text{spin} = 0)) \times \text{BR}(X \rightarrow Z\gamma)$ estimated with only the SM background MC. The production mode is $gg \rightarrow X$. Black lines in upper pad show the expected limit without including experimental systematic uncertainties for spin=0 $X \rightarrow Z\gamma$ signals and the green lines are the limits with all experimental uncertainties included. The red line shows the limits with spurious signal. The curve in middle pad corresponds to the ratio between limits with spurious signal and limits without any systematic uncertainty and indicates the contribution from spurious signal. The curve shown in bottom pad is derived as the ratio of limits with all systematic uncertainties included to limits with only spurious signal uncertainty. This ratio shows how significant the experimental systematic uncertainty are in the limit setting.

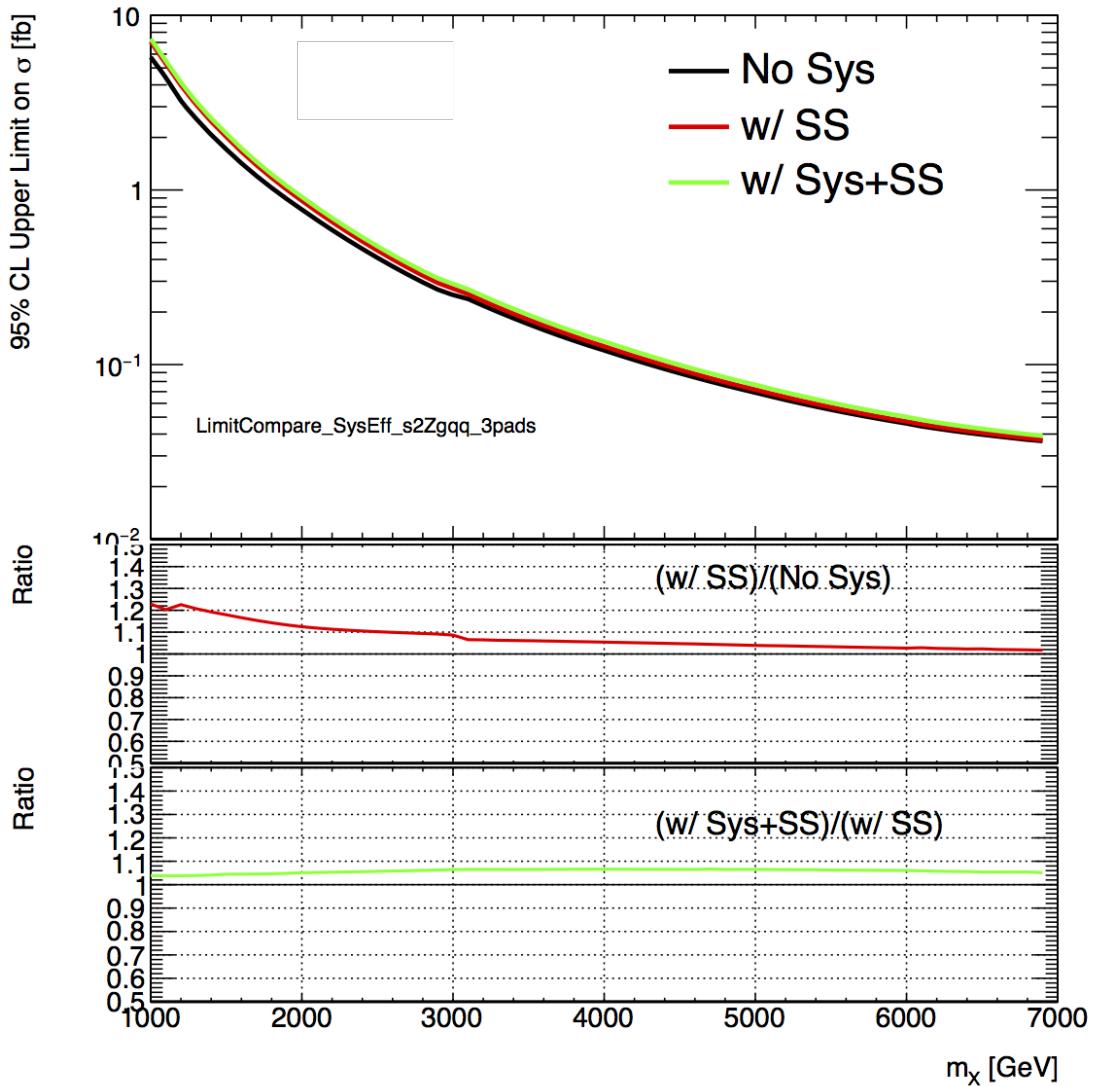


Figure 5.21: This is the expected 95% CL cross section limits for $\sigma(p + p \rightarrow X(\text{spin} = 2)) \times \text{BR}(X \rightarrow Z\gamma)$ estimated with only the SM background MC. The production mode is $q\bar{q} \rightarrow X$. Black lines in upper pad show the expected limit without including experimental systematic uncertainties for spin=2 $qq \rightarrow X \rightarrow Z\gamma$ signals and the green lines are the limits with all experimental uncertainties included. The red line shows the limits with spurious signal. The curves in bottom pads correspond the limit changes wrt. stat. only limits. The curve in middle pad corresponds the ratio between limits with spurious signal and limits without any systematic uncertainty and indicates the contribution from spurious signal. The curve shown in bottom pad is derived as the ratio of limits with all systematic uncertainties included to limits with only spurious signal uncertainty. This ratio shows how significant the experimental systematic uncertainty are in the limit setting.

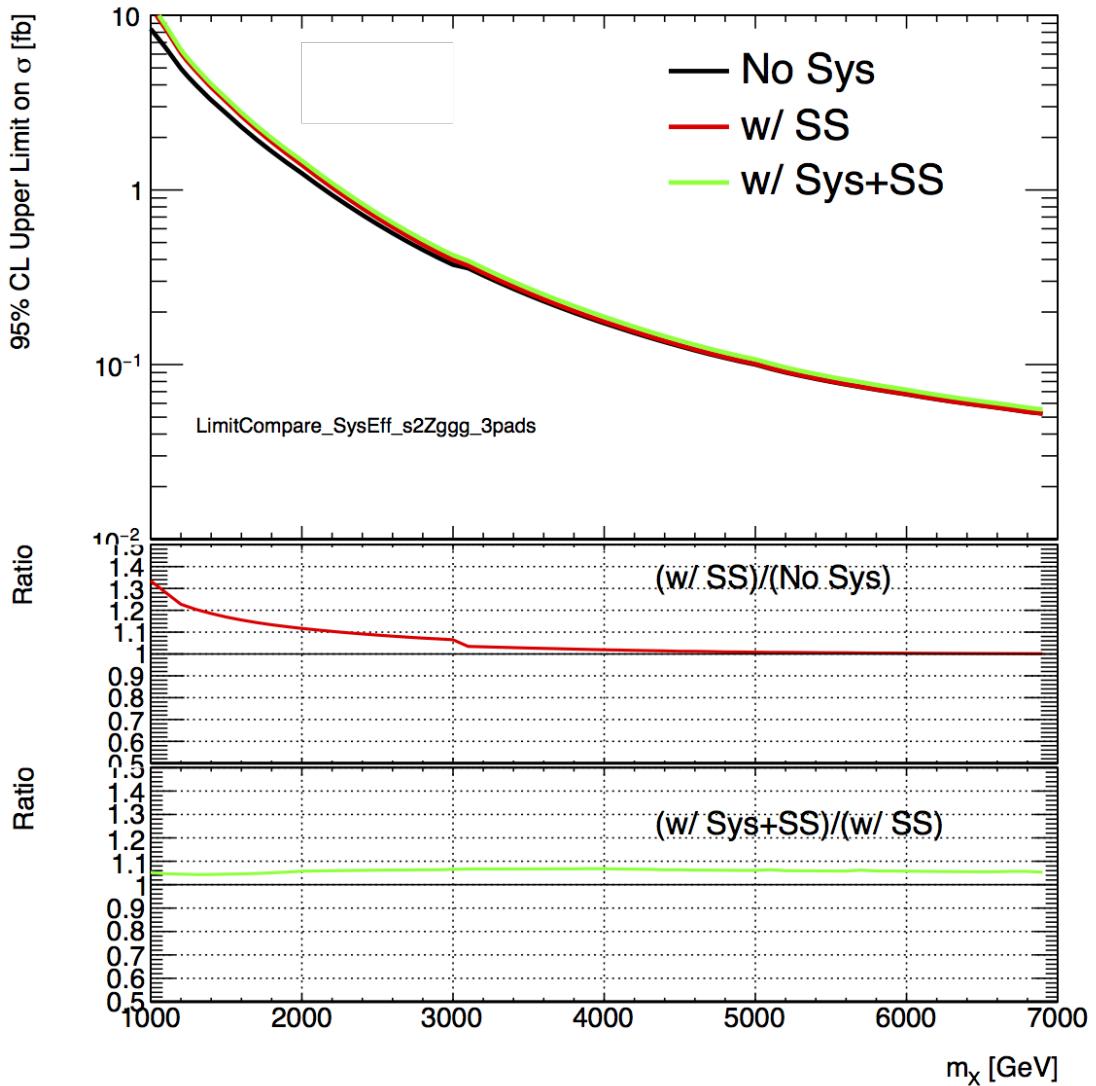


Figure 5.22: This is the expected 95% CL cross section limits for $\sigma(p + p \rightarrow X(\text{spin} = 2)) \times \text{BR}(X \rightarrow Z\gamma)$ estimated with only the SM background MC. The production mode is $gg \rightarrow X$. Black lines in upper pad show the expected limit without including experimental systematic uncertainties for spin=2 $gg \rightarrow X \rightarrow Z\gamma$ signals and the green lines are the limits with all experimental uncertainties included. The red line shows the limits with spurious signal. The curves in bottom pads correspond the limit changes wrt. stat. only limits. The curve in middle pad corresponds the ratio between limits with spurious signal and limits without any systematic uncertainty and indicates the contribution from spurious signal. The curve shown in bottom pad is derived as the ratio of limits with all systematic uncertainties included to limits with only spurious signal uncertainty. This ratio shows how significant the experimental systematic uncertainty are in the limit setting.

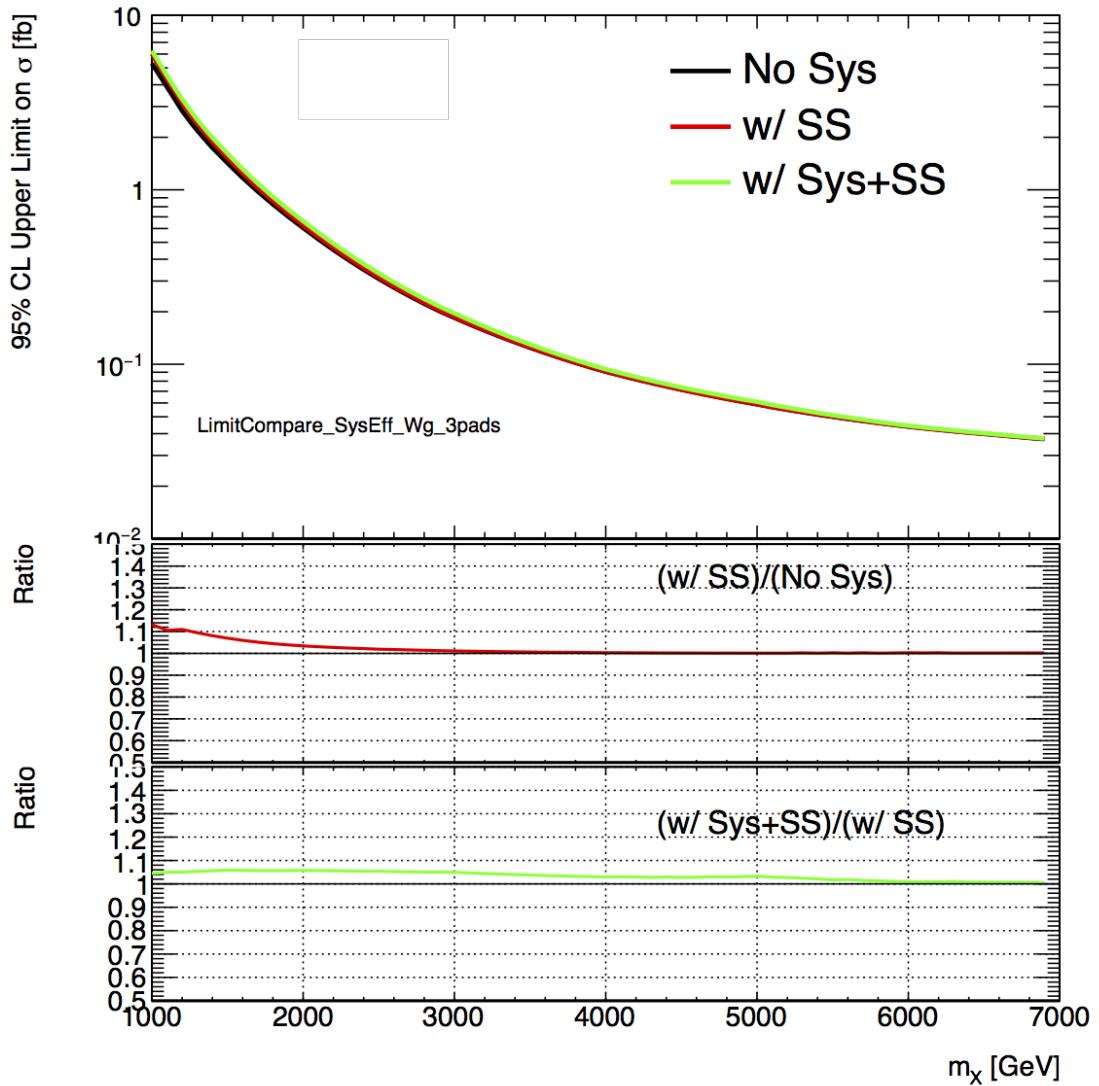


Figure 5.23: This is the expected 95% CL cross section limits for $\sigma(p + p \rightarrow X(\text{spin} = 1)) \times BR(X \rightarrow W\gamma)$ estimated with only the SM background MC. The production mode is $qq' \rightarrow X$. Black lines in upper pad show the expected limit without including experimental systematic uncertainties for spin=1 $W\gamma$ signals and the green lines are the limits with all experimental uncertainties included. The red line shows the limits with spurious signal. The curves in bottom pads correspond the limit changes wrt. stat. only limits. The curve in middle pad corresponds the ratio between limits with spurious signal and limits without any systematic uncertainty and indicates the contribution from spurious signal. The curve shown in bottom pad is derived as the ratio of limits with all systematic uncertainties included to limits with only spurious signal uncertainty. This ratio shows how significant the experimental systematic uncertainty are in the limit setting.

Chapter 6

Results of the $W/Z + \gamma$ heavy resonance search

After all event selection criteria are applied, the number of data events used for the BSM boson searches are summarized in Table 6.1. For each channel and category, distributions of the photon-jet invariant mass $m_{J\gamma}$ are used to search for $Z\gamma$ and $W\gamma$ resonances by employing the analysis techniques described in Chapters 4 and 5. As shown below, no deviations from smoothly falling $m_{J\gamma}$ distributions are observed, consistent with SM predictions. Therefore cross section limits are calculated for the production of BSM $W\gamma$ and $Z\gamma$ resonances with masses between 1.0 and 6.8 TeV.

Table 6.1: Observed number of events in each category for data after all selection criteria have been applied. The dataset luminosity is 139 fb^{-1} and the range of $m_{J\gamma}$ mass is from 800 to 7000 GeV.

Signal Regions	Btag	D2	Vmass	Total
$gg \rightarrow X(\text{spin} = 0) \rightarrow Z\gamma$	436	5659	20728	26823
$gg \rightarrow X(\text{spin} = 2) \rightarrow Z\gamma$	436	10772	32281	43489
$q\bar{q} \rightarrow X(\text{spin} = 2) \rightarrow Z\gamma$	436	5618	18264	24318
$qq' \rightarrow X(\text{spin} = 1) \rightarrow W\gamma$	–	6373	25146	31519

Compared to a similar search done in 2018 with 36 fb^{-1} of data collected with the ATLAS detector during 2015 and 2016, the result of this search is greatly improved over the full $m_{J\gamma}$ range. [47] In the channels with $Z\gamma$ final states, the improvement is about 50% in the low mass region, and by a factor of 3 to 4 in the high mass region. The $W\gamma$ channel is improved by over a factor of 2 over the full mass region. Such an improvement is achieved by not only the increased amount of data collected by

the ATLAS detector, but also by the search methodology re-optimized and improved boson tagging.

This is the first search for $W\gamma$ and $Z\gamma$ resonances with such a large dataset, but there are several previous studies. A publication from CMS on a $Z\gamma$ heavy resonance search using 36 fb^{-1} data [54] presented combined results with both hadronic and leptonic decay modes of the Z boson. In the range of resonance mass below 2 TeV, the sensitivity of the CMS search is similar to the results presented in this thesis. However, in the mass range higher than 2 TeV, this result become more sensitive, and is extended to 6.8 TeV which is higher than the upper bound of 4 TeV for the CMS results. There are also searches from ATLAS for resonances in $Z\gamma$ final states with the Z boson decay to leptons. [55] This search is more sensitive in the low mass region below 1 TeV, but start losing sensitivity quickly after that threshold. Therefore, a combination of leptonic and hadronic channels would be informative.

6.1 Data distributions of $m_{J\gamma}$

Figures 6.1, 6.2, 6.3 and 6.4 show the data $m_{J\gamma}$ distributions. Superimposed are the background only fit functions for spin-0 $Z\gamma$, spin-2 gg $Z\gamma$, spin-2 qq $Z\gamma$ and spin-1 $W\gamma$ signal searches. The local significance of deviations of the data from the background-only function for each bin is also shown in the plots. A good agreement between data and the background fit function is observed.

6.2 Significance scan for $W\gamma$ and $Z\gamma$ resonances

As described in Section 5.7, the *search phase* is performed first and presented as significance p -value scan. Figure 6.5 show the local p -value scan results using data for each mass point (20 GeV per step) and their corresponding significances. The largest local significance in the spin-0 $Z\gamma$ channel appears for $m_X = 3.64 \text{ TeV}$ with

a significance about 2.48σ . For spin-2 gg $Z\gamma$ channel, a signal of mass 3.58 TeV gives the largest local significance of 2.24σ . Similarly, the $m_X = 3.56$ TeV signal hypothesis gives a 1.82σ local significance as the largest one in spin-2 qq $Z\gamma$ channel. For the spin-1 $W\gamma$ signal, the largest significance of 1.52σ occurs at $m_X = 2.82$ GeV.

For the mass points with largest significance in each signal channel, closer studies are described in Appendix. A.6.

6.3 Measured limits on the cross-section for production of BSM $X^\pm \rightarrow W^\pm\gamma$ and $X \rightarrow Z\gamma$ resonances

Since no significant evidence for the existence of BSM signals over the search range is found, the upper limit on cross sections is derived and presented. Figure 6.6 shows the observed limits for the cross section $\sigma(p+p \rightarrow X(\text{spin} = 0)) \times BR(X \rightarrow Z\gamma)$. The solid line in the top plot shows the observed limit which lies around the expected limit calculated from the SM background only MC with data-like statistical uncertainty (assuming no signal observed). The green and yellow bands represent the uncertainty on this expected limit with $\pm 1\sigma$ and $\pm 2\sigma$. As shown in the plot, the observed limit is almost covered by the 2σ band in the full search range, which implies no significant difference between observation from data and expectation from the SM background. Furthermore, the peaks and valleys in the observed limit curve match with the p-value curve from the significance scan. Since the large p-value mass points have been checked separately, there is no concern about the mass points with some deviation between observation and expectation. On the two plots in the bottom of Figure 6.6, the limit for each category is superposed with the combined limit for both expected and observed limits. According to these two plots, the D2 category is dominating over the full range. The Btag category is also impactful at low mass, but stops at 3 TeV

which leaves a small jump in the expected limit curve. Figure 6.7, 6.8 and 6.9 show the cross section limits for the other three signal channels, which have no significant deviation from the background only expectation too. In summary, no evidences or hints are discovered in this search for heavy $W/Z + \gamma$ resonances.

