UNIVERSITY OF CALIFORNIA
SANTA CRUZ

A SEARCH FOR HIGGS BOSON PRODUCTION VIA VECTOR BOSON FUSION IN ASSOCIATION WITH A PHOTON IN THE $H \to b\bar{b}$ CHANNEL WITH THE ATLAS DETECTOR

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PHYSICS

by

Peyton Wells Rose

December 2016

The Dissertation of Peyton Wells Rose
is approved:

Professor Jason Nielsen, Chair

Professor Abraham Seiden

Professor Bruce Schumm

Dean Tyrus Miller
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

A search for Higgs boson production via vector boson fusion in association with a photon in the $H \to b\bar{b}$ channel with the ATLAS detector

by

Peyton Wells Rose

A search for the Standard Model Higgs boson produced via vector boson fusion in association with a photon and decaying to a bottom quark-antiquark pair has been conducted using $12.6\,\text{fb}^{-1}$ of proton-proton collision data collected by the ATLAS detector at a center-of-mass energy $\sqrt{s} = 13\,\text{TeV}$ at the LHC in 2015 and 2016. The search benefits from a large reduction of the non-resonant multijet background relative to similar searches that do not explicitly require a photon and from efficient triggering aided by the presence of the photon. Multivariate techniques trained using simulated Monte Carlo samples are used to enhance the sensitivity of the analysis by constructing regions of phase space with higher expected signal fractions relative to the background. Data-driven techniques are used in those regions to provide a reliable estimate of the di-b-jet invariant mass spectrum of the non-resonant background. The final statistical analysis of the data relies on a profile likelihood fit to the di-b-jet invariant mass distribution, searching for a signal bump in an otherwise smoothly falling distribution. There is no observed excess above the background-only expectation, and the observed (expected) 95% confidence level upper limit on the production cross section times branching ratio for a Higgs boson mass of 125 GeV is 4.0 ($6.0^{+2.3}_{-1.7}$) times the Standard Model expecta-

tion. The observed signal strength is $-3.9^{+2.8}_{-2.7}$ times the Standard Model value. This analysis is also sensitive to $Z + \gamma$ production: the observed (expected) upper limit is 2.0 ($1.8^{+0.7}_{-0.5}$) times the Standard Model expectation with an observed signal strength of 0.3 ± 0.8.

To mom and dad,

for your never-ending support

for my personal and professional endeavors in life.

Acknowledgments

Thank you Jason Nielsen for leading by example with your enthusiasm for physics and dedication to your work. Your support and guidance helped make Santa Cruz a fantastic place to be a graduate student. Thank you to the SCIPP ATLAS group and the broader SCIPP community for being able and willing to answer questions regarding fundamental physics, computing subtleties, and ATLAS analyses. Thank you to the other members of the VBF+γ team–Liaoshan, Song-Ming, Zhijun, and Jake–for your round-the-clock efforts on the analysis during the final push for ICHEP. Lastly, thank you to my office mates, Sheena and Alexander, for being there through the triumphs and tribulations that inevitably come when doing scientific research.

Chapter 1

Introduction

The birth of particle physics began with the Rutherford gold foil experiment, conducted by Geiger and Marsden, which discovered a dense positively charged nucleus at the center of atoms [1]. Since then, experiments have grown in size and complexity as we continue to probe higher energies and smaller distances. Results from these experiments have been interpreted and used to build one of the most predictive and successful theories across all fields of science—the Standard Model (SM) of particle physics. At the current state-of-the-art facility, the Large Hadron Collider (LHC), particle interaction cross section predictions of the SM spanning 10 orders of magnitude agree well with experimental results [2]. Despite this success, the SM is likely not the end of the story. Interpretations of certain astronomical observations require the existence of mysterious Dark Matter, and there are no candidate particles within the SM with the properties necessary to account for these phenomena. Furthermore, some parameters of the SM require an unnatural degree of fine tuning unless they can be protected by

an additional symmetry of nature. These shortcomings are viewed by many particle physicists as hints that new physical phenomena may be within our reach at the LHC. Evidence for new phenomena may be found through the observation of a new resonance or a disagreement in particle production and decay rates relative to SM predictions.

A promising candidate to guide our search for new physical phenomena is the neutral scalar boson discovered at the LHC in 2012 in the search for the SM Higgs boson. The combined global signal strength parameter, which is the measured signal yield across all production and decay modes relative to the prediction for the SM Higgs boson, has been measured by the ATLAS and CMS experiments to be $\mu = 1.09^{+0.11}_{-0.10}$ [3]. This agrees well with the SM prediction for the Higgs boson, though measurements in individual production and decay modes are not measured with the same precision. In particular, individual experiments have not observed the decay of this Higgs boson to fermions. The $H \to \tau^+\tau^-$ decay has been established through a combination of results from ATLAS and CMS [3], but the $H \to b\bar{b}$ decay remains elusive at the LHC. Many theories of physics beyond the SM modify the coupling between the Higgs boson and bottom quarks, so it is important to fully explore this decay mode at the LHC.

Both ATLAS and CMS have published analyses targeting the $H \to b\bar{b}$ decay. The production modes explored in these analyses are associated VH production, vector boson fusion (VBF), and associated $t\bar{t}$ production. The combined ATLAS+CMS signal strength measurement for the $H \to b\bar{b}$ decay is $\mu^{b\bar{b}} = 0.70^{+0.29}_{-0.28}$ [3]. A variation of the VBF production mechanism that has previously not been explored at the LHC is VBF production in association with a photon (VBF+γ). The cross section for this

production mode scales by roughly the fine structure constant relative to the inclusive VBF cross section. Despite the decrease in the expected number of signal events, there are advantages for considering this channel including a dramatic decrease in the number of background events relative to the signal and efficient triggering at the LHC [4]. Furthermore, this analysis can be made orthogonal to the inclusive VBF analysis using a photon veto in the latter with little to no effect on the sensitivity of that analysis.

This dissertation describes the first search for VBF+γ Higgs production at the LHC and targets the $H \to b\bar{b}$ decay. It uses $12.6\,\text{fb}^{-1}$ of proton-proton collision data collected at a center-of-mass energy $\sqrt{s} = 13\,\text{TeV}$ using the ATLAS detector at the LHC in 2015+2016. Part I describes the Standard Model of particle physics, including the Higgs mechanism and details about Higgs boson studies at the LHC. Part II describes both the LHC and the ATLAS detector. Lastly, the data analysis and Higgs boson search results are detailed in Part III.

Part I

Theoretical Motivation

Chapter 2

The Standard Model

The Standard Model of particle physics is a theory describing three of the four fundamental forces in nature and the particles these forces act upon. It emerged in the second half of the 20^{th} century, and describes the strong, weak, and electromagnetic interactions. The other fundamental force, gravity, is not described by this theory. The SM is a $SU(3)_{\text{C}} \times SU(2)_{\text{L}} \times U(1)_Y$ local gauge theory, meaning that its Lagrangian (and therefore the mathematical predictions of the theory) is invariant under local transformations belonging to each respective group. In fact, the gauge principle states that we can deduce the form of an interaction by identifying the relevant symmetry group and requiring the theory to be invariant when acting on the physical states with a generic member of that group [5].

Since its formulation, the SM has had much success, including correct predictions for the existence of many fundamental particles prior to their experimental discovery. These include the $W^{\pm}$ and Z bosons [6], the top quark, and most recently

the Higgs boson [7], which is discussed in detail in Section 2.4. At the Large Hadron Collider, measured particle interaction cross sections spanning 10 orders of magnitude agree well with SM predictions [2]! Despite its widespread success, the SM has several shortcomings. Among these, it does not contain a dark matter candidate, and it requires fine tuning of its parameters. Several theories beyond the SM (BSM) have been proposed to address these limitations, and evidence for these theories could be found through precision measurements that conflict with SM predictions.

2.1 Particles in the Standard Model

The known fundamental particles in nature can be divided into two broad categories: fermions and bosons. Fermions are particles which carry half-integer spin, and are commonly thought of as the matter content in the SM, while bosons have integer spin and are the force mediators in the SM [8].

2.1.1 Bosons

The force carriers in the SM are called gauge bosons, and they all have spin-1 making them vector bosons[1]. They are the photon (γ), corresponding to the electromagnetic interaction; the $W^{\pm}$ bosons, which mediate the charged weak interaction; the Z boson, which mediates the neutral weak interaction; and eight gluons mediating the strong interaction. Both the photon and the gluons are massless; however, due to color

[1] In the rest of this dissertation, the term "vector boson" is used to refer specifically to the massive vector bosons ($W^{\pm}$ and Z).

confinement in the strong interaction, only electromagnetism is a long-range interaction. The $SU(2)$ symmetry of the SM prohibits a mass term in the Lagrangian for the $W^{\pm}$ and Z bosons, though experimentally these particles are observed to have masses of 80 GeV and 91 GeV, respectively. This inconsistency is resolved via the Higgs mechanism, which embeds a complex scalar doublet field in $SU(2)$. This field has a non-zero vacuum expectation value (vev), and the $W^{\pm}$ and Z bosons acquire mass though their interactions with this field. This theory predicts the existence of a massive spin-0 particle called the Higgs boson. The Higgs mechanism is discussed further in Section 2.4.

2.1.2 Fermions

The fermions can be further subdivided into quarks and leptons. Quarks have color charge and thus participate in strong interactions (discussed in Section 2.3), whereas leptons do not. There are three generations of both quarks and leptons, and each generation contains two of each of these types of particles. The first generation of fermions contains up (u) and down (d) quarks which are found in atomic nuclei, as well as electrons (e) which can be found in a cloud surrounding atomic nuclei. The electron neutrino (ν_e) is the other lepton in the first generation. The second and third generation fermions can be thought of as heavier copies of their first generation partners. The exception is neutrinos, where the mass eigenstates are a superposition of the flavor eigenstates, and we do not yet know the mass ordering of the mass eigenstates. The charm (c) and top (t) quarks are the second and third generation copies of the u quark. All three of these particles carry $+\frac{2}{3}$ electric charge. The strange (s) and

bottom (b) quarks are the corresponding partners of the d quark, and these particles all carry $-\frac{1}{3}$ electric charge. Muons (μ) and taus (τ) are the heavier partners of the e, which all carry -1 electric charge. Each of these leptons also has a partner neutrino: ν_μ and ν_τ. The neutrinos are not electrically charged, and thus do not participate in the electromagnetic interaction. Each fermion also has a partner antiparticle, which has the opposite sign electric charge, but is otherwise identical.

Elementary particles have another property, called chirality. Formally, the chirality of a particle is determined by its transformation properties under the Poincaré group. Chirality is closely related to a physical property, helicity, which is the projection of a particle's spin onto its direction of motion. In the limit of massless particles, which move at the speed of light, chirality and helicity are the same. In the SM, the left-chiral quarks and leptons in a given generation are grouped into $SU(2)$ doublets, whereas the right-chiral particles belong to $SU(2)$ singlets. This leads to a rich phenomenology in which the weak interaction couples preferentially to the left-chiral states. This interaction is discussed in more detail in Section 2.2. Fermion mass terms mix the left- and right-chiral particle states, which have different transformation properties under $SU(2)$, thus spoiling the $SU(2)$ gauge symmetry. This would lead us to conclude that the fermions are massless, which is in conflict with experimental observations that they are massive. The Higgs field can be used to resolve this inconsistency through a different mechanism than is used to give mass to the gauge bosons.

2.2 Electroweak theory

In the SM, the electromagnetic and weak interactions are unified into a single electroweak interaction, which is described by the $SU(2)_{\mathrm{L}} \times U(1)_Y$ gauge group [9]. The subscript L is used to denote that the $SU(2)$ transformation acts only on the left chiral particle states. The most general $SU(2)_{\mathrm{L}} \times U(1)_Y$ transformation acting on a doublet state, $\boldsymbol{L}$, is shown in Equation 2.1, where χ and $\boldsymbol{\alpha}$ are local phase transformation parameters that depend on the space-time point x; $\boldsymbol{\tau}$ are the generators of $SU(2)$; and g, g', and Y_L describe the strength of the interactions.

$$\boldsymbol{L} \rightarrow \boldsymbol{L'} = \exp\left(ig'\chi(x)\frac{Y_L}{2} + ig\boldsymbol{\alpha}(x)\cdot\frac{\boldsymbol{\tau}}{2}\right)\boldsymbol{L} \tag{2.1}$$

Similarly, a general U(1) transformation acting on a singlet state, φ_R, is shown in Equation 2.2.

$$\varphi_R^l \rightarrow (\varphi_R^l)' = \exp\left(ig'\chi(x)\frac{Y_R^l}{2}\right)\varphi_R^l \tag{2.2}$$

Due to the local dependence of these gauge transformations on the space-time point, x, the kinetic term in the electroweak Lagrangian (involving a space-time derivative) will acquire additional terms under these transformations, thus spoiling the gauge invariance. This forces us to introduce the covariant derivative, given in Equations 2.3 and 2.4 for $SU(2)_{\mathrm{L}} \times U(1)_Y$ and $U(1)_Y$ transformations, respectively.

$$\mathcal{D}_\mu = \partial_\mu - \frac{1}{2}ig'B_\mu Y_L - \frac{1}{2}ig\boldsymbol{W}_\mu\cdot\boldsymbol{\tau} \tag{2.3}$$

$$\mathcal{D}_\mu = \partial_\mu - \frac{1}{2}ig'B_\mu Y_R^l \tag{2.4}$$

In these equations, B_μ and $\boldsymbol{W}_\mu$ are new gauge fields which transform under $U(1)_Y$ and $SU(2)_\mathrm{L}$, respectively. These fields are chosen to transform in such a way that the combined effect of local $SU(2)_\mathrm{L} \times U(1)_Y$ transformations on the electroweak Lagrangian leaves the theory invariant. The transformation properties of the $\boldsymbol{W}_\mu$ and B_μ fields are given in Equations 2.5 and 2.6, respectively.

$$\boldsymbol{W}_\mu(x) \rightarrow \boldsymbol{W}'_\mu(x) = \boldsymbol{W}_\mu + \partial_\mu \boldsymbol{\alpha}(x) + g\boldsymbol{W}_\mu(x) \times \boldsymbol{\alpha}(x) \tag{2.5}$$

$$B_\mu \rightarrow B'_\mu = B_\mu + \frac{1}{g'}\partial_\mu \chi(x) \tag{2.6}$$

Note that in these equations, $\chi(x)$ and $\boldsymbol{\alpha}(x)$ are the same local phases as in Equations 2.1 and 2.2. Putting it all together, the kinetic terms of the electroweak Lagrangian are given in Equation 2.7.

$$\mathcal{L} = i\bar{\boldsymbol{L}}\gamma^\mu \left(\partial_\mu - \frac{1}{2}ig'B_\mu Y_L - \frac{1}{2}ig\boldsymbol{W}_\mu \cdot \boldsymbol{\tau}\right)\boldsymbol{L} + i\bar{\varphi}^l_R\gamma^\mu \left(\partial_\mu - \frac{1}{2}ig'B_\mu Y^l_R\right)\varphi^l_R \tag{2.7}$$

In this equation, we have omitted terms for the gauge fields. These terms are given in Equation 2.8

$$\mathcal{L}_{gauge\,boson} = -\frac{1}{4}\boldsymbol{F}_{\mu\nu}\boldsymbol{F}^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + \frac{1}{2}M^2\boldsymbol{W}_\mu\boldsymbol{W}^\mu, \tag{2.8}$$

where $\boldsymbol{F}^{\mu\nu}$ and $B^{\mu\nu}$ are field tensors, defined in the following equations.

$$\boldsymbol{F}^{\mu\nu} = \partial^\mu \boldsymbol{W}^\nu - \partial^\nu \boldsymbol{W}^\mu + g\boldsymbol{W}^\mu \times \boldsymbol{W}^\nu \tag{2.9}$$

$$B^{\mu\nu} = \partial^\mu B^\nu - \partial^\nu B^\mu \tag{2.10}$$

Equation 2.8 presents a problem: the first two terms are gauge invariant on their own, whereas the last term is not. This suggests that we should set $M = 0$, which would mean the gauge bosons are massless. As mentioned in Section 2.1.1, this contradicts experimental evidence that the $W^{\pm}$ and Z bosons do, in fact, have mass. We run into a similar problem when we try to include a mass term for fermions, shown in Equation 2.11, in the electroweak Lagrangian.

$$m\bar{\varphi}\varphi = m \begin{pmatrix} \varphi_R^\dagger & \varphi_L^\dagger \end{pmatrix} \begin{pmatrix} \varphi_L \\ \\ \varphi_R \end{pmatrix} = m(\varphi_R^\dagger \varphi_L + \varphi_L^\dagger \varphi_R). \tag{2.11}$$

In Equation 2.11, the fermion fields have been expanded in their chiral basis. Using this basis, it is clear that the mass term mixes the left and right chiral fermion fields. Since these fields transform differently under the electroweak gauge symmetry, this fermion mass term does not preserve the gauge symmetry. This is another apparent shortcoming of electroweak theory since fermions are observed to be massive. The Higgs mechanism resolves this and the gauge boson mass problem and is discussed in Section 2.4.

2.3 Quantum chromodynamics

As mentioned in Section 2.1.2, quarks are differentiated from leptons by their color charge. The color analogy is used because there are three distinct color charges, and combining one of each of the different charges results in a colorless object, similar to combining red, green, and blue light to make white. The theory of colored interac-

tions is called quantum chromodynamics (QCD) and is described by the $SU(3)_{\mathrm{C}}$ gauge group [10]. Each quark is part of a color triplet which transforms under a general $SU(3)$ rotation as shown in Equation 2.12.

$$|q\rangle = \begin{pmatrix} q_r \\ q_g \\ q_b \end{pmatrix} \rightarrow |q'\rangle = \exp\left(i g_s \sum_{k=1}^{8} \eta_k(x) \frac{\lambda_k}{2} \right) |q\rangle \tag{2.12}$$

In this equation, λ_k are the generators of $SU(3)$, η_k are local phases multiplying each of the generators, and g_s is the strong coupling constant. As in electroweak theory, the requirement of invariance under local $SU(3)$ transformations forces us to introduce the covariant derivative, which is defined in Equation 2.13.

$$\mathcal{D}_\mu = \partial_\mu - i g_s G_\mu^k \frac{\lambda_k}{2} \tag{2.13}$$

In this equation, G_μ^k are fields corresponding to the eight gluons. Their transformation property under $SU(3)$ is shown in Equation 2.14, and the kinetic term of the QCD Lagrangian is given in Equation 2.15.

$$G_\mu^k \rightarrow G_\mu'^k = G_\mu^k + \partial_\mu \eta_k(x) + g_s f_{klm} \eta_l(x) G_\mu^m \tag{2.14}$$

$$\mathcal{L}_{QCD} = \bar{q}(i\gamma_\mu \mathcal{D}^\mu - m_q)q - \frac{1}{4} G_k^{\mu\nu} (G_k)_{\mu\nu} \tag{2.15}$$

The $G_k^{\mu\nu}$ in Equation 2.15 is the gluon field tensor, defined in Equation 2.16.

$$G_k^{\mu\nu} = \partial^\mu G_k^\nu - \partial^\nu G_k^\mu + g_s f_{klm} G_l^\mu G_m^\nu \tag{2.16}$$

The strong interaction is unique in that only objects that are neutral (with respect to the strong interaction) are observed as free particles in nature. This means that quarks and gluons, which have net color charges, can never be isolated. This property is referred to as color confinement, and a consequence of this is that quarks are always found in bound states, collectively referred to as hadrons. A bound state of a quark with an antiquark of the same color charge is called a meson, and a bound state of three quarks, each having a unique color charge, is called a baryon. Protons and pions are two examples of hadrons.

2.4 The Higgs mechanism

The Higgs mechanism embeds a complex scalar doublet field in the electroweak sector of the SM to resolve the gauge boson mass problem through electroweak symmetry breaking. As we will see, this field can also be used to give mass to the fermions through a separate mechanism.

2.4.1 Electroweak symmetry breaking

As a complex doublet, the Higgs field, $\mathbf{\Phi}$, has four components, as shown in Equation 2.17.

$$\mathbf{\Phi}(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1^+(x) + i\phi_2^+(x) \\ \phi_1^0(x) + i\phi_2^0(x) \end{pmatrix} \tag{2.17}$$

Three of these four degrees of freedom ultimately become the longitudinal polarizations

of the $W^{\pm}$ and Z bosons, and the other becomes the mass of a scalar particle. The form of the potential for this field is given in Equation 2.18, where μ^2 and λ are both positive numbers.

$$V(\mathbf{\Phi}^\dagger\mathbf{\Phi}) = -\mu^2\mathbf{\Phi}^\dagger\mathbf{\Phi} + \lambda(\mathbf{\Phi}^\dagger\mathbf{\Phi})^2 \tag{2.18}$$

For small displacements from $\mathbf{\Phi} = 0$ the quadratic term in Equation 2.18 dominates, so V is negative; for large displacements the quartic term dominates, and V is positive. This gives the Higgs potential a so-called "Mexican-hat" shape, as shown in Figure 2.1. The key feature of the Higgs potential is that its minimum is not at $\mathbf{\Phi} = 0$, and since

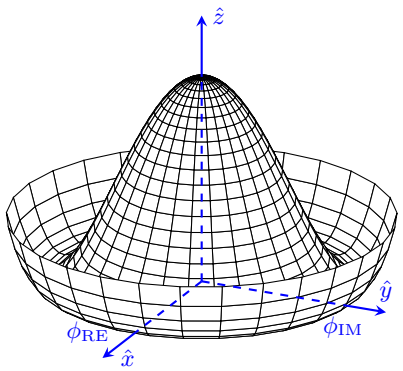


Figure 2.1: The shape of the Higgs potential.

physical states tend towards their lowest energy configuration, the Higgs field acquires a vev. Taking the derivative of V with respect to $\mathbf{\Phi}$ and setting it equal to zero, we find $\mathrm{vev}_{\mathbf{\Phi}} = \frac{1}{\sqrt{2}}v$, where $v \equiv \sqrt{\mu^2/\lambda} = 246\,\mathrm{GeV}$. This value defines a surface corresponding to an infinite number of states that minimize V. The Higgs field must choose one particular value, thus breaking the symmetry of the vacuum. By convention, we choose

a coordinate system such that

$$\langle \mathbf{\Phi} \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{2.19}$$

We can parametrize small perturbations around the minimum of the Higgs potential as

$$\mathbf{\Phi}(x) = \exp\left(i\boldsymbol{\xi}(x) \cdot \boldsymbol{\tau}\right) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \tag{2.20}$$

In Equation 2.20, $H(x)$ is an excitation in the radial direction, and the exponential term is an arbitrary phase. Including the potential, the Lagrangian for the Higgs field is given in Equation 2.21, where $\mathcal{D}_\mu$ is the same $SU(2)$ covariant derivative given in Equation 2.3.

$$\mathcal{L}_{\mathbf{\Phi}} = (\mathcal{D}_\mu \mathbf{\Phi})^\dagger (\mathcal{D}^\mu \mathbf{\Phi}) + \mu^2 \mathbf{\Phi}^\dagger \mathbf{\Phi} - \lambda (\mathbf{\Phi}^\dagger \mathbf{\Phi})^2 \tag{2.21}$$

By evaluating the kinetic term of the Higgs Lagrangian at the vev, we naturally recover mass terms for the gauge bosons, as shown in Equation 2.22.

$$\begin{aligned} &[(\mathcal{D}_\mu \mathbf{\Phi})^\dagger (\mathcal{D}^\mu \mathbf{\Phi})_{|\mathbf{\Phi}=\langle \mathbf{\Phi} \rangle} = \\ &\qquad \frac{g^2 v^2}{8} ((W_\mu^-)^\dagger W^{-\mu} + (W_\mu^+)^\dagger W^{+\mu}) + \frac{1}{2} \begin{pmatrix} W_\mu^{3\dagger} & B_\mu^\dagger \end{pmatrix} \boldsymbol{M}^2 \begin{pmatrix} W^{3\mu} \\ B^\mu \end{pmatrix} + \dots \end{aligned} \tag{2.22}$$

In Equation 2.22, the first term on the right-hand side is the mass term for the $W^\pm$ bosons. Note that we have constructed $W^\pm = \frac{1}{\sqrt{2}}(W^1 \mp iW^2)$ which are the charge eigenstates of the $W^{1,2}$ fields. The second term on the right-hand side mixes the W^3 and B fields through the mass matrix $\boldsymbol{M}$. This matrix has off-diagonal elements, so to

find the physical fields (i.e. mass eigenstates), we diagonalize $\boldsymbol{M}$ finding

$$\boldsymbol{M}_D^2 = \begin{pmatrix} 0 & 0 \\ 0 & \frac{v^2}{4}(g^2 + g'^2) \end{pmatrix}. \tag{2.23}$$

One of the diagonal entries of the mass matrix is 0, corresponding to the photon, and the non-zero entry is a mass term for the Z boson. The masses of the $W^{\pm}$ and Z bosons can be related through the weak mixing angle, defined as $\theta_W = \cos^{-1}\left(\frac{g}{\sqrt{g^2+g'^2}}\right)$. The weak mixing angle also relates the relative amounts of the B and W^3 fields that make up the photon and the Z boson with

$$\gamma = \cos(\theta_W)B + \sin(\theta_W)W^3 \tag{2.24}$$

$$Z = -\sin(\theta_W)B + \cos(\theta_W)W^3. \tag{2.25}$$

2.4.2 The Yukawa couplings

In the previous section, we saw that by embedding a complex scalar doublet in the $SU(2)$ sector of the SM, we could naturally recover mass terms for the gauge bosons. It turns out that we can use this Higgs field to construct gauge invariant terms in the electroweak Lagrangian that mix the left- and right-chiral fermion fields. This is shown explicitly in Equation 2.26.

$$\begin{aligned} \mathcal{L}_{Yukawa} &= -\lambda_e \bar{\boldsymbol{L}} \cdot \boldsymbol{\Phi} e_R - \lambda_d^{ij} \bar{\boldsymbol{Q}}^{\boldsymbol{i}} \cdot \boldsymbol{\Phi} d_R^j - \lambda_u^{ij} \bar{\boldsymbol{Q}}^{\boldsymbol{i}} \cdot \boldsymbol{\Phi}_c u_R^j + h.c. \\ &\{\boldsymbol{\Phi} = \langle \boldsymbol{\Phi} \rangle\} \\ &= -\frac{1}{\sqrt{2}}\lambda_e v \bar{e}_L e_R - \frac{1}{\sqrt{2}}\lambda_d^{ij} v \bar{d}_L^i d_R^j - \frac{1}{\sqrt{2}}\lambda_u^{ij} v \bar{u}_L^i u_R^j + h.c. \end{aligned} \tag{2.26}$$

Each term in Equation 2.26 contains a coupling constant, λ, a left-chiral fermion doublet field, e.g. $\bar{\boldsymbol{L}}$, the Higgs complex doublet field, $\boldsymbol{\Phi}$, and a right-chiral fermion singlet field, e.g. e_R. Under a generic $SU(2)$ transformation, the product of the Higgs field with the fermion singlet field will transform as an $SU(2)$ doublet. Since each term in this equation also contains the left-chiral fermion fields, which will transform in the opposite way under $SU(2)$, these terms all remain invariant under local gauge transformations. Furthermore, when we evaluate each of these terms at the vev for the Higgs field, we explicitly see mass terms as in Equation 2.11, where we identify the mass of the fermion, $m_f = \frac{1}{\sqrt{2}}\lambda_f v$. These terms represent a coupling between the fermion fields and the vev. It should again be emphasized that this mechanism for giving mass to the fermion fields is entirely different from how the Higgs field gives mass to the gauge bosons.

2.4.3 The Higgs boson

Throughout this section, we've seen how we can use a complex scalar doublet field to give mass to gauge bosons and fermions in the SM. We now turn our attention to a prediction of this theory: the existence of a massive scalar particle.

By substituting Equation 2.20 for $\boldsymbol{\Phi}$ in the Higgs potential, we find that

$$V(H(x)) = -\frac{1}{4}\mu^2 v^2 + \mu^2 H^2 + \lambda v H^3 + \frac{1}{4}H^4. \tag{2.27}$$

The first term in Equation 2.27 is a constant and can be ignored. The second term is a mass term for a real particle with mass, $m_H = \sqrt{2\mu^2} = \sqrt{2\lambda}v$. This particle is the SM Higgs boson. Just as we used $H(x)$ to parametrize small radial perturbations of the

Higgs field about its minimum, we identify the Higgs boson as an excitation of the Higgs field. The third and fourth terms represent interactions of the Higgs boson with itself, as shown in Figure 2.2. In addition to the Higgs boson self-couplings, the Higgs boson

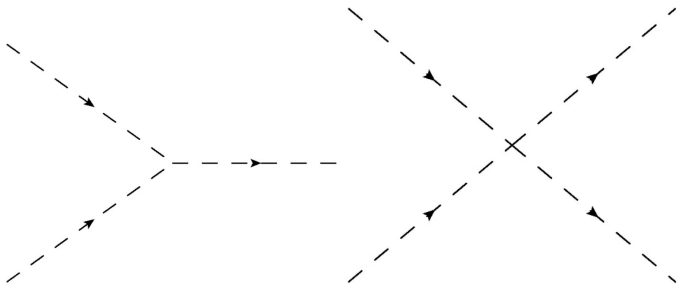


Figure 2.2: Feynman diagrams representing three- and four-Higgs boson couplings

also couples to the massive gauge bosons and the fermions. In Equations 2.22 and 2.26, wherever the massive bosons and fermions couple to the vev (v), we can replace those terms with a coupling to the excitation of the Higgs field, $H(x)$. An important feature of the fermion couplings to the Higgs boson is that these couplings are proportional to the fermion mass. This has important implications for which properties of the Higgs boson we can expect to observe in experimental data.

In July 2012, both the ATLAS and CMS experiments announced the discovery of a particle whose properties are consistent with those of a SM Higgs boson [11] [12]. This particle is observed to have a mass of 125 GeV, and was discovered through its decays to gauge bosons. Individual experiments at the LHC have not yet observed the decay of this particle to fermions, and this observation will be an important milestone in the characterization of this recently discovered boson. In Chapter 3 we discuss the

phenomenology of Higgs boson physics at hadron colliders, including the challenges in measuring fermionic decays of the Higgs boson.

2.5 Beyond the Standard Model

Although nearly all experiments to date agree well with SM predictions, there are compelling reasons to believe that a more complete theory may emerge in the future. Several new theories have been proposed to address the limitations of the SM, but there is not yet any direct experimental evidence to support these BSM theories.

2.5.1 Standard Model limitations

The SM has several shortcomings. Two of these are described here: the hierarchy problem and the lack of a dark matter candidate in the SM.

2.5.1.1 The Hierarchy problem

In its simplest form the hierarchy problem is the lack of an explanation for why gravity is many orders of magnitude weaker than the other fundamental forces. The problem is intimately related to the Higgs sector of the SM, which requires a spectacular degree of fine tuning. Since the Higgs boson is a scalar, its mass squared parameter receives quantum corrections from loop diagrams containing Dirac fermions, as shown in Figure 2.3 [13]. The size of these corrections is given in Equation 2.28, where λ_f is the coupling strength between the Higgs and fermion fields and Λ_{UV} is an ultraviolet momentum cutoff scale interpreted as the energy scale at which new physics affects the

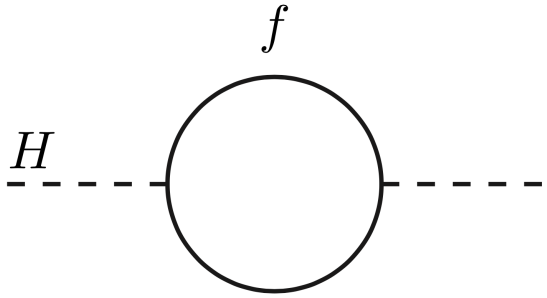


Figure 2.3: One loop quantum correction to the Higgs mass squared parameter from fermion fields.

predictions of the SM.

$$\Delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2}\Lambda_{\mathrm{UV}}^2 + \cdots \tag{2.28}$$

It is expected that the SM will breakdown at the reduced Planck scale $M_{\mathrm{P}} = (8\pi G_{\mathrm{Newton}})^{-1/2} = 2.4 \times 10^{18}\,\mathrm{GeV}$ where a theory describing quantum gravity will be necessary, and if there is no intermediate mass scale for new physics, the size of the corrections in Equation 2.28 will be ≈ 30 orders of magnitude larger than the observed value $m_H^2 \approx -(125\,\mathrm{GeV})^2$. This means that a miraculous cancellation is required between the bare Higgs mass squared parameter and the corrections described in this section to produce the observed value of m_H^2. Such an unlikely fine tuning is seen as a hint that new physics may exist at a scale below M_{P} and possibly within our experimental reach.

2.5.1.2 Dark matter

The existence of matter not described by the SM has been proposed as a solution for some discrepancies between the theory and observation of the motion of astrophysical objects, and is referred to as Dark Matter (DM) [8]. One category of

these discrepancies is galactic rotation curves. The rotational velocity v of an object at radial distance r outside a massive core should scale as $v(r) \propto 1/\sqrt{r}$. In many cases, the observed relationship for objects orbiting outside the luminous part of a galaxy is $v \approx$ const., implying the presence of DM with mass density $\rho(r) \propto 1/r^2$. Other evidence for DM comes from observations of clusters of galaxies, gravitational lensing, and analysis of the power spectrum of the Cosmic Microwave Background. Furthermore, analysis of structure formation in the universe tells us that DM must have been non-relativistic at the onset of galaxy formation, thus ruling out neutrinos as DM candidates because they are too light. Together, these observations suggest a profile of the DM candidate as a stable, weakly interacting, and sufficiently massive particle. There are no SM particles fitting these criteria, so evidence for DM is evidence for BSM physics.

