Search for New Particles Decaying to Diphotons in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV

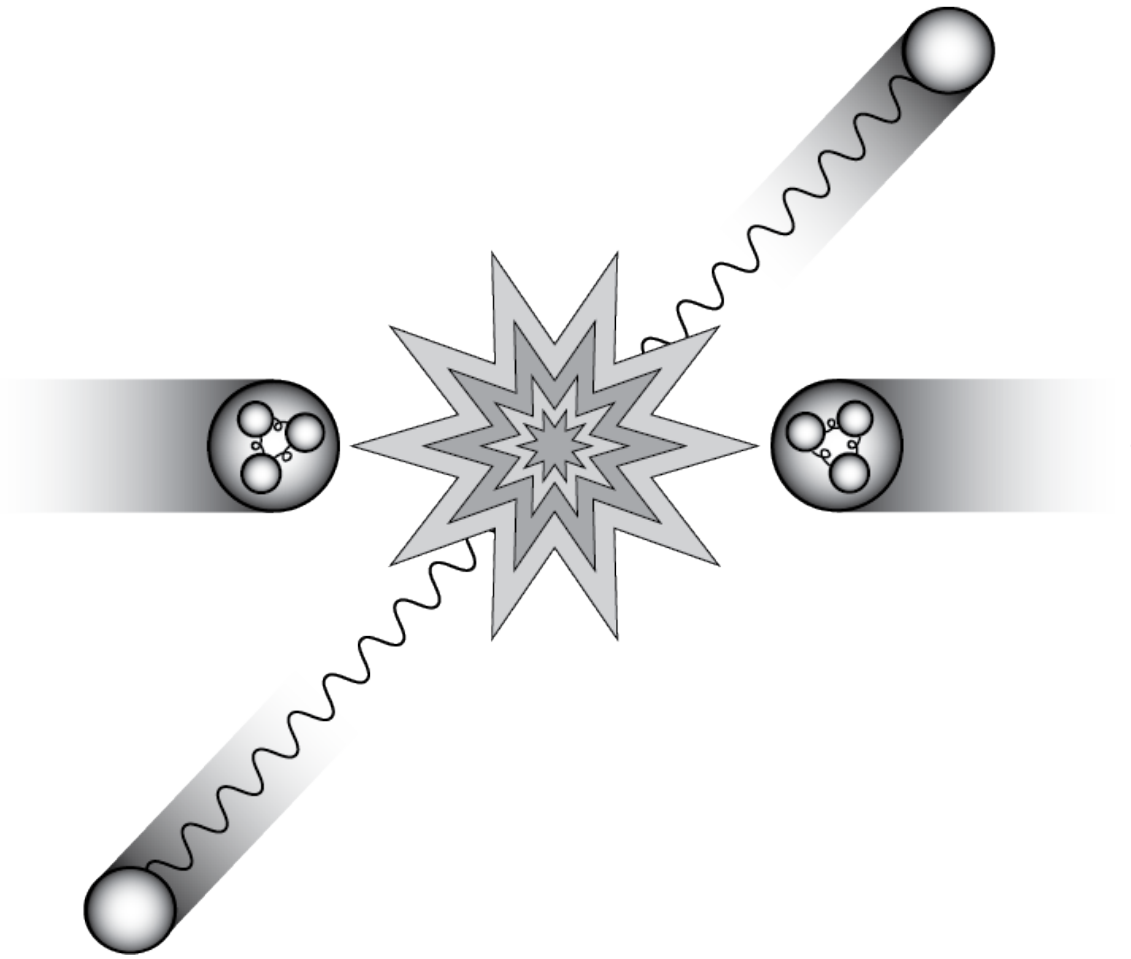
by

Kali Grud

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in The University of Michigan
2021

Doctoral Committee:

Professor Dante Amidei
Professor Eric Bell
Professor Aaron Pierce
Associate Professor Tom Schwarz
Professor Greg Tarle



Cartoon of a proton-proton collision with a diphoton final state.

Kali Grud
cgrud@umich.edu
ORCID iD: 0000-0003-3277-3828

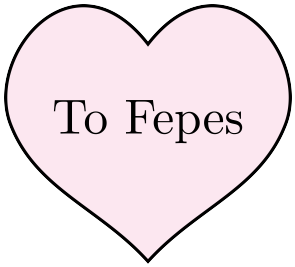

To Fepes

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

LIST OF ABBREVIATIONS

2HDM Two Higgs Doublet Model

2x2D 2×2D sideband method

ALICE A Large Ion Collider Exerpiment

ATLAS A Toroidal LCH ApparatuS

BSM Beyond the Standard Model

BW Breit-Wigner

cdf cumulative distribution function

CL confidence level

CMS Compact Muon Solenoid

COM center of mass

CRT cathode ray tube

CSC Cathode Strip Chambers

CL confidence level

DSCB double sided crystal ball

ECal Electromagnetic Calorimeter

EM electromagnetic

EMB Electromagnetic Barrel Calormeter

EMEC Electromagnetic Endcap Calorimeter

EW electroweak

FCal LAr Forward Calorimeter

FD functional decomposition

FFT Fast Fourier Transform

FSR final state radiation

ggF gluon-gluon fusion

HCal Hadronic Calorimeter

HEC Hadronic Endcap

HLT high level trigger

IBL Insertable B-Layer

ID Inner Detector

ISR initial state radiation

KK Kaluza–Klein

KL Kullback–Leibler

KS Kolmogorov-Smirnov

L1 Level-1

LAr Liquid Argon

LEP Large Electron-Positron Collider

LHC Large Hadron Collider

LHCb LHC-beauty

LINAC2 LINear ACcelerator

LW large width

LO leading order

MC Monte Carlo

MDT Monitored Drift Tubes

MSSM Minimal Supersymmetric Standard Model

NLO next to leading order

NNLO next to next to leading order

NWA narrow width approximation

PD/PIXEL Pixel Detector

PDF parton distribution function

pdf probability density function

PMT photomultiplier tubes

pp proton–proton

PS presampler

PSB Proton Synchrotron Booster

PSyn Proton Synchrotron

RFQ2 Radio Frequency Quadrupole

RMS root mean square

RPC Resistive Plate Chambers

RS Randall-Sundrum

S+B signal plus background

SCT Semi-Conductor Tracker

SS Spurious signal

SM Standard Model

SPS Super Proton Synchrotron

SUSY supersymmetery

TGC Thin Gap Chambers

TileCal Tile Barrel Hadronic Calorimeter

TRT Transition Radiation Tracker

ttH associated Higgs production with $t\bar{t}$

VBF vector boson fusion

VEV vacuum expectation value

QCD Quantum Chromodynamics

QED Quantum Electrodynamics

QFT quantum field theory

ABSTRACT

This thesis presents a search for new resonances decaying to diphotons in 139 fb^{-1} of proton-proton collision data produced by the Large Hadron Collider at a center of mass center of mass (COM) energy of $\sqrt{s} = 13$ TeV. The search was performed in the mass distribution of the detected diphotons using well-tested models for the signal and background line shapes. Functional Decomposition, a new data driven modeling and bump hunting technique inspired by Fourier analysis, was used minimize the uncertainties in the procedure. No new diphoton states are found. New improved 95% confidence level upper limits are placed on the production rate of possible new diphoton states too small to confidently measure with the available data.

CHAPTER I

Introduction

In 1897 the electron was discovered by J. J. Thompson. Since then an effort of tens of thousands of scientists around the world has lead to the discovery and study of an entire world of subatomic particles. In the mid 1970's the *Standard Model of Particle Physics*, a quantum field theory (QFT) which describes the properties and interactions of these particles, took its current form. Over the decades, the Standard Model (SM) has stood up to countless experimental tests, and in 2012 the final missing particle, the Higgs boson, was discovered. However, despite its success in explaining and making predictions about the subatomic world, we know that the SM cannot be complete since there are a number of phenomena in nature that it fails to explain, including quantum gravity, dark matter, and dark energy.

In an effort to explain these phenomena, and others, many new extensions of the SM have been proposed. Most of these extensions include the existence of new particles or new interactions between known particles. A traditional way of searching for new particles is by looking for local excesses, called *resonances*, in an invariant mass spectrum of n objects. If a new particle exists and decays to these n objects, then there would be an excess of events forming a resonance whose invariant mass corresponds to the mass of the decaying particle.

This thesis presents the search for new resonances in the diphoton invariant mass

spectrum. It is performed using the Run-2 data collected at the Large Hadron Collider (LHC) using the ATLAS detector between 2015-2018. This search focuses on narrow width signals, that is signals whose inherent width is much smaller than the resolution of the detector. Two techniques for measuring excesses in the invariant mass spectrum are explored. One technique is performed by fitting the spectrum simultaneously with ad-hoc functions representing the background and signal shapes. The other technique is called functional decomposition (FD) and is inspired by Fourier analysis. It uses a set of orthogonal functions built from integer powers of the exponential function to model a smoothly falling spectrum. A spectrum is modeled as a linear combination of orthogonal functions, which allows any probability density function (pdf) on a semi-infinite interval to be described with arbitrary precision. This ability to describe any invariant mass spectrum with arbitrary precision gives FD an advantage in reducing systematic uncertainties associated with the degree to which a model is capable if capturing the true shape of the underlying pdf. Both delocalized backgrounds and localized signals can be modeled using FD. The delocalized backgrounds are modeled using the lower order terms while localized signals are modeled using mostly the higher order terms.

The remainder of this thesis is organized as follows. Chapter 2 describes the SM and introduces some of its extensions which are needed to describe the existence of new particles. Chapter 3 gives details about the LHC and the ATLAS detector. Chapter 4 describes hows photons are measured by the ATLAS detector. Chapter 5 gives additional details about the production of diphoton events, events which fake diphoton events, and theoretical predictions for new phenomena in diphoton production. Chapter 6 gives an overview of mass resonance search techniques and past results. Chapter 7 describes the simulated samples used in this analysis. Chapter 8 gives details of the analysis event selection, background modeling and signal modeling. Chapter 9 describes in detail FD and how it is used to make a data driven signal

plus background model. Chapter 10 describes the systematic uncertainties in the analysis. Chapter 11 describes the search for new resonances in the diphoton spectrum using FD. Finally, Chapter 12 summarizes and draws conclusions from the results of Chapter 11 and describes future prospects for FD.

CHAPTER II

The Standard Model of Particle Physics

There are four fundamental forces in nature: the strong force, the electromagnetic (EM) force, the weak force, and gravity. The strong force is responsible for holding the nucleus of an atom together. The EM force is responsible for binding electrons to atoms. The weak force is responsible for the radioactive decays of atoms. Gravity is responsible for collecting atoms into large celestial bodies. Some properties of these forces are shown in Table 2.1.

Force	Relatative Strength	Range	Force Carriers
Strong	1	10^{-15} m	Gluons g
Electromagnetic	10^{-2}	∞	Photons γ
Weak	10^{-6}	10^{-18} m	Heavy gauge bosons $W^{\pm}$,Z
Gravitional	10^{-38}	∞	Gravitons G

Table 2.1: Properties of the four fundamental forces

The SM of particle physics is the mathematical framework which explains three of the four fundamental forces (excluding gravity) and how they interact with matter. It describes how *fermions*, the particles of matter, *gauge bosons*, force carrying particles, and a *scalar boson*, an excitation of the quantum field responsible for giving particles mass, interact to create the world we live in.

The remainder of this chapter begins by describing the particles of the known subatomic world and their interactions with each other. Since this thesis presents

a diphoton search, particular attention is paid to describing the properties and interactions of the photon. Next it gives a brief overview of the limitations of our current understanding of the subatomic world, and it ends by describing models for new physics which solve some of these limitations and predict diphoton final states.

2.1 Fermions

Fermions are the matter particles of the Universe. They are spin-1/2 point particles which are divided into two groups, *quarks* and *leptons*. Both quarks and leptons come in three *generations* and six *flavors* which differ by mass, electric charge, and other quantum numbers. In addition fermions have a property known as *chirality* which comes in two values *left handed* and *right handed*. When a fermion is rotated in space the phase of its wavefunction will shift, the sign of the phase shift is determined by the particle's chirality.

2.1.1 Leptons

There are three electrically neutral and three electrically charged leptons. The electrically charged leptons are the electron (e^-), muon (μ^-), tau (τ^-) and their antiparticles. The electron is the lightest of the electrically charged leptons, next is the muon, and finally the tau. Their generations are numbered according to increasing mass, that is the electron is in Generation I, the muon is in Generation II, and the tau is in Generation III. The electrically charged leptons participate in the EM force, the weak force, and the gravitational force.

The electrically neutral leptons are the electron neutrino (ν_e), muon neutrino (ν_μ), tau neutrino (ν_τ) and their antiparticles. Their generations are numbered according to the generation of their associated charged lepton, that is the electron neutrino is in Generation I, the muon neutrino in Generation II, and the tau neutrino in Generation III. Neutrinos participate in the weak force and the gravitational force, however since

Quark	Symbol	Mass
up	u	$2.16^{+0.49}_{-0.26}$ MeV/c^2
down	d	$4.67^{+0.48}_{-0.17}$ MeV/c^2
charm	c	1.27 ± 0.02 GeV/c^2
strange	s	93^{+11}_{-5} MeV/c^2
top	t	172.9 ± 0.4 GeV/c^2
bottom	b	$4.18^{+0.03}_{-0.02}$ GeV/c^2

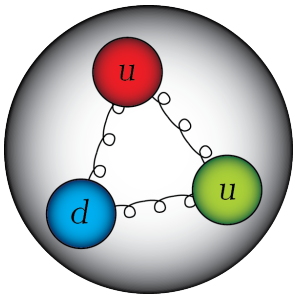
Table 2.2: Masses of the quarks (*Tanabashi, M. et al.*, 2018).

they are electrically neutral they do not participate in the EM force.

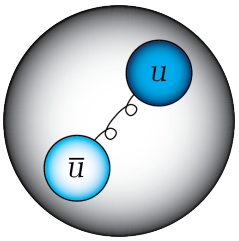
2.1.2 Quarks

Quarks are the only SM particle to participate in all four of the fundamental forces. They come in six flavors: up (u), charm (c), top (t), down (d), strange (s), and bottom (b). The up and down quark make up Generation I, the charm and strange quark make up Generation II, and the top and bottom quark make up Generation III. The u, c, and t quarks are known as *up-type* quarks and have electric charge $Q = +2/3$. The d, s, and b quarks are known as *down-type* quarks and have electric charge $Q = -1/3$. The masses of the quarks are summarized in Table 2.2. The quarks carry what is known as *color charge*, an inherent property of the quark comparable to electric charge. There are three colors red, green, and blue, and three anticolors antired, antigreen, and antiblue. They exist in colorless bound states, known as *hadrons*. There are two types of hardons, *mesons* and *baryons*, illustrated in Figure 2.1. Mesons typically comprise a quark and antiquark pair while baryons are typically composed of three quarks or antiquarks*.

* There are exotic mesons of four or five quarks in a bound state known as tetraquarks and pentaquarks respectively. There have been recent experimental evidence consistent with the existence of these particles (*Aaij et al.*, 2015).



(a) Proton (p)



(b) Pion (π^0)

Figure 2.1: Cartoon of a hadrons. (a) Baryons are composed of three quarks or three antiquarks. Each (anti)quark must carry a different (anti)color to make a colorless state. (b) Mesons are composed of a quark and antiquark of complimentary color and anticolor to make a colorless state.

2.2 Bosons

Bosons are particles whose spin is an integer value. There are five bosons in the SM, the Higgs boson, photon, gluon, and Z and $W^{\pm}$ bosons. The Higgs boson is an electrically neutral massive scalar boson, i.e. it is a spin-0 particle. The photon and gluon are both massless vector bosons, i.e. they are spin-1 particles, while the $W^{\pm}$ and Z bosons make up the massive vector bosons. Table 2.3 shows the properties of the bosons.

Name	Symbol	Mass (GeV/c^2)	Spin	Full Width (GeV/c^2)
Photon	γ	0	1	Stable
W Boson	$W^{\pm}$	80.379 ± 0.012	1	2.085 ± 0.042
Z Boson	Z	91.1876 ± 0.0021	1	2.4952 ± 0.0023
Gluon	g	0	1	Bound in hadrons
Higgs	h	125.10 ± 0.14	0	< 0.013

Table 2.3: Properties of the bosons (*Tanabashi, M. et al.*, 2018).

2.2.1 The Photon

The photon is the particle of light and the massless force carrier for the EM force. It is a spin-1 particle, and since it is massless there is no $S_z = 0$ spin projection. The interaction between photons and matter are described by a theory called Quantum

Electrodynamics (QED). In QED photons couple to fermions which have non-zero electric charge. The Lagrangian of QED is written

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma^{\mu}D_{\mu} - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.1}$$

$$D_{\mu} = \partial_{\mu} + ieA_{\mu} \tag{2.2}$$

$$F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu} \tag{2.3}$$

where γ^{μ} are the Dirac matrices, ψ is the field associated with an electrically charged fermion, $\bar{\psi} = \psi^{\dagger}\gamma^{0}$ is the Dirac adjoint and $\psi^{\dagger}$ is the *Hermitian Conjugate* of ψ, e is the coupling constant equal to the electric charge of the particle, m is the mass of the particle, A_{μ} is the covariant four potential of the EM field.

2.2.2 The $W^{\pm}$ and Z Bosons

The $W^{\pm}$ and Z bosons comprise the massive vector bosons. They are the force carriers of the weak interaction. Because of their large masses, $m_{W^{\pm}} = 80.379 \pm 0.012$ GeV/c^2 and $m_Z = 91.1876 \pm 0.0021$ GeV/c^2, the $W^{\pm}$ and Z bosons have extremely short lifetimes of approximately 10^{-25} seconds and the range of the weak interaction is limited to approximately 10^{-18} meters (*Tanabashi, M. et al.*, 2018).

Only left handed particles couple to the $W^{\pm}$ boson. A particle that absorbs or emits a $W^{\pm}$ boson will change electrical charge and spin by one unit, as well as changing its flavor. For example a d quark that emits a W^{-} will be transformed into a u quark. This leads to nuclear decay since the quarks that make up the nucleus of the atom may absorb or emit a $W^{\pm}$ boson, transforming a proton to a neutron or vice versa. The full decay width the $W^{\pm}$ boson is $\Gamma = 2.085 \pm 0.042$ GeV/c^2 (*Tanabashi, M. et al.*, 2018). It decays to either a lepton and antilepton or to a quark and antiquark of opposing types, that is a up-type and down-type.

The Z boson is its own antiparticle. It couples to both left handed and right

handed fermions but with different strengths. A particle that absorbs or emits a Z boson will not change flavors or electrical charge, only spin and momentum. The full decay width is $\Gamma = 2.4952 \pm 0.0023$ GeV/c^2 (*Tanabashi, M. et al.*, 2018). Since it is electrically neutral it decays to a fermion and its antiparticle to conserve charge.

2.2.3 The Higgs Boson

The Higgs boson is the only fundamental scalar in the SM. It is a quantum excitation of the Higgs field, the field responsible for giving mass to the massive vector bosons and fermions. Discovered in 2012, its mass is $m_h = 125.10 \pm 0.14$ GeV/c^2, and has an extremely narrow decay width of $\Gamma < 0.013$ GeV/c^2 (*Tanabashi, M. et al.*, 2018). The Higgs couples to all the massive elementary particles of the SM and so has many decay channels. It can also decay into massless gauge bosons, but requires an intermediate loop of a virtual quark or massive vector boson, as shown in Figure 2.2. These decay modes enabled the diphoton channel to be used as one of the discovery

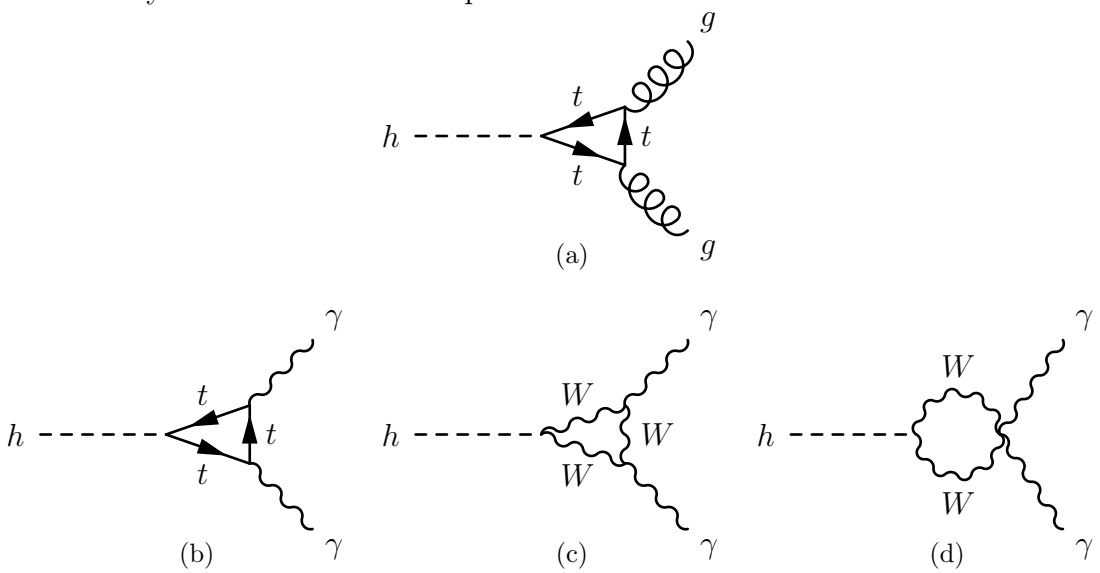


Figure 2.2: Feynman diagrams of the Higgs boson decaying into massless gauge bosons.

channels of the Higgs boson.

2.2.4 The Gluon

Gluons are the particles responsible for holding the nucleus of an atom together. They couple to particles which carry color charge. The gluons themselves carry one unit of color and one unit of anticolor, and so are self interacting. However the *color singlet* state of a gluon does not exist, that is if one were able to measure the color of a gluon, a gluon in the color singlet state would have an equal probability of being red-antired, blue-antiblue, or green-antigreen. Gluons exist in a super position of eight states described by the *color octet*. There are many ways of representing these states but a commonly used list is:

$$\begin{aligned} (r\bar{b} + b\bar{r})/\sqrt{2} \qquad & -i(r\bar{b} - b\bar{r})/\sqrt{2} \\ (r\bar{g} + g\bar{r})/\sqrt{2} \qquad & -i(r\bar{g} - g\bar{r})/\sqrt{2} \\ (b\bar{g} + g\bar{b})/\sqrt{2} \qquad & -i(b\bar{g} - g\bar{b})/\sqrt{2} \\ (r\bar{r} - b\bar{b})/\sqrt{2} \qquad & (r\bar{r} + b\bar{b} + 2g\bar{g})/\sqrt{6} \end{aligned}$$

These states are linearly independent and cannot be combined to form the color singlet state. A visualization of the color octet is shown in Figure 2.3.

2.3 The Standard Model of Particle Physics

The SM is a relativistic QFT that can be written as the product of the symmetry groups

$$\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y \tag{2.4}$$

where $\mathrm{SU}(n)$ are the Special Unitary groups, and $\mathrm{U}(n)$ are the Unitary groups. Each group represents a symmetry, that is a set of transformations which when applied to a physical system leaves it unchanged. $\mathrm{U}(1)_Y$ represents the group of phase rotations

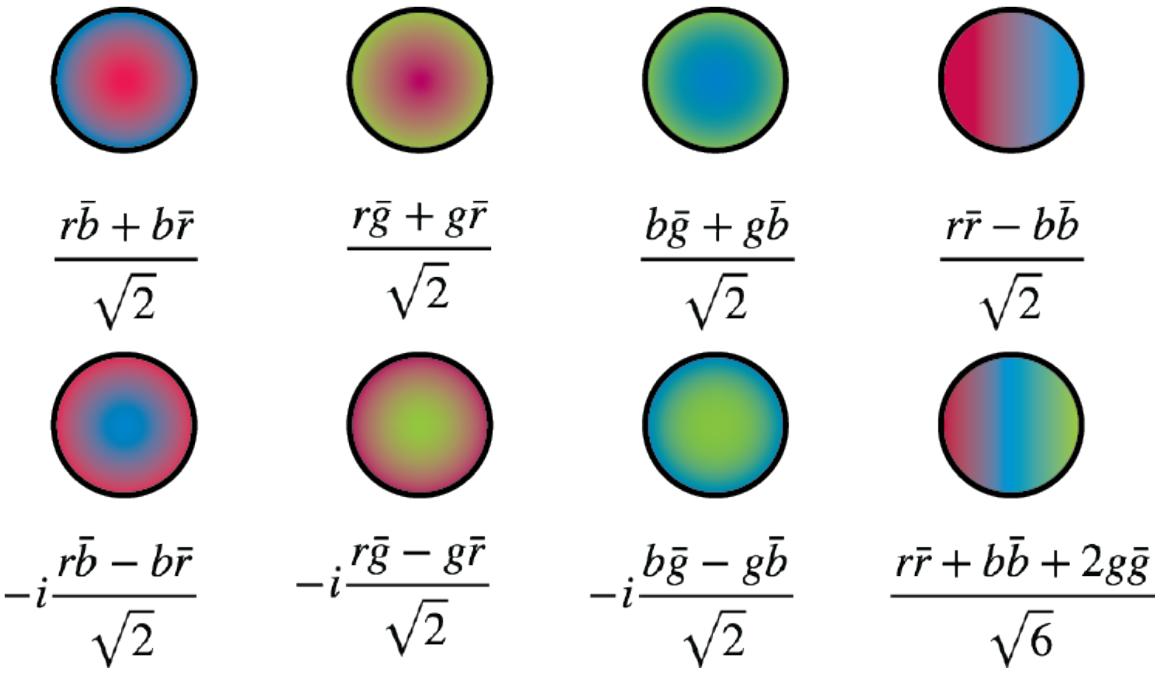


Figure 2.3: Visualization of the color octet.

on a single complex variable, while $\mathrm{SU}(2)_L$ and $\mathrm{SU}(3)_C$ represent the groups of phase rotations on two and three complex variables respectively.

Each symmetry group is associated with a quantum number, sometimes called a charge, as well as one or more gauge bosons. The number of gauge bosons will match the number of generators for a particular group. The $\mathrm{U}(1)_Y$ group's charge, denoted by Y, is called *weak hypercharge*. Since $\mathrm{U}(1)_Y$ has one generator it has one gauge boson called B. One of the $\mathrm{SU}(2)_L$ group's charges, denoted by T, is called *weak isospin*, and because $\mathrm{SU}(2)_L$ has three generators it has three gauge bosons called $W^{a=1,2,3}$. Note that the subscript L here stands for *left* since only left handed fermions obey this symmetery. *Color* charge is associated with $\mathrm{SU}(3)_C$ and comes in six values: *red*, *green*, *blue*, *antired*, *antigreen*, and *antiblue*. Since there are eight generators for $\mathrm{SU}(3)_C$ there are eight gauge bosons for this groups called gluons and are represented by $G^{a=1,2,\ldots,8}$.

The EM and weak interactions are unified in a theory developed by Glashow, Salam, and Weinberg (*Glashow*, 1959)(*Salam and et al*, 1964)(*Weinberg*, 1967). Al-

though these seem like two very different forces, at high enough energies, above about 246 GeV, they merge into a single force known as the electroweak (EW) force.

2.3.1 The Electroweak Interaction

All fermions participate in the EW interaction, however how they participate is quite different. All fermions carry the $\mathrm{U}(1)_Y$ hypercharge Y, but only pairs of left handed fermions, called doublets, transform under the $\mathrm{SU}(2)_L$ weak isospin T. Right handed fermions are electroweak singlets, that is they do not transform under $\mathrm{SU}(2)_L$. The left handed doublets are written

$$\begin{pmatrix} \nu_e \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau_L \end{pmatrix}, \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix}, \tag{2.5}$$

while the right handed singlets are written

$$e_R, \mu_R, \tau_R, u_R, c_R, t_R, d_R, s_R, b_R. \tag{2.6}$$

The physical gauge bosons of the EW interaction, γ, $W^\pm$ and Z boson, are written as combinations of the B and $W^{a=1,2,3}$ bosons

$$W^+ = \frac{1}{\sqrt{2}}\left(W^1 - iW^2\right) \tag{2.7}$$

$$W^- = \frac{1}{\sqrt{2}}\left(W^1 + iW^2\right) \tag{2.8}$$

$$Z = \cos\theta_W \cdot W^3 - \sin\theta_W \cdot B \tag{2.9}$$

$$\gamma = \sin\theta_W \cdot W^3 + \cos\theta_W \cdot B \tag{2.10}$$

where θ_W is known as the *Weinberg angle*.

Through interactions with the $W^\pm$ boson, the doublets are rotated in $\mathrm{SU}(2)_L$ space, and can transform up type elements to down type elements and vice versa.

These interactions along with the interactions between the γ/Z boson and the fermions together constitute the EW force. Figure 2.4 shows the Feynman diagrams for EW interactions.

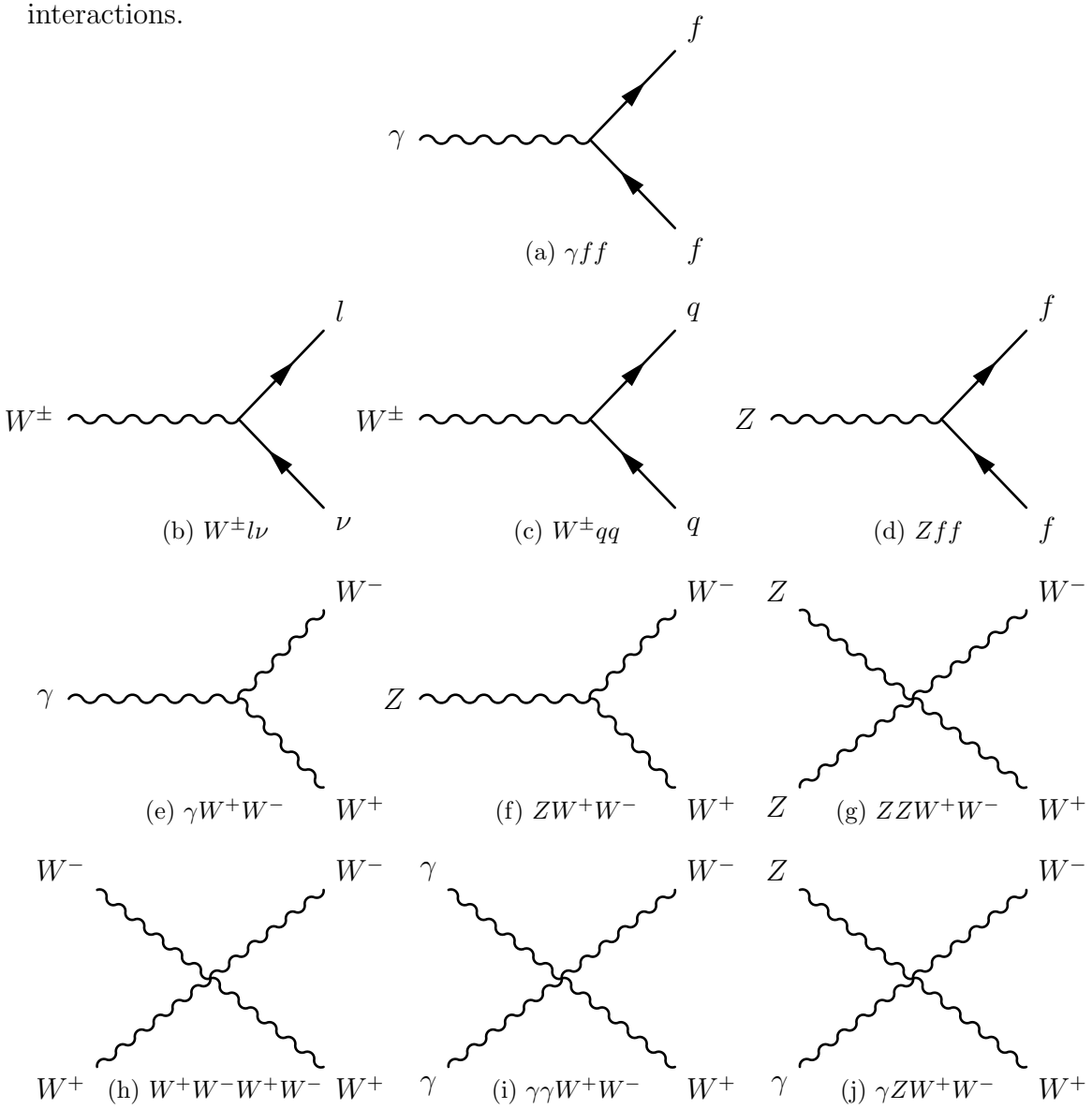


Figure 2.4: Feynman diagrams for EW interactions

The EW charges of the fermions are shown in Table 2.4.

2.3.2 Spontaneous Electroweak Symmetry Breaking

A viable theory to describe the EW interaction must explain the following experimental facts:

	Generation			Electroweak Charge		
	I	II	III	Y	T_3	Q
Quarks	u_L	c_L	t_L	1/3	1/2	2/3
	d_L	s_L	b_L	1/3	-1/2	-1/3
	u_R	c_R	t_R	4/3	0	2/3
	d_R	s_R	b_R	-2/3	0	-1/3
Leptons	ν_{eL}	$\nu_{\mu L}$	$\nu_{\tau L}$	-1	1/2	0
	e_L	μ_L	τ_L	-1	-1/2	-1
	e_R	μ_R	τ_R	-2	0	-1

Table 2.4: Electroweak charges of the SM fermions.

1. γ couples to left handed fermions and right handed fermions with the same strength

2. Z couples differently to left handed fermions and right handed fermions

3. $W^{\pm}$ couples only to left handed fermions

The theory must also be consistent with experimental tests of the relative strengths of fermion-antifermion-vector couplings and the ratio of the $W^{\pm}$ and Z masses. These features are all explained by the sponaneously broken gauge theory known as the Standard Model of Electroweak Interactions. The gauge symmetry breaking can be written

$$\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y \rightarrow \mathrm{U}(1)_{\mathrm{EM}}. \tag{2.11}$$

This is achieved by introducing the Higgs field, a complex scalar $\mathrm{SU}(2)_L$ doublet

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \tag{2.12}$$

The Higgs potential can be written

$$V(\Phi, \Phi^\dagger) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \tag{2.13}$$

since the combination $\Phi^\dagger\Phi$ is a gauge singlet. That is, since under gauge transformations Φ transforms as

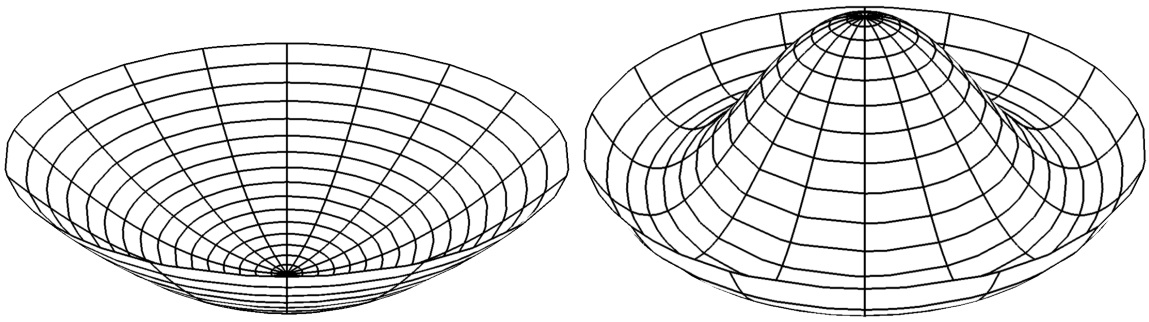
$$\mathrm{SU}(2)_L: \quad \Phi(x) \to \Phi'(x) = e^{i\theta^a(x)\sigma^a/2}\Phi(x) \tag{2.14}$$

$$\Phi^\dagger(x) \to \Phi^{\dagger\prime} = \Phi^\dagger e^{-i\theta^a(x)\sigma^a/2} \tag{2.15}$$

$$\mathrm{U}(1)_Y: \quad \Phi(x) \to \Phi'(x) = e^{i\theta(x)}\Phi(x) \tag{2.16}$$

$$\Phi^\dagger(x) \to \Phi^{\dagger\prime}(x) = \Phi^\dagger e^{-i\theta(x)} \tag{2.17}$$

the product $\Phi^\dagger\Phi$ is constant. In the case that $\mu^2 > 0$ there is a local minimum at $\Phi = \left(\begin{smallmatrix}0\\0\end{smallmatrix}\right)$. However, if $\mu^2 < 0$ then symmetry is broken and the minimum is located in the valley $|\Phi| = \sqrt{-\frac{\mu^2}{2\lambda}}$. The value of this minimum is called the vacuum expectation value (VEV). A visualization of the Higgs potential can be seen in Figure 2.5.



(a) (b)

Figure 2.5: Higgs potential with (a) $\mu^2 > 0$ and (b) $\mu^2 < 0$

Without loss of generality, the VEV of Φ can be chosen to be real and entirely in the electrical neutral component of the Higgs field:

$$\langle 0|\Phi|0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\v\end{pmatrix} \tag{2.18}$$

this is because one can always achieve such an arangement by performing an $\mathrm{SU}(2)_L$ gauge transformation. The Higgs field has two complex scalar degrees of freedom, so

it can be written

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} e^{iG^a(x)\sigma^a/2v} \tag{2.19}$$

and by performing an $\mathrm{SU}(2)_L$ gauge transformation with $\theta^a = -G^a/v$, known as the unitary gauge, it can be written

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{2.20}$$

eliminating G^a, the would be Goldstone bosons, completely.

2.3.2.1 Vector Boson Mass

Using Φ to build the kinetic term of the Higgs Lagrangian density gives

$$\mathcal{L}_{\Phi\,\mathrm{kinetic}} = D^\mu \Phi^\dagger D_\mu \Phi = \frac{1}{2}\partial_\mu h \partial^\mu h + \frac{(v+h)^2}{4}\left[g^2 W_\mu^+ W^{-\,\mu} + \frac{1}{2}(g^2 + g'^2) Z_\mu Z^\mu\right] \tag{2.21}$$

where g and g' are the coupling constants associated with $\mathrm{SU}(2)_L$ and $\mathrm{U}(1)_Y$ respectively. The terms proportional to v^2 give the mass for the Z and $W^\pm$ bosons, while the terms proportional to h and h^2 describe the interactions between the Higgs boson and the massive vector bosons. There is no mass term for photon field A^μ, demonstrating that the photon remains massless, in agreement with the fact that the $\mathrm{U}(1)_{\mathrm{EM}}$ symmetry remains unbroken. The ratio of the $W^\pm$ and Z boson masses can be written as

$$m_W/m_Z = \cos\theta_W \tag{2.22}$$

in agreement with experimental data.

2.3.2.2 Fermion Mass

Another gauge invariant term one can build with the Higgs uses a $\mathrm{SU}(2)_L$ singlet, its corresponding $\mathrm{SU}(2)_L$ doublet, and the Higgs field, and can be written

$$\mathcal{L}_{\text{d Yukawa}} = -y_d \begin{pmatrix} \bar{u}_L & \bar{d}_L \end{pmatrix} \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} d_R + h.c. \tag{2.23}$$

where y_d is a *Yukawa* coupling, and u and d represent up type (i.e. u, c, t, ν_e, ν_μ, ν_τ) and down type (i.e. d, s, b, e, μ, τ) fermions. The $\mathrm{SU}(2)_L$ doublets, $\begin{pmatrix} \bar{u}_L & \bar{d}_L \end{pmatrix}$ and the Higgs field, carry weak hypercharge $Y = +1/2$, while the $\mathrm{SU}(2)_L$ singlet, d_R, carries weak hypercharge $Y = -1$, so the term is an $\mathrm{U}(1)_Y$ singlet. Since the doublets transform under $\mathrm{SU}(2)_L$ as

$$\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \rightarrow e^{-i\theta^a \sigma^a/2} \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{2.24}$$

$$\begin{pmatrix} \bar{u}_L & \bar{d}_L \end{pmatrix} \rightarrow \begin{pmatrix} \bar{u}_L & \bar{d}_L \end{pmatrix} e^{+i\theta^a \sigma^a/2} \tag{2.25}$$

the term is also an $\mathrm{SU}(2)_L$ singlet. Going to the unitary gauge the term becomes

$$\mathcal{L}_{d\,\text{Yukawa}} = -\frac{y_d}{\sqrt{2}}(v+h)(\bar{d}_L d_R + \bar{d}_R d_L) \tag{2.26}$$

$$= -\frac{y_d}{\sqrt{2}}(v+h)\bar{d}d. \tag{2.27}$$

This procedure works the same with up type fermions (other than neutrinos) as well, and can be written more generally as

$$\mathcal{L}_{f\,\text{Yukawa}} = -\frac{y_f}{\sqrt{2}}(v+h)\bar{f}f. \tag{2.28}$$

Immediately one can read off the fermion mass as

$$m_f = \frac{y_f v}{\sqrt{2}}. \tag{2.29}$$

Note that there are no right handed neutrinos in the SM, so one cannot build such a term to obtain neutrino masses.

2.3.2.3 Higgs Mass and Higgs Interactions

Going to the unitary gauge the Higgs potential becomes

$$V(\Phi^\dagger, \Phi) = -\frac{\mu^4}{4\lambda} - \mu^2 h^2 + \lambda v h^3 + \frac{\lambda}{4} h^4. \tag{2.30}$$

The second term is the mass term for the Higgs, and although it can be written

$$m_H = \sqrt{2\lambda v^2}, \tag{2.31}$$

the theory does not predict its mass, since λ is a free parameter. However, from experiment the mass of the Higgs boson is known to be about 125 GeV.

The third and fourth terms are the Higgs self interaction terms giving both three and four Higgs vertices. The term proportional to $h\bar{f}f$ from Equation 2.28 describes how the Higgs interacts with fermions while the terms in Equation 2.21 proportional to h and h^2 describes how the Higgs interacts with the massive gauge bosons. Figure 2.6 shows the Feynman diagrams involving the Higgs boson.

2.3.3 The Strong Interaction

In Quantum Chromodynamics (QCD) quarks carry one unit of the $SU(3)_C$ color charge (antiquarks carry one unit of anticolor) and interact with each other via the exchange of massless gluons. The gluons themselves carry one unit of color and one

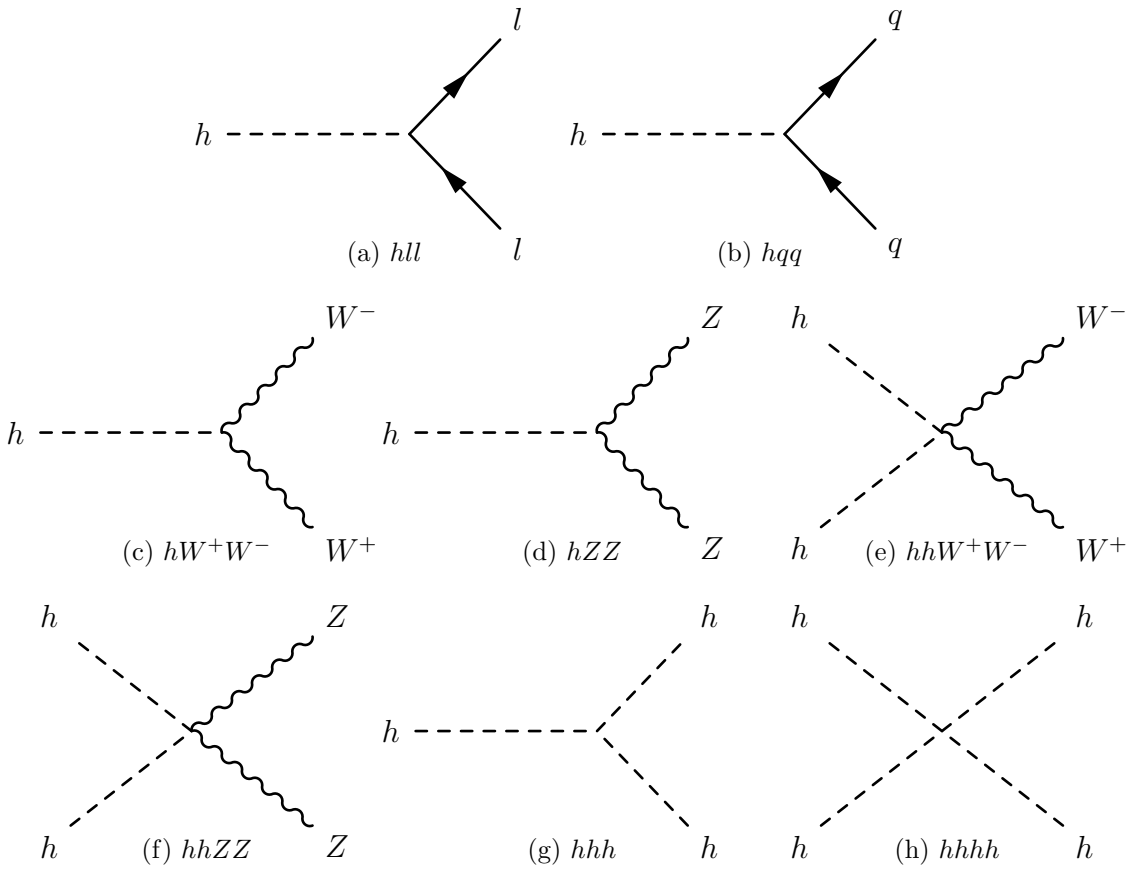


Figure 2.6: Feynman diagrams involving the Higgs boson

unit of anticolor, and so are self interacting. Figure 2.7 shows the Feynman diagrams of the strong force.

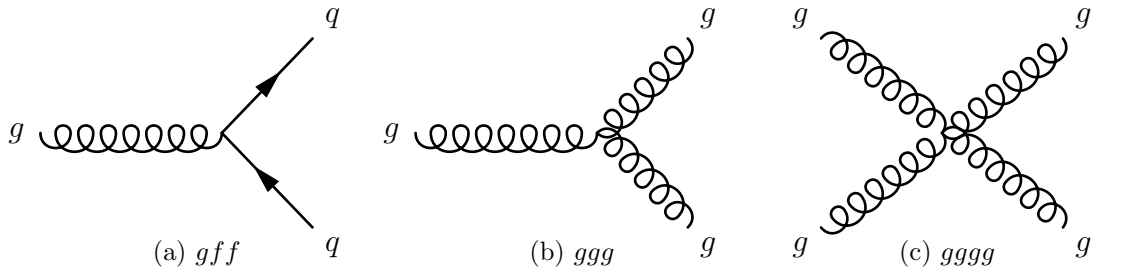


Figure 2.7: Feynman diagrams of QCD

The coupling constant associated with $\mathrm{SU}(3)_C$ is α_s and is related to the four-momentum transfer Q^2 by the relation

$$\alpha_s(Q^2) = \frac{1}{b_0 \ln(Q^2/\Lambda^2_{\mathrm{QCD}})} \tag{2.32}$$

where $b_0 = (11n_c - 2n_f)/12\pi$, Λ^2_{QCD} is the QCD energy scale, n_c is the number of colors and n_f is the number of quark flavors.

For low energies, where $Q^2 < \Lambda^2_{\text{QCD}}$, quarks exist in bound states known as *hadrons*. There are two types of hadrons: *mesons* and *baryons*. Mesons are composed of a quark, carrying color charge, and antiquark, carrying the corresponding anticolor charge (e.g. red and antired), so that the meson itself is colorless. Baryons are comprising a red, green and blue quark (or an antired, antigreen and antiblue antiquark) so that these three quark bound states are also colorless. In fact all hadron *must* be colorless. This is due to a phenomonon called *color confinement*, in which particles carrying net color charge cannot be isolated.

Since gluons themselves carry color charge, when two bound color charged particles are separated in space the gluon field binding them forms a narrow *flux tube* between them and the force between the particles remains constant. As the two particles are separated, the potential energy increases until it becomes energetically favorable to create new quark-antiquark pairs rather than extending the tube further. This leads to a phenomenon known as *hadronization* in which high energy quarks and gluons will form conical sprays of particles, known as *jets*, as they pass through matter.

For high energies, $Q^2 > \Lambda^2_{\text{QCD}}$, the predictions of QCD are computed perturbatively, that is using a power-series expansion in the coupling constant α_s. This is because for large Q^2 the coupling between quarks and gluons weakens and they behave almost like free particles. The simplest perturbative model is known as the leading order (LO) prediction and uses only the first term of the expansion and is represented by tree level Feynman diagrams, that is diagrams with order $\mathcal{O}(\alpha_s^2)$. The next simplest models are the next to leading order (NLO) predicition and the next to next to leading order (NNLO) predicition. These use the first two and three terms of the expansion respectively and are represented by Feynman diagrams with $\mathcal{O}(\alpha_s^3)$ and $\mathcal{O}(\alpha_s^4)$ resprectively.

2.4 Limitations of the Standard Model

Although the SM has been wildly successful at making predictions, the theory is incomplete. There are a number of phenomena in nature which the SM fails to explain.

- Baryon Asymmetry - As per current understanding the Big Bang should have created nearly equal parts matter and antimatter in the early Universe. However, equal parts matter and antimatter are not observed in nature, in fact matter is observed almost exclusively. The SM predicts the violation of CP-symmetry, also known as *charge conjugate parity symmetry*, requires that the physics of a particle is unchanged under a combination of C-symmetry [†] and P-symmetry [‡], observed in K^0 and B^0 mesons, but this can only account for a very small portion of the asymmetry observed.

- Gravity - The SM does not explain the existence of gravity. When one tries to describe gravity as a QFT it becomes non-renormalizable, so an alternative theory must explain its origin. It is believed that the gravitational force is carried by a spin-2 particle called the graviton but as of yet there is no strong experimental evidence for its existence.

- Dark Matter - Cosmological observations of the arms of spiral galaxies, as well as other scientific observations, indicate the existence of matter that only interacts with SM particles through the gravitational force or interacts so weakly that it has not been detected by any known methods. Without dark matter, calculuations show that many galaxies would not have the gravitional force required to maintain the structures observed in nature. It makes up about 27% of

[†] C-symmetry, or charge symmetry, requires that the physics of a paricle remains unchanged if interchanged with its antiparticle.

[‡] P-symmetry, or parity symmetry, requires that the physics of a paricle remains unchanged if its spatial coordinates are inverted.

the Universe, while *normal* matter makes up about 5%. It is called dark matter since it has yet to be directly detected.

- Dark Energy - Experimental measurements of the expansion of the Universe suggest the existence of a form of energy known as *dark energy* which drives the accelerated expansion of the Universer. Dark energy makes up about 68% of the Universe and is believed to permeate all of space driving the expansion of the Universe. It is called dark energy since its fundamental nature is not at this time well understood.

- Neutrino Mass - The SM does not explain neutrino mass. However, from experimental data it is known that neutrinos oscillate flavors. This is only possible if the neutrinos are massive. Thus far no direct experimental measurement of neutrino mass has been made, although measurements of quantities related to the masses, such as the limit on the sum of neutrinos masses have been made.

- Hierarchy Problem - When comparing the forces two protons impose on each other in an atomic nucleus, the weak force is 10^{24} times stronger than gravity. This large difference is not explained by the SM and is considered *unnatural*. This is also called the *Fine Tuning Problem* since it requires that the parameters of a fundamental theory be *fine tuned*.

Although the SM does not explain the existence of these phenomena, there are extensions of the SM and more general theories which predict the existence of new particles and phenomena.

2.5 Beyond the Standard Model

This chapter now turns its focus to describing Beyond the Standard Model (BSM) physics. There are many BSM models, but this thesis will focus on just two models

which are relevant to a diphoton search. Specifically it describes the Two Higgs Doublet Model (2HDM), an extended Higgs sector extension of the SM, and the Randall-Sundrum (RS) graviton, a theory which predicts the existance of extra spatial dimensions and a spin-2 particle which couples to SM particles. These models predict new spin-0 and spin-2 particles that allow for diphoton final states. The search presented in this thesis searches for mass resonances consistent with those predicted by these new particles.

