UNIVERSITY OF CALIFORNIA

SANTA CRUZ

SEARCH FOR THE DECAY OF HIGGS BOSONS AT HIGH TRANSVERSE MOMENTUM 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

Cole M. Helling

September 2021

The Dissertation of Cole M. Helling
is approved:

Professor Jason Nielsen, Chair

Professor Mike Hance

Professor Bruce Schumm

Peter Biehl
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

Search for the Decay of Higgs Bosons at High Transverse Momentum in the $H \to b\bar{b}$ Channel with the ATLAS Detector

by

Cole M. Helling

Presented in this dissertation is a study of Higgs boson production with high transverse momentum and decaying to a pair of bottom quarks. The analysis uses the pp-collision data collected by the ATLAS Detector at the Large Hadron Collider between 2015 and 2018, corresponding to a total integrated luminosity of $136\,\text{fb}^{-1}$. The decay products of the Higgs boson candidate are reconstructed into a single large-radius jet, using the experimental signature of two subjets identified as coming from b-hadrons. The $Z \to b\bar{b}$ cross section is extracted in the same kinematical regions and compared to Standard Model predictions to validate the analysis methods. The large-radius jets receive a mass and momentum correction to account for losses due to heavy hadron decays to muons and neutrinos by adding the muon four-momenta to that of the jet. The template normalizations of the top-quark-related contributions in the signal region are determined by a simultaneous fit to data in a control region rich in semileptonic $t\bar{t}$ events. In the log likelihood fit used to extract the signal strengths, an external constraint, derived from $t\bar{t}$ events with a resolved Wb pair from a hadronically decaying top quark in a dedicated control region, is applied to the jet mass resolution systematic uncertainty nuisance parameter. This external constraint serves to stabilize the fit response and reduce the correlation between the extracted Z signal strength and jet mass resolution parameter. For Higgs boson

production at transverse momentum above 450 GeV, the production cross section is found to be $13 \pm 57\,(\text{stat.}) \pm 22\,(\text{syst.}) \pm 3\,(\text{theo.})$ fb. The differential cross section 95% confidence level upper limits, as a function of the Higgs boson transverse momentum, are $\sigma_H(300 < p_{\text{T}}^H < 450\,\text{GeV}) < 2.9\,\text{pb}$, $\sigma_H(450 < p_{\text{T}}^H < 650\,\text{GeV}) < 89\,\text{fb}$, $\sigma_H(650 < p_{\text{T}}^H < 1000\,\text{GeV}) < 39\,\text{fb}$, and $\sigma_H(p_{\text{T}}^H > 1\,\text{TeV}) < 9.6\,\text{fb}$. All results are consistent with Standard Model predictions.

To my nieces Arianna, Keeli, Eva, and Olivia, and nephews Mason and Holden

My goal in life has been to set an example for you. Continuous effort is far more important to success than base talents or intelligence.

To my twin sister Sunnie and my little sister Mallory

Sunnie, you have a rather unique ability to abstain from forming opinions when you don't have enough information, nurture that and never stop being curious. Mallory, you've come so far in life and you've grown into a person I respect a great deal. I love you both and hope I can live up to the image you have of me.

To my parents, Ken and Lisi

The lessons I've learned have been invaluable. Because of you, I have determination, grit, and a sense of purpose. My childhood instilled in me a drive to succeed through perseverance, thank you. Dad, you can start calling me doctor any day now.

Acknowledgments

I want to thank my advisors: Professors Jason Nielsen and Marco Battaglia. Their guidance in writing this dissertation (mostly) prevented it from becoming an introductory on the Standard Model. Jason's excellent attention to detail and language, along with Marco's big-picture view helped me craft the largest written work of my life; I thank you both greatly for your patience and knowledge over the years.

To my other PhD committee members, Professors Mike Hance and Bruce Schumm, I thank you for helping me over the years. I thank you in advance for reading my dissertation, I know it's long; I hope you see it as a reflection of my interests in particle physics.

Thank you to Tony Affolder and Vitaliy Fedeyev. You were my advisors and friends as I was introduced to Experimental HEP. Vitaliy's approach to research is relentless, always taking a study another step further, squeezing out every ounce of useful knowledge ... even from broken sensors. Tony's encyclopedia-level knowledge of the experiment and where I fit in always helped me see the wider scope of my work. I once told his wife (Hi Kirsten!) the first time I met her that talking to Tony is like drinking from a fire hose; I stand by that. From both of you I learned a great deal, thank you.

I must thank UCSC's ATLAS postdocs: Andrea Sciandra, Matthew Gignac, and Giordon Stark. The three of you have been endlessly kind, answering slews of questions from physics to coding. Your immense competencies have made me realize just how far I have to go in the coming years; thanks for the massive inferiority complex. I also want to thank Luise Poley at UBC. We've worked quite closely on

a number of topics over the years, and I have great respect for you. I look forward to working with you again soon as I turn this next chapter; I leave all glue-related topics to you.

Thank you to my close friends Nico Canzano and Logan Morrison. I appreciate your comments while I hammered out this behemoth. Nico, your comments regarding language use were incredibly helpful. Your friendship over the last six years has improved my life. I almost forgive you for choosing music over physics. Logan, your help with my theory sections saved the day. I hope one day to have a tenth your talent. Have fun being rich someday.

To the rest of my friends in the UCSC grad department, thank you for your conversations and friendship. You'll always be welcome at social hour.

Lastly, I want to thank my little sister Mallory and her husband Brent. I spent the last several months writing this in their basement. Thanks for dealing with the fact that I was always "just two weeks away" from finishing. Thank you, I love you both.

This page has been intentionally left blank.

This one was an accident.

Chapter 1

Introduction

My dissertation work is based on a measurement of the production rate of the Higgs boson decaying to a pair of bottom quarks at high transverse momentum (high p_{T}) using pp-collision data from the ATLAS detector [1, 2]. Interest in the Higgs at high p_{T} has existed for decades [3–5], but it has only been realizable in the last decade or so with the introduction of new, exceptionally powerful colliders. The purpose behind such a search is not only to better understand the role played by the Higgs boson in the Standard Model and to more precisely measure the coupling of the Higgs boson to bottom quarks, but to press further and illuminate the true nature of physics that lies beyond our current knowledge.

The Standard Model of Particle Physics [6–17] is the successful model which incorporates the known matter particles and how they interact, but it has only recently been tested at such high energies. This is not a simple endeavour and a program to study the Higgs boson at high transverse momentum is only possible at the incredible energies of the Large Hadron Collider. The Large Hadron Collider (LHC) [18, 19] at CERN [20] is the largest and most powerful collider ever constructed. There, thousands of physicists and engineers from around the globe rigorously test the Stan-

dard Model at (terrestrially) unprecedented energies. The Higgs boson, discovered in 2012 at CERN [21, 22], answers the problem of the origin of mass for the fundamental particles. But analogous to the way Newtonian Mechanics is seen as a low energy approximation to the more complete General Relativity, the Standard Model is almost certainly a low energy approximation to some more encompassing theory of reality. I therefore think it is appropriate to ask as to whether the Higgs boson, a particle found at the current energy limitations of our most powerful collider, could be used to search for that more complete theory?

To realize the goal of my analysis, we look for a specific signature in which the Higgs boson is boosted to a high transverse momentum by recoiling from another particle produced in the original hard-scatter event, causing the decay products, a pair of b-quarks, to be detected close together in the detector. Though there are several analyses that have performed searches for $H \to b\bar{b}$ in specific production channels such as VH [23, 24], VBF [25, 26], and ttH [27, 28][1], the Boosted $H \to b\bar{b}$ Analysis searches this decay channel inclusively in terms of production channels. In particular, it permits access to the gluon-gluon fusion (ggF) channel, the most common way to produce the Higgs boson. This production channel is particularly difficult due to an unforgivingly high background. Through the combination of this analysis and the measurements from the various analysis channels mentioned above, we can obtain more precise measurements of the $H \to b\bar{b}$ partial decay width, a dominant contribution to the total Higgs boson decay width. At high transverse momentum, the virtual particle loop in ggF responsible for Higgs production becomes sensitive to possible new particles and anomalous couplings, yielding possible hints at new physics seen as deviations between the observed and predicted rates. This

[1] These production modes will be explained in more detail in Section 3.1.

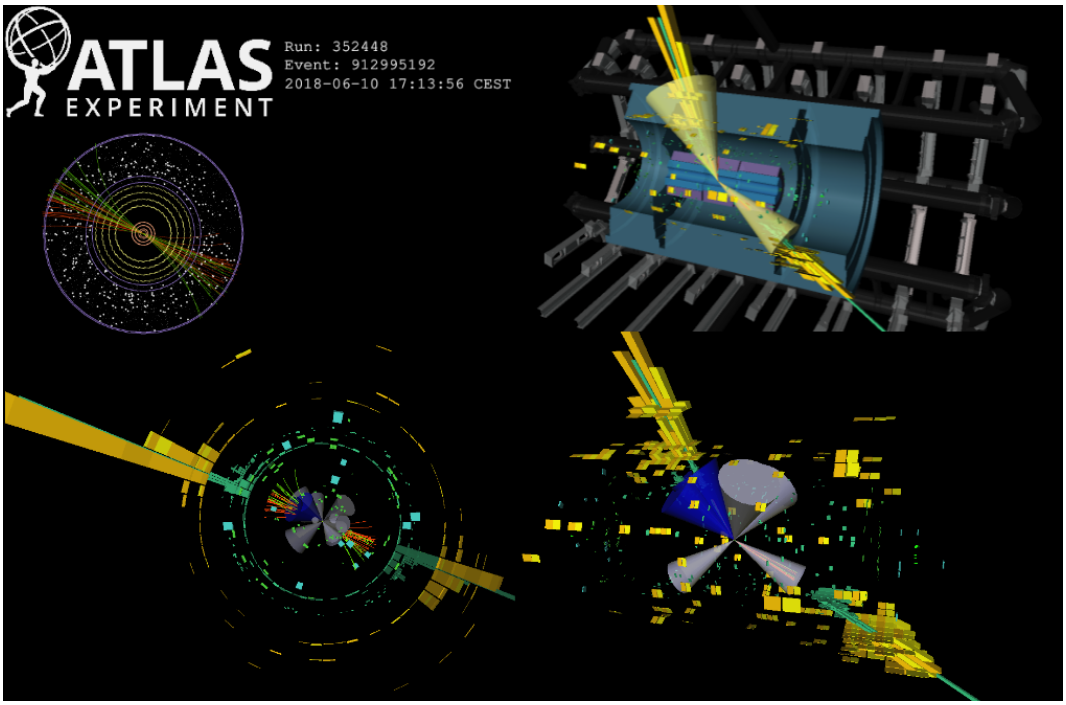


Figure 1.1: Event detected in the ATLAS detector, passing the selection criteria for the leading signal region (SRL) in the Boosted $H \to b\bar{b}$ Analysis for data taken in 2018. The transverse momentum and mass for the candidate jet is 1029 GeV and 121 GeV, respectively. The top left image shows a cutaway of part of the ATLAS detector and two large radius jets; one the Higgs candidate jet (top) and the other the recoil jet (bottom). Shown in the bottom right image are the two b-tagged subjets (blue) and several other non b-tagged jets. The two images on the left are cross-sectional views pointed along the beamline, showing reconstructed tracks (red and green represent charged tracks while yellow are neutral), hits in the calorimeters (green and yellow bars), and small radius jets (blue and gray).

is a particularly challenging task, as there are several other processes with similar final states which constitute a substantial background. At a very basic level, we are looking for two large deposits of energy back-to-back in the detector – one of them the Higgs boson decaying to b-quarks, an example of which can be seen in Figure 1.1.

To interpret the results of this analysis, I begin Chapter 2 by covering our current knowledge of the particles and their interactions in the Standard Model. I will then move on to discuss quantum field theory, the guiding framework for the Standard Model. I will apply this framework to explain electroweak unification to show how the introduction of the Higgs field is essential to understanding how the

masses are obtained by the fundamental particles. In Chapter 3, I will cover Higgs boson production at the LHC, the most prominent decay channels, as well as the latest combined measurements from ATLAS and CMS. Then, how we might use the Higgs boson to discover physics beyond the Standard Model. I will outline the Boosted $H \to b\bar{b}$ channel in Chapter 4 to introduce some of the topics addressed later in the text, highlighting some of the challenges presented in this search. Beyond theoretical considerations, we will see the role played by the Large Hadron Collider, the ATLAS Detector, and its subdetectors in Chapter 5.

In Chapter 6, I will cover how the ATLAS detector uses a robust set of measurements coupled with advanced pattern recognition algorithms to reconstruct hits in the detector into useful physics objects such as muons and jets. The Boosted $H \to b\bar{b}$ Analysis relies on the b-tagging of jets to identify the decay products from the Higgs boson, so I will cover in some detail the algorithms used in identifying the flavor of jets, with emphasis on those used in my analysis. From there, with the requisite knowledge of how the detector allows us to record the data and reconstruct objects for analysis, we will discuss in detail the Boosted $H \to b\bar{b}$ Analysis starting in Chapter 7. The analysis classifies events into various categories (Chapter 8): signal regions for the Higgs boson, control regions for additional constraints, and validation regions for studying the QCD background fit model. Chapter 9 discusses the Monte Carlo simulation and data used in the analysis, including a detailed discussion of the jet mass resolution constraint. The parameters of interest (the Higgs and Z boson rates) are extracted in a likelihood fit which includes the systematic uncertainties from a variety of sources (Chapter 10). I will report the results of the analysis in Chapter 11, and conclude with comparisons between the ATLAS and CMS results

in Chapter 12.

Chapter 2

Standard Model

The Standard Model [6–17] (SM) is an exceptionally successful model describing all of the known particles and their interactions with the exception of the gravitational force. While general relativity [29] successfully describes gravity, shedding light on the movements of planets, stars, and galaxies over vast distances (from millions of kilometers to millions of light-years), the Standard Model delves into the world of the incredibly tiny, with distances on scales of trillionths of a meter and smaller.

As this dissertation revolves primarily around the Higgs boson, I will focus on the unification of the electromagnetic and weak forces and how electroweak symmetry had to be broken through the Higgs Mechanism to imbue the fundamental particles with mass. Though for completeness, and to retrace the path I've taken to understand the SM, I include for the interested reader a section on quantum electrodynamics (QED), the first successful quantum field theory describing the electromagnetic interaction in Appendix A.6. I also include quantum chromodynamics (QCD), the theory describing the strong nuclear force which glues quarks together within the nuclei of atoms, in Appendix A.7.

2.1 Particles

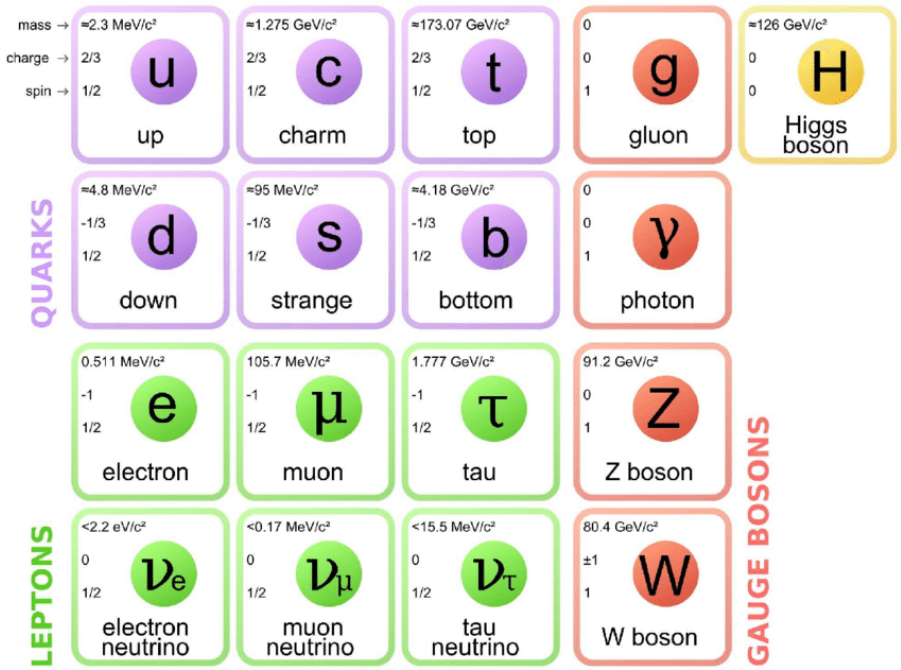


Figure 2.1: Table of the Standard Model particles and their properties [30].

Before discussing the immense complexity, yet beautiful simplicity of the Standard Model (SM), it's helpful to know the players of the game; that is, the particles themselves. Much like how the Periodic Table of Elements, which contains the *fundamental constituents* of chemistry, arranged by their chemical properties and atomic number, we organize the *fundamental particles* based on their quantum numbers such as spin, charge, and mass. The particles of the SM, as seen in Figure 2.1, are placed into two main categories by their intrinsic spin properties, and hence the statistics they obey, called fermions and bosons. Fermions have half-integer spins and obey the Pauli Exclusion Principle [31] so that no two fermions can have the same wave function. They make up all of the everyday matter which we are all accustomed to, formed primarily from electrons and up/down quarks. Bosons have integer spins, with the spin-1 vector bosons comprising the carriers of the electromagnetic, weak,

and strong interactions. The spin-0 boson, known as the Higgs boson, is the only fundamental scalar in the Standard Model. Through interactions with the non-zero value of the Higgs field, the fundamental particles obtain their mass. In addition to the particles in the figure, each particle has an anti-particle with the opposite quantum numbers with the exception of the mass.

With respect to the weak nuclear interaction, the fermions are organized into weak isospin doublets for left-handed (spin and motion in the same direction) and isospin singlets for right-handed (spin and motion in opposite direction) particles. Fermions also come in three generations, with the first generation making up all of the "ordinary" matter. The second and third generations share identical properties to the first except that their masses increase substantially with each generation. We can separate elementary fermions into two groups; leptons and quarks.

The charged leptons, shown at the bottom left of Figure 2.1, come in three types, or *flavors*: the electron e, muon μ, and tau τ with $m_e < m_\mu < m_\tau$, all having an electric charge of -1 (in units of the fundamental electric charge, e). For each flavor of charged lepton, there is a corresponding neutrally charged neutrino: ν_e, ν_μ, and ν_τ. Only left-handed neutrinos have been observed in experiments [32], which means neutrinos cannot get their mass from the Higgs field in the SM [33]. However, experimental evidence from neutrino oscillation experiments [34] show that they do have mass. The charged leptons interact both electromagnetically (because they have electric charge) as well as weakly. Neutrinos are electrically neutral and only participate in weak interactions. The leptons form weak isospin doublets consisting of a charged lepton (a down-type) and its corresponding flavor neutrino (an up-type). Often, we refer to the charged leptons simply as leptons ℓ, with the neutral

leptons as neutrinos ν_ℓ.

The six quarks, shown in the top left of Figure 2.1, are arranged in three generations, with each generation in a weak isospin doublet containing an up-type (up u, charm c, top t) quark with a fractional charge of $+2/3$, and a down-type (down d, strange s, bottom b) quark with charge $-1/3$. The quarks interact not only electromagnetically and weakly, but strongly as well. Due to the nature of the strong force, isolated quarks are never observed; they are typically[1] bound in bosonic quark-antiquark pairs (mesons) or fermionic quark/antiquark triplets (baryons) with integer units of electric charge.

The vector gauge bosons, shown on the right side of Figure 2.1, are the force carriers for the three fundamental forces of the SM. The gluons[2] (g) are the massless force carriers of the strong nuclear force, which act only over an incredibly small distance ($\mathcal{O}(10^{-15}\,\mathrm{m})$), holding quarks together and maintaining the cohesion of protons and neutrons within the nuclei of atoms. Particles participating in the strong interaction have three charges (red, green, and blue), which are referred to *color* because the combination of charges in mesons and baryons combine to produce a neutral (or white) strong charge. The most familiar of the vector gauge bosons is the photon (γ), a massless particle responsible for mediating the electromagnetic force between charged particles, which acts over an infinite range. The W and Z particles are the massive vector bosons which mediate the weak nuclear force, responsible for the radioactive decay of atoms and essential to processes such as the nuclear fission within the Sun. The large masses of the W and Z bosons ensure that the weak force

[1] There have also been recent discoveries of the penta-quark [35] (objects containing four quarks and an anti-quark) and tetra-quark [36] (two quarks and two anti-quarks) by the LHCb Collaboration.