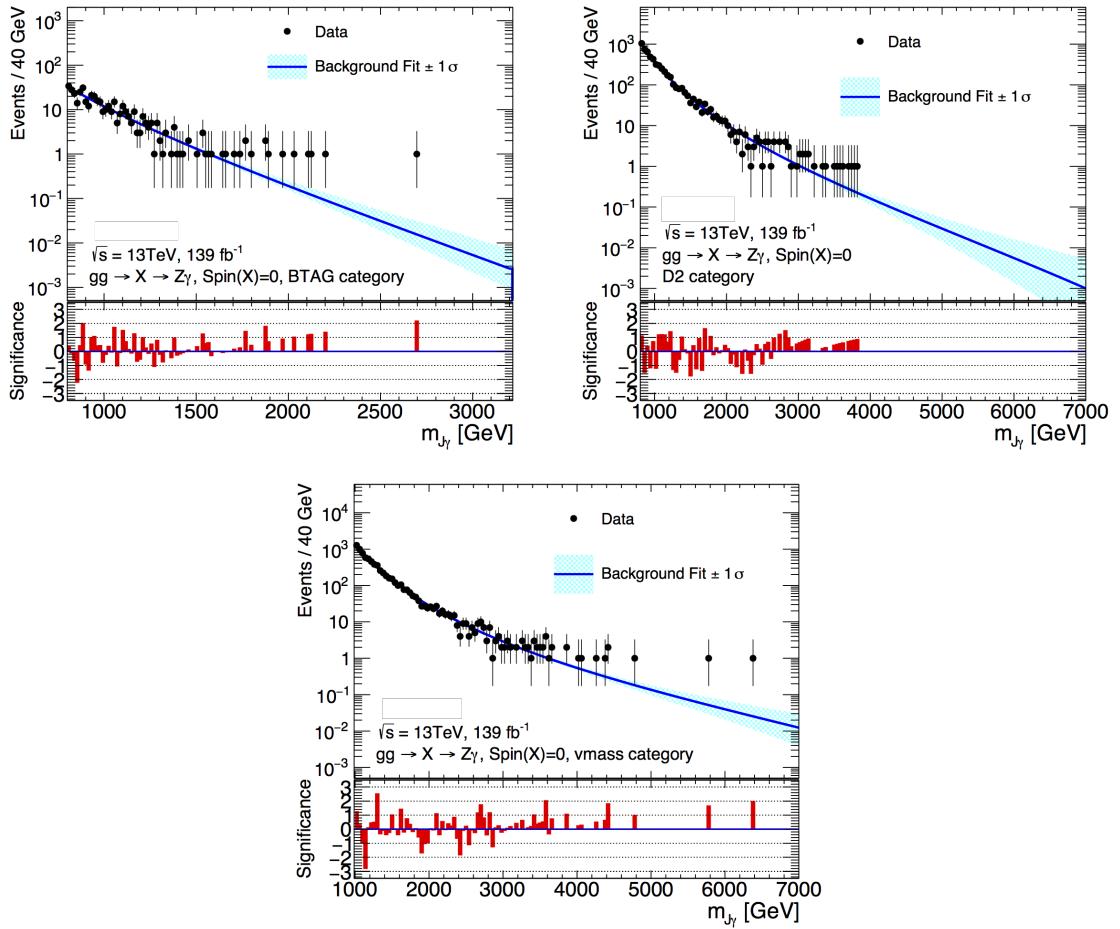


Figure 6.1: Data distributions in spin-0 $Z\gamma$ (a) btag, (b) d2 and (c) vmass categories. Upper pad in each plot also shows the background only fit function (blue line) and its $\pm 1\sigma$ band (light blue band). Lower pad presents the local significance in each bin.

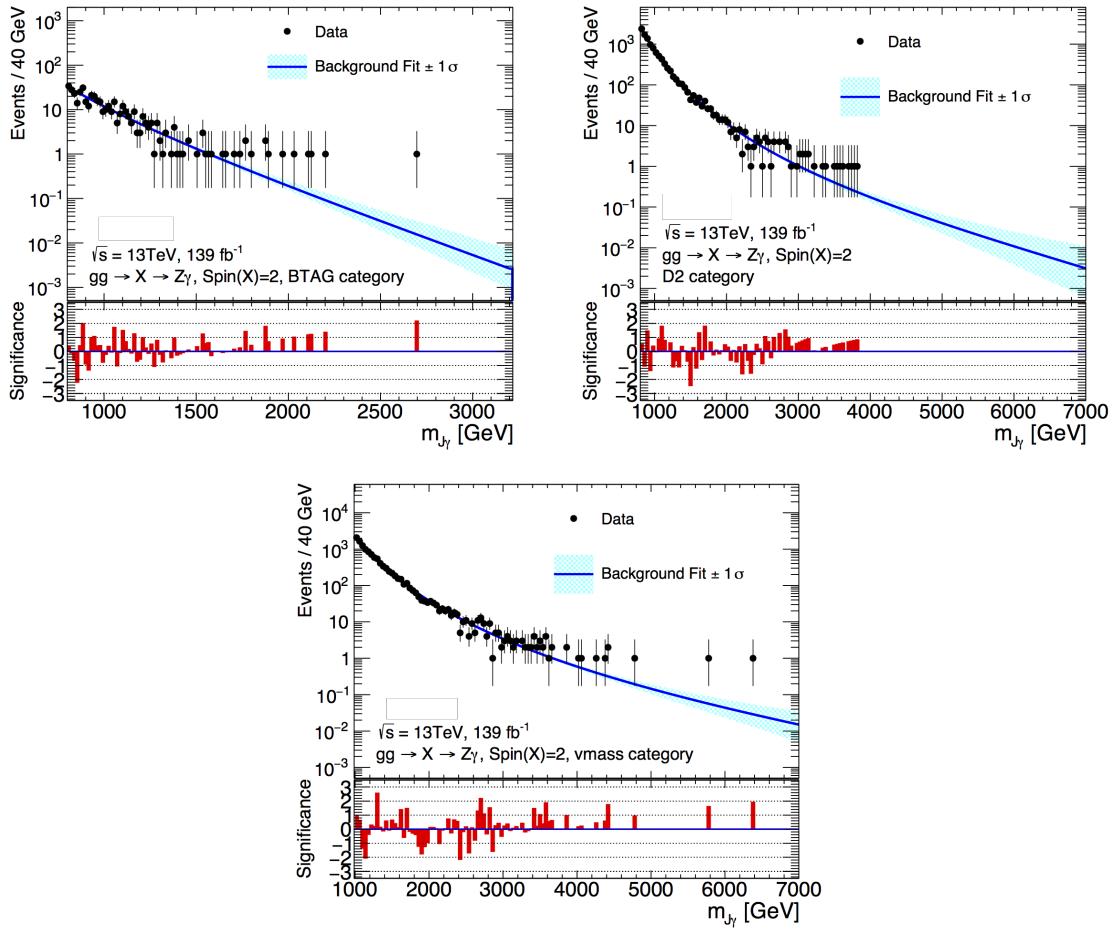


Figure 6.2: Data distributions in spin-2 $gg Z\gamma$ (a) btag, (b) d2 and (c) vmass categories. Upper pad in each plot also shows the background only fit function (blue line) and its $\pm 1\sigma$ band (light blue band). Lower pad presents the local significance in each bin.

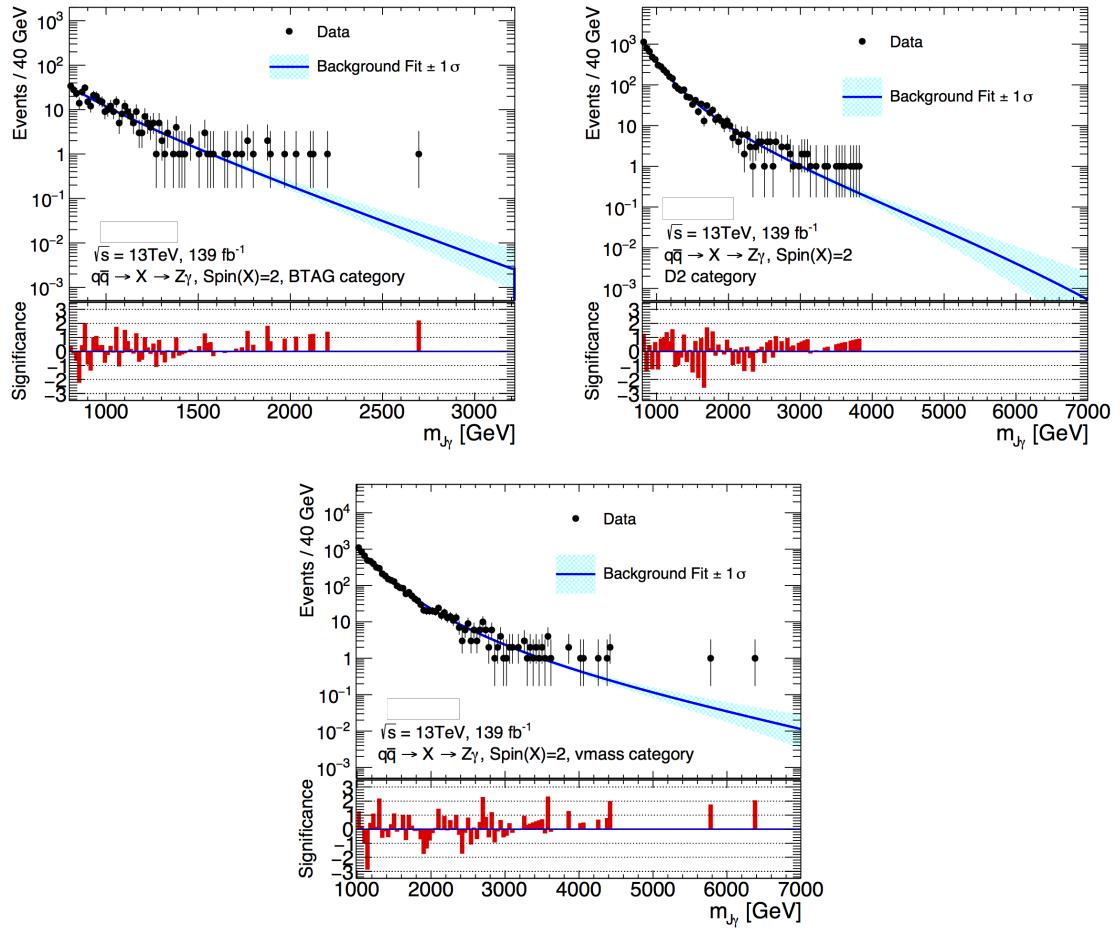


Figure 6.3: Data distributions in spin-2 $q\bar{q}$ $Z\gamma$ (a) btag, (b) d2 and (c) zmass categories. Upper pad in each plot also shows the background only fit function (blue line) and its $\pm 1\sigma$ band (light blue band). Lower pad presents the local significance in each bin.

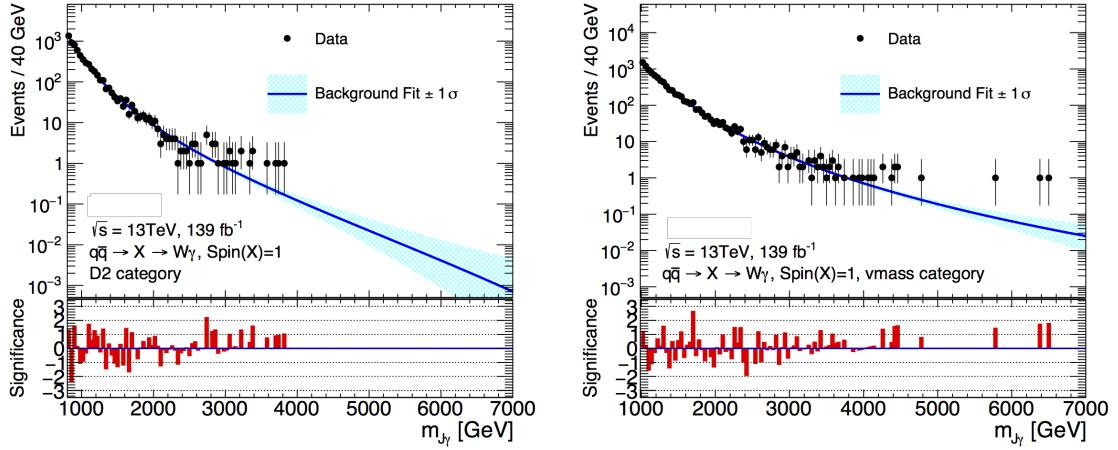


Figure 6.4: Data distributions in $W\gamma$ (a) d2 and (b) vmass categories. Upper pad in each plot also shows the background only fit function (blue line) and its $\pm 1\sigma$ band (light blue band). Lower pad presents the local significance in each bin.

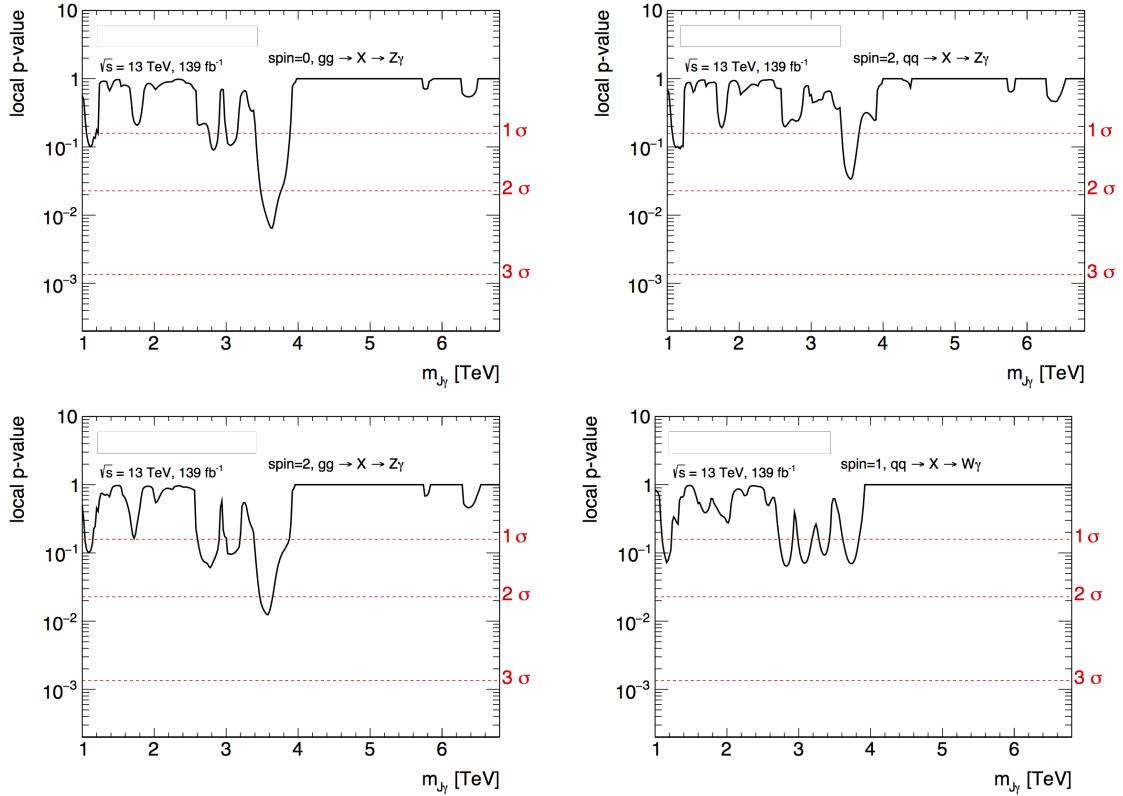


Figure 6.5: Local p -value scan for (a) spin-0 $Z\gamma$, (b) spin-2 qq $Z\gamma$, (c) spin-2 gg $Z\gamma$ and (d) spin-1 $W\gamma$ at different mass points. Dashed gray horizontal lines give the corresponding local significances converted from a normal distribution.

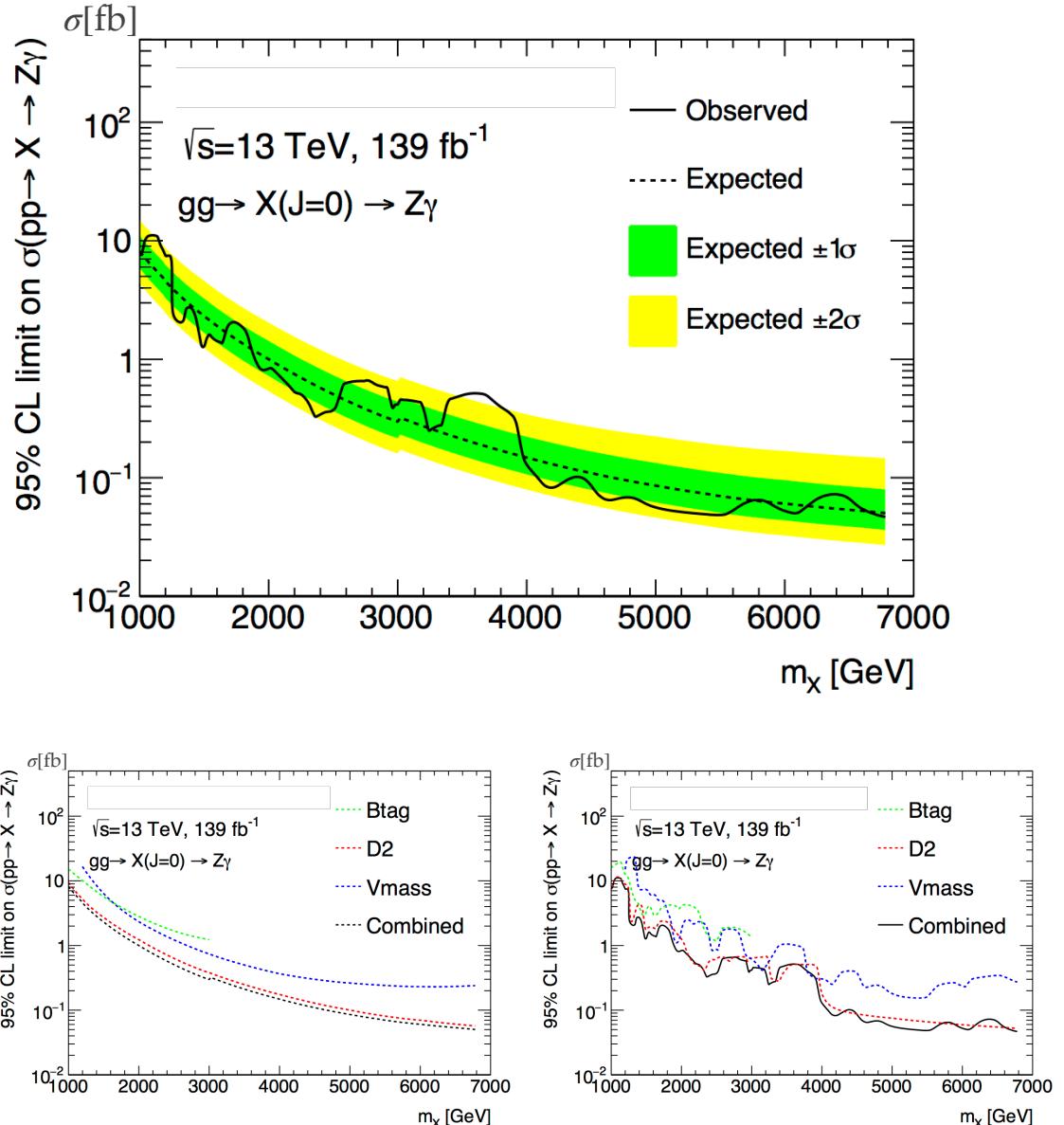


Figure 6.6: 95% CL cross section limits for $\sigma(p+p \rightarrow X(\text{spin}=0)) \times BR(X \rightarrow Z\gamma)$ with the production mode to be $gg \rightarrow X$. The top plot presents the combined observed and expected limits along with the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands for expected limits. Solid line shows the observed limits derived from data and dashed line gives the expected limits from background only MC scaled to data. The bottom left plot gives the expected limits for each category and combining all the categories. The bottom right plot presents the observed limits for each category and combining all the categories.