2.5.2 Physics beyond the SM

Several theories have been proposed to address one or more of the limitations of the SM. A feature common to many of these theories is a modified Higgs sector with possible changes to SM Higgs couplings. Evidence for these theories may present itself through careful measurements of Higgs boson production and decay rates relative to SM predictions. The theories with the most theoretical motivation are supersymmetric theories, discussed next.

2.5.2.1 Supersymmetry

Supersymmetry is the name given to a proposed symmetry in nature that relates fermions and bosons [13]. In supersymmetric theories, all SM particles should have a partner with identical properties, except that of spin which differs by a unit of 1/2. These particles have not been observed, so if nature is described by a supersymmetry, it must be a broken symmetry with the masses of supersymmetric partners not yet within our experimental reach. Supersymmetry contains a solution to the hierarchy problem because the Δm_H^2 correction from Equation 2.28 is cancelled by a (nearly) equal, but opposite-sign correction term from the supersymmetric partners of the SM fermions. Additionally, in many classes of supersymmetric models, the lightest supersymmetric particle is stable and weakly interacting, making it a good candidate for a DM particle.

In the simplest realistic model of supersymmetry, the Higgs sector is described by a two-Higgs-doublet model (2HDM) [8]. There are five physical Higgs bosons in 2HDMs, and the lightest CP-even scalar, h, is often assumed to be the SM Higgs boson. In these theories, the couplings of h to other SM particles are modified relative to the SM values. The tree-level coupling between h and down-type quarks (e.g. bottom quarks), relative to the value for the SM Higgs boson, for some classes of these models, is given in Equation 2.29.

$$\left(\frac{y_{2HDM}}{y_{SM}}\right)_{hb\bar{b}} = -\frac{\sin(\alpha)}{\cos(\beta)} \tag{2.29}$$

In this equation, β is defined such that $\tan\beta$ is the ratio of the vacuum expectation values of each Higgs doublet field, and α is a parameter describing the mixing between

h and the heavier CP-even scalar. For larger values of $\tan\beta$, the coupling strength of h to bottom quarks increases in magnitude relative to the SM expectation, so observing a greater than expected Higgs boson branching ratio to bottom quark-antiquark pairs would be evidence for these classes of BSM theories.

Chapter 3

Higgs Bosons at the LHC

Higgs bosons can only be produced in sufficient quantities for study at particle colliders. After being produced they quickly decay to other elementary particles, so we study them by looking for their decay products amongst other particles resulting from these collisions. The only facility where we are currently able to produce the Higgs boson is the Large Hadron Collider, discussed in Chapter 4. This chapter discusses the relevant Higgs boson production mechanisms at the LHC and the decay modes that are accessible despite the experimental challenges at hadron colliders.

3.1 Production mechanisms

The dominant Higgs boson production mechanisms that are relevant at the LHC are: gluon fusion ($pp \to H$), vector boson fusion ($pp \to qqH$), associated VH production ($pp \to (W/Z)H$), and associated $t\bar{t}$ production ($pp \to t\bar{t}H$) [8]. The cross section is shown for each of these mechanisms as a function of the center-of-mass en-

ergy $\sqrt{s}$ in Figure 3.1, and the leading order Feynman diagrams for each are shown in Figure 3.2.

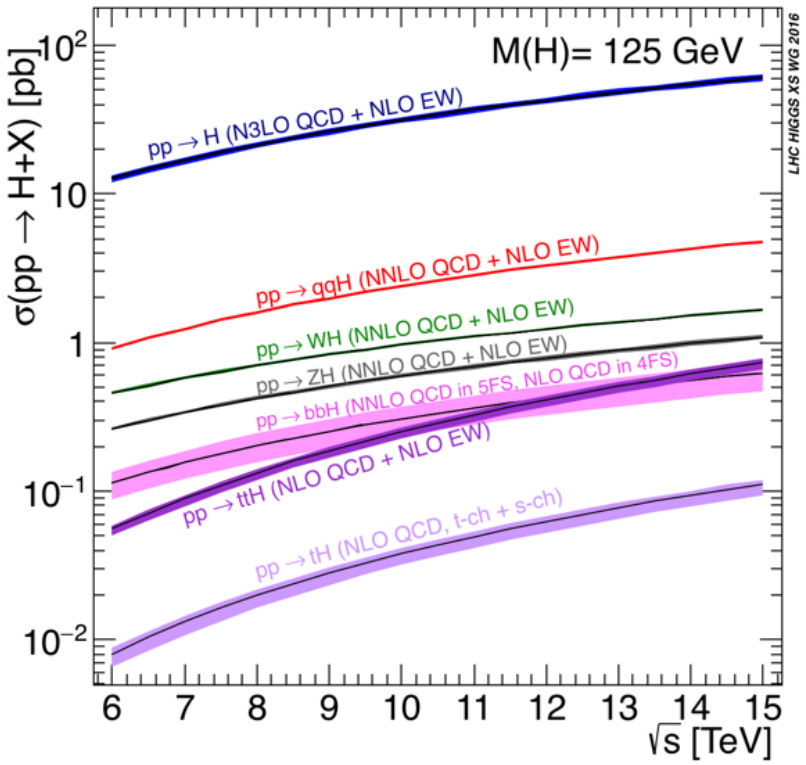


Figure 3.1: Higgs boson production cross sections as a function of $\sqrt{s}$ at the LHC.

Gluon fusion is the dominant production mechanism for the Higgs boson at hadron colliders. Gluons are not massive, so they do not couple directly to the Higgs boson; instead they couple indirectly to the Higgs boson through a quark loop. The dominant contribution to this loop is from top quarks, and contributions from other quarks are suppressed proportional to m_q^2.

VBF Higgs boson production has the second largest cross section at the LHC.

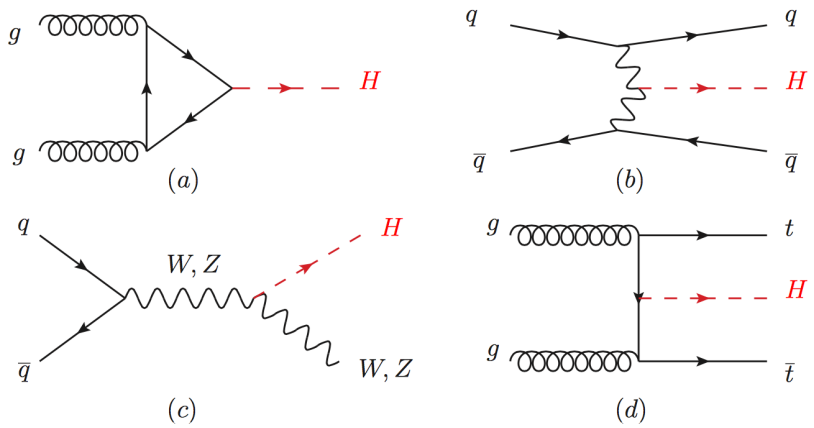


Figure 3.2: Feynman diagrams for Higgs boson production at the LHC: (a) gluon fusion (b) vector boson fusion (c) associated VH production (d) associated $t\bar{t}$ production.

In this mechanism, initial state quarks scatter via a t- or u-channel exchange of a $W^{\pm}$ or Z boson, which then radiates a Higgs boson. The experimental signature for these events is two hard jets in the forward regions of the detector and little hadronic activity between these forward jets due to the color flow in the event.

Higgs boson production in association with a vector boson is an important production mechanism at the LHC due to the decay of the vector boson. The LHC produces an overwhelming number of multijet events, but relatively few leptons. Looking for leptonic decays of the vector bosons helps to reduce the large number of multijet background events in many Higgs boson analyses.

Higgs boson production in association with top quarks is important because it allows direct access to information on the top quark Yukawa coupling, which is not accessible through Higgs boson decays since $m_H < 2m_t$. Gluon fusion provides some access to this coupling, but other heavy particles, both SM and BSM, propagating in

this loop affect the measurement.

3.1.1 Parton distribution functions

As can be seen in the Feynman diagrams, although the LHC collides protons, it is the quarks and gluons (collectively referred to as partons) that participate in the fundamental interactions producing Higgs bosons. When calculating production cross sections at the LHC, we must consider both the hard-scatter probabilities associated with the Feynman diagrams and the probabilities that the initial state partons in those diagrams carry sufficient fractions of the proton's momentum. The latter of these probabilities are described by parton distribution functions (PDFs) which give the probability that a parton i carries momentum fraction x of the proton when probed at scale Q.

3.2 Branching ratios

The expected branching ratios for some decay channels of a 125 GeV Higgs boson are listed in Table 3.1. Though the photon is massless, it couples indirectly

Decay channel	Branching ratio
$H \to \gamma\gamma$	2.27×10^{-3}
$H \to ZZ$	2.62×10^{-2}
$H \to W^+W^-$	2.14×10^{-1}
$H \to \tau^+\tau^-$	6.27×10^{-2}
$H \to b\bar{b}$	5.84×10^{-1}
$H \to Z\gamma$	1.53×10^{-3}
$H \to \mu^+\mu^-$	2.18×10^{-4}

Table 3.1: Expected branching ratios for a 125 GeV Higgs boson.

to the Higgs boson through loops involving massive charged particles, allowing the $H \to \gamma\gamma$ decay mode. Though it has a small branching ratio, this channel is particularly important due to the excellent resolution of photons in the LHC detectors and the relatively small background in this channel. The $H \to ZZ$ decay followed by leptonic decays of both Z bosons is similarly important for the same reasons.

Since fermions and bosons acquire mass through separate mechanisms involving the Higgs field, it is important to measure the Higgs boson couplings to both fermions and bosons. Searches for $H \to b\bar{b}$ and $H \to \tau^+\tau^-$ target the Higgs boson fermionic couplings, while searches for the other Higgs boson fermionic decay modes are not as promising due to the dependence of their couplings on m_f.

3.3 Discovery

Prior to its discovery in 2012, the mass of the Higgs boson was not known. The strategy used to find the Higgs boson was to search for a narrow peak over a smoothly falling distribution, making $H \to \gamma\gamma$ and $H \to ZZ^* \to 4l$ the relevant channels for this search. Given the small branching ratio in each of these channels, the important production mechanisms for these searches are the ones with the largest cross sections: gluon fusion, vector boson fusion, and associated VH production. In the case of $H \to \gamma\gamma$, the dominant backgrounds are SM diphoton production, γ+jet production, and dijet production. The latter two are relevant when one or more jets are misidentified as photons. For $H \to ZZ^* \to 4l$ the dominant background is from continuum $Z^{(*)}/\gamma^*$

production with other important contributions from Z+jets and $t\bar{t}$ production, where charged lepton candidates arise either from decays of heavy-flavor hadrons or from misidentification of jets. Both ATLAS and CMS observed an excess in these channels near $m_{\gamma\gamma(4l)} = 125\,\text{GeV}$, and subsequently focused their low resolution $H \to W^+W^- \to 2l2\nu$ searches on this mass point. This channel enhances the sensitivity of the analysis due to the relatively high branching ratio for $H \to W^+W^-$. The dominant backgrounds in this channel are from non-resonant WW, $t\bar{t}$ and Wt production, as well as Z/γ^*, W+jets, and $W\gamma$ events.

At the time of discovery, ATLAS (CMS) reported an excess of events above the background-only hypothesis with a local significance of 5.9 (5.0) standard deviations [11] [12]. The presence of the $H \to \gamma\gamma$ decay meant that the newly discovered boson had a spin different from one. Figure 3.3 shows as a function of m_H the ATLAS expected and observed p-values (left) and the diphoton mass spectrum in the CMS search (right). In both plots, the disagreement between the data and background-only hypothesis indicates $m_H \approx 125\,\text{GeV}$. These searches, and subsequent analyses with more data, have established the Higgs boson production and decays to gauge bosons. The fermionic decay modes are more difficult to observe, and are discussed next.

3.4 Fermionic decay modes

The most promising searches for the fermionic decay modes of the SM Higgs boson at the LHC include $H \to \tau^+\tau^-$ and $H \to b\bar{b}$ due to their large branching ratios

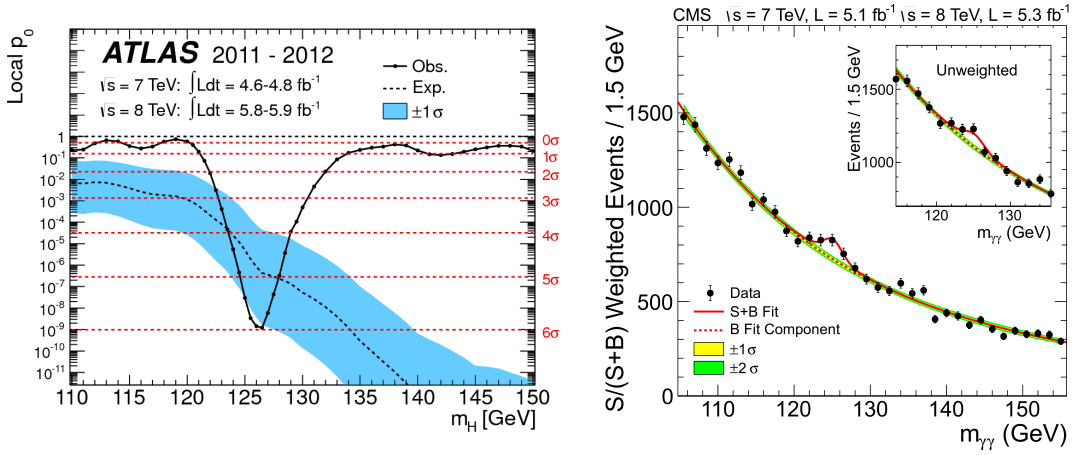


Figure 3.3: Left: expected and observed p-values from the ATLAS Higgs boson discovery for all channels combined. Right: observed diphoton mass spectrum in the CMS experiment from the Higgs boson discovery.

compared to Higgs boson decays to other leptons and quarks.

The $H \to \tau^+\tau^-$ searches target the gluon fusion and VBF production mechanisms and are combined with dedicated searches for VH production. The dominant background to these searches is $Z \to \tau^+\tau^-$ with additional contributions from W+jets and multijets. Individually ATLAS (CMS) measured a 4.5 (3.2) standard deviation significance in their searches for $H \to \tau^+\tau^-$ [14] [15], which when combined yields an observation of that decay with a significance of 5.5 standard deviations. This is the first observation of the Higgs boson decay to fermions [3].

Both ATLAS and CMS have searches for $H \to b\bar{b}$ in three different production modes: vector boson fusion, associated VH production, and associated $t\bar{t}$ production. The dominant gluon fusion production mechanism is not useful for this search due to difficulties in suppressing the overwhelming multijet background. Of these three, the

VH production mechanism is the most promising. It uses the leptonic decay of the vector boson for triggering and to reject the multijet background. The dominant backgrounds for this search are $W/Z+$ jets, $t\bar{t}$, and diboson production. The ATLAS (CMS) result using data from Run 1 observed a significance of 1.4 (2.1) standard deviations in the VH production mode [16] [17]. The analyses targeting the associated $t\bar{t}$ production mechanism are challenging due to difficulties in reconstructing the complex final state. In Run 1 ATLAS (CMS) measured a signal strength μ of 1.4 ± 1.0 (0.7 ± 1.9) in the search for associated $t\bar{t}$ production in the $H \to b\bar{b}$ channel [18] [19]. The VBF production mode is difficult to trigger and analyze due to its all-hadronic final state. In these analyses, the multijet background is suppressed by exploiting the VBF topology of two high-p_T forward jets with a large separation in η that yield a large invariant mass. The Run 1 ATLAS (CMS) analysis found $\mu = -0.8^{+2.3}_{-2.3}$ ($2.8^{+1.6}_{-1.4}$) for $H \to b\bar{b}$ in this production mode [20] [21].

3.5 VBF + photon

A variation of the vector boson fusion production mechanism of the Higgs boson that was not explored in Run 1 of the LHC is VBF production in association with a photon. The leading order Feynman diagrams for this process are shown in Figure 3.4 with $H \to b\bar{b}$. In this production process, the photon can be radiated from either the internal $W^{\pm}$ boson or one of the initial or final state quarks. The dominant background to this search is non-resonant multijet production in association with a

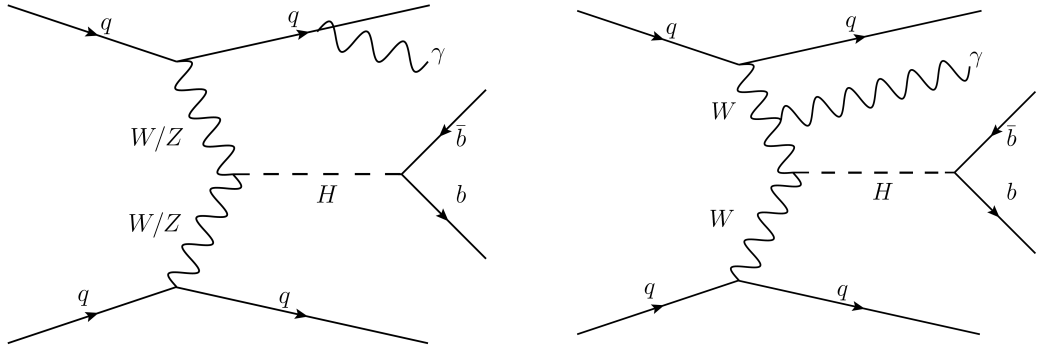


Figure 3.4: Leading order Feynman diagrams for VBF Higgs boson production in association with a photon and $H \to b\bar{b}$.

photon, as shown by the representative Feynman diagrams in Figure 3.5.

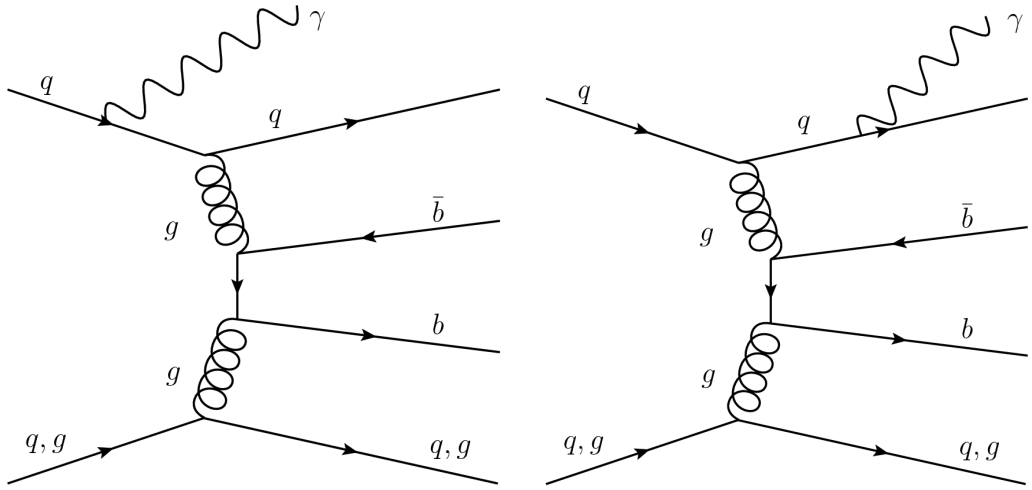


Figure 3.5: Representative leading order Feynman diagrams for the dominant background to VBF+γ production.

Compared to the inclusive VBF analysis, there are several advantages in requiring the photon. First, it provides a clean signature for efficient triggering in the ATLAS detector. The otherwise all-hadronic final state is difficult to trigger on at the

LHC due to its similarity to the large multijet background. Second, the photon naturally suppresses the gluon-induced component of the multijet background since gluons are neutral and don't radiate photons. Finally, there is a destructive quantum interference effect that further suppresses central photon emission in the dominant background process. This interference effect arises in the t-channel diagrams shown in Figure 3.5, between photon emission off the initial state quark radiating a gluon and photon emission off the corresponding final state quark. For the signal process, this destructive interference only affects the ZZ fusion process, while for WW fusion, the interference is additive as a result of the charged current changing the electric charge of the initial-final state quark [4]. These interference effects in the signal process enhance the sensitivity of this channel to the WWH coupling relative to the ZZH coupling.

Overall, these three advantages have the combined effect of reducing the signal cross section by a factor of approximately 100 whereas the dominant background is reduced by a factor of approximately 3000. Studies have shown that analyses in this channel may be competitive with inclusive VBF analyses at the LHC [4]. Part III of this dissertation presents the first analysis in the VBF+γ production mode targeting the $H \to b\bar{b}$ decay.

Part II

Experimental Setup

Chapter 4

The Large Hadron Collider

The Large Hadron Collider [22] is a circular particle accelerator that collides counter-circulating beams of protons and heavy ions, and is located at CERN on the French-Swiss border near Geneva, Switzerland. It has a circumference of 26.7 km and is installed in the tunnel that was formerly constructed for the LEP machine. The design center-of-mass energy is 14 TeV with a peak instantaneous luminosity $L = 10^{34}\text{cm}^{-2}\text{s}^{-1}$.

There are four major experiments at the LHC: ATLAS, CMS, LHCb, and ALICE. Both ATLAS and CMS are general purpose detectors designed to survey the new energy frontier. LHCb specializes in studying physics relating to b-quarks, and ALICE is designed to study heavy ion collisions.

The first proton beams circulated in the LHC on September 10, 2008. Nine days later, an electrical fault resulted in mechanical damage and a breach of liquid helium which is used as a cryogenic coolant. The damage caused by this incident delayed operation of the LHC until November 20, 2009, and on November 30, 2009, the LHC

became the world's highest energy particle accelerator, achieving 1.18 TeV per beam. This first operational run lasted until 2013, and during this time the LHC achieved collision energies of 7 and 8 TeV, which are about half of the design energy of 14 TeV. The landmark discovery during this run was a particle whose properties are consistent with those of a Standard Model Higgs boson. The second operational run began in April 2015 after a two-year shutdown for upgrades of the LHC and the experiments. This run is expected to last through 2018, with a collision energy of 13 TeV. The analysis presented in Part III uses data collected in 2015 and 2016 during Run 2.

4.1 LHC layout and design

A schematic of the LHC layout is shown in Figure 4.1. In this figure the red and blue circles indicate the two separate beams traveling in opposite directions. These beams cross at interaction points located inside the four major experiments and marked by the five-point stars. The LHC is divided into eight octants, and each octant contains a curved section and a long straight section. The experiments are located along the long straight sections of their respective octants. Octant 4 contains the RF cavities used for accelerating the protons, and the beam dump system is located in octant 6. Octants 3 and 7 contain collimation systems used for cleaning the beams.

As was mentioned, the LHC was built in the tunnel formerly occupied by the LEP machine which collided electrons and positrons. The decision to reuse the LEP tunnel was made for the cost savings, as a dedicated tunnel for a proton-proton collider

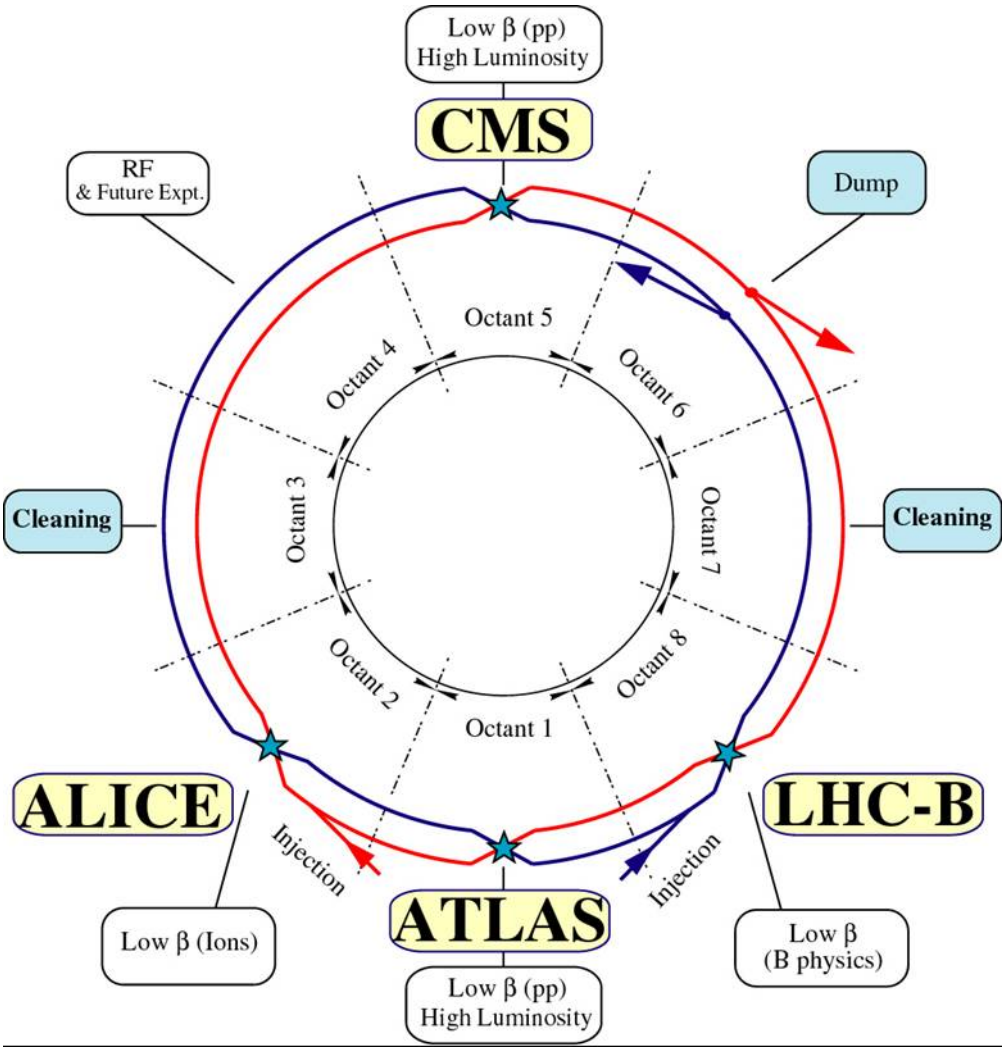


Figure 4.1: Overview schematic showing the relative location of each major experiment at the LHC.

would have been designed with a different geometry. The center-of-mass collision energy for LEP was limited by synchrotron radiation losses which were compensated with RF cavities in each of the long straight sections. The radiated power from synchrotron radiation is proportional to $\frac{1}{m^4}$, where m is the mass of the accelerating particle. Since protons are $\approx$2000 times heavier than electrons, the synchrotron radiation losses are $\mathcal{O}(10^{13})$ times smaller for a proton-proton collider compared to an electron-positron collider. The center-of-mass energy at the LHC is instead limited by our ability to design and build sufficiently strong magnets to bend the proton beam in each curved section. Had a dedicated tunnel been built for the LHC, it would have used longer arcs and shorter straight sections for the same circumference.

LEP collided electrons and positrons, which are identical except for their equal, but opposite, electric charge, so it was able to use a single ring and magnetic field for both counter-circulating beams. The LHC, on the other hand, requires separate magnetic fields for each proton beam. At 3.7 m in diameter, the LEP tunnel is too small to fit two completely separate magnetic rings, so a twin-bore (i.e. "two-in-one") magnet design was chosen. The LHC magnets are built using NbTi superconductors cooled to 2 K using super-fluid helium in a common cryostat for both beams. At the design beam energy of 7 TeV, the dipole magnets must produce a 8.33 T magnetic field.

4.2 Injection system

A chain of four particle accelerators is used to bring protons from rest to a sufficient energy that they can be injected into the LHC. The CERN accelerator complex, including these four accelerators, is shown in Figure 4.2. Hydrogen gas is ionized to produce protons which are first accelerated to 50 MeV using a linear accelerator, Linac 2. A series of circular particle accelerators - the Proton Synchrotron Booster, the Proton Synchrotron, and the Super Proton Synchrotron (SPS) - accelerate the protons to 1.4 GeV, 25 GeV, and 450 GeV respectively. Each of these accelerators was upgraded in order to meet the LHC luminosity and beam quality requirements. From the SPS, proton bunches are injected into the LHC at both octant 2 and octant 8. Each bunch contains $\mathcal{O}(10^{11})$ protons, and the spacing between bunches is 25 ns, resulting in 3564 possible bunch spaces per beam. Some of these bunches must be left empty due to injection and safety requirements, so the nominal number of bunches is 2808 per beam. Once injected, the beams take approximately 20 minutes to accelerate to their maximum design energy 7 GeV.

4.3 Performance

Since the restart of the LHC physics program in 2010 following the electrical fault repairs, the LHC has performed very well. Data have been recorded at three different center-of-mass energies: $\sqrt{s} = 7,\ 8,$ and 13 TeV in 2010-2011, 2012, and 2015-2016, respectively. At $\sqrt{s} = 7$ TeV, the LHC delivered approximately 5.5 fb^{-1} to both

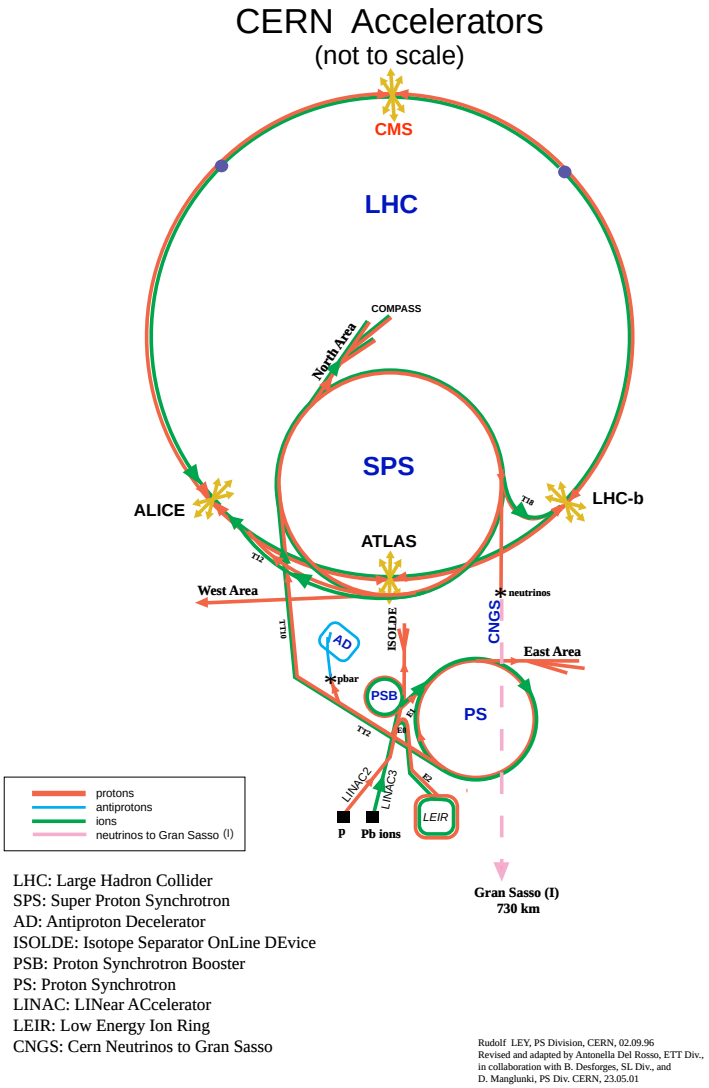


Figure 4.2: Overview of the accelerator layout at CERN.

ATLAS and CMS; at 8 TeV the integrated luminosity delivered to each experiment was roughly 23 fb^{-1}; and thus far at 13 TeV approximately 40 fb^{-1} have been delivered to ATLAS and CMS. The peak instantaneous luminosity so far at $\sqrt{s} = 13\,\text{TeV}$ is $13.7 \times 10^{33}\,\text{cm}^{-2}\text{s}^{-1}$. Figure 4.3 shows the delivered integrated luminosity by year in the ATLAS detector (left), as well as the peak instantaneous luminosity by day in 2016 in the ATLAS detector (right).

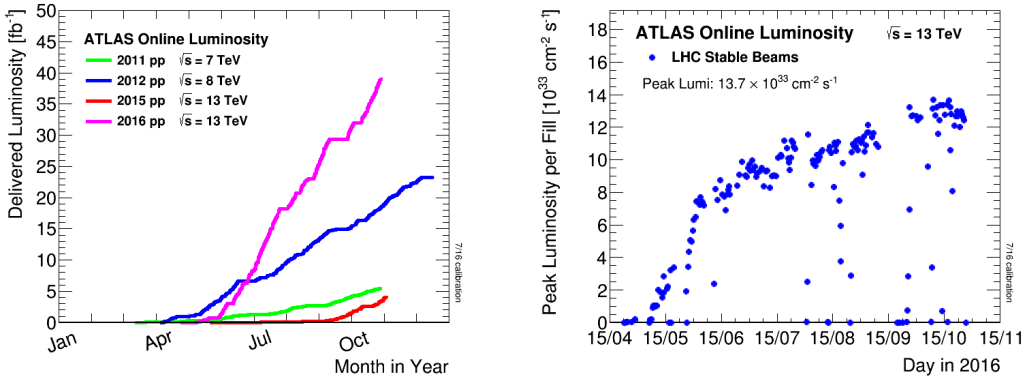


Figure 4.3: Left: delivered integrated luminosity by year in the ATLAS experiment. Right: peak instantaneous luminosity by day in 2016 in the ATLAS experiment.

Chapter 5

The ATLAS Detector

The analysis presented in Part III uses data collected with the ATLAS (A Toroidal LHC ApparatuS) detector [23]. The layout of the ATLAS detector is shown in Figure 5.1. ATLAS is located at Point 1 of the LHC in a cavern 91 m below the surface of the earth. It weighs a total of 7000 metric tons with dimensions of 43 m in length and 22 m in diameter [24].