[2] Oddly, there was a recent joint announcement by the TOTEM [37, 38] and DØ [39] Collaborations confirming the discovery of a "glueball", the odderon [40, 41], which is a state consisting of an odd number of gluons.

only acts over small distances. The Higgs boson, the only spin-0 (scalar) and most recently discovered particle of the SM, gives mass to the fermions and massive vector bosons through the interaction with the Higgs field.

While the familiar particles (the electron, photon, protons and neutrons containing up/down quarks) were discovered by the early and mid-twentieth century, the more massive ones were discovered more recently with the advent of higher-energy colliders. The massive vector gauge bosons, the W [42, 43] and Z [44, 45] bosons, were discovered in 1983 by the UA1 [46] and UA2 [47] experiments at CERN. The bottom quark was discovered in the E288 experiment at Fermilab [48] in 1977 [49], while the top quark wasn't discovered until 1995 by the DØ [50] and CDF [51] Collaborations [52, 53]. The Higgs boson, the last particle of the SM, was discovered in 2012 by the ATLAS and CMS Collaborations at CERN [21, 22]. A full description of the SM particles and their properties can be found in the Review of Particle Physics by the Particle Data Group (PDG) [54] and a list of the fundamental particles, their discovery dates, and original references can be found in Table A.1 of Appendix A.1. Noticeably missing from our list of particles is a force carrier for the gravitational interaction. Though a massless spin-2 tensor boson has been proposed as the mediating particle, it has yet to be discovered[3]. As it stands, the Standard Model of Particle Physics is very robust and most of its parameters have been measured with high precision [54].

To understand the underlying theory of the Standard Model, as well as its consequences, it's necessary to understand the symmetries upon which the theory rests. In the following sections, we will review how conserved quantities result from

[3]The discovery of gravitational waves produced by a binary black hole system by the LIGO and Virgo Collaborations [55] in 2015 are very exciting confirmations of the theory of General Relativity. Whether or not they reveal a quantum nature to gravity still remains to be seen [56].

symmetries and how quantum gauge theory is used to construct the Standard Model.

2.2 Quantum Field Theory

We have expounded the basic properties of our fundamental particles, but have neglected to make clear what a particle actually *is*. While physicists often treat particles as tiny ball-like objects in scattering experiments, where such approximations are appropriate, the mathematical framework of the Standard Model does not treat particles as little hard balls of "stuff," but rather as excitations of quantum fields. A field is a space-time dependent quantity which can be classified as a scalar, vector, spinor, or tensor. Each particle type has its own quantum field and new fields are formed by interactions between these fields. The mathematical description of the Standard Model is laid out in quantum field theory (QFT), which describes all of the known constituents of matter and their interactions. Despite this possibly misleading terminology, it is common to use the term particle interchangeably with its corresponding field excitation.

Born from quantum mechanics in the 1920s out of a desire to create a quantum theory compatible with relativity, QFT provided a limited, yet powerful mathematical framework for making predictions. Electromagnetism was the first force to be treated in this manner with the quantization of the electromagnetic field in quantum electrodynamics (QED) [57]. Among its early successes were an accurate description of the Lamb Shift [58] and the anomalous magnetic moment of the electron $(g_e - 2)$[4] [61]. Despite remarkable triumphs, it suffered crippling pathologies. During its formative years in the early 1930s, it developed a penchant for producing

[4]Recent results at Fermilab [59] have announced a $3.3\,\sigma$ SM deviation ($4.2\,\sigma$ when combined with the previous results at Brookhaven [60]) of the muon's anomalous magnetic moment.

divergences (infinities) in calculations, no doubt leading to many sleepless nights for the developers of the theory. Even with this apparent road to perdition, the use of QFT continued next to the theory of beta decay (caused by the weak interaction), and later the strong interaction. In the late 1940s, through the work of Freeman Dyson [62], Richard P. Feynman [63], Julian Schwinger [64] and Sin-itiro Tomonaga [65], the procedure of renormalization was invented, allowing the infinities to be removed by a redefinition of a finite number of coupling constants and masses. QFT was on a much firmer footing by the 1950s and could be used to make predictions about new particles and interactions; by the 1960s, the electromagnetic and weak interactions were unified. Through spontaneous symmetry breaking, the Higgs field was introduced to explain the masses of fermions and vector gauge bosons. The work performed in the mid-twentieth century led quantum field theory to be a consistent relativistic generalization of the non-relativistic quantum mechanics.

While a full dive[5] into the subject of quantum field theory is unfortunately (or fortunately depending on your disposition) outside the scope of this work, I want to lay out the basics for QFT's implementation in the Standard Model to highlight and motivate the role the Higgs plays in rigorously testing the limits of the SM in the following sections.

2.2.1 Symmetries

Conservation laws had been discovered long before the first particle physicist stepped onto the scene. Conservation laws exist when nature observes certain symmetries, or operations that can be performed on a system that leave it unchanged. This

[5]For a richer appreciation of QFT, please visit your local library and check out the textbooks by Peskin and Schroeder [66] or Schwartz [67]

concept bears repeating; *physical laws are the consequence of symmetries.* For example, a sphere can be observed after being moved across the room (or planet), or after being rotated, or observed at a future time (the day after tomorrow or just before the inevitable heat death of the universe); the sphere's properties remain unchanged. More specifically, the laws of physics are invariant no matter when, where, or in what direction you observe. These three symmetries all relate to well-known conservation laws: Conservation of Momentum (spatial translation invariance), Conservation of Angular Momentum (rotational invariance), and Conservation of Energy (time translation invariance). This can also apply to more abstract symmetries, such as electric charge, weak isospin, or their combination, weak hypercharge (Section 2.3). The relationship between symmetries and conserved quantities was revealed by the exceptional mind of Emmy Noether in 1915[6] [68, 69] with the theorem that now bears her name. Noether's theorem states: *If a Lagrangian has a continuous symmetry, then there exists a current associated with that symmetry that is conserved when the equations of motion are satisfied.* For the interested reader, a brief section on the derivation of Noether's theorem is included in Appendix A.2.

2.2.2 Gauge Theory

Beyond the mathematical formalism of Lagrangian Mechanics[7], the SM relies heavily on gauge theory and the principle of gauge invariance. The freedom to shift or rotate our fields is called a gauge symmetry. This shifting and rotating doesn't occur in space-time, but rather in our definition of the field itself [70]. The symmetry of the system does not refer to the taking of one physical state to another, but instead refers

[6] Published in 1918 [68].

[7] It's possible to use Hamiltonian dynamics, but because the Lagrangian is Lorentz invariant by construction, it is far simpler.

to a mathematical *redundancy* in our description of the system when we move from a global symmetry to a more restrictive local symmetry. This inherent redundancy means that only quantities independent of the gauge can be considered physical; *i.e.* they **do not** depend on our choice of description. The principle of gauge invariance allows us to prevent the physical values from changing simultaneously with the transformation of the fields. To avoid over-counting the redundancies, we must *fix* the gauge, which is performed in a variety of ways depending on what problem we wish to solve. By fixing a gauge, the model often becomes something easier to analyze mathematically. Gauge field theory presents us with a general mechanism for restoring the gauge invariance to our system when we have a symmetry dependent upon space-time. For brevity (and a partial lack of a comprehensive knowledge on the subject), we will not discuss the complicated topic of Group Theory, but the interested reader can check out Reference [71]. For the sake of our discussion, we will use a rather simplified definition, found in Appendix A.4. Beyond that, our initial step is of course to determine the symmetry group, which can be laborious. Assuming we have done this however, we begin by defining some simple Lagrangian:

$$\mathcal{L} = |\partial_\mu \psi|^2 \tag{2.1}$$

with a field ψ that transforms as:

$$\begin{aligned} \psi(x) &\to e^{i\theta_a T^a}\psi(x) && \text{global symmetry} \\ \psi(x) &\to e^{i\theta_a(x) T^a}\psi(x) && \text{local symmetry,} \end{aligned} \tag{2.2}$$

where θ_a ($a \in 1, \ldots, N$) are the numbers which parameterize the group elements and T^a are the generators of the group, represented as Hermitian traceless matrices whose size depends on the symmetry. The exponential term $\exp(i\theta_a T^a)$ is a Unitary matrix U. In the global case, our parameter θ takes on the same value everywhere (*i.e.* globally). This is certainly the simpler case, as the space-time derivative will not act on it. For the local case, θ becomes a function of position and time, which adds a few solvable complications, one of which is that the derivative will act on both terms in Equation 2.2 according to the product rule, manifestly breaking invariance. We will solve this issue shortly by redefining how our gauge fields and derivatives transform. The symmetry must also be a continuous symmetry such that the derivative is defined everywhere and all values for $\theta \in [0, 2\pi]$ correspond to a different symmetry transformation. Once we've modified our symmetry from global to local, we need to define the gauge field whose transformation depends on $\theta(x)$ to ensure invariance under the local group transformation. Suppose we have the gauge fields A^a_μ; these transform as:

$$A^a_\mu \to U A^a_\mu U^{-1} + \frac{i}{g}(\partial_\mu U)U^{-1}, \tag{2.3}$$

where g is the coupling strength of the interaction. The quanta of these fields are also known as gauge bosons, one for each generator of the group. Next, the ordinary derivative operator must be replaced by the covariant derivative so that

$$D_\mu \psi \to e^{i\theta_a T^a} D_\mu \psi, \tag{2.4}$$

or more plainly, so that $D_\mu\psi$ transforms like ψ. The covariant derivative is defined as

$$D_\mu \equiv \partial_\mu + igT^a A^a_\mu \tag{2.5}$$

We can then define the field strength tensor $F_{\mu\nu}$, which is the mathematical object that describes the field in space-time, using the commutator for the covariant derivative

$$F_{\mu\nu} \equiv \frac{i}{g}[D_\mu, D_\nu]. \tag{2.6}$$

From here, following the Totalitarian Principle[8], we can construct our Lagrangian by introducing any terms involving the particle and gauge fields, as well as field strength tensor, provided they satisfy gauge invariance. Then, of the utmost importance, after constructing our theory we must test it in experiments. There will be subtleties for each symmetry group, but the steps outlined above will be our guiding light as we turn to the quantum field theory description of the Standard Model.

2.3 Electroweak Unification

One of the overarching themes in physics is unification. We seek not only to reduce the whole of nature to its base components, but to view these pieces as stemming from a more fundamental principle. James Clerk Maxwell was able to mathematically unite electric and magnetic interactions by recognizing similarities

[8]This principle generally states that everything allowed by the laws of nature must actually exist. It's a phrase often falsely attributed to Murray Gell-Mann and its origins and interpretations are often misstated [72].

in the two in the 1860s [73, 74]. Unifying the four known forces has been something of a Holy Grail for particle physicists for nearly a century, yet a Grand Theory of Everything remains just out of our grasp. Within the Standard Model, there has been success in unifying the electromagnetic and weak forces, coining the electroweak portmanteau. Many attempts were made throughout the 1950s to resolve the issues in weak gauge theory. Success was finally achieved in the 1960s through electroweak unification by Glashow [6], Salam [8], and Weinberg [7] and then electroweak symmetry breaking via the Higgs Mechanism (Section 2.4). In this section, we will cover unification before introducing symmetry breaking via the Higgs Mechanism in Section 2.4 to highlight the need for such a mechanism in the SM.

To incorporate both the electromagnetic and weak interactions, they proposed the $SU(2)_L \otimes U(1)_Y$ gauge group, where $U(1)_Y$ is associated with the weak hypercharge Y, much like the QED $U(1)$ is associated with the electromagnetic charge. Weak hypercharge is related to the electromagnetic charge Q and weak isospin T_3 by:

$$Q = T_3 + \frac{Y}{2}. \tag{2.7}$$

The fields consist of left-handed doublets ψ_L and right-handed singlets ψ_R, which transform as

$$\psi_L \to e^{i\beta_a(x)\tau^a + i\alpha(x)Y}\psi_L, \tag{2.8}$$

$$\psi_R \to e^{i\alpha(x)Y}\psi_R, \tag{2.9}$$

where $\beta_a(x)$ $(a = 1, 2, 3)$ parameterize the weak group elements and $\tau^a = \sigma^a/2$ are the generators of $SU(2)_L$, describing rotation between the members of the doublets with the Pauli spin matrices σ^a:

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.10}$$

As with QED, $\alpha(x)$ parameterize the group element of weak hypercharge, with Y the $U(1)_Y$ generator. To preserve invariance under this symmetry group, we obtain a gauge field for each generator, introducing four massless gauge fields. From $SU(2)_L$ we obtain three massless gauge fields W^a_μ $(a = 1, 2, 3)$, and from $U(1)_Y$ we obtain a single massless gauge field B_μ. These fields transform as

$$W^a_\mu \to W^a_\mu + \frac{1}{g}\partial_\mu \beta^a(x) - \epsilon^{abc}\beta^b(x)W^c_\mu, \tag{2.11}$$

$$B_\mu \to B_\mu + \frac{1}{g'}\partial_\mu \alpha(x). \tag{2.12}$$

The Levi-Civita symbol ϵ^{abc} is the $SU(2)$ structure constant, and g and g' are the coupling constants associated with $SU(2)_L$ and $U(1)_Y$, respectively. Additionally, to respect gauge invariance, we introduce the covariant derivative:

$$D_\mu = \partial_\mu + i\frac{g}{2}\tau_a W^a_\mu + i\frac{g'}{2}Y B_\mu. \tag{2.13}$$

This now allows us to define the field strength tensors for the W and B gauge fields:

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g\epsilon^{abc} W^b_\mu W^c_\nu,$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \tag{2.14}$$

Combining gauge-invariant terms involving the fermion and gauge fields, our Lagrangian becomes

$$\mathcal{L}_{SU(2)\times U(1)} = \sum_{\text{fermions}} i\overline{\psi}\not{D}\psi - \frac{1}{4}W^a_{\mu\nu}W^{\mu\nu}_a - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}, \tag{2.15}$$

Now that we have our Lagrangian, we can see that we still have massless bosons and fermions in our theory, in direct contradiction with expectations from experiment. How do we resolve such an issue when we know that adding explicit mass terms is *not* gauge invariant? For this, we need spontaneous symmetry breaking.

2.4 The Higgs Mechanism: Electroweak Symmetry Breaking

2.4.1 Breaking Symmetry

The solution to our electroweak woes seems to be best summed up as "If it don't work, break it." To give mass to the weak gauge bosons and fermions, we rely on a process generally called spontaneous symmetry breaking, and in our case called electroweak symmetry breaking (EWSB). A symmetry is broken when the vacuum state (or ground state) does not share the symmetry of the underlying physics. A simple way to introduce this is via the Higgs Mechanism, formulated

independently by three groups; Robert Brout and François Englert [75], followed independently by Peter Higgs [76], and later by Gerald Guralnik, Carl Hagen, and Tom Kibble [77][9]. They were all able to demonstrate that by combining a gauge theory with an additional scalar field that spontaneously breaks the symmetry, the gauge bosons may consistently acquire a finite mass. This is accomplished through the introduction of an isospin doublet of complex scalar fields ϕ with hypercharge 1 and weak isospin 1/2:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{2.16}$$

The Lagrangian corresponding to this Higgs field ϕ is

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger D^\mu \phi - V(\phi), \tag{2.17}$$

where the first term is our kinetic term and $V(\phi)$ the potential. Our potential takes the form:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2. \tag{2.18}$$

The parameter λ is assumed to be positive. If $\mu^2 > 0$, we obtain a single ground state at 0, causing the exact symmetry of the Lagrangian to be maintained. For $\mu^2 < 0$, the vacuum state is no longer at zero, but could be any one of an infinite number of minima around zero as seen in Figure 2.2. Specifying one of these minima as the ground state spontaneously breaks the $SU(2)_L \otimes U(1)_Y$ symmetry to $U(1)_{\text{QED}}$ such

[9]The history of the Higgs mechanism is more complicated. Reference [78] and references therein give a more nuanced answer to who invented/discovered what regarding this topic.

that the electromagnetic subgroup is still a true symmetry of the vacuum, implying electric charge is still conserved. Because the electric charge is a conserved quantity, only the neutral scalar field (lower part of doublet in Equation 2.16) can acquire a vacuum expectation value, or vev, which is the scale of the weak interaction.

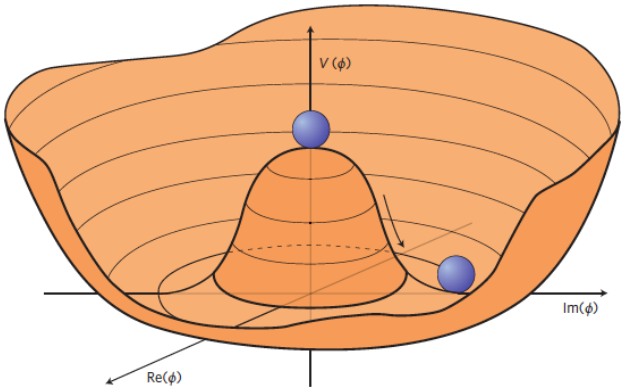


Figure 2.2: Higgs potential in the case $\mu^2 < 0$ with the minimum as $|\phi|^2 = -\mu^2/(2\lambda)$. Choosing any of the minima spontaneously breaks the rotational $U(1)$ symmetry [78].

Solving for the vev, or field minimum of the potential, we find:

$$\langle \phi \rangle = \sqrt{\frac{-\mu^2}{\lambda}} = \frac{v}{\sqrt{2}}, \tag{2.19}$$

where $v \approx 246\,\text{GeV}$ [79]. Without a loss of generality, we are free to rotate ϕ so that $\phi_1 = \phi_2 = \phi_4 = 0$, leaving $\phi_3 = v$ where v is the vacuum expectation value. We can see this by expanding the field about the minimum:

$$\phi(x) = \frac{e^{i\frac{\tau_a \xi^a(x)}{2v}}}{\sqrt{2}} \begin{pmatrix} 0 \\ \\ v + h(x) \end{pmatrix}, \tag{2.20}$$

where $h(x)$ corresponds to radial perturbations around the minima v. These radial perturbations represent a *physical* excitation in the Higgs field, which we observe as a physical Higgs boson H. The real Nambu-Goldstone fields $\xi^a(x)$ [80–82] are

determined by the choice of gauge. According to Goldstone's theorem [81], we should find a massless degree of freedom for every broken generator of the group. These massless modes, or Nambu-Goldstone bosons, are not physical particles, and the local $SU(2)_L$ invariance of the Lagrangian allows us to rotate away the dependence on $\xi(x)$ by proper choice of gauge. We transform $\phi \to e^{i\frac{\tau^a}{2}\beta^a(x)}\phi$, choosing $\beta(x) = -\xi(x)/v$:

$$\begin{aligned}\phi(x) &\to e^{-i\frac{\tau_a\xi^a(x)}{2v}}\frac{e^{i\frac{\tau_b\xi^b(x)}{2v}}}{\sqrt{2}}\begin{pmatrix}0\\ \\ v+h(x)\end{pmatrix}\\ &= \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ \\ v+h(x)\end{pmatrix}.\end{aligned} \tag{2.21}$$

This choice of gauge is called the Unitary gauge, where the ξ degrees of freedom vanish from our scalar field, but reappear as the longitudinal components of the $W^\pm$ and Z bosons. When they acquire mass, they have effectively been "eaten" by the gauge fields, leaving us with only one scalar field, the Higgs boson with mass $m_H^2 = -2\mu^2 = 2\lambda v^2$, with the experimentally measured value of $m_H = 125.10 \pm 0.14\,\text{GeV}$ [79].