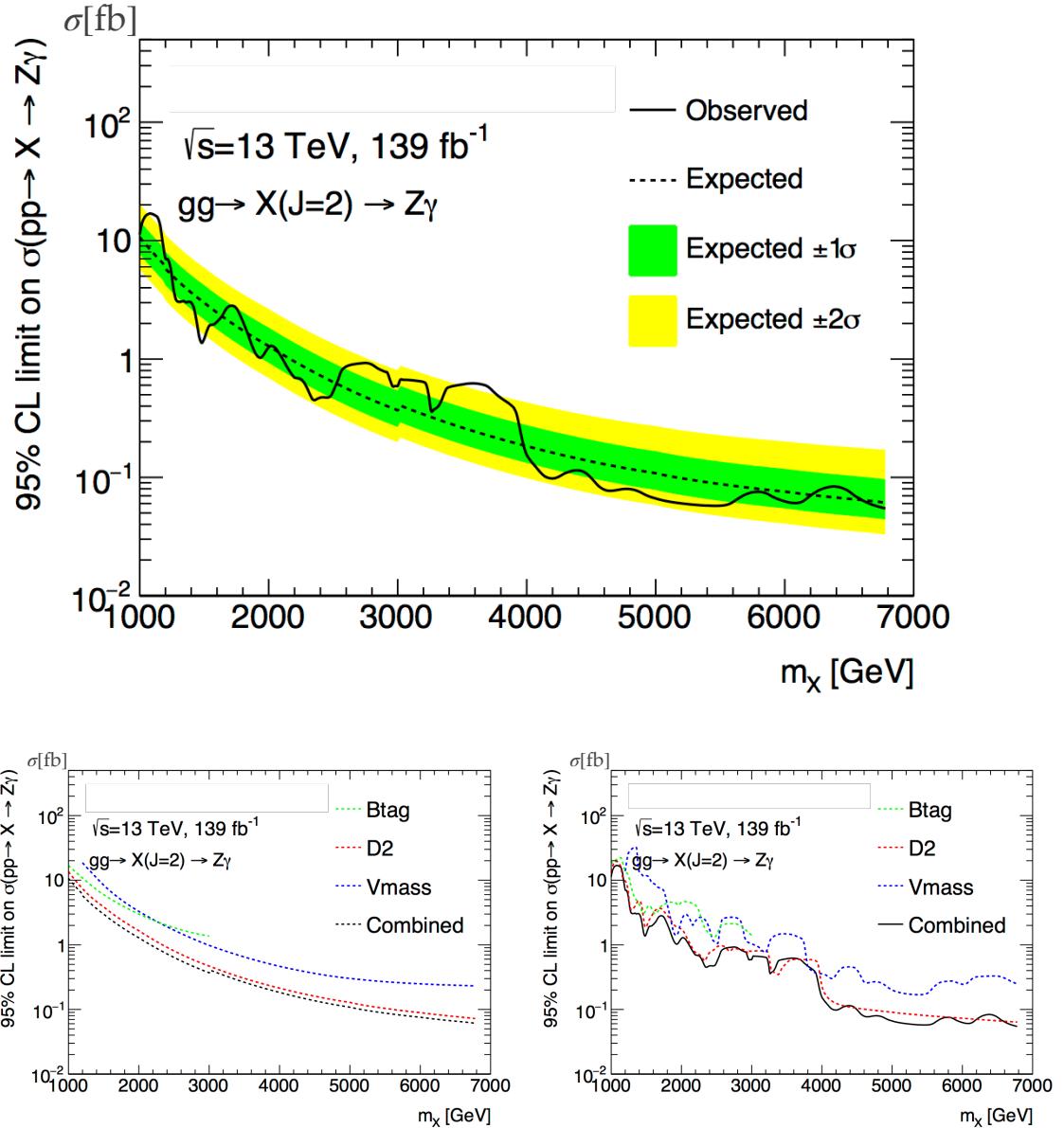


Figure 6.7: 95% CL cross section limits for $\sigma(p+p \rightarrow X(\text{spin} = 2)) \times BR(X \rightarrow Z\gamma)$ with the production mode to be $gg \rightarrow X$. The top plot presents the combined observed and expected limits along with the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands for expected limits. Solid line shows the observed limits derived from data and dashed line gives the expected limits from background only MC scaled to data. The bottom left plot gives the expected limits for each category and combining all the categories. The bottom right plot presents the observed limits for each category and combining all the categories.

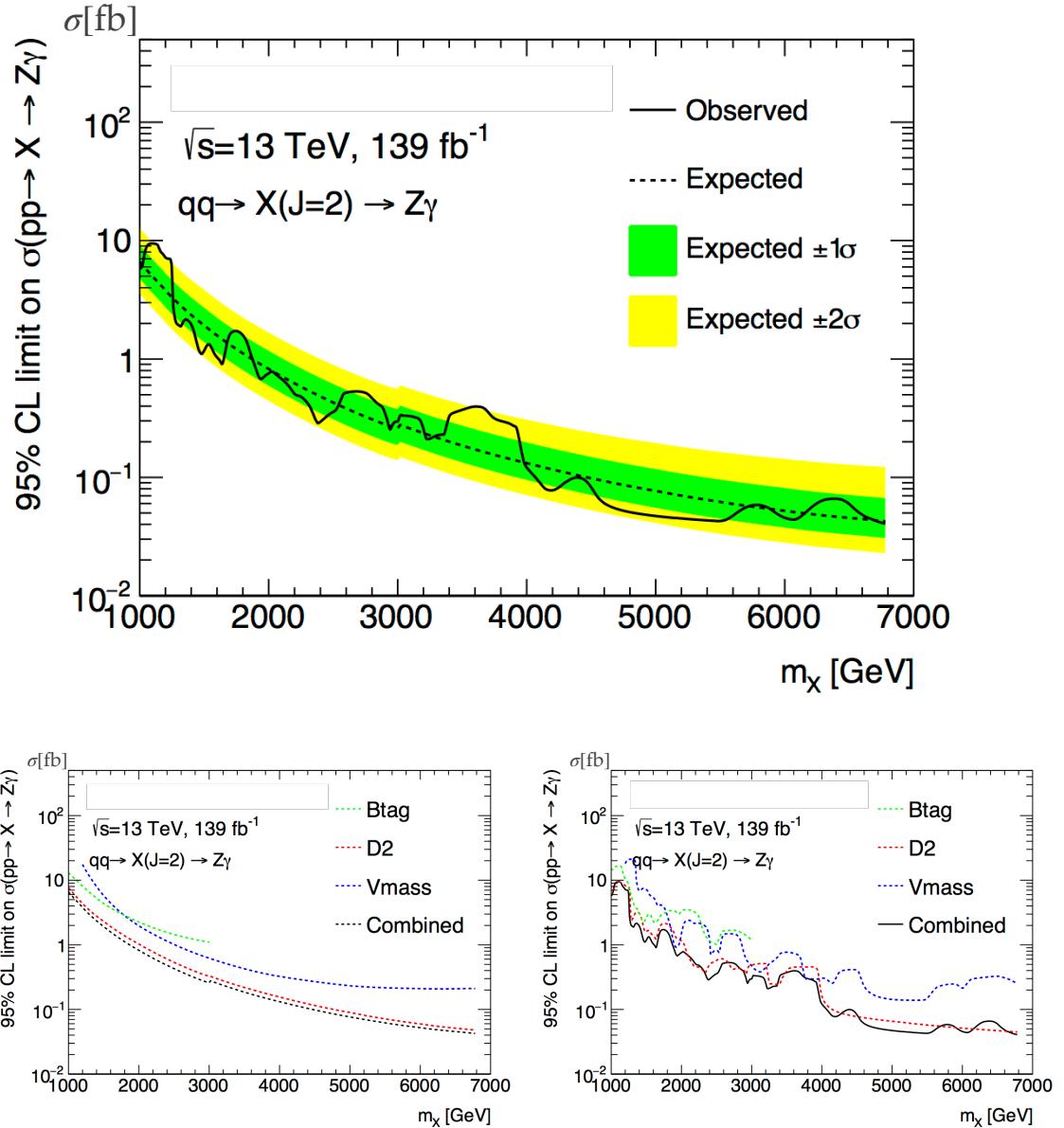


Figure 6.8: 95% CL cross section limits for $\sigma(p+p \rightarrow X(\text{spin} = 2)) \times BR(X \rightarrow Z\gamma)$ with the production mode to be $q\bar{q} \rightarrow X$. The top plot presents the combined observed and expected limits along with the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands for expected limits. Solid line shows the observed limits derived from data and dashed line gives the expected limits from background only MC scaled to data. The bottom left plot gives the expected limits for each category and combining all the categories. The bottom right plot presents the observed limits for each category and combining all the categories.

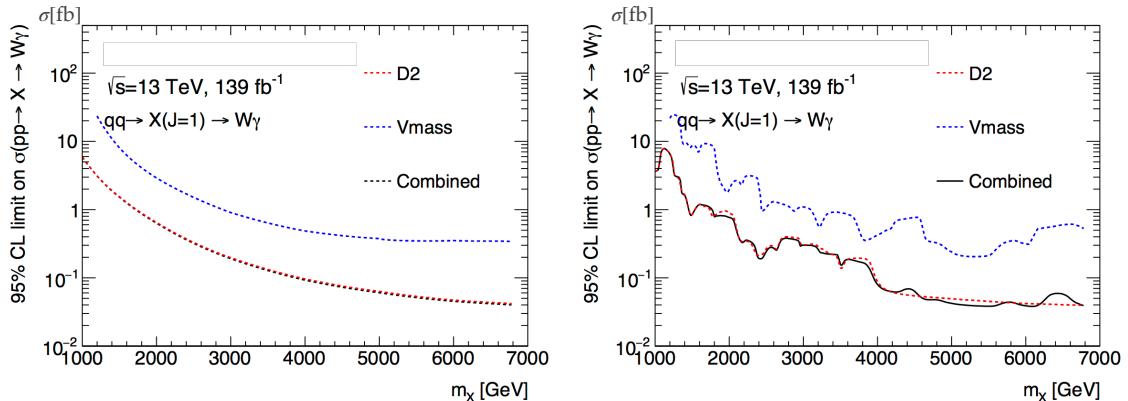
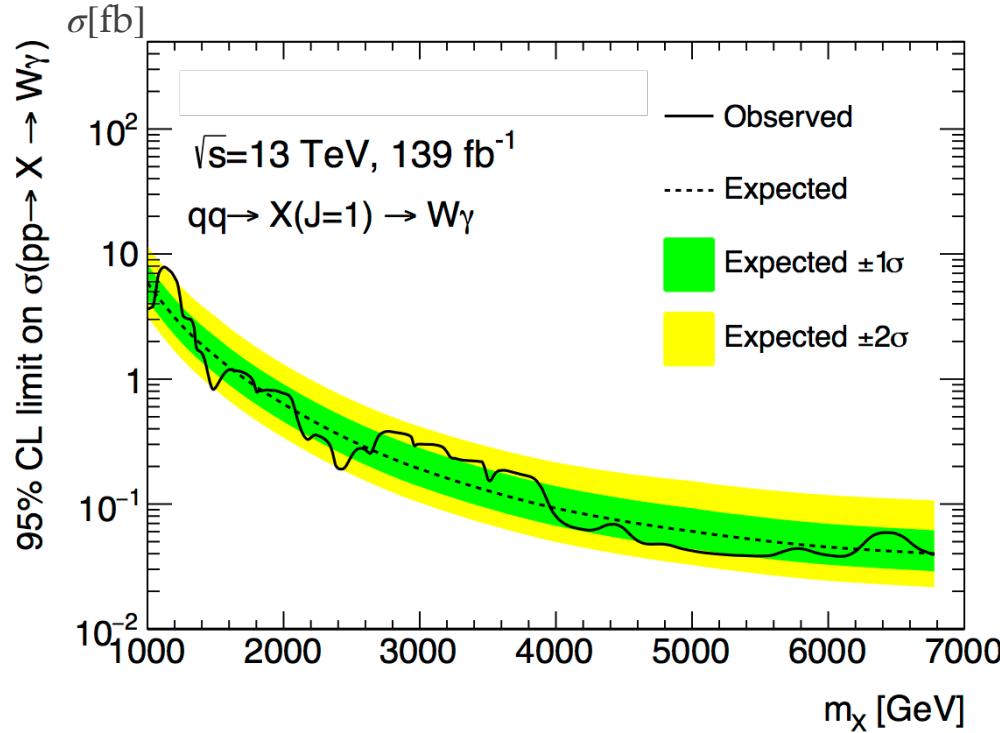


Figure 6.9: 95% CL cross section limits for $\sigma(p + p \rightarrow X^\pm(\text{spin} = 1)) \times \text{BR}(X^\pm \rightarrow W^\pm \gamma)$ with the production mode to be $qq' \rightarrow X$. The top plot presents the combined observed and expected limits along with the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands for expected limits. Solid line shows the observed limits derived from data and dashed line gives the expected limits from background only MC scaled to data. The bottom left plot gives the expected limits for each category and combining all the categories. The bottom right plot presents the observed limits for each category and combining all the categories.

Chapter 7

Conclusions

The search for heavy resonances in final states containing a hadronically decaying W/Z boson and a high energy photon is presented in this thesis. This search is conducted with 139 fb^{-1} of data collected using the ATLAS detector from 2015 to 2018. Energetic photons and boosted large-radius jets within a selected kinematic regime are taken as the main physics objects in this search. Events are further selected and categorized to optimize the sensitivity of the search. BSM signals and SM backgrounds are studied and modeled with simulated samples. Then the statistical model (signal+background) is built, and the probability that resonances exist with certain mass is evaluated. After searching through the full $m_{J\gamma}$ mass range in data, no significant deviation is observed from SM predictions. Therefore, cross section limits for each of the signal channels are set as shown in Figures 6.6 to 6.9.

In this search, the sensitivity is dominated by statistical uncertainty in the high mass region, but also affected by systematics, especially from the background modeling uncertainty estimated using a spurious signal test. The search strategy is designed to cover a broad mass range for BSM resonances. For the high mass region, the main improvement in the future would be more data collected or higher pp collision energy. For the low mass region, the sensitivity can be improved in multiple ways even without larger datasets. Firstly, the spurious signal test uncertainties are mainly due to the statistical limitation of SM simulation samples. Although producing more simulation is limited by ATLAS computational resource, there are clever techniques to reduce the fluctuations. Some ATLAS analyses use a Gaussian process to smear the simulated events, making the distribution smoother. Alternatively, we can replace

our SM background MC with a truth level simulation and a proper transfer function modeling the full-simulated background kinematics. Secondly, the W/Z boson tagging can be improved, which will greatly increase the impact of the D2 category. The D2 category is already the dominant category in all channels with highest purity and signal efficiency. However, current W/Z boson tagging is optimized with longitudinal polarized W/Z bosons, while in this search, only the spin-1 $W\gamma$ channel has longitudinally polarized vector bosons. As shown in Figure 5.7, all the $Z\gamma$ channels have a D2 category efficiency less than the $W\gamma$ channel. If a new tagging is optimized for transversely polarized Z bosons, there will be significant improvement for all $Z\gamma$ channels.

Although my thesis research found no new resonances, it expanded the mass range searched up to almost 100 times the mass of the SM's W and Z bosons. The sensitivity of probing such processes has been improved by a factor of 2 in the low mass region compared to the best result before.[47] This improvement increases up to 4 at the high mass region. However, there are far more topics to be explored in this field. Some of them are closely related to this search, which means we can either include them in further studies, or remove some limitations of this search. One of the potential topics to be explored is dropping the narrow width assumption for signal hypotheses. By looking for broader peaks, more detailed studies of signal kinematics and invariant mass shapes are required. In this search, we have already seen some broad structures in the data invariant mass distribution between 3 to 4 TeV. It is certainly interesting to quantify such broad bumps and see if we can find something. Another shortcoming of this search is its asymmetry in the production and decay modes of the target heavy resonances. All our signal models produce heavy resonances with interactions between hadrons, either quark-quark collisions or gluon-gluon fusions. However, this search looks only into W/Z boson and photon

final states. In this case, it is possible that target heavy resonance interact only with vector bosons, and therefore have completely different kinematic features than our current signal production models. On the other hand, if we look into heavy resonance produced by boson interactions, there will be signatures such as forward jets in the production process. Furthermore, without the hadronic production modes, the target BSM boson is not required to interact with the QCD sector. Therefore, such vector boson fusion channels would allow searches for BSM bosons that interact only with the SM's electroweak sector.

Appendix A

Appendices

A.1 Control Region

A control region (CR) is defined to test the background fitting strategies. This control region is defined with events passing all our baseline selection except for the photon η cut. The signal region (SR) requires the leading photon in event has $|\eta| < 1.37$, and the control region requires photon $|\eta| > 1.52$. The band of photon with $1.37 < |\eta| < 1.52$ is defined on the crack region between barrel and endcap EM calorimeters, and also guarantees the orthogonality between SR and CR. The mass distribution of CR is shown in Figure A.1.

The signal leakage, defined as the total signal efficiency in CR, is shown in Figure A.2. In SR, an additional photon and jet p_T cut is applied after baseline selection which strongly rejects background events. This cut is not applied on CR, so the signal contribute much less in CR than in SR.

Since the CR does not have much larger statistics than data categories, we decide not to split it into categories, but treat it as a single set of events and used to test our background fit functions. In order to test different categories, this CR is rescaled to different categories. This is done by studying the ratio between signal categories and CR with respect to reconstructed mass of photon and jet in MC sample, fit this ratio by pre-defined functions, and then scale the data CR with these fitted functions. The ratio plots and its functional fit are shown in Figure A.3 and A.4. Due to the lack of statistics in Btag category, CR is scaled with a constant factor instead of a function to test fit function for Btag category.

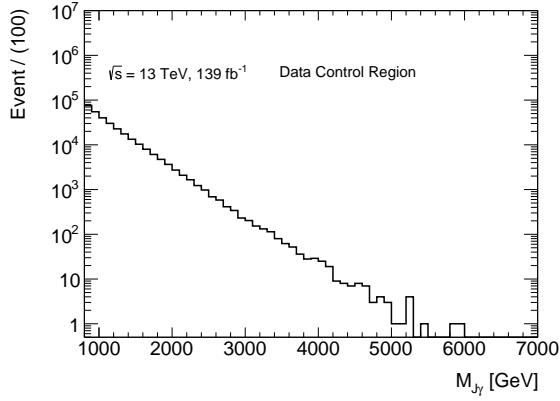


Figure A.1: This plot shows the mass spectrum of the CR dataset. Since the CR is defined as events passed baseline selection except for the primary photon falls in the end-cap, there is no difference for different signal channels.