2.5.1 Extended Higgs Sector

The SM assumes the simplest possible scalar structure, just one $\mathrm{SU}(2)_L$ doublet. At tree level, the experimentally measurable parameter ρ, is given by

$$\rho = \frac{m_W^2}{\cos^2\theta_W m_Z^2}, \tag{2.33}$$

where θ_W is the Weinberg angle and m_Z and m_W are the masses of the Z and $W^\pm$ bosons respectively. This can be rewritten more generally for an $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge theory where there are n scalars multiplets ϕ_i in terms of weak isospin, T_i, weak hypercharge, Y_i, and the VEV's neutral component v_i, as (*Branco et al.*, 2012)

$$\rho = \frac{\sum_{i=1}^{n}\left\{T_i(T_i+1) - \frac{Y_i^2}{4}\right\}v_i}{\frac{1}{2}\sum_{i=1}^{n} Y_i^2 v_i} \tag{2.34}$$

From experiment it is known that $\rho \approx 1$. Both $\mathrm{SU}(2)_L$ singlets with $Y = 0$ and $\mathrm{SU}(2)_L$ doublets with $Y = \pm 1$ give $\rho = 1$, but this is not the only scalar structure which is compatible with this result.

The simplest possible extension to the scalar structure which is consistent with $\rho = 1$ is to add a second doublet of $\mathrm{SU}(2)_L$. This extension, known as the 2HDM, provides eight fields, three of which get *eaten* to give masses to the $W^\pm$ and Z bosons.

The remaining five fields are physical fields, two CP-even neutral scalars h and H, one CP-odd neutral pseudoscalar A, and two CP-even charged scalars $H^{\pm}$.

There are many variations of 2HDM which can be classified as follows:

- Type I (Fermiophobic): All quarks and charged leptons couple to Φ_2 only
- Type II (MSSM[§]-like): Up type quarks couple to Φ_2 while down-type quarks and charged leptons couple to Φ_1
- Type X (Lepton Specific): All quarks couple to Φ_2 while all charged lepton couple to Φ_1
- Type Y (Flipped): Up type quarks and charged leptons couple to Φ_2 while down type quarks couple to Φ_1

Ignoring Higgs self interactions, there are six free parameters in 2HDM, four are the Higgs masses, i.e. m_h, m_H, m_A, $m_{H^{\pm}}$. The other two free parameters are ratio of the two VEVs, $\tan\beta = v_u/v_d$, and the mixing angle, α, which diagonalizes the mass matrix of the neutral CP even Higgses. Two interesting special cases occur when $\sin(\beta-\alpha) \to 0$ and when $\cos(\beta-\alpha) \to 0$. In the former case h has exactly the same couplings as the SM Higgs. In the later case H has exactly the same couplings as the SM Higgs. 2HDM is not the only way to extend the Higgs sector, it is just the simplest.

There are numerous other models of an extended Higgs sector which include the prediction of new particles, many of which are electrically neutral scalars or pseudoscalars. Although the photon is massless and therefore does not couple to the Higgs directly, these electrically neutral scalars and psuedoscalars predicted may allow for diphoton final states through fermion and boson loops.

[§] Minimal Supersymmetric Standard Model (MSSM) is the simplest supersymmetery (SUSY) theory.

2.5.2 Extra Dimensions

2.5.2.1 Kaluza-Klein Theories

The first Kaluza–Klein (KK) theory was developed to unify the electromagnetic and gravitational fields as components of the same higher dimensional field. To achieve this a procedure called *toroidal compactification* is performed. An example of toroidal compactification is achieved by introducing an extra dimension x^5 and periodically identifying it as

$$x^5 \sim x^5 + 2\pi R. \tag{2.35}$$

The space obtained by this procedure can be written $M^4 \otimes S^1$ where M^4 is four-dimension Minkowski space and S^1 is the circle group. The space can be imagined as a five dimensional cylinder of radius R. A massless scalar field $\phi(x^\mu, x^5)$ in such a theory would have quantized momentum in the direction of x^5, that is

$$p^5 = \frac{n}{R}, \tag{2.36}$$

where $n \in \mathbb{Z}$. By expanding the field with a Fourier series $\phi(x^\mu, x^5)$ can be written

$$\phi(x^\mu, x^5) = \sum_{n=-\infty}^{\infty} \phi^n(x^\mu) e^{inx^5/R}, \tag{2.37}$$

and the equation of motion becomes

$$\partial_\mu \partial^\mu \phi^n(x^\mu) = \frac{n^2}{R^2} \phi^n(x^\mu). \tag{2.38}$$

In this way an infinite tower of four dimensional fields ϕ^n with masses $m^2 = n^2/R^2$ is generated (*Gabella*, 2006).

2.5.2.2 Randall-Sundrum Graviton

The Randall-Sundrum (RS) model assumes one extra dimension of space, y, with a warped spacetime metric bounded by two three-dimensional membranes or 3-branes. The 3-branes are separated in 5D space by a distance L in the y direction, and visualized in Figure 2.8. In the RS model SM fields are localized on one 3-brane, the

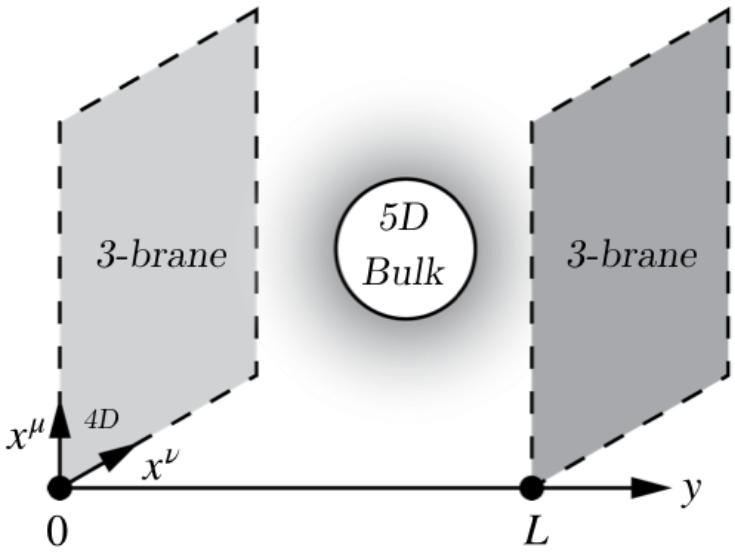


Figure 2.8: Visualization of 3-branes separated in 5D space.

TeV brane, and gravity on the other, the Planck brane[¶]. The extra dimension is not accessible to SM fields, but is to gravity through a massive spin-2 graviton. (*Randall and Sundrum*, 1999).

The warped spacetime metric is given by

$$ds^2 = e^{-2k|y|}\eta_{\mu\nu}dx^\mu dx^\nu - dy^2, \tag{2.39}$$

where k is a constant related to the curvature of the extra dimension. The warped spacetime is only warped in the 5^{th} dimension in such a way that the graviton's probability function drops exponentially across the 5^{th} dimension, leading to gravity

[¶] There are two RS models, RS1 and RS2. Described here is the RS1 model. In the RS2 model there is no TeV brane, and the SM particles are presumed to be on the Planck brane

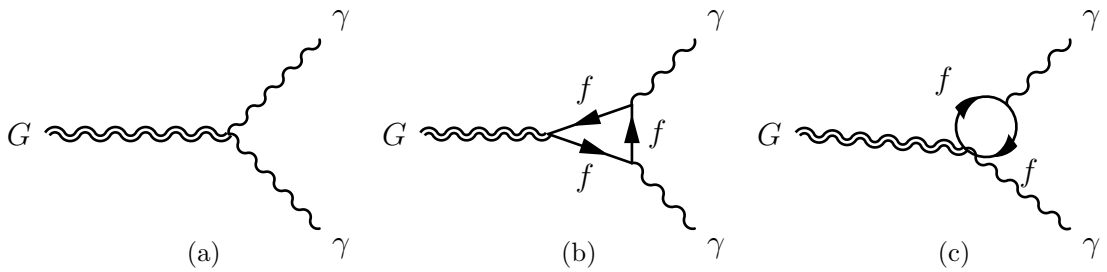


Figure 2.9: Feynman diagrams for the main processes for a graviton decaying to two photons.

being much weaker on the TeV brane than the Planck brane, offering a solution to the Hierarchy Problem..

The mass spectrum of the RS graviton created by the KK modes is given by

$$m_n = ke^{-kL}x_n, \tag{2.40}$$

where x_n are the roots of the Bessel function of order 1, that is $J_1(x_n) = 0$ (*Gabella*, 2006). Figure 2.9 shows the dominant Feynman diagrams for a graviton decaying to two photons.

The search presented in this thesis focuses on looking for RS and Extended Higgs processes which have diphoton final states. More details about the phenomenology and particular tunings of the models used are given in Section 5.4.

CHAPTER III

The Experimental Apparatus

This chapter describes the experimental apparatus used to collect the data used in this thesis. It begins by reviewing the basics of particle accelerators and particle colliders. From there it moves on to describing the LHC and its experiments, in particular the ATLAS detector and its calorimetry system.

3.1 Particle Accelerators

A particle accelerator is a machine that propels charged particles by accelerating them with EM fields. There are both electrostatic accelerators, which use a static electric field to accelerate particles, and electrodynamic accelerators, which use either magnetic induction or oscillating radio frequency to accelerate particles. There are many applications for particle accelerators including cathode ray tube (CRT) televisions, radiation therapy, and scattering experiments.

3.1.1 Particle Beams

Typically in accelerators not just one but many particles are accelerated in groups known as *bunches*. Each bunch contains many particles, for example at the LHC there are approximately 10^{11} particles in a bunch. These bunches can be approximated by Gaussian distributions with widths σ_x and σ_y in the transverse directions and σ_s in

the longitudinal direction. Generally, many bunches are accelerated in a continuous stream producing a *particle beam*.

3.1.1.1 Beam Emittance

Beam emittance, ϵ, is a measure of the spread, in position-momentum space, of the particle coordinates which make up a particle beam. Since measuring the full width is difficult, in practice the root mean square (RMS) width or the area containing a certain fraction of the particles in the beam is typically used. A beam with a small emittance has particles with nearly all the same momentum and are confined to a small space. A beam with a large emittance has particles spread in position space, momentum space, or both. It is often convenient to define the *normalized emittance*

$$\epsilon_n = \beta\gamma\epsilon \tag{3.1}$$

where β is the relativistic speed and γ is the Lorentz factor of the particle. Figure 3.1 shows a visual representation of beam particles in phase space.

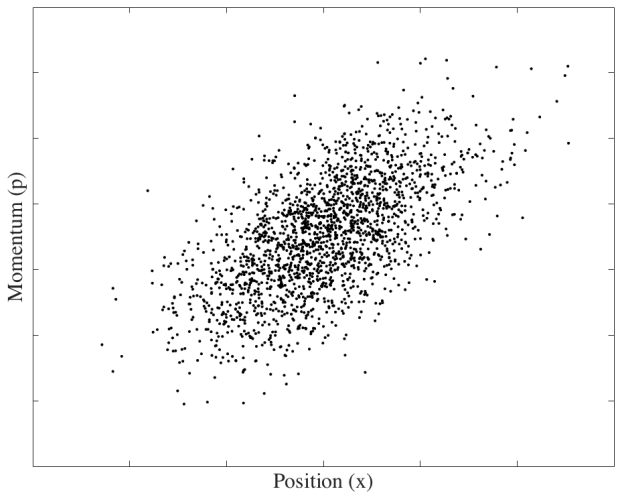


Figure 3.1: A two dimensional Normal Distribution representing beam particles in phase space. The horizontal axis represents position while the vertical axis represents momentum.

3.1.1.2 Beta Function

The *beta function* is a function that describes the transverse size of a particle beam as a function of s, the position along the beam trajectory. It is defined as

$$\beta(s) = \frac{\sigma^2(s)}{\epsilon}, \tag{3.2}$$

where $\sigma(s)$ is the transverse size of the particle beam and ϵ is the beam emittance. Small beta values correspond to a narrow beam while large beta values correspond to a wide beam, as seen in Figure 3.2. The quantity β^* is defined as the beta function at the interaction point. As seen in Figure 3.3, another way to interpret β^* is as the

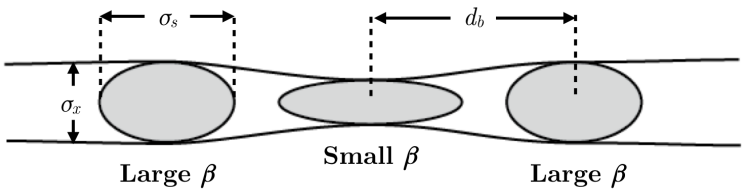


Figure 3.2: Cartoon of particle beam relating the lateral width, σ_x, and the longitudinal width, σ_s, to β where d_b is the distance between bunches in the beam.

distance from the interaction point to the closest point of the beam with twice the width.

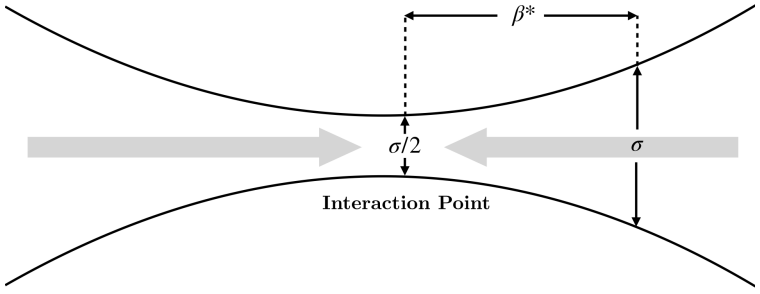


Figure 3.3: A cartoon of the beta function at the interaction point.

3.1.2 Scattering Experiments

The collision that occurs between an incident wave or particle and a target is known as *scattering*. A *scattering experiment* is an experiment in which waves or particles are scattered off of a target or targets and the outgoing waves or particles are measured.

The likelihood of a particular process occurring during a collision is represented by the *cross section*. The number of scatters N_{sc} for a particular process is related to the number of incident particles or waves N_{inc} and the cross section σ by the relation

$$N_{\text{sc}} = N_{\text{inc}} n_{\text{tar}} \sigma, \tag{3.3}$$

where n_{tar} is the number of targets per unit area.

3.2 Particle Colliders

Particle colliders are scattering experiments designed to accelerate and collide beams of particles into each other. They improved on fixed target experiments by increasing the COM energy for the same beam energy. That is, the available energy in a collider goes like beam energy $\sqrt{s} \sim E$ while the available energy in a fixed target experiment goes like the square root of beam energy $\sqrt{s} \sim \sqrt{Em}$.

There are both linear and circular particle colliders. Linear particle colliders are easier to build since they do not require the use of steering magnets to bend the trajectory of a particle, but they cannot reach the same high energies as circular colliders. Particles in circular colliders can reach much higher energies than linear colliders, however there is a limitation to the energy that a particle can acquire in a circular collider. The main limitiation is energy loss synchrotron radiation according

to

$$\frac{dE}{dt} \propto \frac{E^4}{m^4 R}, \tag{3.4}$$

where E and m are the particle's energy and mass respectively, and R is the radius of curvature. This makes it more difficult to maintain high energies for lighter particles than it is for heavier particles. For example, a proton is approximately 2000 times more massive than an electron, so an electron beam loses on the order of 10^{13} times more energy per unit time than a proton beam with the same energy and radius of curvature.

3.2.1 Luminosity

Luminosity is a measure of how much data has been taken in a scattering experiment. The integrated luminosity, L, can be expressed in terms of the *instantaneous luminosty*

$$L = \int_{\mathcal{T}} \mathcal{L} dt, \tag{3.5}$$

where $\mathcal{L}$ is the instantaneous luminosity and $\mathcal{T}$ is the data taking period. The number of events, N, produced from a specific process with cross section σ can be written $N = L\sigma$.

The instantaneous luminosity for two Gaussian beams colliding head on with width σ_x and σ_y containing n_b bunches each can be written

$$\mathcal{L} = \frac{f N_b^2 n_b}{4\pi\sigma_x\sigma_y}, \tag{3.6}$$

where f is the frequency of revolution and N_b is the number of particles in a bunch. However it is difficult to collide bunches head on, so in practice there typically is

some small non-zero crossing angle ϕ. Figure 3.4 shows a cartoon of colliding particle beams. This crossing angle reduces the luminosity by effectively reducing the number

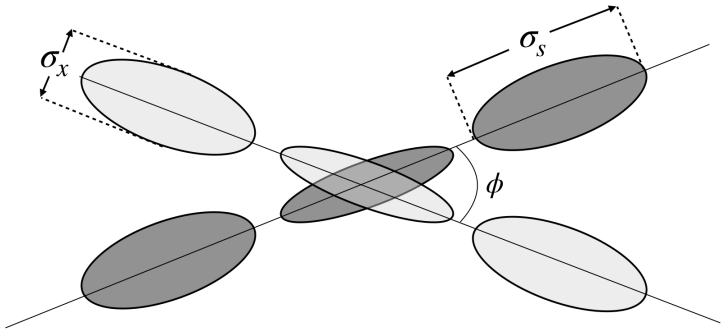


Figure 3.4: A cartoon of a particle beams colliding with crossing angle ϕ, lateral width σ_x, and longitudinal width σ_s.

of particles in a bunch, since not every particle in a bunch will *see* every particle in the bunch it is crossing. To correct for this, $\mathcal{F}$, known as the *geometric luminosity reduction factor*, can be computed

$$\mathcal{F} = \frac{1}{\sqrt{1 + \left(\frac{\sigma_x}{\sigma_s} \tan\frac{\phi}{2}\right)^2}\sqrt{1 + \left(\frac{\sigma_s}{\sigma_x} \tan\frac{\phi}{2}\right)^2}}, \tag{3.7}$$

where σ_s is the bunch length. For small ϕ and $\sigma_s \gg \sigma_{x,y}$, $\mathcal{F}$ becomes

$$\mathcal{F} = \frac{1}{\sqrt{1 + \left(\frac{\sigma_s}{\sigma_x}\frac{\phi}{2}\right)^2}}. \tag{3.8}$$

When this correction factor is applied the luminosity becomes

$$\mathcal{L} = \frac{f N_b^2 n_b}{4\pi\sigma_x\sigma_y}\mathcal{F}. \tag{3.9}$$

This can be rewritten entirely in terms of parameters of the collider

$$\mathcal{L} = \frac{\gamma f N_b^2 n_b}{4\pi\epsilon_n \beta^*}\mathcal{F} \tag{3.10}$$

where γ is the Lorentz factor, ϵ_n is the normalized beam emittance, and β^* is the beta function at the interaction point.

3.2.2 Parton Interactions

At high energies collisions do not occur directly between hadrons, but rather between the quarks and gluons that make them up. In this context these quarks and gluons are known as *partons*. Each parton carries just a fraction of the momentum of its parent particle. How much momentum is carried by a parton is given by its parton distribution function (PDF) and is dependent on the momentum transfer Q. An example of a PDF for a proton with $Q^2 = 10$ GeV^2 is shown in Figure 5.1.

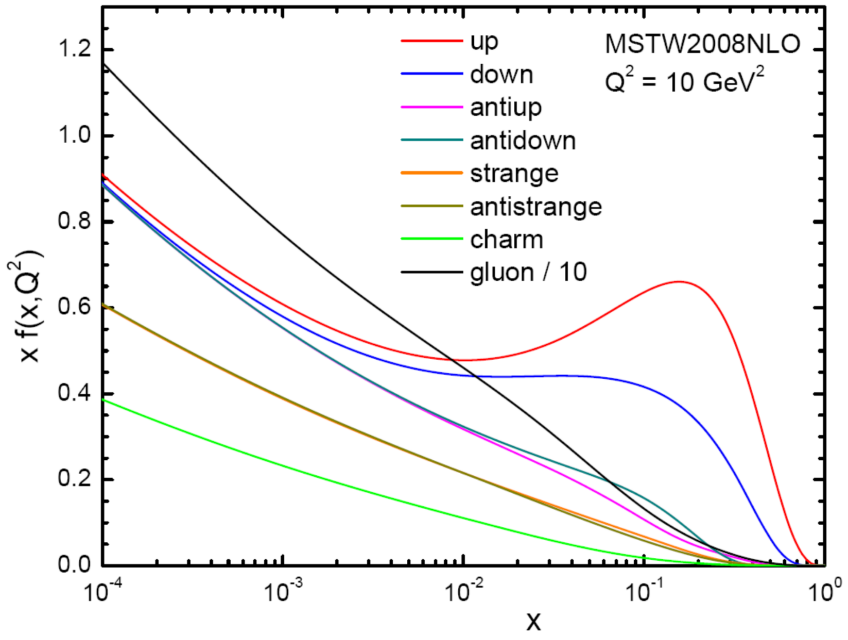


Figure 3.5: The MSTW2008NLO parton distribution function at $Q^2 = 10^2$ GeV, where the horizontal axis is x, the fraction of the hadron's momentum carried by a parton, and the y-axis is x times the probability density $f(x)$.

3.3 The Large Hadron Collider

The LHC is the world's largest particle accelerator sitting 100 meters underground in a 3.8 meter wide circular tunnel originally dug for the Large Electron-Positron Collider (LEP). LEP was the largest lepton collider ever constructed and was designed to make precision measurements and perform other analyses before constructing the LHC. It ran from 1989 to 2001 colliding electrons and positrons at COM energies reaching $\sqrt{s} = 209$ GeV. Around 2001 LEP was dismantled so that the LHC could be constructed in its place.

The LHC was designed to collide both protons and heavy ions with a COM energy on the TeV scale. Along the 27 kilometer circumference sits four experiments: ATLAS, A Large Ion Collider Exerpiment (ALICE), Compact Muon Solenoid (CMS), and LHC-beauty (LHCb). Each of the four experiments is specialized to help explore a different area of physics. ATLAS and CMS are general purpose detectors meant to make precision measurements of the SM and the Higgs boson, and to search for new physics. ALICE studies quark-gluon plasma, a state of matter where quarks and gluons behave almost like free particles. While LHCb is designed to study b quark physics and make precision measurements of CP violation.

3.3.1 Proton Injection Chain

Before reaching the ring of the LHC, protons are ionized then accelerated by a series of particle accelerators known as the *injection chain*. The chain begins with a small volume of hydrogen gas, approximately 8.28×10^{12} H_2 molecules*, which will be the source of the protons. Using an EM field, the electrons are stripped from the gas molecules separating the electrons and protons. In each step of the chain protons are accelerated to higher energies. At some steps, the protons are grouped into a finite

* This quantity of an ideal gas at atmospheric pressure and room temperature would occupy a cube with an edge length of approximately 69.1 microns.

series of bunches separated in space known as a *bunch train*. At some steps bunch trains are combined and grouped into larger bunch trains.

In the first step, protons are brought to an energy of 750 keV and grouped into six bunches by the Radio Frequency Quadrupole (RFQ2). These bunches are then accelerated to energies of 50 MeV after being injected into LINear ACcelerator (LINAC2) and accelerated over a distance of approximately 30 meters. Next they are transferred to the Proton Synchrotron Booster (PSB), a circular accelerator with a diameter of 50 meters, which increases their energies to 1.4 GeV. From there they are injected into another circular accelerator with a diameter of 100 meters, the Proton Synchrotron (PSyn), where they obtain energies of 25 GeV and are grouped into 72 bunches. In the final step before being injected into the ring of the LHC, these 72 bunches are injected into the Super Proton Synchrotron (SPS), a circular accelerator with a diameter of over 1 kilometer. The SPS combines three bunch trains from the PSyn, each containing 72 bunches, into 216 bunches which are accelerated to energies of 450 GeV. Twenty six of these larger bunch trains are injected into the LHC and combined producing two continuous beams of 2,808 proton bunches each. A cartoon illustrating this chain is shown in Figure 3.6.

3.3.2 Beam Pipes

The LHC collides two particle beams which each sit within a *beam pipe* which sit within a shared vacuum. The beams move in opposite directions and cross at four points around the ring in order to initiate particle collisions at the *interaction points* located within the LHC's four experiments. The pressure in the vacuum vessel is approximately 10^{-13} atm, a pressure comparable to that on the Moon's surface. This extreme vacuum is produced in order to remove any gas molecules that the beams' particles could collide with as they travel around the ring.

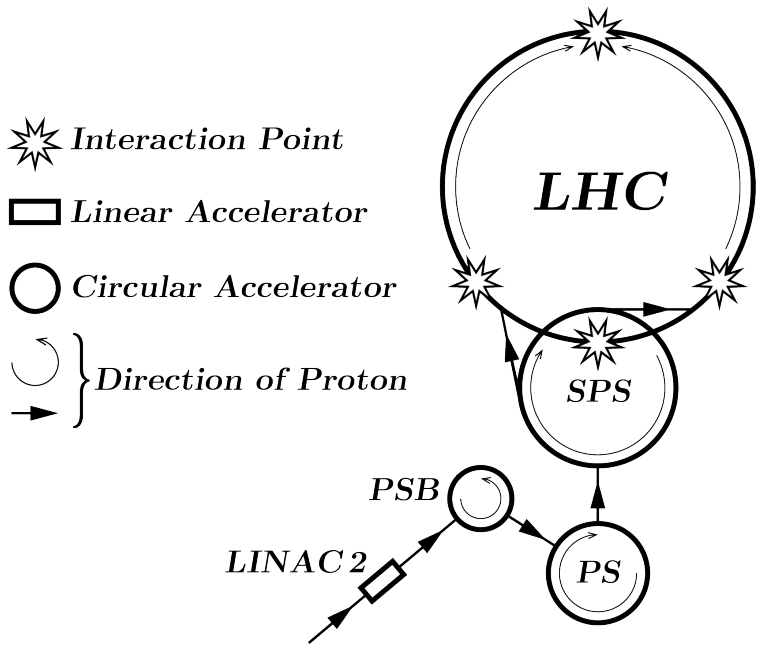


Figure 3.6: Cartoon schematic of the LHC's proton injection chain.

3.3.3 Superconducting Magnet System

The vacuum vessel contains approximately 10,000 liquid helium cooled superconducting magnets made of copper-clad niobium-titanium used to accelerate and focus the beams. Of the 10,000 magnets 1,232 are dipole magnets used to accelerate the protons from 450 GeV to 6.5 TeV as well as maintain those high energies, recal the charged particles which make up the beam lose energy by radiating photons according to Equation 3.4. 392 are quadrupole magnets used to focus the beams, while the remaining magnets are of higher multipole order and are used to correct imperfections in the EM field. The magnets are cooled with about 96 tonnes of superfluid helium-4 to keep them at their operating temperature of 1.9 K.

3.3.4 Proton Beam Conditions

The dipole magnets accelerate protons to 99.999999% of the speed of light, reaching energies of 6.5 TeV per proton and angular speeds leading to 11,245 rotations per second. This acceleration results in a Lorentz factor of $\gamma \approx 6,930$. In order to

acheive these high energies—and power its four experiments—the LHC uses roughly 120 MW of electric power from the French grid while running, which is about 20% of the total energy consumption of Geneva.

Each beam contains $n_b = 2{,}808$ bunches with each bunch containing $N_b = 1.15 \times 10^{11}$ protons. While travelling around the ring, the protons in each bunch are distributed in space according to a three dimensional Gaussian distribution with its transverse widths on the order of millimeters and its longitudinal width on the order of tens of centimeters. At the interaction points, however, the beam is focused with quadrapole magnets until it has a transverse width of about 16 microns. At the interaction points, the normalized transverse beam emittance is approximately $\epsilon_n = 3.75$ μm rad while the beta function is approximately $\beta^* = 0.55$ m.

The bunches are spaced along the beam pipe so that interactions occur ever 25 nanoseconds, leading to a bunch collision rate of 40 MHz. The beams cross at an extremely shallow angle of 285 μrad[†]. Using Equation 3.7, the aforementioned beam widths, and this crossing angle leads to a geometeric luminosity reduction factor of approximately 0.84. That is, the luminosity delivered is 84% of that which would be delivered if the beams were colliding head on.

3.3.4.1 Delivered *pp* Luminosity

Recall Equation 3.10 for the instantaneous luminosity

$$\mathcal{L} = \frac{\gamma f N_b^2 n_b}{4\pi\epsilon_n\beta^*}\mathcal{F}\,.$$

Using this equation, and the aforementioned parameters

[†] The crossing angle is approximately 0.016° in more familiar units.

$$\gamma = 6{,}930 \qquad f = 11{,}245 \text{ kHz}$$
$$N_b = 1.15 \times 10^{11} \qquad n_b = 2{,}808$$
$$\epsilon_n = 3.75 \ \mu\text{m rad} \qquad \beta^* = 0.55 \text{ m}$$
$$\mathcal{F} = 0.84$$

leads to an instantaneous luminosity on the order of $\mathcal{L} = \mathcal{O}\left(10^{34}\right)$ cm^{-2} s^{-1}

3.4 The ATLAS Detector

ATLAS is a general purpose detector at the LHC whose collaboration comprises more than 3000 physicists from 38 countries and 174 universities and laboratories. It has a forward-backward symmetric cylindrical geometry covering nearly 4π in solid angle. It is designed to reconstruct proton-proton collisions at extremely high energies, that is on the TeV scale. The detector, shown in Figure 3.7, has a diameter of 25 meters, a length of 44 meters, and weight of approximately 7000 tonnes.

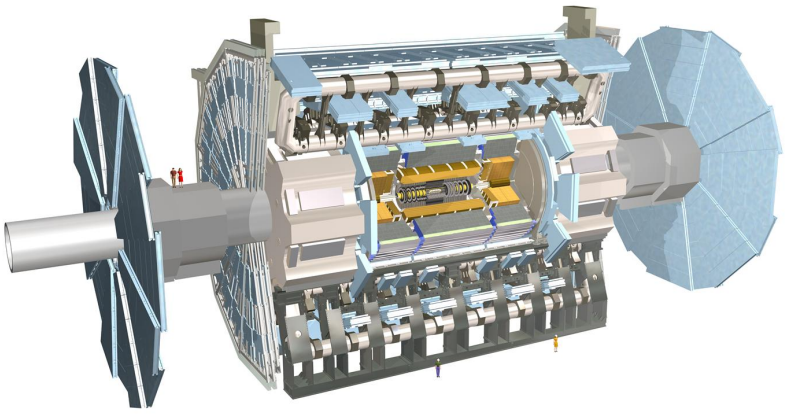


Figure 3.7: The ATLAS detector (*Pequenao*, 2008).

It uses a right hand coordinate system whose positive x-axis points to the center of the LHC ring, positive y-axis points upwards, and z-axis points along the beam

pipe. The azimuthal angle ϕ is measured from the positive x-axis in the x-y plane, and the polar angle θ is measured from the positive z-axis in the z-y plane.

A more convenient way of expressing the polar angle is with pseudorapidity

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] = \frac{1}{2}\ln\left(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z}\right) \tag{3.11}$$

where $\vec{p}$ is the three momentum and p_z is the longitudinal momentum of a moving particle. For very energetic objects the pseudorapidity converges to the rapidity

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) \tag{3.12}$$

where E is the energy of the object. The distance ΔR in (η, ϕ) space is given by

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{3.13}$$

where $\Delta\eta$ and $\Delta\phi$ are the angular separation in η and ϕ respectively.

The experiment is composed of an inner detector, solenoid magnet, toroid magnet, calorimetry system, and muon spectrometer. The inner layer tracks the path of moving charged particles. The calorimetery system is designed to measure the energy a particle loses as it passes through the detector. The muon spectrometer is the outer most layer and is designed to measure the path of muons. By combining the information gained from each detector, not only can a particle's four momentum be measured, but an identification on the type of particle can be made. Figure 3.8 shows a cross sectional view of the ATLAS detector and how various particles interact with it.

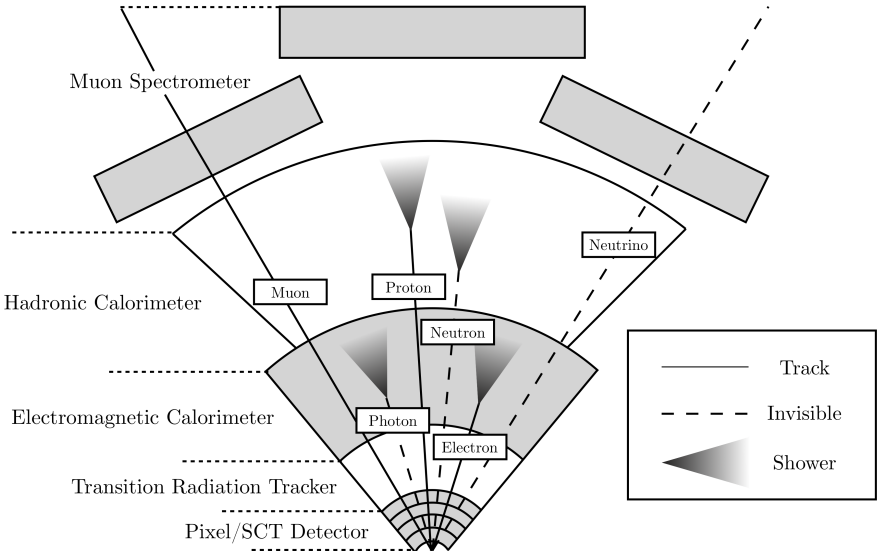


Figure 3.8: The detection of particles by the ATLAS detector.

3.4.1 Inner Detector

The ATLAS Inner Detector (ID) is designed to measure the position and momentum of charged particles. It comprises four main components: the Insertable B-Layer (IBL), Pixel Detector, Semi-Conductor Tracker (SCT), and Transition Radiation Tracker (TRT). The ID is surrounded by a 2 T solenoid magnet in a 7 meter long cylindrical enclosure with radius of 1.15 meters. It is designed to track in the region $-\pi \leq \phi \leq \pi$ and $|\eta| \leq 2.5$. By measuring the trajectory of a track left by an electrically charged particle one can determine it's path and point of origin, known as a *vertex*. Since the ID is in a roughly constant magnetic field pointing in the z direction, the transverse momentum of a particle can be determined by measuring its curvature in the R–ϕ plane where R is the distance from the beam pipe. A schematic of the ID is shown in Figure 3.9.

Figure 3.10 shows the amount of material, in units of radiation length X_0, traversed by a particle as it passes through the ID as a function of η. The radiation length varies for the ID from about $0.5X_0$ to $2.5X_0$. As a consequence approximately 40% of photons will convert to electron-positron pairs before reaching the calorimeters.

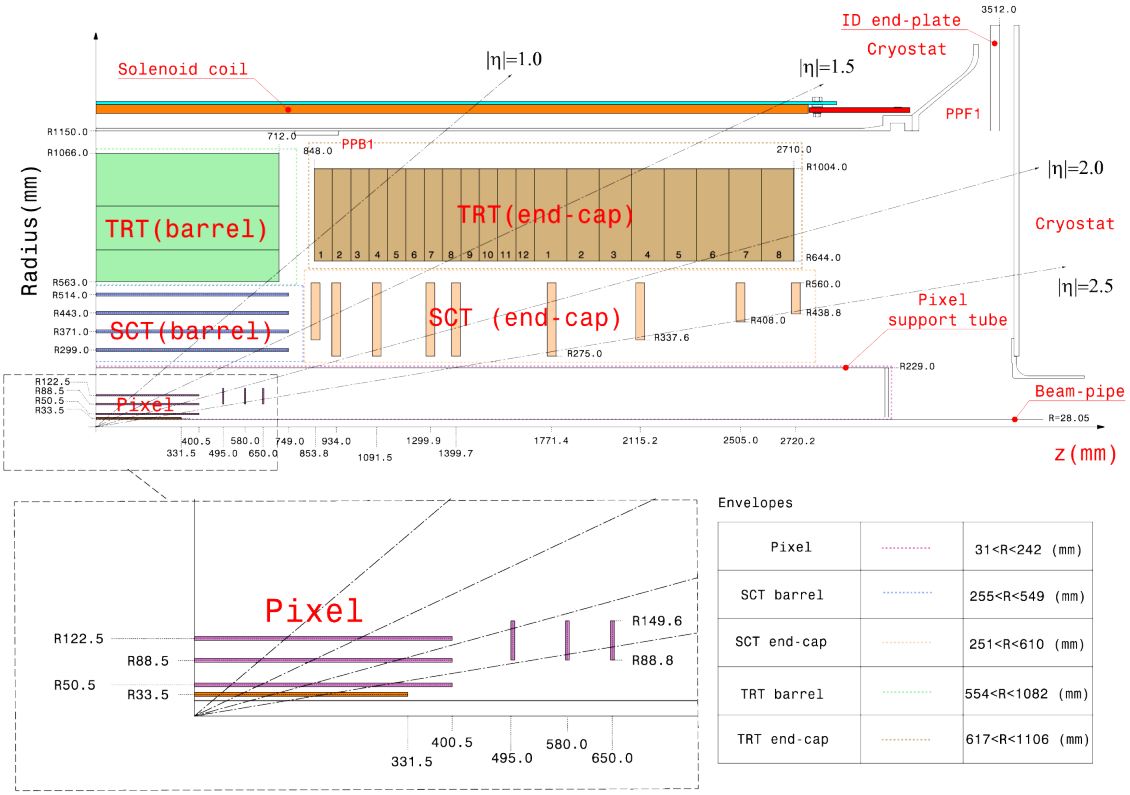


Figure 3.9: Schematic of the ATLAS inner detector (*Abdelouahab et al.*, 2008).

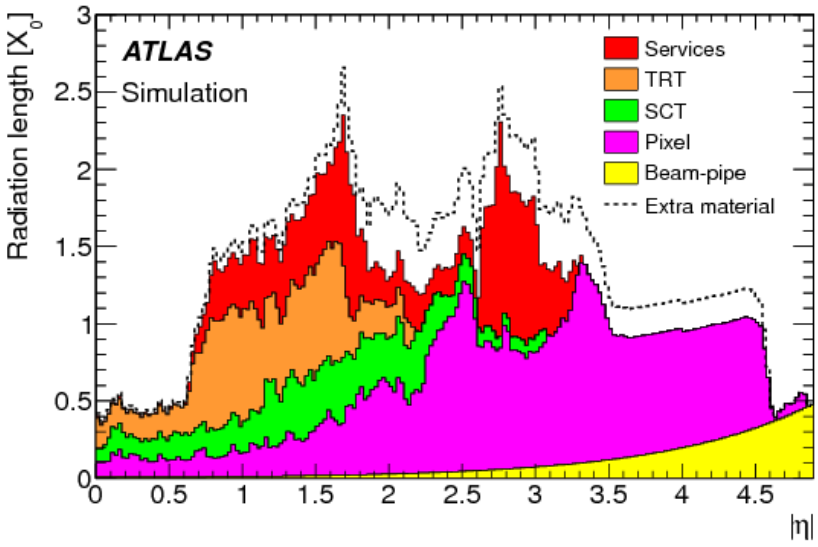


Figure 3.10: The amount of material, in units of radiation length X_0, traversed by a particle as it passes through the ID as a function of η (*ATLAS Collaboration*, 2011)

3.4.1.1 Insertable B-Layer

The Insertable B-Layer (IBL) is the inner most layer of the ID. It was installed during the LHC 2013 shut down to deal with the high radiation and occupancy due to increasing the instantanenous luminosity. It is cylindrial in shape and sits concentrically about the beam pipe with an inner radius of 31 mm and an outer radius of 40 mm. It uses two different silicon sensor technologies with a pixel size of 50 μm $\times$ 250 μm and resolution of 8 μm $\times$ 40 μm.

3.4.1.2 Pixel Detector

The next inner most layer of the ID is the Pixel Detector (PD/PIXEL). It comprises a *barrel* region and two *endcap* regions. The barrel is composed of four cylindrical layers concentrically placed about the beam pipe, while the endcaps entail three disk layers each and sit on either side of the barrel. These layers contain a total of 1744 pixel modules with dimensions 19 mm $\times$ 63 mm each. The sensor area of the pixel modules are made up of 250 μm thick oxygenated n-type silicon wafers and contain 47,232 pixels. Each pixel has a nominal pixel size of 50 μm $\times$ 400 μm and a spacial resolution of 10 μm $\times$ 115 μm. This high precision measurement allows for the reconstruction of displaced vertices from particles such as b-quarks.

3.4.1.3 Semi-Conductor Tracker

Surrounding the Pixel Detector is the Semi-Conductor Tracker (SCT). Like the PD it comprises a barrel region and two endcap regions. The barrel is composed of four cylindrical layers while the endcaps are composed of nine disk layers each. The barrel layers contain 2,112 semiconductor modules while the endcap layers contain 988 each. Each module consists of four silicon strip sensors at a constant pitch of 80 μm. The strip sensors are paired in groups of two to form 768 strips each approximately 12 cm in length. The modules are arranged such that a charged particle originating

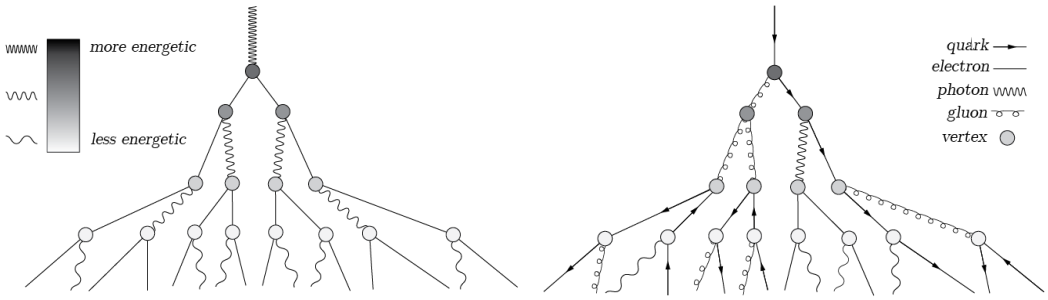
from the beam spot will pass through at least four layers of SCT modules providing a point space resolution of 17 μm $\times$ 580 μm.

3.4.1.4 Transition Radiation Tracker

The final layer of the ID is the Transition Radiation Tracker (TRT). It consists of a barrel region and two endcap regions made of 4 mm diameter polyamide tubes filled with a mixture of Ar, CO_2, and O_2 gas. The barrel region contains 50,000 longitudinally arranged tubes of length 144 cm, while the endcaps contain 320,000 radially arranged tubes with length 32 cm. Running down the axis of each tube is a 31 μm diameter gold plated tungsten wire. The tube wall is held at a voltage of -1.5 kV while the wire is held at ground. When a charged particle passes through the TRT the gas inside is ionized and freed electrons drift to the wire. The drift time is proportional to the Lorentz factor, γ, of the particle and is measured to provide a spatial hit resolution of 130 μm in the plane perpendicular to the wire.

3.4.2 Calorimeters

Next after the ID covering a full 2π azimuthal angle and $|\eta| < 4.9$ is the calorimetry system. It comprises the Electromagnetic Calorimeter (ECal), designed to measure the energy of particles that interact via the EM force, and the Hadronic Calorimeter (HCal), designed to measure the energy of particles that interact via the strong force. The calorimeters are known as *sampling calorimeters* and are composed of alternating layers of *passive material* and *active material*. As particles pass through the calorimeter they interact with the passive material producing lower energy particles. These lower energy particles also interact with the passive material and produce even lower energy particles. This process continues with each new particle producing many lower energy particles until all the energy of the original particle is exhausted. This cascade of particles is known as a *shower*, and is illustrated in Figure 3.11. As the



(a) Electromagnetic Shower (b) Hadronic Shower

Figure 3.11: Cartoon of particle showers.

particles pass through the active material energy is collected via ionization (ECal) or scintillation (HCal). The energy resolution of calorimeters is given by

$$\frac{\sigma(E_0)}{E_0} = \frac{a}{\sqrt{E_0}} \oplus \frac{b}{E_0} \oplus c \tag{3.14}$$

where a is the *sampling term* used to account for stoichastic uncertainty, b is the *electronic noise term* measured in calibration runs, and c is the *constant term* which dominates at high energy. Table 3.1 shows the subsections of the calorimeter and their angular coverage in η.

3.4.2.1 Electromagnetic Calorimeter

The first layer of the calorimeters is the Electromagnetic Calorimeter (ECal). It is the layer of the detector in which photons are primarily measured. It is cylindrical in shape with a length of 6.65 m and radius of 2.25 m. It covers the region $-\pi \leq \phi \leq \pi$ and $|\eta| < 3.2$ and is divided into three subsections: the Electromagnetic Barrel Calormeter (EMB), the Electromagnetic Endcap Calorimeter (EMEC), and the first section of the LAr Forward Calorimeter (FCal). The EMB covers $|\eta| < 1.475$, the EMEC covers $1.375 < |\eta| < 3.2$, while the FCal covers $3.1 < |\eta| < 4.9$. Liquid Argon (LAr) is the active material and lead and copper serve as the passive materials.

The EMB is 6.4 m long with an inner diameter of 2.8 m and outer diameter of

Calorimeter	**Coverage**	**Granularity** ($\Delta\eta \times \Delta\phi$)
ECal		
Presampler	$\lvert\eta\rvert < 1.54$	0.25×0.1
	$1.5 < \lvert\eta\rvert < 1.8$	0.025×0.1
Sampling 1	$\lvert\eta\rvert < 1.4$	0.003×0.1
	$1.4 < \lvert\eta\rvert < 1.475$ (barrel)	0.025×0.025
	$1.375 < \lvert\eta\rvert < 2.5$ (endcap)	0.003 - 0.025×0.1
	$2.5 < \lvert\eta\rvert < 3.2$	0.1×0.1
Sampling 2	$\lvert\eta\rvert < 1.4$	0.075×0.025
	$1.4 < \lvert\eta\rvert < 2.5$	0.025×0.025
	$2.5 < \lvert\eta\rvert < 3.2$	0.1×0.1
TileCal		
Sampling 1-2	$\lvert\eta\rvert < 1.0$ (barrel)	0.1×0.1
	$0.8 < \lvert\eta\rvert < 1.7$ (extended barrel)	0.1×0.1
Sampling 3	$\lvert\eta\rvert < 1.0$ (barrel)	0.2×0.1
	$0.8 < \lvert\eta\rvert < 1.7$ (extended barrel)	0.2×0.1
HEC		
Sampling 1-4	$1.5 < \lvert\eta\rvert < 2.5$	0.1×0.1
	$2.5 < \lvert\eta\rvert < 3.2$	0.2×0.2
FCal		
Sampling 1 (ECal)	$3.1 < \lvert\eta\rvert < 4.9$	$\approx 0.1 \times 0.1$
Sampling 2-3 (HCal)	$3.1 < \lvert\eta\rvert < 4.9$	0.2×0.2

Table 3.1: The coverage and granularity of the calorimeters

4 m. It is divided into two barrels separated by 4 mm at $z = 0$. It comprises 2,048 LAr absorbers and lead samplers. The first layer is the presampler (PS), a single thin layer of argon with no lead absorber meant to correct for energy loss in the ID. The PS is followed by three sampling layers each with progressively larger granularity. Figure 3.12 shows the layers of the EMB.

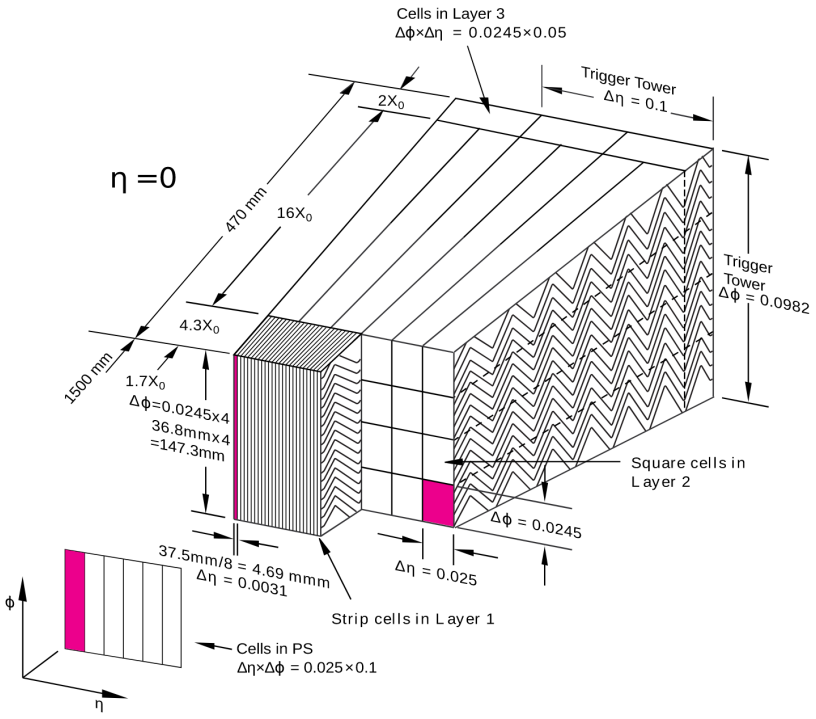


Figure 3.12: The layers of the Electromagnetic Barrel Calorimeter (*Aaboud et al.*, 2019).

The EMEC comprises two wheels on either side of the EMB. Like the EMB it uses LAr absorbers and lead samplers. Where the EMB and EMEC meet is a region of poor resolution known as the *crack region*. The first section of the FCal, known as FCal1, uses LAr absorbers with parallel plates of copper samplers.

The energy resolution of the ECal is:

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus b, \tag{3.15}$$

where $a = \mathcal{O}(10\%)$ and $b = 0.7\%$. Over η the ECal depth is approximately constant

at about 25 radiation lengths.

3.4.2.2 Hadronic Calorimeter

The Hadronic Calorimeter (HCal) surrounds the ECal and sits inside a cylinder with length of 6.1 m and diameter of 8.5 m. It is divided into three sections: the Tile Barrel Hadronic Calorimeter (TileCal), the Hadronic Endcap (HEC), and the final sections of the FCal: FCal2 and FCal3.

The TileCal is divided into the barrel ($|\eta| < 1.0$) and the extended barrel ($0.8 < |\eta| < 1.7$) and comprises steel absorbers and polystyrene scintillating tile samplers. The barrel is 5.8 m long with an inner diameter of 4.56 m and an outer diameter of 8.5 m and the extended barrel is 2.6 m in length with the same inner and outer diameters as the barrel. The regions are divided into 64 modules which are composed of three layers each. After the third layer there are 9,825 photomultiplier tubes (PMT) which amplify the scintillator signal and convert it into an electrical signal. Figure 3.13 shows the layers of the TileCal.

Each endcap of the HEC comprises two wheels located directly behind the EMEC. The wheels were built using 32 wedge shaped modules composed of copper plates to act as a passive material and LAr to act as an active material. The FCal2 and FCal3 detectors are built using tugnsten as the passive material and LAr as the active material.

3.4.2.3 Muon Spectrometer

The final layer of the ATLAS detector is the Muon Spectrometer. It comprises four types of detectors: the Monitored Drift Tubes (MDT), the Resistive Plate Chambers (RPC), the Thin Gap Chambers (TGC) and the Cathode Strip Chambers (CSC). These detectors and their relative orientation can be seen in Figure 3.14. The MDTs are meant to provide precision measurements while the RPC, TGC, and CSC are

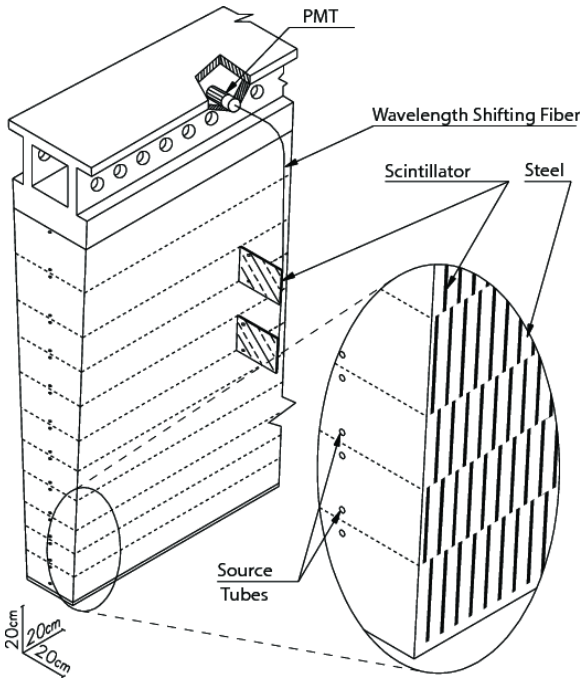


Figure 3.13: The layers of the Tile Barrel Hadronic Calorimeter (*Sotto-Maior Peralva*, 2013).