2.4.2 Vector Boson Masses

Now we turn to how we can give mass to the vector bosons. Let's consider how these mass terms would manifest by combining the four gauge fields (W_μ^a and B_μ)

into a single term V_μ:

$$V_\mu = \begin{pmatrix} W_\mu^1 \\ W_\mu^2 \\ W_\mu^3 \\ B_\mu \end{pmatrix}. \tag{2.22}$$

We could recognize a set of quadratic mass terms as arising from the Lagrangian involving ϕ, which is bilinear in the fields V_μ^a, taking on the compact form

$$|D_\mu \phi_0|^2 \to \frac{1}{2} M_{ab}^2 V_\mu^a V^{\mu,b}, \tag{2.23}$$

where $\phi_0 = \frac{v}{\sqrt{2}}(0 \;\; 1)^T$ such that

$$|D_\mu \phi_0|^2 \to \frac{v^2}{8}\left[g^2 (W_\mu^1)^2 + g^2 (W_\mu^2)^2 + (g W_\mu^3 - g' B_\mu)^2\right]. \tag{2.24}$$

From this, we can quickly read off the values for the squared-mass matrix M^2,

$$M^2 = \frac{v^2}{4}\begin{pmatrix} g^2 & 0 & 0 & 0 \\ 0 & g^2 & 0 & 0 \\ 0 & 0 & g^2 & -gg' \\ 0 & 0 & -gg' & g'^2 \end{pmatrix}. \tag{2.25}$$

By diagonalizing this matrix we obtain the mass eigenvalues and mass eigenstates in terms of the originally proposed fields and couplings, giving us three massive and one massless (*i.e.* the photon) gauge bosons:

$$
\begin{aligned}
W_\mu^\pm &= \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \\
Z_\mu &= \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2+g'^2}} \\
A_\mu &= \frac{gW_\mu^3 + g'B_\mu}{\sqrt{g^2+g'^2}}
\end{aligned}
\tag{2.26}
$$

$$
\begin{aligned}
m_W &= \frac{gv}{2} \\
m_Z &= \frac{v}{2}\sqrt{g^2+g'^2} \\
m_\gamma &= 0.
\end{aligned}
\tag{2.27}
$$

The $W^\pm$ and Z bosons now have mass, which have been measured in experiments with $m_W \approx 80.4\,\text{GeV}$ [83] and $m_Z \approx 91.2\,\text{GeV}$ [84]. We can express the photon and Z as a rotation of the W^3 and B fields by an angle, known as the Weinberg, or weak mixing angle θ_W:

$$
\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}, \tag{2.28}
$$

so that

$$
\begin{aligned}
Z_\mu &= \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu, \\
A_\mu &= \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu.
\end{aligned}
\tag{2.29}
$$

By applying the covariant derivative on the right-handed electron field (e_R), expressed in terms of A_μ and Z_μ

$$D_\mu e_R = \partial_\mu e_R - ig' \cos\theta_W A_\mu e_R + ig' \sin\theta_W Z_\mu e_R, \tag{2.30}$$

and directly comparing with the results obtained from QED:

$$D_\mu e_R = \partial_\mu e_R - ieA_\mu e_R, \tag{2.31}$$

where e is the electromagnetic coupling from QED. We find that

$$e = g' \cos\theta_W. \tag{2.32}$$

Doing the same for the left-handed electron field (e_L):

$$D_\mu e_L = \partial_\mu e_L + i(g' Y_{e_L} \cos\theta_W - \frac{g}{2} \sin\theta_W) A_\mu e_L + \cdots, \tag{2.33}$$

with $Y_{e_L} = -1/2$, compared with QED results which are equivalent to those found in Equation 2.31 yields,

$$e = \frac{g}{2} \sin\theta_W + \frac{g'}{2} \cos\theta_W. \tag{2.34}$$

Combining the results from Equations 2.32 and 2.34 we can express the Weinberg angle θ_W and e in terms of the coupling constants g and g'. This can be seen geometrically in Figure 2.3.

$$\tan\theta_W = \frac{g'}{g} \tag{2.35}$$

$$e = \frac{gg'}{\sqrt{g^2 + g'^2}} \tag{2.36}$$

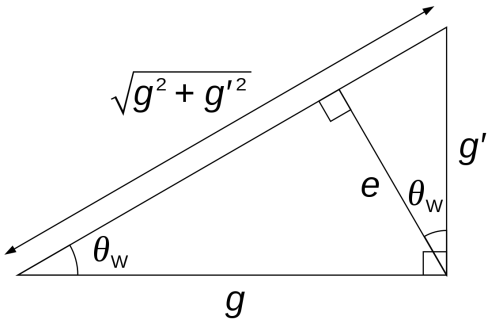


Figure 2.3: Geometrical representation showing the relationships between the weak mixing angle θ_W, the coupling constants g and g', and the electromagnetic coupling e.

2.4.3 Fermion Masses

Without the Higgs Mechanism, fermion masses would break gauge symmetry; left-handed fermions are grouped into doublets and right-handed into singlets, making them incompatible. To mix these left and right chiral states, we again turn to this mechanism which generates mass terms via the Yukawa coupling of the Higgs field to fermions. In the Lagrangian, a Yukawa interaction takes the form:

$$\mathcal{L}_{\text{Yukawa}} = -y_{\text{Yukawa}}\overline{\psi}_{\text{L}}\phi\psi_{\text{R}}, \tag{2.37}$$

where y_{Yukawa}, the coupling constants defining the strength of the interaction between the fermions and the simple scalar field ϕ, give us our fermion mass terms. The Lagrangian is now gauge invariant under $SU(2)_L$ transformations. We rely on two

representations of the Higgs field, with $Y = \pm 1/2$ to separately give mass to down-type quarks and charged leptons, as well as up-type quarks and neutrinos. However, since the SM does not (minimally) contain right-handed neutrinos, no such Yukawa interaction can be built for the neutrinos, which are consequently predicted to be massless at the Lagrangian level [85]. This can be remedied through extensions of the SM such as the see-saw mechanism [86–91], extra dimensions [92], R-parity violating supersymmetry [93], etc. Expressed in terms of left- and right-handed quark and lepton fields, our Yukawa Lagrangian takes the form:

$$\mathcal{L}_{\text{Yukawa}} = Y^u_{ij}\overline{Q}^i_L\tilde{\phi}u^j_R + Y^d_{ij}\overline{Q}^i_L\phi d^j_R + Y^\ell_{ij}\overline{L}^i_L\phi e^j_R + h.c., \tag{2.38}$$

where $\tilde{\phi}$ is the charge conjugate of the Higgs doublet, defined as:

$$\tilde{\phi} = -i\tau_2\phi^* = -\frac{1}{\sqrt{2}}\begin{pmatrix} v + h(x) \\ 0 \end{pmatrix}. \tag{2.39}$$

Defining the fields more explicitly, we have the left-handed doublets for quarks (Q_L) and leptons (L_L) and right-handed singlets for up-type (u_R) and down-type (d_R) quarks and charged leptons (e_R):

$$\begin{aligned} Q^i_L &= \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix} \\ u^i_R &= (u_R), (c_R), (t_R) \\ d^i_R &= (d_R), (s_R), (b_R) \end{aligned} \tag{2.40}$$

$$L_L^i = \begin{pmatrix} \nu_{e_L} \\ \ell_L \end{pmatrix}, \begin{pmatrix} \nu_{\mu_L} \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_{\tau_L} \\ \tau_L \end{pmatrix}$$

$$e_R^i = (e_R), (\mu_R), (\tau_R). \tag{2.41}$$

The matrices of the form Y_{ij} incorporate the Yukawa couplings of the Higgs boson to the up- and down-type quarks and charged leptons. The mass terms are generated from this Higgs-fermion coupling with the form:

$$m_f = \frac{y_f v}{\sqrt{2}}, \tag{2.42}$$

i.e. the diagonalized Yukawa couplings are proportional to the masses of the fermions. Just as with the masses, the couplings are free parameters of the theory and must be determined by experiment. The Yukawa coupling matrices in flavor and mass space cannot be simultaneously diagonalized, so the mass eigenstates are not flavor eigenstates, which introduces mixing between the different flavors. Representing the up-/down-type quarks as a triplet, we can observe that the weak eigenstates q' are a linear combination of the mass eigenstates q:

$$\begin{pmatrix} u' \\ c' \\ t' \end{pmatrix} = U_{L,R} \begin{pmatrix} u \\ c \\ t \end{pmatrix}_{L,R}, \quad \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = U_{L,R} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{L,R}, \tag{2.43}$$

Where $U_{L,R}/D_{L,R}$ are the up/down (left- and right-handed) unitary matrices which diagonalize the quark mass matrices $M_{u,d} = \frac{v}{\sqrt{2}} Y_{u,d}$:

$$U_R^{-1} M_u U_L = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_t \end{pmatrix}, \quad D_R^{-1} M_d D_L = \begin{pmatrix} m_d & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_b \end{pmatrix}, \tag{2.44}$$

The probability of a quark transitioning to another is described in the CKM (Cabibbo-Kobayashi-Maskawa) matrix, which is formed by the unitary transformation that diagonalizes the up/down Yukawa matrices:

$$V_{\text{CKM}} = U_{L,u} U_{L,d}^{\dagger} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{2.45}$$

We will revisit the CKM matrix in Section 4.1, when discussing the decay of bottom quarks.

Next we will explore the phenomenology of the Higgs boson at the Large Hadron Collider (LHC). The LHC is the first particle accelerator that gives us the opportunity to produce and study the Higgs boson at high energies in a laboratory. In the following chapter, I will discuss how the Higgs boson is produced, how it decays, and how it was discovered.

Chapter 3

The Higgs Boson at the Large Hadron Collider

The Large Hadron Collider (LHC) [18, 19], discussed in Chapter 5, is the largest machine ever constructed – designed with the aims to discover the Higgs particle, arguably one of the most important particles in the Standard Model yet discovered, and to study rare events at center-of-mass energies up to 14 TeV [18, 94]. At a particle collider such as the LHC, two counter-rotating beams of protons are accelerated to nearly the speed of light and directed into head-on collisions, producing a host of different particles such as the Higgs boson. The mass of the Higgs boson is a free parameter in the Standard Model, and although lower limits on its mass had been determined by previous experiments at LEP [95, 96] and Tevatron [97, 98], we had no a priori knowledge for it. If the Higgs boson mass did indeed fall within the mass ranges reachable by the LHC, it was expected to be produced about once out of every billion proton-proton collisions, making it a rare process indeed. It would take several years of operation and painstaking analysis techniques to be able to see such a rare process occurring in the midst of such a considerable background. The search

for the Higgs boson, like many other searches, has been compared to looking for a needle in a haystack. I would argue that it is more akin to searching for a particular needle in a mountain of similar needles. To find such an elusive needle we must know how it is produced, as well as how it decays.

3.1 Higgs Production Modes

There are several Higgs production processes accessible at the LHC. The cross sections for these production modes, as a function of the proton-proton center-of-mass energy, can be seen in Figure 3.5. The cross section is a measure of the probability that two particles will collide and react in a certain way, and is expressed in terms of an effective area for collision (in barns, $1\,\text{b} = 10^{-24}\,\text{cm}^2$) [99]. The cross sections at the LHC's center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ are in Table 3.1. Here, we will discuss the largest inclusive contributions to the overall Higgs cross section.

Table 3.1: The SM Higgs boson production cross sections for $m_H = 125\,\text{GeV}$ for $\sqrt{s} = 13\,\text{TeV}$. The predictions for the LHC energies are taken from References [100–103]. The predictions for the ggF channel at the LHC include the latest N3LO results leading to reduced theoretical uncertainties by a factor of 2 compared to the NNLO+NLL results. The total uncertainties are estimated assuming no correlations between α_s and PDF uncertainties [54].

Production cross section (in pb) for $m_H = 125\,\text{GeV}$ at $\sqrt{s} = 13\,\text{TeV}$					
ggF	VBF	WH	ZH	ttH	Total
$48.6^{+4.6\%}_{-6.7\%}$	$3.78^{+2.2\%}_{-2.2\%}$	$1.37^{+2.6\%}_{-2.6\%}$	$0.88^{+4.1\%}_{-3.5\%}$	$0.50^{+6.8\%}_{-9.9\%}$	55.1

Gluon-Gluon Fusion (ggF)

The gluon-gluon fusion (ggF) process [103, 104], $gg \to H$, is the dominant Higgs production mode at the LHC. The Higgs boson is not charged under $SU(3)_c$, but

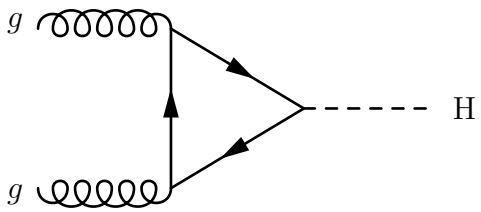


Figure 3.1: Gluon-gluon fusion (ggF)

has a significant coupling (near unity) to top quarks, so the process is mediated indirectly by a virtual fermion loop as seen in Figure 3.1 at leading order (LO). Contributions from lighter quarks within the loop are suppressed proportional to m_q^2. With the top quark nearly 35 times heavier than the next heaviest quark, the bottom quark, the contribution to the overall ggF cross section is top dominated. This allows us to indirectly probe the Higgs coupling to the top quark.

Vector Boson Fusion (VBF)

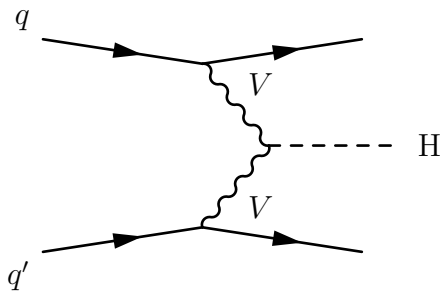


Figure 3.2: Vector Boson Fusion (VBF)

Vector boson fusion (VBF) [103], $qq \to qqH$ as shown at leading order (LO) in Figure 3.2, is the subdominant production mode by which two (anti-)quarks scatter through the exchange of virtual vector bosons V (W/Z) that combine to produce a Higgs boson along with two jets in the forward regions of the detector [105]. The

VBF channel, through its topology, gives us a clean signature that allows us to probe the strength of the non-linear interactions between the Higgs and vector bosons at high energy. It also helps establish the nature of the Higgs; whether it's a composite or elementary particle [54, 106].

Higgs-strahlung (VH)

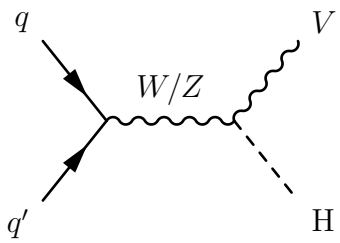
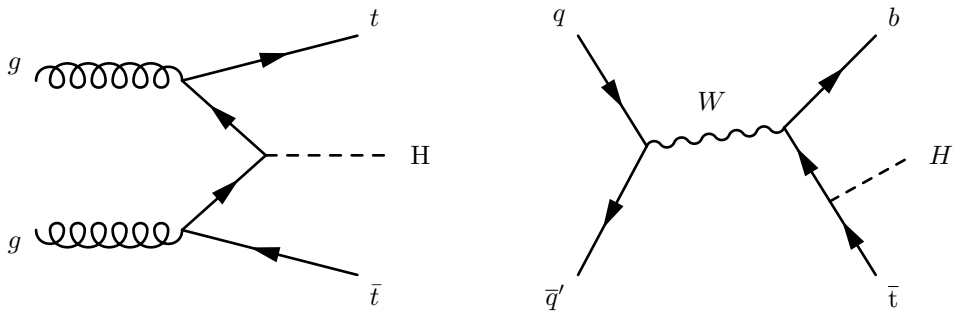


Figure 3.3: Higgstrahlung (VH)

The third leading Higgs production mechanism is in Figure 3.3; Higgs-strahlung (VH) [103], $pp \to VH$, involves a Higgs boson produced in association with a vector boson. While also allowing us to probe the strength of the vector boson coupling to the Higgs boson, this channel benefits from the leptonic decay of the vector bosons, which can lead to an unambiguous event signature. While outside of the scope of this dissertation, this production mode is particularly worthy of study in a lepton collider due to an exceptionally clean signature [107].

Associated Top-Quark Production (ttH and tH)

The last production mode we will discuss is the Higgs produced in association with one or two top quarks (tH and ttH) [103]. While it has a still yet lower cross section compared to VH, it benefits from the unique decay topology of the top quark, which can aid event selection. The tH and ttH production modes provide us



(a) ttH (b) tH

with a way to measure the Higgs boson coupling to top quarks directly. In the case of tH, we even have access to the sign of the top Yukawa coupling [108–117], due to an almost totally destructive interference between two large contributions; one in which the Higgs couples to a space-like W and another where it couples to the top.

3.2 Higgs Decay Modes

The Higgs can decay via a variety of modes. In order to interpret our experimental results, it is key that we be able to compute all of the relevant decay widths (and their uncertainties) of the Higgs boson. As mentioned in Section 2.4.3, the Higgs coupling strengths to fermions are proportional to the masses:

$$g_{Hf\bar{f}} = \frac{m_f}{v}. \tag{3.1}$$

The coupling to vector bosons is not proportional to the masses as with fermions, but to the square of the masses.

$$g_{HVV} = \frac{2m_V^2}{v}. \tag{3.2}$$

The masses are not predicted by the theory, but through independent mass measurements we can determine the expected couplings and compare with data. The branching ratio[1] (BR) is the fraction of parent particles which decay via a specific decay mode to the total number of possible decays as expressed in Equation 3.3.

$$BR(H \to XX) \equiv \frac{\Gamma(H \to XX)}{\sum \Gamma(H \to X_i X_i)}. \tag{3.3}$$

The decay branching fraction for the Higgs boson is proportional to the square of the coupling and a phase-space term. The mass of the Higgs boson also determines the relative fractions for each decay. If the mass of the Higgs boson were larger, for example $m_H \gtrsim 200$ GeV, it would decay almost exclusively to pairs of vector bosons. For fermions, We expect the largest BRs to follow in descending order of the masses[2]. The coupling of the Higgs boson to the top quark is near unity, but with $m_t > m_H$ we do not observe a Higgs boson decaying to a $t\bar{t}$ pair. The largest branching fraction is to bottom quarks, which makes up nearly 58% of the Higgs total decay width. We then expect it to couple more strongly to tau leptons, then charm quarks, etc. Having an accurate measurement for the Higgs boson's decay to b-quarks is a necessity as it not only constrains the coupling of the Higgs boson to bottom quarks, it also helps improve the precision for other decay channels. The reason for this is that all of the observed rates depend upon the product of two partial decay widths (from Higgs production and decay) divided by the overall Higgs boson total width. Given the large branching fraction of $H \to b\bar{b}$, the total width is dominated by this decay channel. Without better constraints on the coupling to b-quarks then, it is possible

[1] Also known as a branching *fraction*

[2] The relationship between the mass and BR for the vector bosons is more complicated because at the Higgs boson mass of $m_H \approx 125\,\text{GeV}$, it cannot decay to two on-shell W or Z bosons. See Section 2.2 of Reference [118] and references therein.

to have an arbitrary shift of all of the observed rates.

Several Higgs decay branching ratios can be found in Table 3.2. All of these decays, provided we have the statistics to properly measure them, give us direct access to the fermion Yukawa couplings. The two seemingly odd decay modes in Figure 3.5 are gg and $\gamma\gamma$. Since the Higgs boson has neither color nor electrical charge, and neither the gluon nor the photon have mass, why are these decay modes present? A loop-induced decay allows the Higgs boson to couple indirectly to gluons (Figure 3.6b) and photons (Figure 3.6a), providing us with indirect information on the Higgs boson coupling to WW, ZZ, and $t\bar{t}$.

Table 3.2: The branching ratios and relative uncertainties [102, 103] for a SM Higgs boson with $m_H = 125$ GeV [54, 79].