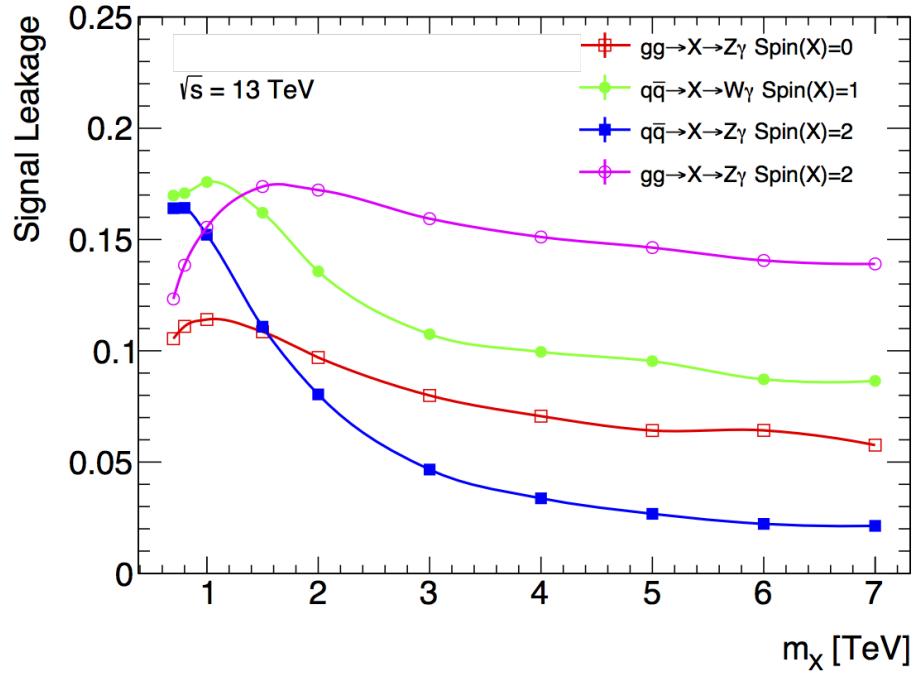


Figure A.2: The signal leakage is shown in this figure. This signal leakage is defined as the fraction of signal events goes into control region compared to the total number of signal events. The curves shown here are strongly dependent on signal η distribution as shown in Figure fig:sigEta

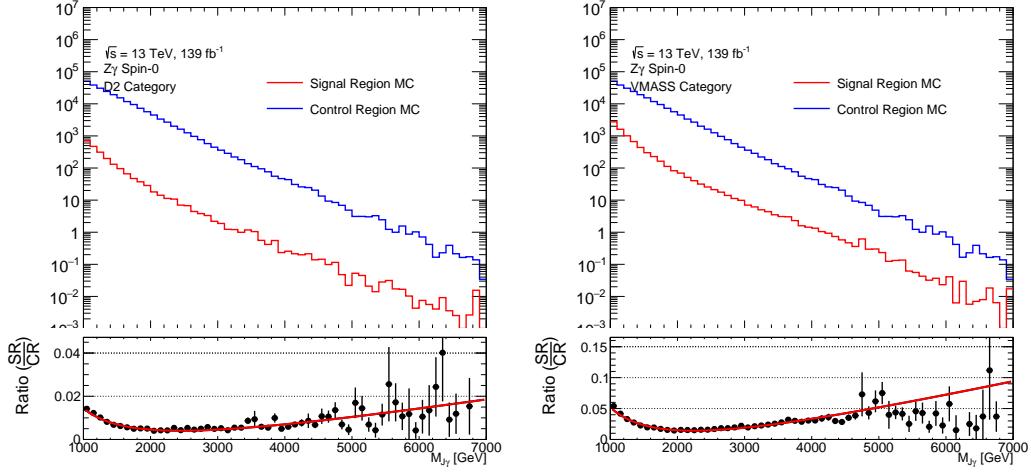


Figure A.3: These three plot shows the ratios and fits between $Z\gamma$ Spin-0 channel categories and the CR in MC. The left one is for D2 category, middle one for Vmass category. The fitted line is a function of $p_0x^{-1} + p_1 + p_2x$, which is used to model the trend of the ratio between signal region and control region. These functions are used to re-weight the CR into samples that statistically compatible with each signal categories, and the re-weighted samples are used to test the performance of background functions.

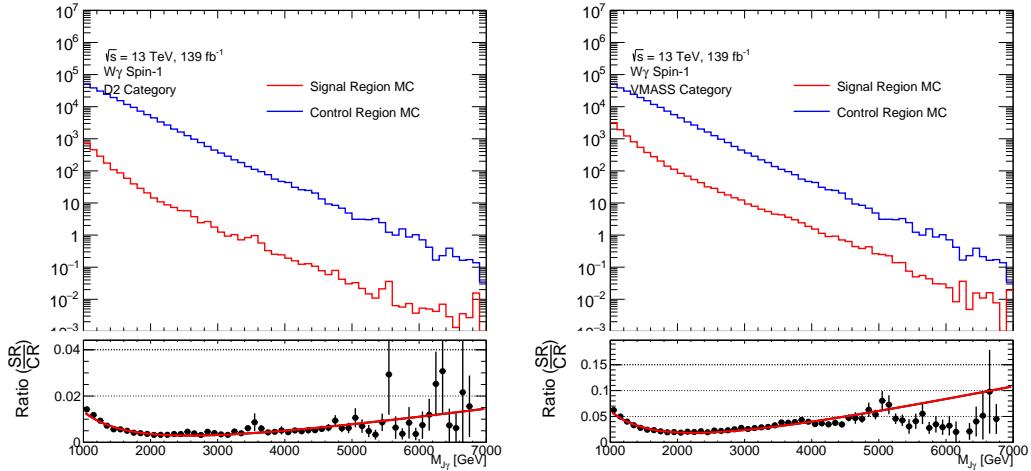


Figure A.4: These three plot shows the ratios and fits between $W\gamma$ Spin-1 channel categories and the CR in MC. The left one is for D2 category, middle one for Vmass category. The fitted line is a function of $p_0x^{-1} + p_1 + p_2x$, which is used to model the trend of the ratio between signal region and control region. These functions are used to re-weight the CR into samples that statistically compatible with each signal categories, and the re-weighted samples are used to test the performance of background functions.

Table A.1: Cutflow for signal samples at Mass = 1 TeV. The pass trigger is for passing the photon trigger we are using (HLT_g140_loose). The photon $p_T > 200$ GeV cut is raising the photon p_T threshold to 200 GeV from trigger threshold at 140 GeV. Photon η is cutting at photon $\eta < 1.37$. Tight Photon and Isolation are efficiencies of photon tags. Jet $p_T > 200$ GeV cut is on jet transverse momentum, and Jet $\eta < 2.0$ cuts on the jet angular distribution. The $\Delta R_{J\gamma} > 1.0$ is for selecting events with photon and jet isolate well with each other. The numbers are the total weight of events pass each selection.

$M(V\gamma)$	$Z\gamma$ spin-0	$W\gamma$ spin-1	$Z\gamma$ spin-2 (qq)	$Z\gamma$ spin-2 (gg)
Pass Trigger	99989.8	99984.4	99991.5	99978
Photon $p_T > 200$ GeV	98194.7	97579.5	99433.6	95439.8
Photon $\eta < 1.37$	85385.7	77916.9	82198.4	78377.5
Tight Photon	81966	75339.3	79603.8	74850.1
Photon Isolation	77067	70756	74800.2	70338.6
Jet $p_T > 200$ GeV	77067	70756	74800.2	70338.6
Jet $\eta < 2.0$	77067	70756	74800.2	70338.6
$\Delta R_{J\gamma} > 1.0$	71604.7	61693	68798.6	62120.2

A.2 Signal Cutflow

The baseline selection has multiple cuts to form a reasonable signal region. Some of the cuts are relatively tight, rejecting lots of signal events, while the others are loose, allowing most of events through it. The Table A.1, A.2 and A.3 shows the number of events passing different cuts for the signal samples corresponding to the resonance mass $m_X = 1, 2$, and 4 TeV respectively. They can provide an impression about how each of the cuts in baseline selection affect the efficiency. It is not only strongly dependent on signal channels, but also on the resonance mass.

A.3 Data/MC Comparison

The MC modeling of selected events are checked comparing few kinematic distributions. The photon p_T , jet p_T , jet η , jet mass distributions are shown in Figs. A.5, A.6, A.7, A.8.

Table A.2: Cutflow for signal samples at Mass = 2 TeV. The pass trigger is for passing the photon trigger we are using (HLT_g140_loose). The photon $p_T > 200$ GeV cut is raising the photon p_T threshold to 200 GeV from trigger threshold at 140 GeV. Photon η is cutting at photon $\eta < 1.37$. Tight Photon and Isolation are efficiencies of photon tags. Jet $p_T > 200$ GeV cut is on jet transverse momentum, and Jet $\eta < 2.0$ cuts on the jet angular distribution. The $\Delta R_{J\gamma} > 1.0$ is for selecting events with photon and jet isolate well with each other. The numbers are the total weight of events pass each selection.

$M(V\gamma)$	$Z\gamma$ spin-0	$W\gamma$ spin-1	$Z\gamma$ spin-2 (qq)	$Z\gamma$ spin-2 (gg)
Pass Trigger	99974.1	99979.3	99976.8	99972.8
Photon $p_T > 200$ GeV	99767.5	99676.8	99942.7	99431.9
Photon $\eta < 1.37$	88894.7	84614.1	91285.2	79871.3
Tight Photon	85410.9	82167.8	88777.2	76504.3
Photon Isolation	78910.7	76137.4	81963.3	71012.2
Jet $p_T > 200$ GeV	78910.7	76137.4	81963.3	71012.2
Jet $\eta < 2.0$	78910.7	76137.4	81963.3	71012.2
$\Delta R_{J\gamma} > 1.0$	76617.4	71475.4	79974.6	67367.1

Table A.3: Cutflow for signal samples at Mass = 4 TeV. The pass trigger is for passing the photon trigger we are using (HLT_g140_loose). The photon $p_T > 200$ GeV cut is raising the photon p_T threshold to 200 GeV from trigger threshold at 140 GeV. Photon η is cutting at photon $\eta < 1.37$. Tight Photon and Isolation are efficiencies of photon tags. Jet $p_T > 200$ GeV cut is on jet transverse momentum, and Jet $\eta < 2.0$ cuts on the jet angular distribution. The $\Delta R_{J\gamma} > 1.0$ is for selecting events with photon and jet isolate well with each other. The numbers are the total weight of events pass each selection.

$M(V\gamma)$	$Z\gamma$ spin-0	$W\gamma$ spin-1	$Z\gamma$ spin-2 (qq)	$Z\gamma$ spin-2 (gg)
Pass Trigger	99977.7	99971.4	99991	99973.6
Photon $p_T > 200$ GeV	99966.2	99949.3	99979.4	99948.8
Photon $\eta < 1.37$	91724.6	88146.3	96298.2	82184.3
Tight Photon	88092.6	85709.7	93400.3	79064.8
Photon Isolation	79048	77223.5	83833.1	71472.5
Jet $p_T > 200$ GeV	79048	77223.5	83833.1	71472.5
Jet $\eta < 2.0$	79048	77223.5	83833.1	71472.5
$\Delta R_{J\gamma} > 1.0$	78406.6	76145.7	83343.1	70624.6

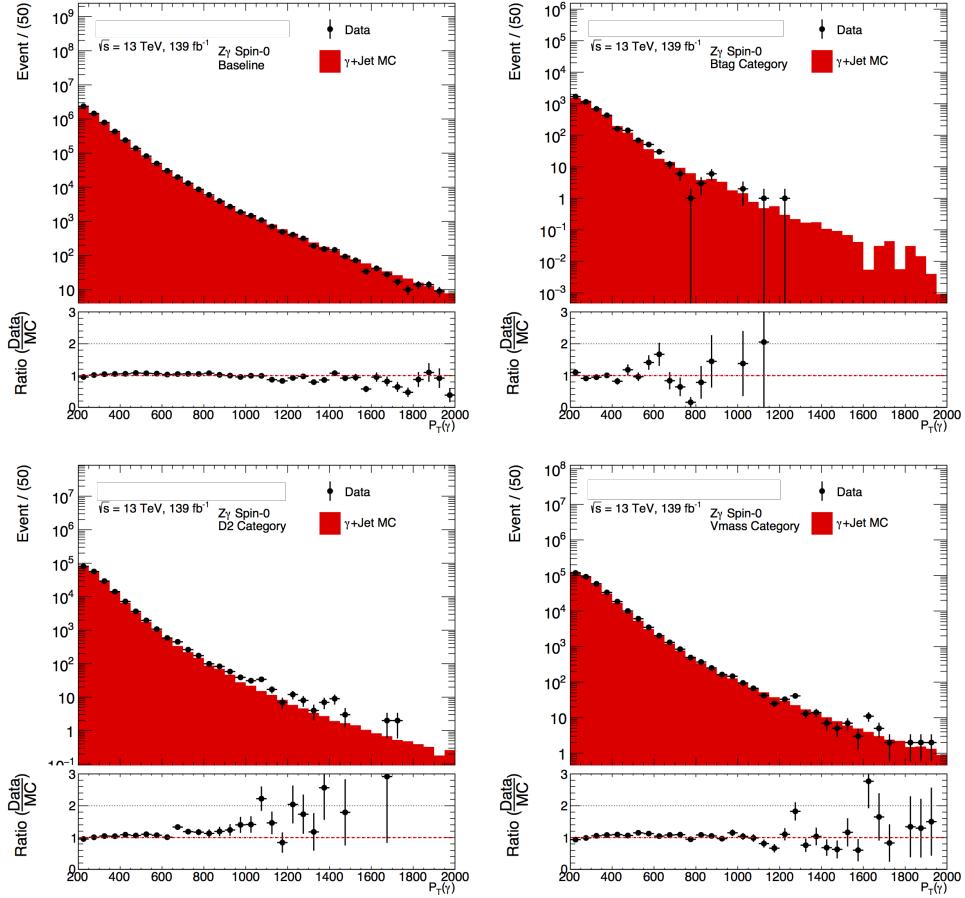


Figure A.5: Photon Transverse Momentum distributions for events passing the baseline cut, and for the events in each of the four categories in $Z\gamma$ channel.

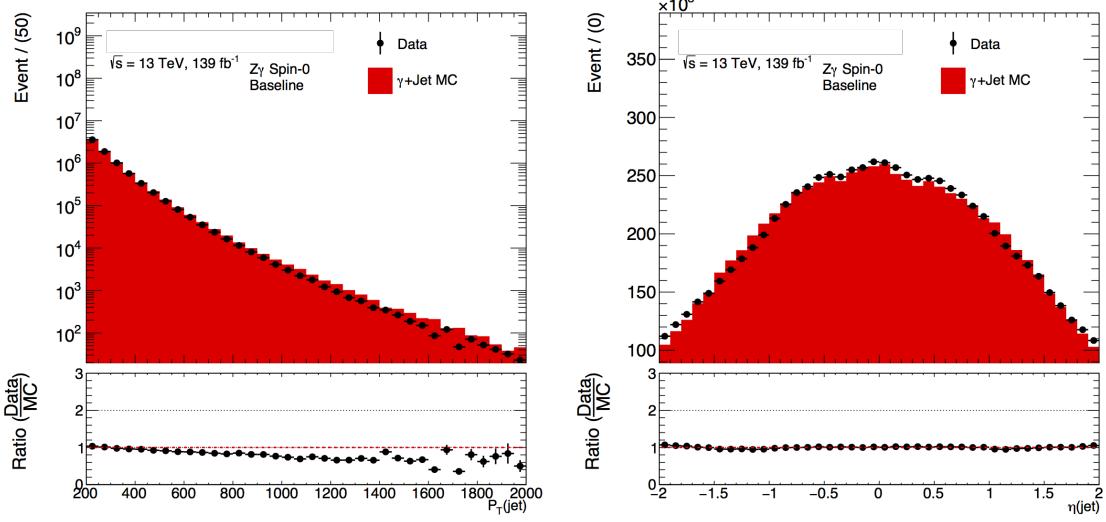


Figure A.6: Jet p_T and jet η distributions for events passing the baseline cut.

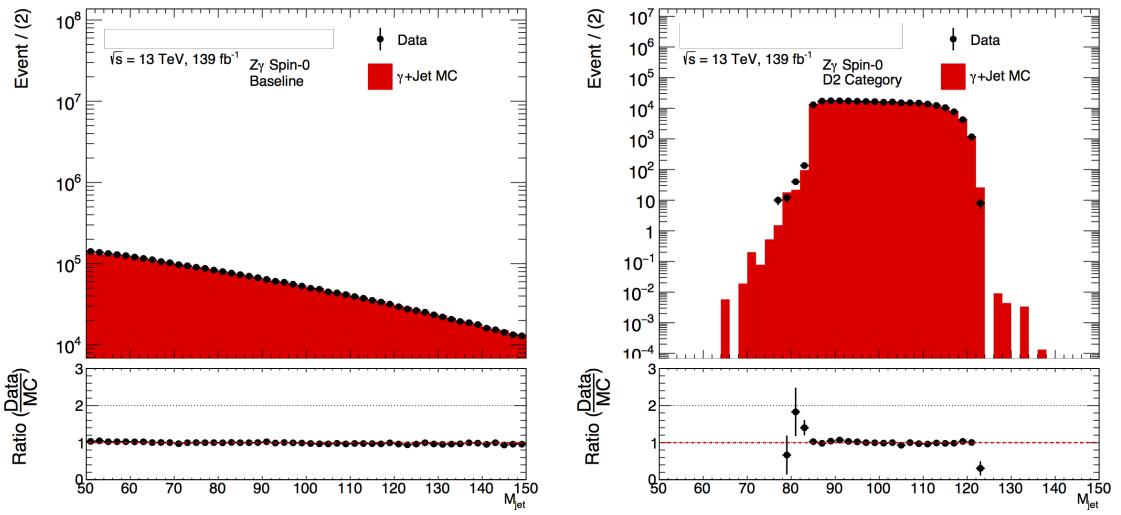


Figure A.7: The top plot shows the jet mass distribution for events passing the baseline cut while the bottom right and for the events passing also the D2 selection criteria of the W (Z) tagger in each of the four categories in $Z\gamma$ channel.

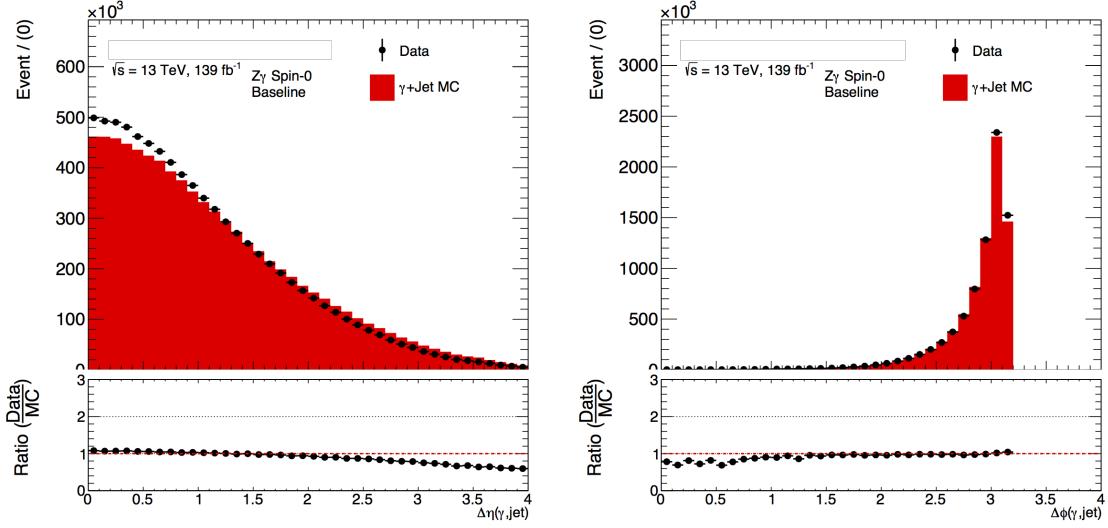


Figure A.8: $\Delta\eta_{\gamma J}$ and $\Delta\phi_{\gamma J}$ for events pass baseline cut.

A.4 Optimization of p_T cuts

With the baseline selection and categorization defined, in order to find the subset with highest sensitivity, we look into the difference in kinematic features between signal and background simulated samples.

After the study of kinematic features of the signal and background MC samples, a difference of photon and jet p_T distribution between signal and background is noticed. The p_T of the signal is in general higher than the p_T of the background. For the signal mass points in the background abundant region between 1 to 4 TeV, by applying an extra photon and jet p_T cut the final limit can be improved by up to 50%.