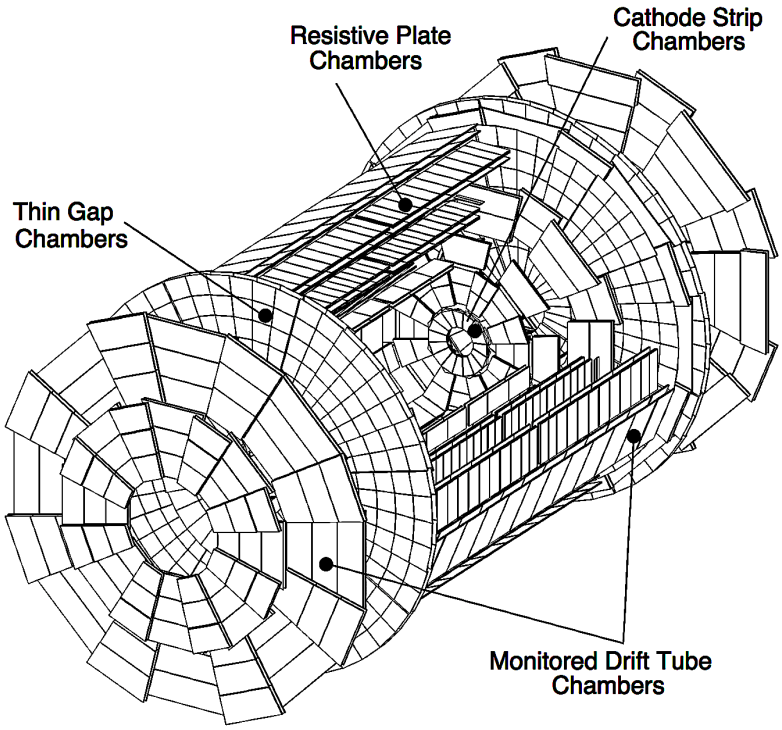


Figure 3.14: The detectors of the Muon Spectrometer (*Camarri et al.*, 1998).

used for triggering on muons and to complement the MDT measurements.

In the barrel region ($|\eta| < 1.7$) there is a combination of MDTs and RPCs arranged in three concentric rings of approximate radii 5 m, 7.7 m, and 10 m. The MDT chambers contain aluminum tubes of 30 mm diameter filled with Ar-CO_2 gas. Along the axis of each tube is a 50 μm gold plated tungsten-rhenium wire is held at a potential of 3 kV while the aluminum wall of the tube is held at ground. As a muon passes through the tube the gas inside is ionized and creates electrons which drift towards the wire. From the *drift time*, that is the time it takes for an electron to drift to the wire, the muon's point of closest approach to the wire can be determined. This provides a spacial resolution of 80 μm and time resolution of less than 1 ns.

The RPCs consist of parallel phenolic-melaminic plastic laminate electrode plates. The plates are separated by 2 mm insulating spacers and the space is filled with a $C_2H_2F_4$ gas mixture. The plates are held at a potential difference of 9.8 kV. Due to the electric field between the plates, as a muon passes through the plates an electron avalanche is produced. The avalanche is read as an electrical signal providing a time resolution of less than 2 ns.

In the endcap regions there is a combination of MDTs, the TGCs, and the CSCs arranged on eight wheels at a distance of $\pm$7.4 m, $\pm$10.8 m, $\pm$14 m, and $\pm$21.5 m from $z = 0$. The first layer is the CSC, which comprises multiwire proportional chambers whose wires run in the radial direction. The wires are held at a voltage of 1.9 kV and provide tracking resolution of 60 μm and timing resolution of less than 40 ns. The next layer is the TGCs which are also multiwire proportional chambers filled with a mixture of CO_2 and n-C_5H_{12} gas. The TGCs have a wire to wire distance of 1.8 mm and wire to cathode distance of 1.4 mm. The wires are held at a voltage of 2.9 kV and provide a time resolution of 4 ns. The final layer is MDTs.

CHAPTER IV

Measurement of photons in the ATLAS Detector

This chapter focuses on how photons are measured with the ATLAS detector. It beings by reviewing the basics of photon-matter interactions. From there it moves to describing how photons are reconstructed and identified in the ATLAS detector. Finally, the chapter ends by describing how the photons are calibrated for use in analyses.

4.1 Photon-Matter Interaction

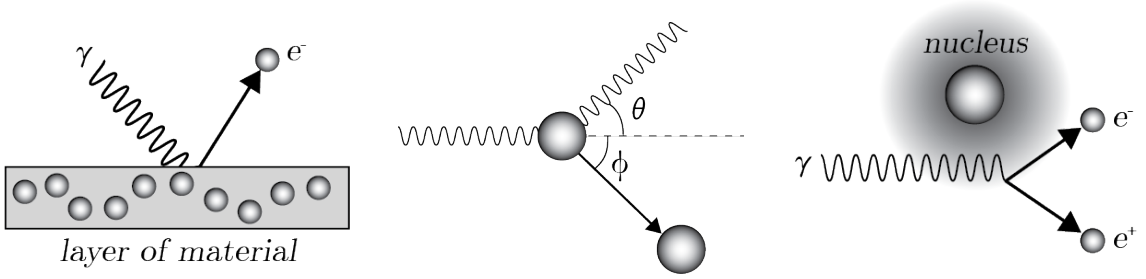
As photons pass through the material of the calorimeter they lose energy due to EM interactions. The intensity of a photon beam as it emerges from a layer of material of thickness x is given by

$$I(x) = I_0 e^{-\mu x} \tag{4.1}$$

where I_0 is the intensity of the photon beam entering the material and μ is the *absorption coefficient* which is related to the absorption cross section σ by

$$\mu = \frac{\sigma N_A \rho}{A} \tag{4.2}$$

where N_A is Avogadro's number, and ρ and A are the density and atomic mass of the material respectively. The primary mechanisms for energy loss, illustrated in Figure 4.1, are the photoelectric effect, Compton scattering, and electron positron pair production. At high energies the cross sections of these three mechanisms are



(a) Photoelectric Effect (b) Compton Scattering (c) Pair Production

Figure 4.1: Cartoons of the main processes responsible for photon losing energy as it passes through matter.

given by

$$\sigma_{\text{pe}} \approx 4\pi r_e^2 \alpha^4 Z^5 \frac{m_e c^2}{E_\gamma} \tag{4.3}$$

$$\sigma_{\text{C}} \approx \pi r_e^2 Z \frac{m_e c^2}{E_\gamma} \ln\left(\frac{2E_\gamma}{m_e c^2}\right) \tag{4.4}$$

$$\sigma_{\text{pp}} \approx 4\alpha r_e^2 Z^2 \left(\frac{7}{9}\log\frac{183}{Z^{1/3}}\right) \tag{4.5}$$

$$\approx \frac{7}{9}\frac{A}{X_0 N_A} \tag{4.6}$$

respectively, where r_e is the classical electron radius, α is the fine structure constant, Z is the atomic number of the nucleus with which the photon is interacting within the material, m_e is the mass of the electron, c is the speed of light, E_γ is the energy of the photon, A is the atomic mass, and N_A is Avogadro's number. For energies less than 100 keV the photoelectric effect dominates as the main source of energy loss while pair production dominates for energies greater than 10 MeV. The photoelectric effect is particularly relevant for detecting photons in a high mass resonance search with ATLAS since all photons considered in the analysis are high energy.

In ATLAS photons produced at the primary vertex can either reach the ECal unconverted or convert to an electron positron pair in the ID. At low $|\eta|$ approximately 20% of photons convert in the ID while at $|\eta| \approx 2.3$ approximately 65% convert in the ID (*Aad et al.*, 2019).

4.2 Photon Reconstruction

Photon reconstruction and electron reconstruction are closely related. Both photons and electrons produce EM showers in the ECal, but since photons do not carry electric charge, they do not interact with the ID and so do not produce tracks. Electrons are identified by an energy deposit in the ECal that is associated with a track. Photons are most simply identified by an energy deposit in the ECal that is not associated with a track. However, this is complicated since a large number of photons in ATLAS will convert to an electron positron pair while still in the ID. These photons are known as *converted photons* and although they do not produce tracks, their decay products, the electron and positron, do. The tracks point back to a secondary vertex known as the *conversion vertex*. Converted photons are identified as energy deposits in the ECal that are associated with a conversion vertex. This section explains this process in further detail focusing on photon reconstruction.

4.2.1 Topo-Cluster Reconstruction

Photon reconstruction begins with *topo-cluster reconstruction* which begins by forming *proto-clusters* in the ECal and HCal using a set of noise thresholds in which the cell initiating the cluster is required to have significance $|\zeta_{\text{cell}}^{\text{EM}}| \geq 4$, where

$$\zeta_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}, \tag{4.7}$$

$E_{\text{cell}}^{\text{EM}}$ is the cell energy at the EM scale and $\sigma_{\text{noise,cell}}^{\text{EM}}$ is the expected cell noise. The expected cell noise includes the known electronic noise and an estimate of the pile-up noise corresponding to the expected average instantaneous luminosity. To suppress the formation of noise clusters, cells from the presampler and the first ECal layer are excluded from initiating proto-clusters. Next, each neighbor cell passing the threshold of $|\zeta_{\text{cell}}^{\text{EM}}| \geq 2$ becomes a seed cell in the next iteration, where it collects its own neighbors in the proto-cluster. In the case that two proto-clusters contain the same cell with $|\zeta_{\text{cell}}^{\text{EM}}| \geq 2$ above the noise cell threshold, the proto-clusters are merged. A crown of nearest neighbor cells is added to the cluster independent of their energy. This set of thresholds is commonly known as 4-2-0 topo-cluster reconstruction. A cell is considered a local maximum when it has $E_{\text{cell}}^{\text{EM}} > 500$ MeV, at least four neighbors, and when none of the cell neighbors has a larger signal. Proto-clusters with two or more local maxima are split into separate clusters.

Photon reconstruction begins with the constructions of *topo-clusters* but only uses the energy from cells in the ECal, except in the transition region of $1.37 < |\eta| < 1.63$, where the energy measured in the presampler and the scintillator between the calorimeter cryostats is also added. This is referred to as the EM energy of the cluster, and the EM fraction f_{EM} is the ratio of the EM energy to the total cluster energy. Only clusters with EM energy greater than 400 MeV are considered, and are referred to as *EM topo-clusters*.

4.2.2 Photon Conversion Reconstruction

Photon conversion vertices use tracks loosely matched to fixed sized clusters as inputs. Both *Si tracks*, tracks with silicon hits, and *TRT tracks*, tracks reconstructed only in the TRT, are used. Two track conversion vertices are reconstructed from two opposite-charge tracks forming a vertex consistent with that of a massless particle. The TRT must determine that the tracks have a high probability of being electron

tracks in order to be used to increase converted photon purity. Next, the conversion vertices are matched to the EM topo-clusters by minimizing the angular separation between the conversion vertex and the EM topo-cluster. If there are multiple conversion vertices matched to a cluster, double-track conversions with two Si tracks are preferred over the double-track conversions without at least one TRT track, followed by single-track conversions. Within each category, the vertex with the smallest conversion radius is preferred.

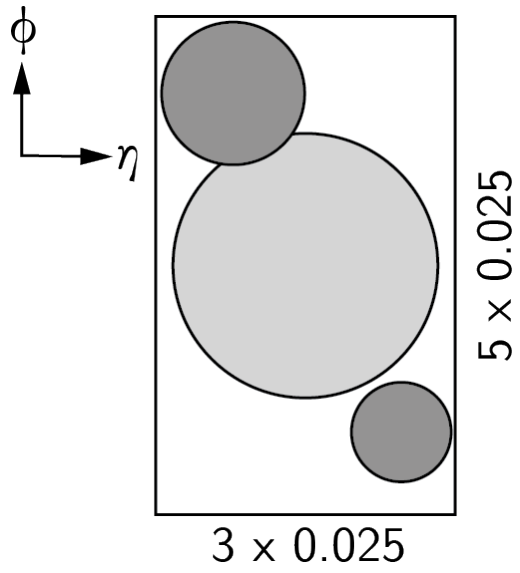
4.2.3 Supercluster Reconstruction

The reconstruction of photon superclusters happens in two stages. In the first stage, EM topo-clusters are tested for use as seed cluster candidates, which form the basis of *superclusters*. In the second stage EM topo-clusters near the seed candidates are identified as satellite cluster candidates and added to the seed cluster to form the final superclusters. These satellite clusters are included in an attempt to capture the energy from bremsstrulung radiation or topo-cluster splitting.

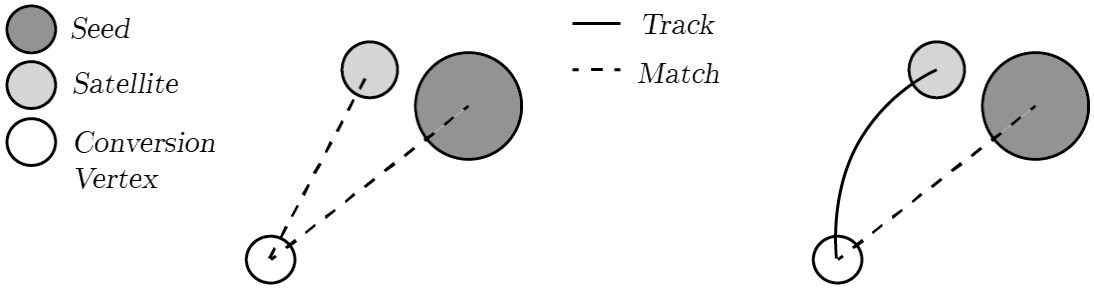
The EM topo-clusters are sorted according to descending E_{T} calculated using the EM energy, and tested one by one. A cluster must have E_{T} greater than 1.5 GeV to qualify as a supercluster seed, and cannot be used as a seed cluster if it has already been added as a satellite cluster to another seed cluster. A cluster is considered a satellite if it falls within a window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ around the seed cluster barycenter*. Photons with conversion vertices made up only of tracks containing silicon hits have a cluster added as a satellite if its best matched track belongs to the matched conversion vertex. The seed clusters with their associated satellite clusters are called superclusters. Figure 4.2 shows a cartoon illustration of the supercluster reconstruction process.

Finally, the supercluster building algorithm is used to assign calorimeter cells to

* Barycenter here refers to the average location in (η, ϕ) space weighted by the energy distribution of the seed cluster.



(a) Add **all clusters within 3 × 5 window** around seed cluster.



(b) Add topo-clusters that have **the same conversion vertex** matched as the seed cluster.

(c) Add topo-clusters with **a track match** that is **part of the conversion vertex** matched to the seed cluster.

Figure 4.2: A cartoon illustration of the construction of a supercluster.

a given supercluster. In most regions only cells from the presampler and the first three LAr calorimeter layers are considered. However, in the transition region of $1.4 < |\eta| < 1.6$, the energy measured in the scintillator between the calorimeter cryostats is also used. The size of each constituent topo-cluster is restricted to a maximal width of 0.075 or 0.125 in the η direction in the barrel or endcap region, respectively, this reduces the superclusters' sensitivity to noise. No restriction is applied in the ϕ direction since interactions between the photon and detector material tend to cause the EM shower to spread in the ϕ direction.

4.2.4 Photon Identification

The superclusters are matched to conversion vertices in the same manner as the EM topo-clusters. The matched or unmatched superclusters are now identified as photons and/or electrons. To identify an object as a photon, discriminating variables are constructed using information from the calorimeters about the EM shower. These variables are divided into three groups: variables involving the first layer of the ECal, variables involving the second layer of the ECal, and variables involving the HCal. The following are descriptions of the discriminating variables and their cartoon representations.

ECal First Layer

- ($w_{s\,\mathrm{tot}}$) Total lateral shower width, $\sqrt{(\Sigma E_i(i - i_{\mathrm{max}})^2)/(\Sigma E_i)}$ where i runs over all cells in a window $\Delta\eta \approx 0.0625$ and i_{max} is the index of the highest energy cell

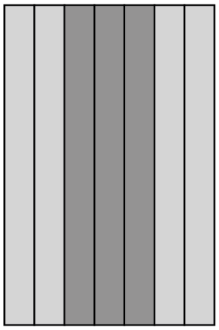
- ($w_{s\,3}$) Later shower width, $\sqrt{(\Sigma E_i(i - i_{\mathrm{max}})^2)/(\Sigma E_i)}$, where i runs over all cells in a window of 3 cells around the highest energy cell

- (f_{side}) Energy fraction outside core of three central cells, within seven cells

- (ΔE_s) Difference between the energy of the cell associated with the second maximum, and the energy reconstructed in the cell with the smallest value found between the first and second maxima

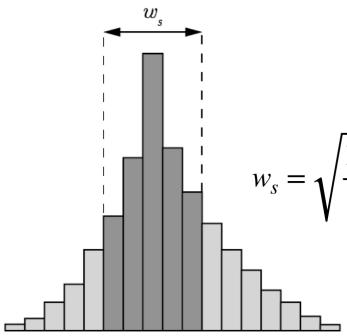
- (E_{ratio}) Ratio of the energy difference between the maximum energy deposit and the energy deposit in a secondary maximum in the cluster to the sum of these energies

- (f_1) Ratio of the energy measured in the first layer of the electromagnetic calorimeter to the total energy of the EM cluster



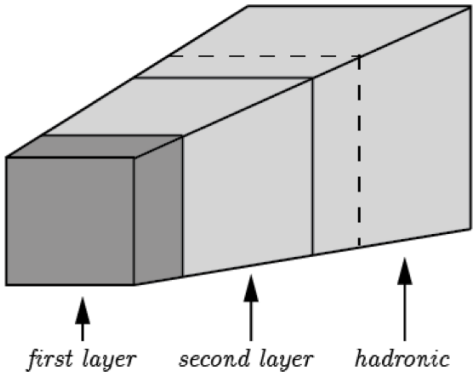
$$f_{\text{side}} = \frac{E_7^{S_1} - E_3^{S_1}}{E_3^{S_1}}$$

(a) The quantity f_{side} is measure of how spread the energy in the first layer of the ECal. The energy in the light gray area represents the numerator while the energy in the dark gray area represents the denominator



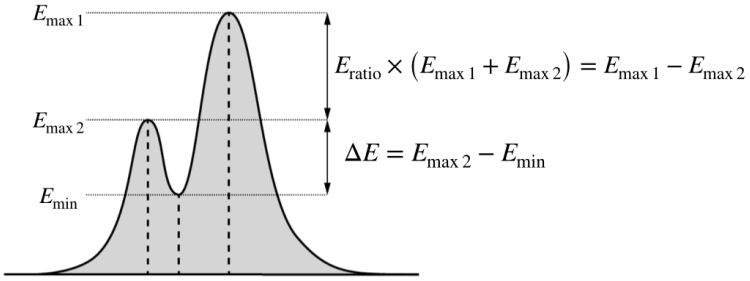
$$w_s = \sqrt{\frac{\Sigma E_i (i - i_{\text{max}})^2}{\Sigma E_i}}$$

(b) The dark gray region represents a measure of the width of the energy deposited in either 3×2 strips ($w_{s\,3}$) or 20×2 strips ($w_{s\,\text{tot}}$)



$$f_1 = \frac{E_{\text{T}}^{S_1}}{E_{\text{T}}}$$

(c) The transverse energy in the dark gray volume represents the numerator and the transverse energy in the light gray and dark gray volumes represents the denominator



(d) The vertical distance between the second local maximum and the local minimum represent ΔE. The ratio E_{ratio} is a measure of the vertical distance between the first and second local maxima relative to their values.

Figure 4.3: Discriminating variables for photon identification involving the first layer of the ECal

ECal Second Layer

- (R_η) Ratio of the sum of the energies of the cells contained in a 3×7 rectangle in (η, ϕ) space, measured in cell units, to the sum of the cell energies in a 7×7 rectangle, both centered around the most energetic cell

- (w_{η_2}) Later shower width where $\sqrt{(\Sigma E_i \eta_i^2)/(\Sigma E_i) - ((\Sigma E_i \eta_i)/(\Sigma E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells

- (R_ϕ) Ratio of the sum of the energies of the cells contained in a 3×3 rectanglein $\eta \times \phi$ space, measured in cell units, to the sum of the cell energies in a 3×7 rectangle, both centered around the most energetic cell

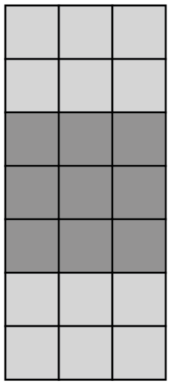
Hadronic Leakage

- ($R_{\text{had}\,1}$) Ratio of E_{T} in the first layer of the hadronic calorimeter to E_{T} of the EM cluster (used over the ranges $|\eta| < 0.8$ and $|\eta| > 1.37$).

- (R_{had}) Ratio of E_{T} in the hadronic calorimeter to E_{T} of the EM cluster (used over range $0.8 < |\eta| < 1.37$).

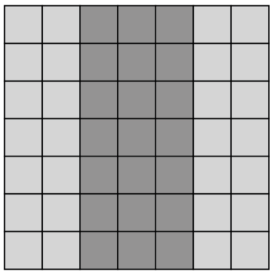
$$* \quad * \quad *$$

Photons and electrons can be identified as either *Loose*, *Medium*, or *Tight*, where Tight represents the most confidence in the identification, while Loose represents the least confidence in the identification. A given supercluster can be identified as both a photon and an electron since they are built independently. In such cases the procedure outlined in Figure 4.6 is applied. The photon or electron will now go through a calibration process before being used in analyses.



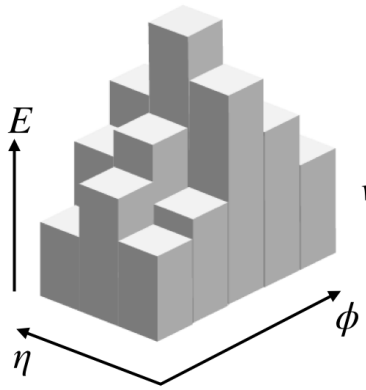
$$R_\phi = \frac{E^{S_2}_{3\times3}}{E^{S_2}_{3\times7}}$$

(a) The quantity R_ϕ is a measure of the central of the energy deposited in the first layer of the ECal is in the ϕ direction. The energy in the dark gray area represents the numerator while the energy in the light gray plus dark gray area represent the denominator.



$$R_\eta = \frac{E^{S_2}_{3\times7}}{E^{S_2}_{7\times7}}$$

(b) The quantity R_η is a measure of the central of the energy deposited in the first layer of the ECal is in the η direction.The energy in the dark gray area represents the numerator while the energy in the light gray plus dark gray area represent the denominator.



$$w_{\eta_2} = \sqrt{\frac{\Sigma E_i \eta_i^2}{\Sigma E_i} - \left(\frac{\Sigma E_i \eta_i}{\Sigma E_i}\right)^2}$$

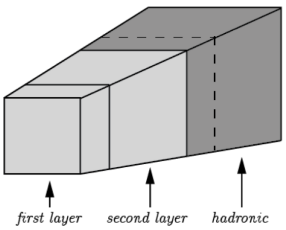
(c) The quantity w_{η_2} is a measure of the width of the energy distribution in the η direction averaged over ϕ in a 3×5 rectangle.

Figure 4.4: Discriminating variables for photon identification involving the second layer of the ECal

4.3 Photon Calibration

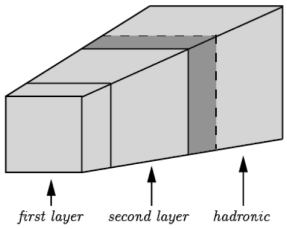
4.3.1 Energy Scale and Resolution Measurements with $Z \to ee$ Decays

Like reconstruction, calibration of photons is closely related to the calibration of electrons. The calibration begins with simulating and measuring the decay $Z \to ee$.



$$R_{\text{had}} = \frac{E_{\text{T}}^{\text{had}}}{E_{\text{T}}}$$

(a) The transverse energy in the dark gray volume represents the numerator and the transverse energy in the light gray and dark gray volumes represents the denominator



$$R_{\text{had 1}} = \frac{E_{\text{T}}^{\text{had 1}}}{E_{\text{T}}}$$

(b) The transverse energy in the dark gray volume represents the numerator and the transverse energy in the light gray and dark gray volumes represents the denominator

Figure 4.5: Discriminating variables for photon identification involving the HCal.

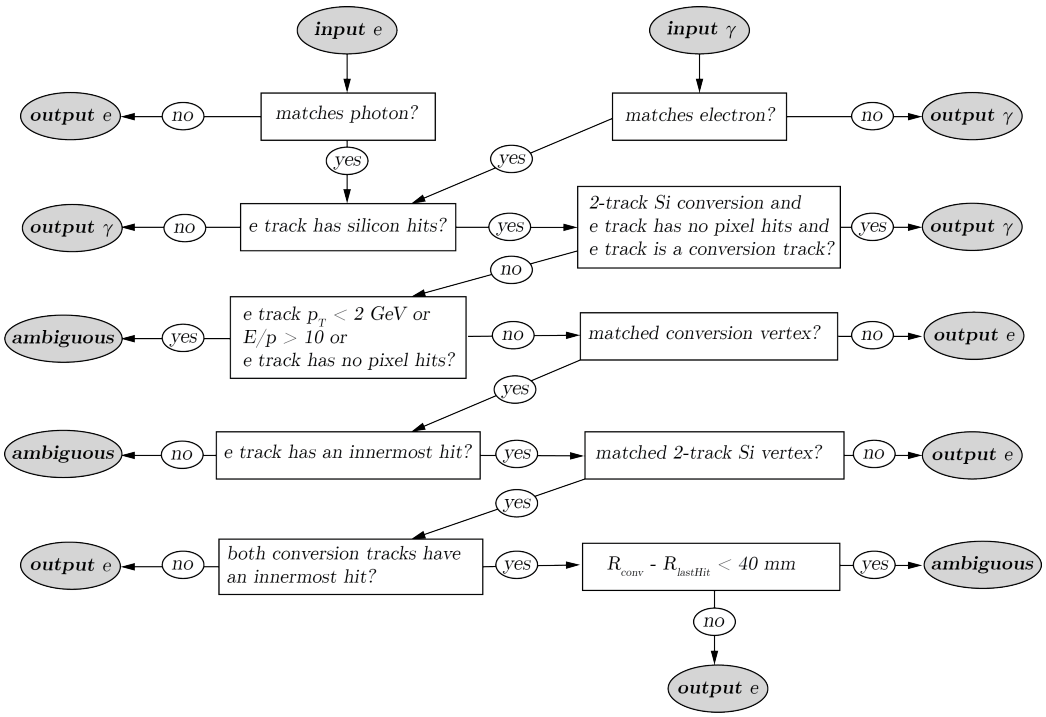


Figure 4.6: Flowchart showing the logic of the ambiguity resolution for particles initially reconstructed both as electrons and photons. An innermost hit is a hit in the functioning pixel nearest to the beam line along the track trajectory, E/p is the ratio of the supercluster energy to the measured momentum of the matched track, R_{conv} is the radial position of the conversion vertex, and R_{firstHit} is the smallest radial position of a hit in the track or tracks that make a conversion vertex (*Aad et al.*, 2019).

The data's energy scale is corrected by dividing by a scale factor, $1 + \alpha_i$, while the

energy resolution is corrected with an additive constant, c_i

$$E^{\text{data,corr}} = \frac{E^{\text{data}}}{1+\alpha_i} \tag{4.8}$$

$$\left(\frac{\sigma_E}{E}\right)^{\text{MC,corr}} = \left(\frac{\sigma_E}{E}\right)^{\text{MC}} \oplus c_i \tag{4.9}$$

where $E^{\text{data,corr}}$ is the corrected energy scale of the data, E^{data} is the unocrrected energy scale of the data, $(\sigma_E/E)^{\text{MC,corr}}$ is the corrected simulated energy resolution, $(\sigma_E/E)^{\text{MC}}$ is the uncorrected simulated energy resolution.

For samples of $Z \to ee$ decays with electrons reconstructed at the (i,j) η region, the effect of the energy scale corrections on the dilepton invariant mass is given in first order by

$$m_{ij}^{\text{data,corr}} = \frac{m_{ij}^{\text{data}}}{1+\alpha_{ij}} \tag{4.10}$$

$$\alpha_{ij} = \frac{\alpha_i + \alpha_j}{2} \tag{4.11}$$

while the difference in the simulated mass resolution is given by

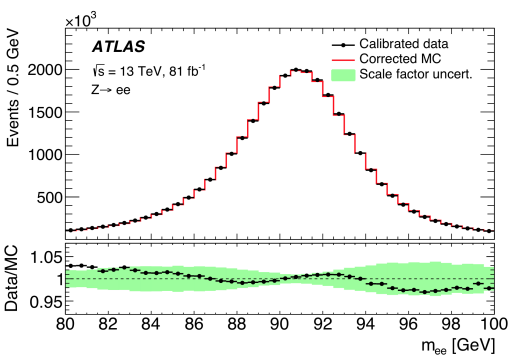
$$\left(\frac{\sigma_m}{m}\right)_{ij}^{\text{MC,corr}} = \left(\frac{\sigma_m}{m}\right)_{ij}^{\text{MC}} \oplus c_{ij} \tag{4.12}$$

$$c_{ij} = \frac{c_i \oplus c_j}{2}. \tag{4.13}$$

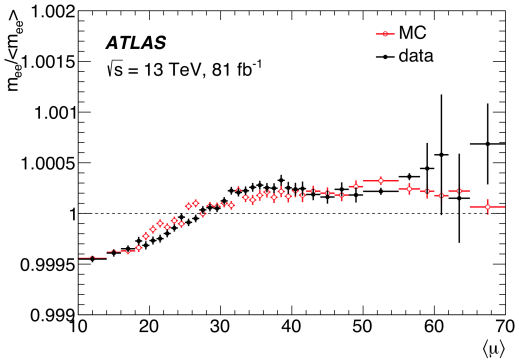
To determine the values of α_{ij} and c_{ij}, the agreement between the invariant mass distributions in the data and simulations are optimized separately for each (i,j) pair. This optimization is performed using two separate methods and the difference is taken as a systematic uncertainty. The first method estimates α_{ij} and c_{ij} by minimizing the χ^2 of the difference between data and simulation templates. To create the templates the mass scale in simulation is shifted by α_{ij} and an extra resolution contribution of c_{ij} is applied. The second method fits a sum of three Gaussian functions to the

data and simulated invariant mass distributions in each (i, j) region. The α_i and c_i are extracted from the differences of the means and widths of the fitted distributions between data and simulations.

Figure 4.7 shows the invariant mass distribution for $Z \to ee$ candidates for data and simulation after the energy scale correction has been applied to the data and the resolution correction to the simulations and the stability of the reconstructed peak position of the dielectron mass distribution as a function of the average number of interactions per bunch crossing for the data collected in 2015, 2016, and 2017.



(a) Comparison between data and simulation of the invariant mass distribution of the two electrons in the selected $Z \to ee$ candidates, after the calibration and resolution corrections are applied. The total number of events in the simulation is normalized to the data. The uncertainty band of the bottom plot represents the impact of the uncertainties in the calibration and resolution correction factors.



(b) Relative variation of the peak position of the reconstructed dielectron mass distribution is $Z \to ee$ events as a function of the average number of interactions per bunch crossing. The error bars represent the statistical uncertainties.

Figure 4.7: Summary of the energy scale and resolution calibration with $Z \to ee$ decays (*Aad et al.*, 2019).

4.3.2 Validation of the Photon Energy Scale with $Z \to \ell\ell\gamma$ Decays

The energy scale corrections extracted from $Z \to ee$ decays, as described in Section 4.3.1, are applied to correct the photon energy scale. Using the radiative decays of the Z boson a data driven validation of the photon energy scale correction is performed. After applying the Z based energy scale corrections, residual energy scale

factors for photon, $\Delta\alpha$, are derived by comparing the mass distribution of the $\ell\ell\gamma$ system in data and simulation. The residual scale factors are applied to the photon energy, and the value of the $\Delta\alpha$ that minimizes the χ^2 comparison between the data and the simulation is extracted.

4.3.3 Energy Scale and Resolution Corrections in Low Pile-Up Data

Energy scale factors are derived for a special low pile-up sample collected in 2017 using the method described in Section 4.3.1 for 24 η regions. Another approach consists of measuring the energy scale factors using high pile-up data and extrapolating the results to the low pile-up conditions. This method is used as validation for the primary method. The explicit dependence of the energy corrections on $\langle\mu\rangle$ and differences between the clustering thresholds used for the two samples are the two main effects considered in the extrapolation.

CHAPTER V

The Two Photon Final State: Standard Model Prediction and New Physics

This thesis now shifts its focus to the diphoton final state. It begins with a review of parton interactions and two body decays and $2 \to 2$ scattering processes. Next, it describes SM predicitions for events with diphoton final states, that is the background of the resonance search. Lastly, it describes in detail the predicitions of some BSM models and results from past searches.

5.0.1 Parton Interactions

At high energies collisions do not occur directly between hadrons, but rather between the quarks and gluons that make them up. In this context these quarks and gluons are known as *partons*. Each parton carries just a fraction of the momentum of its parent particle. How much momentum is carried by a parton is given by its PDF and is dependent on the momentum transfer Q. An example of a PDF for a proton with $Q^2 = 10$ GeV^2 is shown in Figure 5.1.

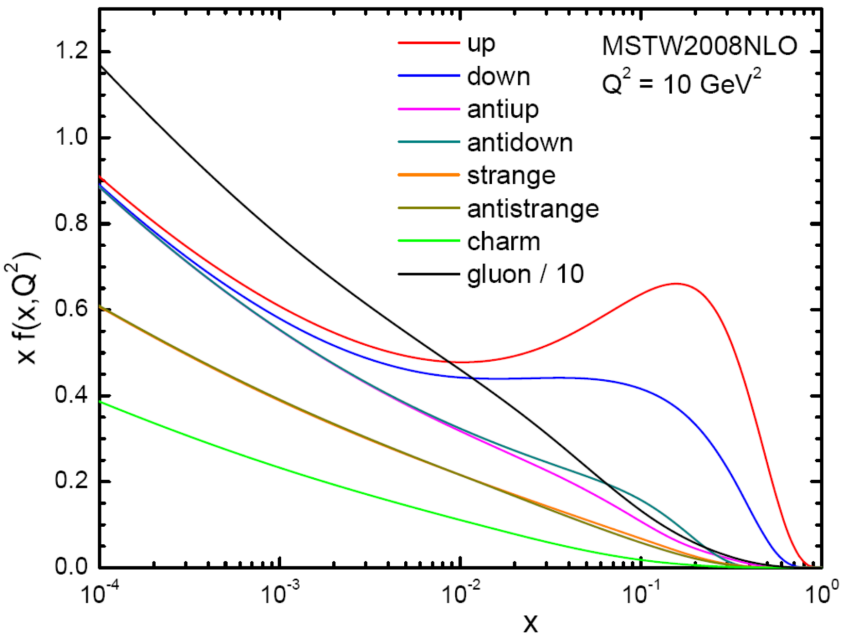


Figure 5.1: The MSTW2008NLO parton distribution function at $Q^2 = 10^2$ GeV, where the horizontal axis is x, the fraction of the hadron's momentum carried by a parton, and the y-axis is x times the probability density $f(x)$.

5.1 Two Body Decays

The lifetime of an individual particle cannot be predicted, however the *decay rate*, Γ, that is the probability per unit time that a given particle will decay, can be determined. Consider an ensemble of $N \to \infty$ identical particles. The change in the number of particles after a time dt is given by

$$dN = -\Gamma N \; dt \tag{5.1}$$

therefore the expected number of particles surviving after time t is given by

$$N(t) = N(0)e^{-\Gamma t}. \tag{5.2}$$

The *lifetime*, the time in takes for a ensemble to become $1/e$ of its original size, is

$$\tau = \frac{1}{\Gamma} \tag{5.3}$$

If there are n decay modes of the inital particle, the total rate is given by

$$\Gamma_{\text{total}} = \sum_{i=1}^{n} \Gamma_i \tag{5.4}$$

the lifetime by

$$\tau = \frac{1}{\Gamma_{\text{total}}} \tag{5.5}$$

and the branching ratios by

$$B_i = \frac{\Gamma_i}{\Gamma_{\text{total}}}. \tag{5.6}$$

The process of a particle with four-momentum $p^\mu = (M, 0, 0, 0)$ decaying to n particles with four-momenta k_i and masses m_i has the differential decay rate

$$d\Gamma = \frac{(2\pi)^4}{2M} |\mathcal{M}|^2 d\Phi_n \tag{5.7}$$

where

$$d\Phi_n = \delta^4 \left(p - \sum_{i=1}^{n} k_i \right) \prod_{i=1}^{n} \left(\frac{d^3 \boldsymbol{k_i}}{(2\pi)^3 2E_i} \right) \tag{5.8}$$

is the n-body Lorentz invariant phase space.

Consider a particle of mass M decaying to two particles of mass m_1 and m_2 with momenta $\boldsymbol{p_1}$ and $\boldsymbol{p_2}$ respectively, as shown in Figure 5.2. In the rest frame of the initial particle, that is the frame in which the initial particle is at rest, the energies

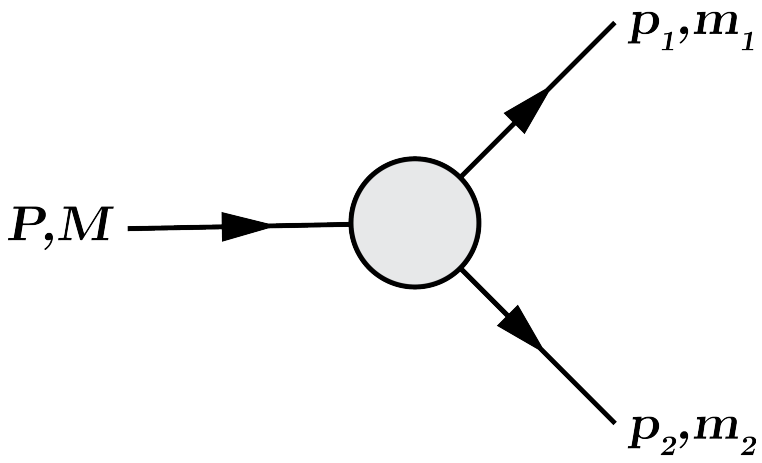


Figure 5.2: Cartoon of a particle of mass M and momentum $\boldsymbol{P}$ decaying into two particles of masses m_1 and m_2 and momenta $\boldsymbol{p_1}$ and $\boldsymbol{p_2}$. The arrows represent the direction of the momentum.

and momenta of the outgoing particles are given by

$$E_1 = \frac{M^2 + m_1^2 - m_2^2}{2M} \tag{5.9}$$

$$E_2 = \frac{M^2 - m_1^2 + m_2^2}{2M} \tag{5.10}$$

$$\begin{aligned} |\boldsymbol{p_1}| &= |\boldsymbol{p_2}| \\ &= \frac{\left[(M^2 - (m_1 + m_2)^2)\,(M^2 - (m_1 - m_2)^2)\right]^{1/2}}{2M} \end{aligned} \tag{5.11}$$

The differential decay rate is given by

$$d\Gamma = \frac{1}{32\pi^2}|\mathcal{M}|^2 \frac{|\boldsymbol{p_i}|}{M^2}\, d\phi\, d(\cos\theta) \tag{5.12}$$

where $i = 1, 2$. In the case that the outgoing particles are massless the energies and momenta simplify to

$$E_1 = E_2 = |\boldsymbol{p_1}| = |\boldsymbol{p_2}| = \frac{M}{2} \tag{5.13}$$

while the differential decay rate becomes

$$d\Gamma = \frac{|\mathcal{M}|^2}{64\pi^2 M}\, d\phi\; d(\cos\theta). \tag{5.14}$$

5.2 $2 \to 2$ Scattering Processes

A $2 \to 2$ scattering process, illustrated in Figure 5.3, is a process in which two incoming particles interact, either elastically or inelastically, leading to two outgoing particles. The outgoing particles can be the same particles as the incoming particles

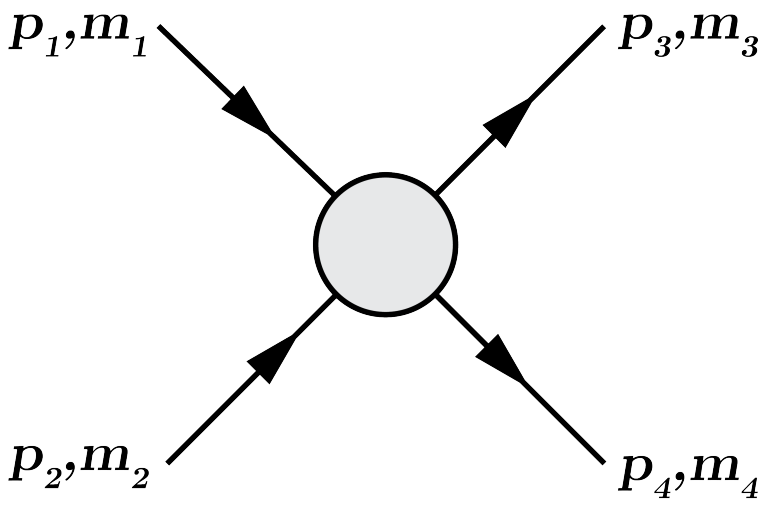


Figure 5.3: Cartoon of a $2 \to 2$ scattering process where the incoming particles have masses m_1 and m_2 and momenta $\boldsymbol{p_1}$ and $\boldsymbol{p_2}$ and the outgoing particles have masses m_3 and m_4 and momenta $\boldsymbol{p_1}$ and $\boldsymbol{p_2}$. The Arrows represent the direct of the momentum.

but with new four momenta, or the incoming particles can be destroyed with two new outgoing particles produced. Consider the process $\varphi_1\varphi_2 \to \varphi_3\varphi_4$ where the particles φ_i have the momenta $\boldsymbol{p_i}$ and energy E_i. The four momenta of the particles are given by $p_i = (E_i, \boldsymbol{p_i})$. The *scattering matrix* S (*Sakurai and Napolitano*, 1964) is related

to the *reduced matrix element* $\mathcal{M}$ by

$$\begin{aligned}\langle p_3 p_4|S|p_1 p_2\rangle = \mathbb{1} - i(2\pi)^4\delta^4(p_1+p_2-p_3-p_4) \\ \times \frac{\mathcal{M}(p_1,p_2;p_3,p_4)}{(2E_1)^{1/2}(2E_2)^{1/2}(2E_3)^{1/2}(2E_4)^{1/2}}\end{aligned} \tag{5.15}$$

where the state normalization is such that

$$\langle p|q\rangle = 2E(2\pi)^3\delta^3(\boldsymbol{q}-\boldsymbol{p}) \tag{5.16}$$

where $E = \left(m_q^2 + |\boldsymbol{p}_q^2|\right)^{1/2}$.

The differential cross section in the COM frame is given by

$$d\sigma = \frac{|\mathcal{M}|^2}{64\pi^2 E_{\text{cm}}^2}\frac{|\boldsymbol{p_i}|}{|\boldsymbol{p_j}|}d\Omega \tag{5.17}$$

where $i = 1, 2$ and $j = 3, 4$.

5.2.1 Kinematics

At tree level, that is processes whose Feynman diagrams do not include any closed loops, there are three distinct $2 \to 2$ scattering categories. Feynman diagrams of the three categories are shown in Figure 5.4. In s-channel scatters, the two incoming

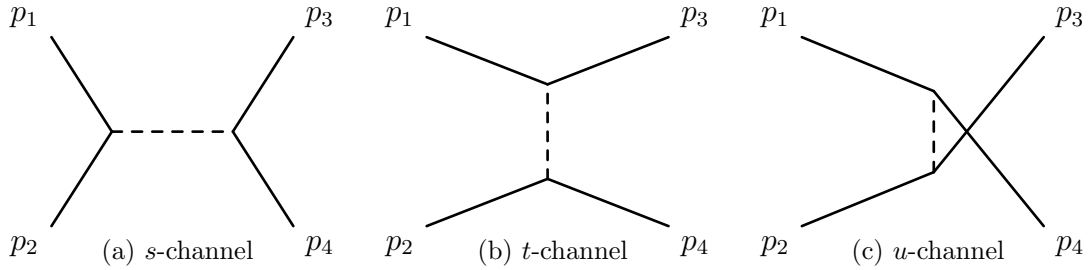


Figure 5.4: Feynman diagrams of the s-channel, t-channel, and u-channel scattering. The solid lines represent the incoming and outgoing particles while the dashed lines represent the exchanged particles. The particles with four-momenta p_1 and p_2 are the incoming particles while the particles with four-momenta p_3 and p_4 are the outgoing particles.

particles annihilate to produce the exchange particle which decays to the outgoing particles. In t-channel scatters, the incoming particles exchange a particle which leads to their outgoing four momenta differing from the incoming four momenta, while u-scatters are the same with the outgoing particles swapped. The Lorentz invariant *Mandelstam variables* are defined as follows.

$$\begin{aligned} s &= (p_1 + p_2)^2 = (p_3 + p_4)^2 \\ &= m_1^2 + 2E_1E_2 - 2\boldsymbol{p_1} \cdot \boldsymbol{p_2} + m_2^2 \end{aligned} \tag{5.18}$$

$$\begin{aligned} t &= (p_1 - p_3)^2 = (p_2 - p_4)^2 \\ &= m_1^2 + 2E_1E_3 - 2\boldsymbol{p_1} \cdot \boldsymbol{p_3} + m_3^2 \end{aligned} \tag{5.19}$$

$$\begin{aligned} u &= (p_1 - p_4)^2 = (p_2 - p_3)^2 \\ &= m_1^2 + 2E_1E_4 - 2\boldsymbol{p_1} \cdot \boldsymbol{p_4} + m_4^2 \end{aligned} \tag{5.20}$$

They represent the four-momenta of the exchange particles in each process and satisfy the following relation.

$$s + t + u = m_1^2 + m_2^2 + m_3^2 + m_4^2 \tag{5.21}$$

Note that since in the COM frame $p_1 + p_2 = (E_1 + E_2, \mathbf{0}) = (E_{\text{cm}}, \mathbf{0})$, then $(p_1 + p_2)^2 = E_{\text{cm}}^2$ which gives

$$\sqrt{s} = E_{\text{cm}} \tag{5.22}$$

where E_{cm} is the COM energy. The energy of the particles in the COM frame is given by

$$E_i = \frac{1}{2\sqrt{s}}(s + m_i^2 - m_j^2) \tag{5.23}$$

where $i = 1, 3$ and $j = 2, 4$ or $i = 2, 4$ and $j = 1, 3$. In the case where the outgoing particles are massless,

$$E_i = |\boldsymbol{p_3}| = |\boldsymbol{p_4}| = \frac{\sqrt{s}}{2} \tag{5.24}$$

where $i = 1, 2, 3, 4$.

The *scattering angle* θ in the COM frame, pictured in Figure 5.5, is defined by

$$\boldsymbol{p} \cdot \boldsymbol{p'} = |\boldsymbol{p}||\boldsymbol{p'}| \cos\theta. \tag{5.25}$$

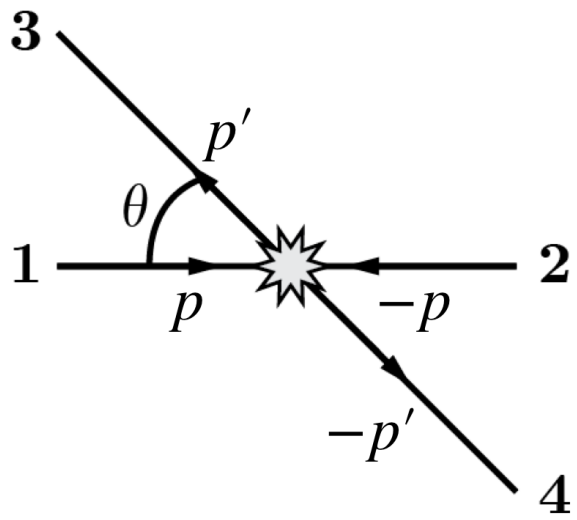


Figure 5.5: Illustration of scattering angle in a $2 \to 2$ scattering process in the center of mass frame.

The angular distribution, assuming a uniform distribution in ϕ, is given by

$$d\Omega = 2\pi \cos\theta \tag{5.26}$$

$$\frac{d\Omega}{dt} = \frac{4\pi s}{\sqrt{\lambda(s, m_1^2, m2^2)}\sqrt{\lambda(s, m_3^2, m_4^2)}} = \frac{\pi}{|\boldsymbol{p}||\boldsymbol{p'}|} \tag{5.27}$$

where $\lambda(a, b, c) = a^2 + b^2 + c^2 - 2ab - 2ac - 2bc$ is the Källén function. The cross

section can be written in terms of s and t as

$$\frac{d\sigma}{dt} = \frac{1}{64\pi s}\frac{1}{|\boldsymbol{p}|^2}|\mathcal{M}|^2. \tag{5.28}$$

5.3 Predicitons for Standard Model Backgrounds

5.3.1 Diphton Pair Production Processes at the LHC

Competing with any signal events, that is events coming from a resonance, is a background many orders of magnitude larger. The background is composed of the *reducible* and *irreducible* background. The irreducible background is composed of events with at least two photons in the final state, that is events with the form $x + x' \to \gamma\gamma + X$ where x and x' are partons and X represent $n \in \mathbb{N}$ particles.

The two main processes contributing to the irreducible background, shown in Figure 5.6, are $q\bar{q} \to \gamma\gamma$ s-channel scattering, known as the Born process, and $gg \to \gamma\gamma$ box diagram with a fermion intermediate state.

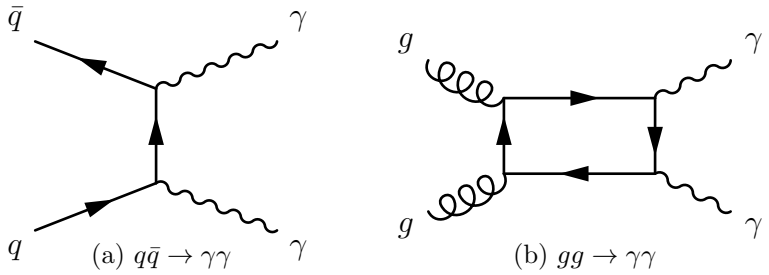


Figure 5.6: Feynman diagrams for the main processes of diphoton production at the LHC: (a) The Born diagram, and (b) the box diagram.