The trajectory of particles within ATLAS is measured with respect to the nominal interaction point. The z-axis is defined by the beam direction such that when the LHC is viewed from above, the counter-clockwise circulating beam indicates positive-z. The $x-y$ plane is transverse to the beam line with the positive x-axis pointing from the interaction point to the center of the LHC ring and the positive y-axis pointing vertically upwards. The azimuthal angle ϕ is the angular distance around the z-axis with $\phi = 0$ corresponding to the x-axis, and θ is the polar angle from the z-axis. The radial distance from the z-axis to a given point is defined as $r = \sqrt{x^2 + y^2}$. The

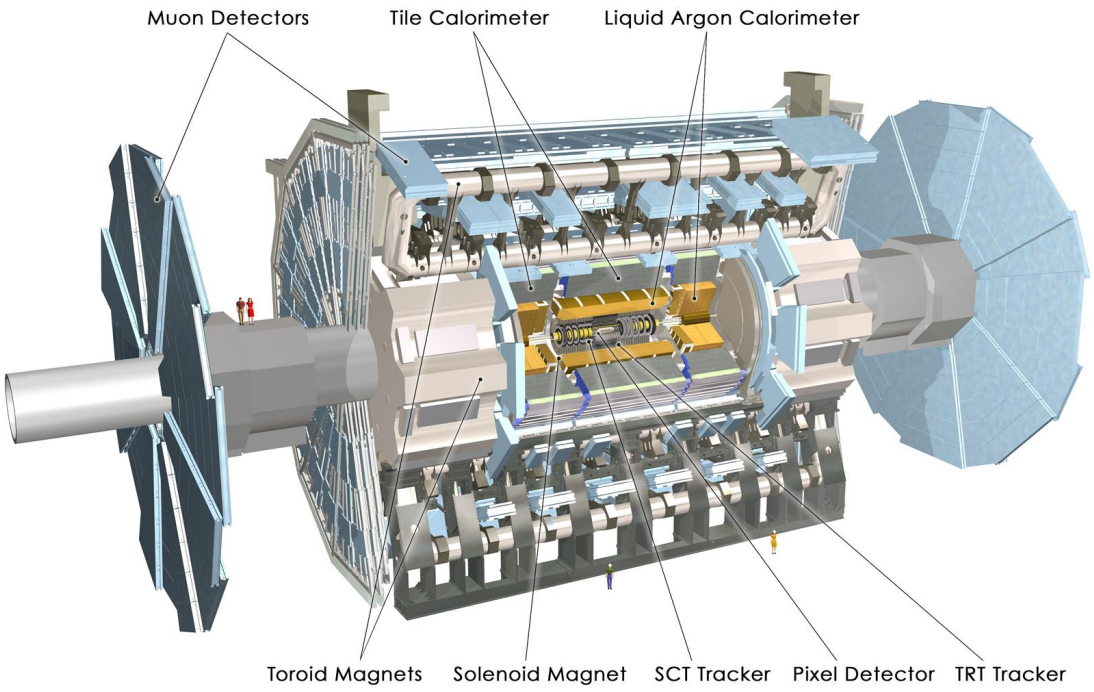


Figure 5.1: Layout of the ATLAS detector

momentum and direction of particles resulting from collisions in ATLAS are described by the following three quantities: p_T, η, and ϕ. $p_T = \sqrt{p_x^2 + p_y^2}$ is the magnitude of the momentum of the particle in the direction transverse to the beam line, and ϕ is the azimuthal angle already discussed. η is called the pseudorapidity and is defined in Equation 5.1.

$$\eta = -\ln[\tan(\theta/2)] \tag{5.1}$$

Pseudorapidity is closely related to another quantity called rapidity which is defined in Equation 5.2. For massless particles, pseudorapidity and rapidity are identical.

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) \tag{5.2}$$

In this equation, E is the energy of the particle and p_z is the component of its momentum along the z-axis. Particle production at the LHC is roughly constant as a function of rapidity, and differences in rapidity are Lorentz invariant under boosts along the z-axis. Nevertheless, pseudorapidity is preferred to rapidity because the former is purely a geometric quantity and is independent of particle energy. Angular separation between different particles in the ATLAS detector is measured by the quantity $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

A number of subsystems compose the ATLAS detector as a whole, and each of these is important in identifying and measuring the energy and trajectory of particles produced by collisions within ATLAS. These components use a mixture of barrel and end-cap geometries, with barrels sensitive to particles at smaller values of $|\eta|$ and forming concentric shells around the interaction point, while end-caps are sensitive to

particles at larger values of $|\eta|$ and constructed as disks perpendicular to the z-axis. The detector closest to the interaction point is the Inner Detector (ID), which performs a non-destructive measurement of the trajectory of charged particles using a solenoid to provide the necessary magnetic field to determine particle momenta. Beyond the ID are the electromagnetic and hadronic calorimeters which measure the energies of all particles except muons and neutrinos. Due to their mass, relatively long lifetime, and because they are not strongly interacting, high energy muons do not shower substantially in the calorimeters, and thus escape the detector. A dedicated muon system, including a toroidal magnet system, provides a redundant measurement of the position and momentum of muons. Neutrinos are not detected in ATLAS, but can be inferred by an imbalance in the total transverse momentum of the detector response. Together, these detectors have approximately 100 million electronic readout channels and utilize roughly 3000 km of cables.

5.1 Inner detector

Three detectors compose what is collectively referred to as the Inner Detector. The ID is designed to measure the momentum of charged particles in the region $|\eta| < 2.5$ by combining information from the three independent detectors. The ID is 5.3 m in length and 2.5 m in diameter and immersed in a 2 T axial magnetic field. The purpose of the magnetic field is to bend charged particle trajectories allowing for measurements of both a particle's momentum and the sign of its electric charge. Using data

collected in 2008 for hardware commissioning, the relative momentum resolution σ_p/p of high-p_T tracks in the ID was measured to be $(4.83 \pm 0.16) \times 10^{-4}\,\text{GeV}^{-1} \times p_T$ [25]. In addition to measuring individual particle tracks, the ID is also important for identifying multi-particle vertices through the extrapolation of tracks back to interaction points. To distinguish the primary interaction vertex from secondary vertices arising due to decays of heavy flavor hadrons away from the interaction point, it is important to have good resolution of both the transverse and the longitudinal impact parameters. The transverse (longitudinal) impact parameter resolution is better than 100 (180) μm and improves with increasing track p_T [26]. The innermost detector is the pixel detector (pixels), followed by the semiconductor tracker (SCT), and finally the transition radiation tracker (TRT). Both the pixels and SCT are precision tracking detectors based on silicon technology, and the TRT provides continuous tracking using gas-based drift tubes. Two different views of the ID are shown in Figure 5.2.

5.1.1 Pixel detector

The ATLAS pixel detector consists of four barrel layers at r = 32.7, 50.5, 88.5, and 122.5 mm extending to $|z|$ = 400.5 mm and six end-caps disks covering r = 88.8–149.6 mm at z = ±495, ±580, and ±650 mm. The innermost pixel barrel, called the Insertable b-Layer (IBL), only extends to $|z|$ = 332 mm and was installed during the LHC shutdown following Run 1. There are 1744 identical pixel sensors, each containing 46080 readout channels. At roughly 80 million total individual pixels, the pixel detector contains over half of the readout channels in ATLAS. Most of the individual pixels

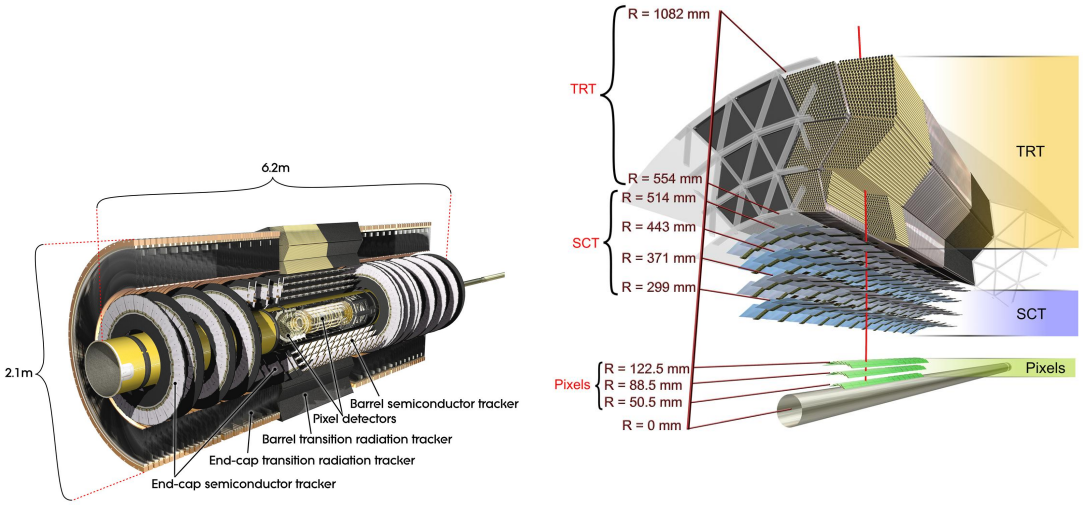


Figure 5.2: Cut-away (left) and cross-section (right) views of the ATLAS inner detector. The innermost pixel layer (IBL) is not shown.

measure $50\times400\,\mu\text{m}^2$, with approximately 10% measuring $50\times600\,\mu\text{m}^2$. Typically four pixel layers are crossed by each track, and the intrinsic measurement accuracy of a single pixel module for the barrel (end-cap) is $10\,\mu\text{m}$ in $r-\phi$ and $115\,\mu\text{m}$ in z (r).

5.1.2 Semiconductor tracker

The ATLAS semiconductor tracker consists of four barrel layers at $r = 299$, 371, 443, and 514 mm extending to $|z| = 749$ mm and 18 end-cap disks of variable radial extent between $z = \pm(854\text{–}2720\,\text{mm})$. Both the barrels and the end-caps are built from double-sided silicon strip modules. In the barrel region, pairs of 6.4 cm strips on each side of the rectangular module are daisy chained together to make effective 12.8 cm strips with a pitch of $80\,\mu\text{m}$. The end-cap uses trapezoidal modules and strips are oriented in the radial direction with a constant separation in ϕ and a mean pitch of $\approx80\,\mu\text{m}$.

Since each strip only provides a measurement in one direction, the strips on opposite sides of a given module are oriented at a 40 mrad angle with respect to each other. The two-dimensional position of a track can be determined by locating the crossing point of opposite sided strips registering a hit. Typically eight strip layers (totaling four space points) are crossed by each track, and the intrinsic measurement accuracy of a single SCT module for the barrel (end-cap) is 17 μm in $r - \phi$ and 580 μm in z (r).

5.1.3 Transition radiation tracker

The transition radiation tracker consists of a barrel region with radial extent r = 563–1066 mm covering $|z| < 712$ mm and an end-cap region with radial extent r = 644–1004 mm covering $z = \pm$(848–2710 mm). Unlike the pixels and SCT which provide tracking in $|\eta| < 2.5$, the TRT coverage only extends to $|\eta| < 2.0$. The TRT is built using 4 mm diameter straw drift tubes and contains roughly 351,000 readout channels. Each tube is filled with a Xe-CO_2-O_2 gas mixture and has at its center a 31 μm diameter gold-plated tungsten wire which detects ionization of the gas. Between the tubes are polymer fibers and foils which allow for transition radiation (TR) to be emitted when charged particles traverse the boundary between materials with different dielectric constants. The amount of TR emitted depends on the Lorentz factor of the charged particle and thus lighter particles (e.g., electrons), which typically have higher Lorentz factors, produce more TR. The TRT uses a two-threshold detection system to identify particles such as electrons, which produce more TR, and thus trigger more high-threshold hits. In the barrel region, the straws have length 144 cm parallel to the

z-axis, and their wires are divided into two halves at $\eta \approx 0$, while the end-cap straws are arranged radially in wheels and have length 37 cm. There are typically 36 TRT hits per track. The TRT only provides tracking information in the $r - \phi$ plane, and the intrinsic drift-time accuracy of a single straw for the TRT is 131 μm.

5.2 Calorimeters

The energies of photons, electrons, and jets are measured using the ATLAS calorimeter system which is located just outside of the ID, and is shown in Figure 5.3. ATLAS uses sampling calorimeters with separate materials for the passive and active layers (vs. homogeneous calorimeters for which the active and passive materials are the same). Sampling calorimeters can be constructed more cheaply and more compactly (than homogeneous calorimeters) because high density materials such as iron and lead can be used as the passive material. ATLAS uses a variety of active+passive material combinations including lead+liquid argon, steel+scintillating tiles, copper+liquid argon, and tungsten+liquid argon. The disadvantage of sampling calorimeters is they typically have a worse energy resolution (compared to homogeneous calorimeters) due to random fluctuations of the shower development in the active vs. passive layers.

ATLAS calorimeters are designed to provide good containment of electromagnetic and hadronic showers. This requirement is not only important for the measurement of particle energies, but also to limit punch-through into the muon system and to infer the presence of weakly interacting particles, such as neutrinos. Neutrinos (as well as

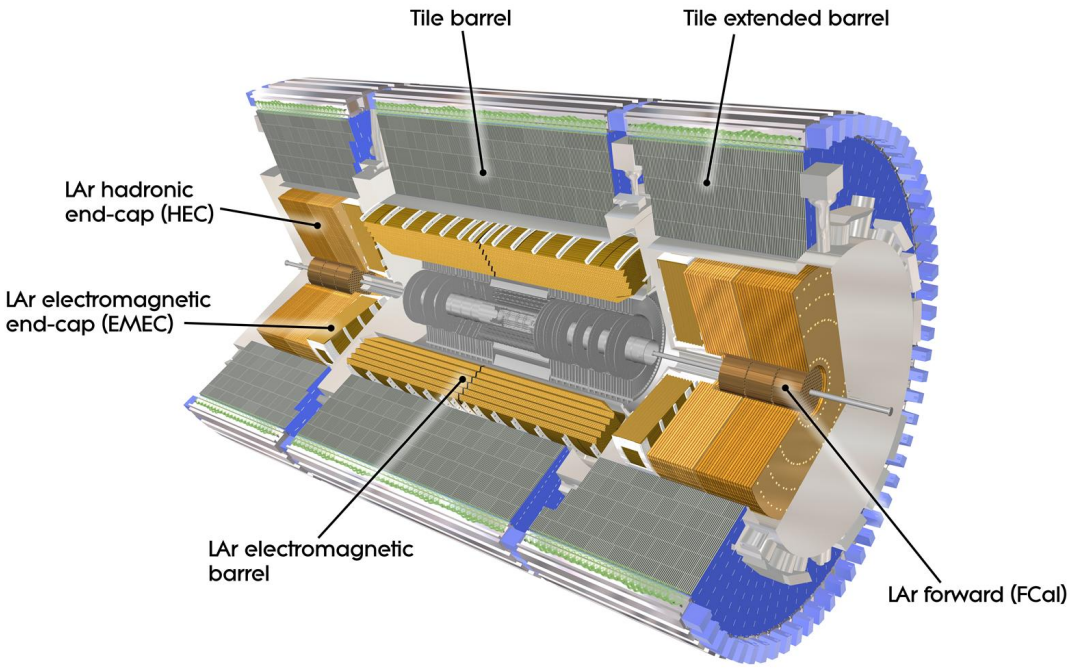


Figure 5.3: Cut-away view of the ATLAS calorimeter system.

some hypothetical particles from BSM scenarios) cannot be directly detected in the ATLAS experiment. Their presence, however, can be inferred using the momentum of other final state objects. Since the initial colliding protons in the LHC do not have any net transverse momentum, conservation of momentum requires that the vector sum of the momenta of all final state objects in a given event should not have a transverse component. If such a component is observed, it can either be attributed to a mis-measurement of the momentum of one or more objects or attributed to an invisible particle such as a neutrino. To reduce spurious transverse momentum imbalance measurements due to particles falling outside the detector acceptance, the calorimeters have near hermetic coverage of the interaction point in the range $|\eta| < 4.9$.

5.2.1 Electromagnetic calorimeter

The innermost calorimeter is the electromagnetic (EM) calorimeter which uses lead+liquid argon technology to shower and measure the energy of electrons and photons, which primarily lose energy due to bremsstrahlung and pair-production, respectively. It consists of a central barrel region covering $|\eta| < 1.475$ and two end-cap regions covering $1.375 < |\eta| < 3.2$. The end-caps are each subdivided into coaxial wheels with each outer wheel covering $1.365 < |\eta| < 2.5$ and each inner wheel covering $2.5 < |\eta| < 3.2$. In the region $|\eta| < 2.5$, the EM calorimeter has three active layers, including a first layer finely segmented [in η], which allows for an accurate position measurement. This first layer is also useful for discriminating photons from overlapping $\pi^0 \to \gamma\gamma$ decays, as the shower shapes for the latter typically exhibit a double peak

structure in the fine segmentation. A sketch of a barrel section of the EM calorimeter is shown in Figure 5.4. The EM calorimeter is roughly 22 radiation lengths deep in the

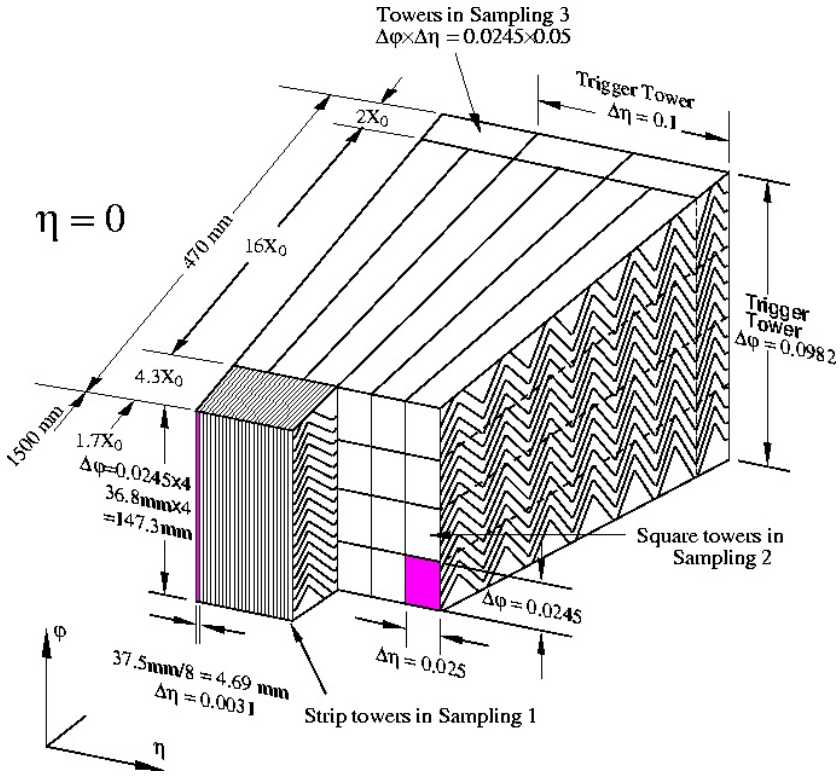


Figure 5.4: Sketch of a barrel section of the EM calorimeter.

barrel sections and at least 24 radiation lengths deep in the end-cap, where a radiation length is the mean distance traveled by an electron before losing all but $\frac{1}{e}$ of its energy due to bremsstrahlung. Both the barrel and end-cap sections use an accordion geometry for the absorber, which allows full coverage in ϕ without any cracks, and the readout electrodes are located in gaps between the absorbers. The EM calorimeter is designed to have an energy resolution $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%$.

5.2.2 Hadronic calorimeter

Several different technologies are used for the hadronic calorimeters in ATLAS, which are located outside of the EM calorimeters. The central region covering $|\eta| < 1.7$ uses steel+scintillating tile and is composed of a central barrel region covering $|\eta| < 1.0$ and an extended barrel on both ends covering $0.7 < |\eta| < 1.7$. The scintillating tiles are read out by wavelength-shifting fibers connected to photomultiplier tubes, as shown by the sketch of a tile calorimeter assembly in Figure 5.5. Including the ID and EM calorimeter, there is a total depth of 9.7 nuclear interaction lengths at $\eta = 0$ between the interaction point and outer edge of the hadronic calorimeter. A nuclear interaction length is the average distance traveled by a hadron before undergoing an inelastic nuclear interaction. In the central barrel (extended barrel), the tile calorimeter is divided into three layers that are approximately 1.5 (1.5), 4.1 (2.6), and 1.8 (3.3) nuclear interaction lengths thick. The tile calorimeter is designed to have an energy resolution $\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%$.

The hadronic end-cap calorimeter uses copper+liquid argon covering $1.5 < |\eta| < 3.2$ and consists of two wheels in each end-cap region located directly behind the EM calorimeter end-caps. Each wheel is further subdivided into two segments in depth for a total of four layers in each end-cap. The hadronic end-cap calorimeter is designed to have an energy resolution $\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%$.

The forward calorimeter is approximately 10 nuclear interaction lengths deep and covers the region $3.1 < |\eta| < 4.9$. It is built using three modules in each end-cap.

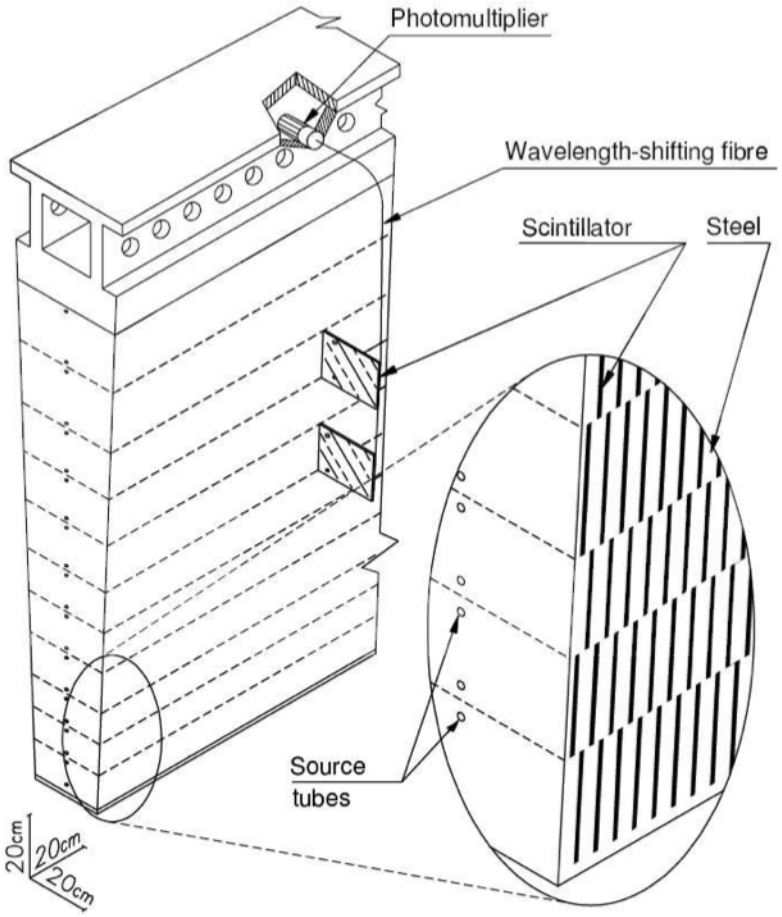


Figure 5.5: Sketch of the the tile calorimeter assembly and readout.

The first module uses copper as the passive medium and is optimised for electromagnetic measurements, while the second and third modules are made of tungsten and designed to measure hadronic activity. The active medium in all modules is again liquid argon. The forward calorimeter is designed to have an energy resolution $\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\%$.

5.3 Muon spectrometer

A dedicated system is designed to measure the momentum of muons in the range $|\eta| < 2.7$ and provide triggering capabilities for muons in $|\eta| < 2.4$. Unlike electrons, high-energy muons primary lose energy in the calorimeters due to ionization rather than bremsstrahlung. The total radiated power from bremsstrahlung is proportional to $\frac{1}{m^4}$ or $\frac{1}{m^6}$ depending on the direction of acceleration relative to velocity. Since muons are roughly 200 times heavier than electrons, they lose substantially less energy in the calorimeters and can escape to be measured in the muon system. Four different types of detectors compose the muon spectrometer, which is shown in Figure 5.6.

The Monitered Drift Tube tracking chambers (MDTs) cover $|\eta| < 2.7$ and use between three and eight layers of drift tubes to achieve an average resolution of 35 μm per chamber. Cathode-Strip Chambers (CSCs) are used in the forward region $2 < |\eta| < 2.7$ as the innermost tracking layer (instead of MDTs) because they can withstand a higher rate and have better time resolution. CSCs are multiwire proportional chambers with a per-chamber resolution of 40 μm in the bending plane and 5 mm in the transverse plane.

The tracking detectors are supplemented by two types of trigger chambers. In

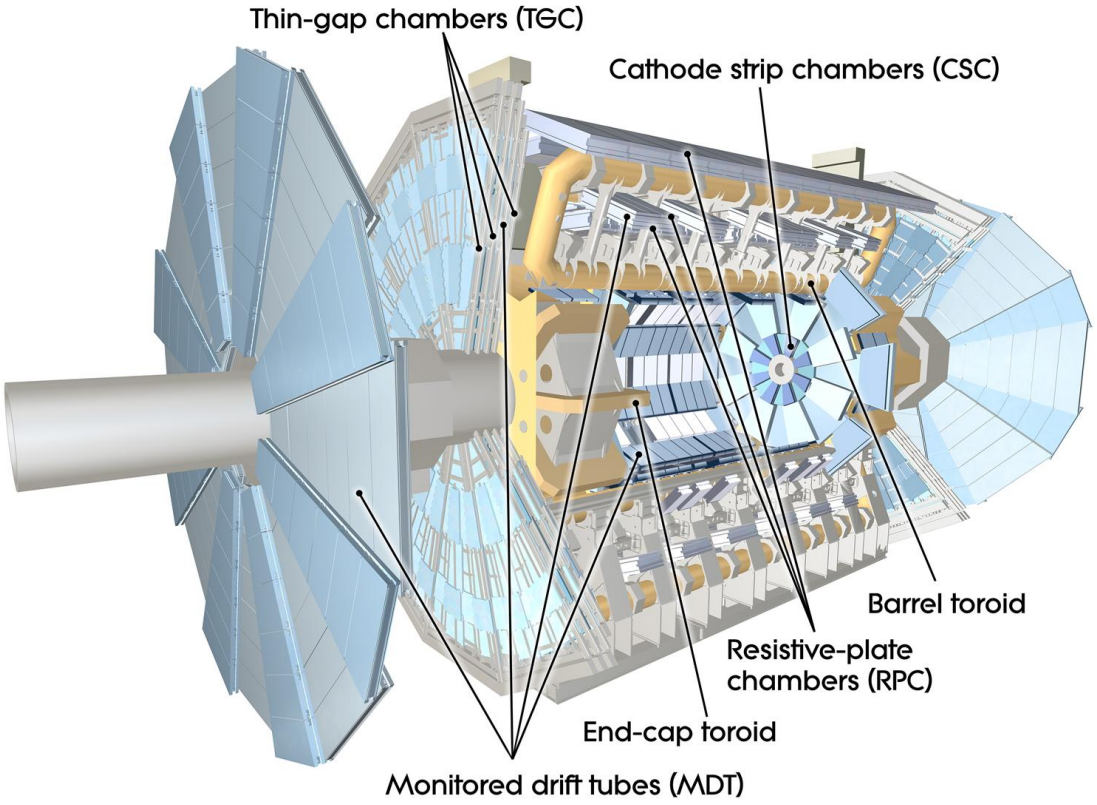


Figure 5.6: Cut-away view of the muon spectrometer.

$|\eta| < 1.05$ Resistive Plate Chambers (RPCs) are used, and in $1.05 < |\eta| < 2.4$ Thin Gap Chambers (TGCs) are used. The RPCs use metallic strips mounted on parallel electrode plates to detect ionization in a gas mixture caused by charged particles, and they have an intrinsic time resolution of 1.5 ns. TGCs are multi-wire proportional chambers with the defining characteristic that their wire-to-cathode distance is smaller than their wire-to-wire distance, and they have an intrinsic time resolution of 4 ns. Combined with the time for signal propagation and electronic processing, both of these chambers are able to deliver signals within 15-25 ns and thus have the ability to tag the beam crossing corresponding to a given particle.

A toroidal magnet system described in Section 5.4.2 provides the bending field for the muon spectrometer.

5.4 Magnet system

A charged particle moving in a magnetic field experiences a Lorentz force ($\mathbf{F} = q\mathbf{v} \times \mathbf{B}$) that depends on its own charge (q) and velocity ($\mathbf{v}$), as well as the direction and strength of the magnetic field ($\mathbf{B}$). The force causes the particle to have a curved trajectory in the presence of this field, and these principles are used by the ID and muon spectrometer to measure the momentum of charged particles by analyzing properties of their tracks within magnetic fields. A solenoid provides the magnetic field for the ID and a toroidal magnet system provides the field for the muon spectrometer.

5.4.1 Central solenoid

The central solenoid is a single-layer Al-stabilized NbTi conductor coil wound inside an Al support cylinder. It is 5.8 m long and 50 cm thick with an inner radius of 1.23 m. It is cooled to a superconducting temperature of 4.5 K and shares the liquid argon calorimeter vacuum vessel to minimize the material thickness in front of the calorimeter. The nominal current is 7.730 kA which produces a 1.998 T magnetic field at the center of the solenoid. The solenoidal field points in the $+z$ direction and has the effect of primarily bending resulting positively-charged particles from collisions in the $-\phi$ direction.

5.4.2 Toroidal magnets

A barrel and two end-cap toroidal magnets are used to bend charged particles within the volume surrounding the calorimeters. The barrel toroid system in 25.3 m long with inner and outer diameters 9.4 m and 20.1 m. The endcap toroids are each 5.0 m long with inner and outer diameters 1.65 m and 10.7 m. Both use a Al-stabilized Nb/Ti/Cu conductor wound into pancake-shaped coils, and consist of eight coils assembled radially and symmetrically around the beam axis. It takes roughly 5 weeks to cool the barrel toroid down to a superconducting temperature of 4.6 K. The magnetic field strengths produced by the barrel and end-cap toroidal magnets are 0.5 T and 1 T, respectively.

5.5 Trigger system

The crossing rate of proton bunches within the ATLAS detector is 40 MHz for Run 2 of the LHC, and there are roughly 25 proton-proton collision events per bunch crossing. Each event occupies 1.6 MB of storage, so if every event were recorded, the data rate would be 1500 TB/s. It is not feasible to store (and then analyze) this many events, so a trigger system is used to select potentially interesting physics events.

For Run 2, ATLAS uses a two-level trigger system [27]. The level-one (L1) trigger is hardware based, and takes coarse granularity input from the calorimeters and muon spectrometer to reduce the event rate from 40 MHz to roughly 100 kHz. The L1 trigger has approximately 2.5 μs to decide whether or not to pass an event to the second trigger level for further processing. The second trigger level in ATLAS is the high level trigger (HLT) which selects events using software algorithms with full granularity detector information as input. The HLT has about 200 ms to decide if an event should be recorded and the HLT output event rate is approximately 1 kHz.

Part III

The Analysis

Chapter 6

Data and Simulated Samples

6.1 Dataset

The data used in this analysis were recorded with the ATLAS detector in $\sqrt{s} = 13\,\text{TeV}$ pp collisions at the LHC in 2015 and 2016. An official Good Runs List (GRL) is used to select only events from luminosity blocks where all necessary detector subsystems were working as expected. This analysis uses an `AllGood` GRL, which requires the inner tracker, calorimeters, and muon spectrometer to have good data quality, as well as the solenoid and toroidal magnets to be operating at their nominal field strengths. A number of additional quality checks are used to reject single events where data from one or more detector subsystems have been corrupted. These checks look for LAr noise bursts, errors in the TileCal, and events affected by the recovery procedure for single event upsets in the silicon tracking detectors. A final check rejects events where some detector information is missing due to the recovery of a subsystem

after detector busy conditions during a run. A dedicated trigger was implemented for this analysis: `HLT_g25_medium_L1EM22VHI_4j35_0eta490_invm700`. The motivation for this trigger is better understood in the context of the event selection, so that discussion is deferred to Section 8.2. The integrated luminosity collected with this trigger in 2015+2016 and used in this analysis is $12.6\,\text{fb}^{-1}$.

6.2 Simulated samples

Signal and background events are modeled by Monte Carlo (MC) simulations. Events are generated at leading order (LO) with MadGraph5_aMC@NLO [28] v2.3.2, using the NNPDF30_lo_as_0130 [29] parton distribution function set in the five-flavor scheme, and interfaced to Pythia 8.210 [30] [31] for the parton shower and hadronisation. The A14 [32] set of optimized parameters for the underlying event (UE) description using the NNPDF2.3LO [33] PDF set, referred to as the "UE tune," is used. Minimum bias events are generated using Pythia 8.186 [34] with the MSTW2008LO [35] PDF set and the A2 [36] tune. These events are overlaid on the hard-scatter interaction, according to the predicted luminosity profile of the recorded data, to model pile-up contributions from both the same bunch crossing and neighboring bunch crossings (referred to as in- and out-of-time pile-up, respectively)[1]. The response of the ATLAS detector to these events is calculated using the full simulation software based on Geant4 [37]. Full simulation is used because of the difficulties in parametrizing the detector response

[1] A residual re-weighting is applied later to correct for differences between the estimated luminosity profile at the time of simulation, and the actual luminosity profile in data.

to very forward jets.

Signal events are those with a final state containing two b-jets (signal jets) coming from the decay of a Higgs boson of mass 125 GeV, a photon, and two additional jets (labelled "VBF jets"). These events contain contributions from both vector boson fusion and associated VH production mechanisms. Contributions from other production mechanisms, e.g. gluon fusion and associated $t\bar{t}$ production (both in association with a photon), were studied at particle level but found to be negligible in the phase space considered in this analysis. The $H\to b\bar{b}$ decay is modeled with PYTHIA.