Decay Channel	Branching Ratio	Rel. Uncertainty
$H \to b\bar{b}$	5.82×10^{-1}	$^{+1.2\%}_{-1.3\%}$
$H \to W^+W^-$	2.14×10^{-1}	$\pm 1.5\%$
$H \to \tau^+\tau^-$	6.27×10^{-2}	$\pm 1.6\%$
$H \to c\bar{c}$	2.89×10^{-2}	$^{+5.5\%}_{-2.0\%}$
$H \to ZZ$	2.62×10^{-2}	$\pm 1.5\%$
$H \to \gamma\gamma$	2.27×10^{-3}	2.1%
$H \to Z\gamma$	1.53×10^{-3}	$\pm 5.8\%$
$H \to \mu^+\mu^-$	2.18×10^{-4}	$\pm 1.7\%$

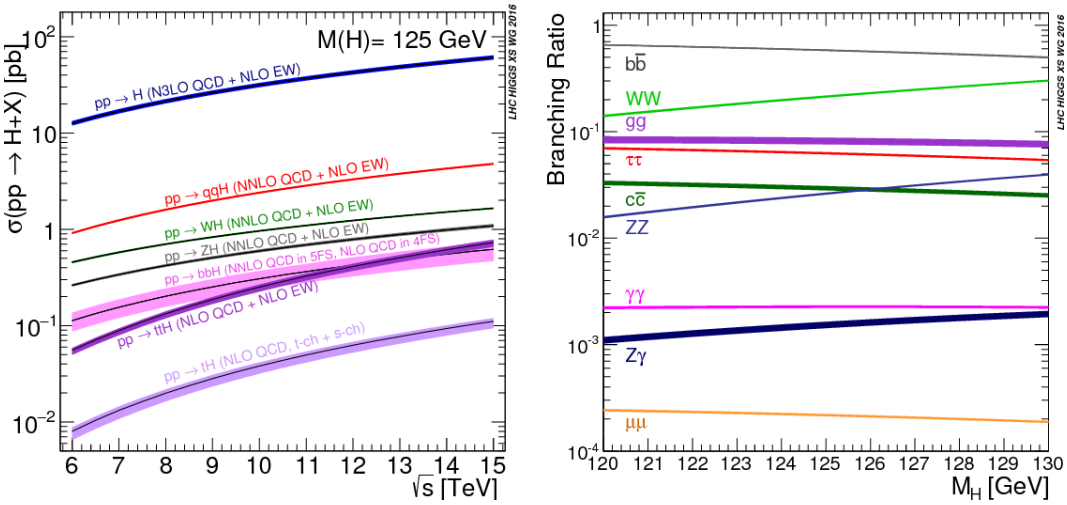
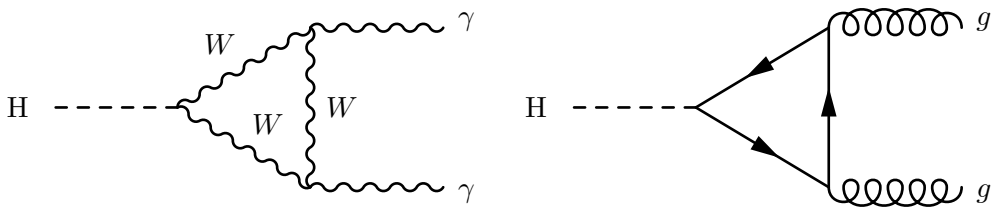


Figure 3.5: Left: Cross sections for various Higgs production modes as a function of the center-of-mass energies for proton-proton collisions. Right: Relative branching fractions of Higgs decays as a function of the Higgs mass [119].



(a) $H \to \gamma\gamma$ (b) $H \to gg$

Figure 3.6: Leading order feynman diagrams for the dominant Higgs boson decay modes to (a) photons via a W boson loop, and to (b) gluons via a top loop.

3.3 Evidence for the Higgs boson

The Higgs boson discovery was announced on July 4th 2012 jointly by the ATLAS [21] and CMS [22] Collaborations. Over 40 years after it was proposed as the linchpin of the Standard Model, the Higgs boson was finally discovered by two of the largest scientific collaborations in history. François Englert[3] and Peter Higgs shared the 2013 Nobel Prize [120] for their role in developing the foundational theory. It should be noted, that despite the fact that Englert and Brout [75] were the first to

[3]Sadly, Robert Brout passed away in May of 2011.

describe the new scalar field and the mechanism of electroweak symmetry breaking, it was Higgs [76] who first made mention of a new massive scalar boson resulting from the broken symmetry. The Higgs boson was observed decaying to a pair of photons (Figure 3.7), and shortly after, a pair of Z bosons decaying to four leptons [21, 22].

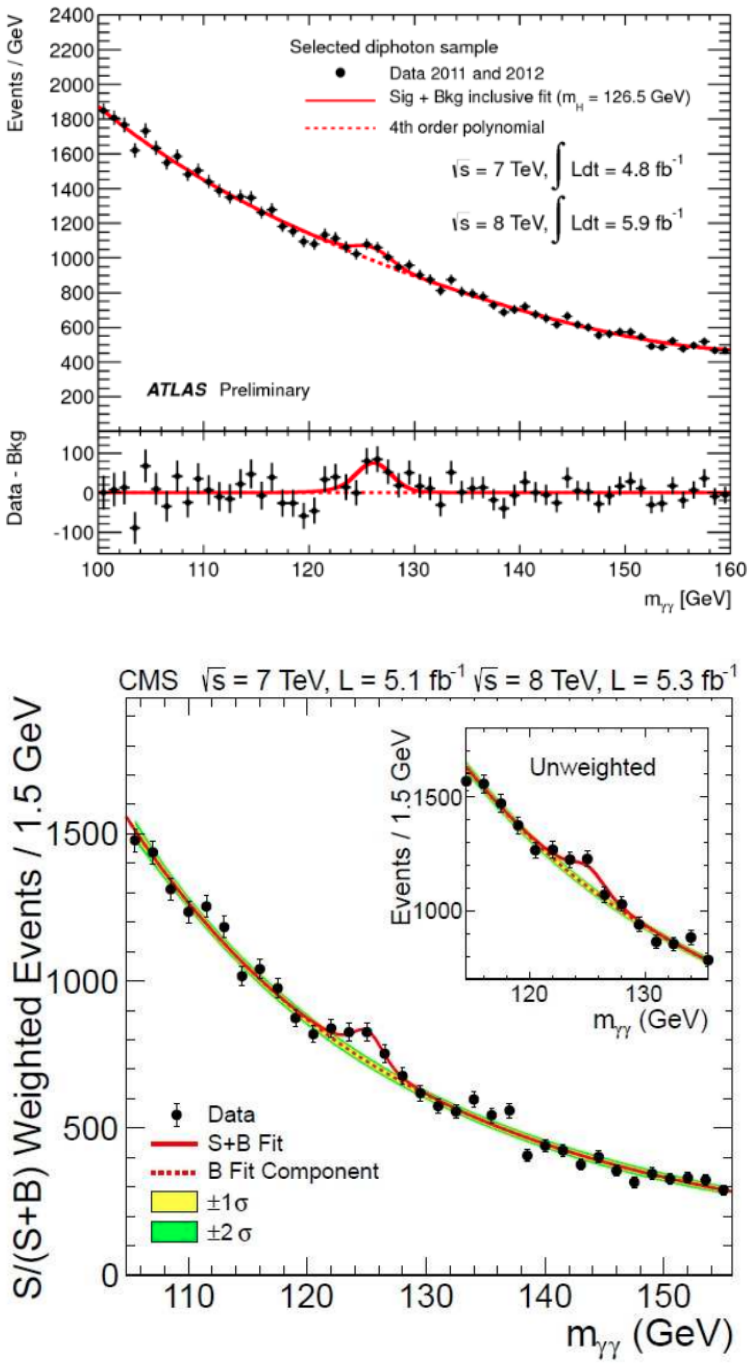


Figure 3.7: ATLAS [21] (above) and CMS [22] (below) $H \to \gamma\gamma$ discovery plots.

Since then, as more data have been collected and more advanced searches performed, the Higgs boson has been observed in multiple production and decay channels; $H \to ZZ$ [121, 122], $H \to WW$ [123, 124], $H \to \tau\overline{\tau}$ [125, 126], $H \to \mu\mu$ [127, 128] (measurements only made recently), $H \to b\bar{b}$ [24, 129], and in association with $t\bar{t}$ pairs [130, 131], all in good agreement with the Standard Model of Particle Physics as seen in Figures 3.8 and 3.9, which show the combined measurements for the main production and decay modes from ATLAS and CMS, respectively. The results reported in the figures, and indeed many analyses in general, are the observed signal strength μ, which is the observed cross section times branching ratio divided by the SM prediction:

$$\mu = \frac{\sigma_{\text{obs}} BR_{\text{obs}}}{\sigma_{\text{SM}} BR_{\text{SM}}}. \tag{3.4}$$

Additionally, combined measurements in the $H \to ZZ^* \to 4$ leptons and $H \to WW^* \to \ell\nu\ell\nu$ channels have shown that the Higgs boson has spin and parity properties as predicted [132–137] by the SM. Conspicuously missing from the ATLAS plot is the $H \to b\bar{b}$ decay mode for the ggF production. This is a particularly challenging channel, and as it is the focus of this dissertation, we will discuss it in general detail in Section 4, and in more detail starting in Chapter 7.

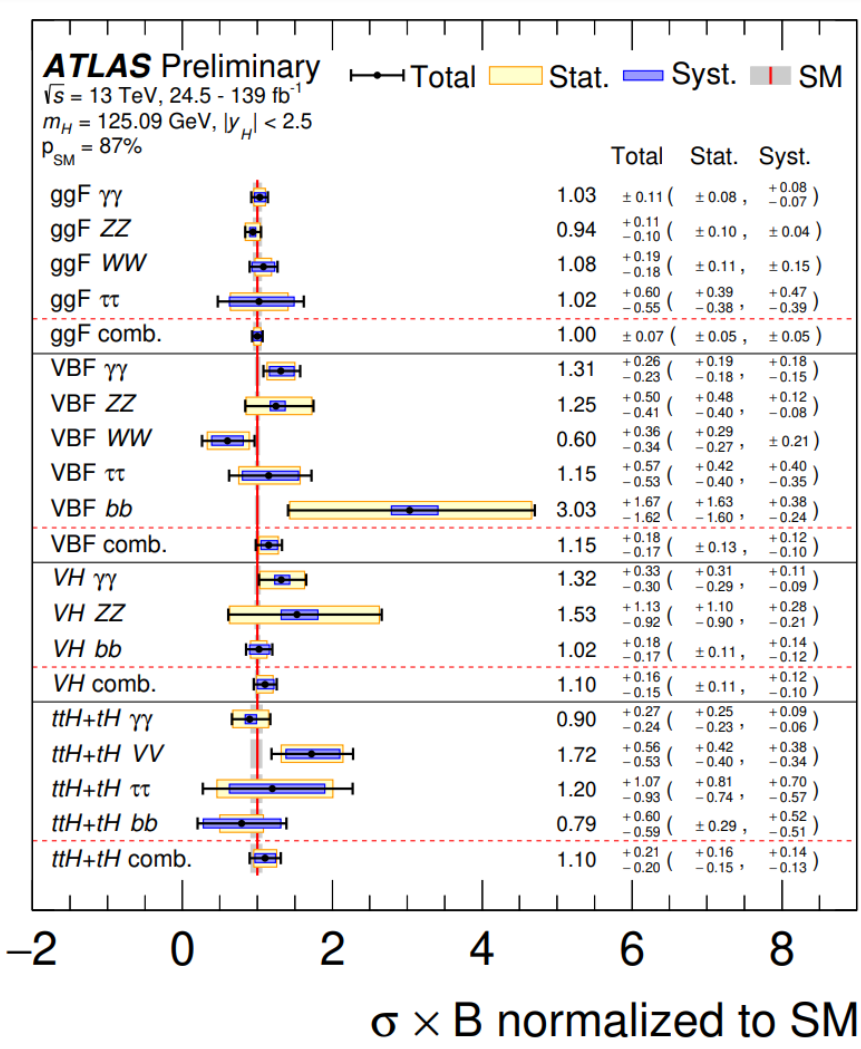


Figure 3.8: Combined ATLAS Collaboration [138] measurements of $\sigma \times \mathrm{BR}$, normalized to their SM predictions for ggF, VBF, VH, and $ttH + tH$ production in each relevant decay mode. The values are obtained from a simultaneous fit to all channels. The cross sections of the ggF, $H \to b\bar{b}$, VH, $H \to WW^*$ and VH, $H \to \tau\tau$ processes are fixed to their SM predictions. Combined results for each production mode are also shown, assuming SM values for the branching fractions into each decay mode. The black error bars, blue boxes and yellow boxes show the total, systematic, and statistical uncertainties in the measurements, respectively. The gray bands show the theory uncertainties in the predictions. The level of compatibility between the measurement and the SM prediction corresponds to a p-value of $p_{\mathrm{SM}} = 87\%$, computed using the procedure outlined in Reference [138].

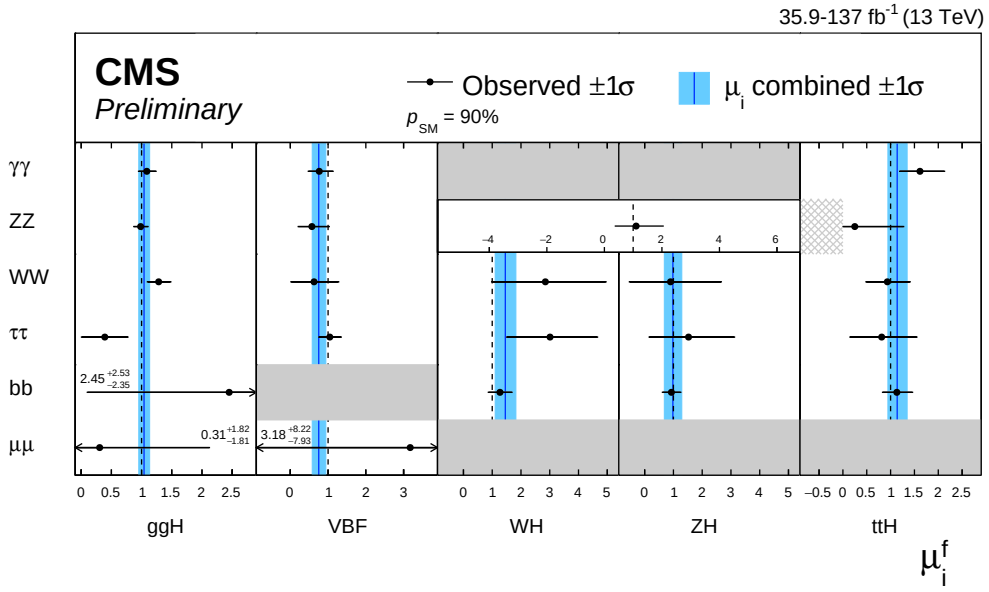


Figure 3.9: Combined CMS Collaboration [139] measurements of $\sigma \times \mathrm{BR}$, normalized to their SM predictions, μ_i^f. The black points and horizontal error bars show the best-fit values and 1σ confidence intervals, respectively. The arrows indicate cases where the confidence intervals exceed the scale of the horizontal axis. The gray filled boxes indicate signal strength modifiers which are not included in the model, while the gray hatched box indicates the region for which the sum of signal and background becomes negative in the fit for μ_{ttH}^{ZZ}. In the $H \to ZZ$ decay mode, a common modifier is fit to the WH and ZH production modes. The measured value and 1σ confidence interval for each production cross-section modifier, μ_i, from the combination across decay channels, is indicated by the blue vertical line and the blue bands, respectively. The indicated p-value is given for the production times decay mode signal strength modifiers. The assumptions used in this fit are described in Reference [139].

3.4 The Higgs Boson: A Tool for Discovery

Despite this tremendous success at the LHC, now is no time to rest on our laurels. We have good evidence that the particle discovered in 2012 is indeed the Higgs boson of the Standard Model, but more data and precise measurements are needed to be absolutely sure. Even with the incredible predictive power of the SM, there remain many fundamental physics questions unanswered. For example, as mentioned, we have no convincing QFT formulation, nor evidence for, a particle nature to the gravitational interaction [140]. The SM does not account for the majority of the mass/energy (dark matter [141] and dark energy [142]) in the universe, nor does it account for the matter/antimatter asymmetry [143].

The physics program of the 20th Century saw a wonderful interplay between experimentalists and theorists; either theorists would predict the existence of new particles which were shortly discovered, or experimentalists would find evidence of new particles which were soon explained within the evolving theoretical framework. After the SM formulation of the 1960s and 1970s, we had a path that was more or less clear; find the remaining quarks, heavy vector gauge bosons of the weak interaction, and finally, the Higgs boson. With that accomplished, the way forward today is far from certain. This is in no way an accusation that either theorists nor experimentalists are lackadaisical, but rather that the physics program of the 21st century is far from clear and that much of the low-hanging fruit has already been picked. There are numerous proposals for discovering physics beyond the SM and we would be remiss in overlooking a promising tool; the Higgs boson itself.

There are many interesting areas where the Higgs can tell us something new.

For example there is the question of the stability of the Higgs potential[4] [144] or the Hierarchy problem[5] [145] to name only a few. While an entire thesis could be devoted to interesting searches for the Higgs boson, some looking to nail down the free parameters of the SM, others looking for the fingerprints of new physics. I will instead focus on the deviations to the Higgs coupling properties, with some emphasis on loop-induced decays. Particular attention has been paid to the couplings of the Higgs boson to various other particles. This is no accident. Many of the searches for SM deviations involve modifications to the coupling properties of the Higgs boson to particles.

To make this concrete, let's consider a popular extension to the Standard Model called the Minimal Supersymmetric Standard Model (MSSM). The MSSM associates a Supersymmetry (SUSY) partner to each gauge boson and chiral fermion of the SM and provides a realistic model of physics at the weak scale [79]. The phenomenology of this minimal extension to the SM is one of the most studied SUSY extensions. Our discussion on this topic will be rather brief, but detailed reviews of the properties and phenomenology of the Higgs sector of the MSSM can be found in References [147–149] and the references therein. For our discussion, we will focus on the subset of the MSSM phenomenology involving the Higgs sector. The MSSM contains the particle content of the Type-II two-Higgs-doublet model (2HDM) extension to the SM and corresponding SUSY partners [149, 150].

Two Higgs doublets, ϕ_1 and ϕ_2, generate masses for the down-type quarks/charged leptons (ϕ_1) and up-type quarks (ϕ_2), respectively [151]. After electroweak symmetry breaking (EWSB) $SU(2)_L \times U(1)_Y \rightarrow U(1)_{\text{em}}$, much like in

[4] I cover this in Appendix A.8 in brief detail. Check out that section and the references therein.
[5] Also covered in Appendix A.9 [145, 146].

Section 2.4, we set

$$\langle \phi_1^0 \rangle \equiv \frac{v_1}{\sqrt{2}}, \quad \langle \phi_2^0 \rangle \equiv \frac{v_2}{\sqrt{2}}. \tag{3.5}$$

The vevs, v_1 and v_2, are connected to the SM vev v, the known mass of the Z boson, and the electroweak couplings:

$$v_1^2 + v_2^2 = v^2 = 4\frac{m_Z^2}{g^2 + g'^2} \approx (246\,\text{GeV})^2. \tag{3.6}$$

The ratio of the vevs is written as:

$$\tan\beta = \frac{v_2}{v_1}. \tag{3.7}$$

The symmetry breaking is similar in both the MSSM and SM, so we expect a set of Nambu-Goldstone bosons [80–82]. After the appropriate gauge fixing, the two complex doublet fields will leave behind five physical Higgs particles: a charged Higgs pair $H^\pm$, one CP-odd neutral scalar A, and two CP-even neutral states h and H, where h is the lightest Higgs boson. Note, up to this point, H has represented the SM Higgs boson. For the remainder of this discussion, h and H are of the MSSM, and h is often compared to the SM Higgs boson. The phenomenology of the Higgs sector depends on the couplings of the Higgs bosons to gauge bosons and fermions. At tree level, the couplings of the two neutral Higgs bosons, h and H, to W and Z bosons are expressed in terms of the angles β and α, the latter diagonalizes the CP-even Higgs bosons' squared-mass matrix:

$$g_{hVV} = g_V m_V \sin(\beta - \alpha), \quad g_{HVV} = g_V m_V \cos(\beta - \alpha), \tag{3.8}$$

where $g_V \equiv 2m_V/v$, for $V = W/Z$ ($g_V m_V$ is the SM hVV coupling). In the limit $\cos(\beta - \alpha) \to 0$, the lightest CP-even Higgs boson h behaves as the SM Higgs boson [152, 153].

At tree level, the neutral component of the first Higgs doublet ϕ_1 couples exclusively to down-type fermion pairs while the neutral component of the second doublet ϕ_2 couples only to up-type fermions [151], defining the type-II 2HDM Higgs-fermion coupling properties [79]. In the MSSM, both neutral Higgs components acquire a vacuum expectation value to provide mass to the fermions. The relationship between the Yukawa couplings h_f (much like the SM Yukawa couplings y_f) and fermion masses m_f is:

$$h_{q_d,\ell} = \frac{\sqrt{2} m_{q_d,\ell}}{v \cos(\beta)}, \quad h_{q_u} = \frac{\sqrt{2} m_{q_u}}{v \sin(\beta)}, \tag{3.9}$$

where q_d are the down-type quarks, ℓ are charged leptons, and q_u are up-type quarks. At large $\tan\beta$, the non-standard Higgs bosons have significantly enhanced couplings to down-type fermions.