Although the $q\bar{q} \to \gamma\gamma$ process is of order α^2_{QED} and the $gg \to \gamma\gamma$ process is of order $\alpha^2_{\text{S}}\alpha^2_{\text{QED}}$, the gluon PDFs are enhanced at the LHC, so the cross section can be $\approx 30\%$ of the $q\bar{q} \to \gamma\gamma$ process. Processes including radiative and virtual corrections increase the diphoton production cross section. In the case of the Born process, it increases the cross section by approximately 30%. Figure 5.7 shows the main $qg \to j\gamma$ processes with an additional photon coming from initial state radiation (ISR) or final

state radiation (FSR). The $qg \to q\gamma\gamma$ processes, although of order $\alpha_S \alpha_{QED}^2$, are also

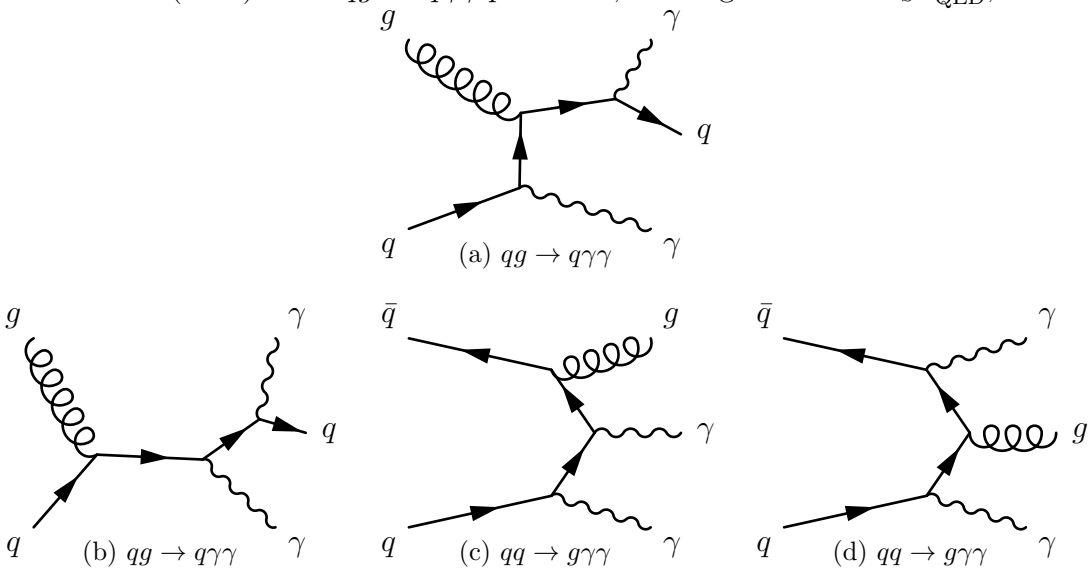


Figure 5.7: Feynman diagrams for the main diphoton processes originating from radiative corrections to $j\gamma$ and γj events.

enhanced due to the large gluon PDFs at the LHC.

Also contributing to the cross section are events where a fragmentation photon is produced. These events can come from $x + x' \to q\gamma + X$ and $x + x' \to qq + X$ processes where the quarks fragment, a process where a quark hadronizes producing new particles, into energetic photons and x and x' are partons. The two main fragmentation processes contributing to the diphoton production cross section are shown if Figure 5.8.

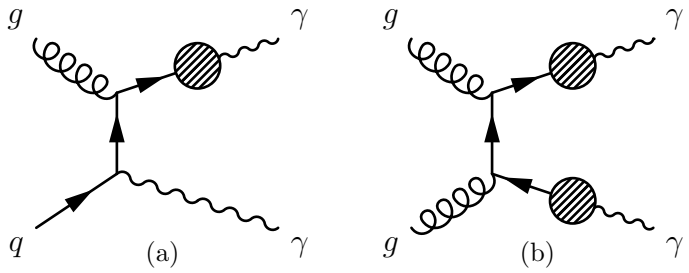


Figure 5.8: Feynman diagrams of diphoton events with (a) one photon coming from fragmentation and (b) both photons coming from fragmentation.

The *reducible background* is mostly comprised of $j\gamma$, γj and jj events misidentified as diphoton events. These events contribute non negligibly to the ATLAS diphoton

mass spectrum. Although the number of misidentified photons constitute just a small fraction of the jets produced at the LHC, on the order a few percent, the number of jets produced far exceeds the number of photons produced. There are approximately two orders of magnitude more $j\gamma$ and γj events than $\gamma\gamma$ events and approximately seven orders of magnitude more jj events than $\gamma\gamma$ events, so these fakes end up contributing significantly to the diphoton mass spectrum. Figure 5.9 shows the main processes that contribute to faked diphoton events.

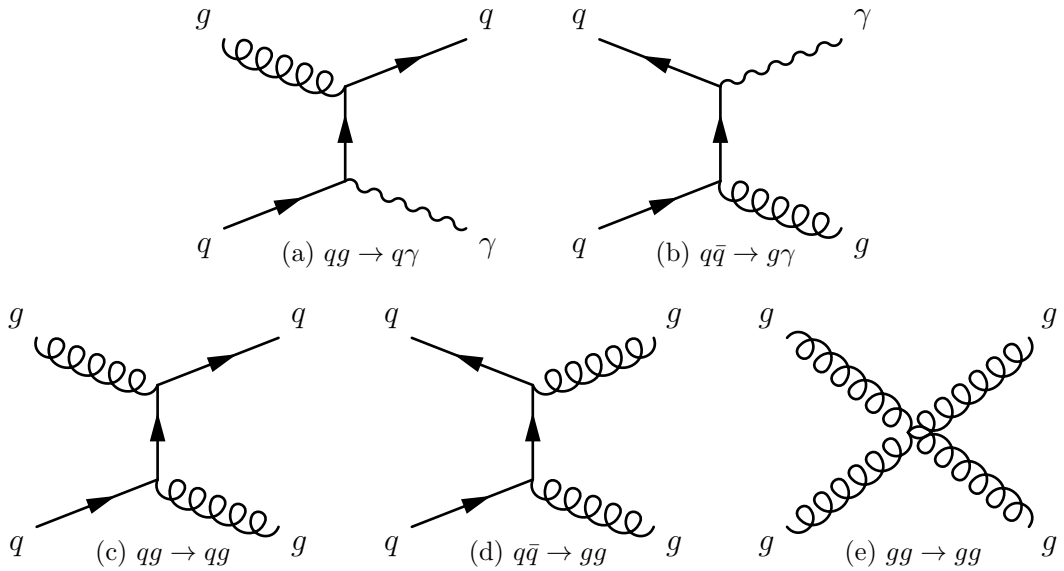


Figure 5.9: Feynman diagrams of the main processes contributing to the reducible background.

5.4 Predictions for Possible New Processes

5.4.1 Graviton Phenomenology

Although the graviton is predicted to be a massless spin-2 boson, KK theories predict the existence of graviton massive resonances, as discussed in Section 2.5.2.1. In KK theories, gravitons appear as a series of separate resonances over a smoothly

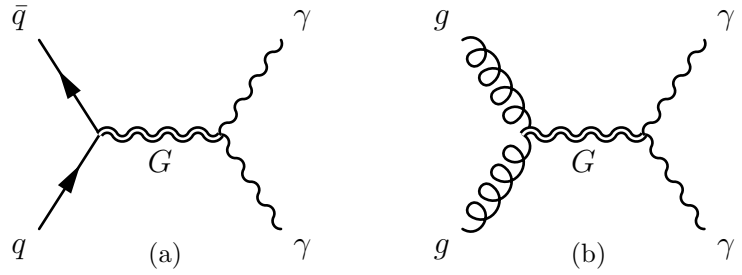


Figure 5.10: Feynman diagrams for the main processes for $x + x' \to G \to \gamma\gamma$ where x and x' are partons.

falling mass spectrum. The interaction Lagrangian is given by

$$\mathcal{L}_{\text{int}} = -\frac{1}{M_{\text{Pl}}} T^{\alpha\beta} h^{(0)}_{\alpha\beta} - \frac{1}{\Lambda_G} T^{\alpha\beta} \sum_{n=1}^{\infty} h^{(n)}_{\alpha\beta} \tag{5.29}$$

where $h^{(0)}_{\alpha\beta}$ and $h^{(n)}_{\alpha\beta}$ are the massless and massive gravitons respectively, M_{Pl} is the Planck scale, $T^{\alpha\beta}$ is the energy-momentum tensor of the matter fields, and Λ_G is the energy scale (*Tang*, 2012). The total cross section for graviton production at hadron colliders is given by

$$\sigma = \int dx_1\, dx_2\, f_{q_1}(x_1, \mu_F) f_{q_2}(x_2, \mu_F) \hat{\sigma}(q_1 q_2 \to G^*; \hat{s}) \tag{5.30}$$

where $f_{q_i}(x, \mu_F)$ is the PDF for a parton with momentum fraction x at the factorization scale μ_F, $\hat{\sigma}$ is the partonic cross section with the initial two partons, q_1,q_2 = q,$\bar{q}$,g, of momentum fraction x_1 and x_2 respectively, and $\hat{s} = x_1 x_2 s$. The two main processes for $G \to \gamma\gamma$ produced at hadron colliders are shown in Figure 5.10.

Figure 5.11 shows the graviton branching ratios as a function of mass. For large masses, that is masses on the TeV scale, the branching ratios are nearly constant. The branching ratios in the high mass limit are shown in Table 5.1. The total decay

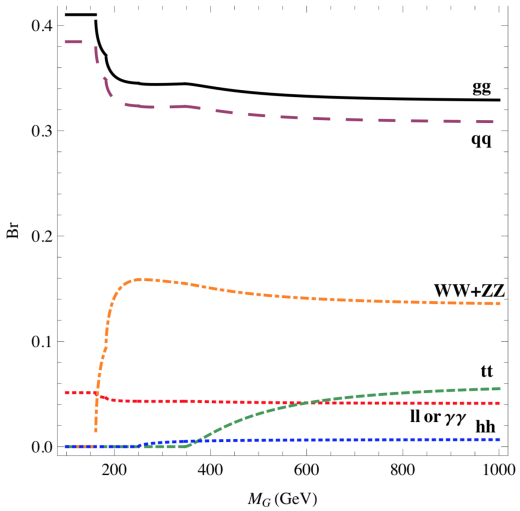


Figure 5.11: Branching ratios for the various decay modes of the graviton as a function of mass where $q = u, d, c, s, b$ and $l = e, \mu$ (*Tang*, 2012).

Decay Mode	Branching Ratio
gg	96/293
$q\bar{q}$	90/293
W^+W^-	26/293
$t\bar{t}$	18/293
ZZ	13/293
$\gamma\gamma$	12/293
$\ell^+\ell^-$	12/293
$\tau^+\tau^-$	6/293
$\nu_e\bar{\nu}_e$	6/293
$\nu_\mu\bar{\nu}_\mu$	6/293
$\nu_\tau\bar{\nu}_\tau$	6/293
hh	2/293

Table 5.1: Branching ratios for the various decay modes of the graviton in the high mass limit where $q = u, d, c, s, b$ and $\ell = e, \mu$ (*Tang*, 2012).

rate for $G \to X$ and $G \to \gamma\gamma$ are given by

$$\Gamma(G \to X) \simeq \frac{293 M_G^3}{960\pi\Lambda_G^2} \tag{5.31}$$

$$\Gamma(G \to \gamma\gamma) \simeq \frac{M_G^3}{80\pi\Lambda_G^2} \tag{5.32}$$

where X is any final state, M_G is the mass of the resonance and Λ_G is the energy

scale (*Tang*, 2012). The number of events where a graviton is produced and decays to two photons is therefore given by

$$N_{pp\to G\to\gamma\gamma} = \sigma(pp \to G) \times \mathrm{Br}(G \to \gamma\gamma) \approx \frac{12}{293}\sigma(pp \to G). \tag{5.33}$$

Although the the graviton decays to photons only about 4% of the time, the diphoton channel is still an attractive discovery channel for its high resolution and clean background. The Higgs boson decays to photons approximately 0.23% of the time, and the diphoton channel was an important discovery channgel.

5.4.1.1 Randall-Sundrum Model

In the Randall-Sundrum (RS) model, discussed in Section 2.5.2.2, the dimensionless couplings of the graviton states $k/\overline{M}_{\mathrm{Pl}}$ and masses m_n are determined by the scale $\Lambda_G = e^{-k\pi L}\overline{M}_{\mathrm{Pl}} \approx \mathcal{O}(\mathrm{TeV})$ where $\overline{M}_{\mathrm{Pl}} = M_{\mathrm{Pl}}/\sqrt{4\pi}$ is the reduced Planck Scale, k is the curvature scale of the extra dimension, and L is the size of the extra dimension. The natural width is related to $k/\overline{M}_{\mathrm{Pl}}$ by

$$\Gamma(G^* \to X) \sim \left(k/\overline{M}_{\mathrm{Pl}}\right)^2 \tag{5.34}$$

where X is any final state. For $k/\overline{M}_{\mathrm{Pl}} < 0.3$, the lightest graviton state is expected to be fairly narrow, that is the width is at most on the order of the ATLAS detector resolution.

Results from past searches are shown in Table 5.2. Figure 5.12 shows the limit plots from the most recent CMS and ATLAS high mass diphoton results.

5.4.2 2HDM Phenomenology

The 2HDM has several variations each with their own couplings and therefore phenomenology. However, in all models the coupling $\lambda_{h/H}^{W/Z}$ of $h(H)$ to W and Z is the

Experiment	$\sqrt{s}$ (TeV)	Luminosity (fb^{-1})	Limit
ATLAS	7	4.9	2.06(1.00) TeV for $k/\overline{M}_{\text{Pl}} = 0.1(0.01)$
ATLAS	13	36.7	4.1 TeV for $k/\overline{M}_{\text{Pl}} = 0.1$
CMS	7	2.2	2.66(1.41) TeV for $k/\overline{M}_{\text{Pl}} = 0.1(0.01)$
CMS	13	35.9	4.1(2.3) TeV for $k/\overline{M}_{\text{Pl}} = 0.1(0.01)$
CDF	1.96	5.4/5.7	1.06 TeV for $k/\overline{M}_{\text{Pl}} = 0.1$
DØ	1.96	5.4	1.05(0.56) TeV for $k/\overline{M}_{\text{Pl}} = 0.1(0.01)$

Table 5.2: Results from past searches of the Randall-Sundrum graviton with $k/\overline{M}_{\text{Pl}} = 0.1$. The ATLAS (*Aad et al.*, 2013) (*Aaboud et al.*, 2017)and CMS (*Chatrchyan et al.*, 2012) (*Sirunyan et al.*, 2018)results are for $G^* \to \gamma\gamma$, while the CDF (*Aaltonen et al.*, 2011) and DØ (*Abazov et al.*, 2010) results are for $G^* \to \gamma\gamma$ and $G^* \to e^+e^-$.

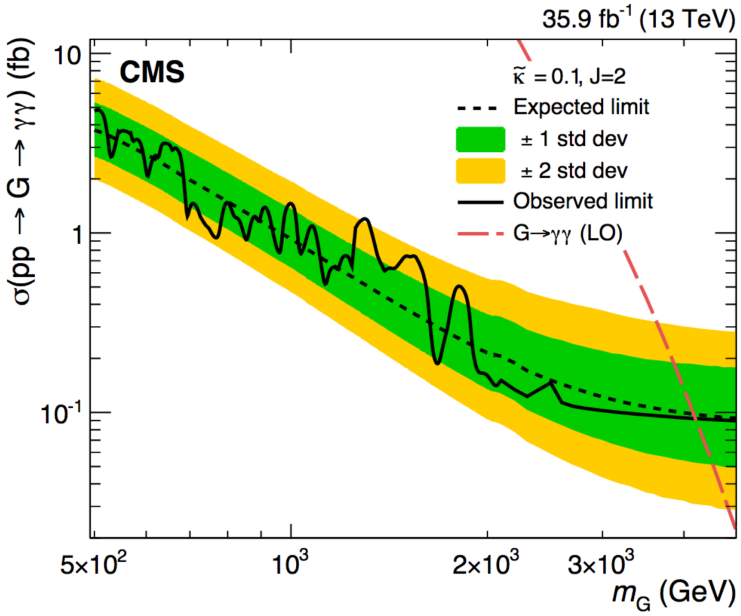
same as the SM, the coupling $\lambda_h^{WW/ZZ}$ of h to WW and ZZ is the same as the SM times $\sin(\alpha-\beta)$, the coupling $\lambda_H^{WW/ZZ}$ of H to WW or ZZ is the same as the SM times $\cos(\alpha-\beta)$, and the coupling of A to vector bosons vanishes. These couplings are summarized in Table 5.3. Table 5.4 shows $\xi_X^Y = \lambda^Y_{X\,\text{2HDM}}/\lambda^Y_{X\,\text{SM}}$ where $\lambda^Y_{X\,Z}$ is the

	$\lambda_{\text{2HDM}}/\lambda_{\text{SM}}$
$\lambda_{h/H}^{W/Z}$	1
$\lambda_h^{WW/ZZ}$	$\sin(\alpha-\beta)$
$\lambda_H^{WW/ZZ}$	$\cos(\alpha-\beta)$
$\lambda_A^{W/Z}$	0
$\lambda_A^{WW/ZZ}$	0

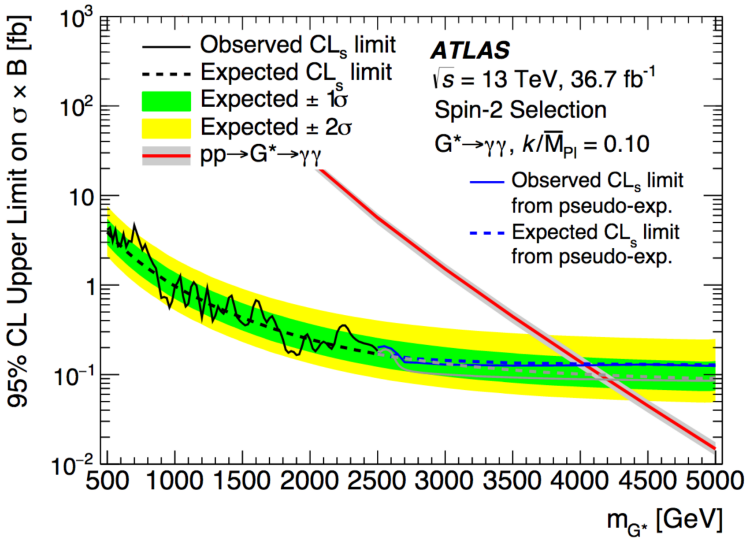
Table 5.3: Summary of the 2HDM particle h, H, and A couplings to vector bosons relative to the SM Higgs h couplings.

Yukawa coupling of X to Y in the model Z.

Figure 5.13 shows the dominant production modes for the SM and the 2HDM CP-even Higgs bosons: gluon-gluon fusion (ggF), vector boson fusion (VBF), Higgsstrahlung, and associated Higgs production with $t\bar{t}$ (ttH). In principle the ggF loop can be made by any quark, however since the coupling of the SM Higgs to fermions is proportional to the fermion mass, only the diagrams with top loops contribute significantly to the cross section.



(a)



(b)

Figure 5.12: Limits on the cross section for $pp \to G^* \to \gamma\gamma$ from CMS (*Sirunyan et al.*, 2018) and ATLAS (*Aaboud et al.*, 2017).

In the SM the partonic cross section for ggF is given by

$$\hat{\sigma}_{\text{ggF}} = m_h^2 \delta(\hat{s} - m_h^2) \frac{G_F \alpha_S^2}{512\sqrt{2}\pi} \left| \sum_q A_{1/2}^{\phi}(\tau_q) \right|^2 \tag{5.35}$$

where G_F is the Fermi constant, α_S is the strong coupling constant, $\tau_q = m_h^2/(4m_q^2)$,

	Type I	Type II	Lepton-Specific	Flipped
ξ_h^u	$\cos\alpha/\sin\beta$	$\cos\alpha/\sin\beta$	$\cos\alpha/\sin\beta$	$\cos\alpha/\sin\beta$
ξ_h^d	$\cos\alpha/\sin\beta$	$-\sin\alpha/\cos\beta$	$\cos\alpha/\sin\beta$	$-\sin\alpha/\cos\beta$
ξ_h^ℓ	$\cos\alpha/\sin\beta$	$-\sin\alpha/\cos\beta$	$-\sin\alpha/\cos\beta$	$\cos\alpha/\sin\beta$
ξ_H^u	$\sin\alpha/\sin\beta$	$\sin\alpha/\sin\beta$	$\sin\alpha/\sin\beta$	$\sin\alpha/\sin\beta$
ξ_H^d	$\sin\alpha/\sin\beta$	$\cos\alpha/\cos\beta$	$\sin\alpha/\sin\beta$	$\cos\alpha/\cos\beta$
ξ_H^ℓ	$\sin\alpha/\sin\beta$	$\cos\alpha/\cos\beta$	$\cos\alpha/\cos\beta$	$\sin\alpha/\sin\beta$
ξ_A^u	$\cot\beta$	$\cot\beta$	$\cot\beta$	$\cot\beta$
ξ_A^u	$-\cot\beta$	$\tan\beta$	$-\cot\beta$	$\tan\beta$
ξ_A^u	$-\cot\beta$	$\tan\beta$	$\tan\beta$	$-\cos\beta$

Table 5.4: Summary of the 2HDM particles h, H, and A couplings to fermions relative to the SM Higgs couplings.

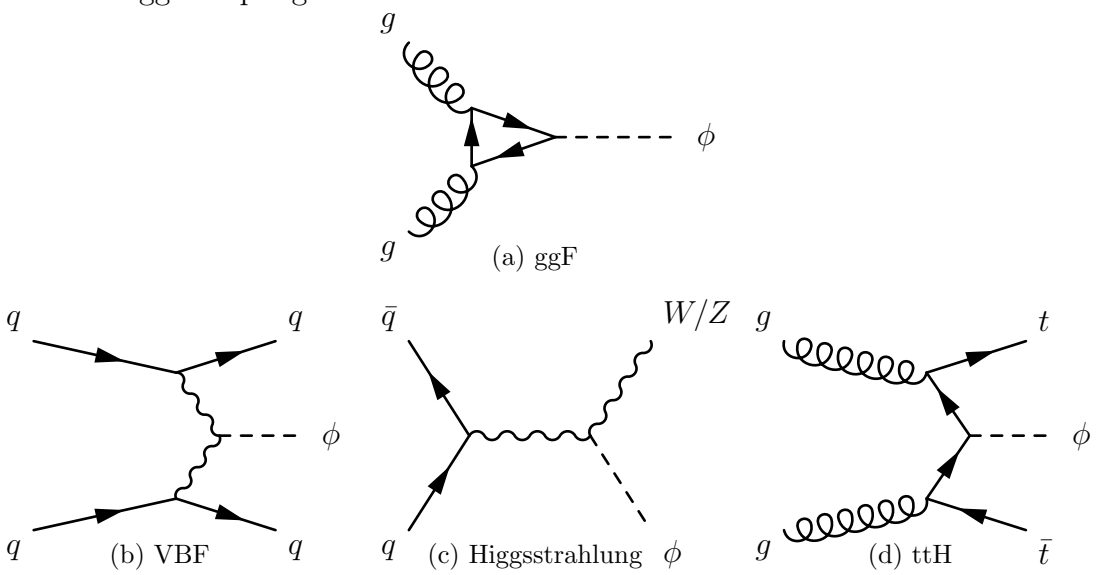


Figure 5.13: Dominant production modes for the 2HDM CP-even Higgs bosons at hadron colliders; (a) and (d) are the dominant production modes for the pseudoscalar Higgs boson at hardron colliders.

and $A_{1/2}^{\phi}(\tau) = 2[\tau + (\tau-1)f(\tau)]/\tau^2$ is the *form factor*, where

$$f(\tau) = -\frac{1}{2}\int_0^1 \frac{dy}{y}\log[1-4\tau y(1-y)] \tag{5.36}$$

$$= \begin{cases} \arcsin^2(\sqrt{\tau}) & \Leftarrow \tau \leq 1, \\ -\frac{1}{4}\left(\log\frac{1+\sqrt{1-1/\tau}}{1-\sqrt{1-1/\tau}} - i\pi\right)^2 & \Leftarrow \tau > 1. \end{cases} \tag{5.37}$$

In the massless quark limit the form factor becomes

$$A_{1/2}^{\phi} \to -[\log(4\tau_q) - i\pi]^2/(2\tau_q). \tag{5.38}$$

In 2HDM the production cross section for a light CP-even Higgs h via ggF is the SM cross section times $(\cos\alpha/\sin\beta)^2$. In the decoupling limit, that is when $\cos(\alpha-\beta) \to 0$, the light CP-even Higgs h has the same cross section as the SM Higgs. In this case, the 125 GeV Higgs boson is the light CP-even Higgs thereby leaving the heavy CP-even Higgs H yet to be discovered. In the limit that $\sin(\alpha-\beta) \to 0$, the heavy CP-even Higgs H has the same cross section as the SM Higgs. In this case, the 125 GeV Higgs boson is the heavy CP-even Higgs thereby leaving the light CP-even Higgs h yet to be discovered.

In Type I and Lepton-Specific 2HDM, the production cross section of the light CP-even Higgs is the SM coupling multiplied by $(\cos\alpha/\sin\beta)^2$ since the top-loop contribution to the amplitude is multiplied by $\cos\alpha/\sin\beta$. In Type II and Flipped 2HDM, the top-loop contribution to the amplitude is modified in the same way. However, if $\tan\beta$ is large then the b quark's Yukawa coupling becomes large, and ggF diagrams with bottom-loops contribute non negligibly. The bottom-loop contribution to the amplitude is multiplied by $-\tan\alpha\tan\beta$. For $m_h = 100$ GeV the cross section is multiplied by the factor $|1+(5-8i)\tan\alpha\tan\beta/100|^2$ and scales like m_h^{-2} for other masses (*Branco et al.*, 2012). By picking the right parameters one can greatly increase the ggF production cross section of Type I and Lepton-Specific 2HDM.

The cross sections for the heavy CP-even Higgs H are modified similarly. In Type I and Lepton-Specific 2HDM the cross section is multiplied by $(\sin\alpha/\sin\beta)^2$. In Type II and Flipped 2HDM the top-loop contribution to the amplitude is multiplied by $(\sin\alpha/\sin\beta)^2$ and the bottom-loop contribution to the amplitude is multiplied by the factor $\cot\alpha\tan\beta$.

The pseudoscalar A has the form factor $A^A_{1/2}(\tau) = 2f(\tau)/\tau$ in ggF production. In addition to this change, the amplitude is multiplied by $\cot\beta$, which blows up for small $\tan\beta$. For $m_A = 100$ GeV the production rate in Type II and Flipped 2HDM is $|1-(3.5-4i)\tan^2\beta/100|^2$ times that of Type I an Lepton-Specific (*Branco et al.*, 2012).

For the light CP-even Higgs h in the VBF or Higgsstrahlung process, the SM cross section is multiplied by $\sin^2(\alpha-\beta)$. For the heavy CP-even Higgs H the SM cross section is multiplied by $\cos^2(\alpha-\beta)$. There are no W^+W^-A or ZZA vertices so the pseudoscalar A cannot be produced via VBF or Higgsstrahlung.

For the light CP-even Higgs h in the ttH process, the SM cross section is multiplied by $(\cos\alpha/\sin\beta)^2$. For the heavy CP-even Higgs H the SM cross section is multiplied by $(\sin\alpha/\sin\beta)^2$. For the pseudoscalar A the SM cross section is multiplied by $\cot^2\beta$.

Because of the possibility for enhanced coupling to the b quark in Type II and Flipped 2HDM, there can be a non-trivial contribution to the production cross section from associated production with $b\bar{b}$. For the light CP-even Higgs the cross section of SM ttH production is multiplied by $(\sin\alpha/\cos\beta)^2(m_b/m_t)^2$. For the heavy CP-even Higgs the cross section of SM ttH production is multiplied by $(\sin\alpha/\sin\beta)^2(m_b/m_t)^2$. For the pseudoscalar A the cross section of SM ttH production is multiplied by $\cot^2\beta(m_b/m_t)^2$.

The decay rates of h, H, and A are modified similarly and are given by

$$\Gamma(H_i\to\gamma\gamma) = \frac{\alpha^2_{\mathrm{QED}}m^3_{H_i}}{256\pi^3v^2}\left(|\mathcal{M}_1|^2+|\mathcal{M}_2|^2\right) \tag{5.39}$$

$$\mathcal{M}_1 = 2\sum_f N_c Q_f^2 C^S_{H_iff}\frac{v}{m_f}A^\phi_{1/2}(\tau_f) - C_{H_iWW}\frac{v}{2m_W^2}A^\phi_1(\tau_W)$$

$$- C_{H_iH^+H^-}\frac{v}{2m^2_{H^\pm}}A^\phi_0(\tau^\pm_H) \tag{5.40}$$

$$\mathcal{M}_2 = 2\sum_f NcQ_f^2 C^P_{H_iff}\frac{v}{m_f}A^A_{1/2}(\tau_f) \tag{5.41}$$

where $H_i \in \{h, H, A\}$, $A_0^\phi = -\tau[1 - \tau f(\tau)]$, $A_1^\phi = -[2 + 3\tau + 3(2\tau - \tau^2)f(\tau)]$, and $A_{1/2}^A = 2f(\tau)/\tau$. Figure 5.14 shows the dominant Feynman diagrams for the CP-even Higgs boson decaying to two photons.

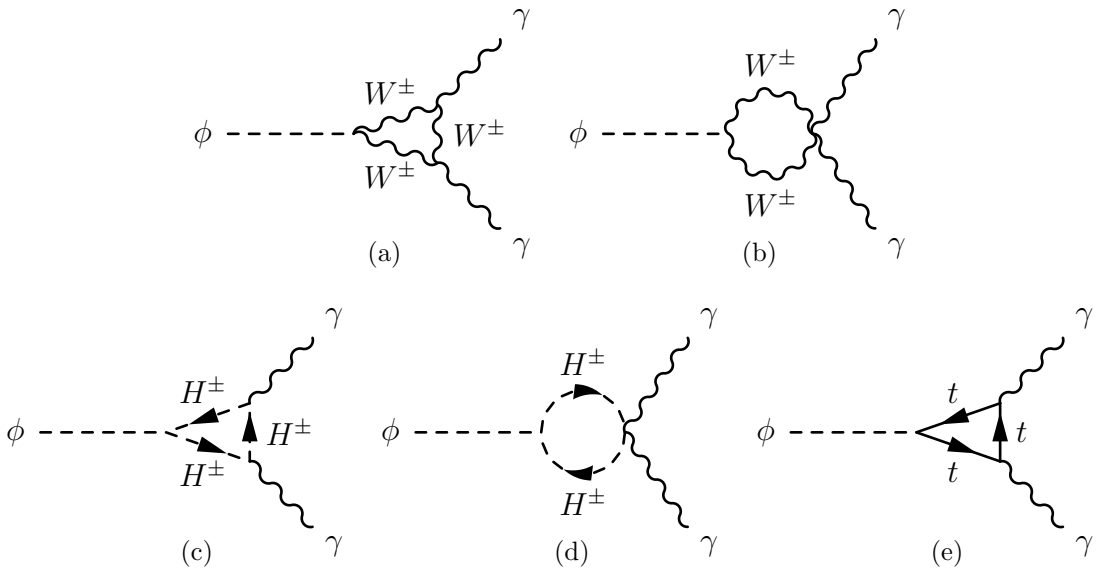


Figure 5.14: Feynman diagrams for the main processes of the decay of the CP-even Higgs bosons of 2HDM; (c), (d), and (e) are the main processes of the decay of the pseudoscalar Higgs boson of 2HDM.

CHAPTER VI

The Reseonance Search Technique and Previous Results

6.1 Bump Hunts

A traditional way of searching for new mass resonances with particle colliders involves looking for local excesses in a mass spectrum, a technique commonly known as a *bump hunt*. Thanks to Lorentz invariance, if the four-momentum $p = (E, \vec{p})$ of a particle is known, the mass of the particle, often called the *invariant mass*, can be easily calculated

$$m^2 = p^2 = E^2 - |\vec{p}|^2 \tag{6.1}$$

in natural units*. From energy and momentum conservation the four-vector of a particle can be determined by adding the four-vectors of its decay products component by component. This allows for the invariant mass of a particle to be determined from measuring the energy and momentum of its decay products.

For simplicity, consider the example of a diphoton mass spectrum produced from the data collected at particle collider. This spectrum is built by taking every event

* Natural units are a set of physical units of measurement which are only based on universal physical constants, e.g. $c = \hbar = 1$

with at least two photons which meet some predetermined conditions and finding the invariant mass of the sum of the four-vectors of the two photons with the largest transverse momenta. It turns out that if nothing interesting is happening in the event, that is the photons come from a soft interaction or processes other than a decaying object, then plotting a large number of these events in a histogram results in a smoothly falling continuum, commonly called the *background*. Now suppose there is a particle X of nominal mass m being produced in the collider with cross section σ and branching fraction $\mathrm{Br}(X \to \gamma\gamma) = f$, that is it decays to diphotons $f \times 100\%$ of the time. If the photons are the decay products of this particle, the invariant mass of the diphoton system $m_{\gamma\gamma}$ will be equal to the mass of the particle, $m_{\gamma\gamma} = m + \delta m$, where δm is deviation from the nominal mass that comes from the inherent width of a decaying particle. Figure 6.1 shows an illustration of a mass resonance with nominal mass m and width $\Gamma \approx \langle \delta m \rangle$. The number of X particles produced by the collider

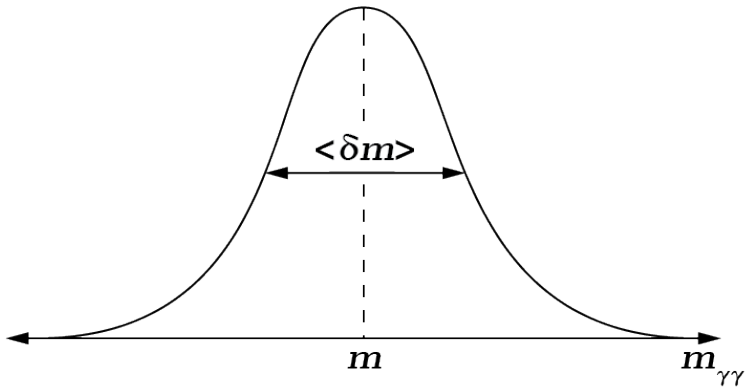


Figure 6.1: Illustration of a resonance with mass m and width $\Gamma \approx \langle \delta m \rangle$.

N_X and number of those which decay to two photons $N_{X\to\gamma\gamma}$ will be given by

$$N_X = L\sigma \tag{6.2}$$

$$N_{X\to\gamma\gamma} = L\sigma f \tag{6.3}$$

where L is the integrated luminosity. This is complicated further since even the

most sophisticated colliders cannot record every event and since the selection criteria for which events to record and use in the analysis will almost certainly exclude some $X \to \gamma\gamma$ events. These shortcomings are captured in a quantity ϵ known as *efficiency*, and is given by

$$\epsilon = \frac{N_{\text{signal}}}{N_{X\to\gamma\gamma}} \tag{6.4}$$

where N_{signal} is the number of $X \to \gamma\gamma$ recorded and used in the $m_{\gamma\gamma}$ spectrum. Combining this with Equation 6.3 gives

$$N_{\text{signal}} = L\sigma f\epsilon. \tag{6.5}$$

Typically the number of events with $m_{\gamma\gamma}$ near m produced by uninteresting events will be orders of magnitude larger than N_{signal}. This creates a difficulty in being able to distinguish a local excess due to a mass resonance from a statistical fluctuation in the mass spectrum. When there is a local excess in the mass spectrum there are statistical and simulation based techniques for estimating the background in the region of the resonance. The deviation from the background model is compared to the expected line shape for new physics. The size of the excess is estimated along with its uncertainty. Traditional benchmarks in high energy physics for evidence and discovery of a new mass resonance are 3σ and 5σ respectively. Chapter IX of this thesis gives the details of how a new data driven technique is used to model the background and search for mass resonances.

6.2 Backgrounds

A conventional way to model a background is to fit the data itself or a Monte Carlo (MC) simulation with a smooth curve. Typically this fit is achieved by binning

the background into a histogram and fitting it with an empirically-derived function. A common choice for a function is the PowLog-n function, defined by

$$f(z;\{a,d,\alpha_i\}) = (1-x^d)^a \cdot x^{\sum_{i=0}^{n}\alpha_i(\log x)^i} \tag{6.6}$$

where $x = z/\sqrt{s}$ and $\sqrt{s}$ is the COM energy of the collider. This function has the property $\lim_{z\to\sqrt{s}} f(z) = 0$. This is a desirable property since the highest diphoton mass possible at a collider is $m_{\gamma\gamma} = \sqrt{s}$. In general the function is very robust and flexible; as more degrees of freedom are needed for a fit, one can simply increase n.

6.3 Signals

Mass resonances are typically modeled using a function that encapsulates the true line shape of the resonance convolved with the detector resolution. A common choice for the dector resolution function is known as the double sided crystal ball (DSCB) (*Amidei*, 2020) function which consists of a Gaussian core with two power law tails:

$$F_{\text{NWA}}(m_{\gamma\gamma}; t, n_{\text{low}}, n_{\text{high}}, \alpha_{\text{low}}, \alpha_{\text{low}}) =$$
$$N \times \begin{cases} e^{-t^2/2} & \Leftarrow -\alpha_{\text{low}} \le t \le \alpha_{\text{high}} \\ \frac{e^{-\alpha_{\text{low}}^2/2}}{\left[\frac{\alpha_{\text{low}}}{n_{\text{low}}}\left(\frac{n_{\text{low}}}{\alpha_{\text{low}}} - \alpha_{\text{low}} - t\right)\right]^{n_{\text{low}}}} & \Leftarrow t < -\alpha_{\text{low}} \\ \frac{e^{-\alpha_{\text{high}}^2/2}}{\left[\frac{\alpha_{\text{high}}}{n_{\text{high}}}\left(\frac{n_{\text{high}}}{\alpha_{\text{high}}} - \alpha_{\text{high}} + t\right)\right]^{n_{\text{high}}}} & \Leftarrow t > \alpha_{\text{high}} \end{cases} \tag{6.7}$$

where $t = (m_{\gamma\gamma} - \mu_{\text{CB}})/\sigma_{\text{CB}}$, μ_{CB} is the peak of the Gaussian core, σ_{CB} is the width of the Gaussian core, N is the normalization parameter, $\alpha_{\text{low}}(\alpha_{\text{high}})$ is the position of the junction between the Gaussian and the power law on the low (high) mass side in units of t, and $n_{\text{low}}(n_{\text{high}})$ is the exponent of the power law. The detector resolution

is usually found with the use of MC simulations of events at the reconstucted level. Very narrow signals, that is signals whose width is much smaller than the detector resolution, are fit with some function, for example the DSCB function, at various resonance masses m_X. The function's parameters are plotted as a function of m_X and fit with polynomials so that one could interpolate the parameters' values at intermediate masses. The detector resolution function for all masses is taken to be the **DCSB!** (**DCSB!**) function with its parameters chosen from these interpolated mass dependent values.

A common choice to model the true lineshape is with the product of a relativistic Breit-Wigner (BW) function and mass dependent factors accounting for the parton luminosity and the matrix elements of the production and decay. This function describing the lineshapes can be validated by comparing it to MC simulations at the parton level. It is convolved with the detector resolution function to give the final signal lineshape. The final signal lineshape can be validated by comparing it to MC simulations on the reconstruction level.

6.4 Previous ATLAS Results

In 2017 there was a previous ATLAS search for new phenomena in the diphoton high mass spectrum. The analysis was performed on 36.7 fb^{-1} data collected at the ATLAS detector between 2016 and 2017. It included both scalar and graviton searches. However, since the analysis presented in this thesis only searches for narrow scalar resonances, the focus of this section will be on the narrow scalar search. The analysis presented in this thesis can be seen as a modified extension of that analysis. Similar selection criteria and signal modeling is used in both searches, however the background modeling techniques differ significantly.

6.4.1 Event Selection

The data sample used uses a the diphoton trigger `HLT_g35_loose_g25_loose`. Each event must contain at least two photon candidates which satisfy the *tight* indentification criteria based on the EM shower shapes. The photon candidates must be central, that is found in the detector region $|\eta| < 2.37$ excluding the transition region $1.37 < |\eta| < 1.52$ between the barrel and the end-cap calorimeters.

Both photon candidates must satisfy isolation requirements using criteria based on the calorimeter only, or on both the inner tracker and the calorimeter. Two variables are constructed to ensure these criteria are met. The first is $E_{\mathrm{T}}^{\mathrm{iso}}$ and is defined as the sum of transverse energy of the topological clusters with positive energy reconstructed in the calorimeter around each photon candidate in a cone of radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$. This sum does not include the contribution from the photon itself and is corrected for the leakage of the photon energy using an event-by-event energy subtraction based on the jet area method (*Cacciari et al.*, 2008)(*Aaboud et al.*, 2017). It is required that $E_{\mathrm{T}}^{\mathrm{iso}} < 0.022 \times E_{\mathrm{T}} + E_{\mathrm{T}}^{\mathrm{iso,th}}$, where the threshold $E_{\mathrm{T}}^{\mathrm{iso,th}}$ is either 2.45 GeV (tight calo-isolation selection) or 7 GeV (loose calo-isolation selection). The second variable is $p_{\mathrm{T}}^{\mathrm{iso}}$ and is defined as the sum of the transverse momenta of all tracks with $p_{\mathrm{T}} > 1$ GeV in a cone of sizes $\Delta R = 0.2$ around each photon candidate, and it is required that $p_{\mathrm{T}}^{\mathrm{iso}} < 0.05 \times p_{\mathrm{T}}$ GeV

The leading photon is required to have $E_{\mathrm{T}} > 40$ GeV while the subleading photon is required to have $E_{\mathrm{T}} > 30$ GeV. No relative E_{T} cut is made in this analysis.

6.4.2 Signal Modeling

Signals are modeled by following the procedure outlined in Section 6.3, that is by convolving the detector resolution with the true lineshape. The detector resolution is modeled with a DSCB function, given by Equation 6.7, while the true lineshape is modeled using the product of a BW function and mass dependent factors ac-

counting for the parton luminosity and the matrix elements of the production and decay processes. The convolution is performed in `RooFit` using the `RooFFTConvPdf` class (*Verkerke and Kirkby*, 2003).

6.4.3 Background Modeling

The background is modeled by following the procedure outlined in Section 6.2. The PowLog-n function, given by Equation 6.6, is chosen to model the background with the d parameter set to 1/2, and $\sqrt{s} = 13$ TeV. This functional form is validated by comparing it to $\gamma\gamma$ MC simulations with varying pdf sets. Extreme cases of the γj and jj contributions to the background are modeled by considering the following functional forms:

$$f(x; k_1, k_2) = x^{k_1(1-\log x)} \times x^{k_2 \log x} \tag{6.8}$$

$$f(x; k_1, k_2, k_3, k_4) = x^{k_1(1-\log x)} \times x^{k_2 \log x} \times \left(1 - \frac{1}{1 + e^{(x-k_3)/k_4}}\right) \tag{6.9}$$

$$f(x; k_1, k_2, k_3, k_4, k_5, k_6) = x^{k_1(1-\log x)} \times x^{k_2 \log x} \times \left(1 - \frac{1}{1 + e^{(x-k_3)/k_4}}\right) \times \left(1 - \frac{1}{1 + e^{(x-k_5)/k_6}}\right) \tag{6.10}$$

The combined sample is used to perform spurious signal tests to further validate the chosen PowLog-n functional form. This builds confidence that the selected functional form can be used to describe the background regardless of its composition.

6.4.4 Results

To search for a narrow scalar resonance, the data was fit simultaneously with a PowLog-n function describing the background and the signal shapes described in Section 6.4.2 for many mass points across the range 180-3000 GeV. At each mass point a yield, N_{signal}, is extracted as well as the local p-value for the background only hypothesis. Figure 6.2 shows the local p-values found as a function of mass. As no

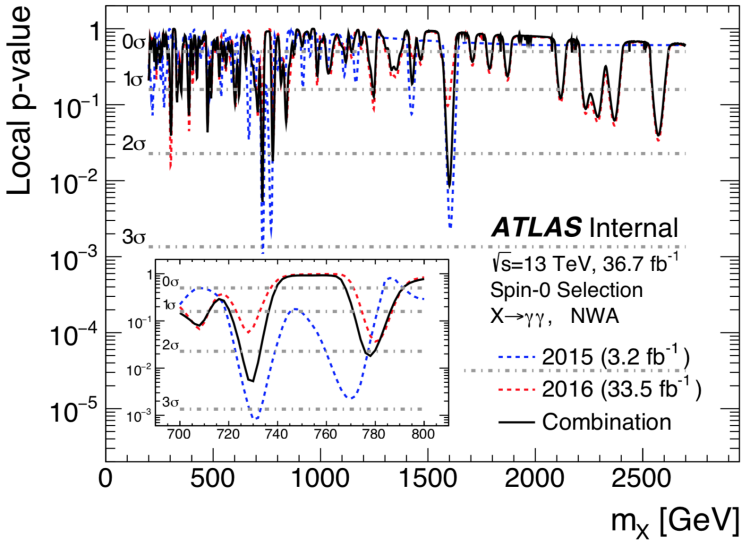


Figure 6.2: Local p-value as a function of resonance mass, m_X, from the 2017 ATLAS narrow width scalar search.

significant excesses were found the focus of the analysis shifted from discovery to setting the 95% confidence level (CL) upper limits on $\sigma_{\text{fid}} \times \text{Br}$ for a narrow width scalar. Figure 6.3 shows the limits as a function of m_X.

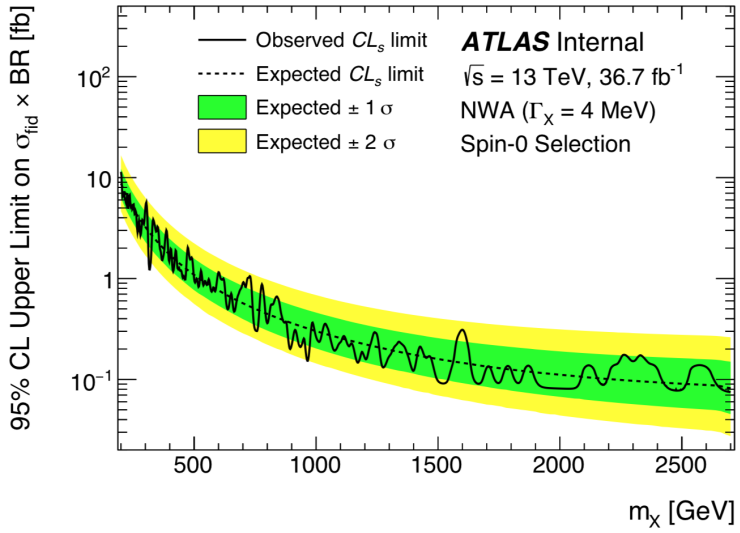


Figure 6.3: The 95% confidence level upper limits on $\sigma_{\text{fid}} \times \text{Br}$ for the narrow width scalar search.

CHAPTER VII

Simulated Samples

This chapter describes the simulated samples used in the analysis presented in this thesis, which includes both MC samples and pseudo-data. These samples are used to model the shape of mass spectra for signal and background processes and to reconstruct the efficiency and accpetance of signal processes. To produce the MC samples, events are first produced at the parton level using MC event generators. Next, they are passed to GEANT4 to simulate the particles passing through the ATLAS detector and reconsturcted using the same analysis chain used for collision data. Using PYTHIA8 with the A14 parameter tune, pileup is included by adding inelastic proton–proton (pp) collisions. Finally, these samples are generated using mc16a, mc16d, and mc16e conditions to account for the different pileup and trigger conditions of the 2015-2016, 2017, and 2018 data, respectively (*Amidei*, 2020). Once the MC samples are generated, their shapes can be used to produce unbinned pseudo-data, or *toy datasets*.

7.1 Background

7.1.1 Monte Carlo Background Samples

Background simulations are produced to assist in both the primary (empirically-derived function) and secondary (FD) analyses. Although the background modeling

in FD is data driven, the shape of the simulated background is used to produce unbinned pseudo-data samples which in turn are used to validate FD as a discovery tool.

Using the pdf set NNPDF3.0 NNLO, and SHERPA and its default tuning for the underlying event, background events from the continuum $\gamma\gamma$ production are generated. The interference effects with the $X \to \gamma\gamma$ signals are not simulated since they are estimated to be negligible. The matrix elements are determined in a two part process. First, the matrix elements are calculated at NLO in the strong coupling constant α_s for zero or one real emission of an additional parton and at LO for two and three additional partons. Next, they are merged with the SHERPA parton shower (*Schumann and Krauss*, 2008) according to the ME+PS@NLO prescription (*Höche et al.*, 2013).

The background is generated in *slices* of $m_{\gamma\gamma}$ in order to provide sufficient statistics in all mass regions. The mass ranges are exclusive to each slice, that is there is no overlap, except between the 175-2000 GeV slice and 1400-2000 GeV slice. However, events with $m_{\gamma\gamma} > 1400$ GeV on the parton level are removed from the 175-2000 GeV slice to prevent double counting. A fast simulation (*Aad et al.*, 2010) is used to generate these samples in which the full simulation of the calorimeter response is replaced with a parameterization (*ATLAS et al.*, 2010). Full simulation is used to produce smaller samples in order to validate the results of the fast simulation.

Generator	$m_{\gamma\gamma}$ range [GeV]	Cross section [pb]	N_{events}
SHERPA+Fastsim	90-175	51.822	50.8M
	175-2000	10.999	35.3M
	1400-2000	3.992×10^{-3}	400k
	2000-∞	703×10^{-6}	400k
SHERPA+Fullsim	175-2000	10.999	14.1M
	2000-∞	703×10^{-6}	400k

Table 7.1: The generators used, mass range, cross section, and number of events in the background MC samples.

7.1.2 Background Pseudo-data

The reweighted MC samples are used to produce unbinned toy datasets with the expected statistics and shape of 140 fb^{-1}. After the event selection is applied, the MC samples are fit with a PowLog-1 function, given by Equation 6.6, which is normalized to unit area, as to produce a pdf. The parameterization of the PowLog-1 fit is shown in Table 7.2.

Parameter	Value
a	8.33
α_0	-2.97
α_1	1.18×10^{-2}

Table 7.2: The parameters extracted from the PowLog-1 fit to the SHERPA $m_{\gamma\gamma}$ distribution in 10 GeV bins

Inversion sampling is used on the pdf to produced 5000 toys each with exactly 433655 events starting at $m_{\gamma\gamma} = 150$ GeV. Figure 7.1 shows one such toy compared to the PowLog-1 shape. The $\chi^2/\text{dof} = 2.22$ is thought to be due to the poor statistics

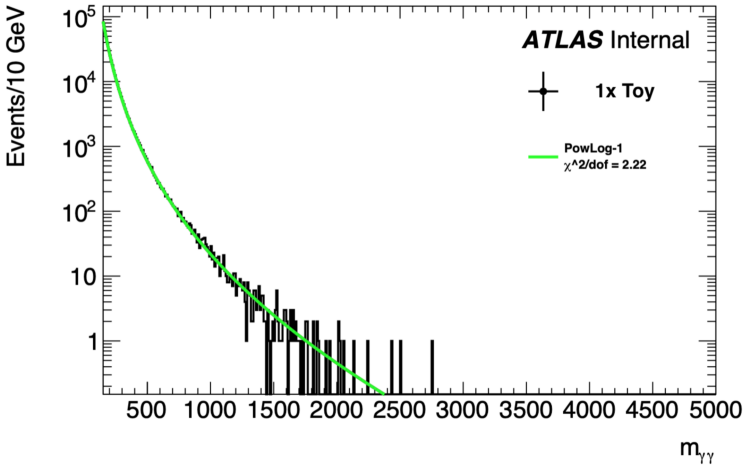


Figure 7.1: One toy dataset's shape compared to the PowLog-1 pdf and the pull distribution.

of the tail.

This chapter now turns to the details of the resonance search in this thesis.

7.2 Scalar

7.2.1 Monte Carlo Samples

The scalar signal models used in this analysis are simulated Higgs-like particles with various masses in the range $m_X \in [200, 5000]$ GeV. The particles are required to decay to two photons, and are simulated using Higgs boson production processes in pp collisions at $\sqrt{s} = 13$ TeV.