Background events containing two b-jets from the decay of a Z boson, a photon, and two additional jets are generated separately for events produced via strong vs. electroweak processes. Events produced via strong processes are generated at order $\alpha_{\mathrm{s}}^2\alpha_{\mathrm{EW}}^3$, while events produced via electroweak processes are generated at order α_{EW}^5. For these samples, the $Z\to b\bar{b}$ decay is modeled with MADGRAPH5_aMC@NLO.

The dominant source of background events in this analysis is from non-resonant multijet production of at least two b-jets, two other jets, and a photon. These events are modeled by requiring the same final state as the signal and $Z+\gamma$ background events, but excluding contributions from diagrams containing on-shell Higgs or Z bosons. Due to the large cross section and low acceptance for this source of background events, it is not feasible to generate a sufficiently large Monte Carlo dataset for use in the final statistical interpretation in this analysis. The events in this sample are used to train a boosted decision tree (BDT), described in Section 8.3, and to optimize the categorization of events in the analysis. A data-driven technique, described in Chapter 9, is used to

estimate the shape and yield of this source of background events for use in the statistical interpretation.

Tables 6.1 and 6.2 summarize the MadGraph5_aMC@NLO syntax and event generation parameters for each of the samples described above.

Sample	Type	Process
HbbjjaSM125	Signal	p p > h a j j
ZbbjjaEWK	Background	p p > z a j j QCD=0, z > b b∼
ZbbjjaQCD	Background	p p > z a j j QED=2 QCD=10, z > b b∼
NonResbbjja	Background	p p > b b∼ j j a $ z h

Table 6.1: Summary of the MadGraph5_aMC@NLO syntax used to generate signal and background processes.

Parameter	Description	Value
lhaid	LHAPDF PDF set ID	263000
ptj	Min. jet p_T	15.0
ptb	Min. b-jet p_T	15.0
pta	Min. photon p_T	15.0
etaj	Max. jet η	-1.0
etab	Max. b-jet η	-1.0
etaa	Max. photon η	3.0
drjj	Min. $\Delta R(jj)$	0.4
drbb	Min. $\Delta R(bb)$	0.4
drbj	Min. $\Delta R(bj)$	0.3
drab	Min. $\Delta R(\gamma b)$	0.4
draj	Min. $\Delta R(\gamma j)$	0.4

Table 6.2: List of MadGraph5_aMC@NLO parameters used during event generation for the MC samples used in this analysis.

Published cross sections are not available for the processes considered here, so the values reported by MadGraph5_aMC@NLO are used. They are shown in Table 6.3. It is important to note that the reported `HbbjjaSM125` cross section in this table is inclusive in Higgs boson decay modes (i.e. it does not include the $H \to b\bar{b}$ branching

ratio). The expected $H \to b\bar{b}$ branching ratio is 57.7% [38] for a 125 GeV Higgs boson.

Sample	x-sec
HbbjjaSM125	66.0 fb
ZbbjjaEWK	58.2 fb
ZbbjjaQCD	2.67 pb
NonResbbjja	767. pb

Table 6.3: Generator reported cross sections for the MC samples used in this analysis.

Chapter 7

Object Selection

This analysis selects events containing photons, calorimeter jets, track jets, and b-tagged jets. These objects are identified using the standard ATLAS procedures, which are described below.

7.1 Calorimeter jets

Calorimeter jets are reconstructed from topologically connected clusters of calorimeter cells (topoclusters), seeded by cells with energy deposits significantly above the noise threshold. These cells are used as inputs to the anti-k_T [39] clustering algorithm with a distance parameter $R = 0.4$. These topoclusters are calibrated to correctly measure the energy deposited by electromagnetic showers, and thus corrections must be applied for hadronic showers. There are roughly four broad stages of corrections applied [40]. First, the effect of pile-up on jet calibration [41] is partially mitigated using an area-based subtraction method, which uses the average energy density in the $\eta \times \phi$

plane, as well as the area of the jet. Two additional corrections account for both in- and out-of-time pile-up, and depend on the number of primary vertices in the event and the average number of interactions per bunch crossing. Second, energy- and $|\eta|$-dependent scale factors are derived from Monte Carlo simulations and collision and test-beam data to correct for calorimeter non-compensation and detector inhomogeneity. Third, a number of global sequential corrections (GSCs) are applied to correct for the dependence of the jet energy scale on the longitudinal and transverse structure of the hadronic shower. Finally, *in situ* corrections, which have been derived from comparisons between γ/Z+jet and multijet processes in data and MC simulations, are applied to account for residual differences in the jet p_T measurement.

Jets selected in this analysis must have $p_T > 40\,\text{GeV}$ and $|\eta| < 4.5$. A number of quality criteria are used to reject "fake jets" coming from background processes rather than high energy objects produced in a collision. These quality criteria are based on the following variables, for which the rigorous definitions are given in Ref. [42]:

- $\langle Q \rangle$: the energy-squared weighted average of the pulse quality of the calorimeter cells in the jet
- $f_{\text{Q}}^{\text{LAr}}$: the fraction of energy in the jet's LAr cells with poor signal shape quality
- $f_{\text{Q}}^{\text{HEC}}$: the fraction of energy in the jet's HEC cells with poor signal shape quality
- E_{neg}: the sum of the energy of all cells with negative energy
- f_{EM}: the fraction of the jet's energy deposited in the EM calorimeter

- f_{HEC}: the fraction of the jet's energy deposited in the HEC
- f_{max}: the fraction of the jet's energy in any single calorimeter layer
- f_{ch}: the ratio of the $\sum p_T$ of charged tracks associated with the jet to the jet's p_T

Jets identified as fake jets in this analysis are labeled BadLoose and satisfy at least one of the following criteria:

- $f_{\mathrm{HEC}} > 0.5$ and $|f_{\mathrm{Q}}^{\mathrm{HEC}}| > 0.5$ and $\langle Q \rangle > 0.8$
- $|E_{\mathrm{neg}}| > 60\,\mathrm{GeV}$
- $f_{\mathrm{EM}} > 0.95$ and $f_{\mathrm{Q}}^{\mathrm{LAr}} > 0.8$ and $\langle Q \rangle > 0.8$ and $|\eta| < 2.8$
- $f_{\mathrm{max}} > 0.99$ and $|\eta| < 2$
- $f_{\mathrm{EM}} < 0.05$ and $f_{\mathrm{ch}} < 0.05$ and $|\eta| < 2$
- $f_{\mathrm{EM}} < 0.05$ and $|\eta| \geq 2$

If none of these conditions are met, the jet is considered a "good jet" likely to be associated with high energy objects produced in a collision.

A final selection requirement, the jet-vertex-tagger (JVT) [43], is introduced to help reject spurious jets originating from pile-up. The JVT uses two track-related variables to calculate the likelihood that a given jet originated from the hard scatter interaction rather than from pile-up. These two variables are corrJVF and $R_{\mathrm{p_T}}$. corrJVF measures the scalar sum of the p_T of tracks associated with the jet and originating from the primary vertex relative to the scalar sum of the p_T of all tracks associated with the

jet, but includes a correction such that it is independent of the number of reconstructed vertices in the event. $R_{\mathrm{p_T}}$ measures the scalar sum of the p_T of tracks associated with the jet and originating from the primary vertex relative to the total jet p_T. Both of these variables are peaked at 0 for pile-up jets, since no or only little p_T from tracks from the hard-scatter vertex is expected. This analysis uses the "medium" JVT working point, which has a signal jet efficiency of 90% and a pile-up fake rate of 1.0% for jets with p_T in the range 20-60 GeV, as measured in simulated dijet events and validated in $Z(\rightarrow \mu\mu)$+jets events in data [43]. The JVT is not applied to jets with $p_T \geq 60\,\mathrm{GeV}$ nor to jets with $|\eta| > 2.4$, beyond the tracker coverage.

7.2 b-tagged jets

Calorimeter jets that are likely to have originated from b-quarks are identified using a multivariate discriminant [44] [45] that takes as inputs the jet kinematics (p_T and $|\eta|$), as well as variables from three algorithms that use tracking information to look for signatures consistent with B-hadron decays. b-tagging can only be applied on jets within the tracking volume ($|\eta| < 2.5$).

Due to the relatively long lifetime of hadrons containing b-quarks ($\approx$1.5 ps), these hadrons travel a non-negligible distance ($\approx$5 mm) from the primary vertex before decaying. Impact-parameter based b-tagging algorithms look for tracks with large impact parameters in the transverse and longitudinal directions (with respect to the primary vertex) that give evidence for a secondary displaced vertex. Secondary vertex

finding algorithms explicitly look for displaced secondary vertices by testing all track pairs within a jet for a two-track vertex hypothesis. Vertices consistent with photon conversions and detector interactions are not kept. Candidate vertices are then fitted using all tracks from that vertex, and properties of those vertices are checked for consistency with B-hadron decays. JetFitter is a multi-vertex reconstruction algorithm that attempts to reconstruct the full B-hadron decay chain.

Outputs from these three classes of algorithms are used as inputs to MV2, which is a boosted decision tree algorithm that further improves the discrimination power between b-, c-, and light-jets. MV2 is trained and cross-validated using independent samples of $t\bar{t}$ events. Three variants of MV2 were released, with the difference being the relative fraction of c- and light-jets used in the training. This analysis uses the MV2c10 variant, which contains 7% c-jets and 93% light-jets. A 77% b-jet efficiency working point was found to be optimal with respect to the trade-off between signal acceptance and background rejection in this analysis. This working point has a c-jet rejection factor of 6 and a light-jet rejection factor of 134. The rejection factor is defined as the multiplicative inverse of the efficiency ($\frac{1}{\text{eff.}}$).

7.2.1 b-jet energy corrections

Two additional energy corrections are applied to b-tagged jets beyond the standard jet calibration procedure [46]. The muon-in-jet correction searches for the closest medium quality muon with $p_T > 4\,\text{GeV}$ and within $\Delta R < 0.4$ of the jet. If such a muon is found, the 4-vector of the reconstructed muon is added to the jet's 4-vector,

and the 4-vector associated with the energy deposited by the muon in the calorimeter is subtracted from the jet's 4-vector. This correction affects only the ≈4% of jets matched to muons meeting these criterion.

A second correction, PtReco, is applied to all b-tagged jets. PtReco attempts to correct for several components of the jet energy that are not accounted for in jet reconstruction, including neutrinos, energy clusters outside of the $\Delta R = 0.4$ cone, and final state radiation. The PtReco correction is derived as a function of the reconstructed jet p_T and in categories based on the leptonic content of the jet by matching reconstructed jets with truth-level jets in MC samples. Within each jet category, the mean of the distribution of the ratio of the truth jet p_T to the reconstructed jet p_T and interpolated across p_T bins is used as the correction.

The overall effect of these two corrections is to improve the resolution of the di-b-jet invariant mass in $H \to b\bar{b}$ decays, as shown in Figure 7.1

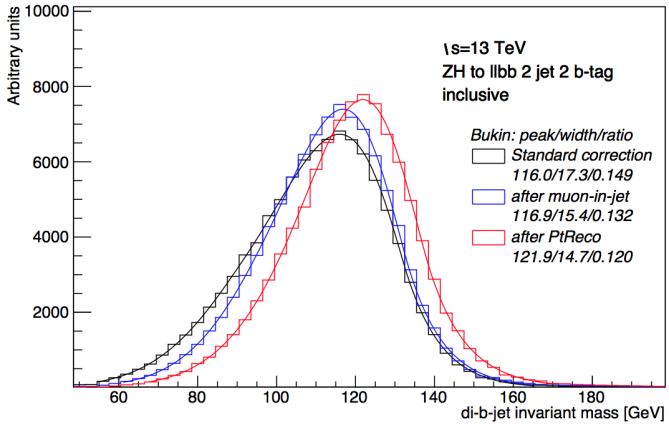


Figure 7.1: Effect of subsequent muon-in-jet and PtReco b-jet energy correction on the $m_{b\bar{b}}$ peak in $ZH \to llbb$ events.

7.3 Track jets

Track jets are reconstructed using tracks [47], rather than calorimeter clusters, as inputs to the anti-k_T algorithm with a distance parameter $R = 0.2$. There must be at least two constituent tracks associated with each track jet. Track jets used in this analysis must have $p_T > 7\,\text{GeV}$ and $|\eta| < 2.5$. Though they are not used for event selection, track jets are used to calculate H_T^{soft}(defined later), which is used in a BDT to discriminate between signal and background events.

7.4 Photons

Photon reconstruction [48] is seeded by clusters of EM calorimeter cells with $p_T > 2.5\,\text{GeV}$ using a sliding window algorithm. Tracks from the ID are then extrapolated to the calorimeter and checked to determine if they can be loosely matched to the calorimeter clusters. The existence (or lack thereof) and properties of these matched tracks are used to decide if a seed cluster corresponds to an unconverted photon, a converted photon, or an electron. From MC simulations, 96% of prompt photons with $E_T > 25\,\text{GeV}$ are expected to be reconstructed as photon candidates, with the other 4% are incorrectly reconstructed as electron candidates.

A number of variables which describe the shower shape in the EM calorimeter and shower energy leakage into the hadronic calorimeter are used for photon identification. These variables have discriminating power because prompt photons typically deposit their energy in a narrow region of the EM calorimeter, whereas photons from

jets have a wider shower shape in the EM calorimeter with additional energy deposits in the hadronic calorimeter due to the presence of extra hadrons near the photon. Pairs of photons produced from decays of neutral pions can be distinguished from prompt photons in the finely segmented strip layer of the EM calorimeter, as the former exhibit a characteristic double peak structure in this first layer. This analysis selects photons passing the tight (rather than loose) identification criteria, which has a strong background rejection factor of ≈5000 and a signal efficiency of 85% for photons with $E_T > 40\,\text{GeV}$. Photons selected in this analysis must have $E_T > 30\,\text{GeV}$ and $|\eta| < 2.37$ (with $1.37 \leq |\eta| \leq 1.52$ excluded due to the transition between the barrel and end-cap sections of the electromagnetic calorimeter).

Photons selected in this analysis are required to pass an additional isolation requirement in the calorimeter. The isolation variable, topoetcone40, is computed by summing all of the topo clusters within a cone of $R = 0.4$ around the photon, but subtracting the central core cells associated with the photon. This quantity is required to be less than the sum of 2.2% of the photon's E_T and an additional offset of 2.45 GeV. In ATLAS, this isolation requirement is called `FixedCutTightCaloOnlyIso`.

7.5 Overlap removal

An overlap removal procedure is used to resolve ambiguities in cases where an object in the ATLAS detector is reconstructed as both a photon and jet. This procedure uses the angular separation variable, $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$. If $\Delta R < 0.2$, the photon and

jet are aligned, and it is more likely that a real photon was also reconstructed as a jet. In this case, the jet is removed. If $0.2 \leq \Delta R < 0.4$, the the object is more likely to be a jet, so the photon is removed.

Chapter 8

Event Selection and Categorization

8.1 Multivariate analysis pre-selection

A multivariate analysis (MVA) is used to improve the discrimination power between signal and background events in this analysis. This section describes the selection of events used as inputs to the MVA. The final state of the signal process in this analysis contains at least four jets and a photon. Two of these jets are b-jets from the decay of the Higgs boson, and another two are forward jets from the VBF production process. The photon may be radiated from one of the virtual $W^{\pm}$ bosons involved in the VBF production mechanism or from an initial or final state quark. Consistent with this signal signature, events selected in this analysis must pass a dedicated trigger (described in Section 8.2) and have at least four jets and a photon. Two of these jets must lie in $|\eta| < 2.5$, where the b-tagging algorithms can be applied. The remaining preselection cuts for this analysis require that we categorize jets into signal jets and VBF

jets. Signal jets are selected by first sorting the selected jets with decreasing MV2c10 weights (i.e. decreasing confidence of being b-tagged), and then selecting the first two MV2c10-sorted jets in $|\eta| < 2.5$. Of the remaining jets, the two that yield the largest invariant mass are chosen as the VBF jets. The p_T-sorted b-jets (VBF-jets) are denoted b_1 and b_2 (j_1 and j_2), where the jet with higher p_T is labeled with "1." The signal jets must be b-tagged using the 77% MV2c10 working point, and the $b\bar{b}$ system is required to have $p_T^{b\bar{b}} > 80\,\text{GeV}$. This latter requirement is used to sculpt the $m_{b\bar{b}}$ distribution of the non-resonant background, which is used in the final statistical interpretation in this analysis, and is discussed further in Section 9.2. Finally, the invariant mass of the VBF jets (m_{jj}) must be greater than 800 GeV, which selects events consistent with VBF signatures. These event selection requirements are summarized in Table 8.1.

Selection	Requirement
Derivation	HIGG5D3
Trigger	`HLT_g25_medium_L1EM22VHI_4j35_0eta490_invm700`
Event quality	pass GRL / no Tile, LAr, SCT, or Core error
Primary Vertex	At least one primary vertex
Photon	≥ 1 photon
Jets	≥ 4 jets ($p_T > 40$ GeV, $\mid \eta \mid < 4.5$)
Central jets	≥ 2 jets in $\mid \eta \mid < 2.5$
Signal jets	two central jets with highest MV2c10 weights
VBF jets	pair of non-signal jets with highest invariant mass
b-jets	2 b-tagged signal jets (MV2c10 at 77% working point)
m_{jj}	$m_{jj} > 800$ GeV
$p_T^{b\bar{b}}$	$p_T^{b\bar{b}} > 80$ GeV

Table 8.1: Summary of the full event selection for the MVA analysis.

8.2 Trigger

A dedicated trigger for this analysis has been implemented in the ATLAS trigger menu, which makes use of combined trigger chains to maximize the signal efficiency while maintaining a low trigger rate. The trigger chain is

`HLT_g25_medium_L1EM22VHI_4j35_0eta490_invm700`.

The level one hardware-based component of this trigger requires an isolated electromagnetic object with $p_T > 22\,\text{GeV}$. Events selected at L1 are used by the high level trigger to perform selections that are as close as possible to the offline selections in physics analyses. The trigger for this analysis requires at HLT level at least one medium-quality photon with $p_T > 25\,\text{GeV}$ and at least four jets with $p_T > 35\,\text{GeV}$ in $|\eta| < 4.9$. Additionally, the maximum invariant mass over all pairs of HLT jets must be greater than 700 GeV.

The performance of this trigger was studied assuming that it could be factorized into three independent components, with the efficiency of each component measured as a function of the most sensitive offline variable. The factorized components and their corresponding offline variables are summarized in Table 8.2. The factorization procedure

Component	Offline variable	Offline threshold
`HLT_g25_medium_L1EM22VHI`	p_T^{γ}	30 GeV
`4j35_0eta490`	$p_T^{j_4}$	40 GeV
`invm700`	m_{jj}	800 GeV

Table 8.2: Decomposition of the combined trigger chain used in this analysis.

assumes that the efficiency of the combined trigger chain can be computed as the prod-

uct of the individual efficiency of each factorized component. This property is referred to as closure, and a closure test was performed using MC samples. The procedures for measuring the efficiency of each component, as well as the closure test, are detailed in the following subsections.

8.2.1 `HLT_g25_medium_L1EM22VHI` efficiency

The efficiency of the photon portion in the combined trigger chain depends on both the L1 and the HLT photon triggers, and is measured as a function of the p_T of the selected photon (p_T^γ). Efficiencies for both of these components were measured by a dedicated team in the ATLAS experiment, and the results are available in [49] and [50]. The efficiency of the HLT component was measured in data using events passing a L1 EM trigger with a substantially lower threshold (`L1_EM7`) without an isolation requirement, and the results are shown in Figure 8.1. `HLT_g25_medium` is fully efficient for photons

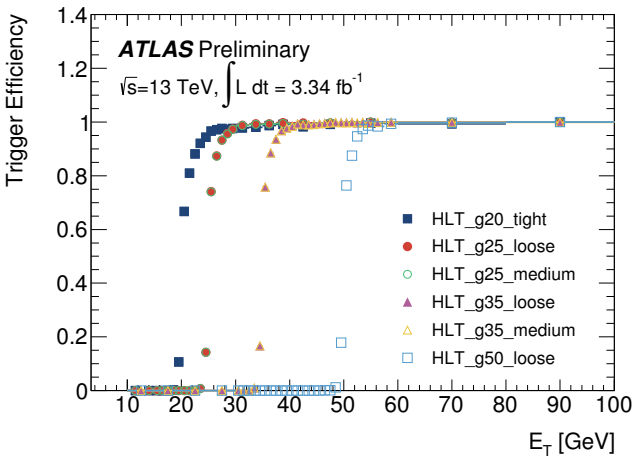


Figure 8.1: Efficiencies of single photon triggers measured using events passing `L1_EM7`.

with $p_T > 30\,\text{GeV}$, independent of the HLT photon identification criteria (i.e. loose vs.

medium).

Figure 8.2 shows the `L1_EM22VHI` trigger turn on curve for both MC and data. A systematic uncertainty is assigned to account for the small efficiency difference be-

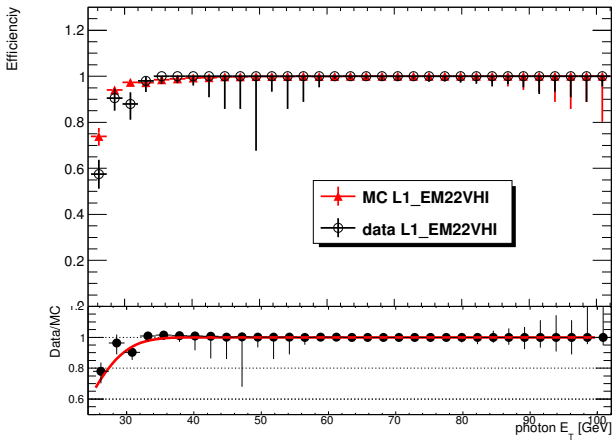


Figure 8.2: Efficiencies of L1_EM22VHI trigger in data and MC

tween data and MC for photons with $p_T < 40\,\text{GeV}$. The uncertainty is applied as an event-by-event scale factor (SF) to the MC samples. The SF is calculated by fitting the ratio of data to MC in Figure 8.2 (the fit is shown by the red line in the lower panel), and evaluating the fit at the value corresponding to the photon p_T in that event. It is also important to point out that the L1 photon component does not yet plateau at 30 GeV, so this corresponds to a tighter requirement than at HLT. The efficiency of the individual L1 and HLT photon trigger components, as well as the total photon trigger efficiency is shown in Figure 8.3 for MC signal events that otherwise satisfy the pre-selection cuts in this analysis.

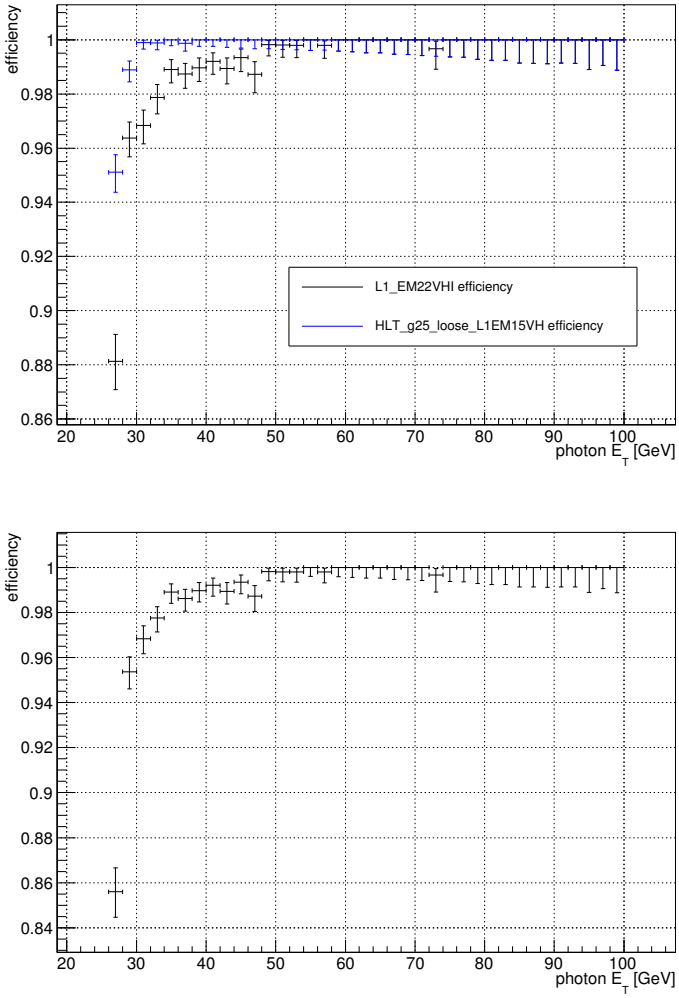


Figure 8.3: The individual (top) and combined (bottom) turn-on curves of `L1_EM22VHI` and `HLT_g25_loose_L1EM15VH` in the signal Monte Carlo sample.

8.2.2 `4j35_0eta490` efficiency

The jet portion of the combined trigger chain requires at least four HLT jets with $p_T > 35\,\text{GeV}$ in $|\eta| < 4.9$. The efficiency of this component is measured as a function of the fourth-leading p_T-ordered selected jet in the event ($p_T^{j_4}$). The events used in the jet efficiency measurement are selected in both data and MC using a set of reference photon triggers. MC events are selected using the `HLT_g25_loose_L1EM15VH` trigger, which is slightly looser than the photon trigger portion of the combined trigger chain in the analysis. In data, a set of pre-scaled single photon triggers with varying L1 and HLT requirements is used in order to maximize the number of events available for study. The L1 photon p_T requirement in these reference triggers ranges from 7-15 GeV, while at HLT the range is 10-120 GeV. The reference triggers have no isolation requirements. The jet efficiency is defined as the number of events passing the event selection and reference trigger and having at least four HLT jets with $p_T > 35\,\text{GeV}$, divided by the number of events passing just the event selection and reference trigger. To avoid bias from the reference trigger, the selected photon in the event is matched to the HLT photon within $\Delta R < 0.07$, and the HLT jets must be separated from the HLT photon by $\Delta R > 0.4$. Figure 8.4 shows the jet component trigger efficiency in both signal MC events and data as a function of $p_T^{j_4}$. The jet component of the combined trigger chain is fully efficient for $p_T^{j_4} > 46\,\text{GeV}$. The data and MC are in good agreement over the entire p_T range, except for a small difference in the 2$^{\text{nd}}$ p_T bin of 42-44 GeV. A systematic uncertainty of 2% is assigned to cover this difference.

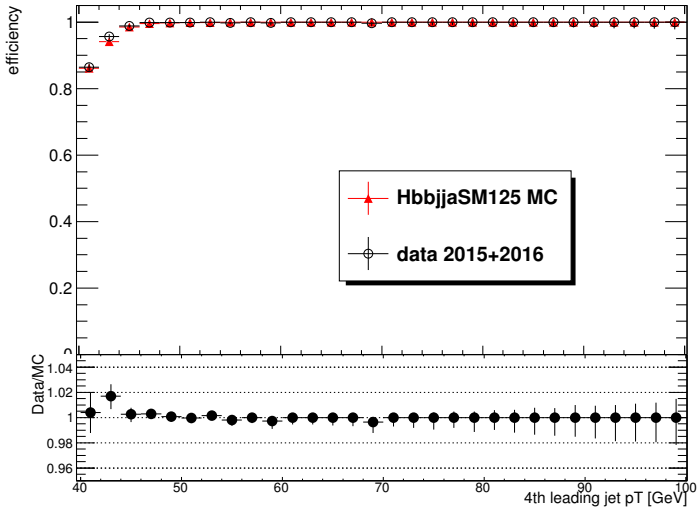


Figure 8.4: `4j35_0eta490` efficiency measured in the photon stream as a function of $p_T^{j_4}$.

8.2.3 `invm700` efficiency

The last portion of the combined trigger chain is the VBF component, and it requires the maximum invariant mass over all pairs of HLT jets to be at least 700 GeV. The efficiency of the VBF component is measured as a function of the invariant mass of the selected VBF jets in the event. A similar method is used here as was used to measure the jet portion of the combined trigger chain: the number of events passing the event selection and reference photon trigger and HLT invariant mass requirement is divided by the number of events passing just the event selection and the reference photon trigger. The HLT jets used in the invariant mass calculation are required to be separated from the photon trigger object (which itself is required to be matched to the selected photon). The VBF component efficiency is shown in Figure 8.5. The VBF component of the combined trigger chain is fully efficient for events with $m_{jj} > 800$ GeV,

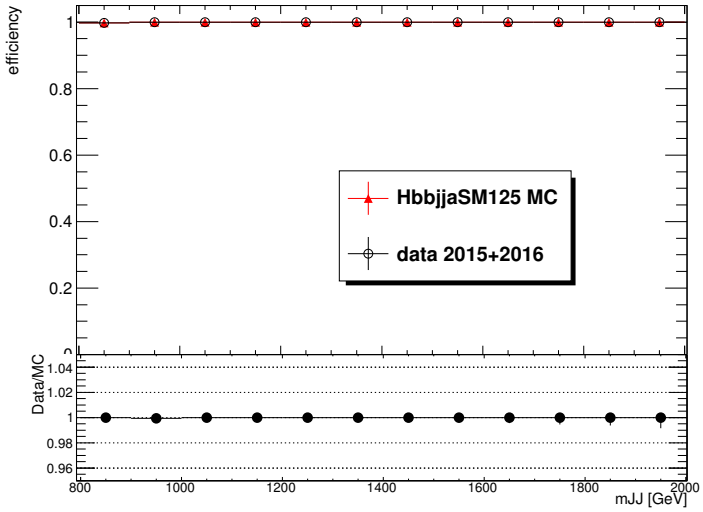


Figure 8.5: `invm700` efficiency measured in photon stream as a function of m_{jj}.

and is well modeled by the MC simulation.

8.2.4 Closure test

The validity of the factorization procedure for computing the combined trigger chain efficiency as the product of the individual photon, jet, and VBF components was checked using the signal MC sample. Initially, a small non-closure was observed and subsequently traced to a correlation between the photon and jet components of the trigger.

A correlation between the photon and jet trigger components exists because a photon with sufficient p_T may compose the four jets that trigger the `4j35_0eta490` component. To correct for this correlation in the closure test, a correlation factor was calculated and is defined as the number of events passing the event selection and

combined trigger chain and containing four isolated (from the photon trigger object) HLT jets with $p_T > 35\,\text{GeV}$, divided by the number of events passing just the event selection and combined trigger chain. Figure 8.6 shows the correlation factor as a function of the p_T of the fourth leading selected jet in both data and MC. The correlation factor is well modeled by the MC simulation, except for a small difference in the first p_T bin of 40-42 GeV.

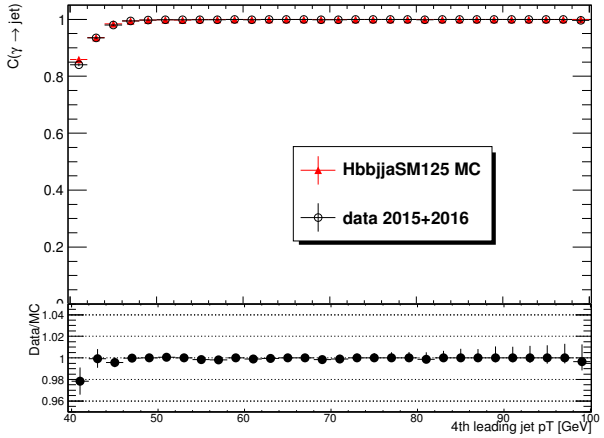


Figure 8.6: Jet-photon correlation factor measured as a function of $p_T^{j_4}$. The inefficiency in the turn-on region shows the probability of a photon triggering the jet trigger component.

The formula for the closure test is defined in Equation 8.1 using the following components:

- The photon component efficiency: $\varepsilon(\text{A})$
- The jet component efficiency: $\varepsilon(\text{B})$
- The VBF component efficiency: $\varepsilon(\text{C})$
- The photon-jet correlation factor: $C(\gamma \to j)$.

$$\text{closure} = \frac{\varepsilon_{\text{combined}}}{\varepsilon_{\text{factorized}}} = \frac{\varepsilon(\text{A} \cap \text{B} \cap \text{C})}{\frac{\varepsilon(\text{A})\varepsilon(\text{B})\varepsilon(\text{C})}{C(\gamma \to j)}} \tag{8.1}$$

The results of the closure test are shown in Figure 8.7 as a function of p_T^{γ}, $p_T^{j_4}$, and m_{jj}. In all cases, the closure is consistent with unity, and thus the factorization procedure is valid in this analysis.