If decays to SUSY particles are kinematically accessible, the SM-like branching fractions of h can be modified. For example, the h boson could decay into a pair of the lightest SUSY particles (the lightest neutralinos χ_1^0), which could become the dominant mode, but would be invisible if R-parity is conserved [154]. If these light superpartners exist and can couple to gluons and/or photons, the h-loop-induced couplings to gg and $\gamma\gamma$ could deviate significantly from the SM predictions [155]. Loop-induced production is of considerable interest at hadron colliders such as the LHC, as the dominant production mechanism for the neutral Higgs boson for moderate values of $\tan\beta$ is via ggF [79]. In the following sections, we will discuss some

of the methods used to encapsulate the effects from BSM physics into modifications of the SM couplings to the Higgs boson. These methods do not specify the new physics model, but have a wide range of applicability in constraining a variety of SM extensions.

3.4.1 Kappa Framework

The kappa framework [102, 156] is a convenient first parameterization for BSM physics in single Higgs boson production. Under this framework, we characterize the coupling properties in terms of a series of Higgs coupling strength modifier parameters κ_i, which are defined as ratios of the couplings of the Higgs boson to particles i to their corresponding SM value [157]. In the SM, under the zero-width approximation [158, 159], we can decompose the cross section as:

$$(\sigma \cdot \mathrm{BR})_{if} = \frac{\sigma_i \Gamma_f}{\Gamma_H}, \tag{3.10}$$

where σ_i is the production cross section of the initial state i. The partial width for the final state f is Γ_f and Γ_H is the total width of the Higgs boson. In the kappa framework, each of the expected SM values in Equation 3.10 is multiplied by the square of the modifier κ_i such that, at leading order, we obtain:

$$(\sigma \cdot \mathrm{BR})_{if} = \frac{\kappa_i^2 \sigma_i^{\mathrm{SM}} \kappa_f^2 \Gamma_f^{\mathrm{SM}}}{\kappa_H^2 \Gamma_H^{\mathrm{SM}}}, \tag{3.11}$$

where κ_H adjusts the SM Higgs width to take into account the modifiers κ_i of the SM Higgs coupling strengths:

$$\kappa_H^2 = \frac{\sum \kappa_i^2 \Gamma_i^{\mathrm{SM}}}{\Gamma_H^{\mathrm{SM}}}. \tag{3.12}$$

If all $\kappa_i = 1$, we have purely SM coupling properties. The values in Equation 3.11 are more useful when represented as the relative rate compared to the SM expectation μ_i^f. The rate is defined as:

$$\mu_f^i \equiv \mu^i \mu_f = \frac{(\sigma \cdot \mathrm{BR})_{\mathrm{observed}}}{(\sigma \cdot \mathrm{BR})_{\mathrm{SM}}} = \frac{\kappa_i^2 \kappa_f^2}{\kappa_H^2}. \tag{3.13}$$

Since the majority of Higgs bosons are produced via gluon fusion, a loop induced process, let us focus more closely on it. At lower energies, if the mass of the fermion is much larger than the Higgs boson mass, the heavy fermion can be integrated out and the coupling between the gluons and the Higgs boson can be described by an effective vertex [160]. Given that the top quark is by far the dominant contribution, this is referred to as the infinite top-quark mass limit [161, 162]. The top Yukawa coupling in the SM is fully determined by the top-quark mass. However, the limit on the top-Yukawa coupling is $y_t/y_t^{\mathrm{SM}} < 1.7$[6] [165], so it is possible to have additional point-like components resulting from BSM physics hiding within our uncertainties.

To describe this additional effective coupling, we consider the following modification of the top Yukawa part of the SM Lagrangian by an extension of the kappa

[6] From $t\bar{t}t\bar{t}$ final state searches performed at ATLAS [163, 164] and CMS [165–167].

framework [168]:

$$\frac{m_t}{v}\bar{t}tH \rightarrow -\kappa_g \frac{\alpha_s}{12\pi v} G^a_{\mu\nu} G^{a,\mu\nu} H + \kappa_t \frac{m_t}{v}\bar{t}tH. \tag{3.14}$$

The first term on the right describes the point-like contribution involving the Higgs boson and gluons and the second term is the modified top Yukawa coupling. For purely SM-like couplings, $\kappa_g = 0$ and $\kappa_t = 1$. With the effective gluon-Higgs coupling, and assuming the heavy top-quark approximation, the Higgs-production cross section is inversely proportional to the square of the sum of the two couplings [168]:

$$\sigma(gg \rightarrow H) \sim (\kappa_g + \kappa_t)^2. \tag{3.15}$$

Even if we could precisely measure the inclusive (in transverse momentum) cross section, we could not disentangle the separate values for κ_g and κ_t. A solution is to use the Higgs boson transverse momentum (p_{T}) observable to measure the differential cross section [169]. This allows us to test the structure of the loop-induced coupling and the underlying Standard Model assumption that the Higgs-gluon coupling is exclusively due to heavy quark loops [170–174]. If we assume that the scale of new physics Λ, responsible for generating our effective Higgs-gluon coupling κ_g, is significantly larger than the top-quark mass, there is a range of transverse momentum $2m_t \ll p_{\mathrm{T}} \ll \Lambda$ such that the BSM contribution to the Higgs-gluon vertex can still be treated as point-like, whereas the top-quark contribution starts being resolved [168]. In this case, we can no longer integrate out the top quark, since the cross section includes not only ratios of the form m_H/m_t, but ratios including the p_{T} as well [175].

In the case of BSM physics which presents itself at some high scale Λ (The TeV scale for example), an effective Lagrangian can be used to describe heavy new physics [168]. In the next section, we will discuss effective field theory (EFT), a useful tool to study the Higgs cross section at high transverse momentum.

3.4.2 Effective Field Theory

Given the lack of BSM found at the LHC thus far, we have good reason to believe that the scale of new physics Λ is much higher than the EW scale, around 1 TeV [176]. The LHC does not have the reach to produce these heavy new particles directly, but their effects are able to manifest in loop-induced processes. EFTs are tools and will not describe the theoretical structure of new physics, but the use of an EFT [177] approach for BSM physics searches offers a model-independent approach which acts as a useful guide for which interactions to focus on by providing us with information about whether or not the couplings of an EFT operator are amplified or suppressed. It allows us to indirectly probe mass scales above the current reach of the LHC, which can aid in decisions regarding future generations of colliders [178].

In an EFT, we introduce higher-dimension operators to the Lagrangian, typically in addition to the SM Lagrangian (SM EFT or SMEFT), of the form [179–181]:

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} + \sum_{d} \sum_{i} \frac{c_i^{(d)}}{\Lambda^{d-4}} \mathcal{O}_i^{(d)}. \tag{3.16}$$

Here the contributions of the d-dimensional ($d > 4$) operators $\mathcal{O}_i^{(d)}$ are suppressed by powers of Λ, the scale of new physics. The couplings $c_i^{(d)}$ are known as Wilson coefficients and encode the virtual effects of the heavy new physics into the low-energy

observables. Not only are the expressions for the operators of higher dimensions not yet known, but using an infinite number of operators is obviously mathematically intractable. By making some basic assumptions, we can greatly reduce this set of terms; by assuming that the baryon B and lepton L numbers are conserved for example, we can exclude all odd-dimension operators, which manifestly break B and L conservation [181]. Additionally, as higher-order operators are suppressed by higher powers of Λ, we can take the leading term, which has dimension-6, so that our EFT Lagrangian is:

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{c_i}{\Lambda^2}\mathcal{O}_i. \tag{3.17}$$

The expression of these operators is not unique, and they are generally formulated in either the Warsaw [181] or SILH [182] operator bases. Much like with the kappa framework, the Wilson coefficients will manifest in the branching fractions as modifications to the SM cross sections. Most strikingly, these modifiers can produce substantial deformations to the Higgs boson differential (in p_{T}) cross section without significantly altering the inclusive rate.

Effective field theories make versatile tools in searches for physics beyond the Standard Model. In an EFT framework, the LHC signal strength measurements can be included in a likelihood fit to place constraints on the masses and couplings of various extensions of the SM. For example, modifications to the Higgs coupling properties (in particular the Higgs couplings to bottom and top quarks and gluons) in two Higgs doublet models (2HDM) [150, 183–191] can be constrained through an EFT interpretation [192]. This is merely one example among many [193–198]. If we assume heavy BSM particles are too heavy to be directly produced at the LHC, a

SMEFT approach allows for a unified description of many possible signals for new physics.

Focus on loop-induced modifications such as gluon-Higgs fusion is warranted; given that we can be sure new physics is not obvious (or we would have found it), we need a large number of events in order to obtain constraints on these EFT parameters. If these effective couplings have non-zero values, we expect new physics to present itself through potentially large changes to the Higgs cross section at high p_{T} scales [199]. We will discuss the ggF-specific operators in more detail in Chapter 11, after presenting the results of the analysis, so that we can make comparisons between the ATLAS and CMS results and gauge our potential sensitivity to the EFT couplings. In the next chapter, we will outline some of the important considerations of the $H \to b\bar{b}$ measurement.

Chapter 4

Overview of the $H \to b\bar{b}$ Measurement

4.1 $H \to b\bar{b}$

A major success of the LHC's Run 2 was the direct observation of the Higgs boson decaying to a pair of bottom quarks in the associated vector boson (VH) production channel, where the vector boson decays leptonically to enable efficient triggering and a reduction of the multijet background [129]. The searches performed after Run 1 by ATLAS [200] and CMS [201] in this channel were not sufficient[1] to establish the coupling of the Higgs boson to bottom quarks. The increase in the center-of-mass energy in Run 2 of the LHC from 7 TeV to 13 TeV proved instrumental for the two experiments to reach the required sensitivity to claim evidence for this decay mode in VH production. The higher center-of-mass energy increased the signal cross sections by more than a factor of 2 [203], allowing both ATLAS [129] and CMS [24] to reach the required sensitivity[2] using the partial Run 2 dataset

[1] The combined searches for $H \to b\bar{b}$ by the ATLAS and CMS Collaborations in Run 1, at $\sqrt{s} =$ 7 and 8 TeV, evaluated for a Higgs boson mass of 125.09 GeV, resulted in an observed (expected) significance of $2.6\,\sigma$ ($3.7\,\sigma$) [202]

[2] The Run 2 observed (expected) signal significance was $4.9\,\sigma$ ($4.3\,\sigma$) and $4.8\,\sigma$ ($4.9\,\sigma$) by ATLAS and CMS, respectively. When combined with the Run 1 measurements at 7 and 8 TeV, an observed (expected) signal significance of $5.4\,\sigma$ ($5.5\,\sigma$) and $5.6\,\sigma$ ($5.5\,\sigma$) was seen by ATLAS and CMS, respectively [79].

of approximately $80\,\text{fb}^{-1}$. Additionally, the LHCb Collaboration has performed a search for $H \to b\bar{b}$ in the VH production channel [204] at a center-of-mass energy of 7 TeV. Using $1.98\,\text{fb}^{-1}$ of Run 1 data, LHCb's sensitivity in this search only put an expected 95% CL exclusion limit of 84 times the SM production rate [79, 204].

ATLAS [205] and CMS [26] performed searches for $H \to b\bar{b}$ in the vector boson fusion (VBF) production mode during Run 1. In the VBF channel, the Higgs boson is accompanied by two light-flavor quarks separated by a large rapidity gap, which provides a clean experimental signature to distinguish Higgs boson production from backgrounds [206]. An analysis using Run 2 data has also been performed by ATLAS in the inclusive channel [207] mentioned above, as well as one in which an extra photon [208, 209] is required [210]. The combination of these analyses has lead to a $2.0\,\sigma$ measurement of VBF $H \to b\bar{b}$ production [207]. Similarly, searches have been performed in the ttH ($H \to b\bar{b}$) channel by both ATLAS [27, 211–213] and CMS [214, 215]. These analyses are sub-divided based on the decay products of the top quarks: the 0-lepton (hadronic), 1-lepton, and 2-lepton channels. An update of the $H \to b\bar{b}$ decay channel was made by CMS with partial Run 2 data ($41.5\,\text{fb}^{-1}$)[3] and reported an observed (expected) significance of 3.7 (2.6) standard deviations [216].

Inclusive searches for the Higgs boson in the gluon-gluon fusion (ggF) production channel with $H \to b\bar{b}$ have limited sensitivity due to an overwhelming background from the inclusive production of $pp \to b\bar{b} + X$ via the strong interaction [79]. Figure 4.1 shows that the inclusive production rate for bottom quark production is roughly 10^7 times that of Higgs boson production via the gluon-fusion process. The background is dominated by gluon splitting to a pair of b-quarks, creating an enormous irreducible background for the Boosted $H \to b\bar{b}$ Analysis. There are sev-

[3] Not combined, but only in the fully hadronic channel [79].

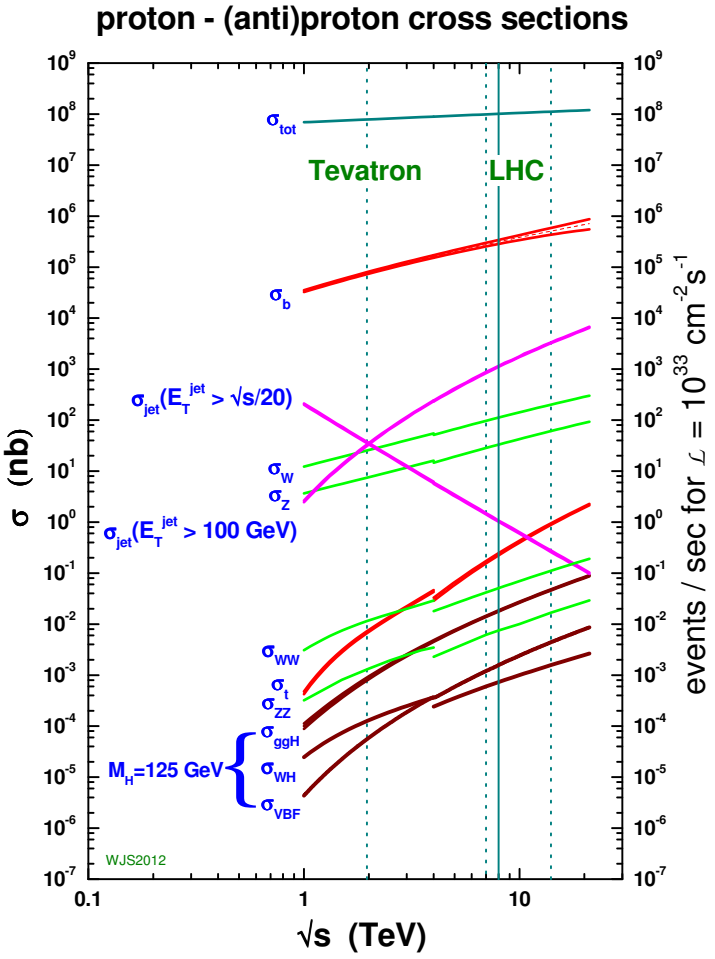


Figure 4.1: Production cross sections for several processes at a hadron collider as a function of the center-of-mass energy [217].

eral other backgrounds, and they will be discussed in more detail in Section 9.4. No meaningful results were produced with the LHC Run 1 dataset. The Run 2 increase in the center-of-mass energy to 13 TeV provided a harder transverse momentum spectrum for the Higgs boson with respect to the QCD background. ATLAS [218], in a previous search in the same channel as the analysis I present in this work, and CMS [219] performed searches for a Higgs boson with high transverse momentum decaying to a pair of bottom quarks in association with an energetic initial state radiation (ISR) jet. The observed significances were $1.6\,\sigma$ and $1.5\,\sigma$ by ATLAS and

CMS, respectively, using partial data from Run 2. The results for the full Run 2 of the LHC have been recently seen by ATLAS [220, 221] (the analysis presented in this dissertation) and CMS [222]. We will cover the results of the ATLAS analysis in Chapter 11, and summarize the latest CMS results in Chapter 12.

A pertinent question to pose is whether or not we are sensitive to this signal. We will end this chapter by defining what is meant by significance and performing a basic estimation of our expected significance. The final significance is extracted using the reconstructed invariant mass distribution of the Higgs boson, which is the best discriminant in an analysis studying $H \to b\bar{b}$ because it is the only relevant variable for which we have a clear a priori expectation. In the following section, we will discuss the objects called jets which are used to represent the four-momenta of the b-quarks and their decay products. Jets offer a useful representation for the hadronic energy produced in proton-proton collisions.

4.2 A Jet in a Nutshell

Protons consist of many partons (quarks and gluons) [223–226], each carrying a fraction of the total proton energy. In a high-energy collision involving two protons, a parton from each proton can transfer a significant amount of momentum, producing additional partons or other particles like W, Z, and Higgs bosons, which themselves can decay to quarks. These high-energy quarks and gluons are not directly observable in the final state of the collision, but instead undergo successive branchings to other particles at small angles [227], producing a collimated stream of quarks and gluons which leave their traces in the detector as seen in Figure 4.2. The initial high-energy quarks and gluons produce a parton shower [228, 229], where partons continually

produce branchings to additional partons with successively lower energies until their energy reaches the hadronization scale Λ_{QCD}, around 1 GeV [230]. The collimated nature of the parton shower results from the collinear divergence of QCD[4] [235, 236]. At this scale, color-charged particles undergo *hadronization* [237] and arrange themselves into color-neutral states called hadrons [235]. We can conceptualize this final state, consisting of a collimated bunch of hadrons, as a jet. A jet acts as a useful proxy for the high-energy quarks and gluons involved in the initial collision. For now, we will keep to this more conceptual view of jets, and a more complete treatment of jets and how they are reconstructed will be covered in Section 6.5.

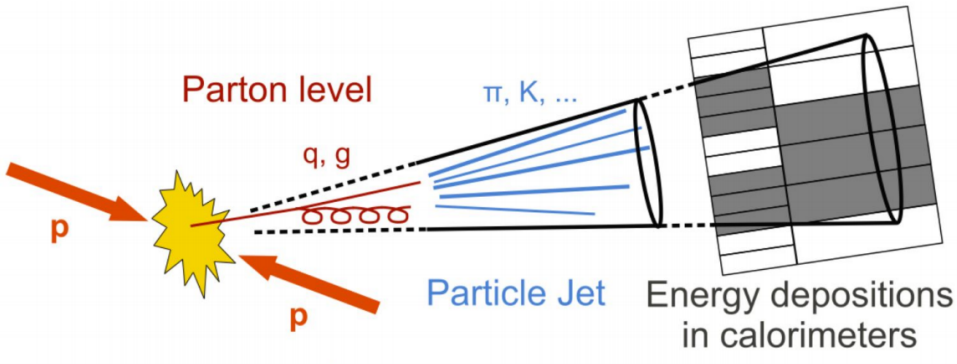


Figure 4.2: Proton-proton collision which produces a collimated cone-shaped stream of quarks/gluons which fragment into additional partons and form hadrons, leaving deposits in the trackers and calorimeters [238].

Measuring the $H \to b\bar{b}$ cross section revolves around our ability to identify jets coming from the decays of heavy hadrons. I believe one of the most interesting topics in experimental High Energy Physics is the study of b- and c-quarks, in part because they are only accessible in colliders, and because they directly connect the verification of the SM and the search for its possible deviations. Deviations found in the study of second- and third-generation particles, those with the largest cou-

[4] Collinear divergences are not actually physical divergences, but are the consequence of the choice of observable [231–234].

plings to the Higgs boson, can provide hints at new effects caused by BSM physics. For this reason, we must have reliably robust algorithms capable of selecting events containing b-quarks while simultaneously keeping backgrounds low to maximize sensitivity; *i.e* efficiency, purity, and background rejection are all critically important parameters to our b-tagging algorithms if we are to measure $H \to b\bar{b}$.