Unlike the conventional constant $\frac{p_T}{m}$ cut, this analysis sets limits up to the extremely high mass regime with very low background. A constant $\frac{p_T}{m}$ cut becomes too strong, which could reduce the sensitivity. Therefore, a further study on the optimal cut value of p_T is conducted, and the trend of this cut versus resonance mass is determined. The dependence is very linear between the optimal p_T cut and resonance mass in the background abundant region, but when the background is reduced to the order

of 1, it becomes more or less constant, regardless of the mass increase. When the background becomes low enough, the optimal p_T cut tends to be very low again, due to the fact that further reducing background changes almost nothing, but sacrifices signal efficiency. The overall trend for the optimal p_T cut value is, firstly increasing linearly with mass, then reaching a plateau, and finally dropping. The third phase is very flexible since changing the cut value by a little will not affect the signal, which has almost all events at very high p_T region, and reducing the background without too much impact, since the total expected background event is already approaching zero.

A.4.1 Optimization Methodology

The final mass dependent p_T cut is determined by finding the optimal cut value and with uncertainty at each of the mass point, and then fitting these cut values by a smooth function. The functional form is chosen to be parabola, both for the optimal cut value distribution trend, and the fact that final cut values for high mass regime are very flexible, so that parabola has enough freedom to model it.

The local optimal cut value is determined by a expected limit scan of the p_T cut, with a 5 GeV interval. The expected limit is calculated by counting the expected background number of events within the resonance mass window of $\pm 2\sigma$ of the signal peak, and then estimating the value with 95% probability of having less number of event observed, and finally calculating the corresponding value of cross section times branching ratio, which can produce this amount of signal events. This scan terminates at 60% signal efficiency point to avoid extreme cases when both signal and background efficiency are extremely low, and statistical fluctuations of the Monte-Carlo samples dominate the limit behavior. The asymmetric uncertainty is assigned by the minimum window covering all possible cut values corresponding to the expected limit within

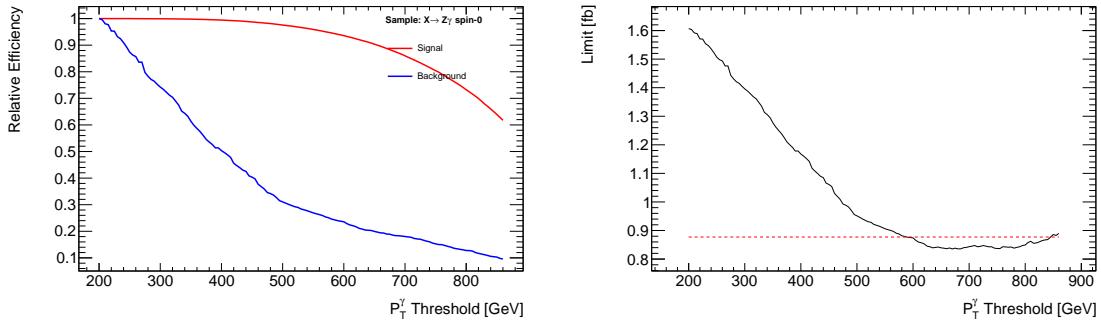


Figure A.9: The p_T cut scan for mass point 2 TeV in the $X(\text{spin} = 0) \rightarrow Z\gamma$ channel. The left plot shows relative efficiency, calculated by dividing the total weighted number of events before and after the corresponding p_T cut is applied. The right plot shows the statistical only expected limit for different cut values. A red dashed horizontal line on the right plot shows the limit value 5% higher than optimal limit, and therefore the crossing points gives an asymmetric uncertainty for this mass point.

5% of the optimal value, as shown in Fig. A.9.

Compared to the conventional constant p_T/m_X cuts, this variational cut provides similar linear kinematic constraint at low mass regime, while automatically become a looser criteria at high mass region where signal and background number of events simultaneously dropped. This cut not only enables our usage of maximum statistics, but also provides us extra smoothing of the MC background shape, prohibiting some high weight event from low photon p_T slice migrating into high mass region. These special events usually appear in the photon p_T 280 GeV to 500 GeV slice, and migrate into the mass region between 1 to 2 TeV, where the mass dependent p_T cut is strongest. Thus, we believe it is worthwhile to sacrifice a small fraction of signal region, for these benefits.

A.4.2 Extrapolation vs. constant cut at high mass region

The p_T cut at high mass can be either extrapolated as a parabola or extended with a constant cut. In order to investigate the difference with these two methods, the

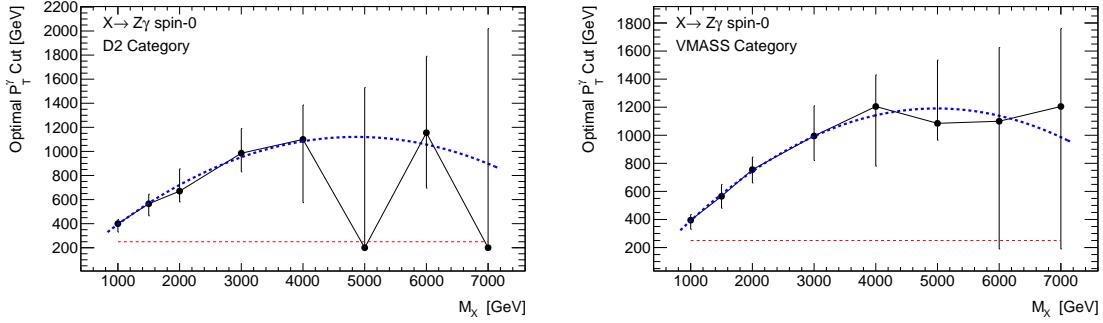


Figure A.10: Photon p_T cut as function of m_X , there is also an underlying jet p_T cut at 80% of the photon cut values. The dot denotes the optimal cut value at each mass point, and the uncertainty here is set by the interval of cut value with the final sensitivity not worse than 95% of the optimal cut value. The blue dashed lines are the fitted parabola functions.

study on different cut functions is conducted. According to the study, the peak of this variable p_T cut is usually at about 5 TeV. Therefore, using constant cut value from 4 TeV means the cut is a little looser, and from 5 TeV means tighter at very high mass. However, the expected limits of D2 and Vmass categories will only be changed by less than 1% because different background rejection at high mass region is not as influential as in low mass region. Thus, the difference in expected limits of these two methods for p_T cut is almost negligible in D2 and Vmass categories. Since the extrapolation method will rapidly loose the cut at high mass, and can potentially produce fake excess, we decide to use a constant p_T cut after the cut function reaching its maximum. This guarantees the cut to be differentially continuous to produce the background mass distribution as smooth as possible.

A.5 Tests of Background Fit

A.5.1 More spurious signal test results

Figures A.11, A.12 and A.13 show the spurious signal results for spin-1 $W\gamma$, spin-2 gg $Z\gamma$ and spin-2 qq $Z\gamma$ channels.

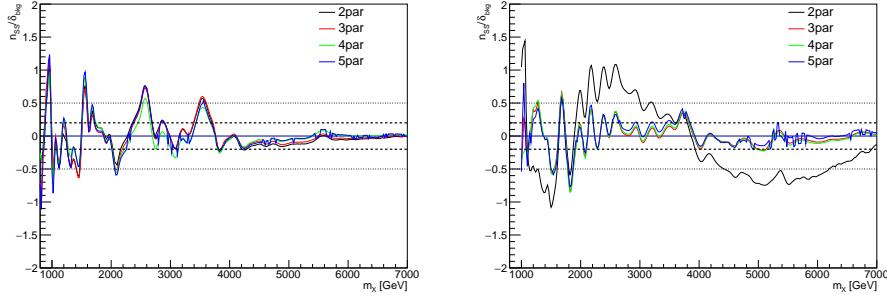


Figure A.11: Spurious signal test on MC simulation samples with different number of parameters in SR (a) $D2$, (b) SR $Wmass$ of spin-1 $W\gamma$ channel.

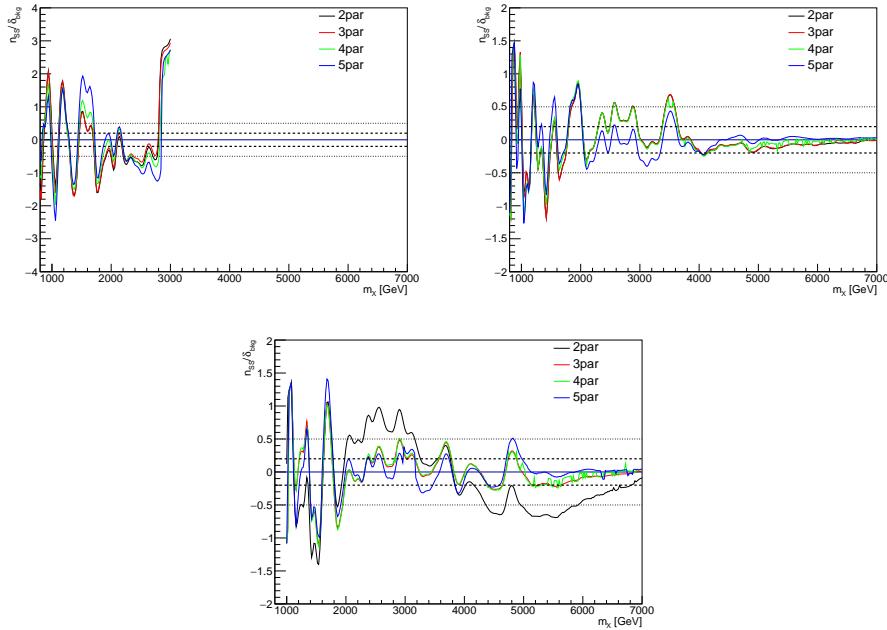


Figure A.12: Spurious signal test on MC simulation samples with different number of parameters in SR (a) $Btag$, (b) $D2$, (c) $Zmass$ of spin-2 $gg Z\gamma$ channel.

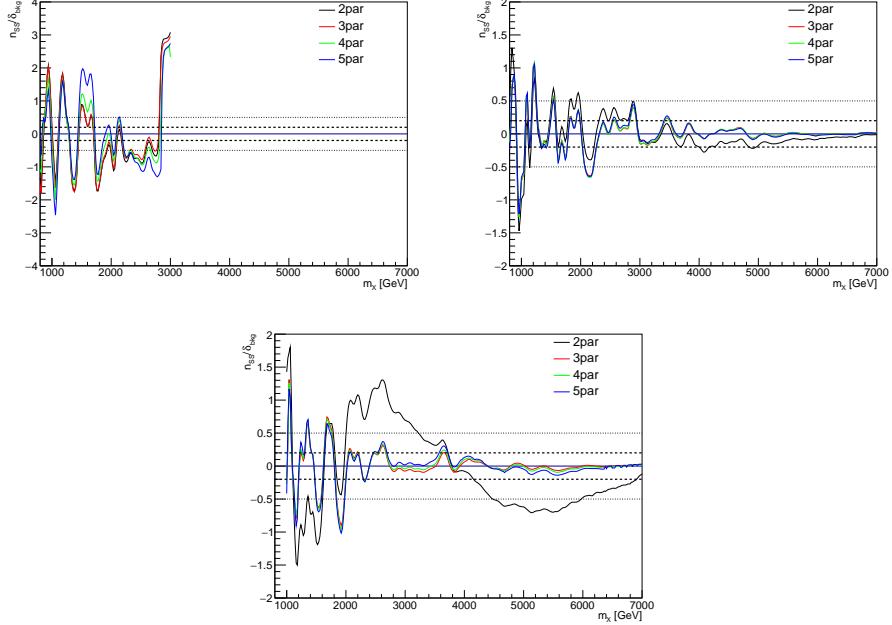


Figure A.13: Spurious signal test on MC simulation samples with different number of parameters in SR (a) $Btag$, (b) $D2$, (c) $Zmass$ of spin-2 qq $Z\gamma$ channel.

Figures A.14, A.15 and A.16 show the number of spurious signal events along the mass spectrum for those three signal channels.

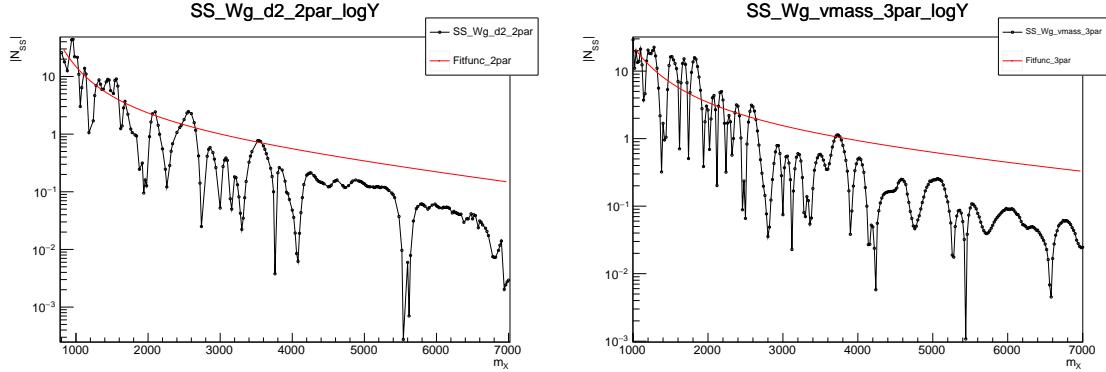


Figure A.14: The raw number of spurious signal events fitted with MC samples in SR (a) $D2$, (b) $Wmass$ of spin-1 $W\gamma$ channel for chosen number of background function parameters. The $|N_{SS}|$ are then parametrized with exponential function to minimize the effect due to MC statistics of spurious signal. In order to have a conservative estimation of spurious signal, double size of parametrized results are treated as spurious signal systematic uncertainty.

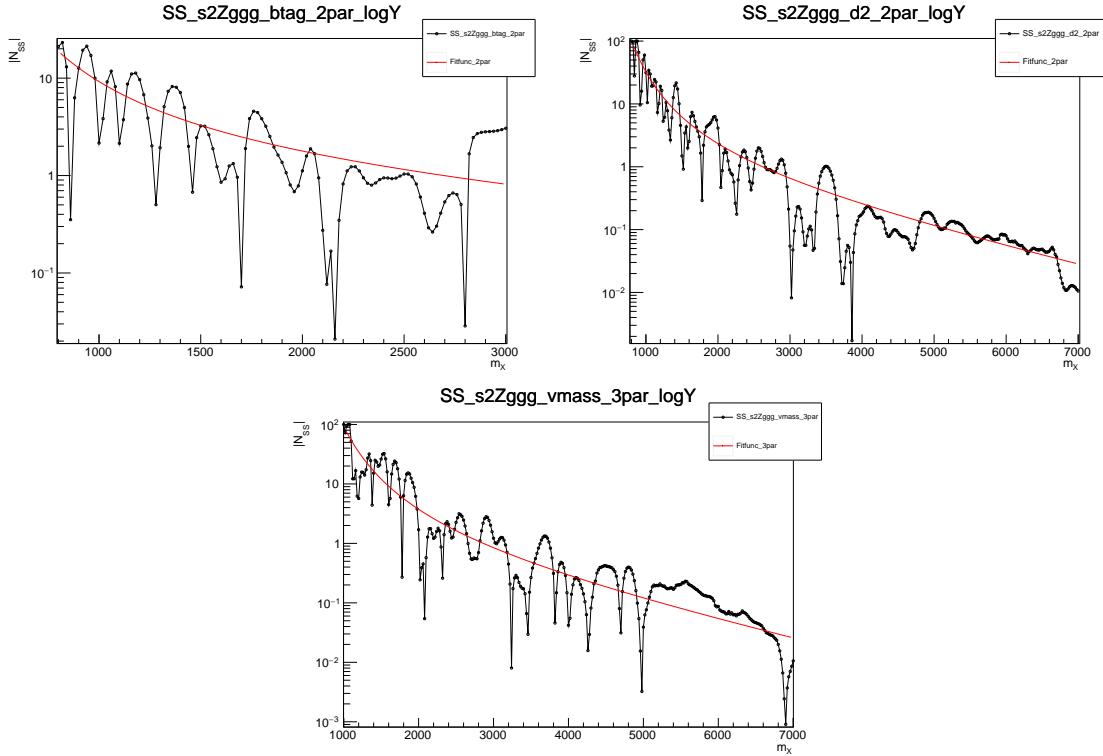


Figure A.15: The raw number of spurious signal events fitted with MC samples in SR (a) $btag$ (b) $D2$, (c) $Zmass$ of spin-2 gg $Z\gamma$ channel for chosen number of background function parameters. The $|N_{SS}|$ are then parametrized with exponential function to minimize the effect due to MC statistics of spurious signal. In order to have a conservative estimation of spurious signal, double size of parametrized results are treated as spurious signal systematic uncertainty.

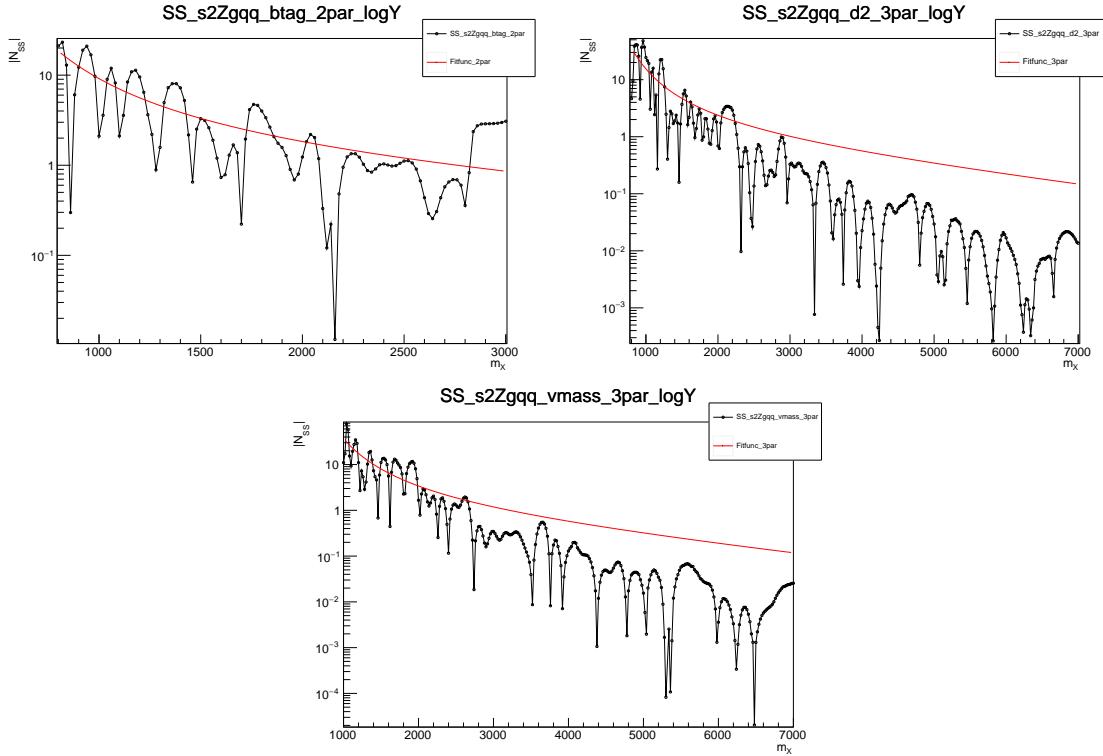


Figure A.16: The raw number of spurious signal events fitted with MC samples in SR (a) $btag$ (b) $D2$, (c) $Zmass$ of spin-2 qq $Z\gamma$ channel for chosen number of background function parameters. The $|N_{SS}|$ are then parametrized with exponential function to minimize the effect due to MC statistics of spurious signal. In order to have a conservative estimation of spurious signal, double size of parametrized results are treated as spurious signal systematic uncertainty.

A.5.2 Spurious signal with data sideband

The spurious signal test is cross checked with forward photon CR as defined in Section 5.3.1. When performing the spurious signal test, instead of categorizing the CR events into different regions according to SR definition, the CR data are reweighted to match $m_{J\gamma}$ distributions in different categories to keep enough statistics in spurious signal test. The reweighting functions are shown in Figures A.3 and A.4.