8.3 Multivariate analysis

To enhance the sensitivity of this analysis, a multivariate discriminant was trained using 11 input variables that were shown to have separation power between the signal and the non-resonant background. These variables were used as inputs to a boosted decision tree, which exploits correlations among these variables to achieve an increased discrimination power between signal and background events beyond what can be achieved using a simple cut-based alternative. The BDT was implemented using the Toolkit for Multivariate Analysis in ROOT (TMVA) [51], and was trained using the `HbbjjaSM125` (signal) and `NonResbbjja` (background) MC samples. The variable with the most discriminating power, $m_{b\bar{b}}$, was excluded from this training because the final background estimation and signal extraction rely on fits to the $m_{b\bar{b}}$ distribution. Additionally, to avoid any potential bias in those fits, the chosen input variables have minimal correlation with $m_{b\bar{b}}$. The 11 input variables, along with a brief explanation of the origin of their separation power, are listed below:

- $\Delta R_{j,\gamma}$: angular separation between each jet and the photon, ($j \in \{j_1, j_2, b_1, b_2\}$)—the photon is likely to be more separated from the selected jets in

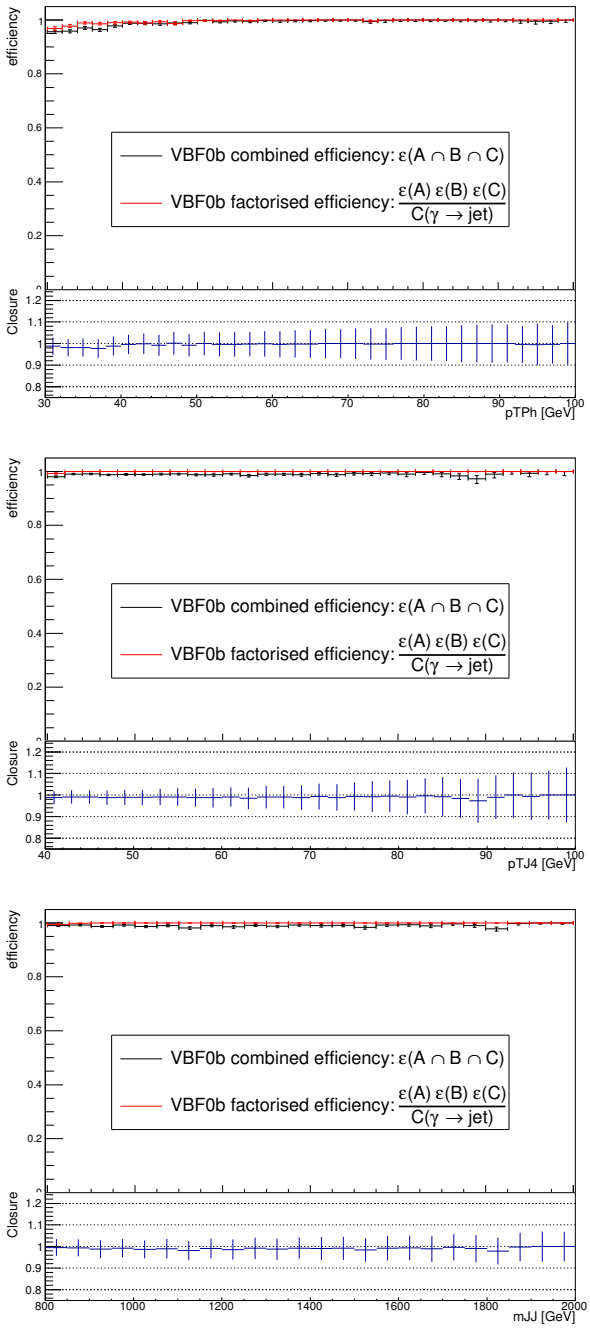


Figure 8.7: Comparison of the combined efficiency and the factorized efficiency of the combined trigger chain as a function of p_T^{γ} (top), $p_T^{j_4}$ (middle), and m_{jj} (bottom), as measured in the signal MC sample.

signal events, where it can be radiated from the internal $W^{\pm}$ boson rather than one of the initial or final state quarks.

- m_{jj}: invariant mass of the VBF jet pair—larger values of m_{jj} are characteristic of VBF signatures.

- $\Delta\eta_{jj}$: η separation between the VBF jets—larger values of $\Delta\eta_{jj}$ are characteristic of VBF signatures.

- j_i width: calorimeter width of each VBF jet ($i = 1, 2$), calculated as the p_T-weighted average angular separation of each jet constituent from the jet axis —gluon jets are generally wider than quark jets. The background contains a mixture of quark and gluon jets, whereas the signal contains only quark jets.

- $p_T^{\text{balance}} = \frac{(\boldsymbol{p}^{b_1}+\boldsymbol{p}^{b_2}+\boldsymbol{p}^{j_1}+\boldsymbol{p}^{j_2}+\boldsymbol{p}^{\gamma})_T}{p_T^{b_1}+p_T^{b_2}+p_T^{j_1}+p_T^{j_2}+p_T^{\gamma}}$: the balance of transverse momentum among the selected final state objects ($\boldsymbol{p}^i$ is the momentum 3-vector for object i)—the transverse momentum in signal events tends to be more balanced among the selected final state objects than in background events.

- centrality$(\gamma) = |\frac{y_\gamma - \frac{y_{j_1}+y_{j_2}}{2}}{y_{j_1}-y_{j_2}}|$: centrality of the photon with respect to the VBF jets—in signal events, the photon is likely to be more centered between the VBF jets than in background events.

- H_T^{soft}: the scalar sum of the p_T of selected track jets. The track jets must be separated from both the photon and VBF jets by $\Delta R > 0.4$, and must not lie in an elliptical region (in the $\eta - \phi$ plane) between the b-jets. This ellipse is

centered on the midpoint between the two b-jets, with a major axis of length $\Delta R(bb) + 0.8$ aligned along the direction connecting the b-jets and minor axis of length 0.8.—hadronic activity between the VBF jets is suppressed in signal events due to the color flow in the event.

Prior to the training, the modeling of the background process for each of these variables was checked in the $m_{b\bar{b}}$ sideband region of this analysis ($m_{b\bar{b}} < 100\,\text{GeV}$ and $m_{b\bar{b}} > 150\,\text{GeV}$), which has a negligible Higgs signal contamination. During these checks, a mis-modeling of the $\Delta\eta_{jj}$ distribution in data was observed, and the treatment of this mis-modeling is discussed in the next section.

8.3.1 $\Delta\eta_{jj}$ reweighting

The left plot in Figure 8.8 shows that the `NonResbbjja` MC sample mis-models the $\Delta\eta_{jj}$ distribution in data. We performed a re-weighting of the $\Delta\eta_{jj}$ distribution in MC, by fitting the ratio of the predicted non-resonant background to the data for the $\Delta\eta_{jj}$ distribution in the $m_{b\bar{b}}$ side-band region. This ratio and the linear fit are both shown in Figure 8.9.

The right plot in Figure 8.8 shows good agreement between data and MC after the re-weighting has been applied. Figure 8.10 shows the background rejection vs. signal efficiency (ROC curve) of the BDT before and after the $\Delta\eta_{jj}$ re-weighting has been applied. With the re-weighting, we see a stronger background rejection for the same signal efficiency (i.e. a boost in BDT performance). The modeling of the other BDT input variables was checked with this re-weighting applied, and is shown in the

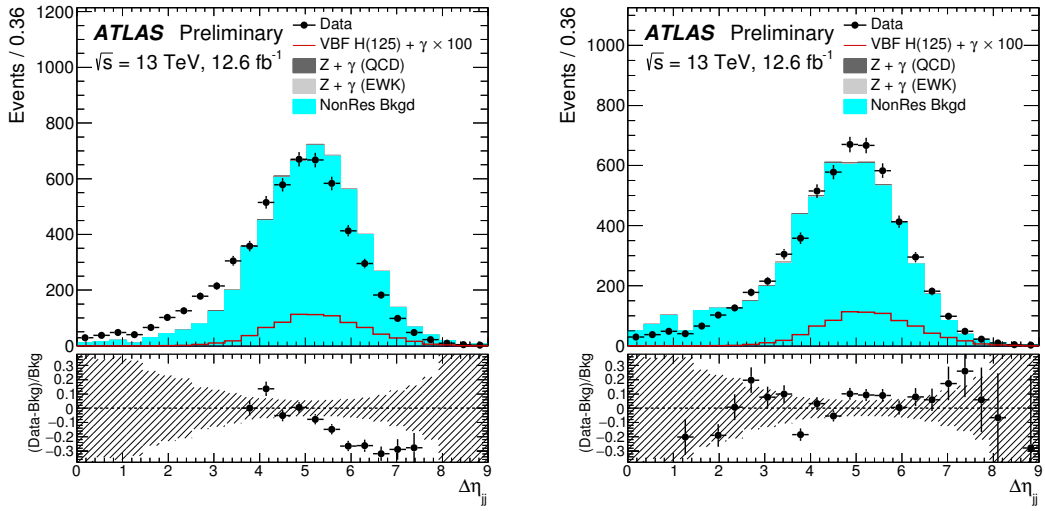


Figure 8.8: Comparison of the $\Delta\eta_{\mathrm{jj}}$ distributions in signal, background, and data before (left) and after (right) reweighting in the $m_{b\bar{b}}$ side-band region. The signal distributions are scaled by a factor of 100. The shaded band in the lower panel shows the statistical uncertainty on the simulated background. Points in the ratio outside the shown range are not displayed.

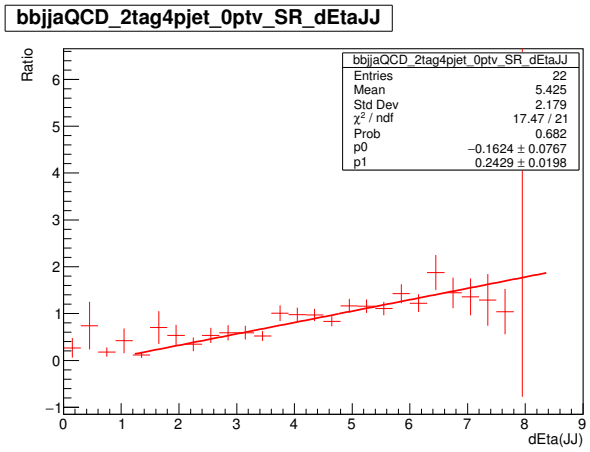


Figure 8.9: Linear fit to the ratio of the predicted non-resonant background to the data for the $\Delta\eta_{jj}$ distribution in the $m_{b\bar{b}}$ side band region.

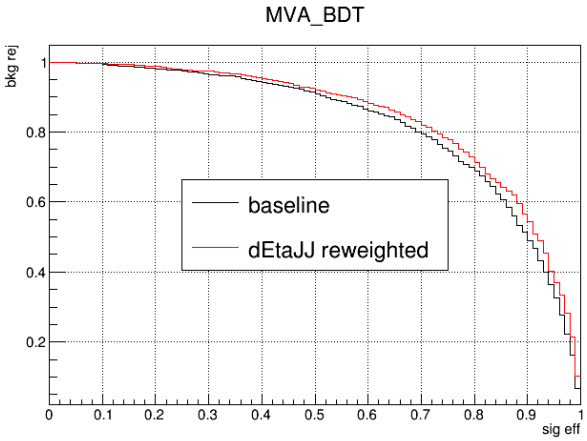


Figure 8.10: Background rejection versus signal efficiency (ROC) plot before (black) and after (red) applying the $\Delta\eta_{jj}$ re-weighting correction to the non-resonant MC background sample.

next section.

8.3.2 Data-MC comparison of BDT input variables after $\Delta\eta_{jj}$ re-weighting

Figures 8.11, 8.12, and 8.13 show that the remaining input variables are reasonably well modeled by the MC. These data vs. MC comparisons are shown with the $\Delta\eta_{jj}$ re-weighting applied. In each of these figures, the shaded band in the lower panel is the relative statistical uncertainty of the simulated background. It should be emphasized that since we use a data-driven technique to estimate non-resonant background contribution for the final statistical analysis, any mis-modelings of the non-resonant background do not affect the integrity of the results presented in Chapter 12. Mis-modelings do, however, impact the BDT training, and thus the sensitivity of the analysis.

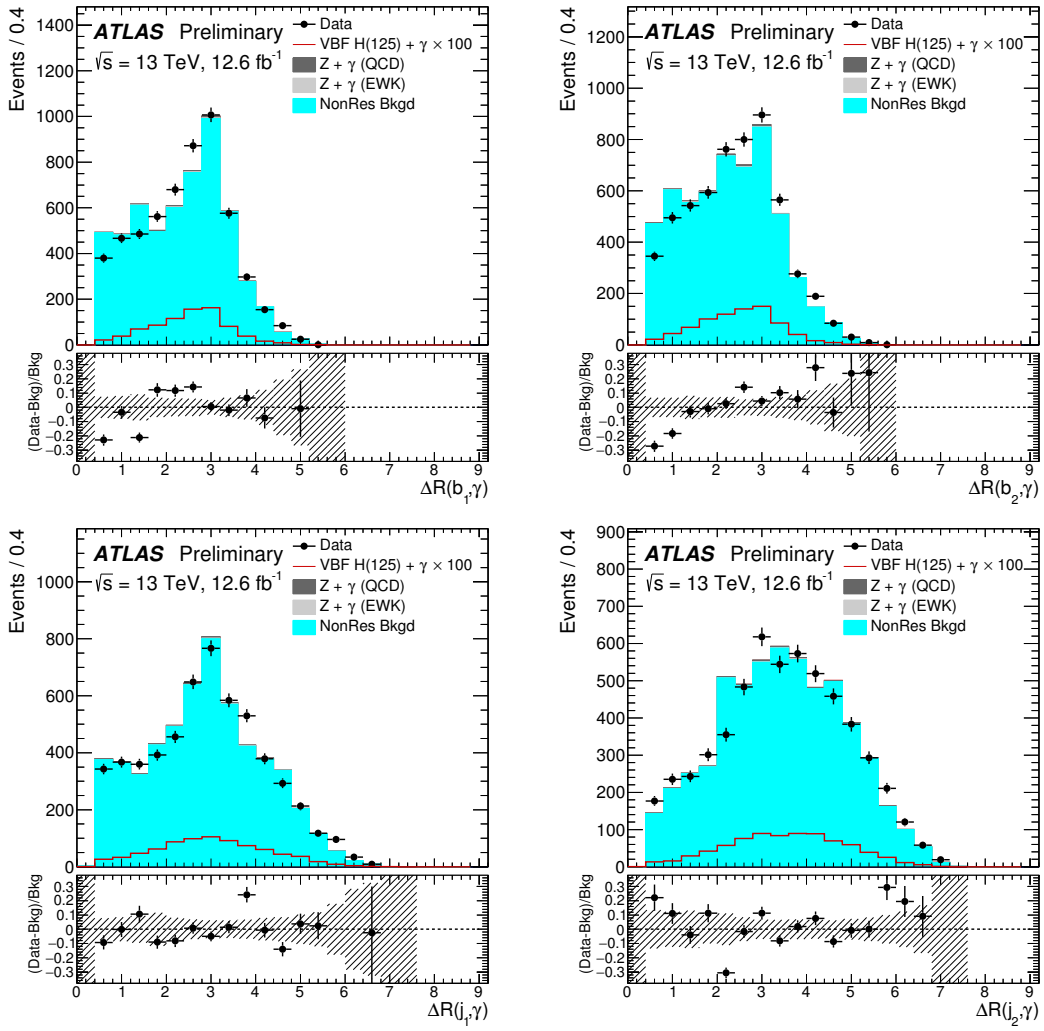


Figure 8.11: Distributions of the input variables for the BDT training, for signal, background, and data after $\Delta\eta_{jj}$ reweighting. The $\Delta R(\text{jet}, \gamma)$ variables represent the distances in (η, ϕ) between the jets and photon in the event. The signal distributions are scaled by a factor of 100. The shaded band in the lower panel shows the statistical uncertainty on the simulated background. Points in the ratio outside the shown range are not displayed.

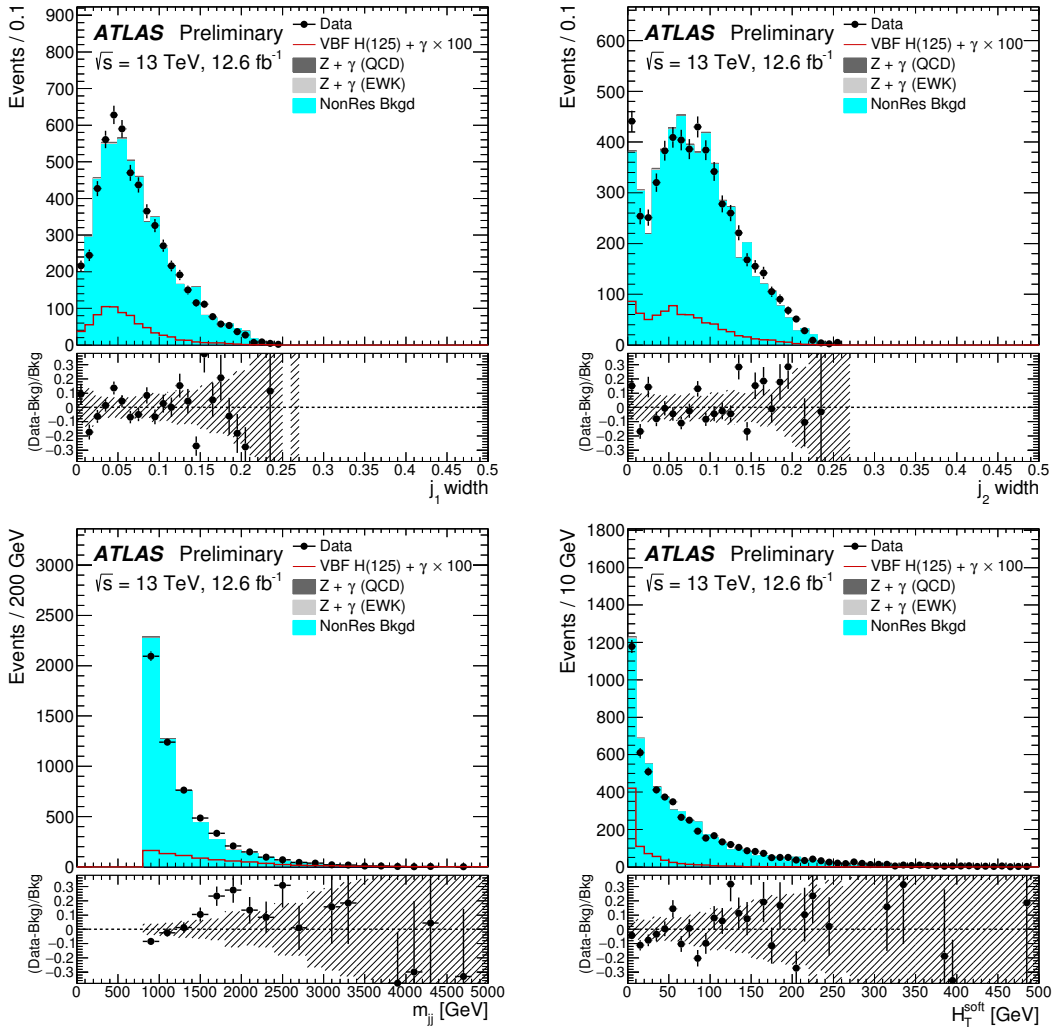


Figure 8.12: Distributions of the input variables for the BDT training, for signal, background, and data after $\Delta\eta_{jj}$ reweighting. The jets labelled j_1 and j_2 are the VBF jets. The signal distributions are scaled by a factor of 100. The shaded band in the lower panel shows the statistical uncertainty on the simulated background. Points in the ratio outside the shown range are not displayed.

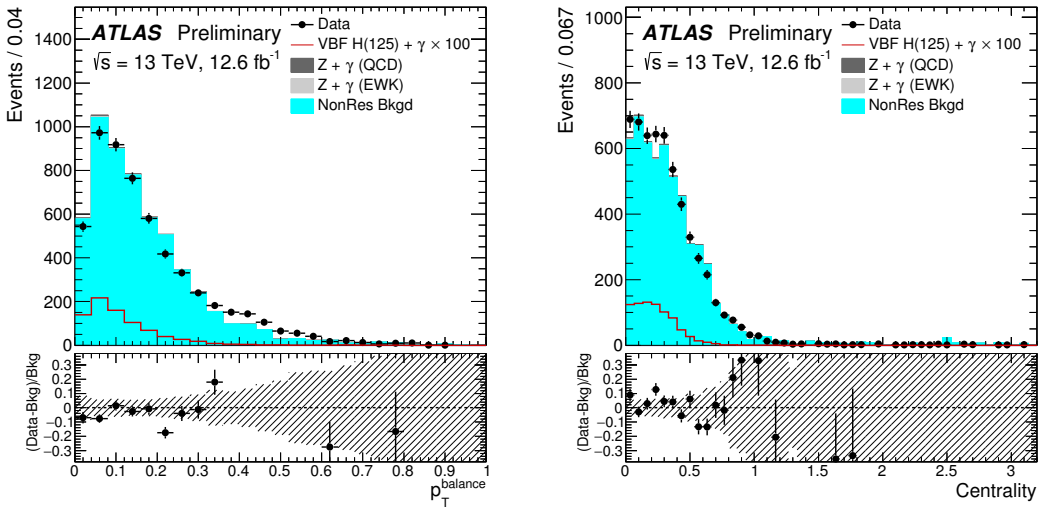


Figure 8.13: Distributions of the input variables for the BDT training, for signal, background, and data after $\Delta\eta_{jj}$ reweighting. The shaded band in the lower panel shows the statistical uncertainty on the simulated background. The signal distributions are scaled by a factor of 100. Points in the ratio outside the shown range are not displayed.

8.3.3 Event categorization

Figure 8.14 shows the expected distribution of BDT weights (w_{BDT}) for both signal (blue) and background (red) events. Signal events are more likely than background events to have higher values of w_{BDT}. This analysis uses the BDT to enhance its sensitivity by dividing events into different categories based on w_{BDT} for the event. The category with the largest expected signal fraction is denoted "High-BDT" and includes events with $w_{\mathrm{BDT}} > 0.1$. This cut value was chosen to maximize the sensitivity of the High-BDT region, while maintaining a sufficient number of background events to perform the data-driven background estimate described in Chapter 9. Additional optimization studies showed that the analysis sensitivity could be slightly improved by

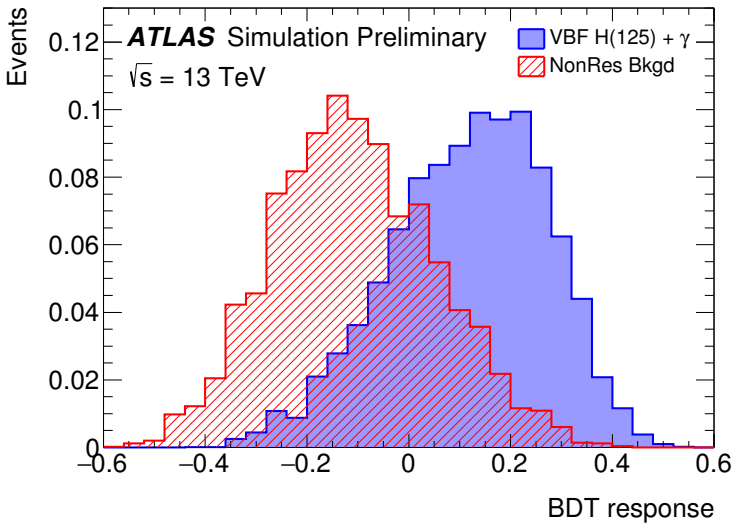


Figure 8.14: Distribution of the BDT output response in simulated $H\gamma jj$ signal events and non-resonant $bb\gamma jj$ background events. The distributions are normalized to unit area.

	HbbjjaSM125	NonResbbjja	ZbbjjaEWK	ZbbjjaQCD	$S/\sqrt{B}$
$w_{\mathrm{BDT}} < -0.1$	0.75	3078	2.43	21.1	0.014
$-0.1 < w_{\mathrm{BDT}} < 0.1$	2.68	2069	4.46	12.7	0.059
$w_{\mathrm{BDT}} > 0.1$	4.56	623.8	6.46	3.17	0.183

Table 8.3: Expected yields and significance for 12.6 fb^{-1} data in the three BDT categories.

sub-dividing events with $w_{\mathrm{BDT}} < 0.1$ into two categories. These BDT categories are denoted "Low-BDT" and "Mid-BDT" with the boundary placed at $w_{\mathrm{BDT}} = -0.1$. Table 8.3 shows the expected yield for each process and in each BDT category for 12.6 fb^{-1} of data.

Chapter 9

Background Estimation

The final statistical procedure in this analysis uses a profile likelihood fit across bins of the histogrammed di-b-jet invariant mass distribution, and is discussed further in Chapter 11. The inputs to this fit are histograms containing the measured yields in data and the expected yields of both the signal and background contributions. The expected yields for the signal and the resonant $Z+\gamma$ background contributions are estimated from MC, whereas a data-driven technique is used to estimate the non-resonant background distributions. The non-resonant background is primarily composed of multijet+γ events, with additional contributions from multijet events with a jet faking a photon and $t\bar{t}$+γ events (where the heavy flavor quark pair ($b\bar{b}$) is non-resonant because each comes from the decay of a different top quark). This chapter discusses the technique used to estimate the non-resonant background.

9.1 Non-resonant background estimation

The non-resonant background contribution is estimated using an unbinned maximum likelihood fit to the di-b-jet invariant mass distribution ($m_{b\bar{b}}$) in data and was implemented using RooFit [52]. The fit is performed separately in each of the three BDT categories. The lower range of the fit is $m_{b\bar{b}} = 50\,\text{GeV}$ in all BDT categories whereas the upper range is 250, 350, and 450 GeV in the High-, Mid-, and Low-BDT categories, respectively. This variable upper range is necessary due to fewer events in the tail of the $m_{b\bar{b}}$ distribution for the higher BDT categories. The probability density function (p.d.f.) used to describe the non-resonant background is a smoothly falling polynomial function. Polynomials of degree 2, 3, and 4 were tested for signal bias using toy MC methods described in Section 9.4, and the 2^{nd} order polynomial, shown to have minimal bias, was chosen for this analysis. When performing the fit in data to extract the non-resonant background shape, both the Higgs and the Z boson mass windows (corresponding to $m_{b\bar{b}} = 80 - 140\,\text{GeV}$) were blinded to prevent signal contamination from biasing the background fit. Figure 9.1 shows the data and background fit in each BDT category.

9.2 $p_T^{b\bar{b}}$ cut

The background fit method assumes that the non-resonant background contribution is smoothly falling over the entire range of the fit. Additionally, the fit needs sufficient constraining data below the blinding window ($m_{b\bar{b}} < 80\,\text{GeV}$), which moti-

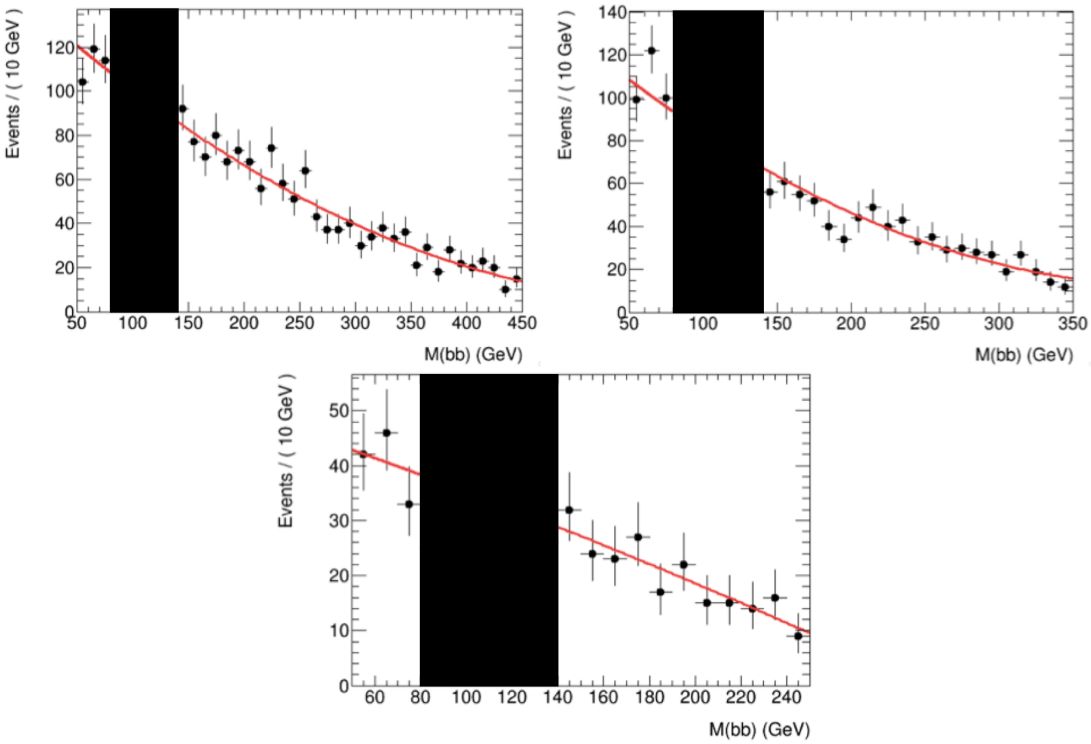


Figure 9.1: Unbinned maximum likelihood fits to data in each BDT category. Top left: Low-BDT, Top right: Mid-BDT, Bottom middle: High-BDT. The black bands in each plot are used to illustrate the region blinded during these fits.

vates the choice of 50 GeV for the lower fit range. However, the jet p_T cut in the trigger and offline selection sculpts the non-resonant background $m_{b\bar{b}}$ distribution, such that without the $p_T^{b\bar{b}}$ cut, the peak of the $m_{b\bar{b}}$ distribution is very close to the Higgs mass peak. If the $b\bar{b}$ system is at rest, the trajectories of the individual b-quarks tend to mimic the decay products of a heavy particle. The $p_T^{b\bar{b}}$ cut is used to discard these types of events, and has the effect of pushing the $m_{b\bar{b}}$ peak towards lower values for the non-resonant background. Figure 9.2 shows the evolution of the `NonResbbjja` $m_{b\bar{b}}$ shape after the MVA pre-selection, but with different cut values on $p_T^{b\bar{b}}$. It was found that for $p_T^{b\bar{b}} > 80$ GeV, the peak of the background $m_{b\bar{b}}$ is sufficiently low, such that the unbinned likelihood fit could be performed with a lower bound of 50 GeV.

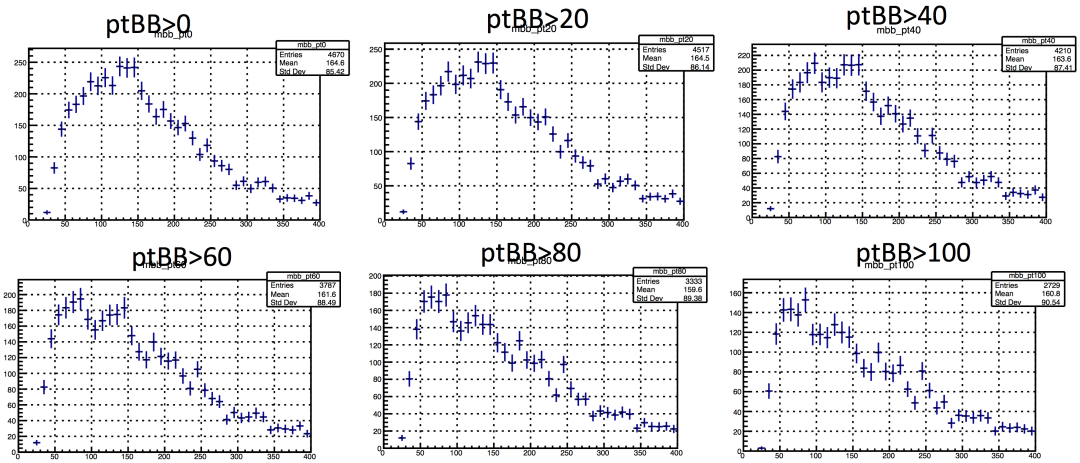


Figure 9.2: Evolution of the `NonResbbjja` $m_{b\bar{b}}$ distribution for increasing $p_T^{b\bar{b}}$ cuts. $p_T^{b\bar{b}} > 80$ GeV is used in this analysis.

9.3 Systematic uncertainties on the non-resonant background

The systematic uncertainty on the non-resonant background fit is determined using toy MC simulations. This method is expected to describe the variation of the non-resonant background yield arising from random fluctuations in the observed data. The procedure makes use of the nominal fits to the $m_{b\bar{b}}$ sideband regions in data for each of the BDT categories. These fits were used to build histograms by integrating the fit function over each histogram bin (bins of width 10 GeV starting at $m_{b\bar{b}} = 50\,\text{GeV}$). These histograms are denoted F_{ref}, and in each BDT category, the following procedure is adopted to determine the uncertainty:

- Generate toy MC simulations based on the real data. In each bin of the $m_{b\bar{b}}$ histogram for data, the simulated number of events is selected from a Gaussian whose mean is the number of events in that bin in data and whose standard deviation is the square root of the number of events in that bin in data.

- Fit each set of toy data using both the nominal fit function (2^{nd} order polynomial) and the alternative fit functions (3^{rd} and 4^{th} order polynomials). Use these fits to build histograms following the method for F_{ref}. For the k^{th} set of toy data, the histogram built using the nominal fit is denoted F_k and the histograms built using the alternative fits are denoted G_k.

- Compute two error histograms for each toy: (1) the difference between the refer-

ence histogram and the nominal toy fit histogram ($dF_k = F_k - F_{ref}$), and (2) the difference between the nominal toy fit histogram and alternative toy fit histogram ($dG_k = G_k - F_k$).

- Compute the error matrices for the independent contributions of the background uncertainty: $F_{i,j} = \frac{1}{N}\sum_{k=1}^{N} dF_k^i dF_k^j$ and $G_{i,j} = \frac{1}{N}\sum_{k=1}^{N} dG_k^i dG_k^j$. Here, k is the toy number and i and j are bin numbers. N is the number of toys.

- Use a principal component analysis to decompose the error matrices into their independent contributions ΔF_i (given by the eigenvectors of the error matrices). The corresponding eigenvalue gives the relative weight of importance of each eigenvector. It was found that the leading two eigenvectors provide a good description of the uncertainties on the non-resonant background fit, so the other eigenvectors are not used in this analysis.