The scalar samples are divided into two categories: narrow width approximation (NWA) and large width (LW). In the NWA the resonance has a fixed width of $\Gamma_X = 4.07$ MeV, the width of the 125 GeV Higgs boson. The ggF production mode of the resonance is chosen as the baseline signal process for signal parameterization, as this is the dominant production mode for a Higgs-like particle. The interference between the $gg \to X \to \gamma\gamma$ process and the continuous QCD diphton production associated with the $gg \to \gamma\gamma$ process is estimated to be negligible, and therefore is neglected in the simulation. Additional samples with other production modes are also considered in the study of the signal shape and efficiency. Several widths are used for the LW scalars, namely 2%, 6%, and 10% of the resonance mass, m_X for the ggF production process only. These samples are produced with widths in the range $\Gamma_X/m_X \in [0.02, 0.10]$. Table 7.3 shows the various production modes and the MC generators used for the scalar samples.

7.2.2 Narrow Width Scalar Pseudo-data

The ggF MC samples for the NWA are used to produce toy datasets which are used in the signal injection studies presented in Section 9.8.1.5. Rather than fitting the samples with a function, the samples are binned and normalized to unit area so that the resulting histogram defines the pdf from which to sample. This luminosity weighted histogram for the $m_X = 1000$ GeV sample is shown in Figure 7.2. The size

Process	Generator	Width (Γ/m_X)
ggF	MG5_aMC@NLO+Pythia8	NWA
ggF	Powheg+Pythia8	NWA
VBF	Powheg+Pythia8	NWA
WH	Pythia8	NWA
ZH	Pythia8	NWA
ttH	Pythia8	NWA
ggF	MG5_aMC@NLO+Pythia8	2%,6%,10%

Table 7.3: The production processes and MC generators used for the scalar samples. The samples are generated for the masses 200, 400, 800, 1000, 1200, 1600, 2000, 2400, 3000, 4000, and 5000 GeV. The pileup configurations mc16a+d+e are generated for 30,000, 40,000, and 70,000 events respectively, to roughly match the proportions of the various pileup conditions in the collider data.

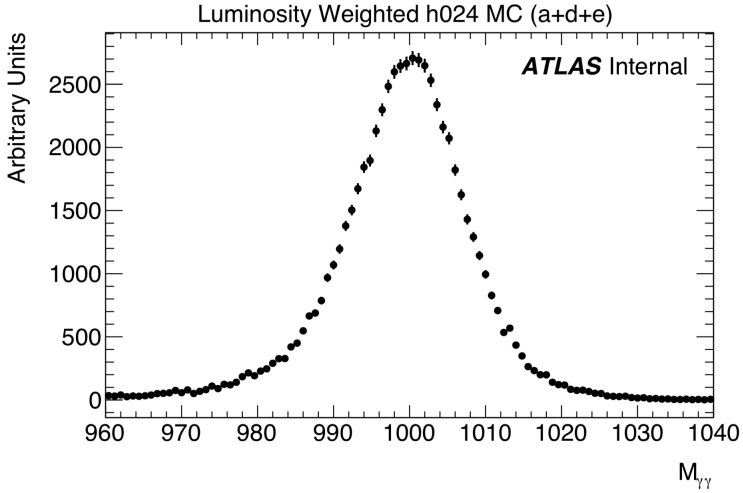


Figure 7.2: Luminosity weighted histogram for the $m_X = 1000$ GeV NWA scalar signal

of these toys are determined by estimating the signal size required to for 1,2,3,4 and 5 σ excesses near the mass of the corresponding signal. For each trial in the signal injection studies a unique toy signal is produced.

7.3 Graviton

The graviton samples used in this analysis were generated using RS graviton production processes in pp collisions at $\sqrt{s} = 13$ TeV, and requires a diphoton final

state. They are generated using PYTHIA8with the NNPDF23LO pdf set and the A14 tune. Only the effect of the first KK excitation is considered. These samples are produces for various coupling values in the range $k/\overline{M}_{\mathrm{Pl}} \in [0.01, 0.3]$ and masses in the range $m_X \in [500, 7000]$ GeV. Table 7.4 gives the full list of the graviton samples.

Generator	Mass [TeV]	$k/\overline{M}_{\mathrm{Pl}}$
PYTHIA8	0.5,1,2,3,4	0.01,0.05,0.1
PYTHIA8	5,6,7	0.01,0.05,0.1,0.2,0.3

Table 7.4: The MC generators used to produce the graviton samples with their masses and coupling strengths $k/\overline{M}_{\mathrm{Pl}}$. In all cases mc16a,d,and e are used to generate 20k, 30k, and 60k respectively in order to roughly match the proportions of the various pileup conditions in the collider data.

CHAPTER VIII

Diphoton Analysis Selection, Background Modeling, and Signal Modeling

Previously in this thesis diphoton production, photon reconstruction, and photon calibration were discussed, the focus shifts now to how the calibrated photons were used in the analysis. This chapter begins by discussing the event selection used in the analysis presented in this thesis. From there it moves to describing the acceptance and efficiency of the chosen selectria criteria. The chapter ends with describing how the signal shapes, now sculpted by the selection, are modeled in the analysis.

8.1 Collision Data and Trigger

The data used in the analysis was recorded with the lowest E_{T} unpreselected diphoton trigger with the least stringent requirements on photon identification that operated during the 25 ns 2015-2018 data-taking period. The high level trigger (HLT) `HLT_g35_loose_g25_loose` was used to obtain the 2015-2016 dataset. This trigger is seeded from a Level-1 (L1) trigger which requires two distinct energy deposits of more than 15 GeV in the ECal. The HLT requires E_{T} thresholds of 35 GeV and 25 GeV for the leading and subleading photons and that associated electromagnetic clusters match the *loose* shower shape criteria. The `HLT_g35_medium_g25_medium_L12EM20VH`

was used to obtain the 2017-2018 dataset. This trigger is seeded from an L1 trigger requiring two energy deposits of more than 20 GeV. The HLT requires E_{T} thresholds of 35 GeV and 25 GeV for the leading and subleading photons and that the clusters match the *medium* shower shape criteria. The tightened selection during the 2017-2018 was enforced to compensate for the increase in pile-up which can be seen in Figure 8.1.

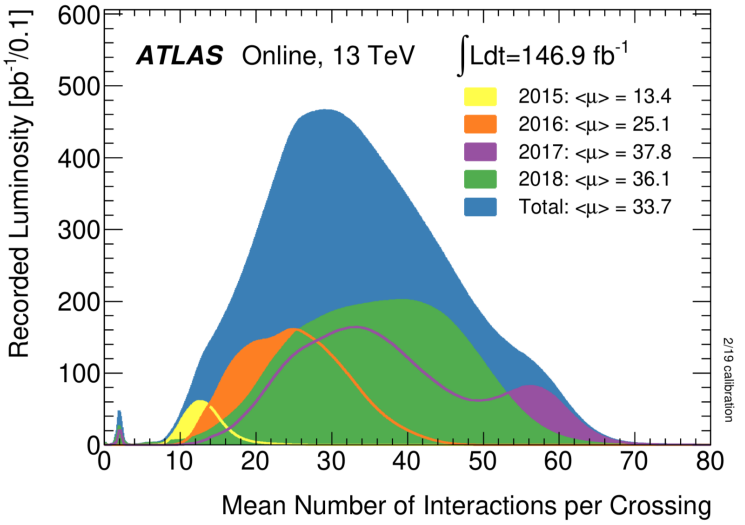


Figure 8.1: Distribution of the number of interactions per bunch crossing, weighted by luminosity for each of the data taking periods. All data recorded by ATLAS during stable beams is shown.

8.2 Event Selection

Photon Identification

As discussed in Section 4.2.4, there are several variables related to the calorimeter shower shape which are used to classify photon candidates. The *loose* selection is based on the shower shape in the second layer of the ECal and the energy deposited in the HCal. The *tight* selection uses the same information plus information from the finely segmented strip layer of the calorimeter. The analysis requires that the leading and subleading photon both satisfy the tight indentification criteria.

Identification Name	Variables Used in Identification
Loose	$R_{\text{had1}}, R_{\text{had}}$, R_{η}, $w_{\eta 2}$
Tight	Loose + R_{ϕ}, w_{s3}, F_{side}, ΔE, $w_{s1,\text{tot}}$, E_{ratio}
LoosePrime-2	Tight − w_{s3}, F_{side}
LoosePrime-3	Tight − w_{s3}, F_{side}, ΔE
LoosePrime-4	Tight − w_{s3}, F_{side}, $\Delta E, w_{s1,\text{tot}}$
LoosePrime-5	Tight − w_{s3}, F_{side}, $\Delta E, w_{s1,\text{tot}}$, E_{ratio}

Table 8.1

Diphoton Vertex

Knowing the precise location of the diphoton production vertex is necessary in order to make precise measurements of the diphoton invariant mass $m_{\gamma\gamma}$ and track-based quantites such as isolation. The location of the vertex is determined using the *photon pointing* method. By combining the trajectories of the two photons, measured using the longitudinal segementation of the calorimeter, the vertex position along the beam axis is determined. In the case of converted photons with track hits in the silicon detectors, the conversion vertex is also used (*Lenzi and Delgove*, 2015). To select the diphoton production vertex from among all the reconstructed primary vertices, a Neural Network algorithm is used. The probabilty of choosing a primary vertex within 0.3 mm of the true vertex is expected to occur more than 80% of the time (*Lenzi and Delgove*, 2015).

Kinematic Selections

Additional kinematic contrainsts are placed on the diphoton system. The leading photon is required to have $E_{\text{T}}/m_{\gamma\gamma} > 0.3$ while the subleading photon is required to have $E_{\text{T}}/m_{\gamma\gamma} > 0.25$. Since the analysis only considers $m_{\gamma\gamma} > 150$ GeV, these requirements at minimum restrict the leading photon to have $E_{\text{T}} > 45$ GeV and the subleading photon to have $E_{\text{T}} > 37.5$ GeV.

Photon Isolation

The leading and subleading photon are both required to be isolated based on criteria from the calorimeter and the inner detector. This criteria is known as the `FixedCutTight` working point. There are two isolation variables considered, $E_{\mathrm{T}}^{\mathrm{iso}}$ and $p_{\mathrm{T}}^{\mathrm{iso}}$. The first variable, $E_{\mathrm{T}}^{\mathrm{iso}}$, is defined as the sum of the transverse energy of the topological clusters with positive energy reconstructed in the calorimeter in a cone of radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$ around the photon (*Lampl et al.*, 2008). After corrections are made, such as subtracting the energy of the core as to not count the energy of the photon itself, it is required that $E_{\mathrm{T}}^{\mathrm{iso}} < 0.022E_{\mathrm{T}} + 2.45$ GeV where E_{T} is the transverse energy of the photon. The second variable, $p_{\mathrm{T}}^{\mathrm{iso}}$, is the scalar sum of the transverse momenta of all tracks with $p_{\mathrm{T}} > 1$ GeV in a cone of radius $\Delta R = 0.2$ (*Lampl et al.*, 2008) around the photon, and it is required that $p_{\mathrm{T}}^{\mathrm{iso}} < 0.05E_{\mathrm{T}}$ after similar corrections are made.

Cutflow

Table 8.2 shows the number of data events from the 139 fb^{-1} sample after each step in the event selection.

8.3 Background Modeling

The strategy for modeling the invariant mass spectrum is employed by summing two spectra, one representing the smooth delocalized backgournd and the other any localized excesses. Each component of the spectra, the background and signal, is modeled with its own ad hoc function. By adjusting the function parameters the signal plus background (S+B) best fit is found and a signal yield is extracted.

Sample	Scalar NWA	Graviton	Data 139 fb^{-1}
	efficiency (abs./rel.)	efficiency (abs./rel.)	event yield
All events (DAOD)	1.000/1.000	1.000/1.000	364.97M
GRL	-	-	357.42M
Trigger presel.	-	-	304.26M
Detector DQ	-	-	304.24M
Primary vertex	1.000/1.000	1.000/1.000	304.24M
2 loose photons	0.767/0.767	0.677/0.677	86.42M
Trigger match	0.727/0.948	0.645/0.953	58.31M
Tight ID	0.677/0.931	0.598/0.927	15.00M
Isolation	0.612/0.904	0.537/0.898	5.93M
$m_{\gamma\gamma} > 150$ GeV	0.612/1.000	0.537/1.000	746896
Rel. E_{T}	0.535/0.874	0.419/0.780	433655

Table 8.2: Effect of the event selection on a NWA scalar and a graviton ($k/\overline{M}_{\mathrm{Pl}} = 0.01$) MC sample generated for $m_X = 1$ TeV and in the data. For the MC samples, the efficiency is shown relative to the total event yield after applying event weights (absolute efficiency) and also relative to the event yield before each selection (relative efficiency). For data, the absolute yields are shown. The initial yields for data at derivation level include a trigger preselection that is the OR of a long list of single photon and diphoton triggers. They also include a duplicate event removal (less than 200 events overall). The "2 loose photons" step includes the kinematic acceptance cuts. The trigger matching indirectly requires medium ID and loose isolation criteria for data 2017-2018 and mc16d-mc16e.

8.3.1 Background Decomposition

8.3.1.1 2×2D Sideband Method

The 2×2D sideband method (2x2D) is used to estimate the relative fractions of the reducible and irreducible backgrounds used in the analysis (*Carminati et al.*, 2012). Two requirements of the signal selection are loosened, namely, the analysis level isolation criteria is removed and the photon identification criteria is relaxed. The observed yield for this sample, W_{tot}, is given by the sum of the diphoton *signal yield*, $W_{\gamma\gamma}$, and the unknown *background yields*, $W_{\gamma j}$, $W_{j\gamma}$, and W_{jj}, that is

$$W_{\mathrm{tot}} = W_{\gamma\gamma} + W_{\gamma j} + W_{j\gamma} + W_{jj}. \tag{8.1}$$

Events are divided into orthogonal categories based on whether each photon in each event passes or fails the isolation and the tight identification selection criteria, giving $2 \times 2^2 = 16$ subsamples. One of these subsamples (both photons pass both criteria) is the signal region used in the analysis, while the other 15 are used as control regions in the 2x2D method. In each of the subsamples, the yield can be written as a function of the signal and background yields, $W_{\gamma\gamma}$, $W_{\gamma j}$, $W_{j\gamma}$, and W_{jj}, the identification and isolation efficiency for prompt photons passing the loosened criteria, and $\epsilon_{\rm ID}$ and $\epsilon_{\rm iso}$, identification and isolation fake rates, $f_{\rm ID}$ and $f_{\rm iso}$. For jj events, the correlations between the isolation distribution, $\xi_{{\rm iso}jj}$, is non-neglible and is in considered. This system of equations can be inverted to give the observed yield in each of the 16 subsamples, for example

$$\begin{aligned} N_{\rm analysis} &= W_{\gamma\gamma}\epsilon_{\rm ID1}\epsilon_{\rm iso1}\epsilon_{\rm ID2}\epsilon_{\rm iso2} \\ &+ W_{\gamma j}\epsilon_{\rm ID1}\epsilon_{\rm iso1}f_{\rm ID2}f_{\rm iso2} \\ &+ W_{j\gamma}f_{\rm ID1}f_{\rm iso1}\epsilon_{\rm ID2}\epsilon_{\rm iso2} \\ &+ W_{jj}f'_{\rm ID1}f'_{\rm iso1}f'_{\rm ID2}f'_{\rm iso2}\xi_{{\rm iso}jj} \end{aligned} \tag{8.2}$$

gives the observed yield in the analysis selection subsample where

- $\epsilon_{\rm ID1(2)}$ is the efficiency of the analysis isolation criteria for the leading(subleading) photon and is determined from simulations;
- $\epsilon_{\rm iso1(2)}$ is the efficiency of the analysis identification criteria for the leding(subleading) photon and is determined from simulations;
- $f_{\rm ID1(2)}$ is the fake rate of the analysis isolation criteria for the leading(subleading) photon and is determined directly from fits to the data;
- $f_{\rm iso1(2)}$ is the fake rate of the analysis identification criteria for the leading(subleading) photon and is determined directly from fits to the data;

- $f'_{\mathrm{ID1(2)}}$ is the fake rate of the analysis isolation criteria for the leading(subleading) photon in jj events and is determined directly from fits to the data;

- $f'_{\mathrm{iso1(2)}}$ is the fake rate of the analysis identification criteria for the leading(subleading) photon in jj and is determined directly from fits to the data.

- $\xi_{\mathrm{iso}jj}$ is the isolation correlation factor between the jets in the jj events and is determined directly from fits the the data.

The observed yields from the remaining 15 subsamples can be defined similarly.

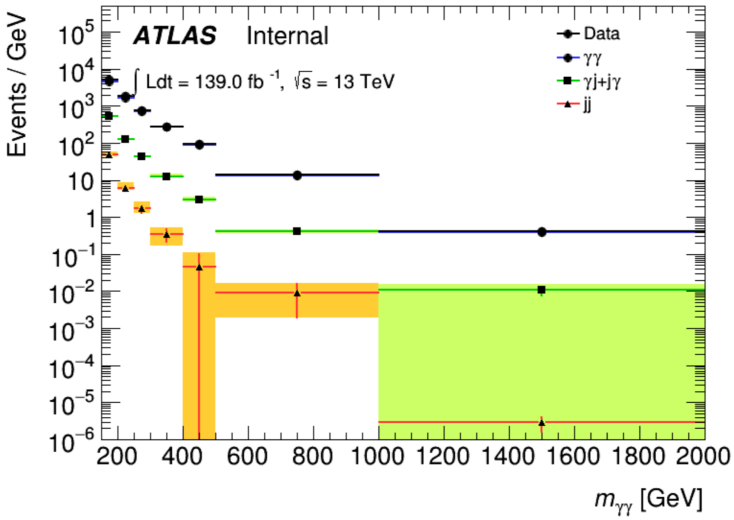
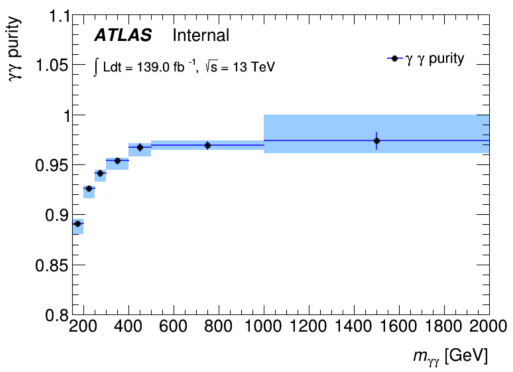


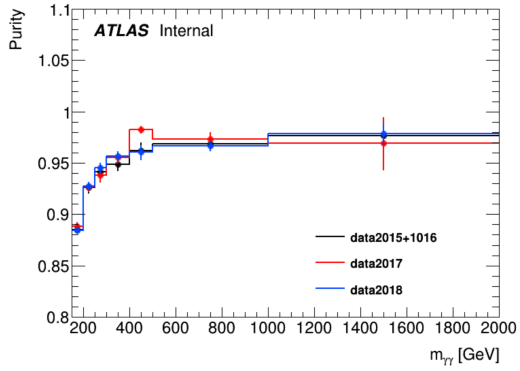
Figure 8.2: The $\gamma\gamma$, $\gamma j + j\gamma$, and jj yields determined by the 2×2D sideband method as a function of the diphoton invariant mass (*Amidei*, 2020).

8.4 Signal Modeling

This section presents the methodology for signal modeling used in the search presented in this thesis. There are three lineshapes considered in the search corresonding to a narrow width scalar, a large width scalar, and a large width graviton. The signals are built by convolving the detector resolution function, a parameterized DSCB, with the true lineshape (*Amidei*, 2020).



(a) Results with statistical errors (error bars) and systematic errors associated with the variation of the loosened identification criteria (rectangles)



(b) Results with statistical error split into data-taking periods

Figure 8.3: Purity of the data sample in prompt diphotons as obtained by the 2×2D sideband decomposition method (*Amidei*, 2020).

8.4.1 Narrow Width Scalar

For the case of a narrow width scalar, the NWA is used. In the NWA for a resonance of mass m_X, the true lineshape is modeled as a Dirac delta function centered around m_X. Therefore convolving it with the detector resolution gives back the detector resolution at mass m_X. The detector resolution function is modeled using a parameterized DSCB function, given by equation 6.7.

The parameters of the DSCB are found empirically using a multistep process. First, the NWA signal samples, described in Section 7.2.1, are fit with the DSCB function, as shown in Figure 8.4. Next, these parameters are parameterized as linear functions of the normalized mass $m'_X \equiv (m_X - 100)/100$ GeV, as shown in Figure 8.9. The normalized mass is used rather than m_X because it leads to better numerical and fit stability. Table 8.3 shows the parameterizations of the NWA DSCB as a function of m'_X and where $\mu_{\text{CB}} = m_x + \Delta m_x$. The parameterized DSCB function is validated by comparing it to the scalar NWA MC samples and can be seen in Figure 8.6.

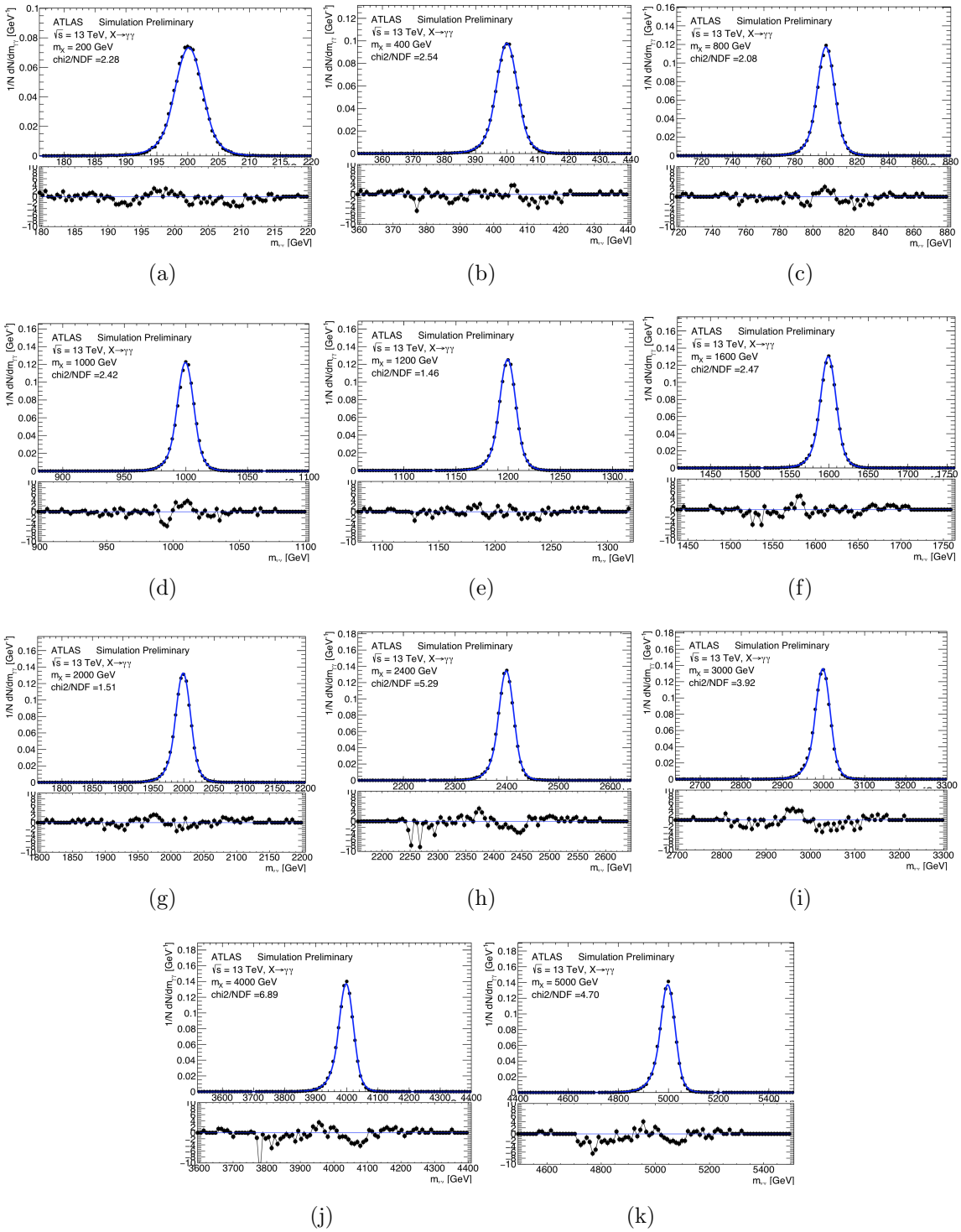


Figure 8.4: Mass spectrum for the NWA as a function of $m_{\gamma\gamma}$ for various m_X. Each mass distribution represents the detector resolution for the NWA and is fit with the DSCB function.

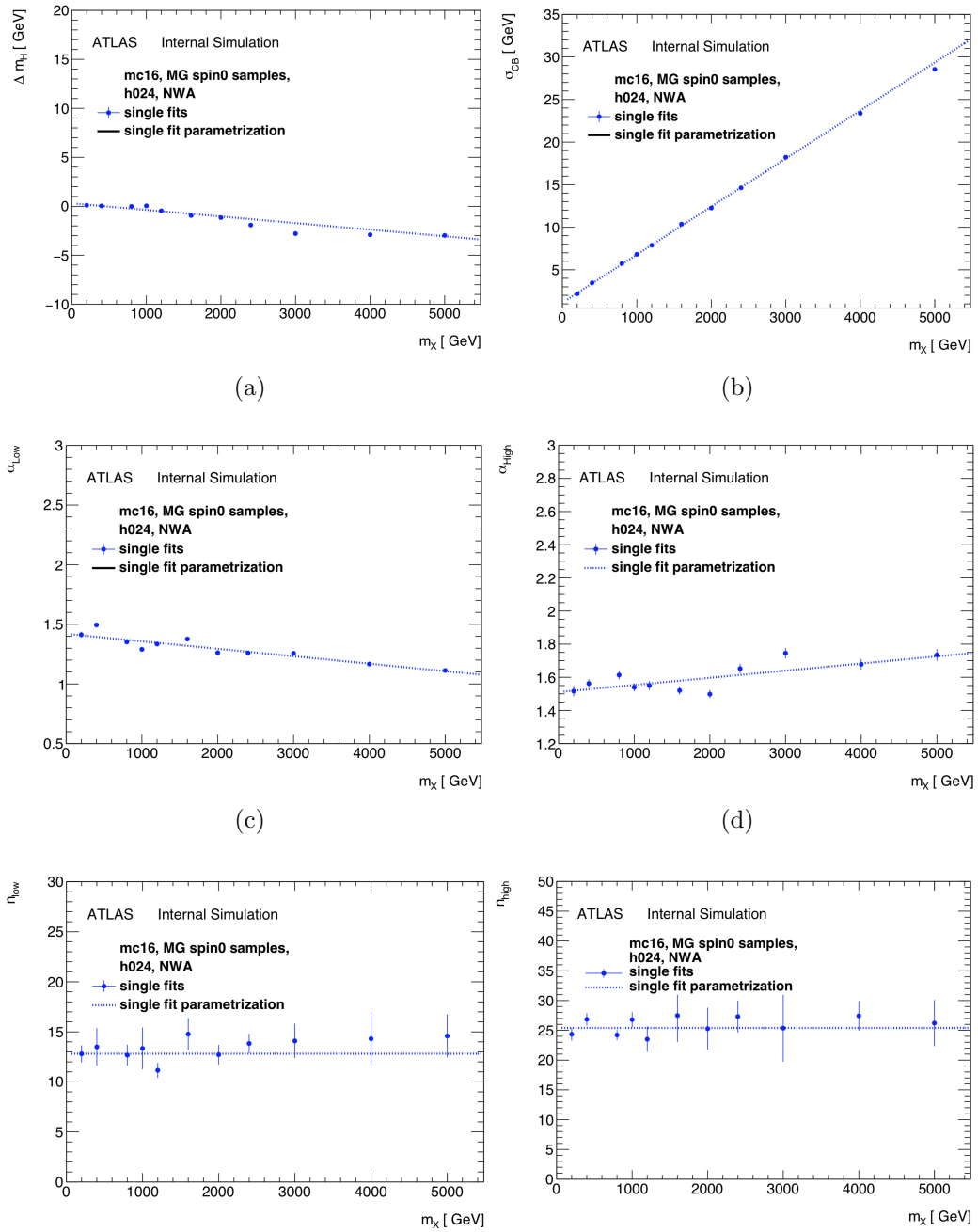


Figure 8.5: The scalar NWA DSCB parameters as a function of m_X. Each plot is fit with a linear function and the intermediate values are interpolated as the straight line.

8.4.2 Large Width Scalar

For the LW scalar signals, the MC samples are constructed by convolving the detector resolution function, that is the parameterized DSCB function for the NWA,

DSCB parameter	Parameterization
Δm_X	$0.238 - 0.0672 m'_X$
σ_{CB}	$1.700 + 0.564 m'_X$
α_{low}	$1.414 - 0.00626 m'_X$
α_{high}	$1.525 + 0.00432 m'_X$
n_{low}	12.0 (const.)
n_{low}	25.1 (const.)

Table 8.3: DSCB function parameters found for the NWA scalar samples in terms of the normalized mass m'_X.

with the lineshape describing the LW resonance:

$$F_{\text{LW}} = F_{\text{NWA}}(m_{\gamma\gamma}; m_X) \cdot m_{\gamma\gamma}^7 \cdot \mathcal{L}_{gg} \cdot F_{\text{BW}}(m_{\gamma\gamma}; \Gamma_X) \tag{8.3}$$

$$F_{\text{BW}} = \left((m_{\gamma\gamma}^2 - m_X^2)^2 + (m_{\gamma\gamma}\Gamma_X)^2\right)^{-1} \tag{8.4}$$

where F_{BW} is the relativistic BW function with width Γ_X and $\mathcal{L}_{gg}$ is the gluon-gluon luminosity as a function of $m_{\gamma\gamma}$. This luminosity is parameterized using the normalized PowLog-0 function, that is

$$f(m_{\gamma\gamma}; d, a, \alpha, N) = N \cdot (1 - x^d)^a \cdot x^\alpha \tag{8.5}$$

where $x = m_{\gamma\gamma}/\sqrt{s}$, $\sqrt{s} = 13$ TeV and $d = 1/3$ is fixed by choice. The free PowLog-0 parameters were found by fitting the pdf set used to generate the samples at NLO and are given in Table 8.4. The convolution is implemented in `RooFit` (*Verkerke and*

Parameter	Value
a	11.6566
α	-2.55713
N	2.09254×10^{-6}

Table 8.4: The PowLog-0 parameters derived from the NNPDF3.0 pdf set for the parton luminosity, $\mathcal{L}_{gg}$, for the LW scalar MC samples.

Kirkby, 2003) using the `RooFFTConvPdf` class, which uses the Fast Fourier Transform (FFT) (*Amidei*, 2020). To validate the convolutions, they are compared to some

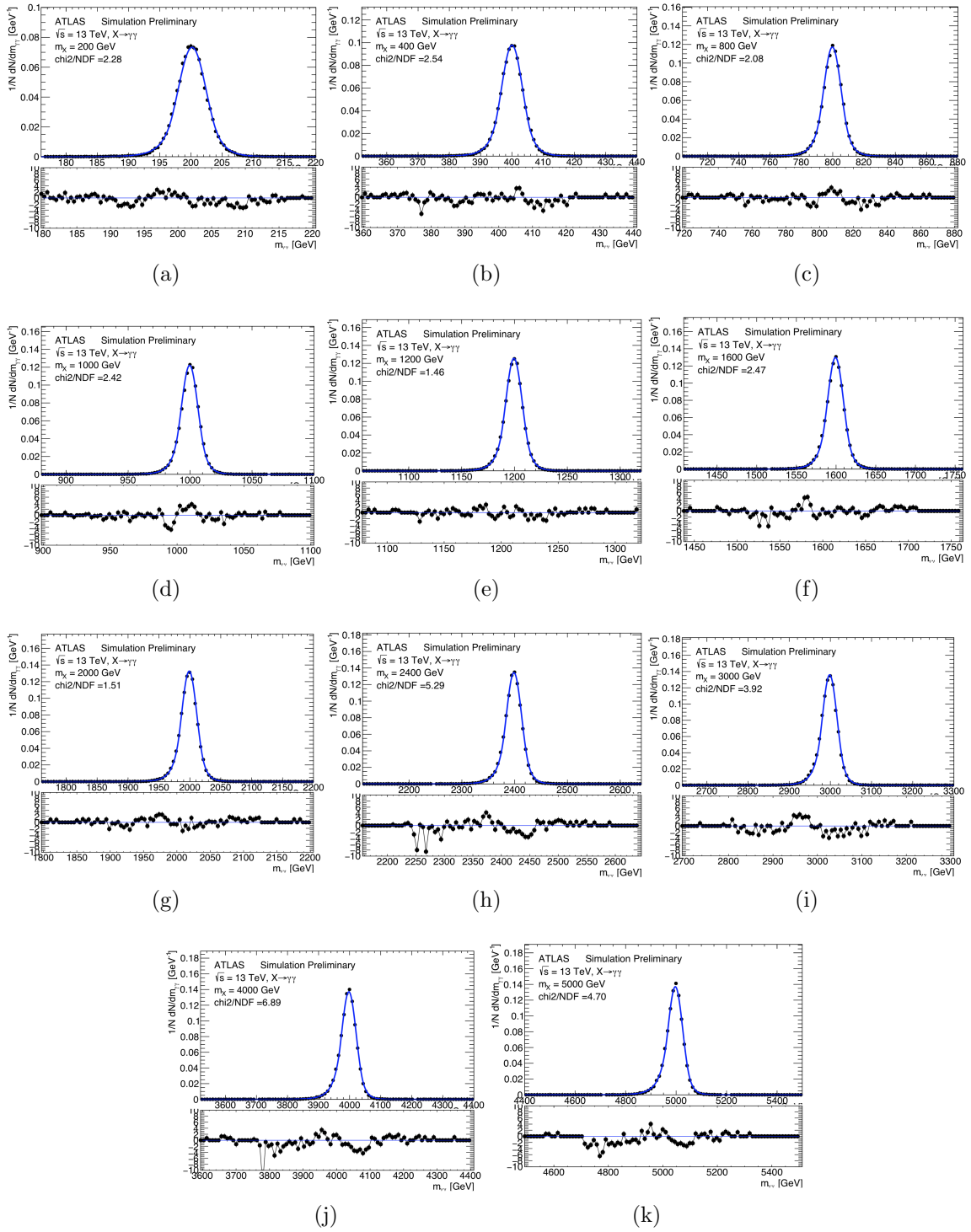


Figure 8.6: Validation plots for the scalar NWA lineshape.

representative mass and width combinations of the MC samples chosen in the low, mid, and high mass regions. These comparisons can be seen in Figure 8.7.

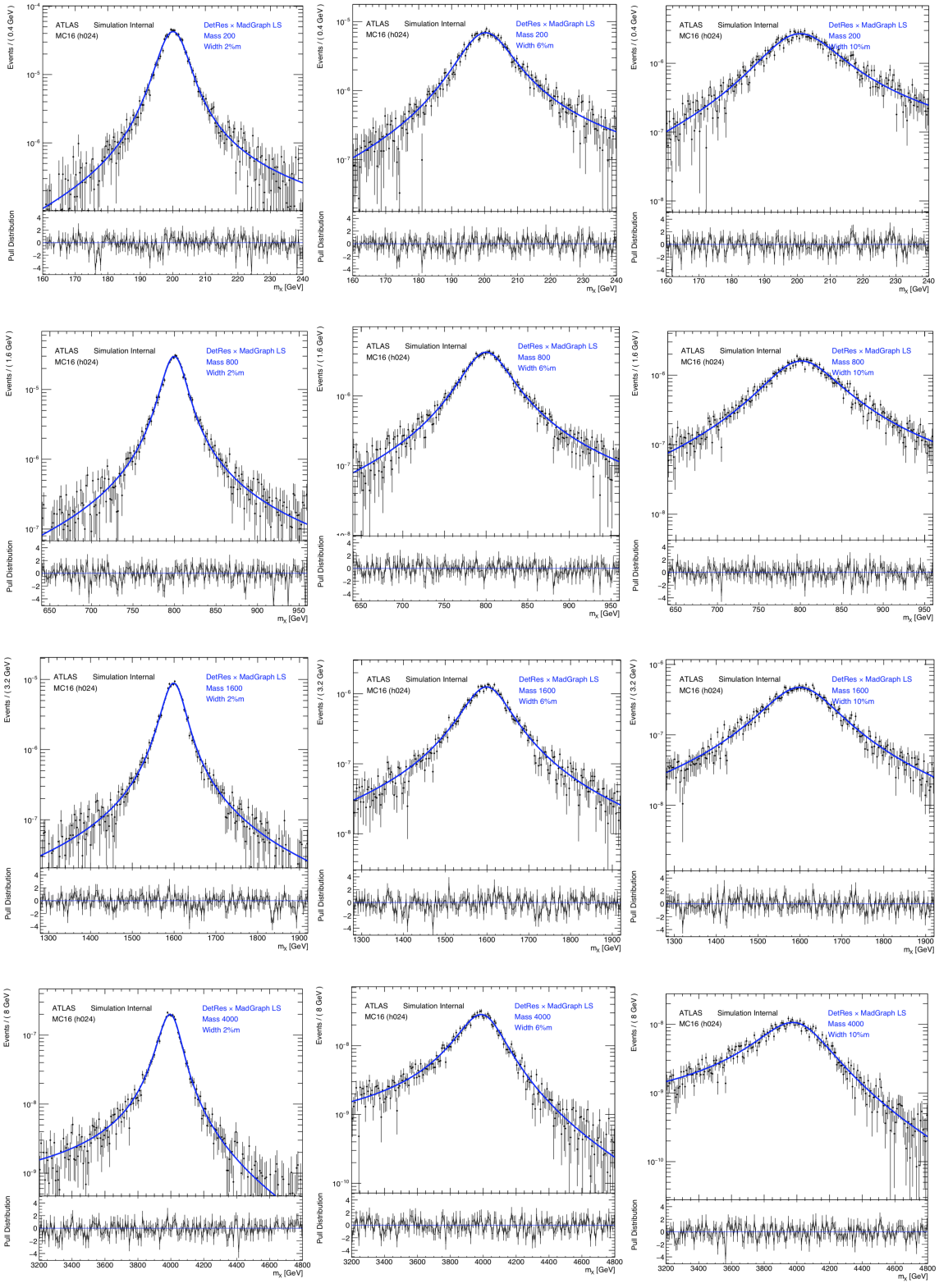


Figure 8.7: Validation plots for the LW scalar signal convolutions

8.4.3 Graviton

The lineshape of the graviton is modeled following the same procedure as the scalar signal. Since the graviton is a spin-2 particle, its decays expected to have a different angular distribution and therefore different kinematics than a scalar, the detector resolution function for the graviton is found using a very narrow, $k/\overline{M}_{\mathrm{Pl}} = 0.01$, resonance. Like with the scalar case, the mass distribution for these samples are fit with a DSCB function, and the DSCB parameters are parameterized as function of the reduced mass. The DSCB function fits can be seen in 8.8. Figure 8.9 shows

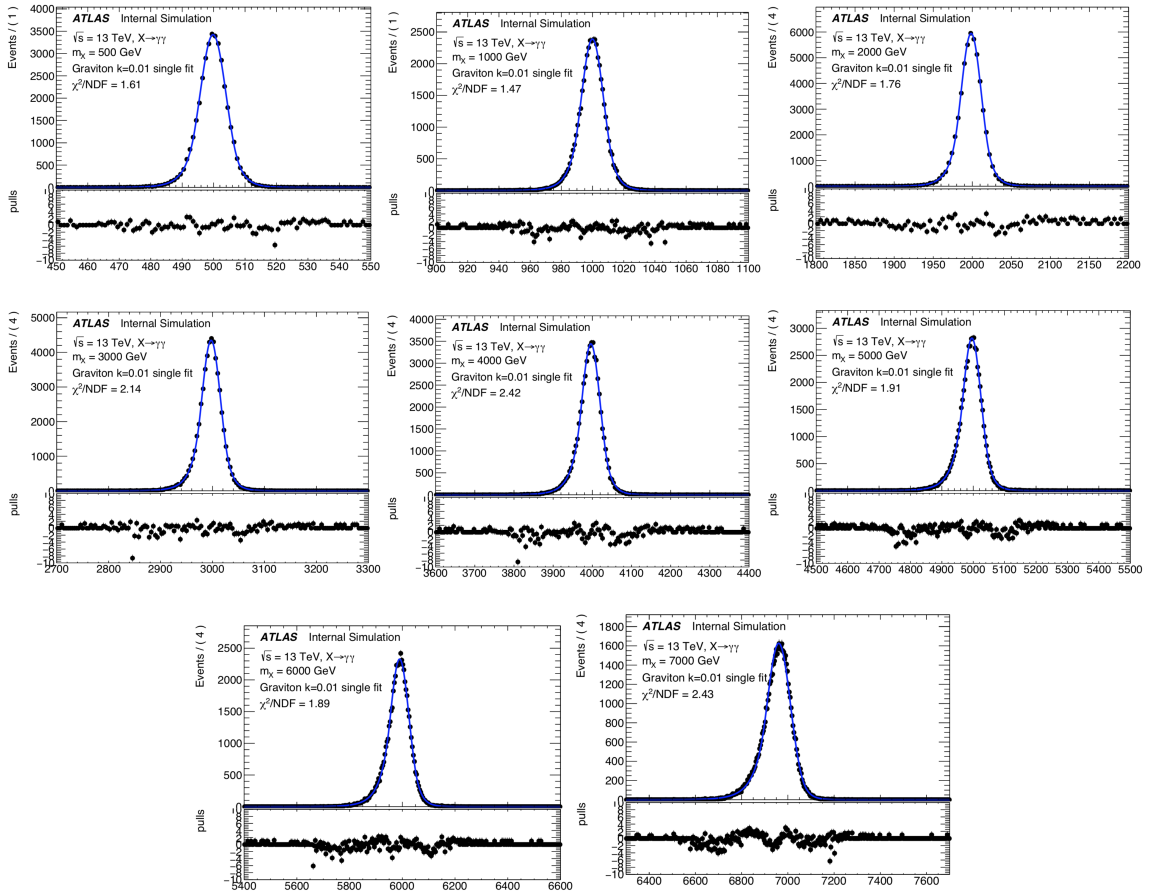


Figure 8.8

the mass dependent fits used for the parameterization of the DSCB parameters. The detector resolution function is now constructed using a DSCB function parameterized

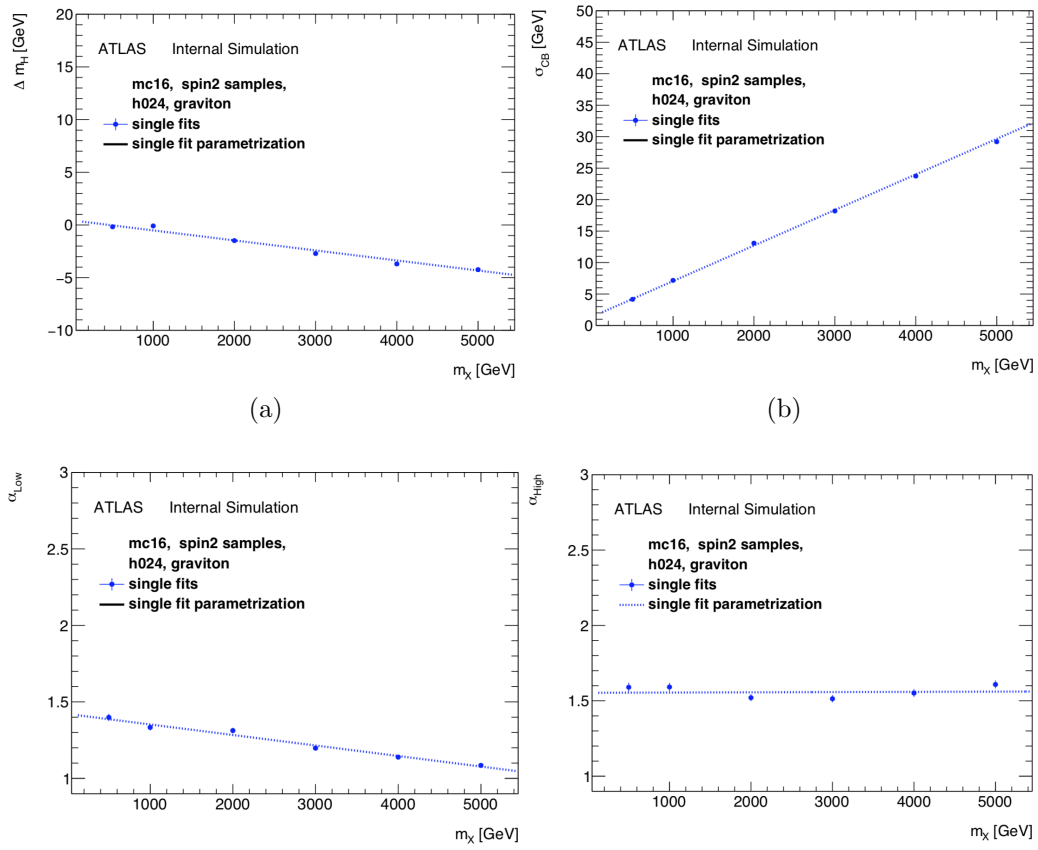


Figure 8.9: The graviton NWA DSCB parameters as a function of m_X. Each plot is fit with a linear function and the intermediate values are interpolated as the straight line.

as shown in Table 8.5. The parameterization is validated by comparing the graviton

DSCB parameter	Parameterization
Δm_X	$0.437 - 0.951 \times 10^{-3} m'_X$
σ_{CB}	$1.404 - 5.649 \times 10^{-3} m'_X$
α_{low}	$1.421 + 6.874 \times 10^{-5} m'_X$
α_{high}	$1.553 - 1.521 \times 10^{-6} m'_X$
n_{low}	13.0 (const.)
n_{low}	25.1 (const.)

Table 8.5

NWA MC sample shapes with the parameterized DSCB. These plots can be seen in Figure 8.10.

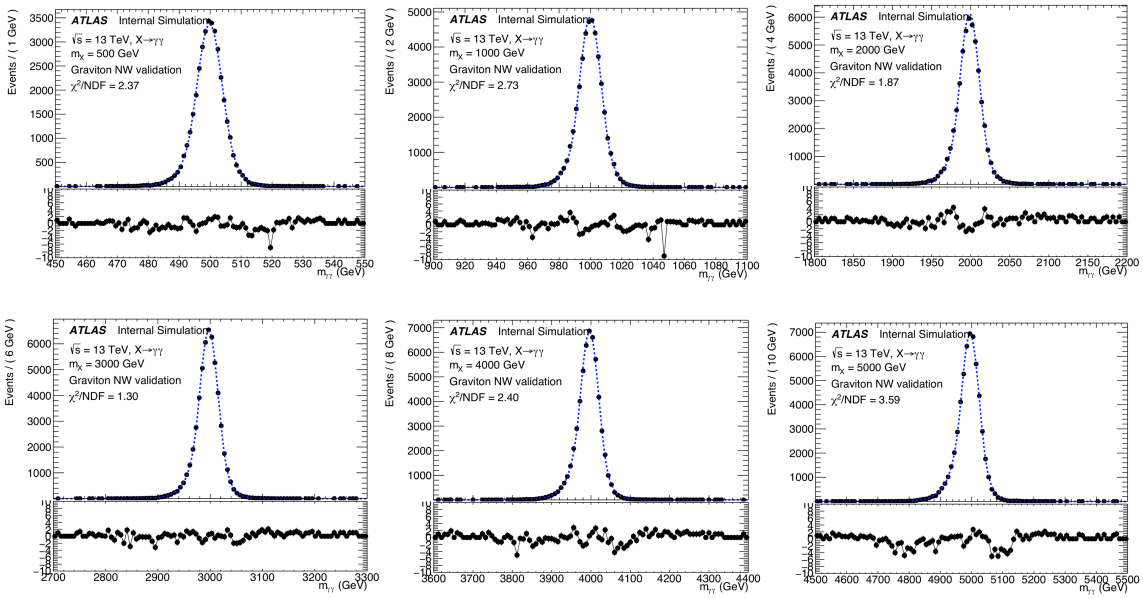


Figure 8.10: Validation plots for the graviton NWA parameterization.

The lineshape of the LW gravitons are given by

$$F_{G^*} \propto m_{\gamma\gamma}^7(\mathcal{L}_{gg} + \alpha\mathcal{L}_{q\bar{q}}) \cdot F_{\text{BW.grav}}(m_{\gamma\gamma}; k/\overline{M}_{\text{Pl}}) \tag{8.6}$$

$$F_{\text{BW.grav}} = \left\{(m_{\gamma\gamma} - m_{G^*}^2)^2 + \left[m_{\gamma\gamma} \cdot 1.44(k/\overline{M}_{\text{Pl}})^2\right]^2\right\}^{-1} \tag{8.7}$$

where $\mathcal{L}_{gg}$ and $\mathcal{L}_{q\bar{q}}$ are the gg and $q\bar{q}$ luminosities repectively, and $F_{\text{BW.grav}}$ is the relativistic BW function for a graviton resonance with coupling $k/\overline{M}_{\text{Pl}}$ and mass m_{G^*}. The convolution is implemented, like in the scalar case, using `RooFit` and are validated by comparing the derived lineshapes to the graviton MC samples. These validation plots can be seeen in Figures 8.11, 8.12, and 8.13.

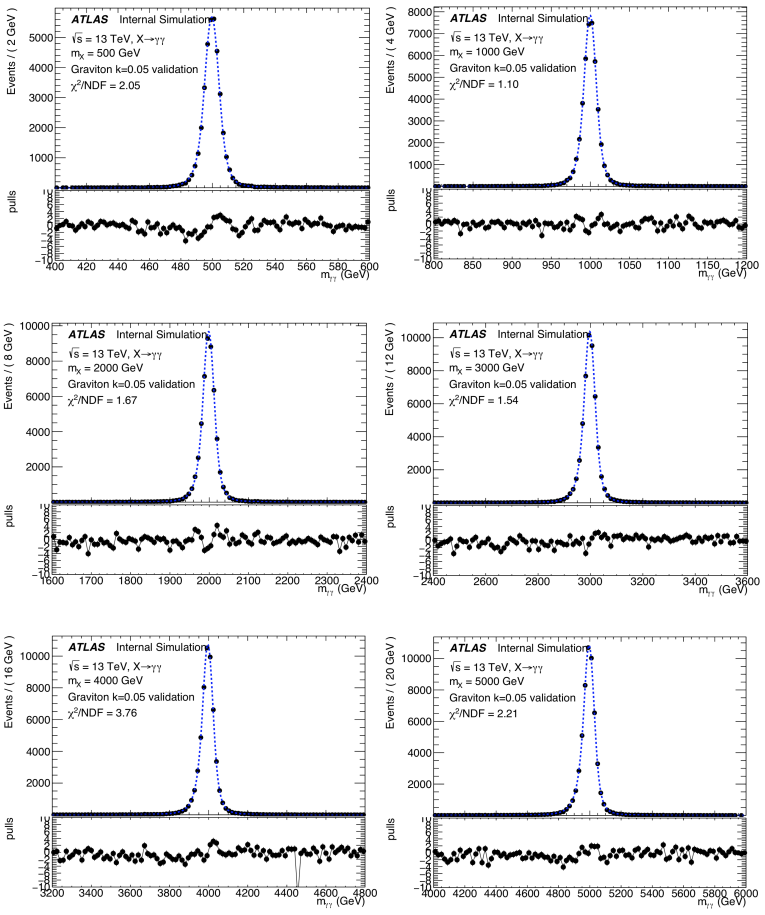


Figure 8.11: Validation plots for the graviton parameterization with $k = 0.05$.