4.3 b-jet Identification (The Anatomy of Beauty)

The Cabibbo-Kobayashi-Maskawa (CKM) matrix [239, 240] describes the mixing between the three different generations of quarks in the SM (See Section 2.4.3). The probability that a bottom quark will decay to a charm quark, the most likely decay mode [241], is proportional to $|V_{cb}|^2$, $|V_{cb}| \approx (41.0 \pm 1.4) \times 10^{-3}$ [242]. The transition $b \to u$ is highly suppressed compared to $b \to c$ by the factor $|V_{ub}/V_{cb}|^2 \sim (0.1)^2$ [243]. Hadrons formed from bottom quarks (or beauty quarks) have some unique properties which can aid in their identification. Bottom quarks and their hadrons are not present in the final state; they predominantly decay to charm quarks (c-hadrons), which themselves decay to lighter-flavored quarks, and off-shell W bosons. W bosons either decay to light quarks, or to a lepton-neutrino pair. The proper lifetimes of b-hadrons are about 1.6 ps ($\langle c\tau \rangle = 450\,\mu$m) [244]. At current LHC energies, the mean flight length of a b-hadron can be on the order of a centimeter [243]. In addition to the lifetime, b-hadrons have the largest mass of any hadrons (top quarks decay too quickly to hadronize [245]), which leads to a high multiplicity of decay products; typically, b-hadrons average five charged particles per decay [246–249]. The large mass and long lifetime of the b-hadrons will cause the tracks left behind by its decay products in the detector to have a non-zero impact

parameter – meaning that the distance of the extrapolated track's closest approach to the hard-scatter vertex is non-zero [250]. Bottom quarks also exhibit different fragmentation properties compared to lighter quarks and tend to carry away a much larger fraction of the energy and momentum than gluons and light jets [246]. The displaced tracks created in b-decays allow us to reconstruct a secondary vertex (SV), seen in Figure 4.3, which provides a strong signature for a b-quark jet. Figure 4.4 shows the differences in the reconstructed invariant mass, decay-length significance, and energy fraction of the SV for b-, c-, and light-jets in simulated $t\bar{t}$ events. The b-tagging algorithms used in the Boosted $H \to b\bar{b}$ Analysis will be covered in more detail in Section 6.4

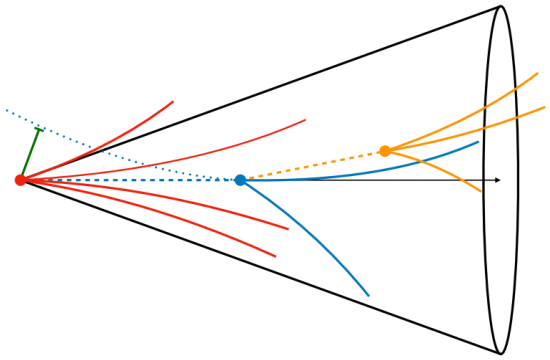


Figure 4.3: Diagram of a typical b-hadron (blue) decay chain seen in the transverse plane, which usually involves a decay to a c-hadron (hadrons containing c-quarks) (yellow) [251]. The blue dotted line illustrates the extrapolation of a single track and the green line is the impact parameter for the track. The black arrow represents the jet axis.

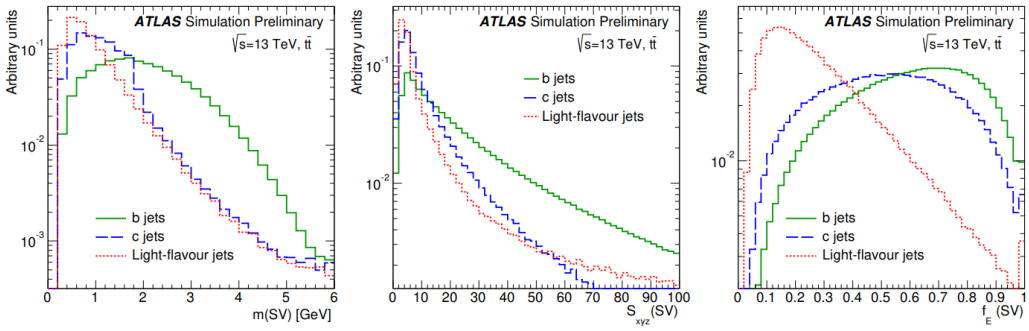


Figure 4.4: Invariant mass of tracks at the secondary vertex (left), the 3D SV decay-length significance (middle), and energy fraction of SV tracks (right) for b-, c-, and light-quark-initiated jets [246].

4.4 The Higgs Boson at High Transverse Momentum

Consider the case of Higgs production via ggF, where a gluon can be emitted in the initial state (called initial state radiation or ISR), providing the Higgs with a higher transverse momentum (Figure 4.5). In this scenario, we call the Higgs boson *boosted*. This boost has a useful feature; at rest, the Higgs decaying into a pair of

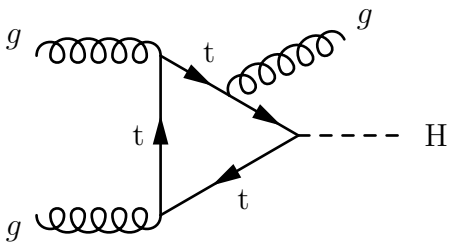


Figure 4.5: Feynman diagram showing the Initial State Radiation (ISR) of a gluon emitted from the top-quark loop.

b-quarks would be detected roughly back-to-back in the detector, a scenario which can be imitated by a variety of other processes ($Z \to b\bar{b}$, $t\bar{t}$, etc). When matching some pair of jets (out of all possible jet pairs) in the event, we run the risk of matching the wrong pair of jets to the heavy resonance. However, at high momentum, the angle between the decay products becomes smaller; small enough to be reconstructed into

the same physics object – a large-radius (large-R) jet, such that mistaking the pair of jets resulting from the decay products becomes very unlikely. An example of a large-R jet can be seen in Figure 4.6 and we will discuss large-R jets in more detail in Section 6.5. The angular separation for the decay products of a massive boosted object like the Higgs boson can be expressed as [236, 252]:

$$\Delta R = \frac{1}{\sqrt{x(1-x)}} \frac{m_H}{p_{T_H}}, \tag{4.1}$$

where x is the momentum fraction carried by the decay products and ΔR is an angular distance metric, to be defined in Section 5.5. In the case of a pair of b-quarks (any pair of particle-antiparticle), the momentum fraction for each parton is $x = 0.5$, such that our opening angle for the boosted decay products becomes:

$$\Delta R = 2\frac{m_H}{p_{T_H}}. \tag{4.2}$$

Searching for the Higgs boson in this kinematical phase space has the advantage of allowing us to search for hints of new physics using the EFT interpretation that we saw in Section 3.4, where we can study the dynamical properties of the Higgs boson. At high p_{T}, we have access to loop effects in Higgs production because the high energy allows us to resolve quasi-point-like couplings. Although we are currently not very sensitive to these effects with the amount of data collected so far, we can begin to glimpse hints at new physics.

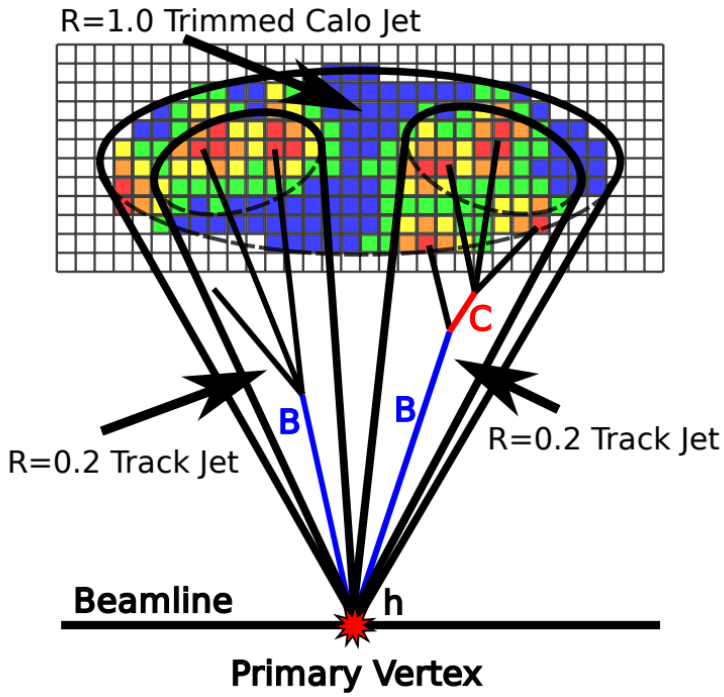


Figure 4.6: Example of a large-radius jet from $H \to b\bar{b}$ [251].

4.5 Significance

Often, we physicists talk about the significance for our signal in experimental measurements or searches for new physics. As with most statistical testing, we will have some null hypothesis (*e.g.* only QCD background) and an alternative hypothesis (*e.g.* QCD background plus some Higgs boson signal). The significance, expressed as some number of sigma, allows us to quantify how statistically likely (or unlikely) it is that we've erroneously rejected the null hypothesis. A high significance means that it's very unlikely that we've falsely excluded the null hypothesis. In High Energy Physics, $5\,\sigma$ is often the "gold standard" of discovery, meaning that a discovery has a roughly 1 in 3.5 million chance of being a statistical fluke. A low significance can still get us very excited[5], but largely because it hints that discovery may be just around the corner as more data arrives. There are many ways to calculate significance S

[5]Sorry 750 GeV diphoton excess [253–256], you will be missed.

in terms of the number of signal n_s and background n_b events. For simplicity, we shall use a definition that is common and simple. Because most measurements in particle physics are counting experiments, they are subject to Poisson statistics with the likelihood function [257]:

$$L(N; n_s, n_b) = \frac{(n_s + n_b)^N}{N!} e^{-(n_s+n_b)} \tag{4.3}$$

The number of observed events N should be compared to the number of expected background events n_b in the null hypothesis ($n_s = 0$). If n_b is sufficiently large, the distribution can be approximated as a Gaussian with average n_b and standard deviation $\sqrt{n_b}$. For an excess in the data, our signal $n_s = N - n_b$, should be compared with the expected standard deviation $\sqrt{n_b}$ [258], yielding the sensitivity S:

$$S = \frac{n_s}{\sqrt{n_b}} \tag{4.4}$$

I should stress the point that in the Boosted $H \to b\bar{b}$ Analysis, the significance is extracted from a fit to data, as outlined in Chapter 10, in a far more complicated scheme than what is discussed here. Still, this simple approach is quite useful.

The number of events we expect to observe for a given process is the cross section for that process times the integrated luminosity L (see Section 5.3). This number is then modified by the acceptance A and efficiencies[6] ϵ associated with our method of search (*i.e.* mass range, trigger efficiency, b-tagging efficiency, etc.).

$$n_{\text{events}} = \sigma L A \epsilon. \tag{4.5}$$

[6] We can (and do) have multiple efficiencies to consider. The efficiencies are multiplicative such that for i efficiencies, our overall efficiency is $\epsilon = \prod_i \epsilon_i$

Let's take a moment to discuss the expected number of events, for ggF (the dominant Higgs production mode) and QCD dijet events, so we can make a back-of-the-envelope calculation of the expected significance. While these cross sections can be computed theoretically, we can simplify our lives by using MC as our starting point to make some rough calculation of our expected significance, which I define as:

$$S = \frac{n_s}{\sqrt{n_b}} = \frac{N^{\mathrm{MC}}_{ggF} \cdot BR(H \to b\bar{b}) \cdot A_{ggF} \cdot \epsilon_{b-\mathrm{tag}}}{\sqrt{N^{\mathrm{MC}}_{QCD} \cdot BR(g \to b\bar{b}) \cdot A_{QCD} \cdot \epsilon_{b-\mathrm{tag}}}}. \tag{4.6}$$

Here, N^{MC}_X $(X = ggF, QCD)$, are the number of events satisfying some mass/p_{T} threshold of the trigger (in this case, a leading jet with $p_{\mathrm{T}} > 450\,\mathrm{GeV}$ and $m > 60\,\mathrm{GeV}$). The trigger is addressed in general in Section 6.1 and the analysis-specific triggers are outlined in Table 7.3 of Section 7.2. At this stage, the events are not required to contain jets of any particular flavor, so we will use the relative branching fractions to estimate the number of events we would be interested in. The current Run 2 luminosity recorded by ATLAS is approximately $139\,\mathrm{fb}^{-1}$ [259]. With that luminosity, the expected numbers of events passing this trigger requirement for ggF and QCD events are $N^{\mathrm{MC}}_{ggF} \simeq 3250$ and $N^{\mathrm{MC}}_{QCD} \simeq 570{,}000{,}000$, respectively. This alone produces an expected significance of approximately 0.14. We now take into consideration the branching ratios for the Higgs boson and gluons to produce $b\bar{b}$ pairs. For the Higgs boson this is 0.584 [79] and for $g \to b\bar{b}$ it is approximately 2.2×10^{-3} [260–264]. We must also take into account the efficiency for tagging a jet originating from a b-hadron, which we'll estimate as 70% [246]; the efficiency for tagging two would then be approximately 50%. Lastly, to avoid counting the events well outside of the expected Higgs-mass window, we should restrict our mass range.

The resolution on the Higgs boson decaying to $b\bar{b}$ pairs is roughly 10% [265], so it is reasonable to restrict our mass range to where we expect the reconstructed mass peak to be, between about 105 and 140 GeV, which also allows us to largely avoid the long mass tails created by $V+$jets and top-quark backgrounds. The acceptance for ggF and QCD dijet events in this range is 0.20 and 0.07, respectively. The estimated expected significance with these values is roughly $0.9\,\sigma$. This may not be the $5\,\sigma$ we want, but it is helpful to ground our expectations to something reasonable. Additionally, this estimate should be taken with a grain of salt as we only considered one of the backgrounds, a single production mode, and made several simplifying assumptions. If the end result of the analysis indeed has a relatively low significance, we would not yet be capable of extracting strong constraints on the Wilson coefficients in an EFT interpretation, but comparisons may still be interesting.

In the following chapters, we will cover the ATLAS detector at the LHC. To produce a statistically significant number of interesting events (events with Higgs bosons in our case), the LHC collides a large number of protons together simultaneously, adding a considerable number of interactions in each event. The extra energy and tracks present in the event adds a significant layer of difficulty, which will require us to understand the effects of the high luminosity at the LHC. In the ATLAS detector, we take a snapshot of an event, which is composed of a multitude of hits in various detector subsystems. Any analysis hinges on precise high-granularity sensors which use robust pattern recognition to reconstruct the detector hits into physics objects such as photons, electrons, hadrons, etc. This requires excellent track and vertex resolution for tracks created by charged particles, and the ability to measure the energy of particles produced in a collision. We must rely on sophisticated al-

gorithms to tag the flavor of hadrons based on their unique decay characteristics. Correctly identifying b-flavored decays with a low fake rate (and high efficiency) is crucial in obtaining the highest possible significance.

Chapter 5

The ATLAS Detector at the Large Hadron Collider

5.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [18, 19], at the forefront of discovery in particle physics, is the world's largest science facility ever constructed. It is operated by the European Center for Nuclear Research, or CERN [20], founded in the aftermath of WWII in 1954. CERN is an international organization established to promote fundamental research in particle physics and consists of over 12,000 users from more than 23 member states, 9 associate member states, as well as several observer and non-member states. The LHC is a giant ring – 26.7 km in circumference and located about 100 m below ground on the French-Swiss border (Figure 5.1) – at the base of the Jura mountain range to its northwest and the Geneva Airport at its southeast. Its primary tasks were to further probe the SM, discover the Higgs boson, and discover new particles. To carry out this ambitious program, the LHC hosts four large experiments: ATLAS [266], CMS [267], ALICE [268], and LHCb [269]. As its name

implies, the LHC is designed to accelerate and collide hadrons, namely protons and lead ions, to observe rare processes with energies previously only present in the early universe.

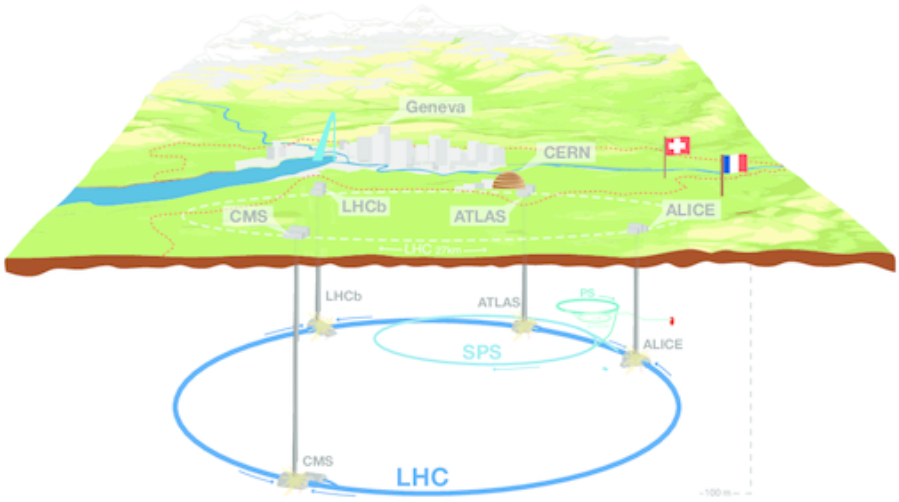


Figure 5.1: The LHC ring on the border between France and Switzerland near Geneva. The four main experiments: ATLAS, CMS, ALICE, and LHCb are located at points along the circumference. Also shown is the Super Proton Synchrotron (SPS), the last accelerator in the accelerator chain which injects protons into the LHC [270].

5.2 Layout and Design

The LHC forms a large ring, 26.7 kilometers in circumference, capable of accelerating counter-rotating beams of protons, with an energy of 6.5 TeV per beam, to an unprecedented center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$. It is not actually arranged as a circle, but an octagon with rounded arcs where the charged-particle beams are bent along their trajectories. Protons are accelerated along straight sections of the ring in counter-circulating beams by the use of radio frequency (RF) cavities [271] and steered by superconducting dipole magnets [272], cooled to 2 K using liquid helium [273]. The beams cross at the locations of the major experiments where the products of the collisions can be measured.

To bring protons up to speed, hydrogen atoms are stripped of their electrons and sent through a series of progressively larger accelerators (see Figure 5.2) – analogous to the changing of gears in a car's transmission. First they are injected from the linear accelerator (LINAC4 [274]), which accelerates the protons to 160 MeV and into the Proton Synchrotron Booster (PSB) [275] to bring them up to 1.4 GeV. From there, they are injected into the Proton Synchrotron (PS) [276], where they reach the center-of-mass energy of 25 GeV. In the penultimate stage, protons are accelerated to 450 GeV in the Super Proton Synchrotron (SPS) [277]. Finally, they are injected into the LHC, where they circulate for roughly 20 minutes to reach their maximum center-of-mass energy of 13 TeV. The beam lifetime for the LHC is approximately 10 hours [278].

The proton beams are directed around the LHC ring using 1232 15-m-long niobium-titanium (NbTi) super-conducting dipole magnets [280] (Figure 5.3) that bend the protons' trajectories with an 8.33 T magnetic field [272]. To keep the transverse width of the proton bunches in a tightly packed beam, 392 quadrupole magnets, each 5–7 m long, are used to squeeze the protons close together into tightly packed bunches.