- Construct two additional "systematic uncertainty histograms" in each BDT category: $F_{ref}+\Delta F_1$ and $F_{ref}+\Delta F_2$. These additional histograms correspond to nuisance parameters in the profile likelihood fit discussed in Chapter 11. The difference between each of these histograms and F_{ref} is the 1σ uncertainty on the non-resonant background estimate for the corresponding nuisance parameter.

The nominal and systematic uncertainty histograms corresponding to the above procedure for estimating the non-resonant background contribution in each BDT category are shown in Figure 9.3.

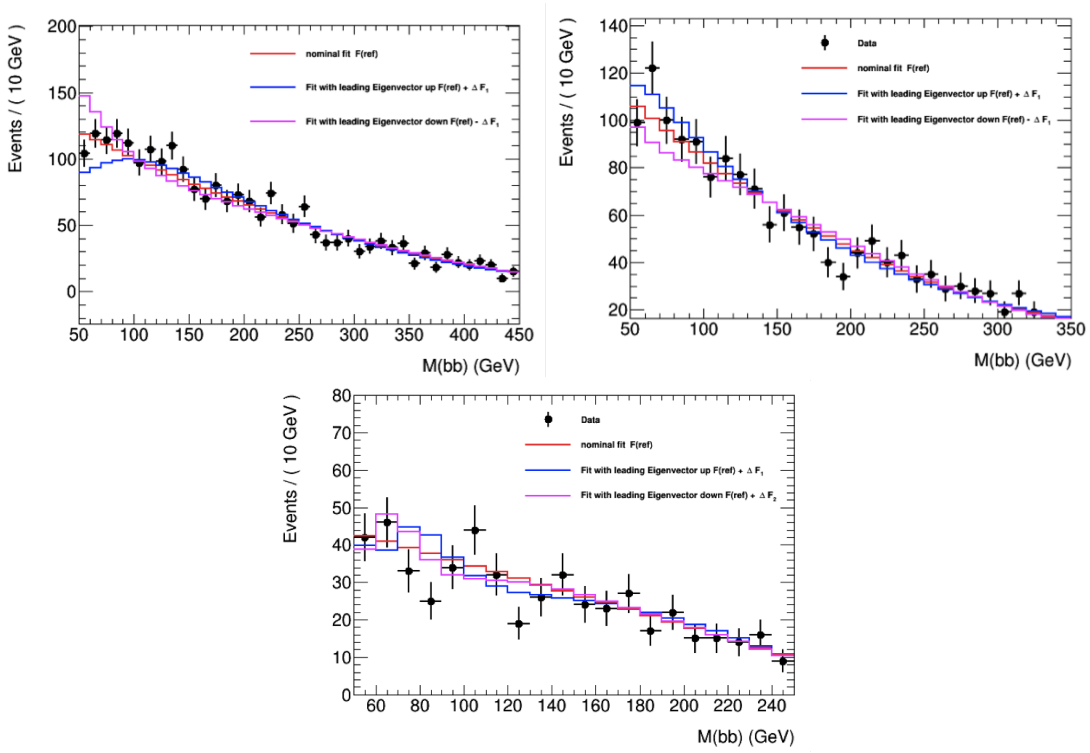


Figure 9.3: Nominal and systematic histograms of the non-resonant background estimation in this analysis. Top left: Low-BDT, Top right: Mid-BDT, Bottom middle: High-BDT

9.4 Linearity test

A linearity test has been performed to check for potential bias in the non-resonant background estimation procedure. Under this procedure, 10000 sets of toy MC data have been generated according to the nominal background fit shape and yield. For each of these toy datasets, additional events are generated according to the shape of the resonant $H + \gamma$ and $Z + \gamma$ MC samples, with varying signal strengths, μ_H and μ_Z, ranging from 0 to 5. A fit to these data is performed using a polynomial p.d.f. to describe the non-resonant background and a Crystal Ball [53] p.d.f. to describe the signal and resonant background contributions. During these fits, the parameters of the Crystal Ball functions are fixed to parameters derived from MC simulation, but their normalizations are allowed to float. Figure 9.4 shows an example fit to one of the toy datasets.

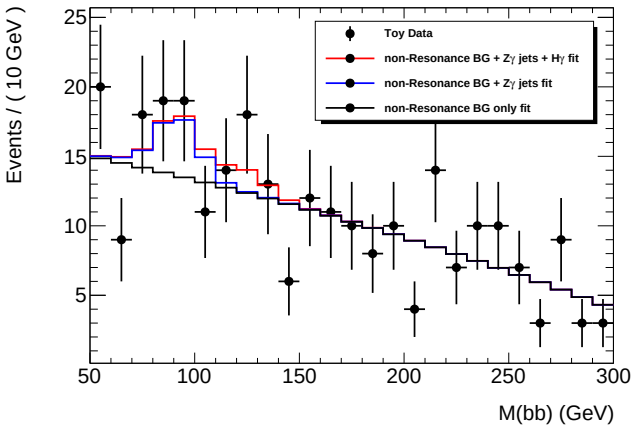


Figure 9.4: Example fit to toy MC data for the linearity test.

Figure 9.5 contains histograms showing the results of all the toy fits, where

the horizontal axis in each plot is the measured signal strength and each color line corresponds to a particular injected signal strength. Results for $H + \gamma$ ($Z + \gamma$) are shown on the left (right) of this figure.

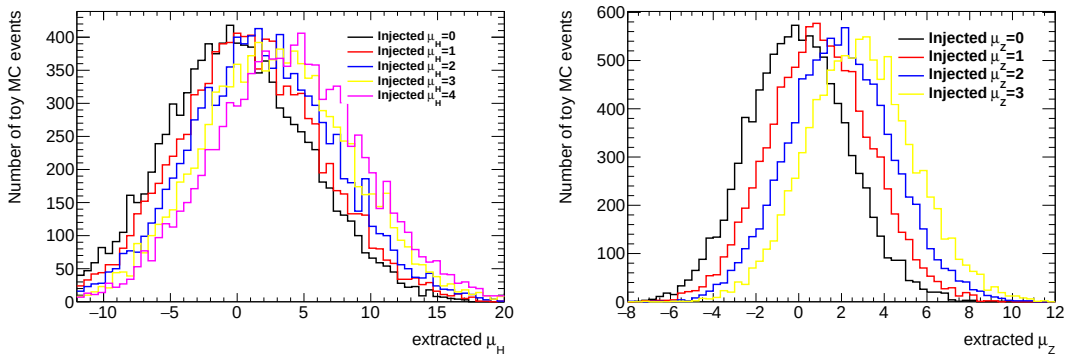


Figure 9.5: Distribution of extracted signal strengths for varying injected signal strengths. Left: μ_H, right: μ_Z.

The results of the linearity test are shown in Figure 9.6, where the horizontal axis in each plot corresponds to the injected signal strength, and the vertical axis shows the mean of the distribution of corresponding extracted signal strengths. The different colored lines on the left (right) plot are for different values of the $Z + \gamma$ ($H + \gamma$) signal strength in the test. These tests show good linearity between the injected and extracted signal strength with less than 5% bias resulting from the fit function, and less than 5% bias due to the correlation between μ_H and μ_Z.

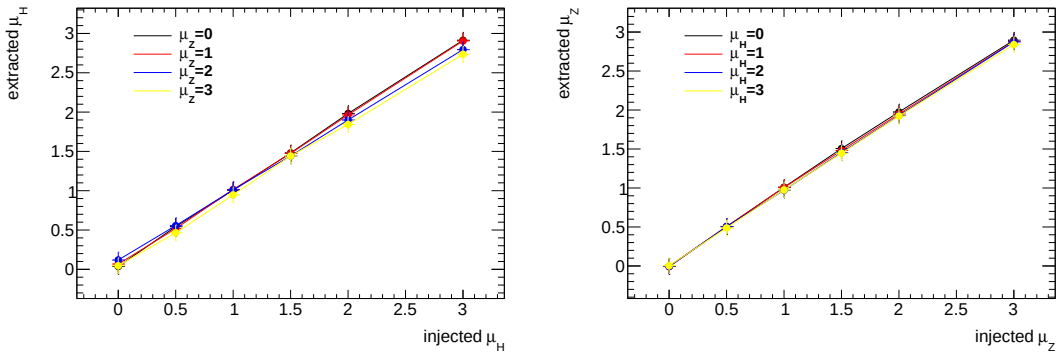


Figure 9.6: Linearity test between the injected and extracted signal strength. Left: μ_H, right: μ_Z.

Chapter 10

Systematic Uncertainties

Systematic uncertainties on the non-resonant background estimation were discussed in Section 9.3. The major systematic uncertainties for the resonant background contributions ($Z+\gamma$) and signal expectations are divided into experimental uncertainties and theory uncertainties. These uncertainties affect both the overall yield for each sample and (by propagation to the BDT input variables) the relative yield in each BDT category. Each of these systematic uncertainties is associated with a nuisance parameter in the likelihood function (Equation 11.1). The nuisance parameters corresponding to an overall yield uncertainty simply scale the likelihood fit input histograms for a particular sample by the same factor in all BDT categories. Nuisance parameters corresponding to uncertainties affecting the relative yield in each BDT category work somewhat differently. When propagated through the analysis, these relative yield uncertainties translate into uncertainties on both the normalization and the shape of the $m_{b\bar{b}}$ distribution in each BDT category. The uncertainties are parametrized by "sys-

tematic uncertainty histograms," which describe the change in the $m_{b\bar{b}}$ distribution in each BDT category for a particular sample when the corresponding nuisance parameter is shifted by one standard deviation.

10.1 Theory uncertainties

10.1.1 QCD scale uncertainty

Uncertainties on the sample cross section and acceptance for the signal and resonant background MC samples due to the choice of renormalization and factorization scale (μ_R and μ_F respectively) were evaluated by varying the choice of scale for both parameters together by factors of 0.5 (scale-down) and 2.0 (scale-up) during the event generation stage. The cross sections of the scale-up and scale-down samples are both compared with the nominal sample cross section, and the larger of the two differences is used as the uncertainty. These results are summarized in Table 10.1. Note that the nominal cross sections in this table are different from those in Table 6.3 due to slightly different parameters during event generation (the cross sections in Table 6.3 correspond to the MC samples used in the BDT training and statistical interpretation). These minor differences do not affect the relative cross section uncertainties, which are reported in the last column of Table 10.1.

An analysis performed at "truth-level" has been implemented to evaluate the uncertainty on the acceptance due to the choice of scale. This truth-level analysis is performed on MC events that have been showered using Pythia, but have not under-

Sample	Nom. X-sec	Scale Up X-sec	Scale Down X-sec	Assigned Unc.
HbbjjaSM125	58.6 ± 0.3	56.3 ± 0.3	60.9 ± 0.2	3.9%
ZbbjjaEWK	66.0 ± 0.1	64.5 ± 0.1	67.6 ± 0.1	2.4%
ZbbjjaQCD	2687 ± 6.5	2205 ± 6.1	3311 ± 6.9	23.2%

Table 10.1: MadGraph5 calculated cross sections for variations of the renormalization and factorization scales. All cross section are in [fb].

gone detector simulation and event reconstruction. At truth-level, photons are identified in the Pythia event record by selecting particles with HepMC [54] status code 1 (i.e. particles that are not decayed further by Pythia) and Particle Data Group (PDG) [8] identification number 22. Truth-level jets are formed using particles with HepMC status code 1, but excluding neutrinos and muons, as inputs to the anti-k_T clustering algorithm with distance parameter $R = 0.4$. A truth-level jet is b-tagged if it can be matched (within $\Delta R < 0.4$) to a B-hadron in the Pythia event record. The truth-level analysis is designed to mimic the MVA preselection of the full analysis; however, there is a significant difference in the acceptance between the two analyses, primarily due to the different b-tagging strategies. These differences on their own are not important. A change in the choice of QCD scale may affect the jet p_T spectrum, so it is important to determine if these effects propagate similarly to the full analysis. This check was performed by raising the jet p_T threshold by 5 GeV in both the truth-level analysis and the full analysis. The same relative change in acceptance was observed for both analyses, thus validating the truth-level procedure for evaluating the acceptance uncertainty. Further details on the truth-level analysis, including this validation, can be found in Appendix A.

Events from both the nominal and the scale-varied samples are passed through

Sample	Nom. Acc.	Scale Up Acc.	Scale Down Acc.	Assigned Unc.
HbbjjaSM125	3.56 %	3.64 %	3.57 %	2.2 %
ZbbjjaEWK	2.55 %	2.61 %	2.66 %	4.3 %
ZbbjjaQCD	0.68 %	0.89 %	0.48 %	31 %

Table 10.2: Truth-level analysis acceptances for nominal and scale-varied samples.

the truth-level analysis, and the acceptance is calculated as the fraction of events passing the selection cuts. Both the scale-up and scale-down sample acceptances are compared with the acceptance of the nominal sample, and the larger of the two differences is used as the uncertainty. These results are shown in Table 10.2.

10.1.2 Parton distribution function uncertainty

Uncertainties due to the choice of parton distribution function have been evaluated using a procedure to re-weight the MC samples from the nominal PDF to a set of error PDFs which capture the uncertainty of the fits used to produce the PDF set. Under this procedure, a weight (w_{PDF}) is assigned to each event in the nominal sample, and the formula for these weights is given in Equation 10.1.

$$w_{\mathrm{PDF}} = \frac{\mathrm{PDF}_{\mathrm{new}}(x_1, f_1, Q) * \mathrm{PDF}_{\mathrm{new}}(x_2, f_2, Q)}{\mathrm{PDF}_{\mathrm{nom}}(x_1, f_1, Q) * \mathrm{PDF}_{\mathrm{nom}}(x_2, f_2, Q)} \tag{10.1}$$

In this equation, $\mathrm{PDF}_{\mathrm{new}}$ ($\mathrm{PDF}_{\mathrm{nom}}$) is the PDF we are re-weighting to (from), x_i are the PDG identification numbers of the initial state partons ($i = 1, 2$), f_i are the momentum fractions of the initial state partons, and Q is the scale at which the PDF is evaluated.

Once these weights have been computed for all events in a sample, that sample's

cross section with the new PDF can be calculated by summing the weights over all events, dividing by the total number of events, and multiplying by the nominal cross section. The acceptance of a sample for a new PDF can be calculated by summing the PDF weights of selected events, and dividing by the sum of the PDF weights over all events. Formulas for calculating the cross section and acceptance after PDF re-weighting are given in Equations 10.2 and 10.3, respectively.

$$\sigma_{\mathrm{PDF_{new}}} = \sigma_{\mathrm{PDF_{nom}}} * \frac{\sum_{\text{all events}}(w_{\mathrm{PDF}})}{\text{Num. Events}} \tag{10.2}$$

$$\text{acc.} = \frac{\sum_{\text{selected events}}(w_{\mathrm{PDF}})}{\sum_{\text{all events}}(w_{\mathrm{PDF}})} \tag{10.3}$$

Using the PDF re-weighting procedure described above, the cross section and acceptance (relative to the values using the nominal PDF) have been calculated for the 100 error members of PDF4LHC_nlo_mc PDF set. Figures 10.1 and 10.2 show the results of this PDF study.

The total PDF uncertainty on the cross section (acceptance) used in this analysis is evaluated by calculating the standard deviation of the set of relative differences from the nominal cross section (acceptance) across all 100 error members of the PDF set. These uncertainties are shown in Table 10.3.

10.1.3 Parton shower uncertainty

The Pythia showering program, which is used to perform the parton shower for the MC samples, has several free parameters which are set using tunes to data.

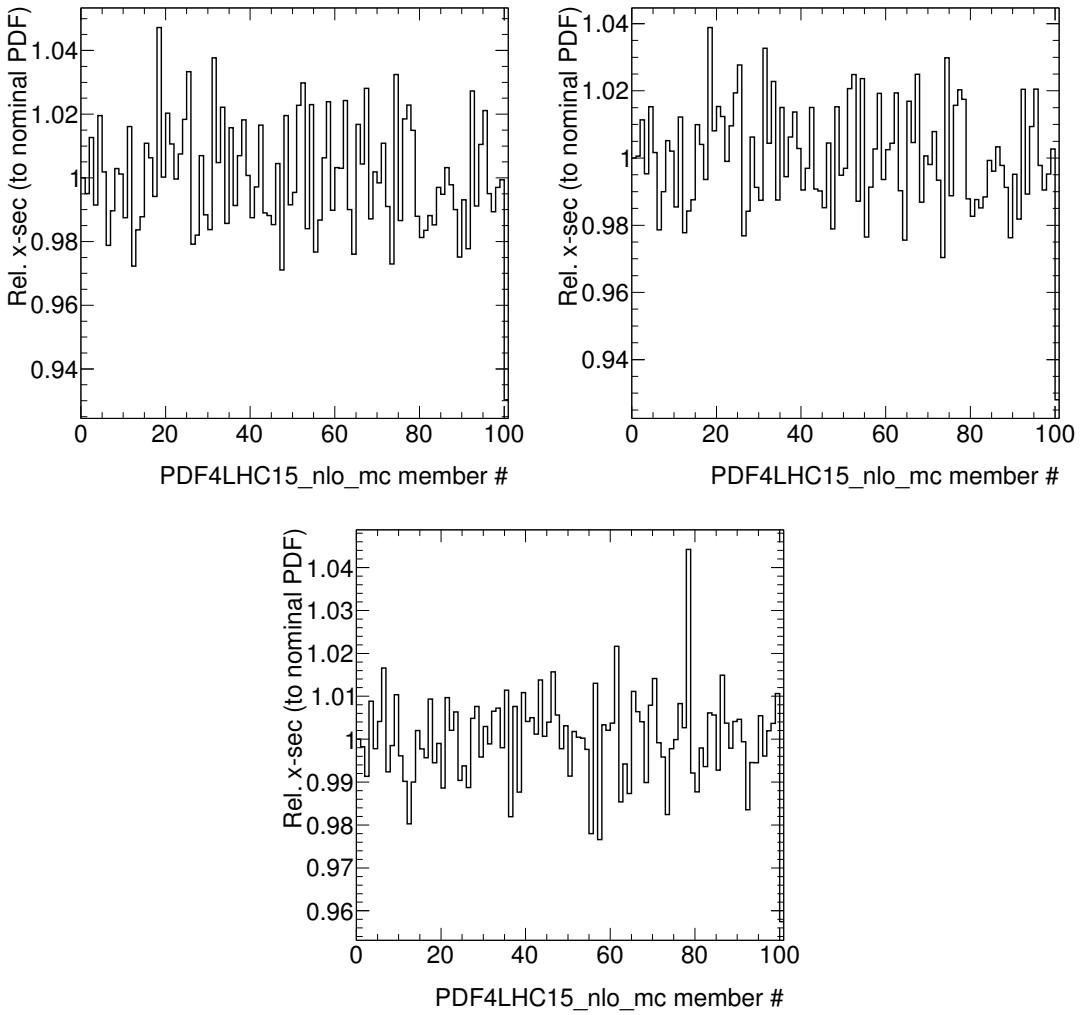


Figure 10.1: Ratio of the cross section of the PDF re-weighted samples to the nominal sample. Top left: `HbbjjaSM125`, Top right: `ZbbjjaEWK`, Bottom middle: `ZbbjjaQCD`.

Sample	X-sec Unc.	Acc. Unc.
`HbbjjaSM125`	1.8 %	1.0 %
`ZbbjjaEWK`	1.6 %	1.3 %
`ZbbjjaQCD`	1.1 %	1.2 %

Table 10.3: PDF uncertainties on the cross section and acceptance for samples used in this analysis.

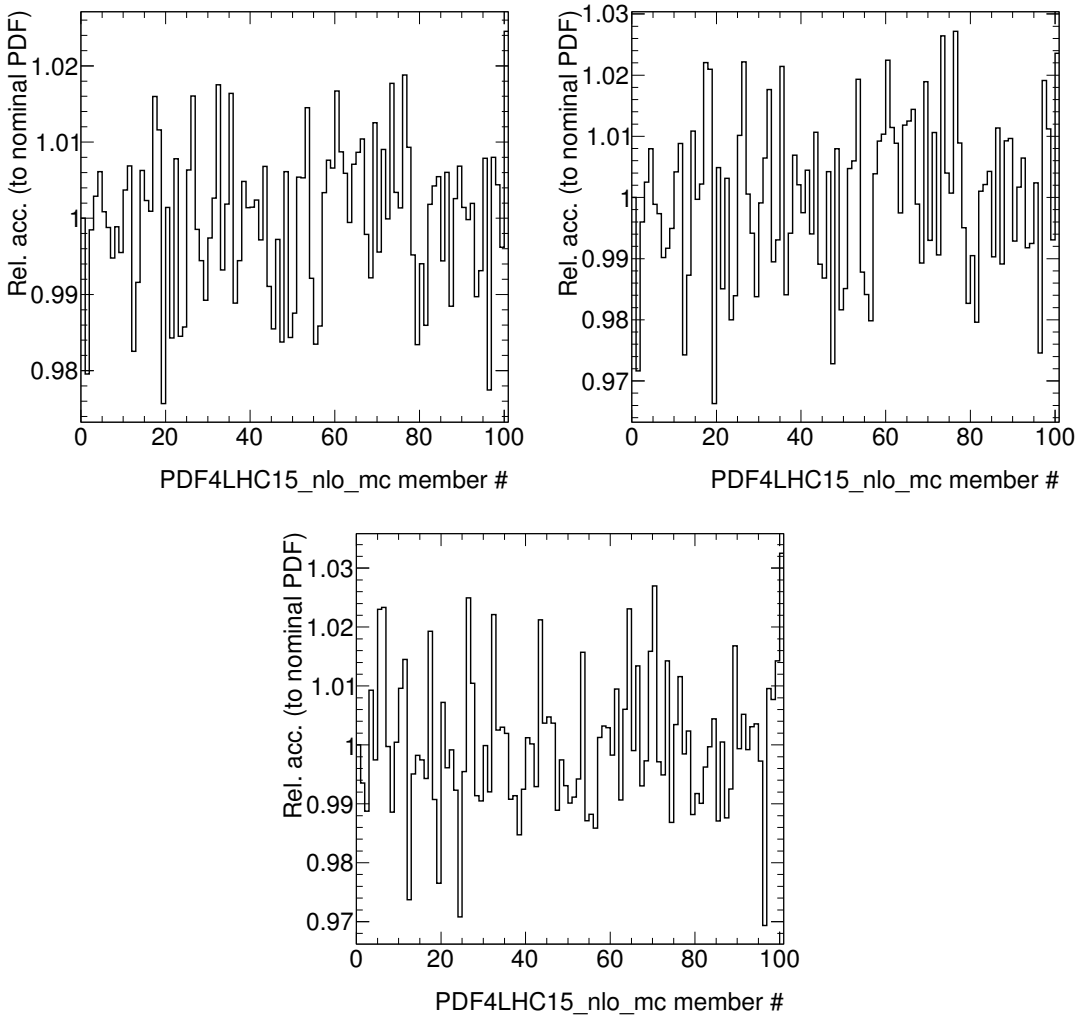


Figure 10.2: Ratio of the acceptance of the PDF re-weighted samples to the nominal sample. Top left: `HbbjjaSM125`, Top right: `ZbbjjaEWK`, Bottom middle: `ZbbjjaQCD`.

Sample	Acc. Unc. (%)
HbbjjaSM125	4.8
ZbbjjaEWK	5.1
ZbbjjaQCD	8.8

Table 10.4: Parton shower uncertainties on the acceptance for samples used in this analysis

Theory uncertainties on the parton shower are studied by generating samples with different parameter tunes. Ref. [32] details a set of five pairs of eigentune variations for the A14 tune series which are expected to provide good coverage of the experimental and modelling uncertainties implicit in the tuning. New samples were generated for all five pairs of these eigentune variations, and the acceptances were calculated using the truth-level analysis previously mentioned. The assigned uncertainty is the maximum deviation of the acceptance for any of the systematic samples from the nominal sample, and the results are summarized in Table 10.4.

10.1.4 Propagation of uncertainties through the BDT

In the previous sections, we examined the effect of QCD scale, choice of PDF, and parton shower tune on the cross section and acceptance for each MC sample. These same sources of uncertainty can also affect the relative fraction of events in each BDT category in the analysis through their impact on the distributions of BDT input variables. Using the same truth-level analysis as was used to evaluate the acceptance uncertainty, we compared BDT input variable distributions between the nominal and systematic samples for events passing the truth-level analysis cuts. Both the forward jet kinematics (m_{jj} and $\Delta\eta_{jj}$) and p_T^{balance} were found to be sensitive to the choice of QCD

scale for Higgs signal events and electroweak $Z+\gamma$ events. For QCD $Z+\gamma$ events, only p_T^{balance} was found to be sensitive to the choice of QCD scale. p_T^{balance} was also found to be sensitive to the parton shower eigentune variations for Higgs and electroweak $Z+\gamma$ events, while no systematic effect was observed for QCD $Z+\gamma$ events. Figures 10.3, 10.4, 10.5, and 10.6 show the truth-level distributions of these input variables for which systematic effects were observed.

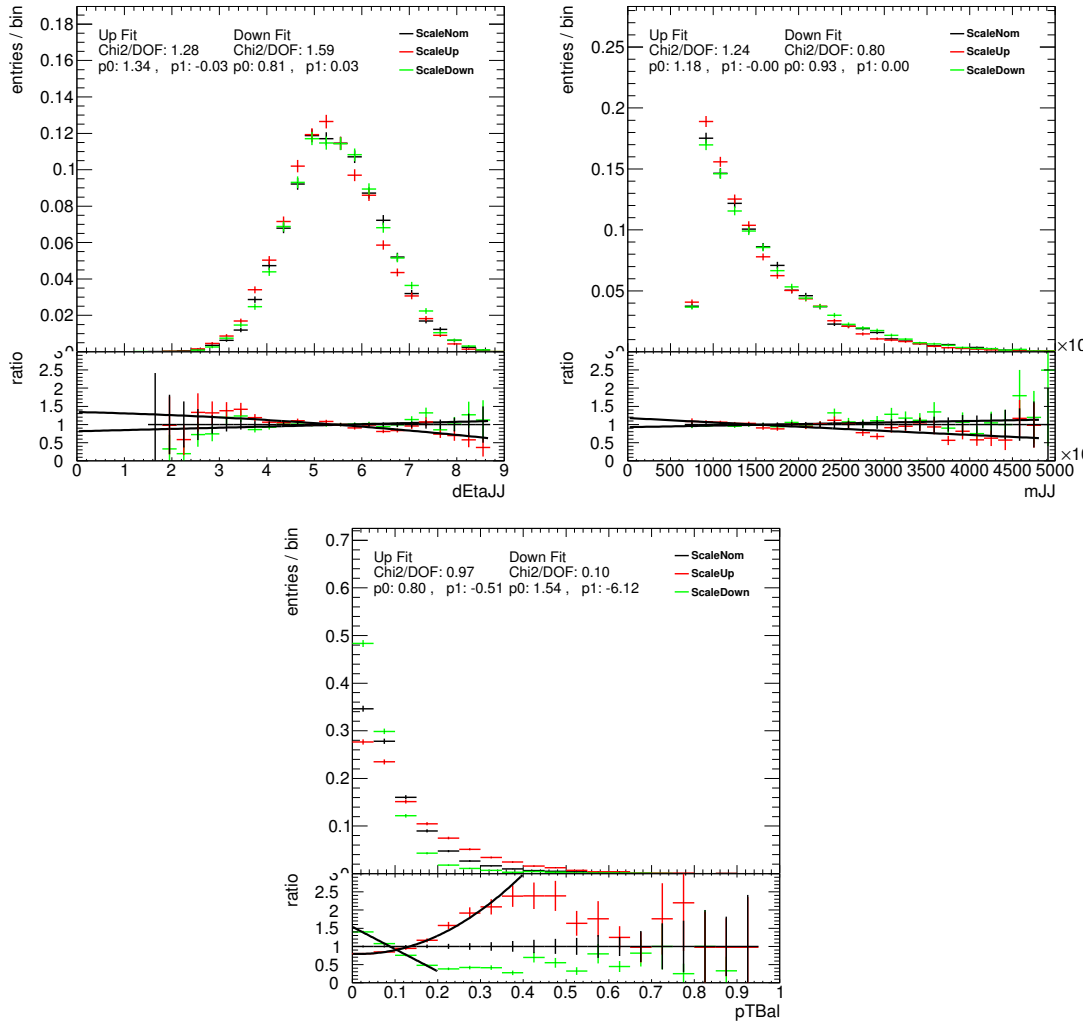


Figure 10.3: Effect of QCD scale uncertainties on BDT input variables for `HbbjjaSM125`.

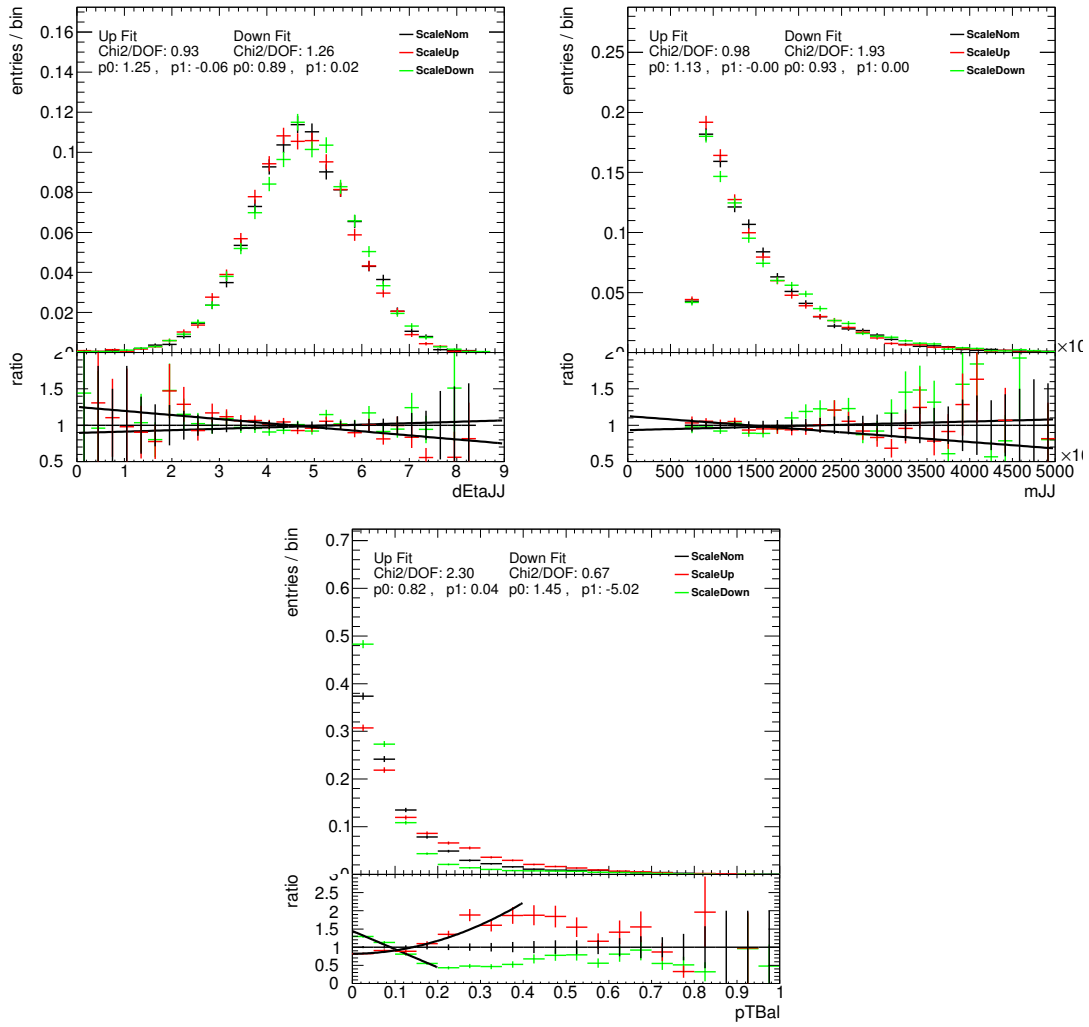


Figure 10.4: Effect of QCD scale uncertainties on BDT input variables for `ZbbjjaEWK`.