Figure A.17 shows the comparison of spurious test results (N_{ss}) with CR data and MC sample for spin-0 $Z\gamma$ d2 and vmass categories. A good agreement is observed between N_{ss} from data and MC simulated samples. The similar results for three other channels are shown in Figure A.20, A.18 and A.19.

To understand whether current spurious signal uncertainty is enough, the comparison between spurious signal parameterization ($2\times$ fitted shape) and data CR fit residual are shown in Figure A.21, A.22, A.23 and A.24 respectively.

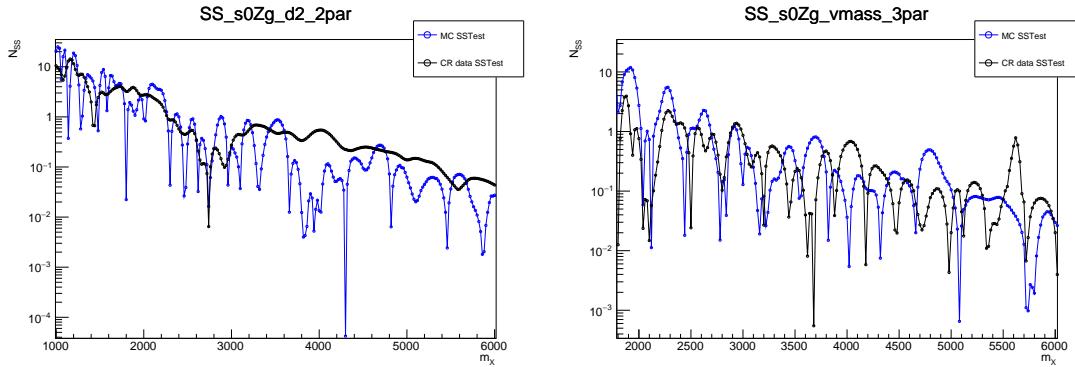


Figure A.17: Comparison of N_{ss} results with data control region (black line) and MC samples (blue line) for spin-0 $Z\gamma$ d2, vmass categories.

A.5.3 Spurious signal test with fractional MC samples

The MC samples are divided two statistical independent parts according to even and odd event counting, thus each part contains 50% of total MC events. The

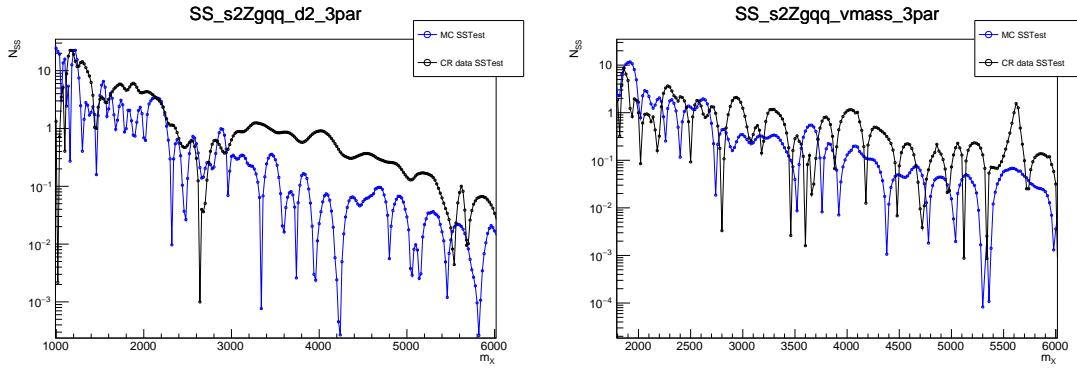


Figure A.18: Comparison of N_{ss} results with data control region (black line) and MC samples (blue line) for spin-2 qq $Z\gamma$ d2, vmass categories.

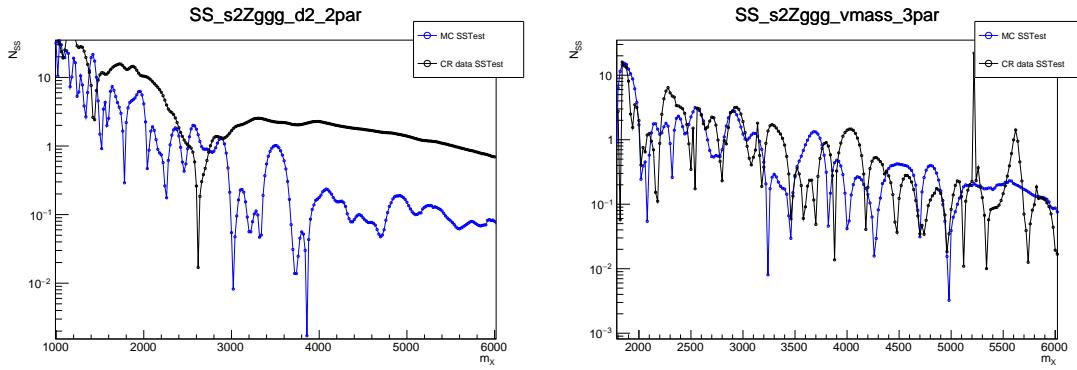


Figure A.19: Comparison of N_{ss} results with data control region (black line) and MC samples (blue line) for spin-2 gg $Z\gamma$ d2, vmass categories.

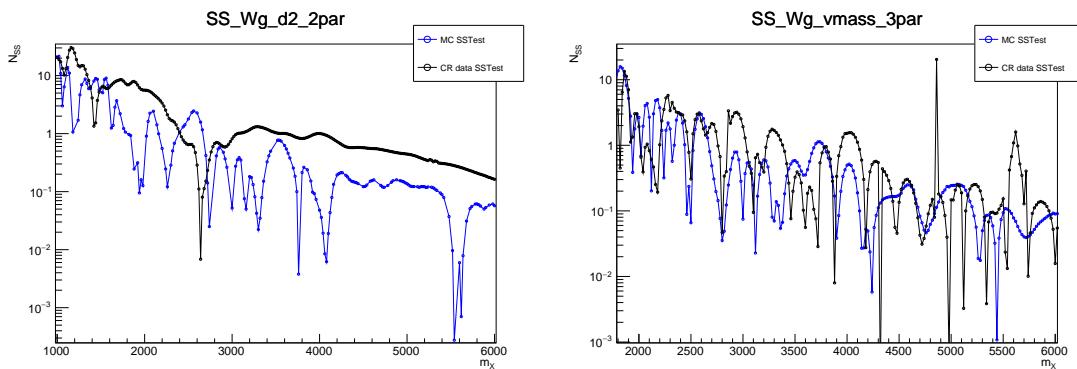


Figure A.20: Comparison of N_{ss} results with data control region (black line) and MC samples (blue line) for spin-1 $W\gamma$ d2, vmass categories.

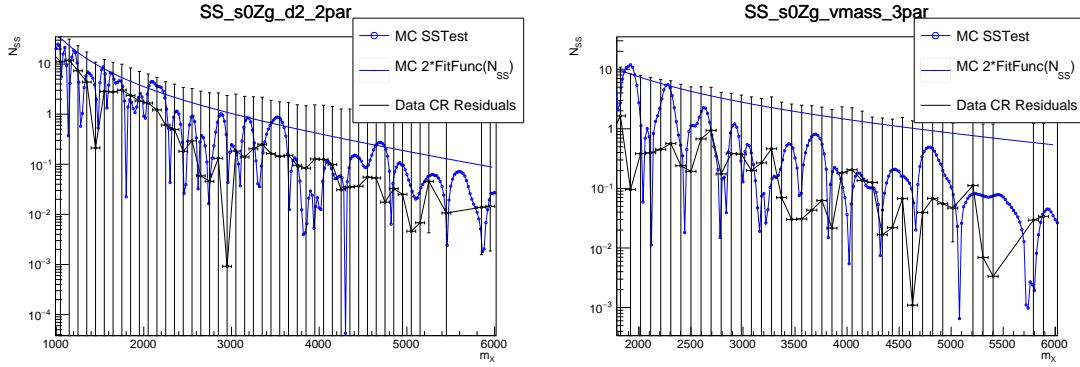


Figure A.21: Comparison of spurious signal parameterization results ($2 \times$ fitted shape) to residual distributions between data control region data and their fitted shapes for spin-0 $Z\gamma$ channel. Basically, the spurious signal can cover the difference between data points and function shape.

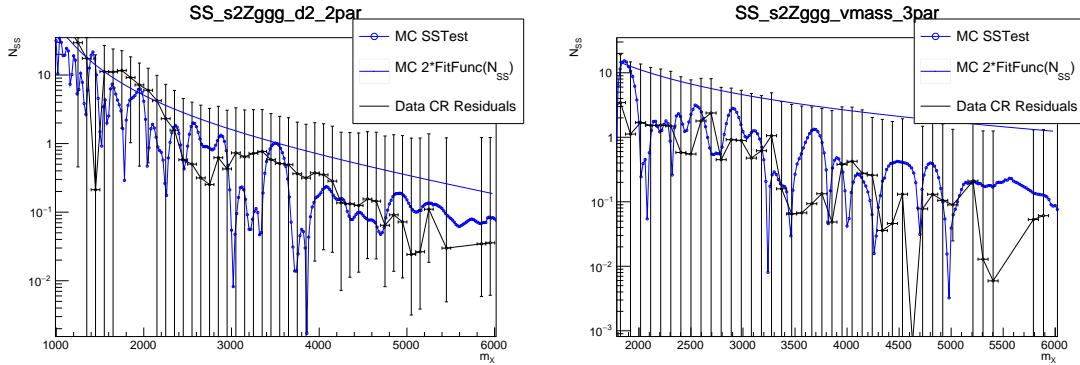


Figure A.22: Comparison of spurious signal parameterization results ($2 \times$ fitted shape) to residual distributions between data control region data and their fitted shapes for spin-2 gg $Z\gamma$ channel. Basically, the spurious signal can cover the difference between data points and function shape.

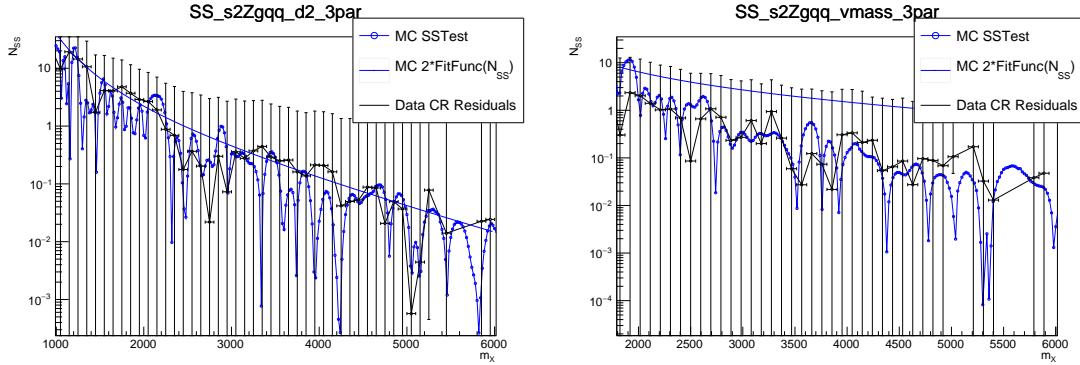


Figure A.23: Comparison of spurious signal parameterization results ($2 \times$ fitted shape) to residual distributions between data control region data and their fitted shapes for spin-2 qq $Z\gamma$ channel. Basically, the spurious signal can cover the difference between data points and function shape.

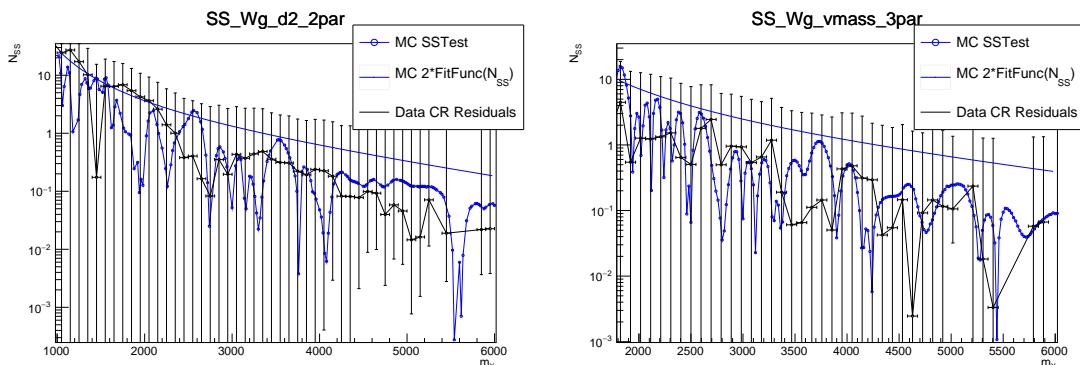


Figure A.24: Comparison of spurious signal parameterization results ($2 \times$ fitted shape) to residual distributions between data control region data and their fitted shapes for spin-0 $Z\gamma$ channel. Basically, the spurious signal can cover the difference between data points and function shape.

spurious signal test is then repeated with each subset. The results are shown in Figures A.25, A.26, A.27 and A.28 for spin-0 $Z\gamma$, spin-1 $W\gamma$, spin-2 qq $Z\gamma$ and spin-2 gg $Z\gamma$ signals, respectively. In general, no significant statistical correlations observed between spurious signal results tested with different MC samples and the spikes appear in different positions. These results indicates that the spikes and deeps in spurious signal test are due to statistical fluctuations.

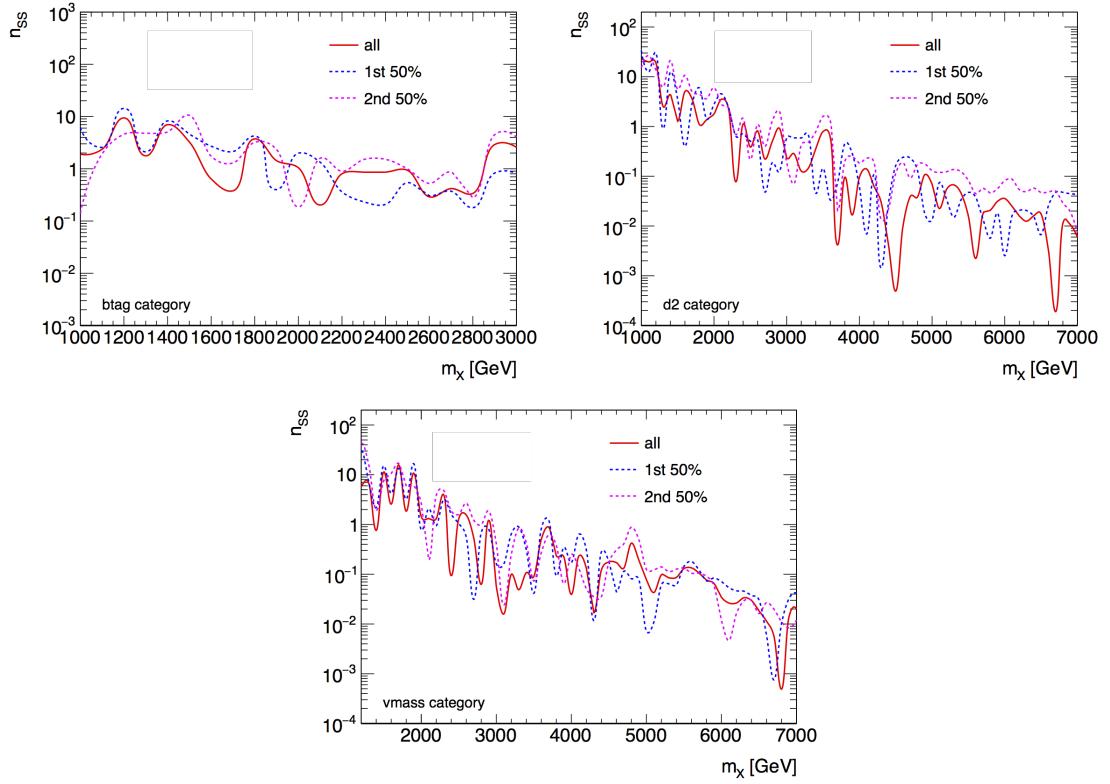


Figure A.25: Comparison of spurious signal test results with full MC statistics samples and partial statistics samples for spin-0 $Z\gamma$ (a) btag, (b) d2, (c) vmass categories. In general, the spikes appear randomly along the whole mass range and no strong statistical correlations observed.

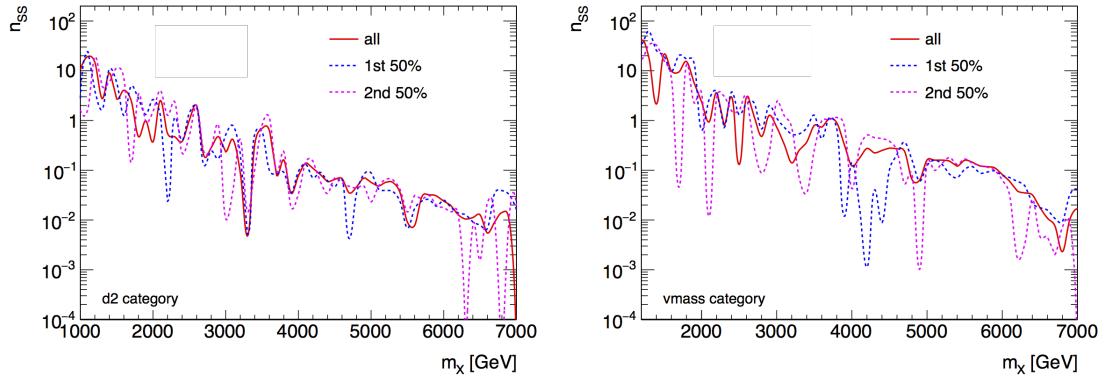


Figure A.26: Comparison of spurious signal test results with full MC statistics samples and partial statistics samples for spin-1 $W\gamma$ (a) d2, (b) vmass categories. In general, the spikes appear randomly along the whole mass range and no strong statistical correlations observed.

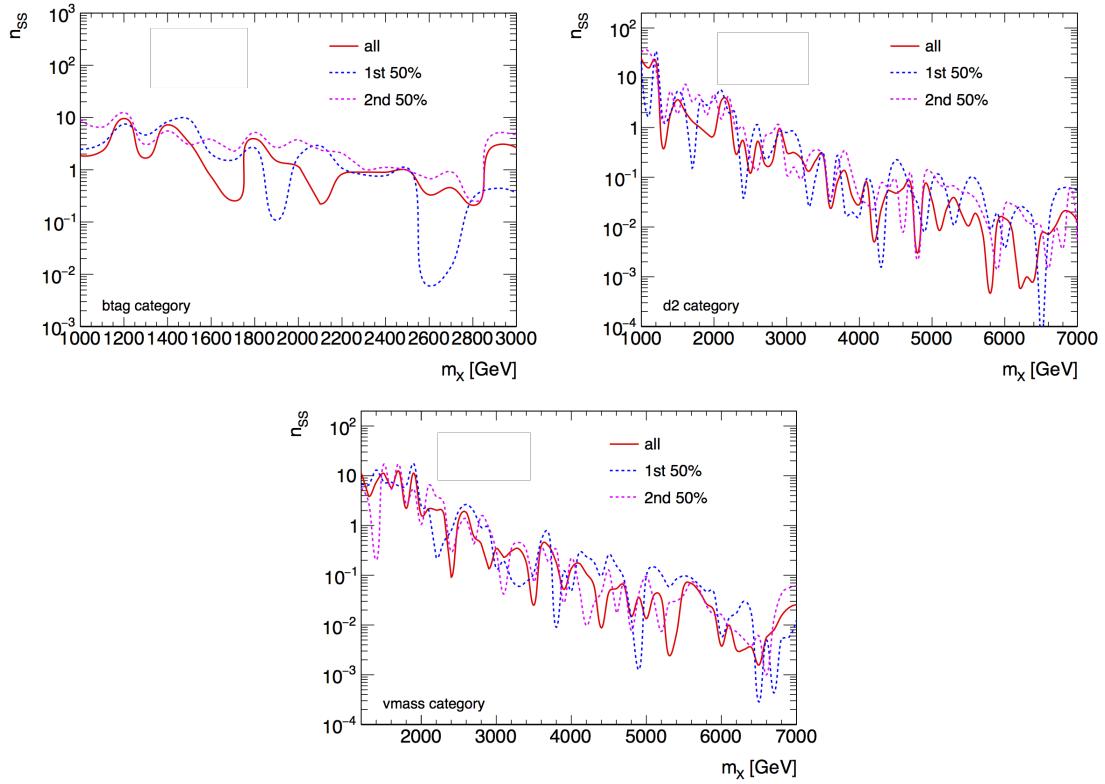


Figure A.27: Comparison of spurious signal test results with full MC statistics samples and partial statistics samples for spin-2 $qq Z\gamma$ (a) btag, (b) d2, (c) vmass categories. In general, the spikes appear randomly along the whole mass range and no strong statistical correlations observed.