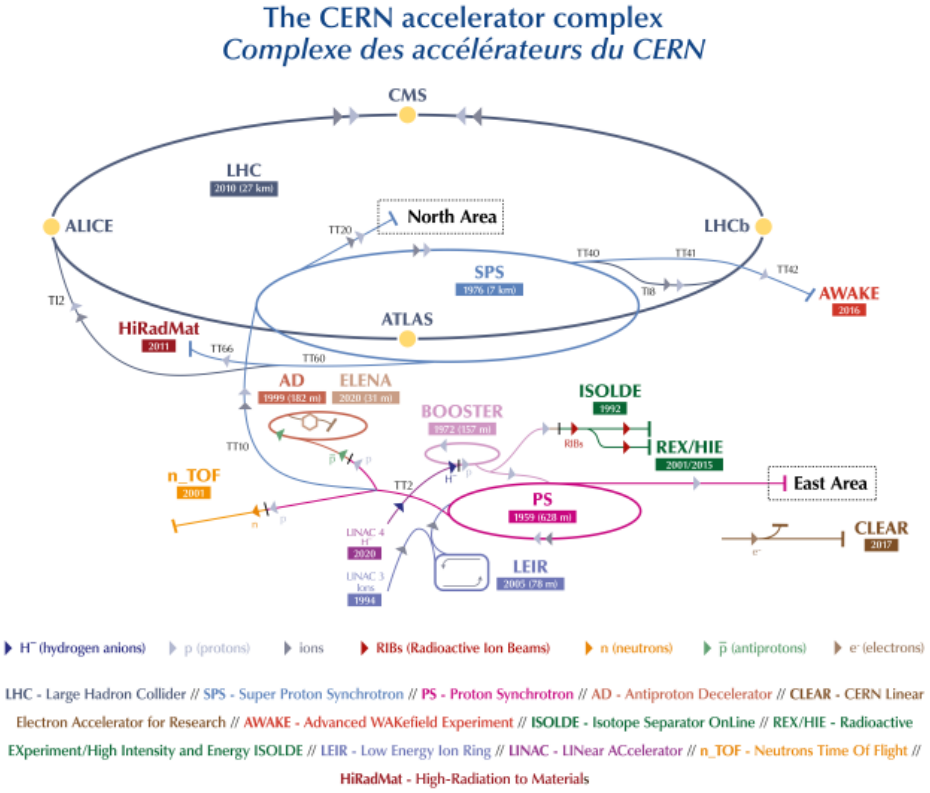


Figure 5.2: Diagram of the LHC accelerator complex [279]. The smaller accelerators are chained together to bring the protons up to a sufficient center-of-mass energy before entering the LHC.

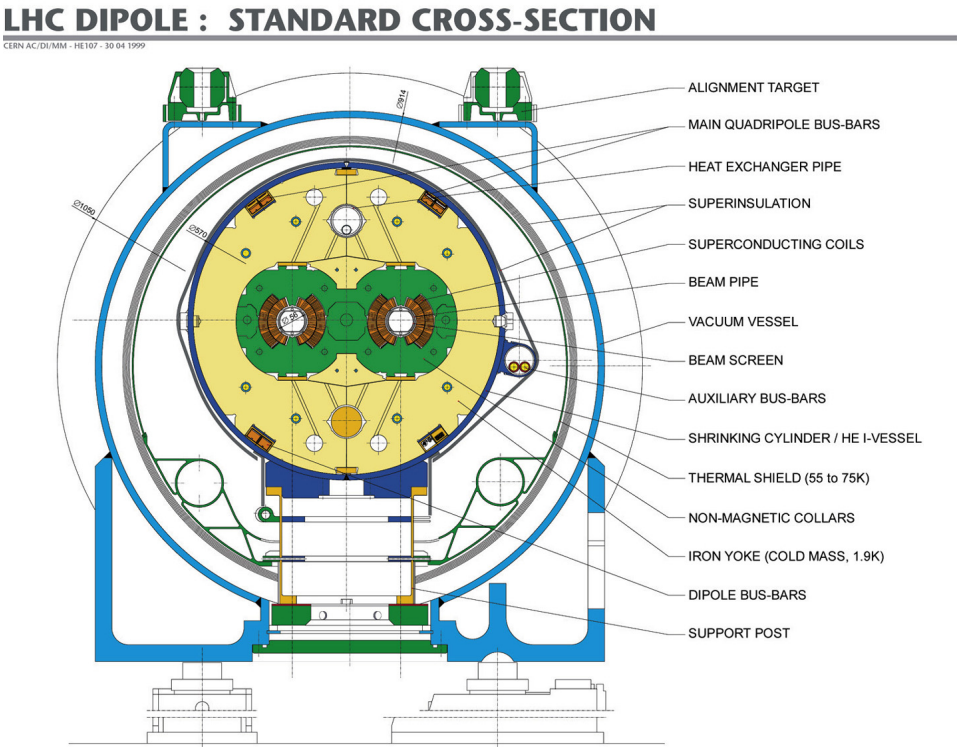


Figure 5.3: Cross section of an LHC dipole magnet with the major components labeled [281].

5.3 Luminosity and Pileup

At the core of the LHC's success, as well as a source of significant challenges, is its incredible luminosity. At a basic level, the luminosity is simply the particle flux, or number of particles passing through some space per unit time. The rate R of production for a particular process is determined by both the machine's instantaneous luminosity $\mathcal{L}$ and the cross section σ for that process:

$$R_{\text{process}} = \mathcal{L} \cdot \sigma_{\text{process}} \tag{5.1}$$

The rate observed for any process will be lower than the rate produced depending on our efficiency for observing it. It's useful to describe luminosity in two aspects: the instantaneous and integrated luminosities, $\mathcal{L}$ and L, respectively. The instantaneous luminosity, as the name implies, is the luminosity at any given moment. The beams themselves are formed from small, mostly Gaussian-like distributions of protons which collide head-on[1] in bunches every 25 ns, allowing us to express the instantaneous luminosity as [282]:

$$\mathcal{L} = \frac{N_1 N_2 f N_b}{4\pi\sigma_x\sigma_y}, \tag{5.2}$$

where N_1 and N_2 correspond to the number of particles per bunch in each of the beams, f is the frequency of revolution around the ring, $\sigma_{x/y}$ are the transverse widths of the proton bunches, and N_b is the number of bunches in the beam. At the LHC, bunches are arranged into *trains* with a 25 ns spacing in time. A bunch train is typically composed of several bunches, which may be separated from each

[1] Not exactly, but at a small crossing angle of about 300 µrad [282].

other by a smaller number of empty bunches. Approximately 2544 bunches in total, each like a packet containing 1.1×10^{11} protons with transverse widths of 16.6 µm, revolve around the ring with a fixed frequency of just over 11 kHz [19]. These values together produce an instantaneous luminosity of approximately $10^{34}\,\mathrm{s}^{-1}\,\mathrm{cm}^{-1}$. Some of the parameters change somewhat from year to year, and the values for Run 2 beam parameters are included in Table 5.1 [283]. One parameter shown in the table, but not in Equation 5.2 is β^*. The beta function is a parameter that characterizes the transverse width of the beams. Essentially, it is the distance over which the beam width is reduced by a factor of two. So a smaller β^* corresponds to a "squeezing" of the beam and thus a higher luminosity. By integrating over the instantaneous luminosity over time, we arrive at the total, integrated luminosity:

Table 5.1: Selected LHC parameters for pp collisions for LHC Run 2, at a center of mass energy $\sqrt{s} = 13\,\mathrm{TeV}$ in 2015-2018. The values shown represent the peak accelerator performance during normal physics operation. During 2017, the LHC was operated in two modes: standard 25 ns bunch trains with long trains, and "8b4e" (8 bunches with protons followed by 4 empty bunches). The instantaneous luminosity was leveled by beam separation to about $\mathcal{L}_{\mathrm{peak}} = 16 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ for part of the 8b4e period. The early 2015 data, consisting of $0.1\,\mathrm{fb}^{-1}$ integrated luminosity, which was delivered using a 50 ns spacing is not included [283].

Parameter	2015	2016	2017	2018
Maximum number of colliding bunch pairs (N_b)	2232	2208	2544/1909	2544
Bunch spacing (ns)	25	25	25/8b4e	25
Typical bunch population (10^{11} protons)	1.1	1.1	1.1/1.2	1.1
β^* (m)	0.8	0.4	0.3	0.3–0.25
Peak luminosity $\mathcal{L}_{\mathrm{peak}}$ ($10^{33}\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$)	5	13	16	19
Peak number of inelastic interactions/crossings ($\langle\mu\rangle$)	16	41	45/60	55
Luminosity-weighted mean inelastic interactions/crossings	13	25	38	36
Total delivered integrated luminosity (fb^{-1})	4.0	38.5	50.2	63.4

$$L_{\mathrm{int}} = \int \mathcal{L}\,dt, \tag{5.3}$$

The cumulative luminosity vs. time delivered (recorded) to (by) ATLAS can be seen in Figure 5.4 [259].

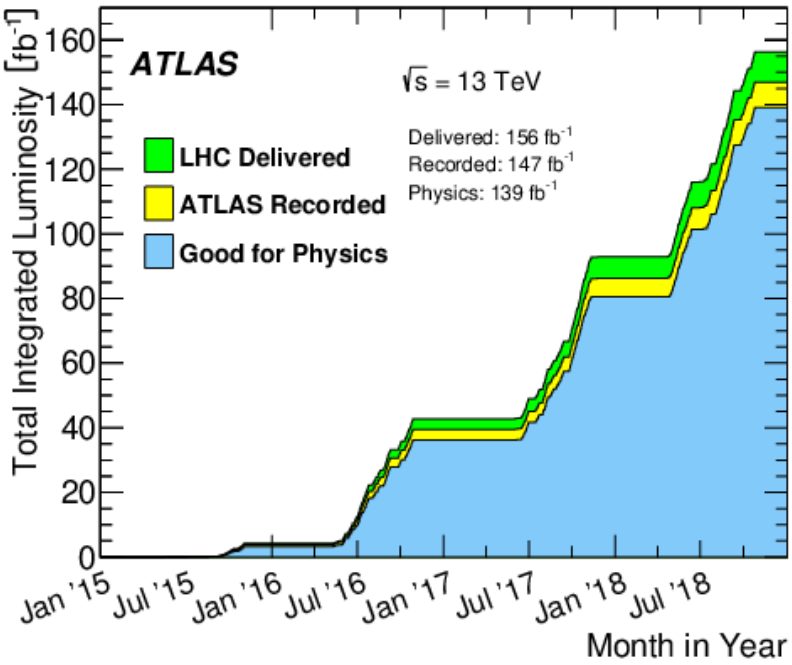


Figure 5.4: Cumulative integrated luminosity delivered to (green) and recorded by (yellow) ATLAS between 2015 and 2018 during stable beam pp collision data-taking at $\sqrt{s} = 13\,\text{TeV}$. Also shown is the cumulative integrated luminosity certified for physics analysis usage (blue) [284].

In any bunch crossing, the hardest interaction, involving two partons from two protons, is the event of interest. To be able to detect rare processes in the detector, it is crucial that the luminosity be high enough so that we may observe a statistically significant number of high-energy collisions in a reasonable amount of time. More data allows us to study the properties of the SM in greater detail and gives us the opportunity to observe rare new phenomenon. These high intensities produce a dilemma however; the number of proton-proton interactions per bunch crossing increases with an increase in luminosity. We call this pileup, and it can manifest in two forms: in-time pileup results from multiple interactions that occur within the same bunch crossing, and out-of-time pileup results from the effects outside of the

bunch crossing caused by the long integration times of the electronics and various detector systems. Pileup tends to contribute mainly at low transverse momentum and produces a diffuse background of particles around the event of interest, distorting the final state observables' kinematics and substructure (especially in the case of jets).

Pileup is characterized as the number of interactions μ – whose value is subject to Poisson fluctuations – and the time-averaged pileup as $\langle\mu\rangle$. The mean number of interactions per bunch crossing is calculated from the instantaneous per-bunch luminosity as [285]:

$$\mu = \frac{\mathcal{L}_{\text{bunch}}\sigma_{\text{inel}}}{f_r}, \tag{5.4}$$

where $\mathcal{L}_{\text{bunch}}$ is the per-bunch instantaneous luminosity, σ_{inel} is the inelastic cross section (taken to be approximately 80 mb for 13 TeV collisions [286–288]), and f_r is the LHC revolution frequency. The pileup distributions for data taken in Run 2 between 2015 and 2018 can be seen in Figure 5.5.

In the last part of Run 2, the ATLAS and CMS experiments achieved an average pileup of $\sim$ 35 and $\sim$ 77 simultaneous pp collisions, respectively [290, 291]. Each additional pileup collision at 13 TeV adds an average p_{T} of $\sim$ 900 MeV per unit area in the rapidity-azimuthal ($y-\phi$) plane[2] [292], adding extra energy into each recorded event. This can become significant, especially in the Boosted $H \to b\bar{b}$ Analysis, which uses jets with a large area. Beyond affecting the kinematics of a jet, pileup produces additional hits in the detector, adding many extra tracks and primary vertices. This can adversely affect jet flavor-tagging algorithms, which are critical in identifying

[2]The detector coordinates are discussed shortly.

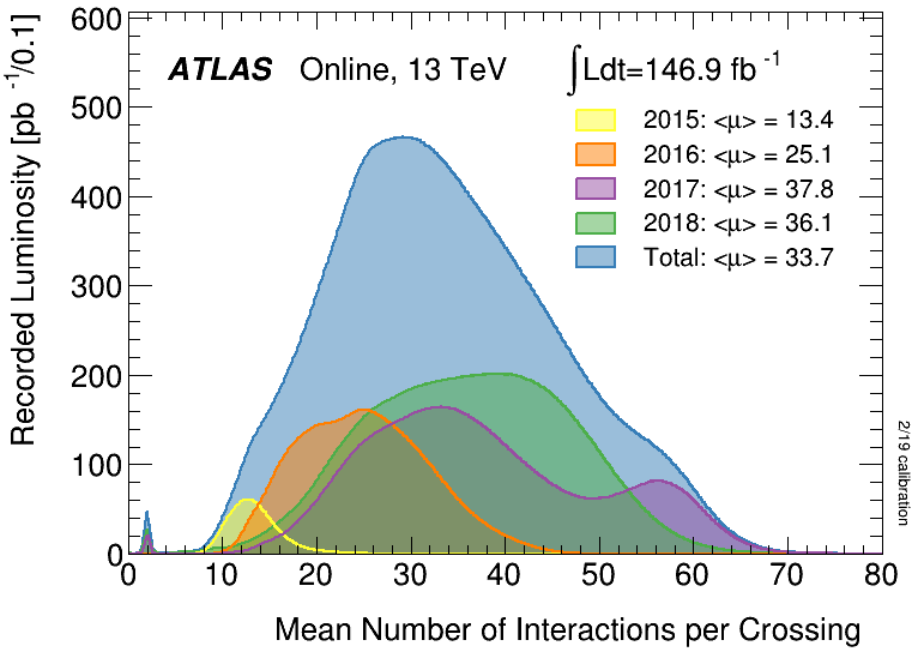


Figure 5.5: Luminosity-weighted distribution of the mean number of interactions per crossing for the ATLAS Run 2 pp-collision data at $\sqrt{s} = 13\,\text{TeV}$ center-of-mass energy. The mean number of interactions per crossing corresponds to the mean of the Poisson distribution of the number of interactions per crossing calculated for each bunch [289].

the b-jets from the Higgs boson. The mitigation strategies to reduce pileup in jets are discussed in more detail in Section 6.5.3 as well as in Reference [292]. Flavor tagging will be discussed in Section 6.4.

5.4 ATLAS Detector Overview

The ATLAS (**A** **T**oroidal **L**HC **A**pparatu**S**) detector [1, 2], shown to scale in Figure 5.6, is a general-purpose detector with a cylindrical geometry located at Point 1 of the LHC. It is 46 m long, 25 m tall, and weighs in at about 7000 metric tons – making the titanic name ATLAS apropos indeed. It is composed of several detector subsystems surrounding the interaction point where protons collide, giving it an almost hermetic, 4π solid-angle coverage.

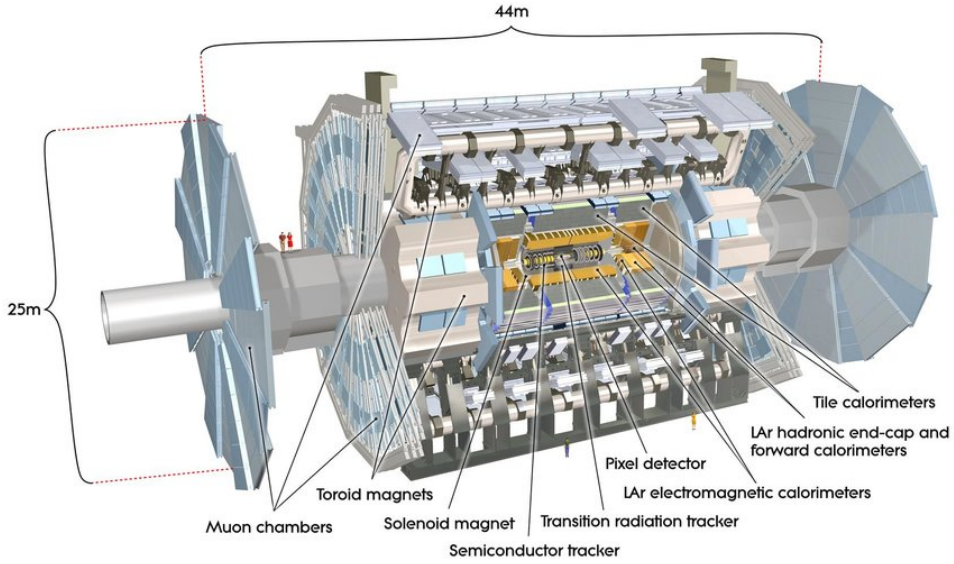


Figure 5.6: Cutaway diagram showing the ATLAS detector and its subsystems. For scale, see the human sized figures (if you can find them) [293].

Protons are sent hurtling around the LHC ring by the magnets discussed in Section 5.2 to collide at the center of the ATLAS detector. In all, over 100 million readout channels provide us with an enormous flux of data through more than 3000 km of cabling. The following sections in this chapter give a brief overview of ATLAS detector and its subsystems, which can be found in considerable detail in References [1, 2].

5.5 Coordinates

ATLAS uses a right-handed coordinate system which is centered on the nominal interaction point (*i.e.* the detector's center). The x-axis points towards the center of the LHC ring and the y-axis upwards towards the surface so that the x-y plane (or transverse plane) is perpendicular to the beam axis. This leaves the $+z$ direction to lie along the beam axis, going counter-clockwise with the LHC ring if viewed from

above. The detector's cylindrical geometry and back-front symmetry make it useful to use the polar angle θ, which is the angle measured from the beamline, and the azimuthal angle ϕ, which lies in the transverse plane. The momenta measured in the ATLAS detector is split into a longitudinal component p_z, and transverse component p_{T} ($p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}$). The transverse momentum has been referenced multiple times so far, but we have yet to see why it is such an important variable. In experiments where the particle velocities are so close to the speed of light, we prefer to define quantities which do not change depending on whether it is in the lab frame or the frame boosted along the beam, or longitudinal axis (z). The momentum $\vec{p}$ is not the relevant measure of momentum when events can have an arbitrary boost along the beam axis, which can be difficult to precisely measure. The p_{T} is Lorentz invariant for boosts along the beam axis, and allows us to relate directly to observables like the invariant mass. The transverse momentum is also used to define the vertex where the hard-scatter occurred. Other geometric quantities invariant under a boost are also desirable, so we define an object's rapidity y, a measurement of the velocity of a particle parallel to the beam axis, in terms of the energy and longitudinal momentum:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right). \tag{5.5}$$

This is a useful quantity because differences in rapidity are Lorentz invariant, whereas differences in θ are not. For highly relativistic particles, such as those seen in the LHC, this quantity can be difficult to obtain because the total momentum of the individual partons in proton-proton collisions is not well known. Fortunately, at high energy the rapidity is equivalent to the pseudorapidity in the massless limit,

such that $E \gg m \implies E \approx |\vec{p}|$. The pseudorapidity is expressed more simply in terms of a geometrical quantity, the polar angle θ, which is far easier to measure. The pseudorapidity η is defined as

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \tag{5.6}$$

Here, $\eta = 0$ perpendicular to the beam axis and $\eta \to \infty, -\infty$ as $\theta \to 0, \pi$. This is shown in Figure 5.7.