We used a re-weighting method to propagate these truth-derived uncertainties associated with the BDT input variables through the full analysis chain. The ratios of the truth distributions of the systematic samples to the nominal sample were fit using either a linear or quadratic function, which was then used to calculate a weight for each event in the full analysis. In some cases the fit is only valid in a particular

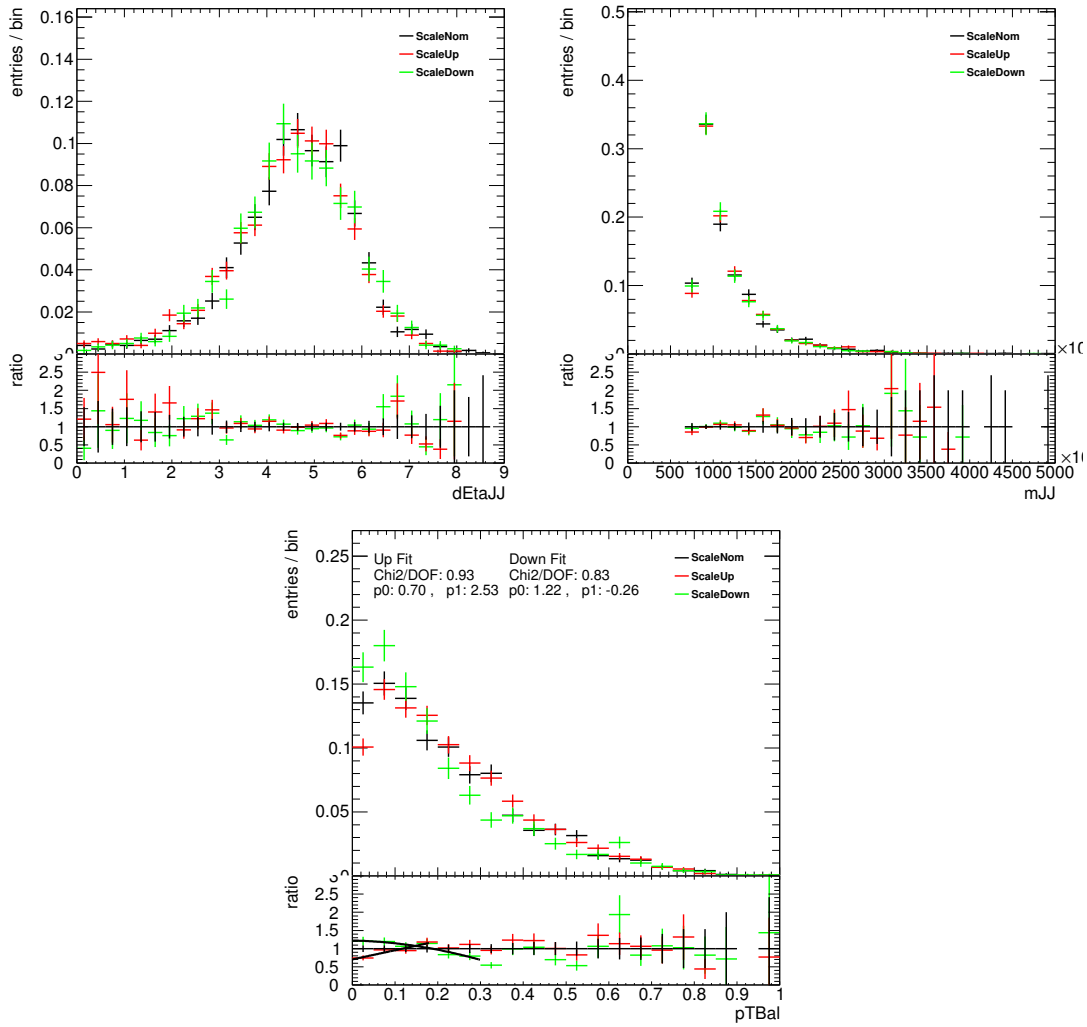


Figure 10.5: Effect of QCD scale uncertainties on BDT input variables for `ZbbjjaQCD`.

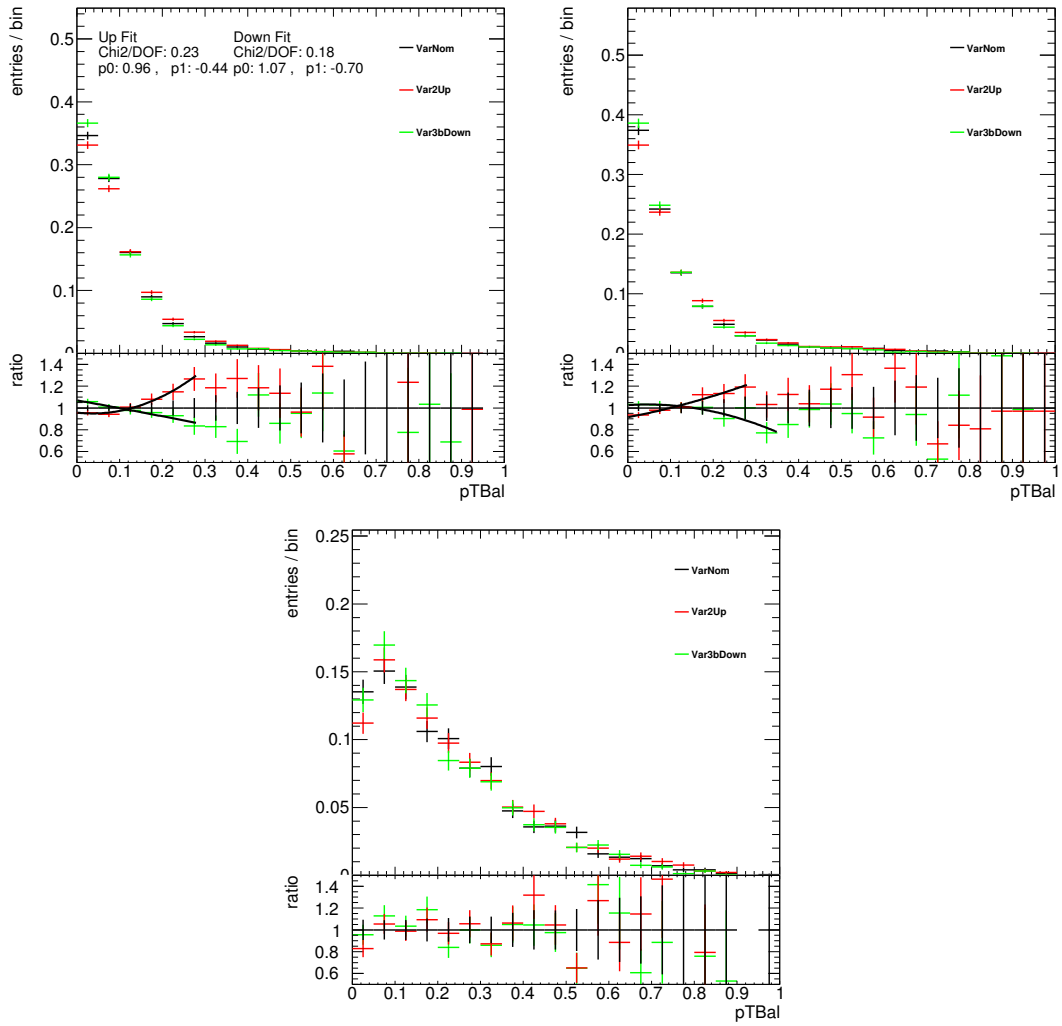


Figure 10.6: Effect of parton shower uncertainties on BDT input variables for `HbbjjaSM125` (top left), `ZbbjjaEWK` (top right), and `ZbbjjaQCD` (bottom middle).

range, and outside this range, a constant value is used as the weight. For example, in the left side of Figure 10.6, the fit is only appropriate for $p_T^{\text{balance}} < 0.28$, so for $p_T^{\text{balance}} > 0.28$, we use the value of the fit at 0.28 (i.e. $w(p_T^{\text{balance}} > 0.28) = w(p_T^{\text{balance}} = 0.28)$). Above 0.28, the ratio is relatively flat within the statistical uncertainty, so this approach is justified. Although this re-weighting does not affect the categorization of individual MC events, it does affect the weighted fraction of MC events in each BDT category. Each re-weighting function is scaled such that the sample normalization at the level of the MVA preselection in the full analysis is not affected by the re-weighting. Figure 10.7 shows an example of the truth-derived p_T^{balance} re-weighting applied to the full analysis (left), as well as its effect on the w_{BDT} distribution (right). These weights

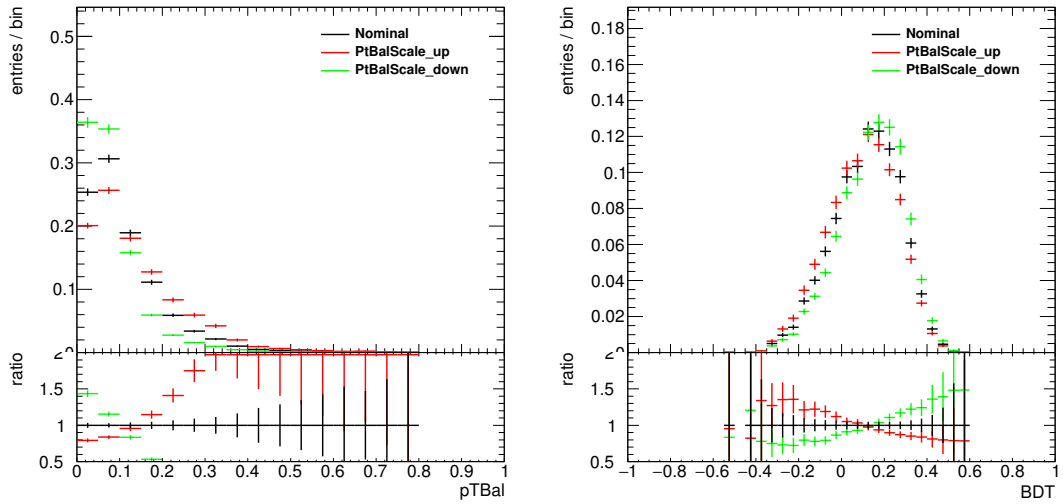


Figure 10.7: Truth-derived p_T^{balance} re-weighting applied to the full analysis for the `HbbjjaSM125` sample (left), and its effect on the BDT weight distribution (right).

are ultimately propagated to the $m_{b\bar{b}}$ distributions and used to construct systematic uncertainty histograms describing the theory uncertainties on the event categorization.

10.2 Experimental uncertainties

10.2.1 Luminosity uncertainty

The preliminary uncertainty on the combined 2015+2016 integrated luminosity is 2.9%. It is derived, following a methodology similar to that detailed in Ref [55], from a preliminary calibration of the luminosity scale using $x - y$ beam-separation scans performed in August 2015 and May 2016.

10.2.2 Pileup uncertainty

Pileup modelling uncertainties are estimated by re-weighting simulated events so that the average number of interactions per crossing $\langle\mu\rangle$ varies by +7% and -16%. The weights are propagated through the full analysis chain and used to construct systematic uncertainty histograms describing the pileup uncertainty.

10.2.3 Jet uncertainties

Uncertainties on the jet energy scale (JES) and jet energy resolution (JER) have been established based on Run 1 MC and data-based calibrations, with additional corrections for differences in beam conditions, the ATLAS detector, simulation, and object reconstruction between Run 1 and Run 2 [40]. These uncertainties cover the p_T and η dependence of JES and JER, pileup conditions, flavor response, global sequential corrections, and high p_T jets. In total 74 JES uncertainty terms are required to correctly account for all correlations in the jet calibration. This analysis uses a reduced set of

19 nuisance parameters which are treated as uncorrelated independent components. The reduction scheme uses a principal component analysis to construct a new set of independent uncertainty sources. A good approximation of the full set of nuisance parameters is obtained by selecting the subset of principal components with the most significant contributions to the overall uncertainty [56]. In the case of JER, the dominant uncertainties all have a similar shape, and thus these uncertainties are combined into a single nuisance parameter.

The modeling of quark jet width has been checked using dedicated MC and data samples. Events consistent with $t\bar{t}$ production and a subsequent semi-leptonic decay were selected, and the pair of jets with invariant mass closest to the $W^{\pm}$ mass were used to study the quark jet width. The dependence of the modeling on jet p_T, η, and the event pileup was checked by dividing the quark jets into different categories for each variable:

- jet p_T : 20-40 GeV, 40-80 GeV, and >80 GeV
- jet η: $|\eta| < 1.0$, $1.0 < |\eta| < 2.5$, $|\eta| > 2.5$
- Number of primary vertices: $n_{\mathrm{PV}} \leq 9$, $n_{\mathrm{PV}} \geq 10$.

Figure 10.8 shows some example jet width distributions, and Figure 10.9 shows the mean jet width as a function of p_T in selected kinematic regimes. In all cases, shifting the jet width by $\pm 10\%$ in MC is sufficient to cover any mis-modelings of this variable, and this is the procedure adopted by this analysis.

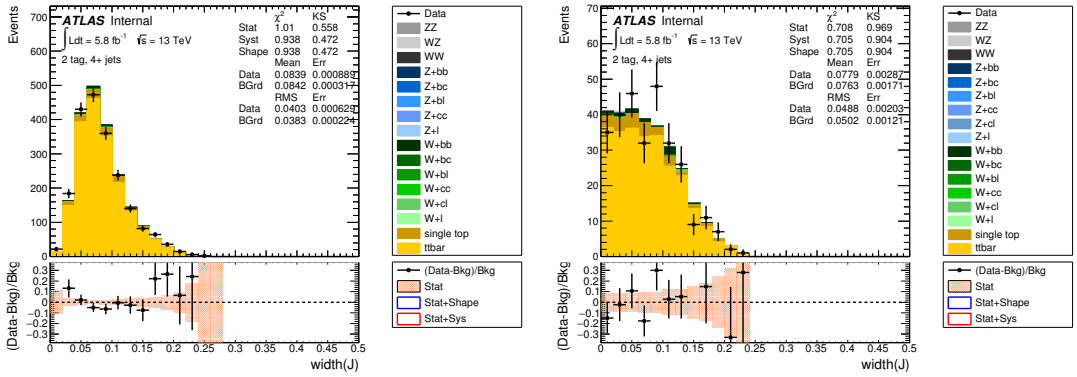


Figure 10.8: Example quark jet width distributions in selected $t\bar{t}$ events. Left: $p_T > 80\,\text{GeV}, 1.0 \leq |\eta| \leq 2.5, n_{\text{PV}} \geq 10$. Right: $20 \leq p_T \leq 40\,\text{GeV}, |\eta| > 2.5, n_{\text{PV}} \leq 9$.

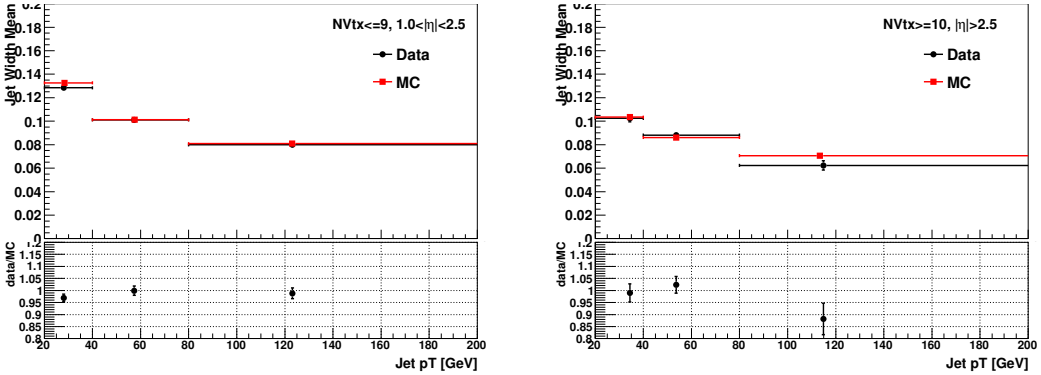


Figure 10.9: Mean jet width as a function of p_T in different kinematic regimes.

The sideband region in this analysis contains a mixture of quark and gluon jets, and is well modeled by MC simulation as shown in Figure 8.12 (top).

Jet uncertainties are propagated through the full analysis chain to construct systematic uncertainty histograms for the corresponding nuisance parameters.

10.2.4 Flavor tagging uncertainties

b-tagging (in)efficiency scale factors are calculated jet-by-jet and applied to MC events as a whole to correct for differences in b-tagging efficiencies between data and MC. A combination of muon-based methods and $t\bar{t}$-based methods are used to calculate these scale factors [57]. Systematic uncertainties affecting these methods arise from:

- how well MC simulation models heavy flavor production, decays, and fragmentation
- JES, JER, and pileup uncertainties
- heavy flavor contamination in light flavor control samples
- soft-muon tagging criterion
- how well MC simulation models $t\bar{t}$ production (generator dependence, initial state radiation, heavy flavor fragmentation)
- background subtraction techniques

In total, ≈160 systematic variations are needed to properly account for all sources of uncertainty and the correlations among them. To reduce the burden of accounting for this large number of variations, a principal component analysis was performed which reduces this number to a set of 14 variations expected to provide a good description of the uncertainties associated with b-tagging. b-tagging uncertainties are propagated through the full analysis chain to construct systematic uncertainty histograms for the corresponding nuisance parameters.

10.2.5 Photon uncertainties

This analysis is only weakly sensitive to the energy scale and resolution of the photon, so a simplified correlation model is used for EM scale and resolution uncertainties. Only two systematic variations are used: one for scale and one for resolution.

The first step in photon identification in MC samples applies a correction to the electromagnetic shower shape. This is because the electromagnetic showers in simulation are usually narrower than in the data, so the distributions of shower shape variables used by photon ID algorithms must be shifted. The uncertainty on photon efficiency scale factors is computed from the absolute value of the relative difference between the MC photon efficiencies before and after applying the shower shape shifts [58].

Before evaluating photon isolation in MC samples, a data-driven correction is applied to topoetcone40. The systematic uncertainty on the photon calorimeter isolation is evaluated by switching off this data-driven correction, and re-evaluating the photon isolation.

Photon uncertainties are propagated through the full analysis chain to construct systematic uncertainty histograms for the corresponding nuisance parameters.

10.2.6 $H_T^{\mathbf{soft}}$ uncertainty

Many ATLAS analyses include as an object the negative vector sum of the p_T of reconstructed objects in the event ($E_{\mathrm{T}}^{\mathrm{miss}}$). The calculation of this quantity includes a track-soft-term (TST) which comprises all tracks not associated with high-p_T physics objects, and a standard procedure exists for calculating the TST experimental uncertainties [59]. Given the similarity between the definitions of the $E_{\mathrm{T}}^{\mathrm{miss}}$ TST and H_T^{soft}, we assign the TST uncertainty on H_T^{soft} in this analysis. The TST uncertainties depend on the transverse momentum of the vector sum of the high-p_T physics objects in the event. This quantity is called $\overrightarrow{p}_{\mathrm{T}}^{\mathrm{hard}}$ and is defined in Equation 10.4.

$$\overrightarrow{p}_{\mathrm{T}}^{\mathrm{hard}} = (\overrightarrow{p}_{\gamma} + \overrightarrow{p}_{b_1} + \overrightarrow{p}_{b_2} + \overrightarrow{p}_{j_1} + \overrightarrow{p}_{j_2})_{\mathrm{T}} \tag{10.4}$$

The TST uncertainties are split into scale and resolution components. The resolution component is further divided into contributions parallel and perpendicular to $\overrightarrow{p}_{\mathrm{T}}^{\mathrm{hard}}$. Figure 10.10 shows the TST uncertainties applied on $H_{\mathrm{T}}^{\mathrm{soft}}$ in the signal MC sample. These H_T^{soft} uncertainties are propagated through the full analysis chain and used to construct systematic uncertainty histograms for the corresponding nuisance parameters.

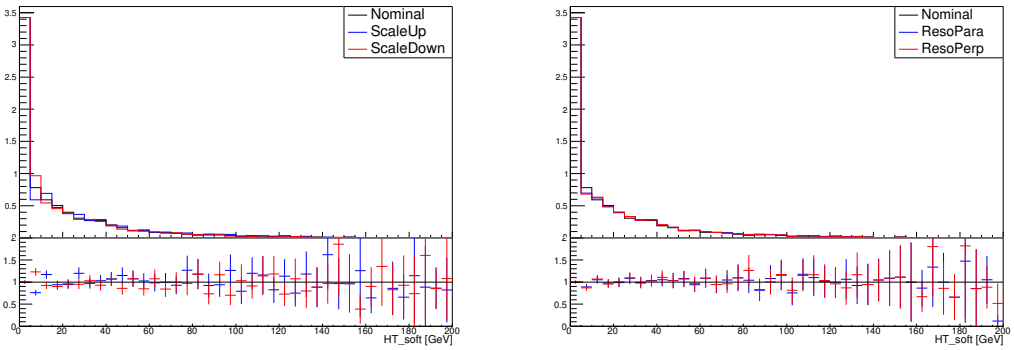


Figure 10.10: Effect of missing transverse momentum TST uncertainties on $H_{\mathrm{T}}^{\mathrm{soft}}$ in the signal MC sample. The vertical axis in each plot is the expected number of signal events per bin for $12.6\,\mathrm{fb}^{-1}$.

Chapter 11

Statistical Analysis

A statistical procedure based on a likelihood function is used to quantify the results of this search. The inputs to this procedure are: (1) $m_{b\bar{b}}$ histograms containing the measured data yields in each BDT category, (2) $m_{b\bar{b}}$ histograms containing the expected signal and background yields in each BDT category, (3) parameters describing normalization uncertainties on the expected yields due to uncertainties described in Chapter 10, and (4) $m_{b\bar{b}}$ histograms describing shape+normalization uncertainties on the expected $m_{b\bar{b}}$ distributions due to uncertainties discussed in Chapters 9 and 10. Using the likelihood function and these inputs, three statistical analyses are performed. A discovery test is used to measure the compatibility of the observed data with the background-only hypothesis. Data that are sufficiently discrepant from the background-only hypothesis would indicate a discovery. In the absence of discovery, a procedure is used to set upper limits on the signal cross section relative to the SM prediction. Lastly, the best fit value of the signal cross section relative to the SM prediction is obtained

by maximizing the likelihood function over all parameters. This chapter describes the likelihood function, its configuration, and how it is used to perform the three statistical tests in this analysis.

11.1 Likelihood function definition

The likelihood function is built as the production of Poisson probability terms across all analysis bins with constraint terms for systematic uncertainties implemented as nuisance parameters. The likelihood function is defined in Equation 11.1.

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j(\boldsymbol{\theta}) + b_j(\boldsymbol{\theta}))^{n_j}}{n_j!} \, e^{-(\mu s_j(\boldsymbol{\theta}) + b_j(\boldsymbol{\theta}))} \prod_{l=1}^{P} \mathcal{G}_X(\theta_l | \bar{\theta}_l, \sigma_{\theta_l}) \prod_{i=1}^{Q} f_X(\theta_i | \bar{\theta}_i, \sigma_{\theta_i}) \tag{11.1}$$

In this equation, the index j runs over all $m_{b\bar{b}}$ bins in the three BDT regions; μ is the parameter of interest, defined as the ratio of the measured and expected signal cross section; s_j and b_j are the expected number of signal and background events, respectively, in the j^{th} bin, and n_j is the number of measured data events in that bin. Both s_j and b_j are functions of the nuisance parameters (NP), $\boldsymbol{\theta}$. The NPs describe the systematic uncertainties discussed in previous chapters, and are parametrized by Gaussian ($\mathcal{G}_X$) or log-normal (f_X) priors. Log-normal priors are used for normalization systematics (e.g. cross section and acceptance uncertainties) to maintain the normalization positive. Each prior constrains a NP to its nominal value ($\bar{\theta}$) within its associated uncertainty ($\sigma_{\bar{\theta}}$). In total, there are 74 NPs in this analysis.

11.2 Fit configuration

Two fit configurations are used in this analysis: a default configuration is used to search for $H(\rightarrow b\bar{b}) + \gamma$, and an alternative configuration is used to search for $Z(\rightarrow b\bar{b}) + \gamma$. In both configurations the non-resonant background contribution is estimated using the method described in Chapter 9. In the default configuration, `HbbjjaSM125` models the signal of interest while `ZbbjjaEWK` and `ZbbjjaQCD` model the resonant backgrounds. In the alternative configuration, both `ZbbjjaEWK` and `ZbbjjaQCD` are considered signal samples, while `HbbjjaSM125` is a background source. In Equation 11.1, signal and background samples are differentiated in the fit by the signal strength parameter μ which multiplies the signal yield in each analysis bin. In the alternative configuration, a single signal strength parameter multiplies the expected yield for both $Z + \gamma$ contributions.

Both configurations use a common set of analysis bins, which were optimized for the $H(\rightarrow b\bar{b}) + \gamma$ search. These bins are defined by:

- w_{BDT} : three bins between -1.0 and 1.0 with boundaries at -0.1 and 0.1
- $m_{b\bar{b}}$: bins of width 10 GeV starting at $m_{b\bar{b}} = 50$ GeV. The upper fit range depends on the BDT bin with 250, 350, and 450 GeV used in the High-, Mid-, and Low-BDT bins, respectively.

The following systematic uncertainties are considered in the fit:

- QCD scale uncertainties on the cross section, acceptance, and BDT categorization of simulated samples

- PDF uncertainties on the cross section and acceptance of simulated samples
- Parton shower uncertainties on the cross section, acceptance, and BDT categorization of simulated samples
- Uncertainties on the non-resonant background estimation due to the choice of fit function and eigenvariations of the fit parameters
- Photon-related systematics on energy scale and resolution, and on identification and isolation efficiency
- Reduced set of jet systematics relating to energy calibration
- Reduced set of b-tagging systematics
- $H_{\mathrm{T}}^{\mathrm{soft}}$ uncertainty derived from the $E_{\mathrm{T}}^{\mathrm{miss}}$ TST uncertainty
- Trigger efficiency uncertainty
- Integrated luminosity uncertainty
- Jet width uncertainty.

11.3 Statistical tests

11.3.1 Best fit μ

The best fit value of the signal strength parameter is denoted $\hat{\mu}$ and is calculated by maximizing the likelihood function with respect to all NPs and μ ($\hat{\boldsymbol{\theta}}$ are

the corresponding set of NPs that maximize $\mathcal{L}$). Each term under the first product in Equation 11.1 is maximized when the expected number of signal and background events ($\mu s + b$) is equal to the number of events in data (n). Thus, in maximizing $\mathcal{L}$, the fit adjusts the values of μ and $\boldsymbol{\theta}$ to bring the measured and expected yields into better agreement. However the 2$^{\text{nd}}$ and 3$^{\text{rd}}$ factors in that equation contain penalty terms which decrease the likelihood when $\boldsymbol{\theta}$ are shifted away from their nominal values. In this way, there is an interplay between bringing the measured and expected yields into better agreement whilst not shifting the NPs too far from their nominal values. The uncertainty on $\hat{\mu}$ is calculated by varying μ up and down until the natural log of the likelihood function shifts by one-half.

11.3.2 Discovery test

The discovery test is based on the likelihood ratio defined in Equation 11.2.

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}}_{\boldsymbol{\mu}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{11.2}$$

In this equation, the denominator is the maximized value of $\mathcal{L}$ over all NPs and μ, whereas the numerator is the maximized likelihood value over all NPs for a given value of μ. $\hat{\hat{\boldsymbol{\theta}}}_{\boldsymbol{\mu}}$ denote the set of NPs that maximize $\mathcal{L}$ for a particular μ. For this test, the null hypothesis (H_0) is defined as the background-only hypothesis, and the alternative (H_1) includes both signal and background. The discovery test is used to measure the compatibility of the observed data with H_0, by calculating a p-value representing the probability of observing data as discrepant or more than our collected data, under

the assumption of H_0. The test statistic (r_0) used in this discovery test is defined in Equation 11.3.

$$r_0 = \begin{cases} -2 \ln \lambda(0), & \hat{\mu} > 0 \\ +2 \ln \lambda(0), & \hat{\mu} < 0 \end{cases} \tag{11.3}$$

We use the expected distribution of the test statistic under H_0 (denoted $f(r_0|0)$) to calculate the p-value, as shown in Equation 11.4.

$$p_0 = \int_{r_{0,obs}}^{\infty} f(r_0|0) \, \mathrm{d}r_0 \tag{11.4}$$

A small p-value indicates the observed data is poorly described by the background-only hypothesis. This p-value is often converted to an equivalent Z-score (i.e. "number of sigma"), representing the number of standard deviations away from the mean of a Gaussian distribution such that integrating the upper tail would yield the same probability. Mathematically,

$$Z = \Phi^{-1}(1 - p_0), \tag{11.5}$$

where Φ is is the cumulative distribution of the standard Gaussian function. The statistical significance of a measurement is expressed as the Z-score, and in particle physics, 3σ is considered evidence for new phenomena and 5σ is the threshold for discovery. The expected significance for this analysis is calculated using Equation 11.4 with the median of the distribution of r_0 under the assumption of H_1 as the lower bound of integration.

11.3.3 Exclusion limits

In the absence of discovery, we will set upper limits on the signal strength parameter, μ, using the CLs method [60]. An alternative likelihood ratio ($\tilde{\lambda}(\mu)$) and test statistic ($\tilde{r}_\mu$) are used, and these are defined in Equations 11.6 and 11.7, respectively.

$$\tilde{\lambda}(\mu) = \begin{cases} \frac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}}_{\boldsymbol{\mu}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})}, & \hat{\mu} \geq 0 \\ \frac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}}_{\boldsymbol{\mu}})}{\mathcal{L}(0, \hat{\hat{\boldsymbol{\theta}}}_{\mathbf{0}})}, & \hat{\mu} < 0 \end{cases} \tag{11.6}$$

$$\tilde{r}_\mu = \begin{cases} -2 \ln \tilde{\lambda}(\mu), & \hat{\mu} \leq \mu \\ +2 \ln \tilde{\lambda}(\mu), & \hat{\mu} > \mu \end{cases} \tag{11.7}$$

Using the expected distribution of $\tilde{r}_\mu$ under both the signal + background hypothesis ($f(\tilde{r}_\mu|\mu)$) and the background-only hypothesis ($f(\tilde{r}_\mu|0)$) we can compute two additional p-values. p_μ is the probability of rejecting the μ hypothesis, when it is true (type-I error), and p_b is the probability of failing to reject the μ hypothesis when $\mu = 0$ is correct (type-II error). $1-p_b$ is referred to as the power of the test, and is the probability that the test correctly rejects the μ hypothesis when $\mu = 0$ is correct. p_μ and p_b are defined in Equations 11.8 and 11.9, respectively.

$$p_\mu = \int_{\tilde{r}_{\mu,obs}}^{\infty} f(\tilde{r}_\mu|\mu) \mathrm{d}\tilde{\mathrm{r}}_\mu \tag{11.8}$$

$$p_b = \int_{-\infty}^{\tilde{r}_{\mu,obs}} f(\tilde{r}_\mu|0) \mathrm{d}\tilde{\mathrm{r}}_\mu \tag{11.9}$$

Both p_μ and p_b are used in the CLs method to set an upper limit on μ. The CLs ratio is defined in Equation 11.10, and is a measure of our ability to distinguish that value of

μ from the best fit value $\hat{\mu}$.

$$\text{CLs} = \frac{p_\mu}{1 - p_b} \tag{11.10}$$

Using this method, we can exclude values of μ or greater at $(1 - \alpha)\%$ confidence when $\text{CLs}(\mu) < \alpha$. The purpose of the denominator in Equation 11.10 is to provide a more conservative exclusion limit in cases where the experimental sensitivity to a given value of μ is very low. In those cases, as p_μ decreases, so does $1 - p_b$, and the $\text{CLs}(\mu) < \alpha$ condition is less likely to be satisfied. The expected exclusion limit can be found by replacing the finite integration bounds in Equations 11.8 and 11.9 with the median of the distribution of $\tilde{r}_\mu$ under the background-only hypothesis.

Chapter 12

Results

12.1 Expected and measured yields

The measured and expected di-b-jet invariant mass distributions in the Low-, Mid-, and High-BDT regions are shown in Figure 12.1. In those plots, the black points are the measured yield in data, the light blue contribution to the histogram stack shows the non-resonant background estimate from the polynomial fits in the $m_{b\bar{b}}$ sideband region, the light (dark) gray shows the MC estimated electroweak (QCD) $Z + \gamma$ contribution, and the un-stacked red line histogram shows the expected $H + \gamma$ signal for $\mu = 10$. The black points in the lower panel of each plot show the ratio of data to the total background estimate, and the blue lines represent the total uncertainty on the estimated background. In both the Low- and the Mid-BDT regions, the observed number of events are consistent with the expectations, while a deficit of events is seen in the High-BDT region at the bin with the most expected signal events.

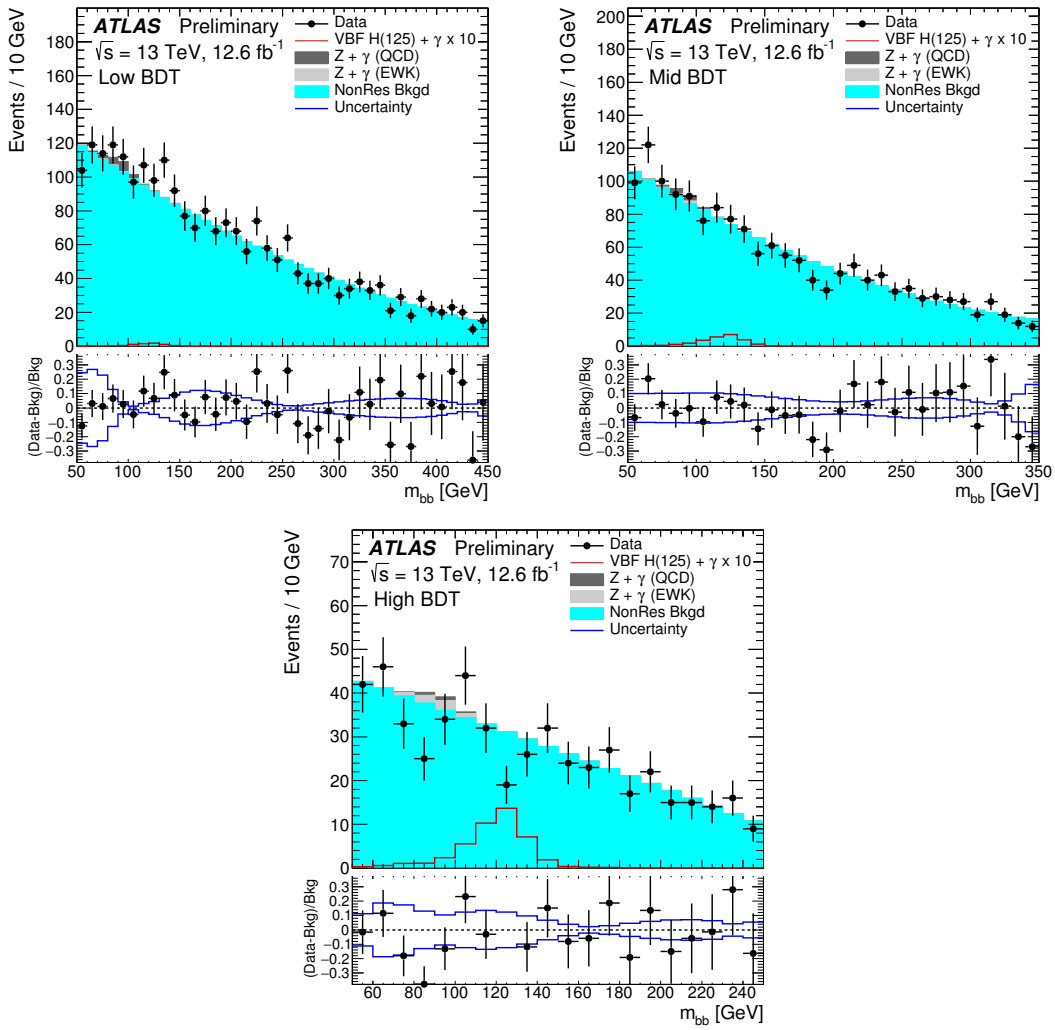


Figure 12.1: Invariant mass distributions for each of the three BDT regions considered in the likelihood fit.