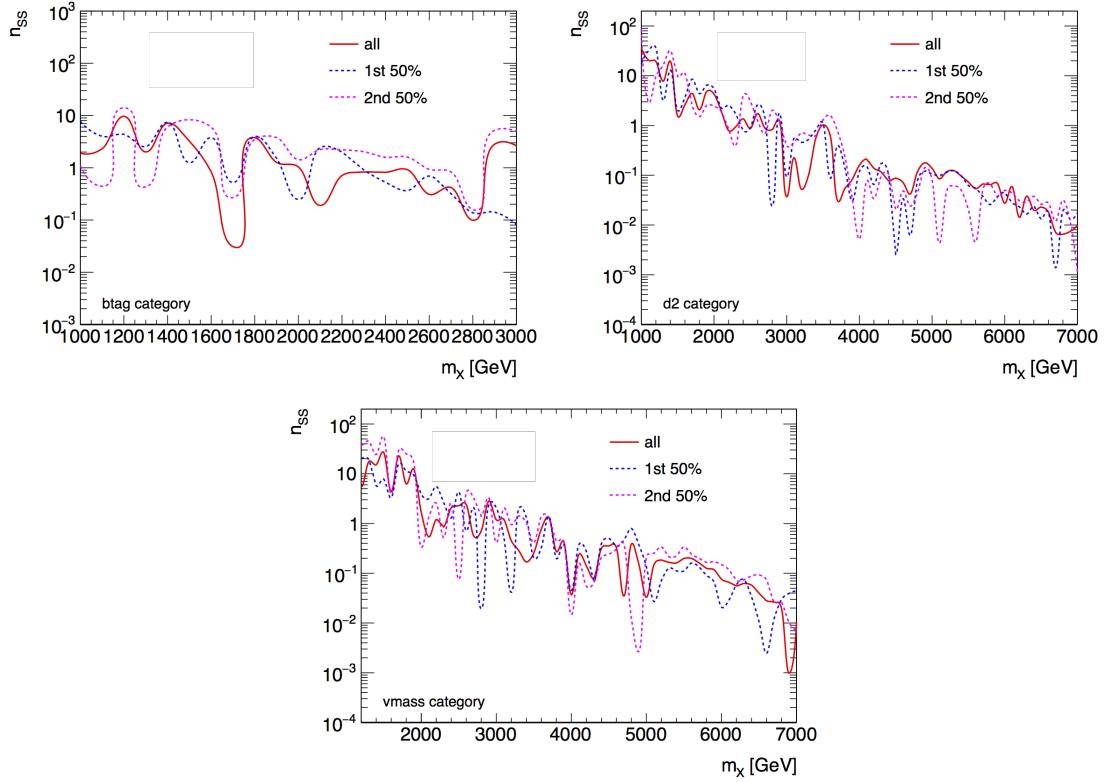


Figure A.28: Comparison of spurious signal test results with full MC statistics samples and partial statistics samples for spin-2 gg $Z\gamma$ (a) btag, (b) d2, (c) vmass categories. In general, the spikes appear randomly along the whole mass range and no strong statistical correlations observed.

A.6 Checks for the maximum significance mass points

The comparison between data and signal+background fit for the mass points with large significance are checked. Here the mass point with largest local significance listed as:

- spin-0 $Z\gamma$ $m_X = 3.64$ TeV with $Z = 2.49\sigma$
- spin-2 gg $Z\gamma$ $m_X = 3.58$ TeV with $Z = 2.24\sigma$
- spin-2 qq $Z\gamma$ $m_X = 3.56$ TeV with $Z = 1.82\sigma$
- spin-1 $W\gamma$ $m_X = 2.82$ TeV with $Z = 1.52\sigma$

A.6.1 Distributions and Fits

Figure A.29 shows the distributions of spin-0 $Z\gamma$ d2 and vmass categories with background+signal fit applied for $m_X = 3640$ GeV signal hypothesis. Figure A.30 presents the distributions of signal and background shapes including data points in spin-2 gg $Z\gamma$ channel with $m_X = 3580$ GeV signal hypothesis signal+background fit performed, Figure A.31 gives similar distributions for spin-2 qq $Z\gamma$ channel with $m_X = 3560$ GeV and Figure A.32 presents that for spin-1 $W\gamma$ with $m_X = 3120$ GeV.

A.6.2 Pull and Ranking

Pull and ranking for large significance mass points are checked. The results are shown in Figures A.33 for spin-0 $Z\gamma$ $m_X = 3640$ GeV, spin-2 gg $Z\gamma$ $m_X = 3580$ GeV, spin-2 qq $Z\gamma$ $m_X = 3560$ GeV and spin-1 $W\gamma$ $m_X = 2820$ GeV mass points. No major pull and constraint for nuisance parameters is observed.

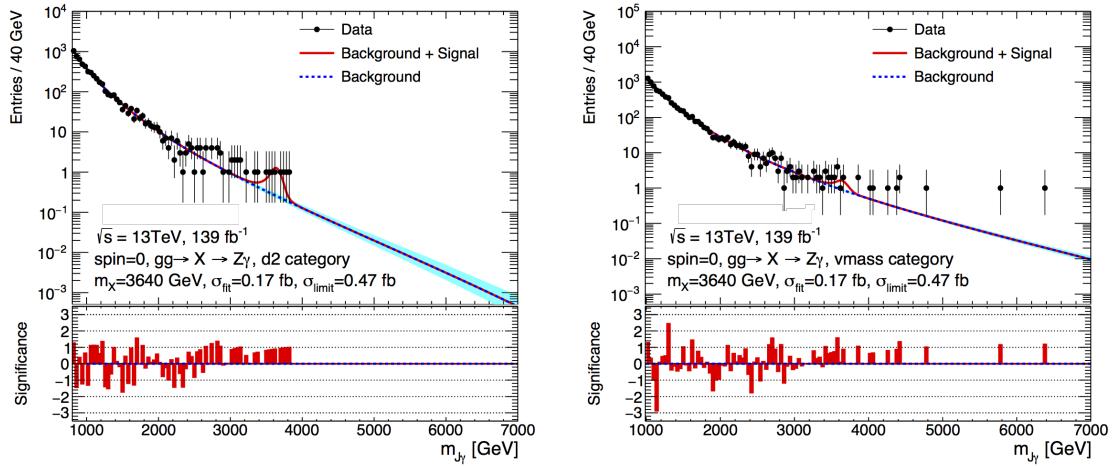


Figure A.29: Data distributions with $m_X = 3640$ GeV signal included for (a) spin-0 $Z\gamma$ d2 category and (b) spin-0 $Z\gamma$ vmass category. Dashed blue lines and light blue bands show the background pdf distributions and their $\pm 1\sigma$ error bands after performing all categories combined background+signal fit to data. Red solid lines in top pads present the signal+background pdf. The bars in low pads present the local significance for each bin.

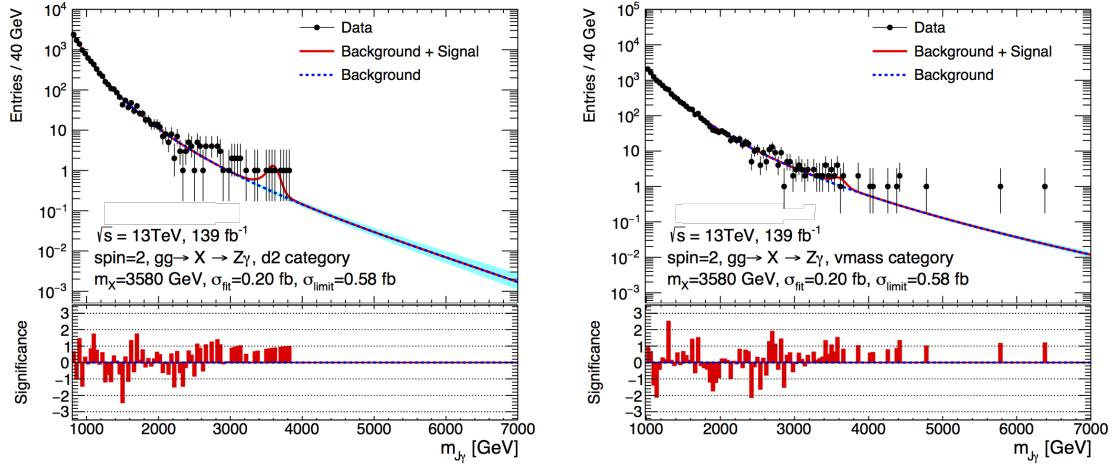


Figure A.30: Data distributions with $m_X = 3580$ GeV signal included for (a) spin-2 gg $Z\gamma$ d2 category and (b) spin-2 gg $Z\gamma$ vmass category. Dashed blue lines and light blue bands show the background pdf distributions and their $\pm 1\sigma$ error bands after performing all categories combined background+signal fit to data. Red solid lines in top pads present the signal+background pdf. The bars in low pads present the local significance for each bin.

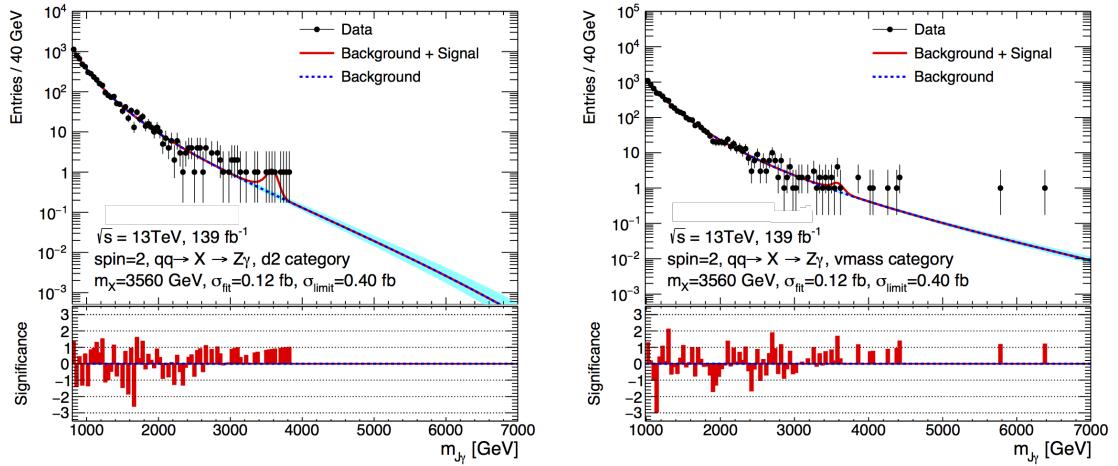


Figure A.31: Data distributions with $m_X = 3560$ GeV signal included for (a) spin-2 $qq Z\gamma$ d2 category and (b) spin-2 $qq Z\gamma$ vmass category. Dashed blue lines and light blue bands show the background pdf distributions and their $\pm 1\sigma$ error bands after performing all categories combined background+signal fit to data. Red solid lines in top pads present the signal+background pdf. The bars in low pads present the local significance for each bin.

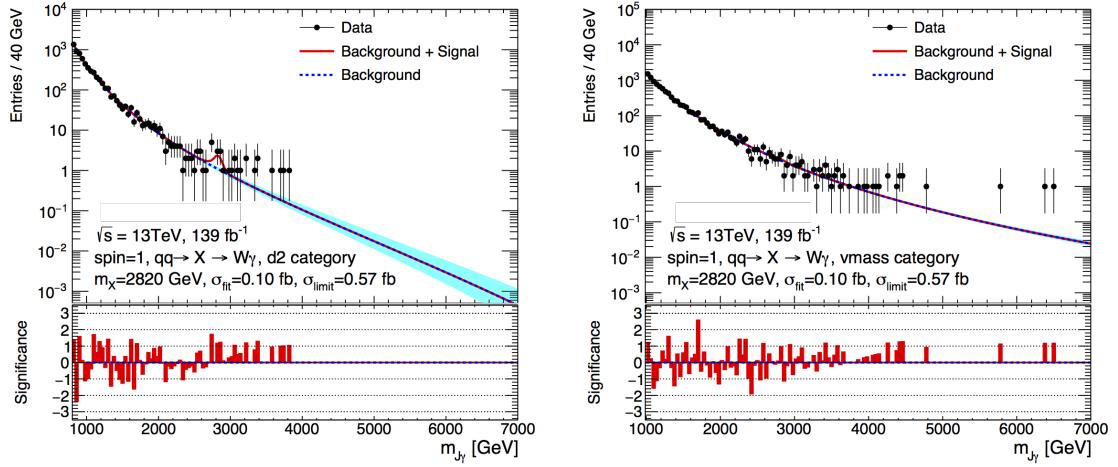


Figure A.32: Data distributions with $m_X = 2820$ GeV signal included for (a) spin-2 $qq Z\gamma$ d2 category and (b) spin-2 $qq Z\gamma$ vmass category. Dashed blue lines and light blue bands show the background pdf distributions and their $\pm 1\sigma$ error bands after performing all categories combined background+signal fit to data. Red solid lines in top pads present the signal+background pdf. The bars in low pads present the local significance for each bin.

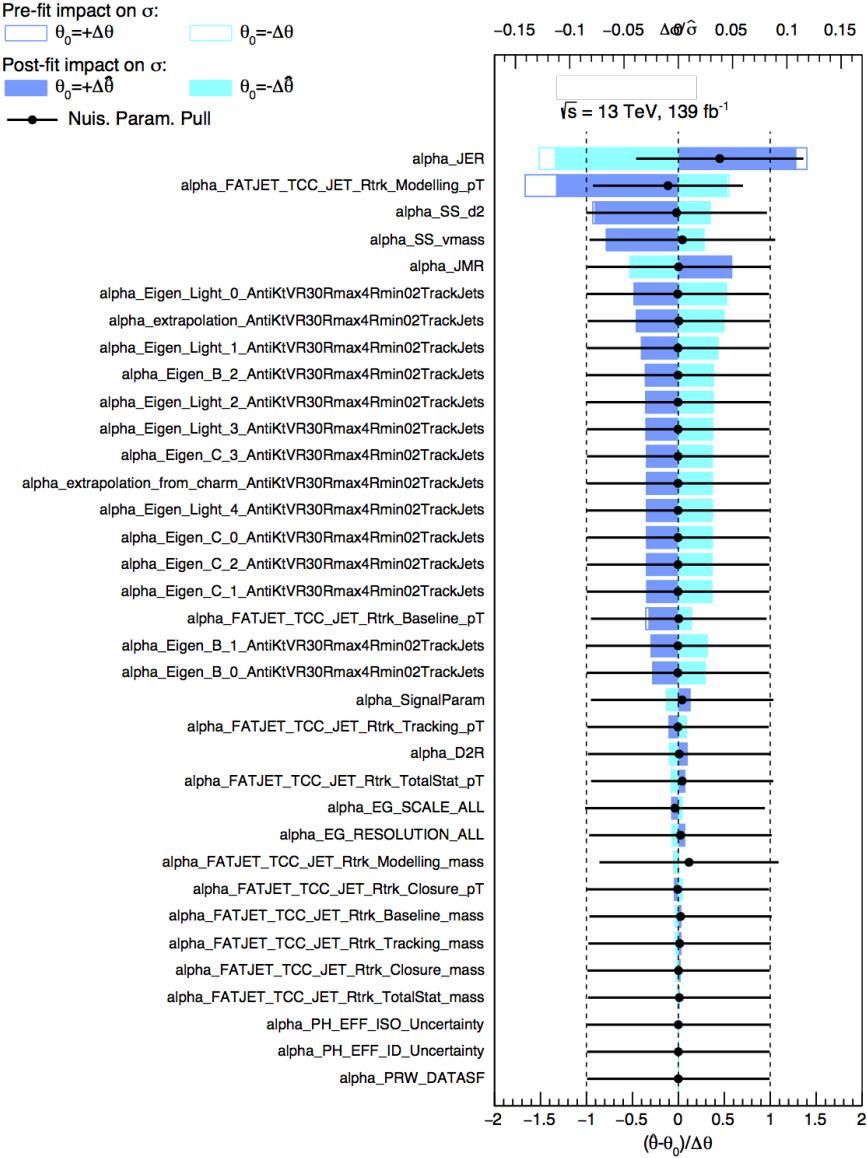


Figure A.33: Pull and ranking distributions for spin=0 $X \rightarrow Z\gamma$ $m_X = 3640 \text{ TeV}$ signal. The solid (empty) dark blue hands are the postfit (prefit) impact on σ when the NP is shifted by $+1 \sigma(\theta)$ (+1), while the solid (empty) light blue hands are the postfit (prefit) impact on σ when the NP is shifted by $-1 \sigma(\theta)$ (+1). The NP values are shown in solid point with error bars. In order to well present the systematic uncertainty impacts on fitted cross section, the impacts are shown as relative values as $\Delta\sigma/\sigma$ (top label).

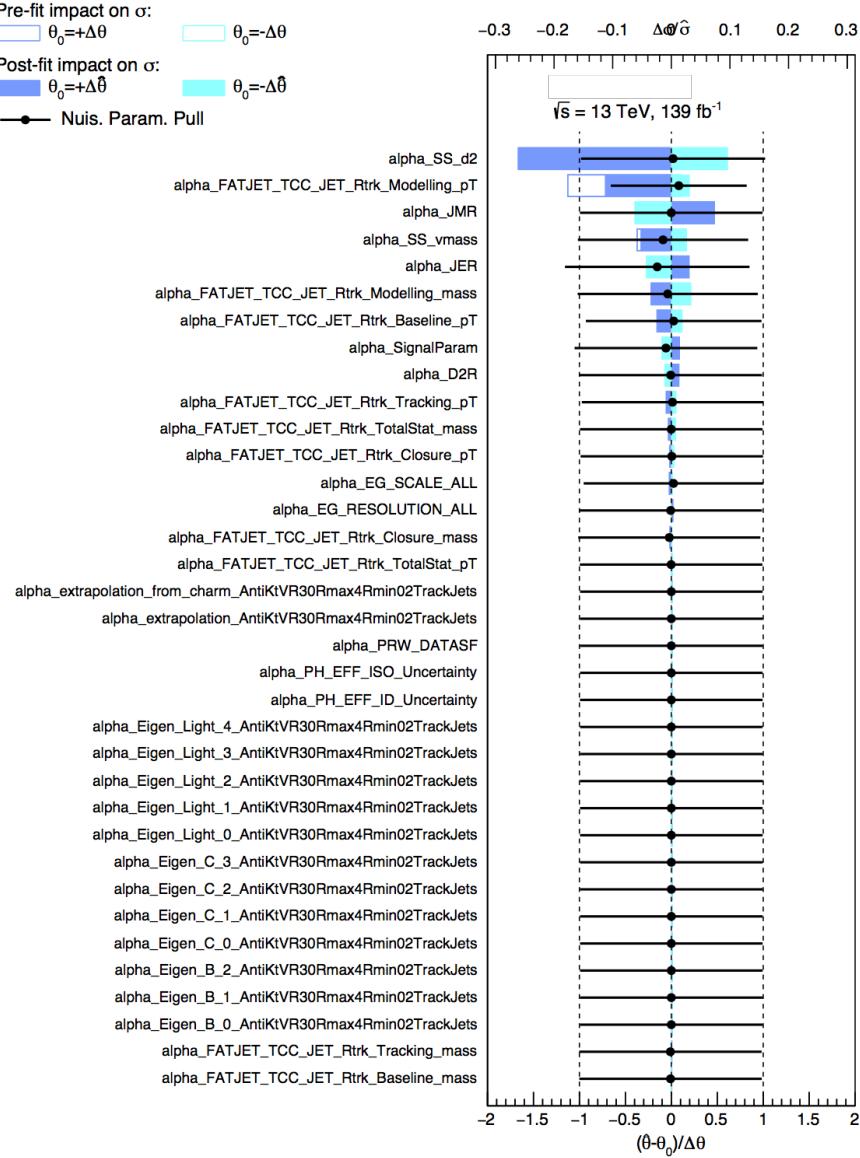


Figure A.34: Pull and ranking distributions for spin=1 $X \rightarrow W\gamma$ $m_X = 2820 \text{ GeV}$ signal. The solid (empty) dark blue hands are the postfit (prefit) impact on σ when the NP is shifted by $+1 \sigma(\theta)$ ($+1$), while the solid (empty) light blue hands are the postfit (prefit) impact on σ when the NP is shifted by $-1 \sigma(\theta)$ ($+1$). The NP values are shown in solid point with error bars. In order to well present the systematic uncertainty impacts on fitted cross section, the impacts are shown as relative values as $\Delta\sigma/\sigma$ (top label).