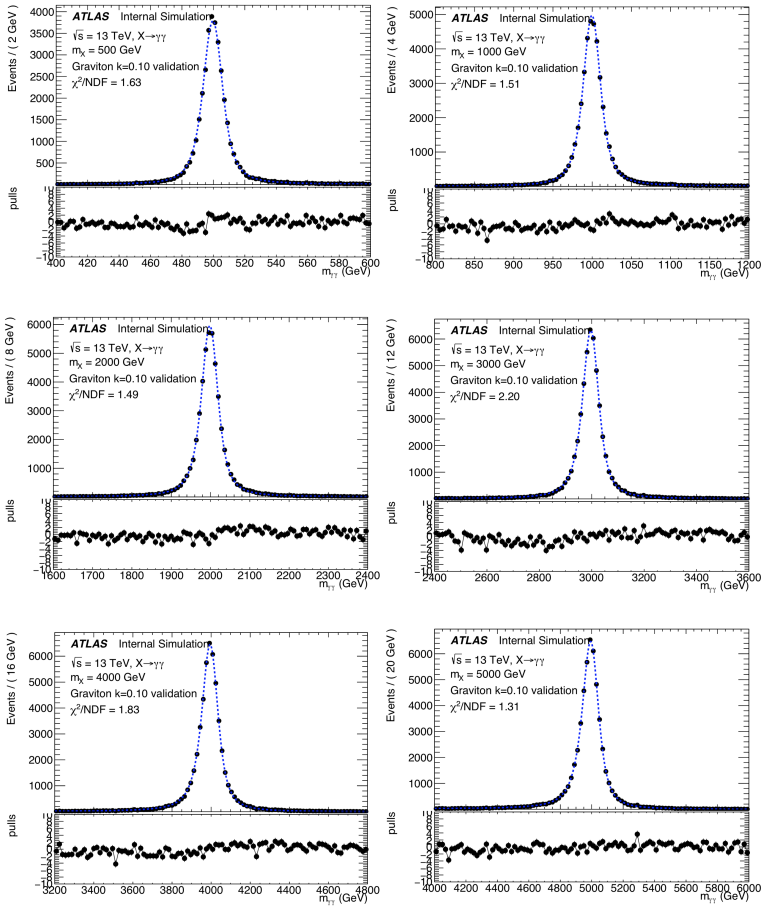


Figure 8.12: Validation plots for the graviton parameterization with $k = 0.10$.

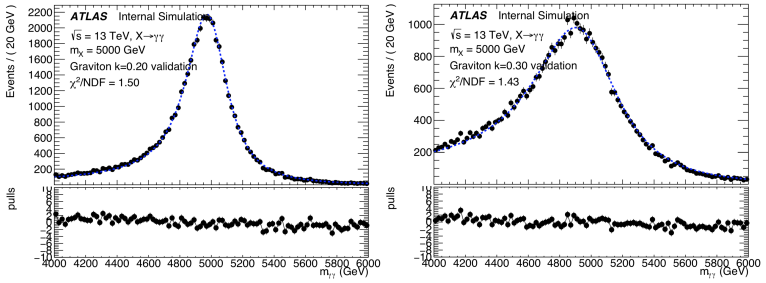


Figure 8.13: Validation plots for the graviton parameterization with $k = 0.20$ and $k = 0.30$ and mass $m_X = 5000$ GeV.

CHAPTER IX

Functional Decomposition

FD is a new data driven modeling technique inspired by Fourier analysis. It uses a sum of orthogonal functions to represent smoothly falling spectra with or without local features. Lower order terms in the sum are used to represent the delocalized background while higher order terms represent localized structures such as resonances from decaying particles. In this way an S+B model for a spectrum can be built. Once an S+B model is constructed, deviations from the model are compared to a given lineshape, and an estimate for the size of the deviation is extracted. FD is particularly well suited for modeling narrow resonances on a smoothly falling background, making it an ideal tool to search for new particles in the diphoton channel (*Edgar et al.*, 2018).

9.1 Functional Decomposition Fundamentals

Suppose $\mathcal{F}$ is the set of all real-valued functions $f(z)$ defined on $[z_0, \infty)$ with the properties

$$\lim_{z\to\infty} f(z) = 0 \tag{9.1}$$

$$\int_{z_0}^{\infty} f(z)\, dz = M \tag{9.2}$$

where z_0 is a constant. Now suppose one had a set of orthonormal functions $\{E_n\}$ complete on the interval $[z_0, \infty)$ with respect to $\mathcal{F}$. That is, any function $f \in \mathcal{F}$ could be represented as a linear combination of the functions E_n. This linear combination can be written

$$f(z) = M\bar{f}(z) \tag{9.3}$$

$$\bar{f}(z) = \sum_{i=0}^{\infty} c_n E_n(z) \tag{9.4}$$

where the coefficients c_n are determined by averaging E_n over $\bar{f}(z)$ on $[z_0, \infty)$

$$c_n = \int_{z_0}^{\infty} \bar{f}(x) E_n(z)\ dz \tag{9.5}$$

$$= \frac{1}{M} \int_{z_0}^{\infty} f(z) E_n(z)\ dz. \tag{9.6}$$

If f represents a spectrum of M events, then as $M \to \infty$ the quantity $\bar{f}$ approaches the pdf of the spectrum. Therefore an approximation of the underlying pdf can be written

$$\bar{f}(z) \approx \sum_{i=i}^{\mathcal{N}} c_n E_n(z) \tag{9.7}$$

with the Hilbert space decomposition

$$\tilde{\boldsymbol{f}} = \langle c_0\ ,\ c_1\ ,\ \ldots\ ,\ c_{\mathcal{N}} \rangle \tag{9.8}$$

and covariance matrix

$$\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{f}} nm} = \int_0^{\infty} dz\ f(z) E_n(z) E_m(z) - \tilde{\boldsymbol{f}}_n \tilde{\boldsymbol{f}}_m, \tag{9.9}$$

where $\tilde{\boldsymbol{f}}_i$ is the i^{th} entry of $\tilde{\boldsymbol{f}}$, i.e. $\tilde{\boldsymbol{f}}_i \equiv c_i$. The function $f(z)$ can now be written compactly with Einstein summation as

$$f(z) = \tilde{\boldsymbol{f}}^n E_n(z) \tag{9.10}$$

where $\tilde{\boldsymbol{f}}_n = \tilde{\boldsymbol{f}}^n$ and the covariance matrix as

$$\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{f}}nm} = \int_0^\infty dz\ \tilde{\boldsymbol{f}}^i E_i(z) E_n(z) E_m(z) - \tilde{\boldsymbol{f}}_n \tilde{\boldsymbol{f}}_m \tag{9.11}$$

$$= \hat{\boldsymbol{I}}_{inm} \tilde{\boldsymbol{f}}^i - \tilde{\boldsymbol{f}}_n \tilde{\boldsymbol{f}}_m \tag{9.12}$$

where

$$\hat{\boldsymbol{I}}_{inm} = \int_0^\infty dz\ E_i(z) E_n(z) E_m(z) \tag{9.13}$$

acts on $\tilde{\boldsymbol{f}}^i$.

9.2 Orthonormal Exponentials

Since most high energy physics spectra are smoothly falling, the function $f(z) = N \exp(-\alpha z)$ can often be used as a first order approximation for these spectra. Armed with this knowledge, a set of orthogonalized exponential functions are chosen as a basis for FD. Beginning with the non-orthogonal parent functions

$$F_n(z) = \sqrt{2} e^{-nz} \tag{9.14}$$

a set of functions orthogonal with respect to the L^2 inner product

$$\langle f, g \rangle = \int_0^\infty dz \; f(z)g(z) \tag{9.15}$$

is constructed called the *orthonormal exponentials*. More concisely, a set of functions is built with the form

$$E_n(z) = \sum_{m=1}^{n} \hat{\boldsymbol{d}}_{nm} F_m(z) \tag{9.16}$$

where

$$\hat{\boldsymbol{d}}_{nm} = \sqrt{n}(-1)^{n+m} \left(\frac{2m}{n+m} \right) \prod_{i=1}^{m-1} \frac{m+i}{m-1} \prod_{i=m+1}^{n} \frac{i+m}{i-m} \tag{9.17}$$

are constants chosen such that

$$\langle E_n, E_m \rangle = \delta_{nm}. \tag{9.18}$$

Although there is a closed form, calculating $\hat{\boldsymbol{d}}_{nm}$ directly is prohibitively computationally expensive so an alternative approach is required. The following recursion relations are computationally efficient compared to calculating $\hat{\boldsymbol{d}}_{nm}$ directly or through an othogonalization scheme such as Gram-Schmidt.

$$E_0(z) = 0 \tag{9.19}$$

$$E_1(z) = \sqrt{2}e^{-z} \tag{9.20}$$

$$E_{n+1}(z) = \frac{1}{\phi_{2n+1}} \left(4e^{-z} E_n(z) - \frac{2}{\phi_{2n}^2} E_n(z) - \phi_{2n-1} E_{n-1}(z) \right) \tag{9.21}$$

$$\phi_n = \sqrt{1 - \frac{1}{n^2}} \tag{9.22}$$

These recursion relations are derived by considering the quantity $\hat{\boldsymbol{d}}_{n(m+1)}/\hat{\boldsymbol{d}}_{nm}$. The increased efficiency comes from the lack of integrals or other complex computations, a function E_{n+1} is evaluated at z simply by considering the values of $E_n(z)$ and $E_{n-1}(z)$.

The orthonormal exponentials are complete with respect to pdfs on the interval $[0, \infty)$. The completeness can most easily be seen by mapping $\sqrt{2}F_n = e^{-nz}$ on $[0, \infty)$ to the polynomials $F_n^\star(z) = \sqrt{2}y^n$ on $(0, 1]$ and relying on the completeness of $F_n^\star$. A proof for the completeness of the orthonormal exponentials can be found in Appendix A.

9.3 Power-law Transformation

Most physics spectra are not on the interval $[0, \infty)$, rather they are on the interval $[x_0, \infty)$ where x_0 is a constant. To make matters worse, simply shifting x by x_0 usually results in a spectrum which takes many terms to reasonably model. To reduce the number of terms needed to model a lineshape, a transformation $z = T(x; \bar{\theta})$ is applied to the input variable x. This transformation must invertibly map $[x_0, \infty) \to [0, \infty)$, be continuous, and render z dimensionless. The power-law transformation

$$z = \left(\frac{x - x_0}{\lambda} \right)^\alpha \tag{9.23}$$

meets these requirements and introduces three hyperparameters, x_0, the *lower edge* with dimensions of x, λ, a *length scale* with the dimensions of x, and α, a dimensionless *power-law exponent.* Figure 9.1 shows a plot of the power-law transformation applied to the function $g(z) = g_0 e^{-z}$ where g_0 is a constant versus the input variable x. If $g(x)$ represents a pdf, then holding λ fixed and increasing $\alpha \to \alpha + \delta\alpha$ decreases the likelihood of measuring $x > \lambda$ and increases the likelihood of measuring $x < \lambda$. On the other hand, holding α constant and increasing $\lambda \to \lambda + \delta\lambda$ results in a uniform scaling which *stretches* the curve, that is $x \to \frac{\lambda}{\lambda + \delta\lambda} x$.

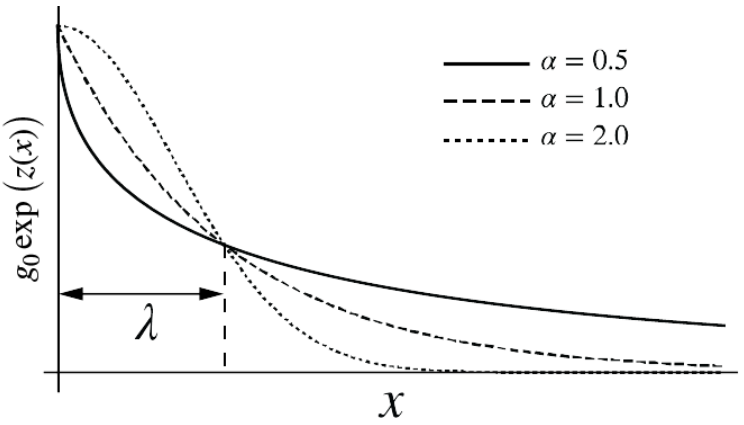


Figure 9.1: Plot of the power-law transformation for a fixed λ and several values of α applied to a decaying exponential.

Intuitively, one can think of the three hyperparameters in the following ways. The lower edge, x_0, shifts x, that is $x \to x - x_0$. Small increases to λ correspond to small increases in the spread of the distribution. Small changes to α correspond to small changes on the shape of the distribution, particularly the tail.

9.4 Decomposing a Dataset

A set of M measurements $\{x_i\}$ is transformed to the set $\left\{ z_i \middle| z_i = \left(\frac{x_i - x_0}{\lambda}\right)^{\alpha} \right\}$ and represented as a sum of Dirac delta functions

$$f(z) = \sum_{i=1}^{M} \delta(z - z_i). \tag{9.24}$$

It is decomposed and written as

$$\bar{f}(z) = \lim_{\mathcal{N} \to \infty} \sum_{n=0}^{\mathcal{N}} c_n E_n(z) \tag{9.25}$$

where

$$c_n = \frac{1}{M}\int_{z_0}^{\infty} f(z)E_n(z)\ dz \tag{9.26}$$

$$= \frac{1}{M}\int_{z_0}^{\infty} \sum_{i=1}^{M} \delta(z - z_i)E_n(z)\ dz \tag{9.27}$$

$$= \frac{1}{M}\sum_{i=1}^{M} E_n(z_i). \tag{9.28}$$

This is an exact representation of the data and requires an infinite number of terms. A choice for the size of the basis, typically $\mathcal{N} = 2048$, must be chosen to obtain a useable approximation of the underlying pdf $\bar{f}$. Plugging Equation 9.28 into Equation 9.9 gives the covariance matrix

$$\hat{\boldsymbol{\Sigma}}_{nm} = \frac{1}{M}\sum_{i=1}^{M} E_n(z_i)E_m(z_i) - \tilde{\boldsymbol{f}}_n\tilde{\boldsymbol{f}}_m. \tag{9.29}$$

9.5 Constructing a Model

A background only model $B(z)$ can be constructed using the first $N \in \mathbb{N}$ moments, that is

$$B(z) = \sum_{n=0}^{N} c_n E_n(z) \tag{9.30}$$

or in Hilbert space notation

$$\tilde{\boldsymbol{B}}_n = \begin{cases} \tilde{\boldsymbol{f}}_n & \Leftarrow n < N \\ 0 & \Leftarrow n \geq N. \end{cases} \tag{9.31}$$

Now suppose one wants to include a signal of size $M_s = sM$ and with lineshape $\tilde{\boldsymbol{S}}$, where M is the total number of events in the spectrum. Since $\tilde{\boldsymbol{S}}$ is a localized

structure it is described with mostly the higher order moments. However, the lower order moments, that is $\tilde{\boldsymbol{S}}_n$ for $n < N$, are in general non-zero. To prevent *double counting*, $\tilde{\boldsymbol{B}}$ is adjusted to account for the first N moments of $\tilde{\boldsymbol{S}}$, that is

$$\tilde{\boldsymbol{B}}_n = \begin{cases} \tilde{\boldsymbol{f}}_n - s\tilde{\boldsymbol{S}}_n & \Leftarrow n < N \\ 0 & \Leftarrow n \geq N. \end{cases} \tag{9.32}$$

An S+B model can be built by adding the signal, scaled appropriately, to the adjusted background

$$\tilde{\boldsymbol{\Omega}}_n = \tilde{\boldsymbol{B}}_n + s\tilde{\boldsymbol{S}}_n. \tag{9.33}$$

This can be generalized to multiple signals $S_{(m)}$ with signal sizes $M_{s_{(m)}} = s_{(m)}M$

$$\tilde{\boldsymbol{\Omega}}_n = \tilde{\boldsymbol{B}}_n + s^{(m)}\tilde{\boldsymbol{S}}_{(m)n} \tag{9.34}$$

where $\tilde{\boldsymbol{B}}_n = \tilde{\boldsymbol{f}}_n - s^{(m)}\tilde{\boldsymbol{S}}_{(m)n}$, and the raised and lowered (m) imply summation.

9.6 Estimating Signal Parameters

Since FD is primarily a resonance search tool, perhaps the most important part of it is its signal extraction procedure. The idea is to construct a set of estimators $\left\{\tilde{\boldsymbol{\omega}}_{(n)}\right\}$, each corresponding to a signal, such that $\langle \omega_{(n)}, B\rangle = 0$ and $\langle \omega_{(n)}, S_{(m)}\rangle = \delta_{nm}$, ensuring the overlap between the data and each of the the estimators is the size of its corresponding signal, that is $\langle \tilde{\boldsymbol{\omega}}_{(n)}, f\rangle = s_{(m)}$. There are many choices for these estimators which satisfy these properties, however they are not equal in their usefulness. A key difference between these choices is the entropy $H = \log\det(2\pi e\sigma)$

of the covariance matrix given by

$$\sigma_{(nm)} = \tilde{\boldsymbol{\omega}}^{i}_{(n)} \hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{f}}ij} \tilde{\boldsymbol{\omega}}^{j}_{(m)}. \tag{9.35}$$

and where the factor $2\pi e$ is a normalization convention. The set is *optimal* if the entropy of σ is minimized with respect to the set of all possible linear unbiased estimators $\tilde{\boldsymbol{\epsilon}}_{(n)i}$ given by

$$\tilde{\boldsymbol{\epsilon}}_{(n)i} = \hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{l}}ij} \tilde{\boldsymbol{S}}^{j}_{(n)} \tag{9.36}$$

where $\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{l}}jk}$ is the covariance defined from the *lower* moments, that is

$$\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{l}}jk} = \sum_{i=0}^{N-1} \tilde{\boldsymbol{f}}_i \hat{\boldsymbol{I}}_{ijk}. \tag{9.37}$$

To minimize the entropy, N^2 Langrange multipliers $\eta_{(ij)}$ are introduced to produce the objective function

$$\mathcal{L} = \log\det\left(2\pi e\sigma\right) - \eta^{(nm)} \left(\tilde{\boldsymbol{\omega}}_{(n)i} \tilde{\boldsymbol{f}}^{i}_{(m)} - \delta_{nm} \right). \tag{9.38}$$

The signal parameter $s_{(n)}$ is found by taking the inner product of the estimator, $\tilde{\boldsymbol{\epsilon}}_{(n)i}$, with the data, that is

$$\eta^{-1}_{(nm)} = \langle \epsilon_{(n)}, S_{(m)} \rangle = \tilde{\boldsymbol{S}}^{i}_{(n)} \hat{\boldsymbol{\Sigma}}^{-1}_{\tilde{\boldsymbol{l}}ij} \tilde{\boldsymbol{S}}^{j}_{(m)} \tag{9.39}$$

$$s_{(n)} = \eta^{(\ k)}_{(n\ )} \langle \epsilon_{(k)}, f \rangle. \tag{9.40}$$

In practice $\eta_{(n)}$ is found numerically. The full procedure for constructing these estimators can be found in Appendix D. To convert the signal parameters to yields, they are simply multiplied by the total number of events in the dataset.

9.7 Optimizing Hyperparameters

Choosing hyperparameters is a major challenge of FD. Hyperparameter choice greatly effects the number of terms needed to capture the shape of the spectrum as well as FD's ability to detect a signal. In principle any hyperparameter choice can be used to construct a model, however most hyperparameter choices are impractical. Most choices will require many moments to construct a reasonable background model leaving fewer moments for measuring signal yields. To help combat this problem an *objective function*, $\mathscr{L}$, is introduced consisting of a *likelihood* term, $\mathcal{L}$, designed to measure the data's deviation from the model and a *penalty* term, $\mathcal{P}$, designed to measure the background model's deviation from a prior distribution, $\tilde{\boldsymbol{p}}$. Recall from Sections 9.5 and 9.3 that the hyperparameters N, x_0, λ, and α have been introduced. Only the lower edge, x_0, is specified. Therefore a function with the form

$$\mathscr{L} = \mathcal{L}\left(\tilde{\boldsymbol{f}} \middle\| \tilde{\boldsymbol{\Omega}}\right) + \mathcal{P}\left(\tilde{\boldsymbol{B}} \middle\| \tilde{\boldsymbol{p}}\right) \tag{9.41}$$

is minimized with respect to the hyperparameters α, λ, and N. Intuitively the likelihood term picks the best α and λ while the penalty term decides where to truncate the series. Of course in reality, these two terms are correlated so each contribute to the choice of all three hyperparameters to some degree.

The log likelihood for a multivariate normal distribution of measuring $\tilde{\boldsymbol{f}}$ given the underlying pdf $\tilde{\boldsymbol{\Omega}}$ with covariance $\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{\Omega}}}$ is given by

$$\log L = -\frac{1}{2}\left\{\log\left(|\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{\Omega}}}|\right) + \left(\tilde{\boldsymbol{f}} - \tilde{\boldsymbol{\Omega}}\right)^{\top} \hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{\Omega}}}^{-1}\left(\tilde{\boldsymbol{f}} - \tilde{\boldsymbol{\Omega}}\right) + k\log(2\pi)\right\} \tag{9.42}$$

where $\tilde{\boldsymbol{\Omega}}, \tilde{\boldsymbol{f}} \in \mathbb{R}^k$. Since the moments are assumed to be normally distributed about their means, the likelihood function is chosen to be a modified log likelihood of the

multivariate normal distribution

$$\mathcal{L}\left(\tilde{\boldsymbol{f}}\middle\|\tilde{\boldsymbol{\Omega}}\right) = \frac{1}{2}\left\{\log\left(|\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{f}}}|\right) + \left(\tilde{\boldsymbol{f}} - \tilde{\boldsymbol{\Omega}}\right)^{\top} \hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{f}}}^{-1}\left(\tilde{\boldsymbol{f}} - \tilde{\boldsymbol{\Omega}}\right)\right\}. \tag{9.43}$$

Note that the term $k\log(2\pi)$ has been omitted since it is constant for any choice of α, λ, and N. Note also that the covariance matrix $\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{\Omega}}}$ has been replaced with $\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{f}}}$ since $\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{\Omega}}} \approx \hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{f}}}$ and in practice $\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{f}}}$ has been previously calculated.

The Kullback–Leibler (KL) divergence of $\tilde{\boldsymbol{a}}$ from $\tilde{\boldsymbol{b}}$ for a multivariate normal distribution with covariances $\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{a}}}$ and $\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{b}}}$ is given by

$$D_{\mathrm{KL}}\left(\tilde{\boldsymbol{a}}\middle\|\tilde{\boldsymbol{b}}\right) = \frac{1}{2}\left\{\mathrm{Tr}\left(\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{b}}}^{-1}\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{a}}}\right) + \left(\tilde{\boldsymbol{b}} - \tilde{\boldsymbol{a}}\right)^{\top} \hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{b}}}^{-1}\left(\tilde{\boldsymbol{b}} - \tilde{\boldsymbol{a}}\right) - k + \log\frac{|\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{b}}}|}{|\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{a}}}|}\right\} \tag{9.44}$$

where $\tilde{\boldsymbol{a}}, \tilde{\boldsymbol{b}} \in \mathbb{R}^k$ (*Kullback and Leibler*, 1951) (*Soch and Allefeld*, 2016). The penalty term is chosen to be the KL divergence of $\tilde{\boldsymbol{B}}$ from $\tilde{\boldsymbol{p}}$. Without choosing a specific form, the prior $\tilde{\boldsymbol{p}}$ is taken to be weak, that is $\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{p}}}$ is large compared to $\hat{\boldsymbol{\Sigma}}_{\tilde{\boldsymbol{B}}}$ and has the equivalent statistical strength $j = k = N$ where N is the number of background moments in $\tilde{\boldsymbol{B}}$. This choice results in the first two terms of the KL divergence vanishing, and the log determinant term approaching $N\log\frac{M}{N}$ where M is the number of events in the dataset (*Edgar et al.*, 2018). Therefore the penalty term can be written

$$\mathcal{P}\left(\tilde{\boldsymbol{B}}\middle\|\tilde{\boldsymbol{p}}\right) = \frac{N}{2}\log\left(\frac{M}{Ne}\right). \tag{9.45}$$

To optimize the objective function $\mathscr{L}$ first a grid search is performed at several hundred (α, λ) points, and at each grid point $\mathscr{L}$ is evaluated for $N = 1\,,\,2\,,\,\ldots\,,\,N_{\max}$. Next, a gradient descent algorithm is performed beginning at the global minimum found in the grid search. The gradient descent is performed in *flattened* space, that is at each point (α, λ) the value is taken to be the minimum value with respect to N (*Edgar et al.*, 2018).

9.8 Validating Functional Decomposition as a Search Technique

9.8.1 Signal Injection

This section gives details about the *signal injection* studies performed for this analysis. The objective of these studies is to test FD's ability to find an NWA scalar signal in the expected $m_{\gamma\gamma}$ distribution for 140 fb^{-1} of data. It is performed using backgrounds with the expected shape and statistics of 140 fb^{-1} of data and narrow width scalar signal samples with sizes corresponding to excesses between 0–5σ. Signal injection studies for gravitons and LW scalars are an ongoing effort and therefore will not be presented in this thesis.

9.8.1.1 Methodology

Many statistically independent background toys, described in Section 7.1.2, are each combined with a toy signal of size N_{inj} generated with the method described in Section 7.2.2. These combined samples are used as inputs for FD. First the combined toys are modeled using the procedure outlined in Section 9.5. Next, the number of signal events, N_{ext}, is extracted using the signal estimators described in Section 9.6. The number of signal events extracted, N_{ext}, is compared to the number of signale events injected, N_{inj} and is fit with a linear function.

Since the analysis searches for resonances in the range $m_{\gamma\gamma} \in [150, 5000]$ GeV the toy backgrounds are injected with toy signals whose mass is within that range. At a given resonance mass, m_{inj}, the size of the injected signal, N_{inj}, is varied such that they are within the 0–5σ significance range. The values of N_{inj} are approximated with To estimate B_{res}, the PowLog-1 fit parameterization from Table 7.2 is used assuming

a binning of 10 GeV. The background in each signal region is defined such that

$$B_{\text{res}} = B \times \frac{4\sigma_{\text{std}}^{signal}}{10 \text{ GeV}} \tag{9.46}$$

where B is the number of background events per 10 GeV bin. The results are cross checked with the reweighted SHERPA MC samples and are found to be consistent (*et. al.*, 2019). Table 9.1 shows the estimated number of signal events rounded to the nearest integer corresponding to $n\sigma$ excesses where $n \in [1, 5]$.

$$N_{\text{inj}}(n\sigma) = n\sigma_{\text{stat}}^{\text{toy}} = n\sqrt{B_{\text{res}} + S_{\text{res}}} = \frac{n^2 + n\sqrt{n^2 + 4B_{\text{res}}}}{2} \tag{9.47}$$

where the substitution $S_{\text{res}} = N_{\text{inj}}$ is made to extract the significance (*et. al.*, 2019). Table 9.1 shows the number of events rounded to the nearest integer corresponding to excesses in the range 1-5 for the improved approximation.

$m_{\gamma\gamma}$ [GeV]	B $\left[\frac{\text{count}}{10 \text{ GeV}}\right]$	$\sigma_{\text{std}}^{\text{signal}}$ [GeV]	N_{inj}				
			1σ	2σ	3σ	4σ	5σ
400	1555.7	3.911	50	101	153	206	260
800	66.11	6.760	14	29	45	62	81
1000	21.89	8.235	9	19	30	43	57
1200	8.48	9.645	6	14	22	32	44
1600	1.73	12.46	3	8	14	22	32

Table 9.1

For each mass m_{inj}, 10 values of N_{inj} are injected into the background corresponding $n\sigma$ excesses where $n \in [0, 5]$. For each combination of m_{inj} and N_{inj} the signals are injected into 500 background toys, for a total of 5000 statistically independent samples for each m_{inj}. Each of these samples is used as input to FD and an S+B decomposition is made with $m_X = m_{\text{inj}}$. The number of signal events, N_{ext}, given by FD is compared to N_{inj} with the expectation that for each m_{inj}-N_{inj} combination N_{ext} will be normally distributed about m_{inj} with a width approximately equal to $\sqrt{B_{\text{res}}}$.

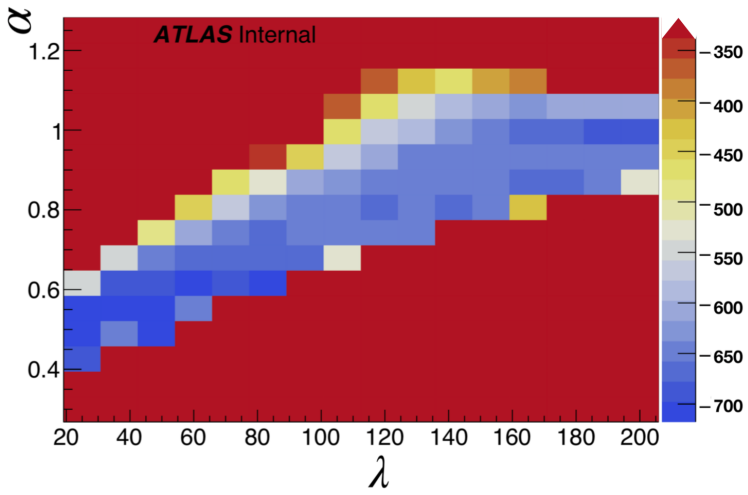


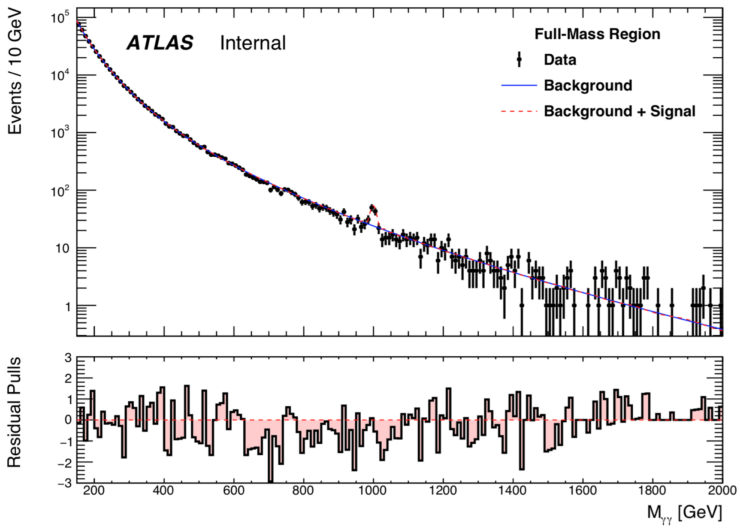
Figure 9.2: Hyperparameter landscape of the likelihood for an S+B decomposition of a background toy injected with a signal toy. Note that the N axis is flattened in the sense that for each point in α and λ, the N corresponding to the best likelihood *at that point* is shown.

Deviations from this expectation are assigned as a bias in the FD methodology (*et. al.*, 2019).

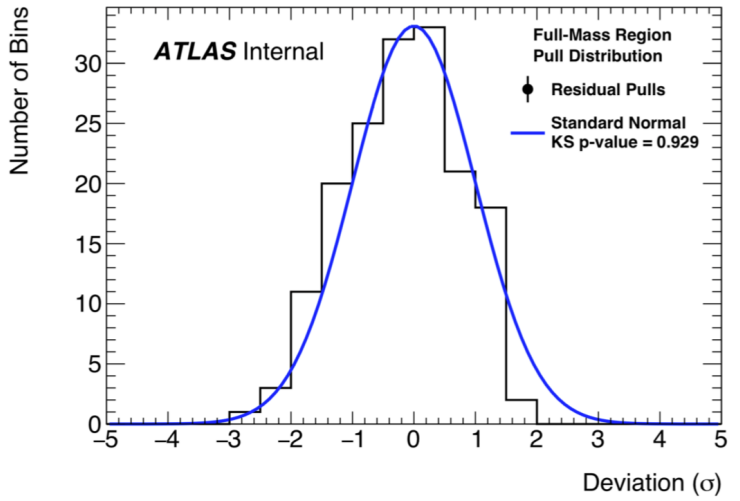
9.8.1.2 One Pseudo-Experiment with $m_{\text{inj}} = 1000$ GeV and $N_{\text{inj}} = 40$

The process of performing a pseudo-experiment is best understood through an example. One background toy is injected with a signal toy of mass $m_{\text{inj}} = 1000$ GeV and size $N_{\text{inj}} = 40$ events, corresponding to a 3.8σ excess. The hyperparameter landscape, found while minimizing the penalized likelihood function with respect to $\alpha \in [0.3, 1.25]$, $\lambda \in [25, 200]$, and $N \in \{1, 2, \ldots, 24\}$ as described in Section 9.7, can be seen in Figure 9.2. Note that the values shown are the likelihoods found during a course grid scan. The best point found in this scan is then used as the initial point in a gradient descent algorithm which gives the final hyperparameter choice. Figure 9.3 shows the full S+B decomposition compared to the input toys and its pull distribution. The pull distribution is compared to a standard normal and shows good agreement with the Kolmogorov-Smirnov (KS) p-value calculated to be 0.929.

The signal size is determined using the signal estimators described in Section 9.6.



(a)



(b)

Figure 9.3: An S+B decomposition of a toy background plus a 1000 GeV toy signal with size N_{inj} =40.

Figure 9.4 shows the DSCB estimator and signal lineshape used to extract N_{ext}.

9.8.1.3 Many Pseudo-Experiments with $m_{\text{inj}} = 1000$ GeV and $N_{\text{inj}} = 40$

The process for one pseudo-experiment, described in Section 9.8.1.2, is repeated 500 times, each time using a statistically independent background toy. The N_{ext} distribution for these trials can be seen in Figure 9.5. The distribution is fit with a

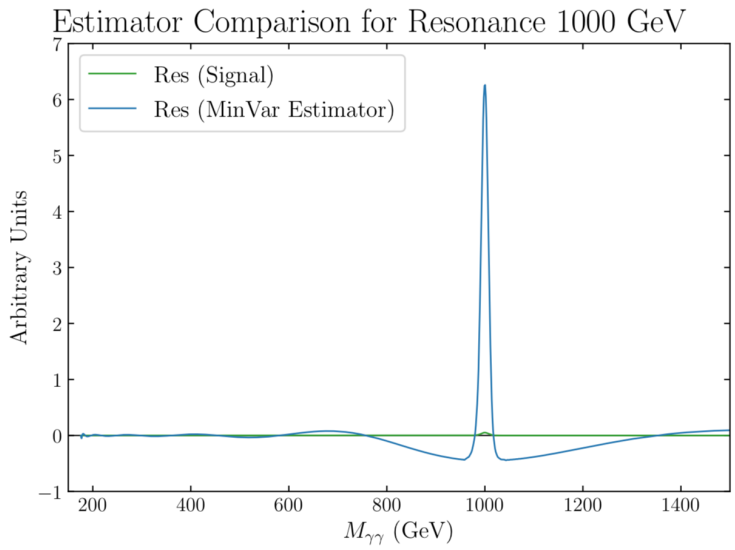


Figure 9.4: The DSCB lineshape and estimator for m_{inj} =1000 GeV.

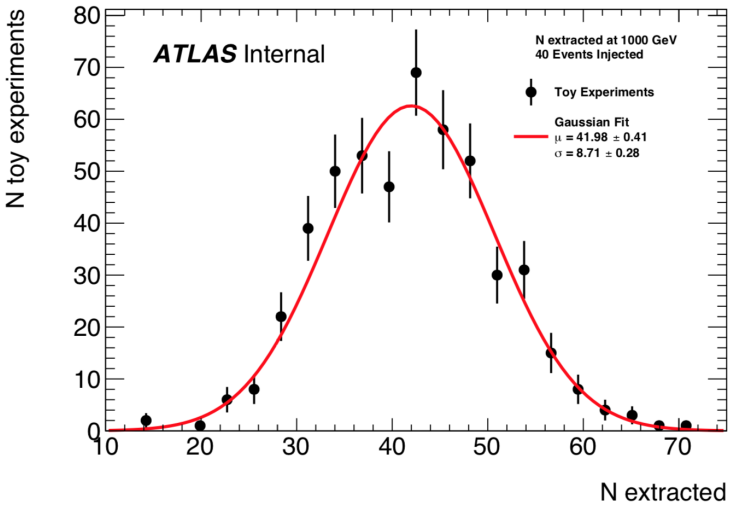


Figure 9.5: The N_{ext} distribution from 500 pseudo-experiments corresponding to a resonance mass $m_{\text{inj}} = 1000$ GeV and number of injected events $N_{\text{inj}} = 40$

Gaussian with mean $\mu = 41.98 \pm 0.41$ and width $\sigma = 8.71 \pm 0.28$. As mentioned earlier, the deviation of the mean from N_{inj} is taken to be a bias in FD's methodology. It is expressed as a fraction of N_{inj}, $(1 - N_{\text{ext}}/N_{\text{inj}})$, and is found to be approximately -5% for this case. The width σ is found to be consistent with the expected uncertainty, $\sigma_{\text{std}}^{\text{signal}}$, by comparing its value to that found in Table 9.1.

9.8.1.4 Many Pseudo-Experiments with $m_{\text{inj}} = 1000$ GeV and Various Values of N_{inj}

The pseudo-experiment process for $m_{\text{inj}} = 1000$ GeV with $N_{\text{inj}} = 40$, described in Section 9.8.1.3, is repeated for various values of N_{inj}, namely 0, 5, 10, 15, 20, 25, 30, 35, 40, and 45 events. These values of N_{inj} corresponded to local excess in the 0-5σ range. Table 9.2 shows the means and widths of the various N_{ext} distributions found using a Gaussian fit. Figure 9.6 shows the linearity plot, that is N_{ext} vs. N_{inj}, for

N_{inj}	N_{ext}	
	μ	σ
0	0.54 ± 0.42	8.85 ± 0.31
5	4.65 ± 0.36	7.86 ± 0.25
10	10.94 ± 0.37	8.08 ± 0.29
15	15.97 ± 0.42	8.76 ± 0.31
20	20.96 ± 0.40	8.27 ± 0.30
25	26.28 ± 0.45	9.60 ± 0.38
30	31.01 ± 0.39	8.49 ± 0.29
35	36.40 ± 0.43	8.88 ± 0.33
40	41.98 ± 0.41	8.71 ± 0.28
45	46.93 ± 0.40	8.64 ± 0.31

Table 9.2: The mean and standard deviation of the N_{ext} distributions for $m_{\text{inj}} = 1000$ GeV

$m_{\text{inj}} = 1000$ GeV, for $m_{\text{inj}} = 1000$ GeV. The plot is fit with a straight line and the slope is found to be 1.04 ± 0.01 and the y-intercept to be 0.24 ± 0.23, which agrees with the ideal case of with a slope of one and y-intercept of zero.

9.8.1.5 Many Pseudo-Experiments with Various Values of m_{inj} and N_{inj}

The process described in Section 9.8.1.4 is repeated for several values of m_{inj}: 400, 800, 1000, 1200, and 1600 GeV. The linearity plots for each of these values of m_{inj} are summarized in Table 9.3 (*et. al.*, 2019).

Figure 9.7 shows the bias in the N_{ext} extraction for all mass hypothesis tested.

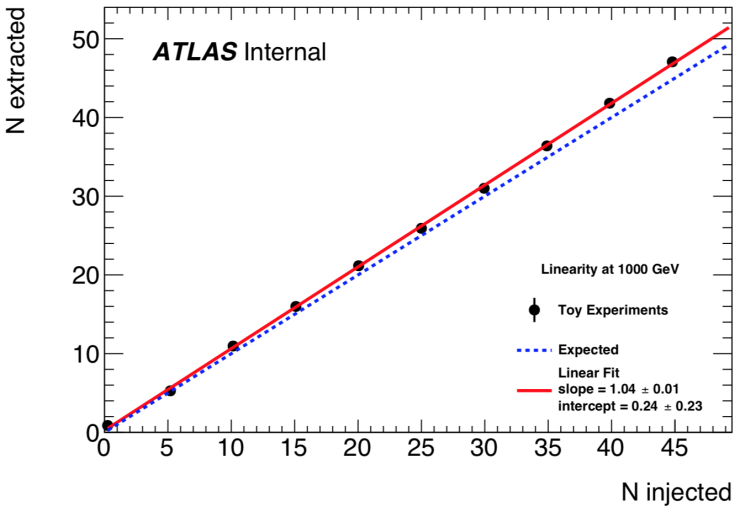


Figure 9.6: Linearity of N_{ext} at $m_{\text{inj}} = 1000$ GeV

$m_{\gamma\gamma}$ [GeV]	Slope	y-Intercept
400	1.02 ± 0.01	-3.99 ± 1.32
800	1.02 ± 0.01	1.16 ± 0.53
1000	1.04 ± 0.01	0.24 ± 0.23
1200	1.02 ± 0.01	0.07 ± 0.19
1600	1.01 ± 0.01	-0.22 ± 0.10

Table 9.3: Summary of the linearity plots for $m_{\text{inj}} = 400$, 800, 1000, 1200, and 1600 GeV

9.8.2 Spurious Signal

This section describes the *spurious signal studies* used for the analysis presented in this thesis. Spurious signal (SS), a systematic uncertainty associated with a model's inability to capture the true shape of a spectrum, is introduced in section 6.2. The goal of the SS studies is to quantify the average size of a model's deviation from the true lineshape of a spectrum.

Since FD can model any pdf on the interval $[x_0, \infty)$, the method is quite adaptive and is expected to capture all the features of the diphoton mass spectrum. The error from mismodeling the underlying pdf of the diphoton mass spectrum is therefore expected to be small compared to the statistical uncertainty. The statistical uncertainty,

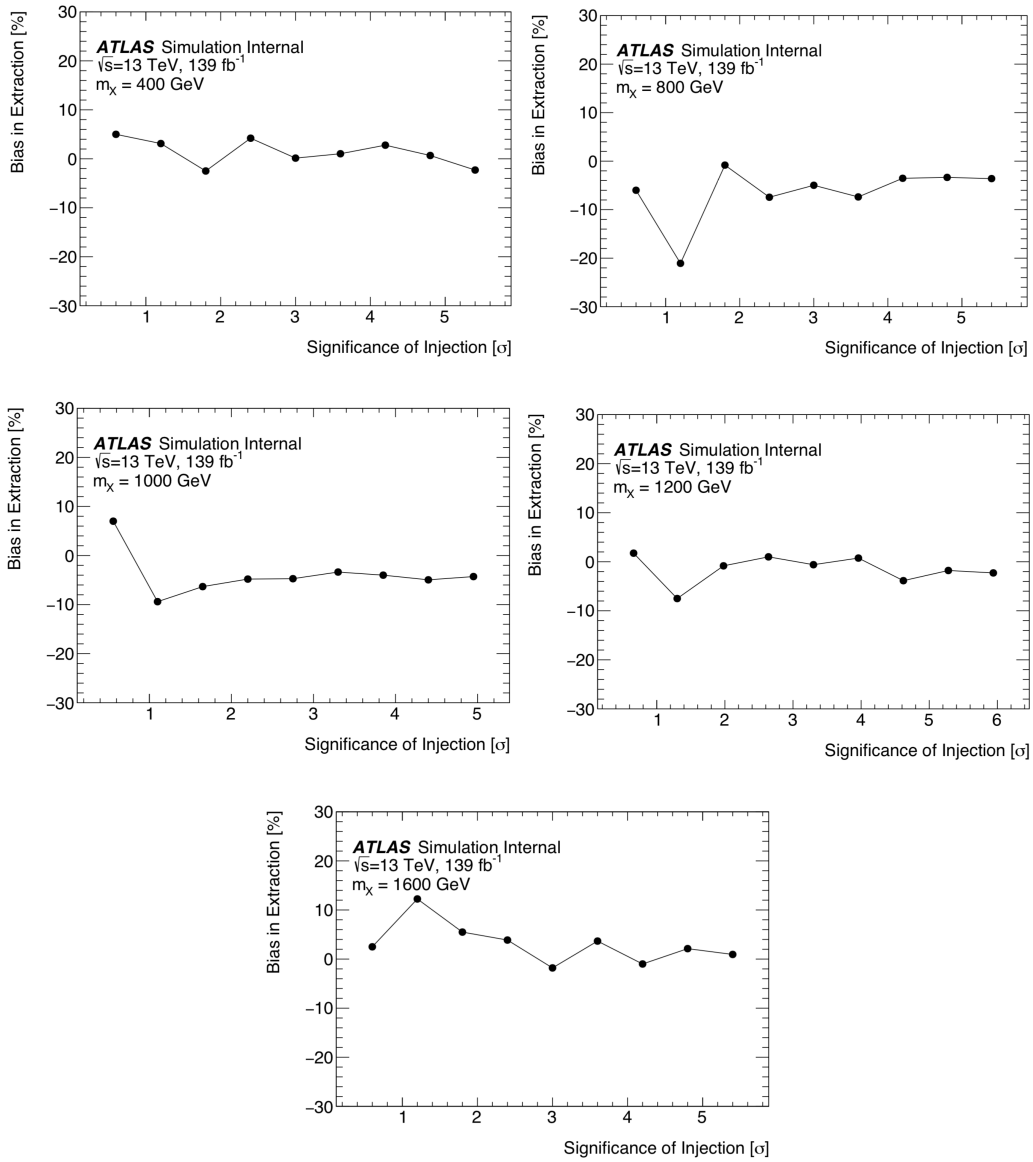


Figure 9.7: The bias, $(1 - N_{\text{ext}}/N_{\text{inj}})$, extracted for $m_{\text{inj}} = 400,\ 800,\ 1000,\ 1200$ and 1600 GeV

$\sigma_{\text{stat}}^{\text{data}}$, is expected to go like the square root of the background, that is

$$\sigma_{\text{stat}}^{\text{data}} \approx \sqrt{B} = \left(\sum_{\{i|m_i \in \Delta\}} w_i \right)^{1/2} \tag{9.48}$$

where Δ is the interval $[m_s - \sigma_s, m_s + \sigma_s]$, m_s is the signal mass, σ_s is the signal width, w_i is the weight for event i, and B is the number of events in Δ (*et. al.*, 2019).

This thesis presents two approaches for determining the SS. The first approach uses an *Asimov dataset* (*Amidei*, 2020), that is a representative dataset with no statistical fluctations, to find the SS directly. The second approach uses an ensemble of toy datasets with the statistics of 140 fb^{-1} to find the average signal size extracted by averaging out the impact of statistical fluctations in the spectrum. Like with the signal injection studies, described in Section 9.8.1.5, the SS for the LW scalar and graviton signals are a work in progress, and so will not be presented in this thesis.

9.8.2.1 Asimov Study

This study is performed by searching for narrow signals in a sample free from statistical fluctuations. The parameterized NWA DSCB function, described in Section 8.4.1, is used as the signal shape. The background is modeled using a PowLow-n function, given by Equation 6.6, where $d = 1/3$ by choice and $\sqrt{s} = 13$ TeV. The parameters of the function are found by fitting the MC template described in Section 7.1.1. To determine which value of n to use in the PowLog-n function, the MC template is fit using Equation 6.6 for $n \in [1, 10]$ and choosing the fit which produces the highest χ^2 probability. The χ^2 probabilities for these fits can be seen in Table 9.4. The parameterization of the chosen fit, the PowLog-6 function, is shown in Table 9.5.

n	χ^2 Probability
1	2.45×10^{-8}
2	7.18×10^{-5}
3	1.11×10^{-3}
4	7.85×10^{-2}
5	1.10×10^{-1}
6	1.68×10^{-1}
7	1.41×10^{-1}
8	1.48×10^{-1}
9	1.41×10^{-1}
10	1.31×10^{-1}

Table 9.4: χ^2 probability of the PowLog-n function while varying n

The PowLog-6 fit to the MC template can be seen in Figure 9.8.

Parameter	Value	Error
a	8.97	5.6×10^{-2}
α_0	-2.8185	1.23×10^{-2}
α_1	9.6367×10^{-3}	1.29×10^{-3}
α_2	-2.9597×10^{-4}	2.3×10^{-4}
α_3	-1.6492×10^{-4}	5.6×10^{-5}
α_4	-1.1894×10^{-4}	7.52×10^{-6}
α_5	-2.1581×10^{-5}	1.84×10^{-6}
α_6	-6.6266×10^{-6}	3.49×10^{-7}

Table 9.5: Parameters and their errors of the PowLog-n parameterization with $n = 6$

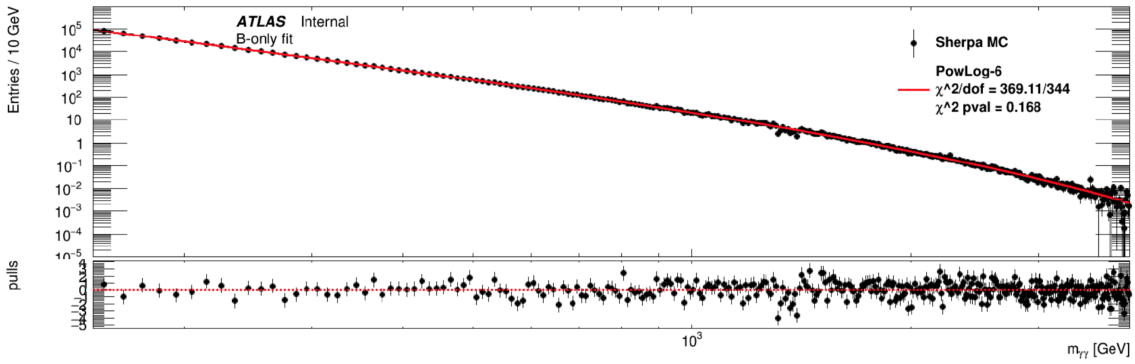


Figure 9.8: Background only fit of the Sherpa MC with the PowLog-6 function.

Systematic shape variations on the background are described in the empirically-derived function version of the search presented in this thesis (*Amidei*, 2020). Included in these systematics are shape variations from varying $f_{\gamma\gamma}$ the γj contribution, different choices of QCD factorization and renormalization scale, and different choices of PDF sets to describe the parton content of the proton (*et. al.*, 2019). The ratio between the nominal and systematically varied templates are summarized in Figure 9.9.

Although FD in principle can decompose any functional form f with the property $\lim_{x\to\infty} f(x) = 0$ on the interval $[x_0, \infty)$, the software developed for FD requires a discrete set of measurements represented as a sum of Dirac delta functions, similar to a *Dirac*

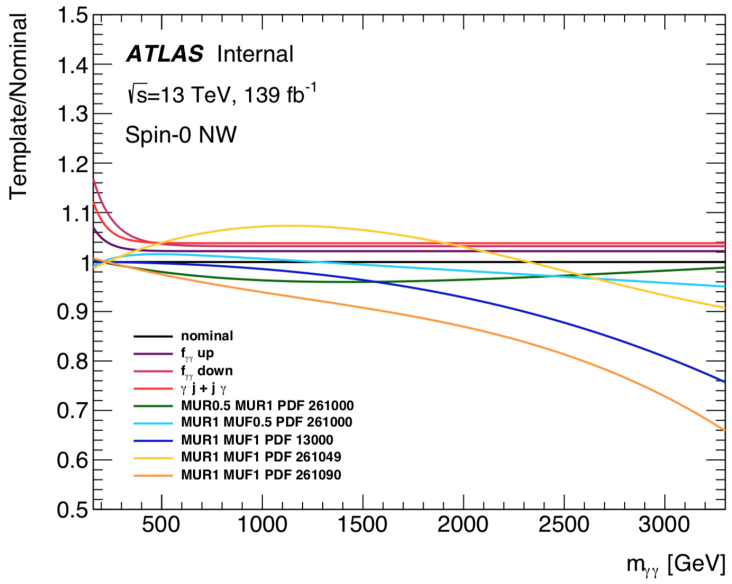


Figure 9.9: Systematic shape variations considered in the asimov spurious signal study.