12.2 Results from the statistical analysis

Results from the likelihood fit to the $m_{b\bar{b}}$ distributions are shown in Table 12.1. Due to the deficit of events near the Higgs peak in the High-BDT region, the observed signal strength is negative, and a stronger-than-expected limit is set on the signal strength in this analysis. The observed signal strength is consistent with the SM expectation value of $\mu = 1$ at the level of 2σ. In the search for $Z+\gamma$ with $Z \to b\bar{b}$, a mild excess above the background-only hypothesis is observed, and the data are consistent with the SM expectation value $\mu = 1$ at the level of 1σ.

Result	$H(\to b\bar{b}) + \gamma jj$	$Z(\to b\bar{b}) + \gamma jj$
Expected significance	0.4	1.3
Expected p-value	0.4	0.1
Observed p-value	0.9	0.4
Expected limit	$6.0\ ^{+2.3}_{-1.7}$	$1.8\ ^{+0.7}_{-0.5}$
Observed limit	4.0	2.0
Observed signal strength μ	$-3.9\ ^{+2.8}_{-2.7}$	0.3 ± 0.8

Table 12.1: Results from the statistical interpretation of search results for the three BDT bins combined.

The dominant systematic uncertainties on the measurement of μ are shown in Table 12.2. These are ranked by the shift in the fitted value of μ when the corresponding nuisance parameter is varied by $\pm 1\sigma$, and only absolute uncertainties greater than 0.08 are included in the table. Of those systematic uncertainties, the non-resonant background fit uncertainties have the largest impact on the fitted value of μ. These uncertainties depend on the statistical uncertainty of the data and are expected to decrease in size for future iterations of this analysis that have access to more data.

It is also important to note that the power of this analysis is limited by the

Uncertainty source	$\Delta\mu$
Non-resonant background uncertainty in Mid-BDT region	0.22
Non-resonant background uncertainty in High-BDT region	0.21
Non-resonant background uncertainty in Low-BDT region	0.17
Parton shower uncertainty on $H + \gamma$ acceptance	0.16
QCD scale uncertainty on $H + \gamma$ cross section	0.13
$H \to b\bar{b}$ branching ratio uncertainty	0.13
Jet energy uncertainty from calibration across η	0.10
Jet energy uncertainty from flavour composition in calibration	0.09
Integrated luminosity uncertainty	0.08

Table 12.2: Summary of systematic uncertainties affecting the measurement of the Higgs boson signal strength μ.

number of data events. The absolute uncertainties listed in Table 12.2 are significantly smaller than the overall uncertainty on μ, which is roughly ± 2.7. The dominant contribution to the uncertainty on μ is the statistical uncertainty on the data, which will decrease as more ATLAS data is added to the analysis in the future. Similarly, the dominant three systematic uncertainties, which are derived using toy MC methods applied to real data, will also decrease with more data.

12.3 Conclusion

A search for Higgs boson production via vector boson fusion in association with a photon has been performed targeting the $H \to b\bar{b}$ decay. No excess of events above the background-only expectation was observed in this analysis, and the observed and expected 95% CL upper limits on the Higgs boson production cross section times branching ratio are 4.0 and $6.0^{+2.3}_{-1.7}$ times the Standard Model expectation. This analysis is not yet at the level of sensitivity needed to provide a stringent test of the Standard

Model, but has great future potential. In particular, only a small subset of the expected LHC Run 2 data set has been analyzed, and this analysis was built "from scratch" without a Run 1 counterpart. A significant effort was expended to implement a dedicated trigger, generate new MC samples, and build the analysis framework, as well as design the overall analysis strategy. Going forward, some of the efforts that were focused in these areas can be used to study changes that have the potential to enhance the overall analysis sensitivity. Current efforts are focused on a combination with an inclusive VBF $H \to b\bar{b}$ analysis, and those results can be expected in the not-too-distant future.

Bibliography

[1] Rutherford, Ernest , "The scattering of α and β particles by matter and the structure of the atom," *Philosophical Magazine Series 6*, vol. 21, no. 125, pp. 669–688, 1911.

[2] D. Gillberg, U. Blumenschein, and J. Kretzschmar, "Updated $\sqrt{s}$ Summary Plot of STDM, TOP and HIGGS Cross Sections," Tech. Rep. ATL-COM-PHYS-2016-280, CERN, Geneva, Mar 2016.

[3] G. Aad *et al.*, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV," *JHEP*, vol. 08, p. 045, 2016.

[4] E. Gabrielli *et al.*, "Higgs Boson Production in Association with a Photon in Vector Boson Fusion at the LHC," *Nucl. Phys.*, vol. B781, pp. 64–84, 2007.

[5] C. C. Tully, *Elementary particle physics in a nutshell.* Princeton, NJ: Princeton Univ. Press, 2011.

[6] S. Weinberg, "A Model of Leptons," *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, Nov 1967.

[7] P. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys. Rev. Lett.*, vol. 13, pp. 508–509, Oct 1964.

[8] C. Patrignani *et al.*, "Review of Particle Physics," *Chin. Phys.*, vol. C40, no. 10, p. 100001, 2016.

[9] H. Georgi, *Weak interactions and modern particle theory.* Dover Books on Physics, New York, NY: Dover, 2009.

[10] G. Dissertori, I. Knowles, and M. Schmelling, *Quantum Chromodynamics: High Energy Experiments and Theory.* International series of monographs on physics, Clarendon Press, 2003.

[11] G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC ," *Physics Letters B*, vol. 716, no. 1, pp. 1 – 29, 2012.

[12] S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC ," *Physics Letters B*, vol. 716, no. 1, pp. 30 – 61, 2012.

[13] S. P. Martin, "A Supersymmetry primer," 1997. [Adv. Ser. Direct. High Energy Phys.18,1(1998)].

[14] G. Aad *et al.*, "Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector," *Journal of High Energy Physics*, vol. 2015, no. 4, p. 117, 2015.

[15] S. Chatrchyan *et al.*, "Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons," *JHEP*, vol. 05, p. 104, 2014.

[16] G. Aad *et al.*, "Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated $(W/Z)H$ production with the ATLAS detector," *JHEP*, vol. 01, p. 069, 2015.

[17] S. Chatrchyan *et al.*, "Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks," *Phys. Rev.*, vol. D89, no. 1, p. 012003, 2014.

[18] G. Aad *et al.*, "Search for the Standard Model Higgs boson decaying into $b\bar{b}$ produced in association with top quarks decaying hadronically in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *JHEP*, vol. 05, p. 160, 2016.

[19] V. Khachatryan *et al.*, "Search for a Standard Model Higgs Boson Produced in Association with a Top-Quark Pair and Decaying to Bottom Quarks Using a Matrix Element Method," *Eur. Phys. J.*, vol. C75, no. 6, p. 251, 2015.

[20] M. Aaboud *et al.*, "Search for the Standard Model Higgs boson produced by vector-boson fusion in 8 TeV pp collisions and decaying to bottom quarks with the ATLAS detector," CERN-EP-2016-076. 2016.

[21] V. Khachatryan *et al.*, "Search for the standard model Higgs boson produced through vector boson fusion and decaying to $b\bar{b}$," *Phys. Rev.*, vol. D92, no. 3, p. 032008, 2015.

[22] L. Evans and P. Bryant, "LHC Machine," *Journal of Instrumentation*, vol. 3, no. 08, p. S08001, 2008.

[23] G. Aad *et al.*, "The ATLAS Experiment at the CERN Large Hadron Collider," *Journal of Instrumentation*, vol. 3, no. 08, p. S08003, 2008.

[24] ATLAS Collaboration, "ATLAS Fact Sheets." http://atlasexperiment.org/fact_sheets.html, 2006. [Online; Accessed November 2016].

[25] G. Aad *et al.*, "The ATLAS Inner Detector commissioning and calibration," *Eur. Phys. J.*, vol. C70, pp. 787–821, 2010.

[26] ATLAS Collaboration, "Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ TeV," Tech. Rep. ATL-PHYS-PUB-2015-018, CERN, Geneva, Jul 2015.

[27] A. Ruiz-Martinez *et al.*, "The Run-2 ATLAS Trigger System," Tech. Rep. ATL-DAQ-PROC-2016-003, CERN, Geneva, Feb 2016.

[28] J. Alwall *et al.*, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *JHEP*, vol. 07, p. 079, 2014.

[29] R. D. Ball *et al.*, "Parton distributions for the LHC run II," *Journal of High Energy Physics*, vol. 2015, no. 4, p. 40, 2015.

[30] T. Sjostrand *et al.*, "An Introduction to PYTHIA 8.2," *Comput. Phys. Commun.*, vol. 191, pp. 159–177, 2015.

[31] T. Sjstrand, S. Mrenna, and P. Skands, "Pythia 6.4 physics and manual," *Journal of High Energy Physics*, vol. 2006, no. 05, p. 026, 2006.

[32] ATLAS Collaboration, "ATLAS Run 1 Pythia8 tunes," Tech. Rep. ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov 2014.

[33] R. D. Ball *et al.*, "Parton distributions with LHC data," *Nuclear Physics B*, vol. 867, no. 2, pp. 244 – 289, 2013.

[34] T. Sjstrand, S. Mrenna, and P. Skands, "A brief introduction to PYTHIA 8.1," *Computer Physics Communications*, vol. 178, no. 11, pp. 852 – 867, 2008.

[35] A. D. Martin *et al.*, "Parton distributions for the LHC," *The European Physical Journal C*, vol. 63, no. 2, pp. 189–285, 2009.

[36] ATLAS Collaboration, "Further ATLAS tunes of PYTHIA6 and Pythia 8," Tech. Rep. ATL-PHYS-PUB-2011-014, CERN, Geneva, Nov 2011.

[37] S. Agostinelli *et al.*, "Geant4–a simulation toolkit," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250 – 303, 2003.

[38] J. R. Andersen *et al.*, "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties," CERN-2013-004, FERMILAB-CONF-13-667-T. 2013.

[39] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-kt jet clustering algorithm," *Journal of High Energy Physics*, vol. 2008, no. 04, p. 063, 2008.

[40] ATLAS Collaboration, "Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV," Tech. Rep. ATL-PHYS-PUB-2015-015, CERN, Geneva, Jul 2015.

[41] G. Aad *et al.*, "Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector," CERN-PH-EP-2015-206. 2015.

[42] ATLAS Collaboration, "Selection of jets produced in 13 TeV proton-proton collisions with the ATLAS detector," Tech. Rep. ATLAS-CONF-2015-029, CERN, Geneva, Jul 2015.

[43] ATLAS Collaboration, "Tagging and suppression of pileup jets," ATLAS-CONF-2014-018. 2014.

[44] ATLAS Collaboration, "Expected performance of the ATLAS b-tagging algorithms in Run-2," Tech. Rep. ATL-PHYS-PUB-2015-022, CERN, Geneva, Jul 2015.

[45] ATLAS Collaboration, "Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run," Tech. Rep. ATL-PHYS-PUB-2016-012, CERN, Geneva, Jun 2016.

[46] A. Buzatu, "Decoupled b-jet-energy corrections (muon-in-jet, neutrino-in-jet, and

transverse momentum) for $H \to b\bar{b}$ searches for ATLAS in early Run II," Tech. Rep. ATL-COM-PHYS-2016-035, CERN, Geneva, Jan 2016.

[47] ATLAS Collaboration, "Early Inner Detector Tracking Performance in the 2015 data at $\sqrt{s} = 13$ TeV," Tech. Rep. ATL-PHYS-PUB-2015-051, CERN, Geneva, Dec 2015.

[48] M. Aaboud *et al.*, "Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data," CERN-EP-2016-110. 2016.

[49] ATLAS Collaboration, "2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data," Tech. Rep. ATL-DAQ-PUB-2016-001, CERN, Geneva, Mar 2016.

[50] M. Backes *et al.*, "Electron/photon trigger efficiency plots for ICHEP2016," Tech. Rep. ATL-COM-DAQ-2016-086, CERN, Geneva, Jul 2016.

[51] A. Hoecker *et al.*, "TMVA: Toolkit for Multivariate Data Analysis," *PoS*, vol. ACAT, p. 040, 2007.

[52] W. Verkerke and D. P. Kirkby, "The RooFit toolkit for data modeling," *eConf*, vol. C0303241, p. MOLT007, 2003. [,186(2003)].

[53] T. Skwarnicki, "A study of the radiative cascade transitions between the upsilon-prime and upsilon resonances," DESY-F31-86-02.

[54] M. Dobbs and J. B. Hansen, "The HepMC C++ Monte Carlo event record for High Energy Physics," *Comput. Phys. Commun.*, vol. 134, pp. 41–46, 2001.

[55] M. Aaboud *et al.*, "Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC," CERN-EP-2016-117. 2016.

[56] G. Aad *et al.*, "Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector," *Eur. Phys. J.*, vol. C75, p. 17, 2015.

[57] G. Aad *et al.*, "Performance of b-Jet Identification in the ATLAS Experiment," *JINST*, vol. 11, no. 04, p. P04008, 2016.

[58] C. Anastopoulos *et al.*, "Photon identification pre-recommendations for run 2," Tech. Rep. ATL-COM-PHYS-2015-496, CERN, Geneva, Jun 2015.

[59] T. J. Khoo *et al.*, "Performance of algorithms that reconstruct missing transverse momentum in $\sqrt{s} = 8$ TeV proton–proton collisions in the ATLAS detector," Tech. Rep. ATL-COM-PHYS-2015-341, CERN, Geneva, Apr 2015.

[60] A. L. Read, "Presentation of search results: the CLs technique," *Journal of Physics G: Nuclear and Particle Physics*, vol. 28, no. 10, p. 2693, 2002.

[61] W. Lampl *et al.*, "Calorimeter Clustering Algorithms: Description and Performance," Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr 2008.

Appendix A

Truth Analysis

A truth-level analysis has been implemented to evaluate theoretical uncertainties on the MC samples due to renormalization and factorization scale uncertainties, PDF uncertainties, and parton shower uncertainties. In this Appendix, the details of the truth-level analysis are provided, along with a validation.

- Photon Selection
  - TruthPhotons collection
  - $p_T > 30\,\text{GeV}$
  - $|\eta| < 2.5$
  - exclude $|\eta| > 1.37$ and $|\eta| < 1.52$ (crack region)
- Jet selection
  - AntiKt4TruthWZJets collection

  - $p_T > 40\,\text{GeV}$
  - $|\eta| < 4.5$

- Overlap removal
  - Remove jets overlapping with photons in $\Delta R(j, \gamma) < 0.2$
  - Remove photons overlapping with jets in $0.2 \leq \Delta R(j, \gamma) < 0.4$

- Signal jet selection
  - Sort jets by TruthLabelDeltaR_B (smallest first)
  - Pick the first two jets in $|\eta| < 2.5$ as the signal jets
  - Pick the two remaining jets with the largest invariant mass as the VBF jets

- Event selection
  - Require at least 1 photon and 4 jets (2 of which must be central)
  - Require both signal jets to have TruthLabelDeltaR_B<0.4
  - $m_{jj} > 800\,\text{GeV}$
  - $p_T^{b\bar{b}} > 80\,\text{GeV}$

To validate the truth selection procedure, a simple cut-flow has been performed and compared with the corresponding cuts at reco-level. The result is shown in Table A.

Columns 3 and 5 in Table A show the efficiency for each cut. While there is relatively good agreement in the selection efficiency for the m_{jj} and $p_T^{b\bar{b}}$ cuts between the truth-level and reco-level analyses, there is a discrepancy at pre-selection level and

Cut	N Truth Events	% of previous (truth)	N Reco Events	% of previous (reco)
Initial	1250000	-	768000	-
1 photon and 4 jets	148856	11.9	64941	8.5
2-tag	66471	44.7	18728	28.8
$m_{jj} > 800\,\text{GeV}$	30340	45.6	8977	47.9
$p_T^{BB} > 80\,\text{GeV}$	24098	79.43	7375	82.1

Table A.1: Cut flow table for truth-level and reco-level analyses. Columns 3 and 5 show the efficiency for each cut.

another due to b-tagging. The difference due to b-tagging is expected since the reco-level analysis uses the MV2c10 multivariate discriminant, while the truth-level analysis simply does ΔR matching to B-hadrons. We also see a sizeable difference due to the pre-selection requirement of having at least four jets ($p_T > 40\,\text{GeV}$) and one photon ($p_T > 30\,\text{GeV}$). As changes in the renormalization and factorization scales may affect the jet p_T spectrum, we need to check whether changes at truth-level would propagate similarly to reco-level. To do this, we perform a new selection with the jet p_T threshold raised to 45 GeV, and compare the change in acceptance at truth-level and reco-level. The result is shown in Table A.

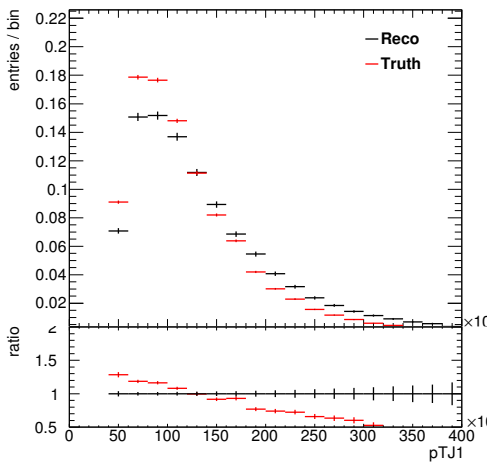
Cut	N Truth ($p_T^j > 40$)	N Truth ($p_T^j > 45$)	N Reco ($p_T^j > 40$)	N Reco ($p_T^j > 45$)	$\frac{p_T^j>45}{p_T^j>40}$ (truth)	$\frac{p_T^j>45}{p_T^j>40}$ (reco)
Initial	1250000	1250000	768000	768000	-	-
1 photon and 4 jets	148856	108050	64941	44029	72.6	67.8
2-tag	66471	48306	18728	13392	72.7	71.5
$m_{jj} > 800\,\text{GeV}$	30340	22942	8977	6797	75.6	75.7
$p_T^{bb} > 80\,\text{GeV}$	24098	19412	7375	5868	80.6	79.6

Table A.2: Change in yields at both truth-level and reco-level level when raising the jet p_T threshold from 40 to 45GeV. Columns 6 and 7 show the fraction of events retained for each cut with the higher threshold.

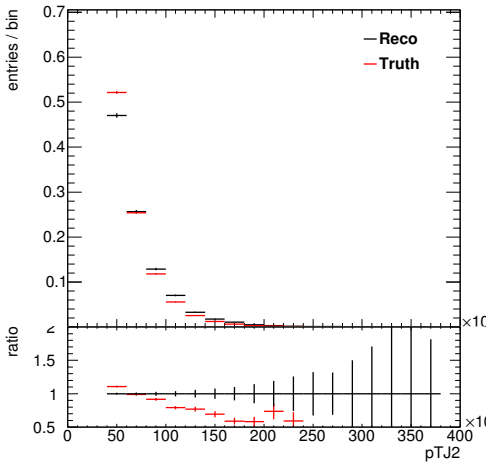
The last two columns in Table A show the fraction of events retained for each cut with the higher jet p_T threshold. The numbers are very similar for both truth-level and reco-level, thus suggesting that changes in the jet p_T spectrum at truth level due

to scale variations would propagate similarly to reco-level.

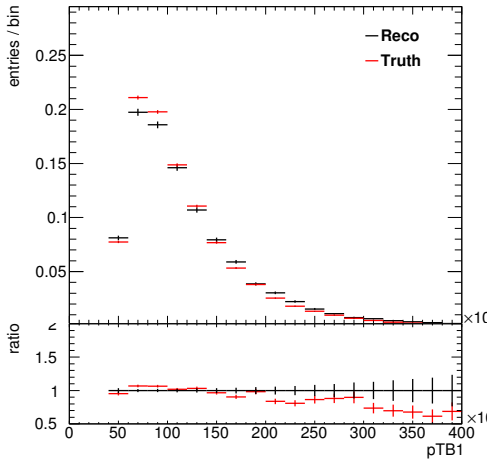
Lastly, the shape of many event variables have been compared between truth-level and reco-level. This comparison is done at "2-tag level", so the m_{jj} and $p_T^{b\bar{b}}$ cuts have not been applied. These plots are shown in Figures A.1, A.2, A.3, A.4, A.5, and A.5. The shapes of all variables are very similar between truth-level and reco-level level with the exception of p_T^{balance} and the p_T of the objects.



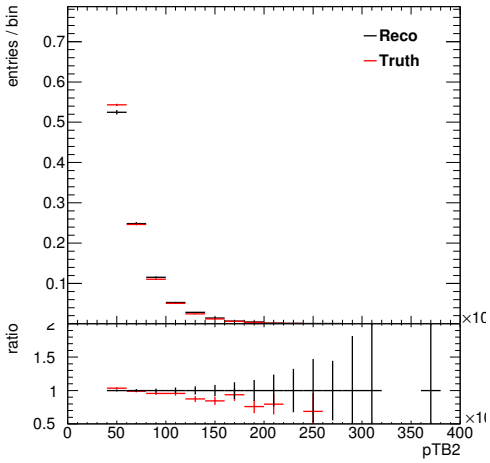
(a) $p_T^{j_1}$



(b) $p_T^{j_2}$

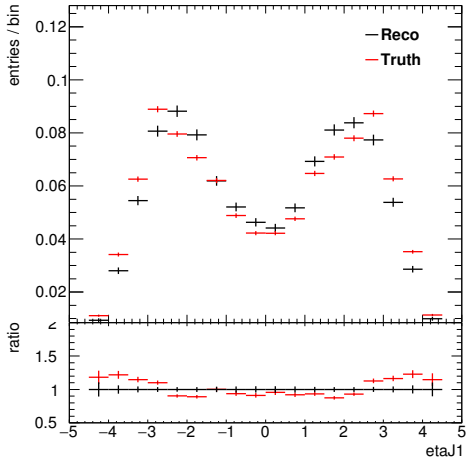


(c) $p_T^{b_1}$

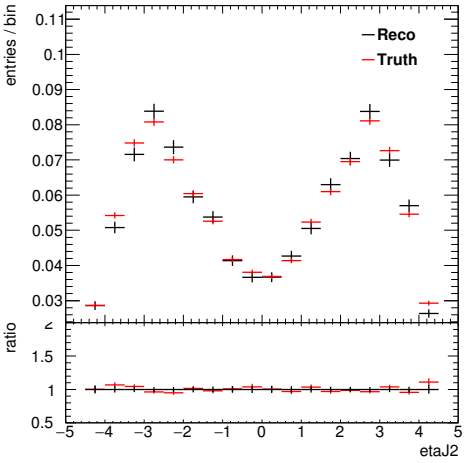


(d) $p_T^{b_2}$

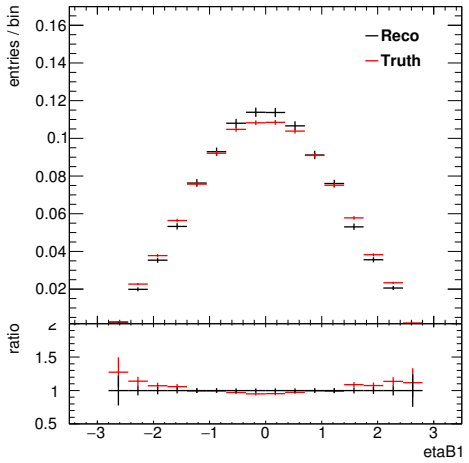
Figure A.1: Comparison of jet p_T between truth-level and reco-level analyses.



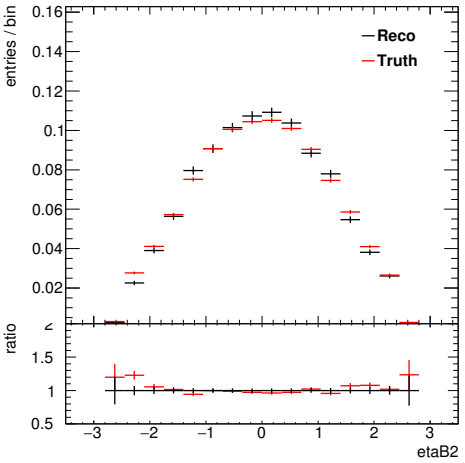
(a) η_{j_1}



(b) η_{j_2}

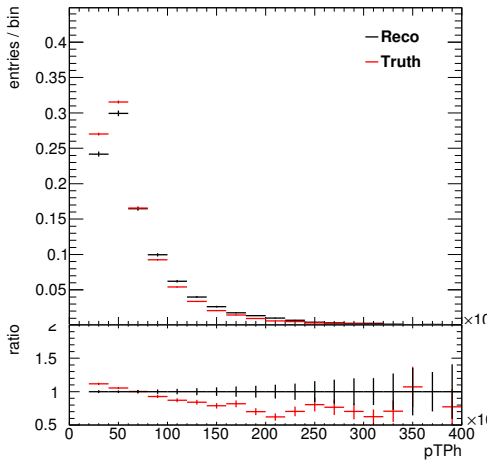


(c) η_{b_1}

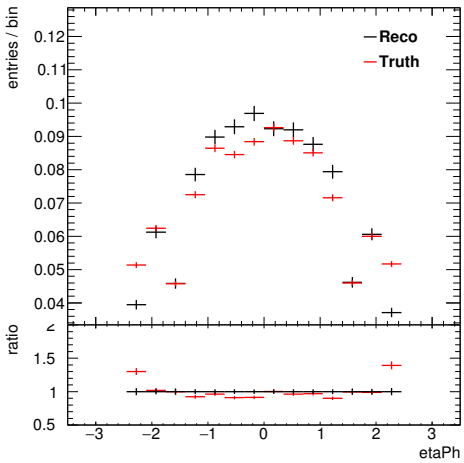


(d) η_{b_2}

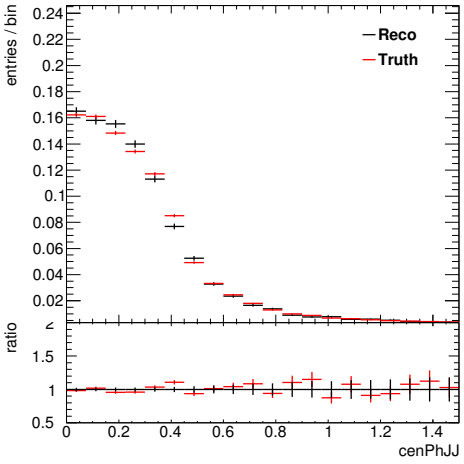
Figure A.2: Comparison of jet $|\eta|$ between truth-level and reco-level analyses.



(a) p_T^γ

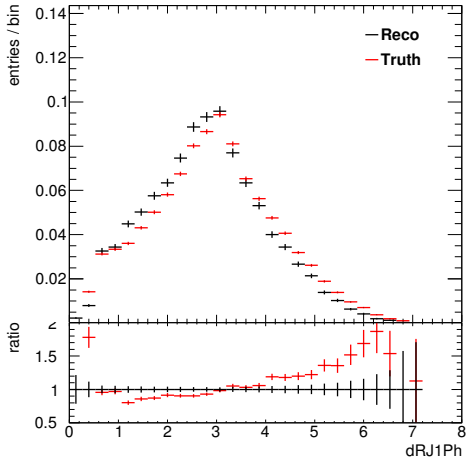


(b) η_γ

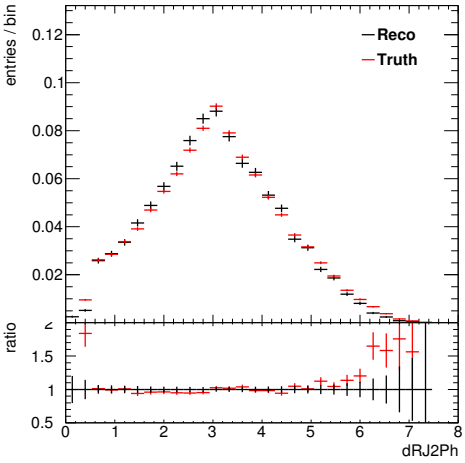


(c) centrality(γ)

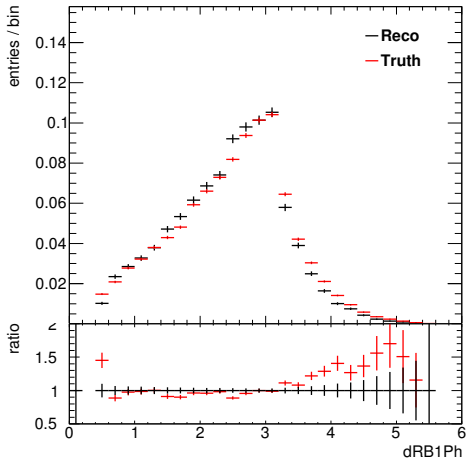
Figure A.3: Comparison of photon kinematics between truth-level and reco-level analyses.



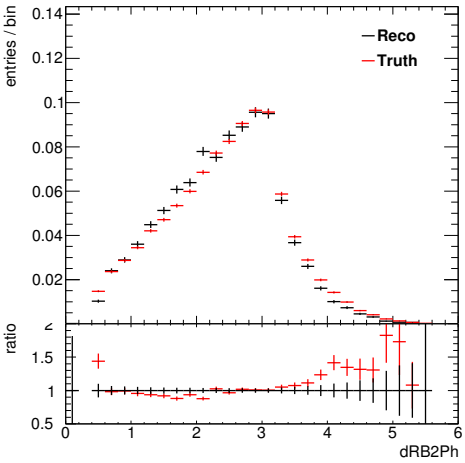
(a) $\Delta R(j_1, \gamma)$



(b) $\Delta R(j_2, \gamma)$

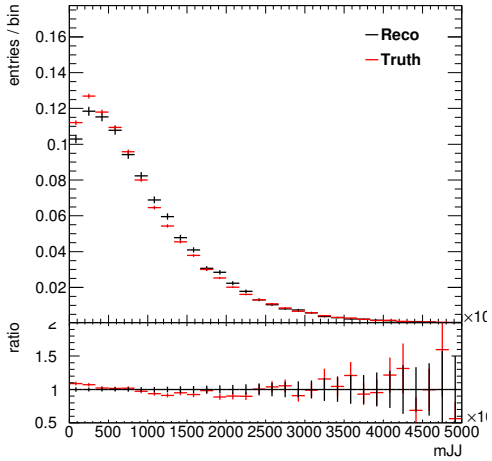


(c) $\Delta R(b_1, \gamma)$

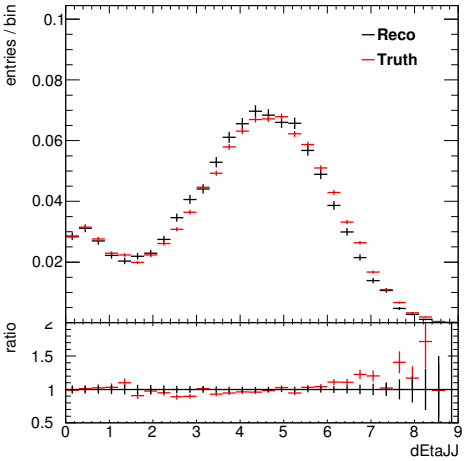


(d) $\Delta R(b_2, \gamma)$

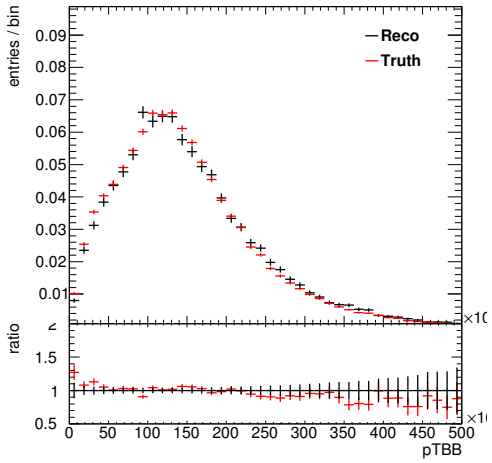
Figure A.4: Comparison of $\Delta R(jet, \gamma)$ between truth-level and reco-level analyses.



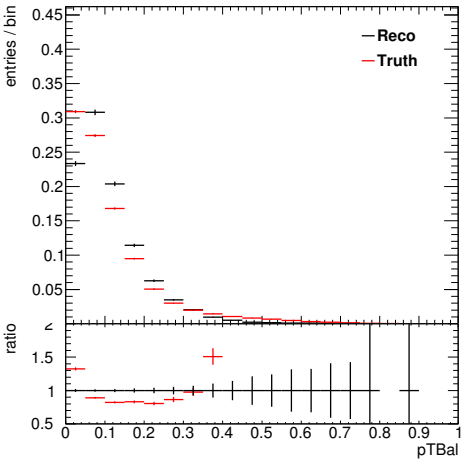
(a) m_{jj}



(b) $|\Delta\eta(jj)|$



(c) $p_T^{b\bar{b}}$



(d) p_T^{balance}

Figure A.5: Comparison of event variables between truth-level and reco-level analyses.