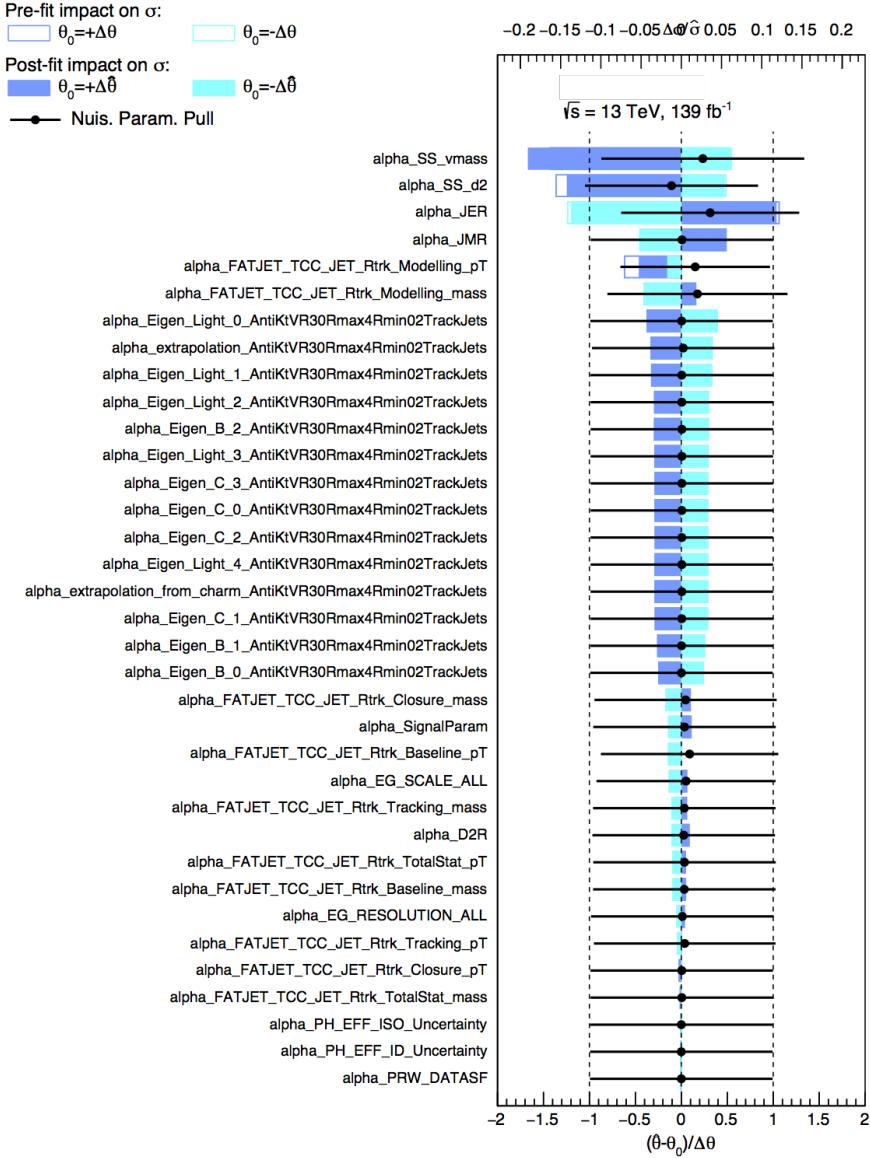


Figure A.35: Pull and ranking distributions for spin=2 $gg \rightarrow X \rightarrow Z\gamma$ $m_X = 3580 \text{ GeV}$ signal. The solid (empty) dark blue hands are the postfit (prefit) impact on σ when the NP is shifted by $+1 \sigma(\theta)$ (+1), while the solid (empty) light blue hands are the postfit (prefit) impact on σ when the NP is shifted by $-1 \sigma(\theta)$ (+1). The NP values are shown in solid point with error bars. In order to well present the systematic uncertainty impacts on fitted cross section, the impacts are shown as relative values as $\Delta\sigma/\sigma$ (top label).

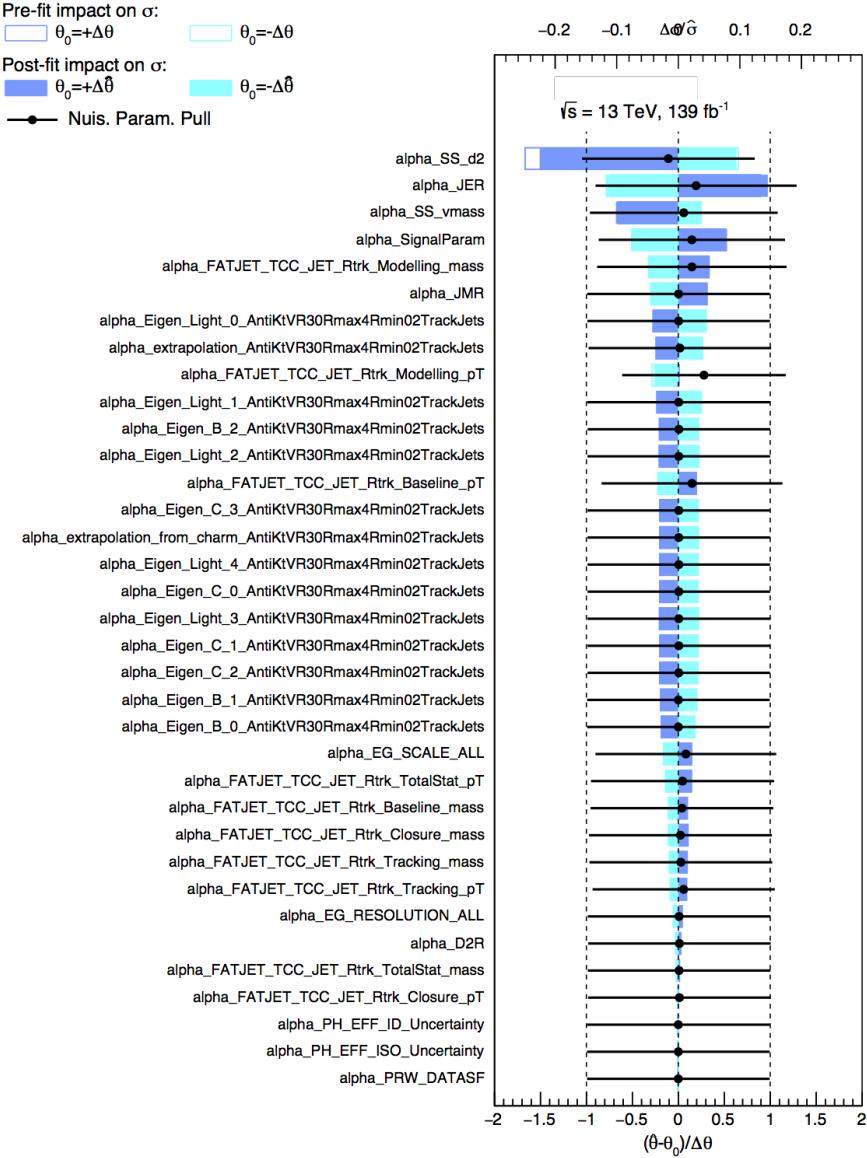


Figure A.36: Pull and ranking distributions for spin=2 $qq \rightarrow X \rightarrow Z\gamma$ $m_X = 3560 \text{ GeV}$ signal. The solid (empty) dark blue hands are the postfit (prefit) impact on σ when the NP is shifted by $+1 \sigma(\theta)$ (+1), while the solid (empty) light blue hands are the postfit (prefit) impact on σ when the NP is shifted by $-1 \sigma(\theta)$ (+1). The NP values are shown in solid point with error bars. In order to well present the systematic uncertainty impacts on fitted cross section, the impacts are shown as relative values as $\Delta\sigma/\sigma$ (top label).

Bibliography

- [1] Particle Data Group et al. “Review of particle physics”. In: *Progress of Theoretical and Experimental Physics* 2020.8 (2020), p. 083C01.
- [2] H Weyl. “Gravitation and Electricity, Sitzungsber. K”. In: *Preuss. Akad. Wiss. Berlin* 465 (1918), p. 1918.
- [3] David Griffiths. *Introduction to elementary particles*. John Wiley & Sons, 2008.
- [4] W Noel Cottingham and Derek A Greenwood. *An introduction to the standard model of particle physics*. Cambridge university press, 2007.
- [5] Peter W Higgs. “Broken symmetries and the masses of gauge bosons”. In: *Physical Review Letters* 13.16 (1964), p. 508.
- [6] François Englert and Robert Brout. “Broken symmetry and the mass of gauge vector mesons”. In: *Physical Review Letters* 13.9 (1964), p. 321.
- [7] Gerald S Guralnik, Carl R Hagen, and Thomas WB Kibble. “Global conservation laws and massless particles”. In: *Physical Review Letters* 13.20 (1964), p. 585.
- [8] Georges Aad et al. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Physics Letters B* 716.1 (2012), pp. 1–29.
- [9] *Standard Model Summary Plots Spring 2020*. Tech. rep. ATL-PHYS-PUB-2020-010. Geneva: CERN, May 2020. URL: <http://cds.cern.ch/record/2718937>.
- [10] Steven Weinberg. “A model of leptons”. In: *Physical review letters* 19.21 (1967), p. 1264.
- [11] Po Bagnaia et al. “Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN pp Collider”. In: *Physics Letters B* 129.1-2 (1983), pp. 130–140.
- [12] Ian Low, Joseph Lykken, and Gabe Shaughnessy. “Singlet scalars as Higgs boson imposters at the Large Hadron Collider”. In: *Physical Review D* 84.3 (Aug. 2011). ISSN: 1550-2368. DOI: 10.1103/physrevd.84.035027. URL: <http://dx.doi.org/10.1103/PhysRevD.84.035027>.
- [13] Duccio Pappadopulo et al. “Heavy vector triplets: bridging theory and data”. In: *Journal of High Energy Physics* 2014.9 (Sept. 2014). ISSN: 1029-8479. DOI: 10.1007/jhep09(2014)060. URL: [http://dx.doi.org/10.1007/JHEP09\(2014\)060](http://dx.doi.org/10.1007/JHEP09(2014)060).
- [14] Nima Arkani-Hamed, Andrew G. Cohen, and Howard Georgi. “Electroweak symmetry breaking from dimensional deconstruction”. In: *Physics Letters B* 513.1-2 (July 2001), 232?240. ISSN: 0370-2693. DOI: 10.1016/s0370-2693(01)00741-9. URL: [http://dx.doi.org/10.1016/S0370-2693\(01\)00741-9](http://dx.doi.org/10.1016/S0370-2693(01)00741-9).

- [15] Lyndon Evans and Philip Bryant. “LHC machine”. In: *Journal of instrumentation* 3.08 (2008), S08001.
- [16] Georges Aad et al. “The ATLAS experiment at the CERN large hadron col- linder”. In: *Jinst* 3 (2008), S08003.
- [17] Roman Adolphi et al. “The CMS experiment at the CERN LHC”. In: *Jinst* 803 (2008), S08004.
- [18] R Perin. “The superconducting magnet system for the LHC”. In: *IEEE Transactions on Magnetics* 27.2 (1991), pp. 1735–1742.
- [19] P Grafström and W Kozanecki. “Luminosity determination at proton colliders”. In: *Progress in Particle and Nuclear Physics* 81 (2015), pp. 97–148.
- [20] G Aad et al. “ATLAS pixel detector electronics and sensors”. In: *Journal of Instrumentation* 3.07 (2008), P07007.
- [21] A Ahmad et al. “The silicon microstrip sensors of the ATLAS semiconductor tracker”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 578.1 (2007), pp. 98–118.
- [22] E Abat et al. “The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance”. In: *Journal of Instrumentation* 3.02 (2008), P02013.
- [23] *ATLAS calorimeter performance: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: <http://cds.cern.ch/record/331059>.
- [24] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: <https://cds.cern.ch/record/331061>.
- [25] *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: <https://cds.cern.ch/record/331062>.
- [26] Richard Wigmans. “Advances in hadron calorimetry”. In: *Annual Review of Nuclear and Particle Science* 41.1 (1991), pp. 133–185.
- [27] Joao Victor da Fonseca Pinto. “Ring-shaped Calorimetry Information for a Neural Egamma Identification with ATLAS Detector”. In: *J. Phys.: Conf. Ser.* Vol. 762. ATL-DAQ-PROC-2016-007. ATL-COM-DAQ-2016-015. 2016, p. 012049.
- [28] Aranzazu Ruiz-Martinez et al. *Electron and photon trigger efficiency plots using the full 2016 dataset*. Tech. rep. ATL-COM-DAQ-2017-015. Geneva: CERN, Mar. 2017. URL: <https://cds.cern.ch/record/2254973>.

- [29] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. “The anti-ktjet clustering algorithm”. In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), 063?063. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2008/04/063. URL: <http://dx.doi.org/10.1088/1126-6708/2008/04/063>.
- [30] *Improving jet substructure performance in ATLAS using Track-Calorimeter Clusters*. Tech. rep. ATL-PHYS-PUB-2017-015. Geneva: CERN, July 2017. URL: <https://cds.cern.ch/record/2275636>.
- [31] Andrew J. Larkoski, Ian Moult, and Duff Neill. *Analytic Boosted Boson Discrimination at the Large Hadron Collider*. 2017. arXiv: 1708.06760 [hep-ph].
- [32] Andrew J. Larkoski, Ian Moult, and Duff Neill. “Power counting to better jet observables”. In: *Journal of High Energy Physics* 2014.12 (Dec. 2014). ISSN: 1029-8479. DOI: 10.1007/jhep12(2014)009. URL: [http://dx.doi.org/10.1007/JHEP12\(2014\)009](http://dx.doi.org/10.1007/JHEP12(2014)009).
- [33] Andrew J. Larkoski, Ian Moult, and Duff Neill. “Non-global logarithms, factorization, and the soft substructure of jets”. In: *Journal of High Energy Physics* 2015.9 (Sept. 2015). ISSN: 1029-8479. DOI: 10.1007/jhep09(2015)143. URL: [http://dx.doi.org/10.1007/JHEP09\(2015\)143](http://dx.doi.org/10.1007/JHEP09(2015)143).
- [34] Dinos Bachas et al. *Search for heavy WW/WZ/ZZ resonances in semi-leptonic final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. ATL-COM-PHYS-2018-1549. Geneva: CERN, Oct. 2018. URL: <https://cds.cern.ch/record/2646593>.
- [35] *Expected performance of the ATLAS b-tagging algorithms in Run-2*. Tech. rep. ATL-PHYS-PUB-2015-022. Geneva: CERN, July 2015. URL: <https://cds.cern.ch/record/2037697>.
- [36] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *The European Physical Journal C* 71.2 (Feb. 2011). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1554-0. URL: <http://dx.doi.org/10.1140/epjc/s10052-011-1554-0>.
- [37] Isaac Asimov. *Franchise, volume 1*. 1990.
- [38] Alexander L Read. “Presentation of search results: the CLs technique”. In: *Journal of Physics G: Nuclear and Particle Physics* 28.10 (2002), p. 2693.
- [39] Stefano Frixione, Paolo Nason, and Carlo Oleari. “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”. In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].
- [40] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. “PYTHIA 6.4 Physics and Manual”. In: *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175 [hep-ph].

- [41] S. Agostinelli et al. “GEANT4: A Simulation toolkit”. In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.
- [42] T. Gleisberg et al. “Event generation with SHERPA 1.1”. In: *JHEP* 02 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].
- [43] Tanju Gleisberg and Stefan Hoeche. “Comix, a new matrix element generator”. In: *JHEP* 12 (2008), p. 039. DOI: 10.1088/1126-6708/2008/12/039. arXiv: 0808.3674 [hep-ph].
- [44] Fabio Cascioli, Philipp Maierhofer, and Stefano Pozzorini. “Scattering Amplitudes with Open Loops”. In: *Phys. Rev. Lett.* 108 (2012), p. 111601. DOI: 10.1103/PhysRevLett.108.111601. arXiv: 1111.5206 [hep-ph].
- [45] Federico Buccioni, Stefano Pozzorini, and Max Zoller. “On-the-fly reduction of open loops”. In: *Eur. Phys. J.* C78.1 (2018), p. 70. DOI: 10.1140/epjc/s10052-018-5562-1. arXiv: 1710.11452 [hep-ph].
- [46] Ana Rosario Cueto Gomez, Claudia Glasman, and Juan Terron. *Measurement of isolated-photon plus two-jet production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. ATL-COM-PHYS-2019-037. Geneva: CERN, Jan. 2019. URL: <https://cds.cern.ch/record/2655001>.
- [47] Morad Aaboud et al. “Search for heavy resonances decaying to a photon and a hadronically decaying Z/W/H boson in p p collisions at $s = 13$ TeV with the ATLAS detector”. In: *Physical Review D* 98.3 (2018), p. 032015.
- [48] M. Oreglia. “A Study of the Reactions $\psi' \rightarrow \gamma\gamma\psi$ ”. PhD thesis. SLAC, 1980. URL: <http://www-public.slac.stanford.edu/sciDoc/docMeta.aspx?slacPubNumber=slac-r-236.html>.
- [49] T. Aaltonen et al. “Search for new particles decaying into dijets in proton-antiproton collisions at $s^{**}(1/2) = 1.96$ -TeV”. In: *Phys. Rev. D* 79 (2009), p. 112002. DOI: 10.1103/PhysRevD.79.112002. arXiv: 0812.4036 [hep-ex].
- [50] Richard G Lomax. *Statistical concepts: A second course*. Lawrence Erlbaum Associates Publishers, 2007.
- [51] Georgios Choudalakis and Diego Casadei. “Plotting the differences between data and expectation”. In: *The European Physical Journal Plus* 127.2 (2012), p. 25.
- [52] *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*. Tech. rep. ATLAS-CONF-2019-021. Geneva: CERN, June 2019. URL: <http://cds.cern.ch/record/2677054>.
- [53] Morad Aaboud et al. “In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector”. In: *The European Physical Journal C* 79.2 (2019), p. 135.

- [54] A. M. Sirunyan et al. “Search for Z resonances using leptonic and hadronic final states in proton-proton collisions at $\sqrt{s} = 13$ TeV”. In: *Journal of High Energy Physics* 2018.9 (Sept. 2018). ISSN: 1029-8479. DOI: 10.1007/jhep09(2018)148. URL: [http://dx.doi.org/10.1007/JHEP09\(2018\)148](http://dx.doi.org/10.1007/JHEP09(2018)148).
- [55] Morad Aaboud et al. “Search for heavy resonances decaying to a Z boson and a photon in pp collisions at $s = 13$ TeV with the ATLAS detector”. In: *Physics Letters B* 764 (2017), pp. 11–30.