Comb. To accomodate this limitation, the parameterized PowLog-6 function is taken to be a pdf and its distribution is approximated as a sum of Dirac delta functions. This is achieved by first splitting the mass region $[150, 5000]$ GeV into discrete bins. The integral of the PowLog-6 function in each bin is calculated and used as a weight for a Dirac delta function whose central value is that of the bin's center. Table 9.6 gives the binning used in this process. Figure 9.10 shows the decomposition of the delta

$m_{\gamma\gamma}$ range [GeV]	Number of bins
150-175	2,000,000
175-2000	1,825,000
2000-5000	40,000

Table 9.6: Binning used to approximate a continuous distribution with a sum of weighted Dirac delta functions.

function representation using $N = 16$ background moments. Note that the feature in the residual pulls in the high mass tail is thought to be due to the relatively low number of bins used in this region. Although this implies poor agreement on the tail, the number of events expected in the high mass region is very small, and therefore it

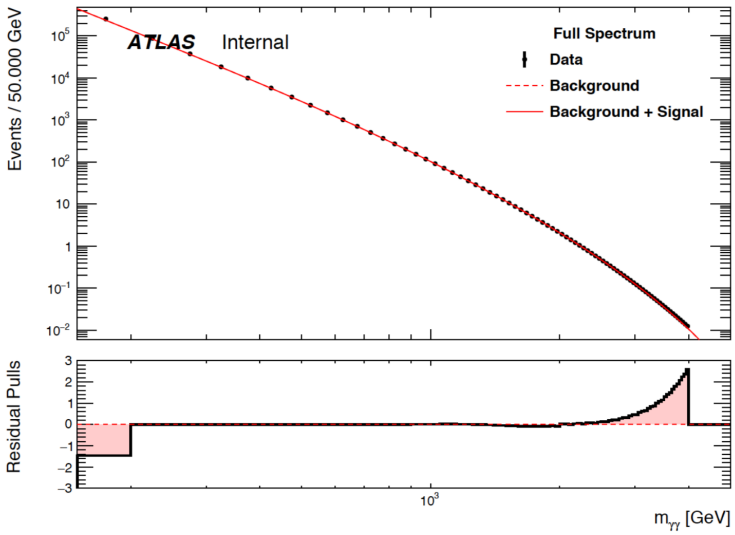


Figure 9.10: An FD S+B decomposition of the nominal asimov background using 16 moments.

is not expected to impact the results significantly.

Signal yields are extracted across the interval $[160, 3500]$ GeV using the `RooFit` framework by floating the normalization on the background and signal shapes to find the best fit. The nine Asimov curves, that is the nominal and eight systematics, are each decomposed into an FD S+B model. For each of these decompositions the number of background moments N is varied from between 10 and 23 moments. Figure 9.11 shows the extracted number of SS events, $|N_{\mathrm{SS}}|$, using 16 background moments. The local maxima are fit with a PowLog-2 function to approximate an envelope which encompasses the maxima. This envelope is taken to be a conservative estimate on the SS uncertainty for a decomposition with 16 background moments. Figure 9.12 shows the parameterized SS uncertainties, in number of events, found using this process for all tested values of N. Figure 9.13 shows the relative spurious signal, that is the number of spurious signal events $|N_{\mathrm{SS}}|$ over the uncertainty δS, for the case where FD uses $N = 16$ background moments. Figure 9.14 shows the parameterized relative spurious signal, $|N_{\mathrm{SS}}|/\delta\mathrm{S}$, which contributes to the systematic uncertainties of the resonance search entering the statistical model as a nuisance

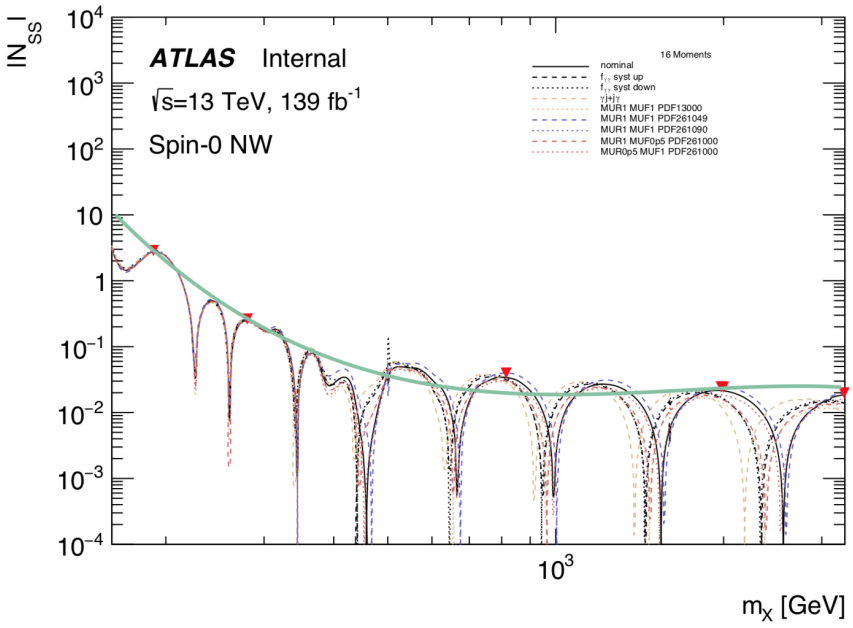


Figure 9.11: Spurious signal in number of events with $N = 16$ background moments for several systematic variations of the simulated background.

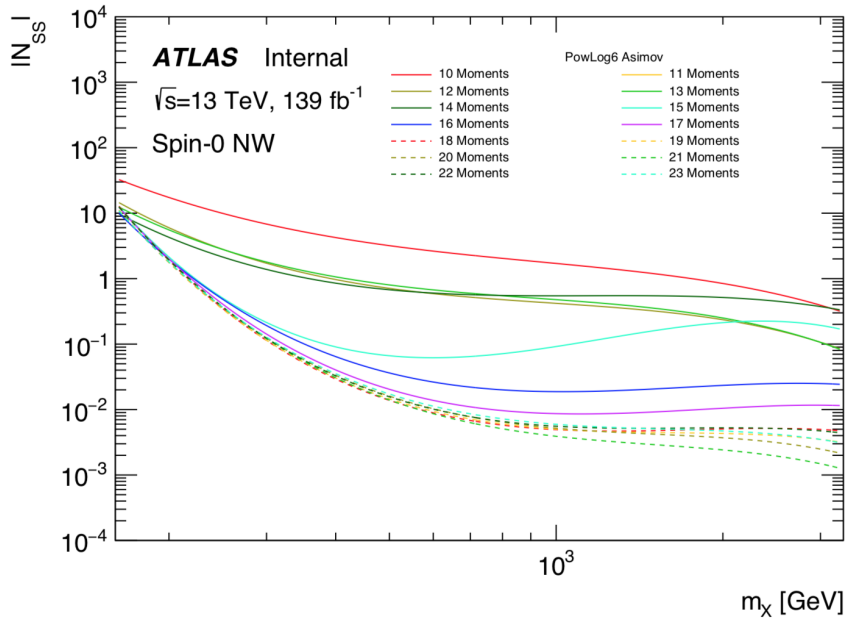


Figure 9.12: Parameterized spurious signal in number of events for 10-23 background moments

parameter.

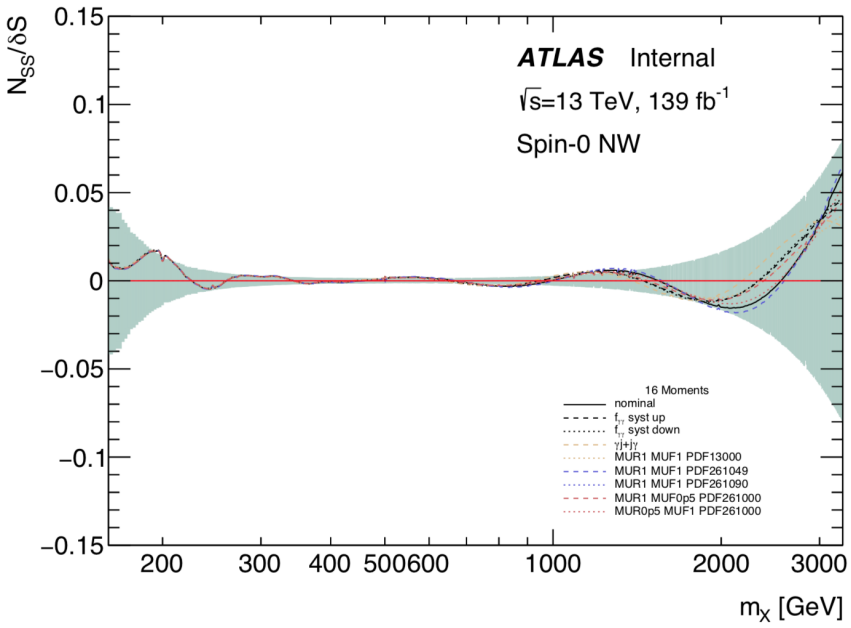


Figure 9.13: Relative spurious signal for the nominal case and the eight systematics with 16 background moments

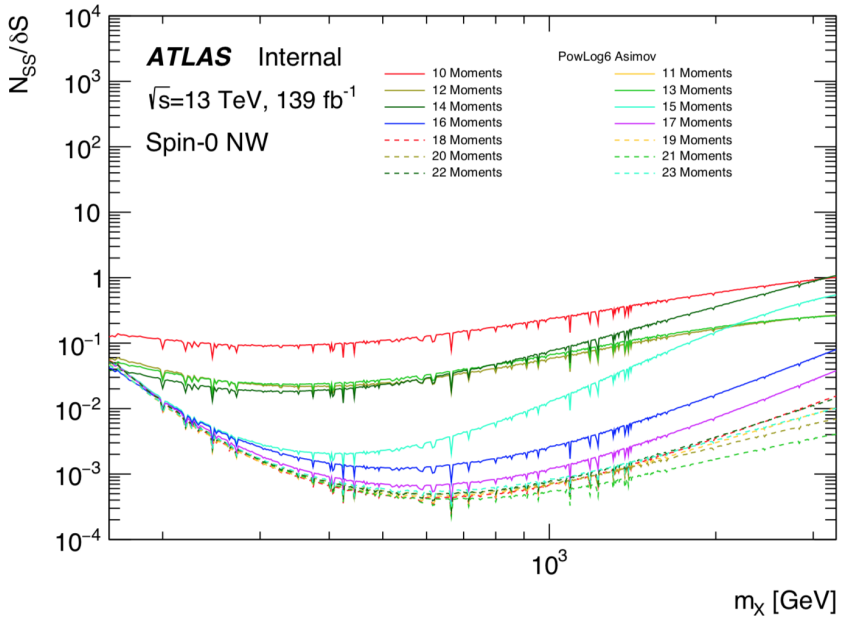


Figure 9.14: Parameterized relative spurious signal for the nominal case and the eight systematics for all integer values of $N \in [10, 23]$.

CHAPTER X

Systematic Uncertainites

10.1 Mass Scale and Resolution Uncertainties

The uncertainty in the diphoton mass scale and mass resolution is primarily influenced by the photon energy scale and energy resolution uncertainties. To quantify their impact on the diphoton mass scale and mass resolution, first the photon energy scale and energy resolution of MC signal samples are varied according to the uncertainties discussed in Section 4.3. Since the photon energy scale and energy resolution are each varied up and down there are four new invariant mass distributions for each of the four variations: photon energy scale varied up, photon energy scale varied down, photon energy resolution up, photon energy resolution down. Each of these mass distributions are compared to the nominal distribution and the deviation is used to determine their contribution to the uncertainty of the mass scale and mass resolution.

10.1.1 Photon Energy Scale Uncertainty

To quantify the impact of the photon energy scale uncertainty on the diphoton mass scale, the energies of the photons in the signal MC samples are first systematically varied. The `ElectronPhotonFourMomentumCorrection` package, provided by the ATLAS e/γ Group, provides variations by the using the `es2017_R21_v1` cor-

rection model and the `1N_v1` correleation model (*ATLAS* e/γ, 2020a). Figure 10.1 shows the nominal $m_{\gamma\gamma}$ distribution for the simulated signal with mass $m_X = 1$ TeV compared to the $m_{\gamma\gamma}$ distributions obtained by using photons whose energy scale has been systematically varied up and down.

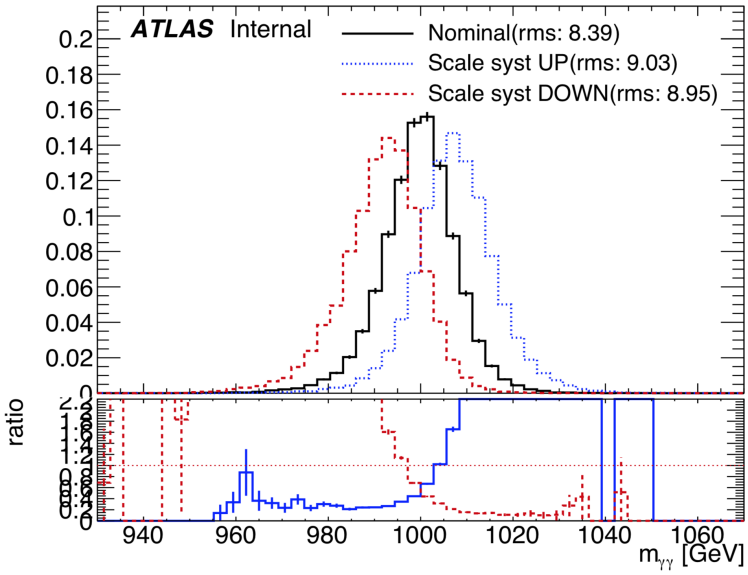


Figure 10.1: Invariant mass distribution of the $m_X = 1$ TeV simulated signal obtained with nominal and systematically varied energy scale calibration. The difference in the mean of the distributions with respect to the nominal are used to estimate the systematic uncertainty associated to the photon energy scale (*Amidei*, 2020).

This is repeated for several mass on the interval $m_X \in [200, 4000]$ GeV. Since $m^2_{\gamma\gamma} \propto E_1 E_2$, the change in the invariant mass can be estimated to first order as an overall shift in the distribution. Figure 10.2 shows the difference, and the relative difference, between the mean value of these distributions and m_X as a function of m_X.

The relative difference at a given m_X is taken to be the relative uncertainty, $\delta\mu_{\mathrm{CB}}^{\mathrm{up/down}}$, of the parameter μ_{CB} in the DSCB signal parameterization.Table 10.2 shows the relative uncertainty parameterized as an exponential function.

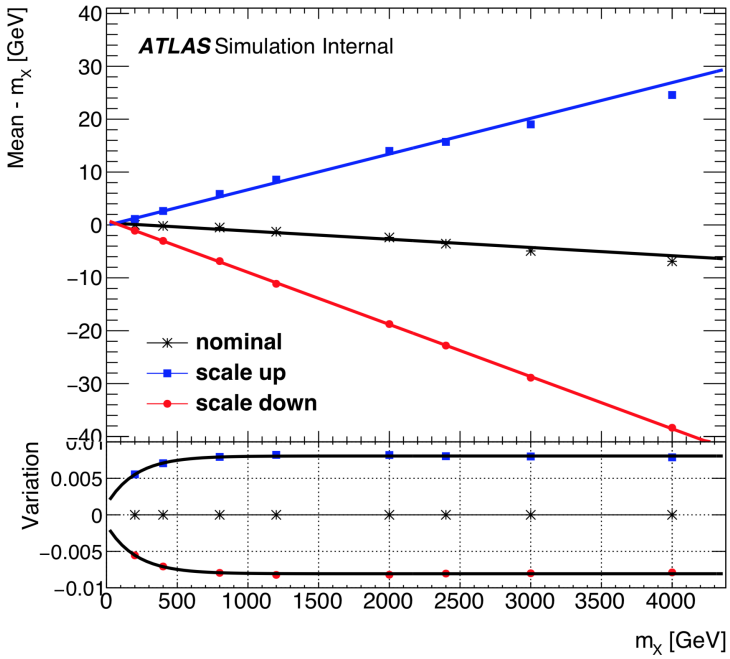


Figure 10.2: The mean of the invariant mass distribution as a function of m_X for the NW scalar samples for the nominal and systematically varied photon energy scales. The lower panel shows the difference between the nominal and shifted $m_{\gamma\gamma}$ distributions normalized by m_X (*Amidei*, 2020).

Variation	Parameterization
$\delta\mu_{\mathrm{CB}}^{\mathrm{up}}/\mu_{\mathrm{CB}}$	$0.0081 - 0.0066\exp(-0.0048 m_X)$
$\delta\mu_{\mathrm{CB}}^{\mathrm{down}}/\mu_{\mathrm{CB}}$	$-0.0081 + 0.0067\exp(-0.0048 m_X)$

Table 10.1: Parameterization of the systematic variation on the μ_{CB} parameter of the NW scalar signal model

10.1.2 Photon Energy Resolution Uncertainty

Recall Equation 3.14

$$\frac{\sigma(E_0)}{E_0} = \frac{a}{\sqrt{E_0}} \oplus \frac{b}{E_0} \oplus c \tag{3.14}$$

where E_0 is the energy of the photon, $\sigma(E_0)$ is the uncertainty in the energy, a is the sampling term, b is the noise term, and c is the constant term. Photons in the energy range considered in this analysis have a negligible noise term, so the equation is simplified by setting $b = 0$, and the equation for the relative uncertainty in the

photon energy resolution becomes

$$\frac{\sigma(E_0)}{E_0} = \frac{a}{\sqrt{E_0}} \oplus c. \tag{10.1}$$

To quantify the impact of the photon energy resolution uncertainty on the diphoton mass resolution, first the photon energy resolution of a signal MC sample is systematically varied. In the *resolution up* variation both resolution parameters a and c are systematically varied up, while in the *resolution down* variation both the resolution parameters are systematically varied down. Like with the photon energy scale determination described in Section 10.1.1, the variations are provided by the `ElectronPhotonFourMomentumCorrection` package (*ATLAS* e/γ, 2020a) using the same correction and correlation models. Figure 10.3 show the nominal $m_{\gamma\gamma}$ distribution for the simulated signal with mass $m_X = 1$ TeV compared to the $m_{\gamma\gamma}$ distributions obtained by using systematically photons.

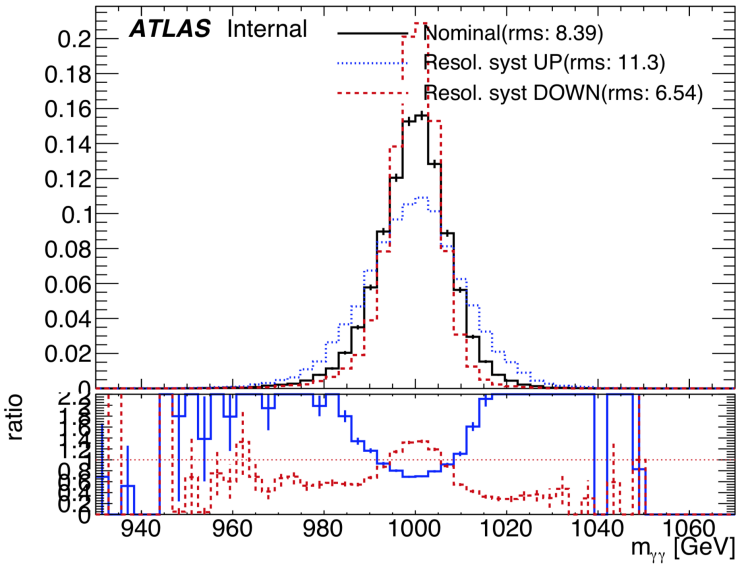


Figure 10.3: Invariant mass distribution for the simulated NWA signal with $m_X = 1000$ GeV for the nominal and systematicically varied photon energy resolution.

The width of the distributions are approximated by the smallest interval in $m_{\gamma\gamma}$ which contains 68% of the distribution. This interval, Q_{68}, is estimated by assuming

the distributions are roughly Gaussian and taking the difference between the 84^{th} and the 16^{th} percentile of the distribution. This is repeated several times for various $m_X \in [200, 4000]$ GeV. Figure 10.4 shows the Q_{68} (and the Q_{68} normalized by the Q_{68} of the nominal distribution) as a function of m_X.

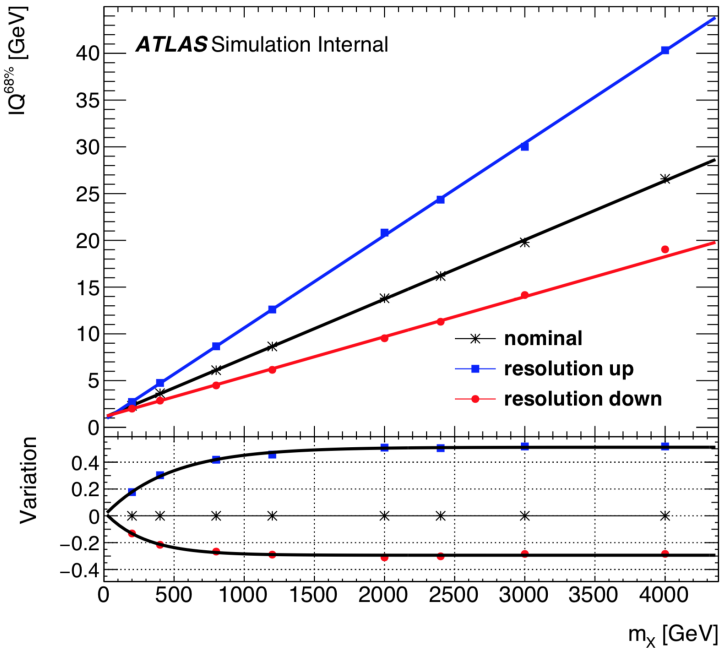


Figure 10.4: Q_{68} of the simulated NWA signal with $m_X = 1000$ GeV for the nominal and systematicically varied photon energy resolution is shown in the upper pannel with the lower pannel shows the Q_{68} normalized by the Q_{68} of the nominal case.

The normalized Q_{68} for the resolution up and resolution down variations are parameterized as exponential functions, and taken to be the relative uncertainty, $\delta\sigma_{\text{CB}}/\sigma_{\text{CB}}$, on the parameter σ_{CB}. Table 10.2 shows the parameterization for the resolution up and resolution down variations.

Variation	Parameterization
$\delta\sigma_{\text{CB}}^{\text{up}}/\sigma_{\text{CB}}$	$0.511 - 0.512\exp(-0.002m_X)$
$\delta\sigma_{\text{CB}}^{\text{down}}/\sigma_{\text{CB}}$	$-0.294 + 0.325\exp(-0.003m_X)$

Table 10.2: Parameterization of the systematic variation on the μ_{CB} parameter of the NW scalar signal model

10.2 Signal Yield Uncertainties

The systematic uncertainties on the signal yield are derived from several sources. Details about these sources of error are given below:

Photon Identification

The photon identification efficiency for the data taking periods used in this analysis is measured using three data driven methods providing data-to-MC ratios called *scale factors* (*Petit*, 2017). The values are provided by the ATLAS e/γ Group via the `PhotonEfficiencyCorrection` tool (*ATLAS* e/γ, 2020b). The uncertainties on these scale factors are propagated to the analysis by applying the corresponding variation.

Photon Isolation

The systematic uncertainty of the signal yield is obtained by applying a shift to the calorimeter isolation and a p_{T} dependent shift to the track isolation (*De Vivie De Regie et al.*, 2017). The shifts on the correction factor, C_X, from these two sources are calculated independently then added in quadrature.

Photon Trigger

The trigger efficiency is measured with data for single-photon triggers and is found to be well modeled in the MC simulation. For the diphoton triggers used in the analysis, a bootstrap method is used to estimate the efficiency, resulting in an efficiency for the full dataset of about 98% for the 125 GeV Higgs boson, with an associated uncertainty from this estimation of about 0.5%. This estimation is taken to be the systematic uncertainty assigned to the photon trigger. As the efficiency is expected to be even higher for larger diphoton mass signals, and given the good modeling of the trigger efficiency by the MC simulation, this assumption can be considered sufficiently conservative. (*Amidei*, 2020)

Pile-up Reweighting

A variation in the pile-up reweighting of the simulation is performed to cover the difference between the predicted and measured inelastic pp cross-section in the fiducial volume of the detector. This variation is performed by shifting the μ distribution derived from the data by $\pm 3\%$ before reweighting the MC sample. The uncertainty associated with this reweighting affects the signal yield by up to 2% and the effect on the signal shape is found to negligible. (*Amidei*, 2020)

Physics Bias

The current analysis makes no assumption on the possible production modes of the hypothetical scalar. As the ggF, VBF, top-associated, and vector boson associated production modes have significantly different kinematics that lead to a significant change in the correction factor, a systematic uncertainty is assigned to account for the "physics bias" introduced by the assumed kinematics of the signal model. This uncertainty is taken from the envelope of the difference in the correction factor estimated from the different production modes. In practice, only the difference between the ggF, VBF and top-associated production modes is considered as the differences between the vector-boson associated production mode and ggF production mode are negligible. (*Amidei*, 2020)

Photon Energy Scale/Resolution

Apart from the impact on the signal model, the photon energy scale and resolution uncertainties may impact the determination of the signal yield through the correction factor. The energy scale and resolution relate to the migrations in and out of the fiducial volume. Such effects are insignificant in this analysis, so the correction factor is hardly affected by this source of uncertainty. Figure 10.5 shows the systematic uncertainty assigned on the correction factor to account for the physics bias introduced by the assumed signal model.

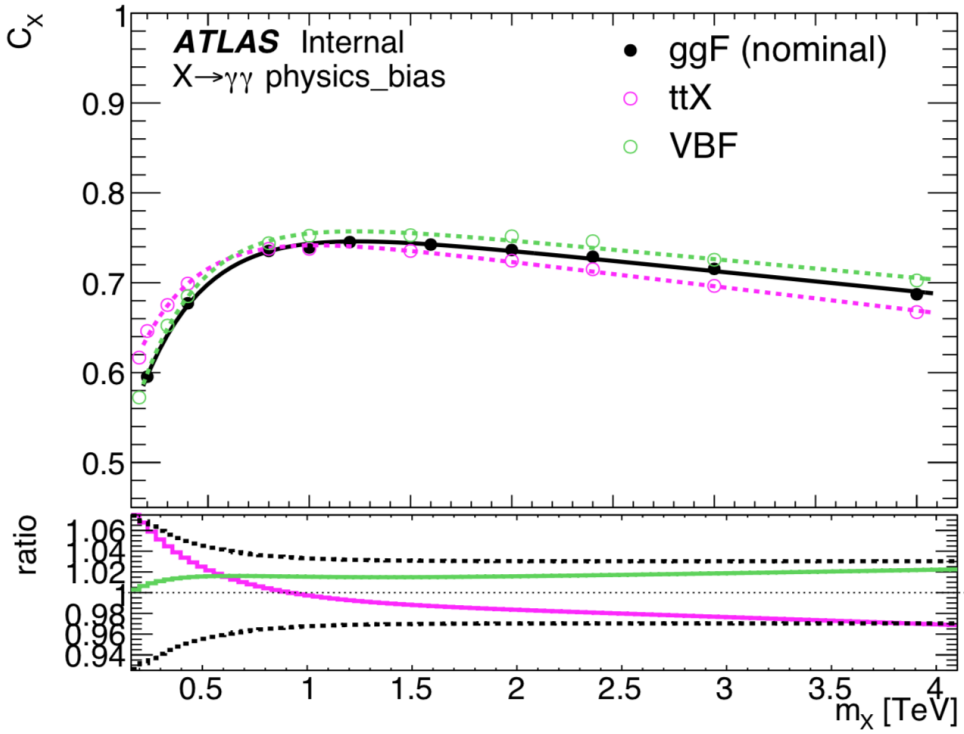


Figure 10.5: Effect of different signal kinematics associated with the assumed production mode on the correction factor. The difference observed between the ggF and ttX and VBF production modes is used to define an envelope to serve as systematic uncertainty associated with the assumption of the production mode, shown as a function of m_X with the dashed black lines in the ratio panel.

CHAPTER XI

Seach for New Phenomena with Diphoton Final States

Due to numerical challenges in implementing FD as the search method and pressure to complete the analysis from an internally imposed deadline, the analysis was completed with a standard method which was able to move forward quickly to meet the deadline. Although not used as the search method, FD was still instrumental in the analysis. FD was used for smoothing background templates used to determine the uncertainties on the spurius signal. The idea being that since FD can model any curve on semi-infinite interval with a finite area underneath, simulated background templates are replaced with an FD background only model of the template. The assumption is that the true shape of the template's pdf is sufficiently represented while eliminating any statistical fluctuations from the simulation, giving a more accurate representation of the PowLog-n's ability to capture the shape of the underlying pdf.

This chapter describes the spin-0 and spin-2 high mass diphoton resonance search procedure and results. The spin-0 search explores the volume 400 GeV $\leq m_X \leq$ 2800 GeV and $0 \leq \Gamma_X/m_X \leq 0.1$ while the spin-2 search explores the volume 500 GeV $\leq m_X \leq$ 5000 GeV and $0.01 \leq k/\overline{M}_{\mathrm{Pl}} \leq 0.1$. In both cases, the diphoton

mass spectrum is modeled with

$$\Omega(z;m) = f(z;\{a,d,\alpha_i\}) + N \cdot F_{X/G^*}(z;m) \tag{11.1}$$

where F_{X/G^*} is a DSCB and represents the signal shape, Equation 6.7, f is a PowLog-n function and represents the background, Equation 6.6, m is the mass of the hypothesized resonance, and N is the number of signal events.

This chapter continues by describing how the likelihood of a given fit is determined accounting for each systematic uncertainty effecting the observed number of events. Next, it describes how the local p-value, local significance, and global significance of the background only hypothesis, as well as the upper 95% CL on the scalar (graviton) fiducial production cross section times the branching ratio $\mathrm{Br}(X(G^*) \to \gamma\gamma)$ are computed in this analysis. Finally, it discusses the results and presents the upper 95% CL as a function of m_X and Γ_X/m_X $\left(k/\overline{M}_{\mathrm{Pl}}\right)$ for the spin-0 (spin-2) search.

11.1 Statistical Procedure

11.1.1 Background Only Fit of the Diphoton Mass Spectrum

The background is fit with a PowLog-1 function, given by Equation 6.6

$$f(z;\{a,d,\alpha_i\}) = (1-x^d)^a \cdot x^{\sum_{i=0}^{n}\alpha_i(\log x)^i} \tag{6.6}$$

with $n = 1$ and $d = 1/3$ leading to

$$f(z;\{a,\alpha_0,\alpha_1\}) = (1-x^{1/3})^a \cdot x^{\alpha_0+\alpha_1\cdot\log x}. \tag{11.2}$$

Table 11.1 gives the parameters of the PowLog-1 function for the background only hypothesis.

Fit Parameter	Value
a	10.468 ± 0.156
α_0	-1.8375 ± 0.0675
α_1	0.1378 ± 0.0087

Table 11.1: Parameters for the PowLog-1 function fit to the diphoton invariant mass spectrum (*Amidei*, 2020).

The background only fit plotted with the data is shown in Figure 11.1. Excluding bins with fewer than ten entries ($m_{\gamma\gamma} > 1.4$ TeV) since they tend to bias the estimation, the reduced χ^2 for the background only fit is $\chi^2/N_{\text{dof}} = 37.2/54$, corresponding to a p-value of $p(\chi^2) = 0.96$.

11.1.2 Accounting for Systematic Uncertainties

Each uncertainty is accounted for in the S+B fits by using nuisance parameters contrained by Gaussian penalty terms in the likelihood function

$$\begin{aligned}\mathscr{F}\left(m_{\gamma\gamma};\sigma,m_X,\alpha_X,N_b,\boldsymbol{a},\boldsymbol{\theta}\right) = f_X\left(m_{\gamma\gamma};\boldsymbol{x}_X\left(m_X,\alpha_X\right),\theta_\sigma\right)\cdot N_X\left(\sigma;m_X,\boldsymbol{\theta}_{N_X},\theta_{\text{SS}}\right)\\ + f_b\left(m_{\gamma\gamma};\boldsymbol{a}\right)\cdot N_b\end{aligned} \tag{11.3}$$

where f_X and f_b are the pdfs for a given signal X and background b respectively, N_X and N_b are the corresponding yields, σ is the fiducial cross section times the branching ratio of the hypothetical resonance, $\boldsymbol{a}$ are background shape parameters, $\boldsymbol{x}_X$ are the DSCB parameters, θ_{SS} is the spurious signal systematic, and $\boldsymbol{\theta}_{N_X}$ is the set of nuisance parameters for the systematic uncertainties impacting N_X which are listed below:

- θ_{lumi} is the uncertainty on the intergrated luminosity of the data sample;
- $\theta_{\text{eff},X}$ is the systematic uncertainty on the photon identification on the resonance;
- $\theta_{\text{iso},X}$ is the systematic uncertainty on the photon isolation on the resonance;

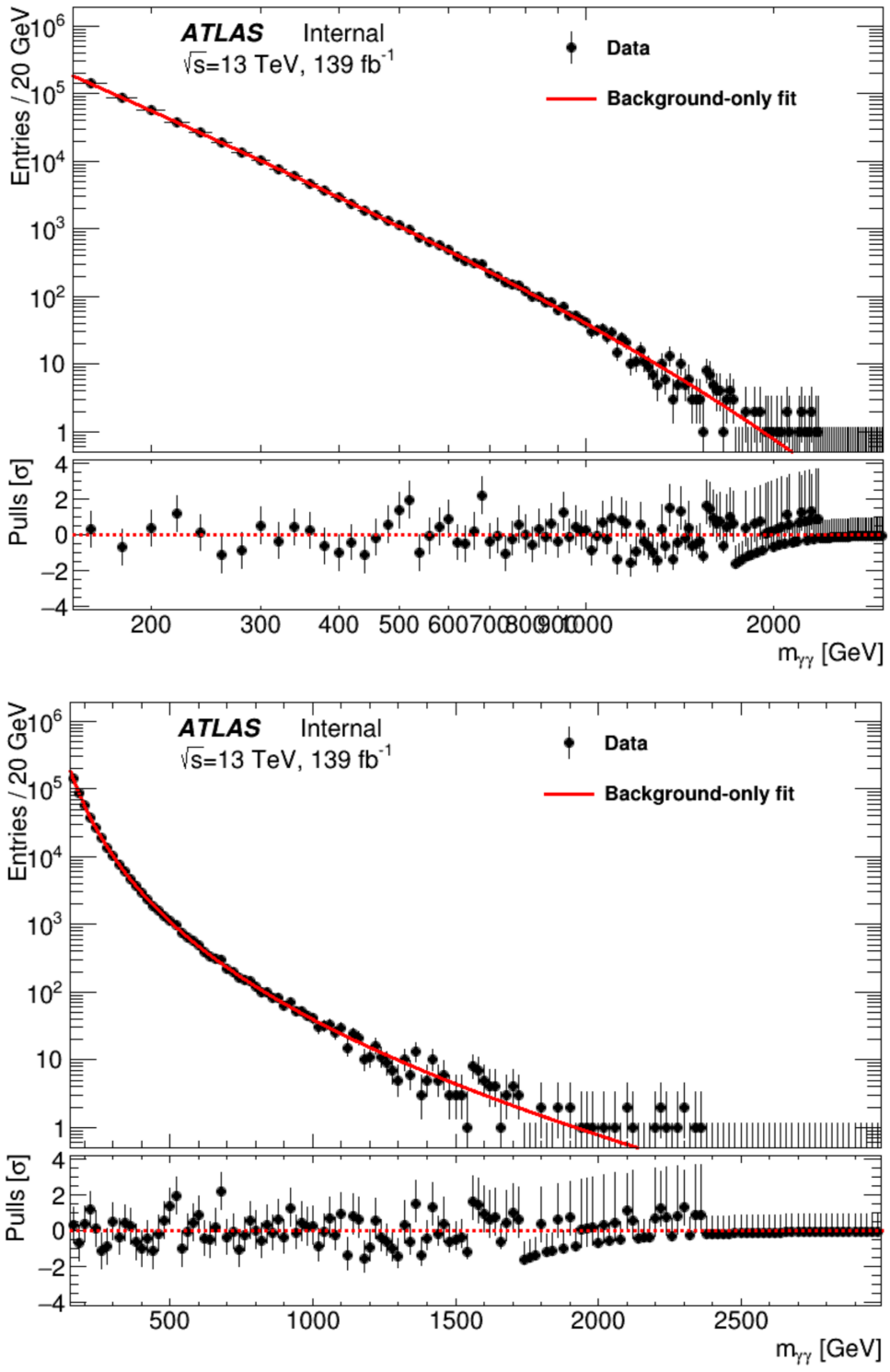


Figure 11.1: The diphoton invariant mass spectrum plotted with the background only fit along with three NWA signals at 400 GeV, 1000 GeV, and 2000 GeV.

- θ_{SS} is the spurious signal systematic;
- θ_{ES} is the photon energy scale systematics;
- θ_{ER} is the photon energy resolution systematics;
- θ_{C_X} is the production mode uncertainty on the C_X factor.

The quantity N_b is a free parameter in the fit, while N_X is parameterized as

$$N_X\left(\sigma, m_X, \boldsymbol{\theta}_{N_X}, \theta_{\mathrm{SS}}\right) = \sigma \cdot L \cdot C_X\left(m_X\right) \cdot \prod_{k}^{\dim \boldsymbol{\theta}_{N_X}} K_k\left(\theta_k\right) + \delta_{\mathrm{SS}} \cdot \theta_{\mathrm{SS}} \tag{11.4}$$

where L is the integrated luminosity of the sample, $C_X(m_X)$ is the correction factor at the mass m_X, $\delta_{\mathrm{SS}} = |N_{\mathrm{SS}}|$ and θ_{SS} are the values of the Spurious signal (SS) (see Section 9.8.2.1) and its associated nuisance parameter, the index k runs over the set of systematic uncertainties impacting N_X, and where K_k is a function characterizing the effect on the k^{th} normalization systematic given by

$$K_k(\theta_k) = \left[r_k(m_X)\right]^{\theta_k} \tag{11.5}$$

with

$$r_k(m_X) = \begin{cases} \frac{N_{+k}(m_X)}{N_X(m_X)} \Leftarrow \theta_k > 0 \\ \frac{N_X(m_X)}{N_{-k}(m_X)} \Leftarrow \theta_k < 0 \end{cases} \tag{11.6}$$

where $N_X(m_X)$ is the number of nominal signal events in the m_X simulated sample, and $N_{\pm k}$ is the number of signal events for the k^{th} up (down) systematic. This normalization ensures that $\theta_k = \pm 1$ corresponds to the $\pm 1\sigma$ variations used to define the uncertainties. To avoid numerical problems at $\theta_k = 0$ the expression for r_k used in the computation is interpolated smoothly between the cases $\theta_k > 0$ and $\theta_k < 0$ using the `RooStats::HistFactory::FlexibleInterpVar` class, with modifications

to allow mass-dependent values for the uncertainties.

$$\mathcal{L}\left(\sigma, m_X, \alpha_X, N_b, \boldsymbol{a}, \boldsymbol{\theta}\right) = e^{-(N_X+N_b)} \left[\prod_{i=1}^{n} \mathscr{F}\left(m_{\gamma\gamma_i}; \sigma, m_X, \alpha_X, N_b, \boldsymbol{a}, \boldsymbol{\theta}\right)\right] \times \prod_{k=1}^{\dim\theta} \exp\left(-\frac{1}{2}\left(\theta_k - \theta_k^{\text{aux}}\right)^2\right) \tag{11.7}$$

11.1.3 The Look-Elsewhere Effect

In the scalar (graviton) analysis, a search is performed for a resonance of unknown mass m and width Γ (coupling $k/\overline{M}_{\text{Pl}}$), and signal size μ on a smoothly falling background distribution with nuisance parameters $\hat{\nu}$. A test statistic, q_0, is constructed by comparing the likelihood of the background only hypothesis to the likelihood of the S+B hypothesis

$$q_0(m, \Gamma) = -2 \log \frac{\mathcal{L}\left(0, m, \Gamma, \hat{\nu}'\right)}{\mathcal{L}\left(\hat{\mu}, m, \Gamma, \hat{\nu}\right)} \tag{11.8}$$

where Γ represents the scalar width or graviton coupling $k/\overline{M}_{\text{Pl}}$, $\hat{\nu}'$ are the parameters for the best fit background only model and $\hat{\mu}$ and $\hat{\nu}$ are the parameters for the best fit signal plus background model for a specified mass m and width Γ.

In the asymptotic limit, the q_0 test statistic has a distribution that is defined by $\chi^2(N_{\text{dof}} = 1)$. The significance in this case is $Z = \sqrt{q_0}$ and the probability is $p_0 = 1 - \Phi(\sqrt{q_0})$ where Φ is the cumulative distribution function (cdf) (*Cowan et al.*, 2011).

The test statistic $q_0(m, \Gamma)$ is computed at each mass point for a given width, Γ, in the search and used to compute the local significance Z_0^{local}. The quantity $q_{Z_{\max}}$ is selected by locating the maximium Z_0^{local} value, $Z_{0,\max}^{\text{local}}$, and identified as its corresponding q_0 value.

Since both width and mass are being varied in the search, there are two addi-

tional degrees of freedom which provide a larger parameter space to be explored thus increasing the possibility of containing a *deeper* fit minimum. This effect results in larger Z_0^{local} values than the corresponding p_0 probability would suggest and is known as the *look-elsewhere effect*.

11.1.4 Global Significance with Psuedo Experiments

Using fits to background only pseudo experiment ensembles, the global significance, Z_0^{global}, is computed as a function of Z_0^{local}. The procedure begins with generating 1,000 pseudo experiments whose inital parameters are taken from the PowLog-1 function described in Section 7.1.2. The global observables of the function are randomly varied according to a Gaussian pdf with a mean value equal to the profiled value of the corresponding nuisance parameter. Values of the experimental observables (N and $m_{\gamma\gamma}$) are generated randomly from the background only pdf and the Poisson distribution.

Next, the maximum local significance in the search volume is computed. As it would be computationally expensive to determine Z_0^{local} for each mass point and width combination in the analysis, a smaller subset of points are randomly selected from each psuedo experiment and the minimum log likelihood computed within that subset is taken to be $Z_{0,\text{max}}^{\text{local}}$ for the corresponding pseudo experiment. To determine the size of the subsets used, the median significance and standard deviation of the significance of the toys are plotted against the size of the subsets. These plots are shown in Figure 11.2 for both the spin-0 and spin-2 searches and show that after the size of the subsets reaches about 50, that these distributions appear to be asymptotic. Therefore 50 is taken to be an adequately large subset size which is still significantly smaller than $1201 \times 21\,(1151 \times 19)$, the number of mass points $\times$ number of width points, of the entire scalar (graviton) search volume, reducing the computational cost of the optimization by approximately $99.80\%\,(99.77\%)$.

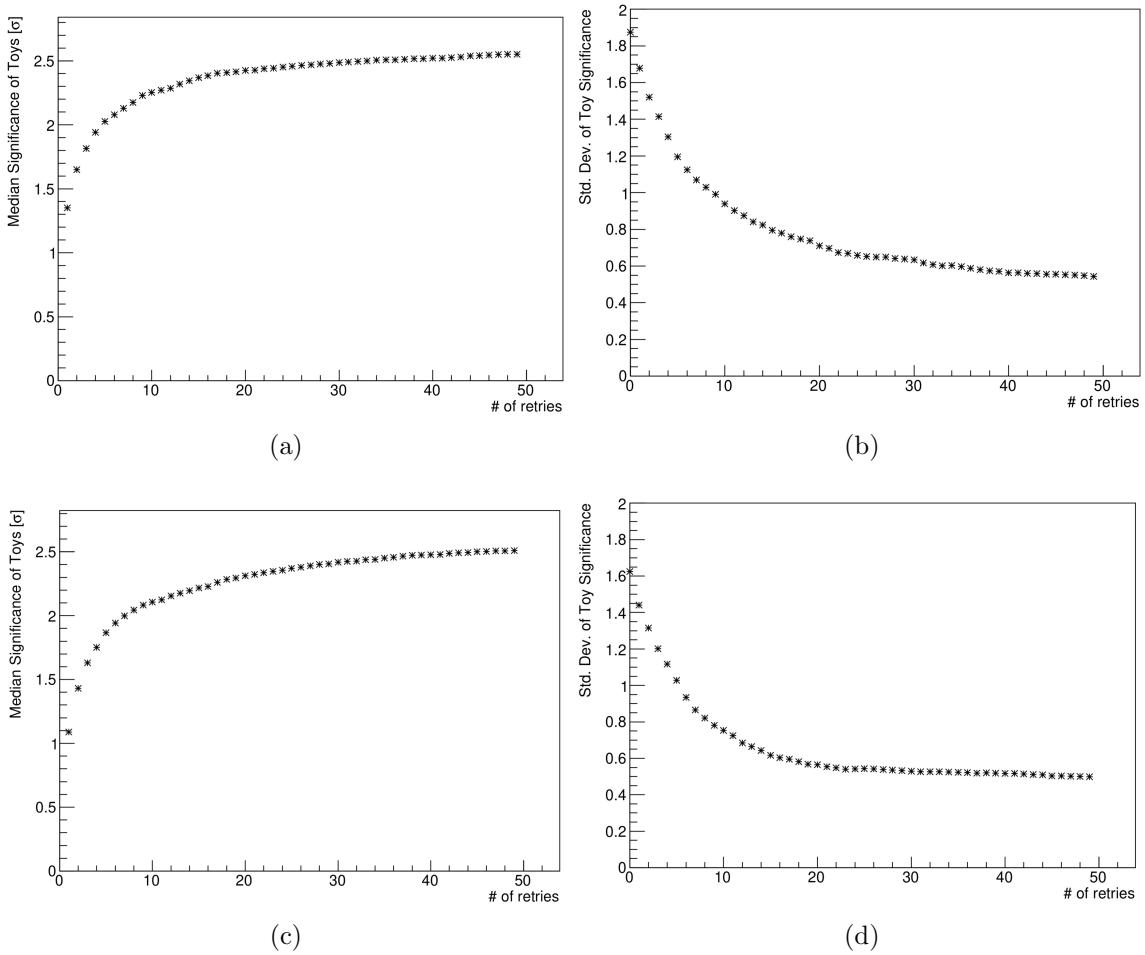
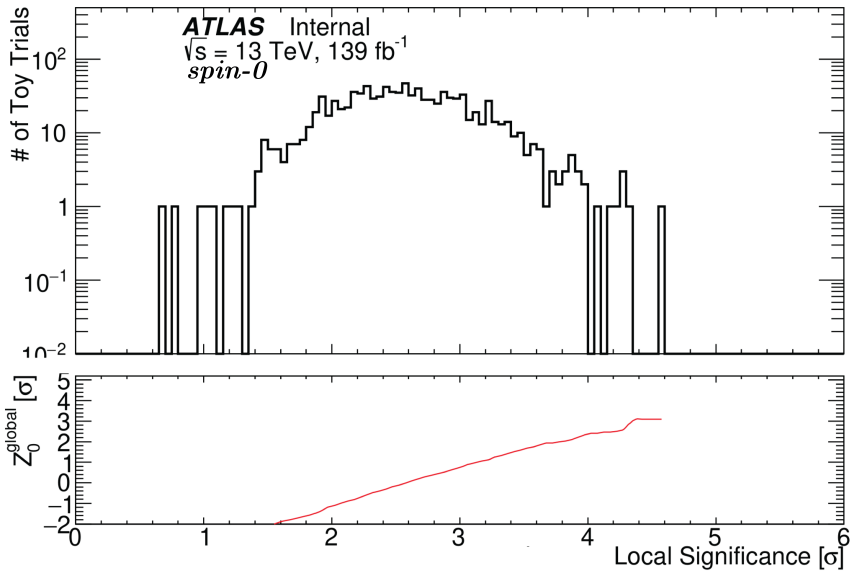


Figure 11.2: The median and standard deviation of the local significance Z_0^{local} as a function of subset sizes for the (a,b) spin-0, and the (c,d) spin-2 analyses (*Amidei*, 2020).

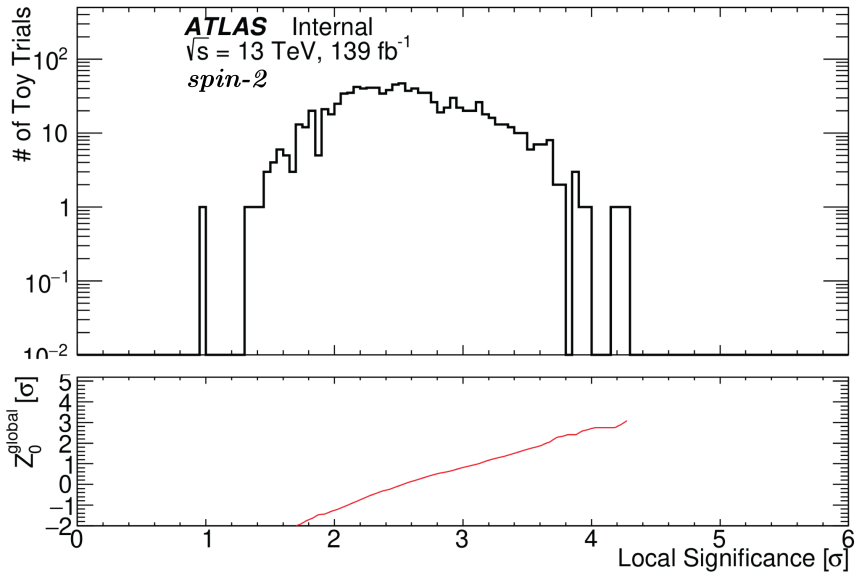
The global significance corresponding to a given $Z_0^{\text{local,observed}} = \widetilde{Z}$ is given by

$$Z_0^{\text{global, observed}} = \int_{\widetilde{Z}}^{\infty} H(Z_0^{\text{local}})\, dZ_0^{\text{local}} \tag{11.9}$$

where H represents the Z_0^{local} distribution. The Z_0^{local} distribution for the scalar and graviton analyses are shown in Figure 11.3 for 1,000 pseudo experiments and a subset size of 50. These empirically derived distributions are used as H for the determination of the global significance. Errors for the global significance are given by the Binomial



(a)



(b)

Figure 11.3: The local significance Z_0^{local} distributions for 1,000 pseudo datasets for the (a) spin-0 and (b) spin-2 analyses (*Amidei*, 2020).

distribution based on the number of pseudo experiments and the local observed p-value, $p_0^{\text{local,observed}}$.

11.1.5 Confidence Limits

The probability of observing d events given an expected measurement of $s + b$ signal plus background events is given by the Poisson distribution

$$\mathcal{P}(d; s, b) = \frac{(s+b)^d}{d!} e^{-(s+b)}. \tag{11.10}$$

A test statistic X can be constructed which discriminates signal-like outcomes from background-like outcomes. An optimal choice for the statistic is given by

$$X = \prod_{i=1}^{n} X_i \tag{11.11}$$

where i indexes the mass hypothesis and

$$X_i = \frac{\mathcal{P}(d = d_i; s = s_i, b = b_i)}{\mathcal{P}(d = d_i; s = 0, b = b_i)} \tag{11.12}$$

$$= \frac{(s_i + b_i)^{d_i}}{b_i^{d_i}} e^{-s_i} \tag{11.13}$$

is the ratio of the S+B and background only Poisson distribution for the i^{th} mass hypothesis, s_i is the estimated signal, b_i is the estimated background, d_i is the number of observed candidates.

The confidence level for excluding the S+B hypothesis with test statistic X_{observed} is

$$\text{CL}_{s+b}(X_{\text{observed}}) = \mathscr{P}_{s+b}(X \leq X_{\text{observed}}) \tag{11.14}$$

with

$$\mathscr{P}_{s+b}(X \leq X_{\text{observed}}) = \sum_{\widetilde{d}_i \in \Delta} \frac{e^{-(s_i+b_i)} (s_i + b_i)^{\widetilde{d}_i}}{\widetilde{d}_i!} \tag{11.15}$$

where $\Delta = \left\{ \widetilde{d}_i \middle| X_i \leq X_{\text{observed}} \right\}$ is the set of all possible outcomes, $\widetilde{d}_i$, where the test statistic is smaller than that of the observed outcome. The confidence level for the background alone is given by

$$\text{CL}_b = \mathscr{P}_b \left(X \leq X_{\text{observed}} \right) \tag{11.16}$$

where the probabilities assume there is no signal present. The Modified Frequentist confidence level CL_s is given by

$$\text{CL}_s = \text{CL}_{s+b} / \text{CL}_b. \tag{11.17}$$

However, in practice evaluating Equation 11.17 is cumbersome, being on the order of $\mathcal{O}\left(n^m\right)$ where n is the number of channels and m is the number of possible outcomes in that channel. As a result, the upper limit on the expected and observed confidence levels are computed based on the asymptotic formulas to determine the cross section value corresponding to the 95% CL exclusion (*Cowan et al.*, 2011).

11.2 Results

Tables 11.2 and 11.3 summarize the systematic uncertainties associated with this analysis which are discussed in Chapter 10.

Source	Uncertainty
Signal Yield	
Luminosity (2015–2018)	$\pm 1.7\%$
Trigger	$\pm 0.5\%$
Photon Identification	$\pm 0.5\%$
Isolation Efficiency	$\pm 1.5\%$
Pile-up Reweighting	$1 \pm 0.03 \exp(-2m_X[\text{TeV}])\,\%$ $\Rightarrow \pm(0.2\text{–}2.0)\%$
Scalar Production	$(1 \pm 0.03) \pm 0.08 \exp(-3.35m_X[\text{TeV}])\,\%$ $\Rightarrow \pm(3.0\text{–}7.0)\%$
Photon Energy Scale	negligible
Photon Energy Resolution	negligible
Signal Modeling	
Photon Energy Resolution	$^{+14\%}_{-9.3\%}$ (at $m_X = 200$ GeV) $^{+51\%}_{-29\%}$ (at $m_X = 2000$ GeV)
Photon Energy Scale	$\pm 0.5\%$– $\pm 0.6\%$
Pile-up Reweighting	negligible

Table 11.2: A summary of the systematic uncertainties associated with the signal yield and signal modeling which effect the observed number of events.

Source		
Background		
Spurious Signal (spin-0)		
NWA	144 events	(at $m_{G^*} = 160$ GeV)
	0.04 events	(at $m_{G^*} = 2800$ GeV)
$\Gamma_X/m_{G^*} = 2\%$	107 events	(at $m_{G^*} = 400$ GeV)
	0.14 events	(at $m_{G^*} = 2800$ GeV)
$\Gamma_X/m_{G^*} = 6\%$	223 events	(at $m_{G^*} = 400$ GeV)
	0.38 events	(at $m_{G^*} = 2800$ GeV)
$\Gamma_X/m_{G^*} = 10\%$	437 events	(at $m_{G^*} = 400$ GeV)
	0.50 events	(at $m_{G^*} = 2800$ GeV)
Spurious Signal (spin-2)		
$k/\overline{M}_{\text{Pl}} = 0.01$	4.71 events	(at $m_{G^*} = 500$ GeV)
	0.04 events	(at $m_{G^*} = 2800$ GeV)
$k/\overline{M}_{\text{Pl}} = 0.05$	19.0 events	(at $m_{G^*} = 500$ GeV)
	0.09 events	(at $m_{G^*} = 2800$ GeV)
$k/\overline{M}_{\text{Pl}} = 0.10$	31.2 events	(at $m_{G^*} = 500$ GeV)
	0.20 events	(at $m_{G^*} = 2800$ GeV)

Table 11.3: A summary of the systematic uncertainties associated with the background modeling which effect the observed number of events.

11.2.1 Scalar Results

Based on the procedure outline in Section 11.1, the 95% CL on the parameter σ_{fid} times the branching ratio BR is computed every 2 GeV on the interval $m_{\gamma\gamma} \in [400, 2800]$ GeV and steps of 0.5% on the interval $\Gamma_X/m_X \in [0, 10]\%$. Figure 11.4 shows the local p-value as a function of m_X for various signal widths.

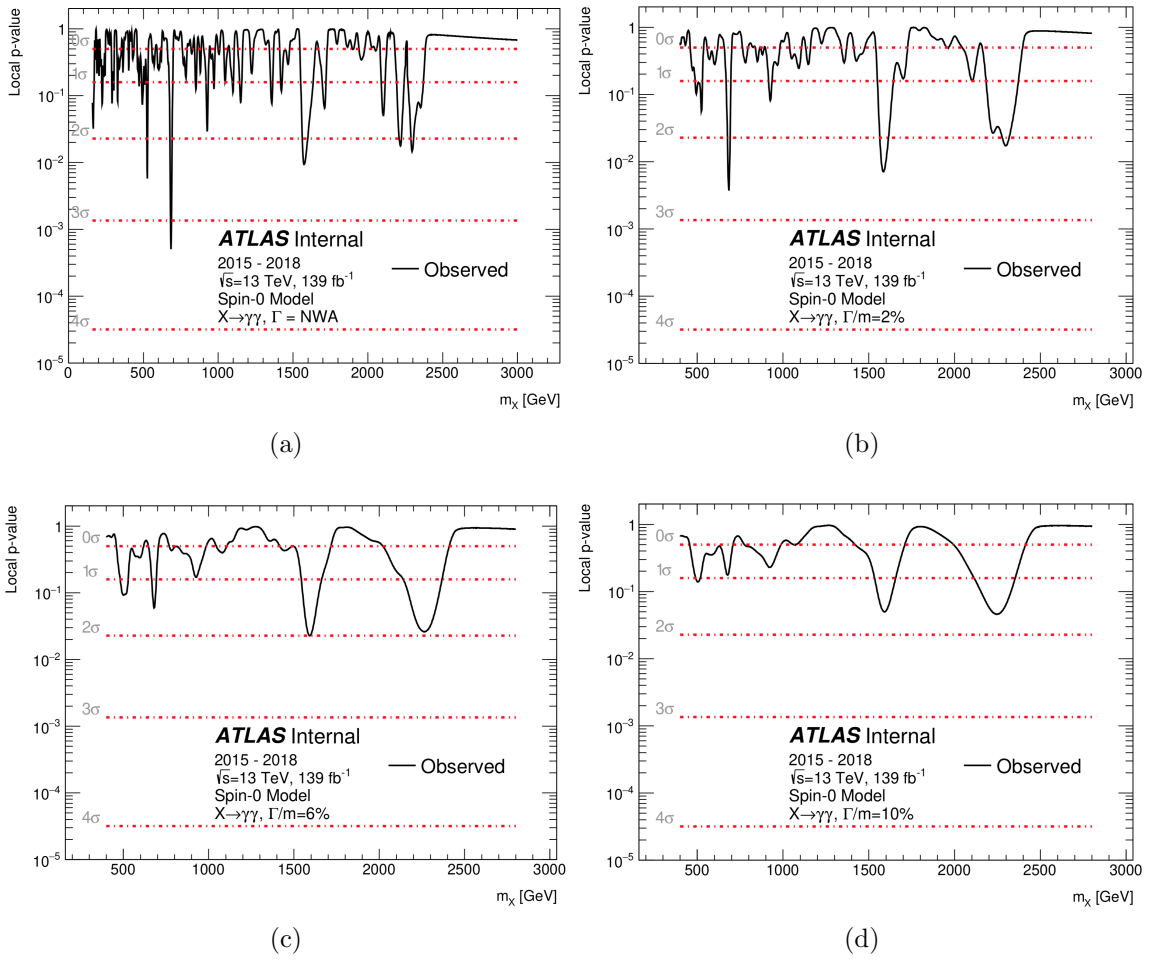


Figure 11.4: The local p-value as a function of m_X for the scalar search for various widths: (a) NWA, (b) $\Gamma_X/m_X = 2\%$, (c) $\Gamma_X/m_X = 6\%$, (d) $\Gamma_X/m_X = 10\%$.

Figure 11.5 shows a two-dimensional plot of local significance as a function of width Γ_X/m_X and mass m_X. The most signficant excess was observed for the $m_X = 684$ GeV mass hypothesis for the NWA model, corresponding to a 3.29σ local significance and a $(1.30 \pm 0.06)\,\sigma$ global significance.

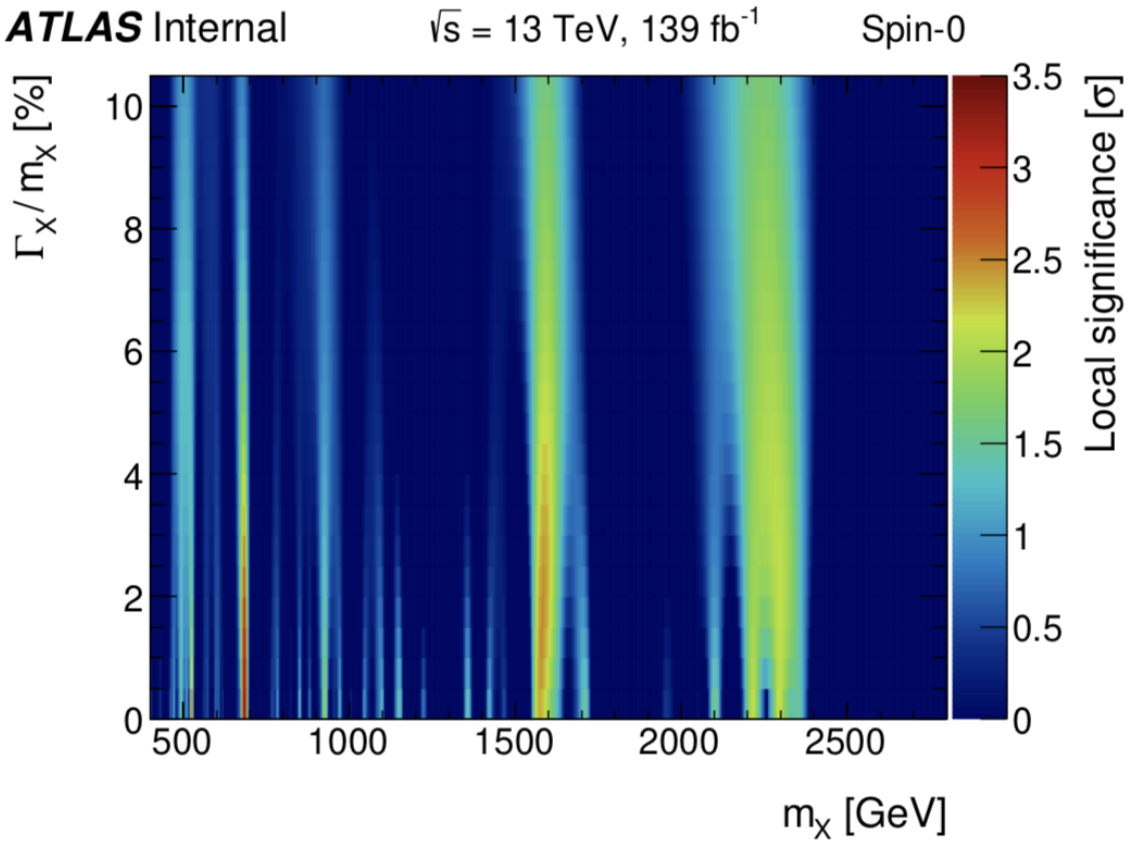


Figure 11.5: The local significance Z_0^{local} as a function of width Γ_X/m_X and invariant mass m_X.

Figure 11.6 shows several plots of the 95% CL on the parameter σ_{fid} times the branching ratio BR as a function of m_X for and several widths Γ_X/m_X, while the expected and observed limits as a function of width and invariant mass are shown in Figure 11.7.

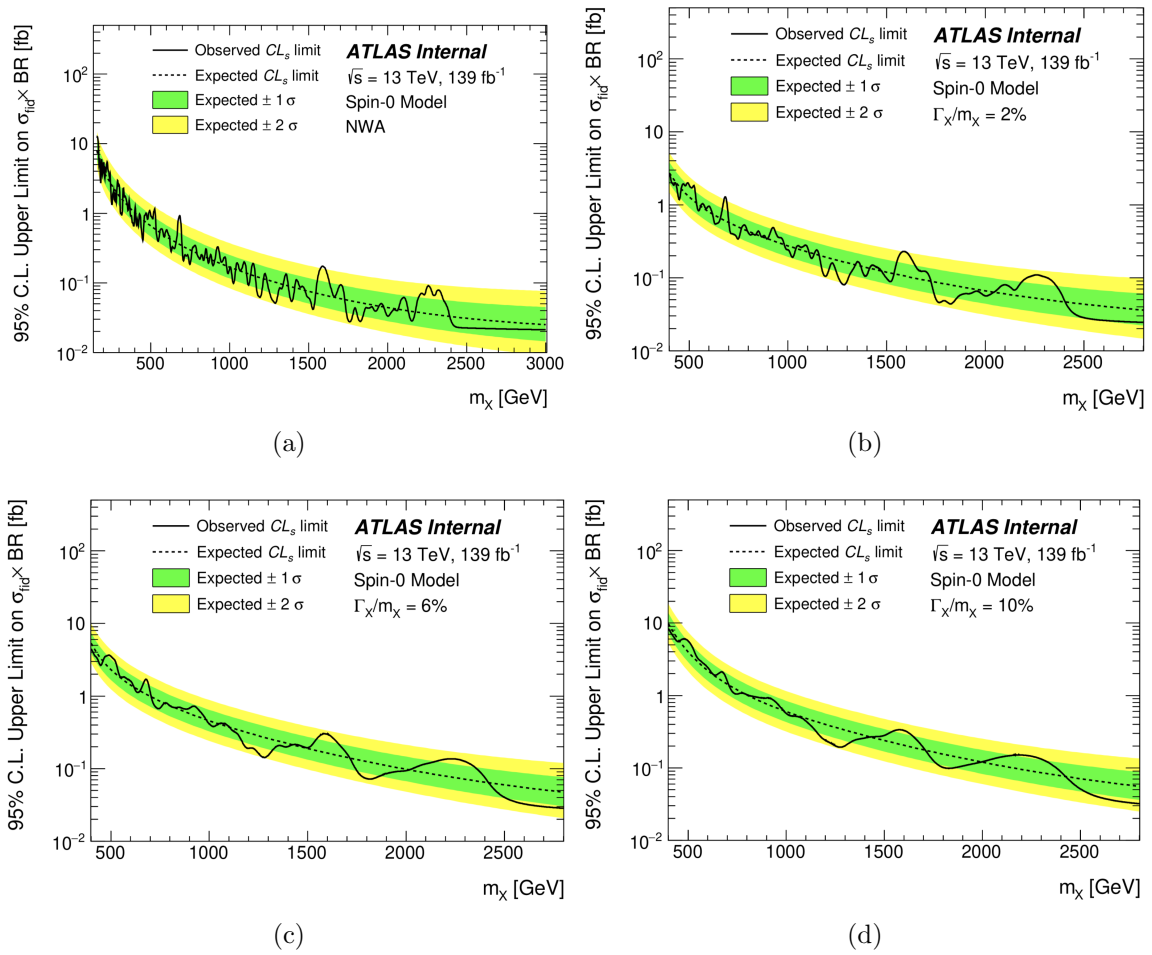
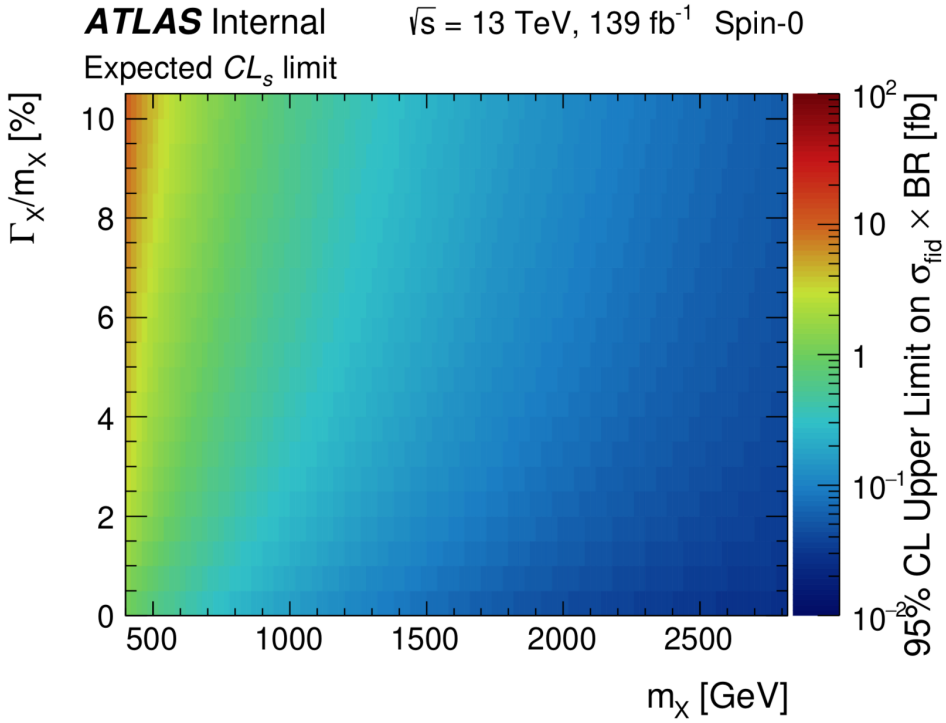
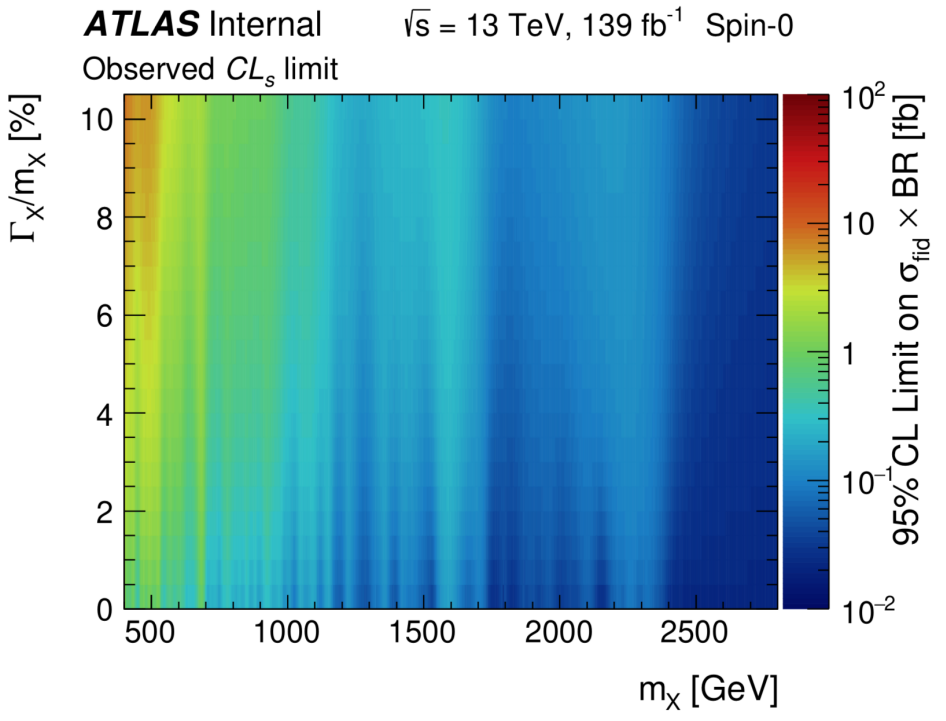


Figure 11.6: Expected and observed limit on the fiducial production cross section times branching ratio, $\sigma_{\text{fid}} \times \text{BR}$, as a function of the resonance mass m_X for various widths: (a) NWA, (b) $\Gamma_X/m_X = 2\%$, (c) $\Gamma_X/m_X = 6\%$, (d) $\Gamma_X/m_X = 10\%$.



(a)



(b)

Figure 11.7: Expected and observed limit on the fiducial production cross section times branching ratio, $\sigma_{\text{fid}} \times \text{BR}$, as a function of the resonance mass m_X and width Γ_X/m_X.

11.2.2 Graviton Results

Based on the procedure outline in Section 11.1, the 95% CL on the parameter σ_{fid} is computed every 2 GeV on the interval $m_{\gamma\gamma} \in [500, 5000]$ GeV and steps of 0.005 on the interval $k/\overline{M}_{\text{Pl}} \in [0.01, 0.1]$. Figure 11.8 shows the local p-value as a function of $m_X = m_G$ for various couplings $k/\overline{M}_{\text{Pl}}$.

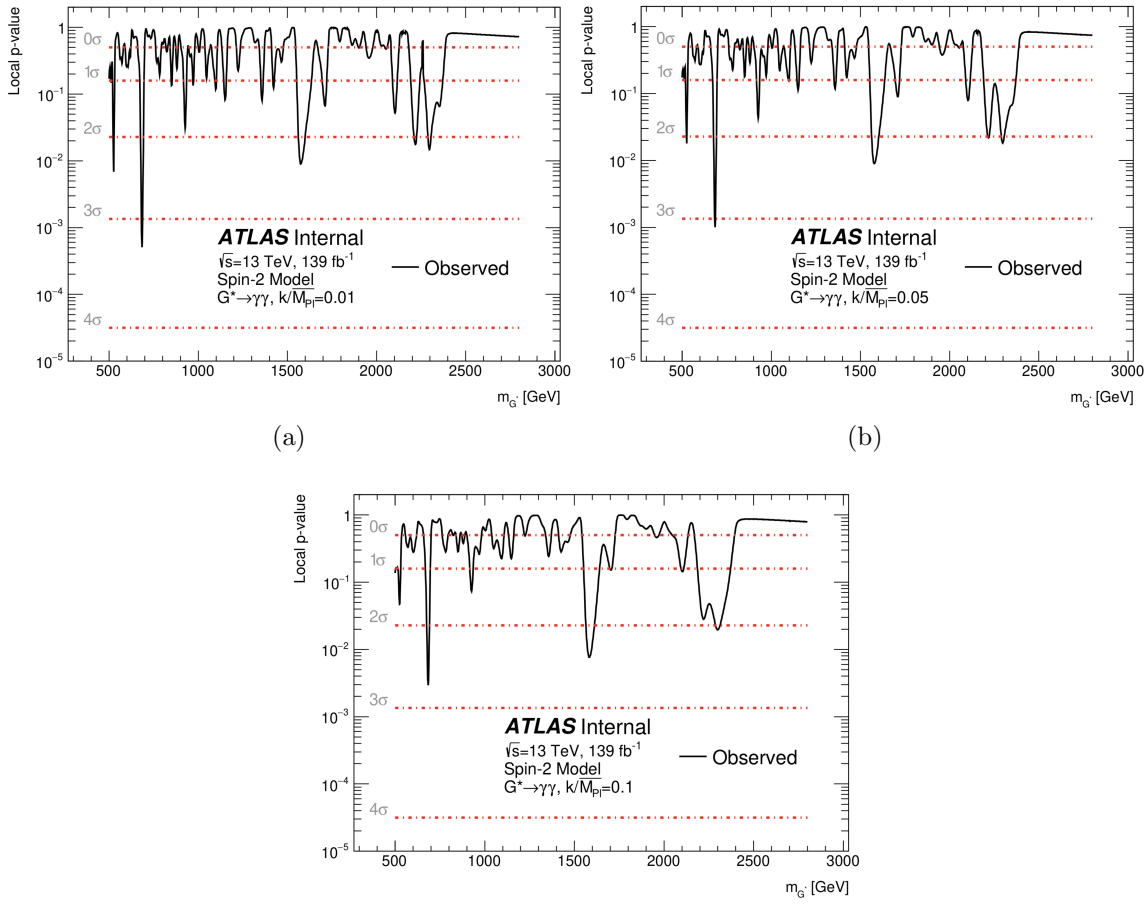


Figure 11.8: The local p-value as a function of m_G for the graviton search for various couplings: (a)$k/\overline{M}_{\text{Pl}} = 0.01$, (b) $k/\overline{M}_{\text{Pl}} = 0.05$, (c) $k/\overline{M}_{\text{Pl}} = 0.10$.

Figure 11.9 shows a two-dimensional plot of local significance as a function of width Γ_X/m_G and mass m_G. The most signficant excess was observed for the $m_G =$ 684 GeV mass hypothesis and coupling $k/\overline{M}_{\text{Pl}}$ =0.01, corresponding to a 3.29σ local significance and a $(1.36 \pm 0.06)\,\sigma$ glocal significance.

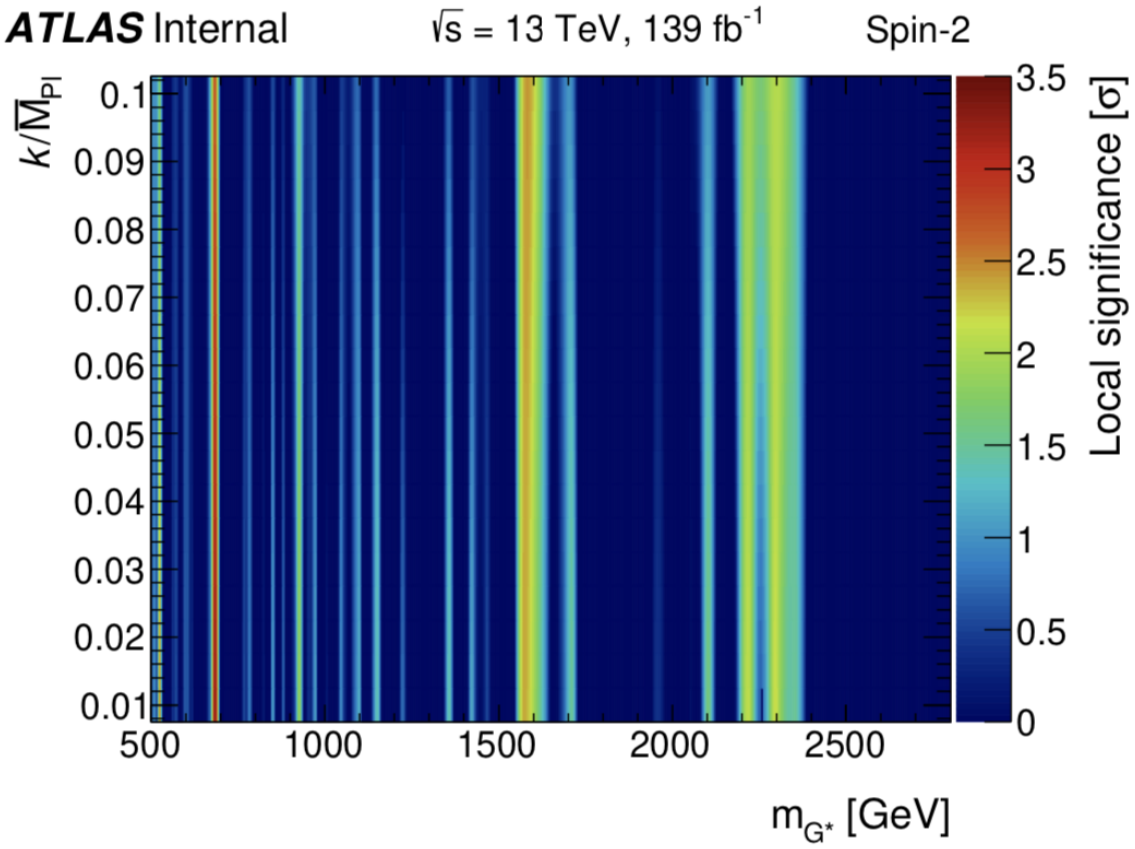


Figure 11.9: The local significance Z_0^{local} as a function of coupling $k/\overline{M}_{\text{Pl}}$ and invariant mass m_G.

Figure 11.10 shows several plots of the 95% CL on the parameter σ_{fid} times the branching ratio BR as a function of m_G for and several couplings $k/\overline{M}_{\text{Pl}}$, while the expected and observed limits as a function of coupling and invariant mass are shown in Figure 11.11.

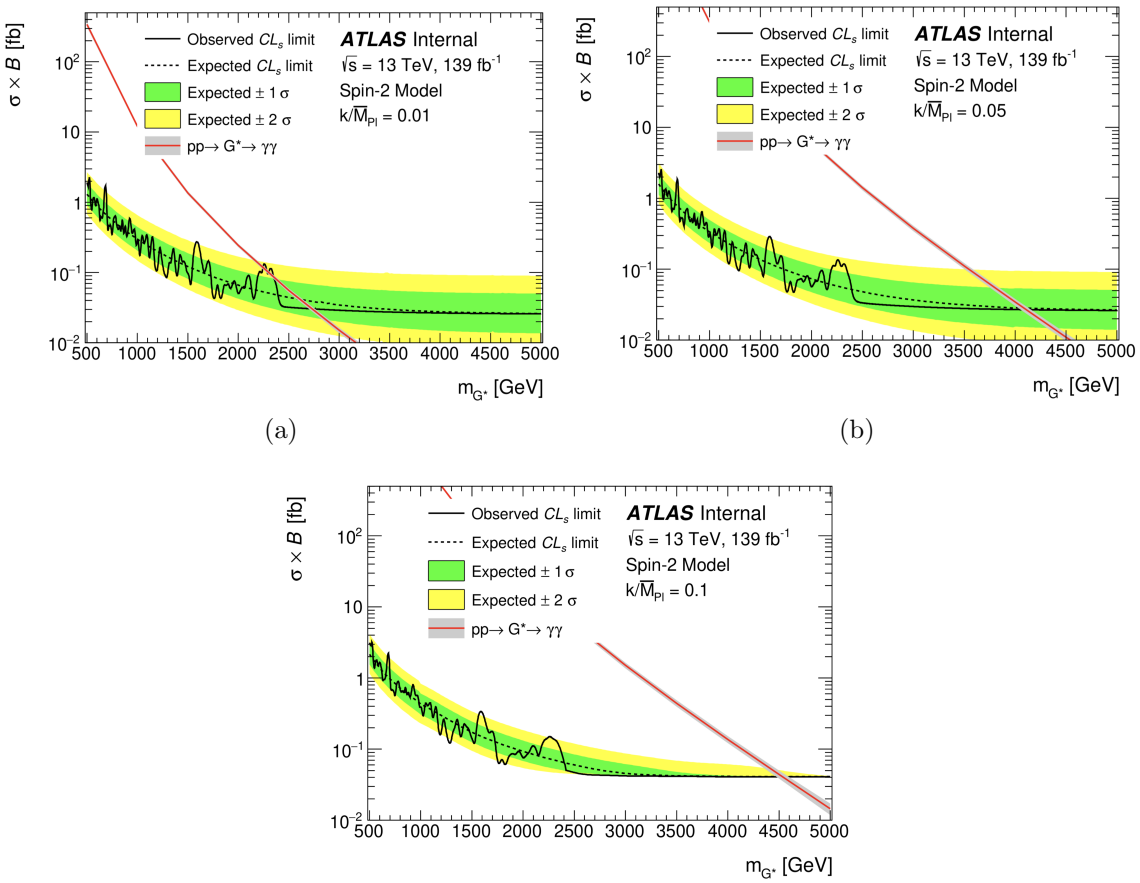
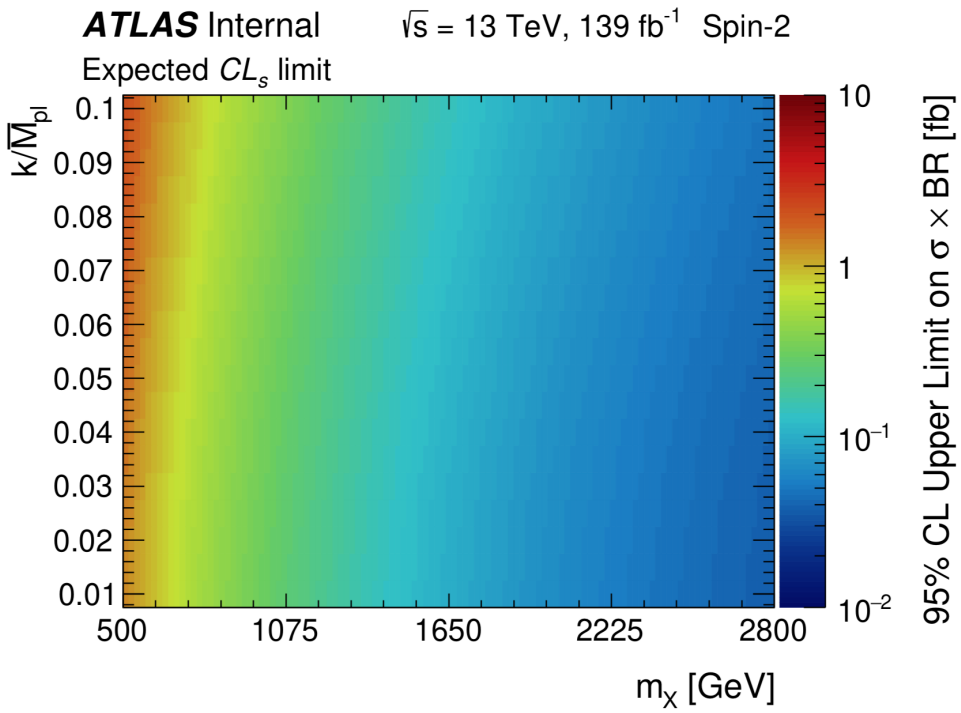
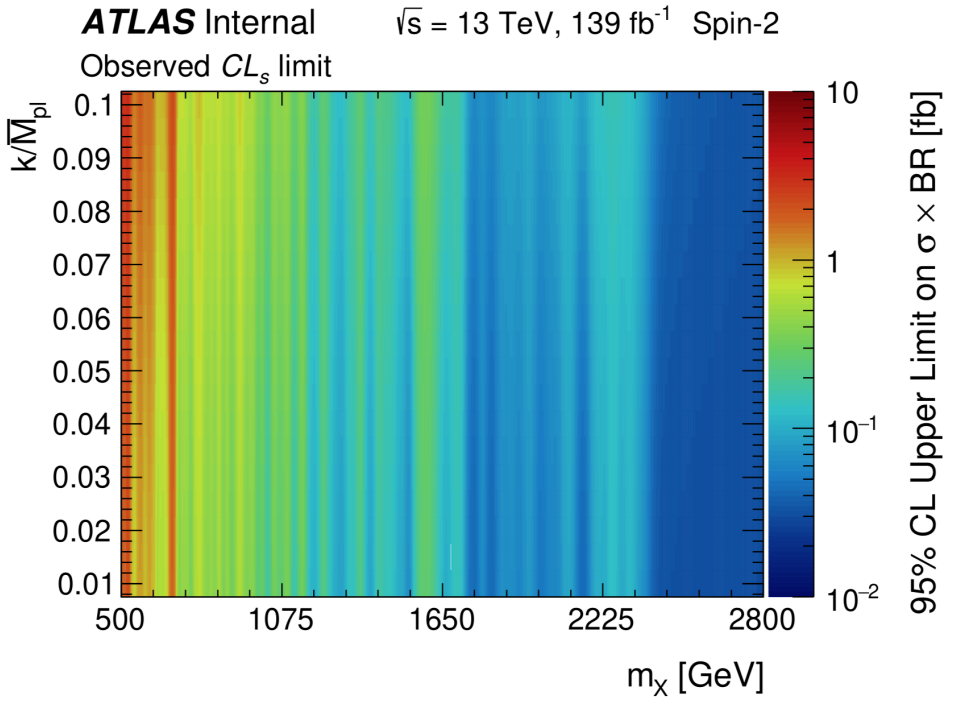


Figure 11.10: Expected and observed limit on the fiducial production cross section times branching ration $\sigma_{\text{fid}} \times \text{BR}$ as a function of the resonance mass m_G for various couplings: (a) $k/\overline{M}_{\text{Pl}} = 0.01$, (b) $k/\overline{M}_{\text{Pl}} = 0.05$, (c) $k/\overline{M}_{\text{Pl}} = 0.10$.



(a)



(b)

Figure 11.11: Expected and observed limit on the fiducial production cross section times branching ration $\sigma_{\text{fid}} \times \text{BR}$ as a function of the resonance mass m_G and coupling $k/\overline{M}_{\text{Pl}}$.

11.2.3 Summary of Observed Upper Limits

Table 11.4 summarizes the results of this analysis, showing the upper limit on the fiducial cross section times branching ratio for both the scalar and graviton search for various widths and couplings.

Width/Coupling	$\sigma_{\text{fid}} \times \text{Br}\,(X/G^* \to \gamma\gamma)$	
	Spin-0	
	$m_X = 400$ GeV	$m_X = 2800$ GeV
NWA	1.1 fb	0.03 fb
$\Gamma_X/m_X = 2\%$	2.5 fb	0.03 fb
$\Gamma_X/m_X = 6\%$	4.4 fb	0.03 fb
$\Gamma_X/m_X = 10\%$	8.3 fb	0.04 fb
	Spin-2	
	$m_{G^*} = 500$ GeV	$m_{G^*} = 5000$ GeV
$k/\overline{M}_{\text{Pl}} = 0.01$	1.9 fb	0.04 fb
$k/\overline{M}_{\text{Pl}} = 0.05$	2.3 fb	0.04 fb
$k/\overline{M}_{\text{Pl}} = 0.10$	3.2 fb	0.04 fb

Table 11.4: Summary of the limit on the fiducial cross section times branching ratio.

CHAPTER XII

Conclusion

The analysis presented in this thesis comprises two high mass diphoton final state resonance searches. The searches are conducting using the 139 fb^{-1} of Run 2 data collected with the ATLAS experiment at the LHC from 2015–2018. The spin-0 search is optimized for Higgs-like scalars with masses above 200 GeV, while the spin-2 search is optimized for resonances predicted by the Randall-Sundrum (RS) graviton model with masses above 500 GeV.

As no globally significant excess was observed, the data collected are consistent with the background expectation of the Standard Model (SM). The observed 95% CL upper limit on the fiducial cross section times branching ratio for the scalar narrow width approximation (NWA) ranges from about 12.5 fb at 160 GeV to 0.03 fb at 2800 GeV, while the observed 95% CL upper limit for the graviton with coupling $k/\overline{M}_{\text{Pl}}$ =0.1 ranges from about 3.2 fb at 500 GeV to 0.04 fb at 5000 GeV. The RS1 graviton with $k/\overline{M}_{\text{Pl}} = 0.1$ is excluded for masses below 4500 GeV, with $k/\overline{M}_{\text{Pl}}$ =0.05 is excluded for masses below 4200 GeV, and with $k/\overline{M}_{\text{Pl}}$ =0.01 is excluded for masses below 2200 GeV.

APPENDICES

APPENDIX A

Completeness of the Exponentials

Let $f(z)$ be a real valued function defined on the interval $[0, \infty)$ with the property

$$\lim_{x \to \infty} f(z) = 0 \tag{A.1}$$

and let $z = T(y)$ be the transformation $z = -\log y$. $T(y)$ bijectively maps the exponentials $\sqrt{2}F_n(z) = e^{-nz}$ on $[0, \infty)$ to the polynomials $F_n^\star(z)$ on $(0, 1]$

$$F_n^\star(y) = \sqrt{2}y^n \tag{A.2}$$

and maps the inner product

$$\langle f^\star, g^\star \rangle = \int_0^\infty \frac{dy}{y}\ f^\star(y) g^\star(y). \tag{A.3}$$

By the completeness of the polynomials, $f^\star(y) = f(-\log y)$ can be represented

$$f^\star(y) = \sum_{n=0}^{\infty} a_n^\star y^n. \tag{A.4}$$

However, since $\lim_{y \to 0} f^\star(y) = 0$ the constant term a_0 must vanish. Therefore $f(z)$ can

be written

$$f(z) = \sum_{n=1}^{\infty} a_n F_n(z) \tag{A.5}$$

where $a_n = a_n^{\star}/\sqrt{2}$ (*Edgar et al.*, 2018).

APPENDIX B

General Hyperparameter Transformation Matrix

$$f(x) = \tilde{\boldsymbol{f}}^{\star\, n} E_n(z^\star) \tag{B.1}$$

$$\tilde{\boldsymbol{f}}_n = \hat{\boldsymbol{\mathcal{M}}}_{nm} \tilde{\boldsymbol{f}}^{\star\, m} \tag{B.2}$$

$$\hat{\boldsymbol{\mathcal{M}}}_{nm} = \int_0^\infty dz\ E_n'(z) E_i(z) \frac{dz}{d\beta} \tag{B.3}$$

$$\hat{\boldsymbol{\mathcal{M}}} = \exp\left[\hat{\boldsymbol{\mathcal{T}}}\right] \tag{B.4}$$

$$\hat{\boldsymbol{\mathcal{T}}} = \int_0^\infty dz\ E_n'(z) E_m(z) \frac{dz}{d\beta} \tag{B.5}$$

APPENDIX C

Hyperparameter Transformation Matrix for the Power-Law Transformation

$$z = \left(\frac{x - x_0}{\lambda} \right)^{\alpha} \tag{C.1}$$

$\alpha = \alpha(\beta)$ $\lambda = \lambda(\beta)$

$$\frac{dz}{d\beta} = \frac{1}{\alpha}\frac{d\alpha}{d\beta} z \log z - \frac{\alpha}{\lambda}\frac{d\lambda}{d\beta} z \tag{C.2}$$

$$\frac{1}{\alpha}\frac{d\alpha}{d\beta} = c \tag{C.3}$$

$$-\frac{\alpha}{\lambda}\frac{d\lambda}{d\beta} = s \tag{C.4}$$

$\alpha = \alpha^{\star} e^{\beta c}$

$$\frac{d}{d\beta}\log\lambda = -\frac{s}{\alpha^\star}e^{-\beta c} \tag{C.5}$$

$$\log\lambda + C = \frac{s}{\alpha^\star c}e^{-\beta c} \tag{C.6}$$

$$\log\frac{\lambda}{\lambda^\star} = \frac{s}{\alpha^\star c}\left(e^{-\beta c} - 1\right) \tag{C.7}$$

$$c = \log\frac{\alpha}{\alpha^\star} \tag{C.8}$$

$$s = -\frac{\alpha c}{e^c - 1}\log\frac{\lambda}{\lambda^\star} \tag{C.9}$$

$$\hat{\boldsymbol{\tau}}_{nm} = \int_0^\infty dz\ E_n'(z)E_m(z)\left(cz\log z + sz\right) \tag{C.10}$$

$$= -\sum_{i=1}^\infty\sum_{j=1}^\infty d_{ni}d_{mj}i\int_0^\infty dz\ e^{-(i+j)z}\left(cz\log z + sz\right) \tag{C.11}$$

$$\int_0^\infty dz\ e^{-nz}z = \frac{1}{n^2} \tag{C.12}$$

$$\int_0^\infty dz\ e^{-nz}z\log z = \frac{1-\gamma-\log n}{n^2} \tag{C.13}$$

APPENDIX D

Optimal Signal Estimators

$$F(z) = c^{(m)} f_{(m)}(z) \tag{D.1}$$

$$\langle \omega_{(n)}, f_{(m)} \rangle = \delta_{nm} \tag{D.2}$$

$$\sigma_{(nm)} = \int_0^\infty dz \; F(z)\omega_{(n)}(z)\omega_{(m)}(z) - \left[\int_0^\infty dz \; F(z)\omega_{n)}(z)\right]\left[\int_0^\infty dz \; F(z)\omega_{(m)}\right] \tag{D.3}$$

$$= \tilde{\boldsymbol{\omega}}^i_{(n)} \hat{\boldsymbol{\Sigma}}_{ij} \tilde{\boldsymbol{\omega}}^j_{(m)} \tag{D.4}$$

$$\mathcal{L} = \log\det(2\pi e \sigma_{(nm)}) - \eta^{(nm)} \left(\tilde{\boldsymbol{\omega}}_{(n)i} \tilde{\boldsymbol{f}}^i_{(m)} - \delta_{nm}\right) \tag{D.5}$$

BIBLIOGRAPHY

BIBLIOGRAPHY

Aaboud, M., et al. (2017), Search for new phenomena in high-mass diphoton final states using 37 fb[1] of protonproton collisions collected at $\sqrt{s} = 13$ TeV with the ATLAS detector, *Physics Letters B*, *775*, 105125, doi:10.1016/j.physletb.2017.10.039.

Aaboud, M., et al. (2019), Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016, *The European Physical Journal C*, *79*(3), doi:10.1140/epjc/s10052-019-6650-6.

Aad, G., et al. (2010), "the atlas simulation infrastructure", *The European Physical Journal C*, *70*(3), 823874, doi:10.1140/epjc/s10052-010-1429-9.

Aad, G., et al. (2013), Search for extra dimensions in diphoton events from proton-proton collisions at $\sqrt{s} = 7$ TeV in the ATLAS detector at the LHC, *New Journal of Physics*, *15*(4), 043,007, doi:10.1088/1367-2630/15/4/043007.

Aad, G., et al. (2019), Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data, *Journal of Instrumentation*, *14*(12), doi:10.1088/1748-0221/14/12/p12006.

Aaij, R., et al. (2015), Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ decays, *Physical Review Letters*, *115*(7), doi:10.1103/physrevlett.115.072001.

Aaltonen, T., et al. (2011), Search for New Dielectron Resonances and Randall-Sundrum Gravitons at the Collider Detector at Fermilab, *Physical Review Letters*, *107*(5), doi:10.1103/physrevlett.107.051801.

Abazov, V. M., et al. (2010), Search for Randall-Sundrum Gravitons in the Dielectron and Diphoton Final States with 5.4 fb[1] of Data fromppCollisions at $\sqrt{s} = 1.96$ TeV, *Physical Review Letters*, *104*(24), doi:10.1103/physrevlett.104.241802.

Abdelouahab, A., et al. (2008), Combined performance tests before installation of the ATLAS semiconductor and transition radiation tracking detectors, *JINST*, *3*, doi:10.1088/1748-0221/3/08/P08003.

Amidei, D. e. a. (2020), High-mass diphoton resonance search using 139 fb^{-1} of 13 TeV pp collision data with the ATLAS detector, *Tech. rep.*, CERN.

ATLAS, C., et al. (2010), The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim, *Tech. Rep. ATL-PHYS-PUB-2010-013*, CERN, Geneva.

ATLAS Collaboration (2011), Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data, *Eur. Phys. J. C*, *72*(arXiv:1110.3174. CERN-PH-EP-2011-117), doi:10.1140/epjc/s10052-012-1909-1.

ATLAS e/γ (2020a), Electron Photon Four Momentum Correction, `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ElectronPhotonFourMomentumCorrection`.

ATLAS e/γ (2020b), Photon efficiencies for run 2, `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PhotonEfficiencyRun2`.

Branco, G., P. Ferreira, L. Lavoura, M. Rebelo, M. Sher, and J. P. Silva (2012), Theory and phenomenology of two-Higgs-doublet models, *Physics Reports*, *516*(1-2), 1102, doi:10.1016/j.physrep.2012.02.002.

Cacciari, M., G. P. Salam, and G. Soyez (2008), The catchment area of jets, *Journal of High Energy Physics*, *2008*(04), 005005, doi:10.1088/1126-6708/2008/04/005.

Camarri, P., et al. (1998), RPCs for the ATLAS Level-1 muon trigger:Test beam results, *Tech. Rep. ATL-MUON-98-234. ATL-COM-MUON-98-027*, CERN, Geneva.

Carminati, L., et al. (2012), Measurement of the isolated di-photon cross section inv 4.9 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, *Tech. Rep. ATL-COM-PHYS-2012-592*, CERN, Geneva.

Chatrchyan, S., et al. (2012), Search for Signatures of Extra Dimensions in the Diphoton Mass Spectrum at the Large Hadron Collider, *Physical Review Letters*, *108*(11), doi:10.1103/physrevlett.108.111801.

Cowan, G., K. Cranmer, E. Gross, and O. Vitells (2011), Asymptotic formulae for likelihood-based tests of new physics, *The European Physical Journal C*, *71*(2), doi:10.1140/epjc/s10052-011-1554-0.

De Vivie De Regie, J.-B., et al. (2017), ATLAS electron, photon and muon isolation in Run 2, *Tech. Rep. ATL-COM-PHYS-2017-290*, CERN, Geneva, this note contains the Moriond 2017 recommendations. It will be updated when new recommendations become available.

Edgar, R., D. Amidei, C. Grud, and K. Sekhon (2018), Functional Decomposition: A new method for search and limit setting.

et. al., G. (2019), Study of Performance of Functional Decomposition in the High Mass Diphoton Resonance Search, *Tech. rep.*, CERN.

Gabella, M. (2006), The randall-sundrum model, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.562.4965.

Glashow, S. L. (1959), The renormalizability of vector meson interactions, *Nuclear Physics*, *10*, 107–117.

Höche, S., F. Krauss, M. Schnherr, and F. Siegert (2013), "qcd matrix elements + parton showers. the nlo case", *Journal of High Energy Physics*, *2013*(4), doi: 10.1007/jhep04(2013)027.

Kullback, S., and R. A. Leibler (1951), On Information and Sufficiency, *The Annals of Mathematical Statistics*, *22*(1), 79–86, doi:10.1214/aoms/1177729694.

Lampl, W., et al. (2008), Calorimeter Clustering Algorithms: Description and Performance, *Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003*, CERN, Geneva.

Lenzi, B., and D. Delgove (2015), Selection of the diphoton production vertex using multivariate techniques for $H \to \gamma\gamma$ and other analyses, *Tech. Rep. ATL-COM-PHYS-2015-1321*, CERN, Geneva.

Pequenao, J. (2008), Computer generated image of the whole ATLAS detector, https://cds.cern.ch/record/1095924.

Petit, E. (2017), Combination of photon identification measurements with 13 TeV data recorded in 2015 and 2016., *Tech. Rep. ATL-COM-PHYS-2017-070*, CERN, Geneva.

Randall, L., and R. Sundrum (1999), Large mass hierarchy from a small extra dimension, *Physical Review Letters*, *83*(17), 33703373, doi:10.1103/physrevlett.83.3370.

Sakurai, J., and J. Napolitano (1964), *Modern Quantum Mechanics*, second ed., Addison Wesley.

Salam, A., and et al (1964), Electromagnetic and weak interactions, *Physics Letters*, *13*(2), 168–171.

Schumann, S., and F. Krauss (2008), A parton shower algorithm based on catani-seymour dipole factorisation, *Journal of High Energy Physics*, *2008*(03), 038038, doi:10.1088/1126-6708/2008/03/038.

Sirunyan, A., et al. (2018), Search for physics beyond the standard model in high-mass diphoton events from proton-proton collisions at $\sqrt{s} = 13$ TeV, *Physical Review D*, *98*(9), doi:10.1103/physrevd.98.092001.

Soch, J., and C. Allefeld (2016), Kullback-Leibler Divergence for the Normal-Gamma Distribution.

Sotto-Maior Peralva, B. (2013), Calibration and Performance of the ATLAS Tile Calorimeter, *Tech. Rep. arXiv:1305.0550*, ATLAS.

Tanabashi, M. et al. (2018), Review of Particle Physics, *Phys. Rev. D*, *98*, doi: 10.1103/PhysRevD.98.030001, 030001.

Tang, Y. (2012), Implications of LHC searches for massive graviton, *Journal of High Energy Physics*, *2012*(8), doi:10.1007/jhep08(2012)078.

Verkerke, W., and D. Kirkby (2003), The RooFit toolkit for data modeling.

Weinberg, S. (1967), A Model of Leptons, *Phys. Rev. Lett.*, *19*, 1264–1266.