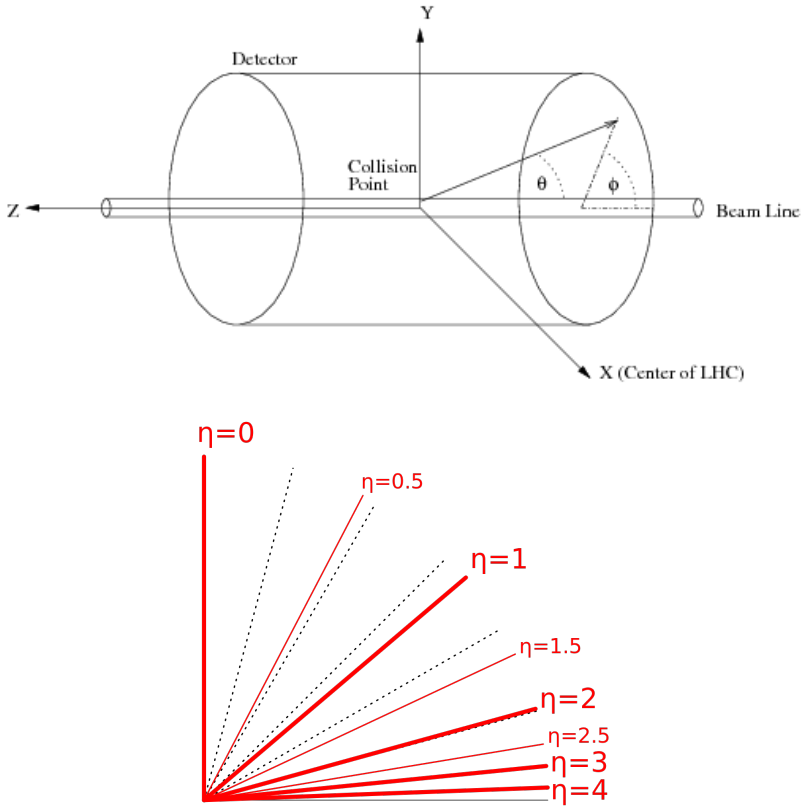


Figure 5.7: (Above) Diagram showing the polar angle θ and azimuthal angle ϕ with respect to the collision point [294]. (Below) Plot showing η for various polar angles [295] (black dashed lines in 15 degree increments).

The angular position of an object is often expressed in terms of the pseudorapidity and the azimuthal angle ϕ, which allows us to define an angular distance

metric ΔR:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{5.7}$$

Because ϕ lies in the transverse plane, ΔR is also Lorentz invariant[3] for boosts along the z-direction.

5.6 ATLAS Sub-Detectors

The ATLAS detector is made up of multiple detector subsystems, arranged in cylindrical layers around the interaction point where protons collide. We can group these subsystems into 3 main categories, starting with the inner-most and working outwards:

- The Inner Detector (ID) [296, 297]: embedded in a 2 T axial magnetic field, the three subsystems of the ID are designed to precisely reconstruct the position and momentum of charged particles.
- Calorimeter Systems (ECal and HCal) [298–300]: absorb photons, electrons, and hadrons to measure their energy and position.
- Muon Spectrometer (MS) [301]: embedded in a 4 T (maximum) toroidal magnetic field. Provides additional position and momentum measurements for muons (with the ID), which largely escape absorption in the calorimeters.

Each detector subsystem has a layered, cylindrical geometry around the interaction point known as the *barrel*, providing the best resolution in the central region

[3]For massless particles, ΔR is Lorentz invariant for longitudinal boosts. For massive particles with high transverse momentum, it is approximately invariant (*i.e.* $y \approx \eta$)

of the detector. At the ends of the ATLAS detector, the subsystems are layered with circular *endcaps* which provide the hermiticity needed to account for all of the energy and momentum of the collision. Barrel components tend to be easier to describe, as their geometry tends to be more uniform. In what follows, there will be more attention paid to barrel regions and we will not fully describe every characteristic of each subdetector. Rather, we will briefly outline what each system is responsible for and highlight the most important aspects for our purposes, namely the spatial and energy resolution, where appropriate. More details can be found in the ATLAS Technical Design Report (TDR) [1, 2], the subdetector TDR References in bullet point above, and the summary tables in Appendices B.1–B.2.

5.6.1 The Inner Detector

The ID is composed of three separate detector subsystems shown in Figure 5.8: the Silicon Pixel Tracker and the Insertable B-Layer (IBL), the Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT). These three systems all rely on detecting the ionization caused by a charged particle traversing the detector. The Pixel Tracker and SCT are both made of silicon under a bias voltage so that a charged particle in the doped silicon bulk will create electron-hole pairs through ionization. The TRT exploits the ionization of a gas inside a straw tube and transition radiation at the many interfaces of the inhomogeneous material. All three components are immersed in a 2 T solenoidal magnetic field [302] so that charged particles curve in the transverse direction, allowing us to measure not just their position, but their momentum. The ID has a transverse impact parameter resolution of $22.1 \pm 0.9\,\mu\text{m}$ and a relative momentum resolution of $\sigma_{p_\text{T}}/p_\text{T} = (4.83 \pm 0.16) \times 10^{-4}\,\text{GeV}^{-1} \times p_\text{T}$ for high momentum tracks [303].

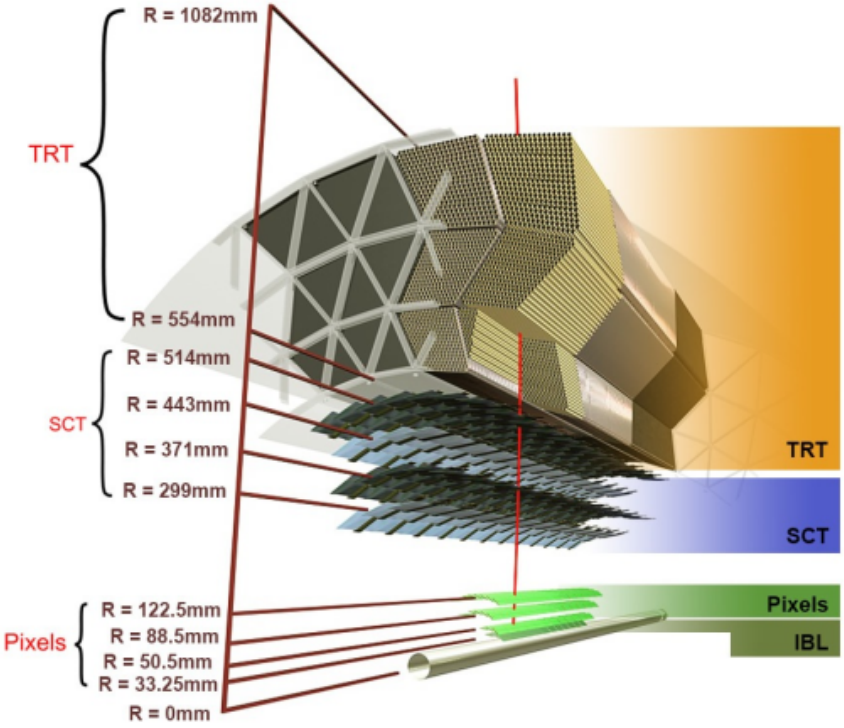


Figure 5.8: A radial view of the ID showing a charged particle (red vertical line) traversing through it [304].

Pixel Detector

The pixel system, highlighted green in Figure 5.8, has the highest granularity of all detector subsystems. This is critical for pattern recognition algorithms to reconstruct tracks and vertices in the high occupancy and pileup environments so close to the interaction. The innermost pixel layer ($r = 33.25\,\text{mm}$) is the Insertable B-Layer (IBL) [305–307], which consists of 14 staves covering the region $|\eta| < 3.03$ that contain over 12 million silicon pixels of size $(r-\phi) \times z = 50\,\mu\text{m} \times 250\,\mu\text{m}$, each with a spatial resolution[4] of $(r-\phi) \times z = 8\,\mu\text{m} \times 40\,\mu\text{m}$. The IBL's smaller pixel size, compared to the pixel detector, greatly improves position measurements and aids in identifying primary and secondary vertices, the latter of which is essential in

[4]The upper limit for resolution of a 2D uniform distribution is area/12, but better resolution can be achieved by using charge measurement (and charge sharing) to determine which pixel the particle was closer to, as well as additional measurements in other layers of the detector. See Ref [308] for more details.

identifying tracks coming from b-quarks. The addition of this layer in Run 2 improved the resolution of the longitudinal and transverse impact parameters by more than 40% for tracks with $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$ [309]. The original pixel detector [310] consists of three barrel layers ($50.5 < r < 122.5\,\mathrm{mm}$) comprised of 1744 pixel-sensor modules ($10\,\mathrm{cm}^2$) with each module containing 46,080 pixels $(r-\phi) \times z = 50\,\mu\mathrm{m} \times 400\,\mu\mathrm{m}$ in size, providing a spatial resolution of $(r-\phi) \times z = 10\,\mu\mathrm{m} \times 115\,\mu\mathrm{m}$ in the barrel region.

Semiconductor Tracker (SCT)

The Semiconductor Tracker (SCT) [296, 297], highlighted in blue in Figure 5.8, is designed to provide between four and nine precision spatial measurements per track. Combined with the measurements in the IBL and pixel, the SCT measurements contribute to precise determination of the transverse momentum, impact parameter, and vertex for charged particle tracks. The SCT is composed of 4088 two-sided silicon micro-strip modules arranged in four barrel layers ($299 < r < 514\,\mathrm{mm}$) [311] and two endcaps with nine wheels each. Barrel SCT modules ($6.36\,\mathrm{cm} \times 6.40\,\mathrm{cm}$) have 768 strips with an $80\,\mu\mathrm{m}$ pitch. The strips are oriented along the z-direction, providing greater resolution in the $r-\phi$ plane. To improve the resolution in the longitudinal direction, the double-sided modules are positioned not with the strips on each side parallel, but with one side parallel to the beam and the other at a small rotation of 40 mrad, which allows for a better position measurement along the length of the strip. The spatial resolution of the SCT is $(r-\phi) \times z = 17\,\mu\mathrm{m} \times 580\,\mu\mathrm{m}$ in the barrel layers[5].

[5]The endcaps are a bit more complicated. For more information, see Refs [312, 313].

Transition Radiation Tracker (TRT)

Highlighted in orange in Figure 5.8 is the outer-most layer ($554 < r < 1082\,\text{mm}$) of the ID. The Transition Radiation Tracker (TRT) [296, 297] consists of stacks of 4 mm diameter, 144 cm (37 cm) long polyamide tubes in the barrel (endcap) filled with a mixture of gases: 70% Xe, 27% CO_2, and 3% O_2. At the center of the tube is a 31 µm-diameter tungsten wire with a thin gold plating (0.5–0.7 µm). When a charged particle traverses the straw tube (kept at $-1.5\,\text{kV}$), it ionizes the gas and the released electrons are collected on the anode wire at ground. The TRT provides a spatial position resolution of 130 µm in ϕ but does not give a precision measurement along the z-direction, the direction the tubes are aligned with. An important feature of the TRT is its particle identification capabilities. When an ultra-relativistic charged particle crosses a boundary between two media with different dielectric constants [314], their probability to emit transition radiation [315, 316] depends upon their Lorentz γ-factor, meaning that for a fixed momentum, a light charged particle is more likely to emit transition radiation photons than a heavier mass charged particle [317]. By exploiting this effect, the TRT can distinguish between electrons and pions [318].

5.6.2 Calorimetry

As seen in Figure 5.9, the Liquid-Argon (LAr) and Tile Calorimeter systems are located outside the ID. The sampling calorimeters are designed to provide an energy and spatial measurement by absorbing [319] a particles' energy through interactions which produce electromagnetic (electrons and photons) and hadronic (pions and kaons) showers that are measured as they travel through alternating layers of active

and passive material. These incoming particles create a cascade of progressively lower-energy particles by interacting with the dense passive material, creating a particle shower which continues until the energy is depleted. The active layers collect the energy via ionization in the LAr, or Electromagnetic Calorimeter (ECal), and via scintillation in the Tile, or Hadronic Calorimeter (HCal)[6]. There are five calorimeter systems, whose details are described in Table 5.2, some of which are used to capture the showers of both strongly and electromagnetically interacting particles. The dense materials used to absorb the energy of particles cause only a fraction of the energy to be measured by the sensors. This non-compensated loss requires calibration of the measured energy by studying the calorimeter response, which is discussed in more detail in Section 6.5.3.

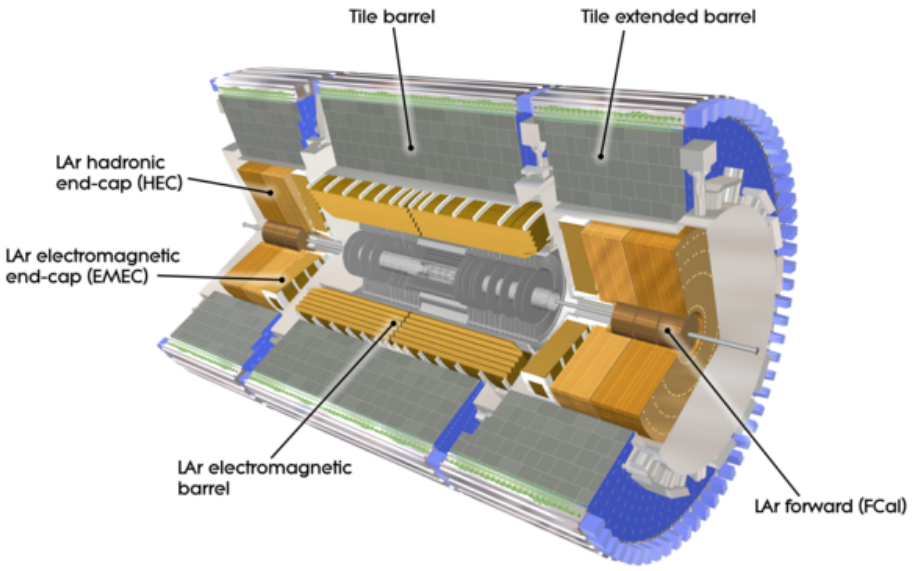


Figure 5.9: Cut-away diagram of the ATLAS calorimeter system [266].

[6]Ecal/LAr, as well as Hcal/Tile, are used somewhat interchangeably throughout the text. The LAr systems are those that use LAr as the active material, which captures mostly electromagnetic showers. The Tile Cal uses scintillator as the active material, and is better for capturing hadronic showers.

Table 5.2: Summary table of the calorimeter systems, the active/absorber material, the coverage in $|\eta|$, and the number of readout channels in each [298–300, 320]. The FCal uses a tungsten-copper combination to capture both shower types. The overlap in the Tile is due to two systems, the barrel and extended barrel.

Detector	Shower	Active	Absorber	Coverage	Channels
EMB	EM	LAr	Lead	$\|\eta\| < 1.475$	99,712
EMEC	EM	LAr	Lead	$1.375 < \|\eta\| < 3.2$	62,208
Tile	Had	Scintillator	Steel	$\|\eta\| < 1.0$	5760
				$0.8 < \|\eta\| < 1.7$	4092
HEC	Had	LAr	Copper	$1.5 < \|\eta\| < 3.2$	5632
FCal	Had	LAr	Tungsten	$3.2 < \|\eta\| < 4.9$	754
	EM	LAr	Copper		1008
Total					179,166

Electromagnetic Calorimeter

The ECal [298, 299] is located just outside the TRT ($1.5 < r < 2.0\,\text{m}$). The relative resolution of the ECal improves as the energy deposited increases and is the quadratic sum of three terms: the stochastic (a), noise (b), and constant (c) coefficients [321, 322]:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \tag{5.8}$$

$$= \frac{10\%}{\sqrt{E}} \oplus \frac{170\,\text{MeV}}{E} \oplus 0.7\%,$$

$$\approx \frac{10\%}{\sqrt{E}}. \tag{5.9}$$

The stochastic term is the result of the proportionality between the calorimeter response and the number of track segments present in the shower and is typically dominant at low energies. The noise term accounts for the electronic noise in the readout electronics. The constant term includes effects that are not energy dependent, including effects from the calorimeter geometry and the evolution of the par-

ticle shower. The sampling calorimeters are made of alternating passive layers of lead and active layers of Liquid-Argon (LAr), arranged in an accordion geometry as seen in Figure 5.10, which provides a full azimuthal coverage with no gaps. Using LAr is useful because of its radiation hardness [323], and the lead makes an excellent material for absorbing the energy of electromagnetically interacting particles [324]. Lead has a high Z (82 protons), which gives it a relatively short radiation length ($X_0 \approx 5.6\,\text{mm}$ [325]), or the characteristic length describing the energy decay of a beam of particles. The ECal is divided into two main parts – the EM Endcap Calorimeter (EMEC) and the EM Barrel Calorimeter (EMB). A third part forms the first section of the Forward Calorimeter (FCal1). The EMB, shown in Figure 5.10, consists of three segmented layers as we move out radially from the center of the detector; the first of which is formed of strip cells, finely segmented in η, with a resolution of $\Delta\eta \times \Delta\phi = 0.0031 \times 0.0245$, allowing for precise position measurements to aid in discrimination between single γ and $\pi^0 \to \gamma\gamma$ events [321]. The second layer is composed of square $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ cells, and contains most of the energy from the electromagnetic shower. The third consists of cells with a resolution of $\Delta\eta \times \Delta\phi = 0.05 \times 0.0245$, catching the remainder of the shower. Not shown in the figure is an additional thin layer, which comes before the three mentioned above. This layer is a pre-sampler, which covers $|\eta| < 1.8$, contains no absorber, and aids in energy corrections due to the non-compensated losses in the layers beyond.

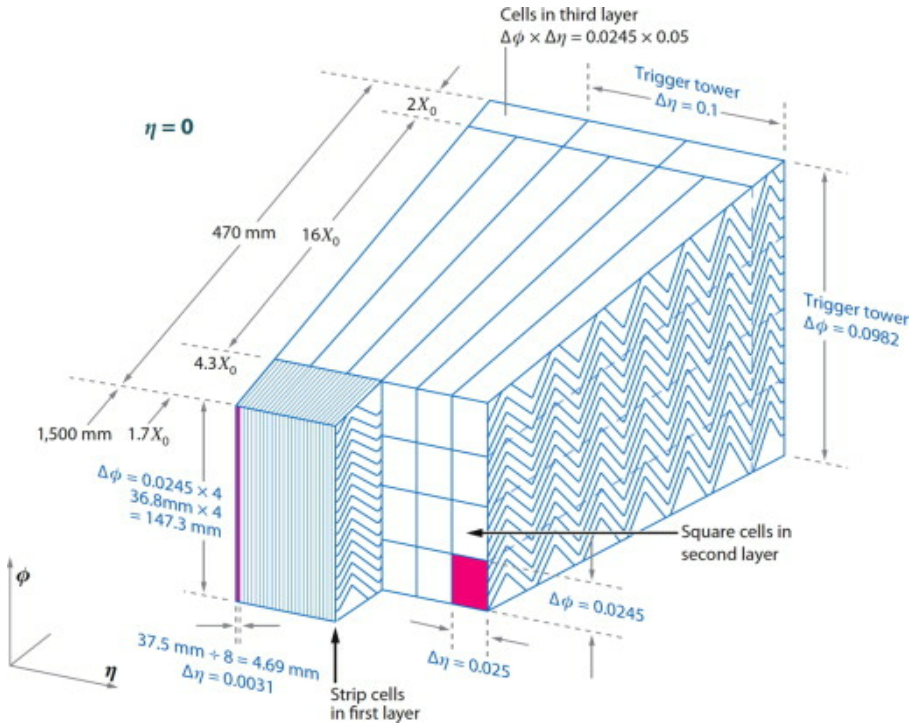


Figure 5.10: Sketch of a section of the EM Barrel Calorimeter, showing the three layers in their accordion geometry [326].

Hadronic Calorimeter

Beyond the ECal ($2.25 < r < 4.25\,\text{m}$) is the Hadronic Calorimeter (HCal) [298, 300], which captures the energy of particles that are either not fully captured by the ECal or are neutral and pass through unobstructed. Similar to the ECal, the HCal's relative resolution improves with increasing energy [321, 322]:

$$\frac{\sigma_E}{E} = \frac{52.9\%}{\sqrt{E}} \oplus 5.7\%. \tag{5.10}$$

The electronic noise is negligible in the HCal [322]. The HCal consists of three systems. In the central region is the TileCal, which is further divided into two regions. Most central is the barrel ($|\eta| < 1.0$) and beyond that is the extended-barrel ($0.8 < |\eta| < 1.7$). These regions are divided azimuthally into 64 modules split into 3 layers. The next system is the LAr Hadronic Endcap (HEC) Calorimeter. Finally, the

Forward Calorimeters, FCal2 and FCal3 cover the most forward regions. The TileCal uses steel as its absorber, and polystyrene scintillating tiles as the active medium and is depicted in Figure 5.11. The HEC uses copper for its absorber and LAr as the active material and the FCal2 and FCal3 use copper and tungsten for their absorbers, respectively. At the edge of the modules of the calorimeter, wavelength-shifting fibers connect the photomultiplier tubes (PMT) to the tiles. This allows the scintillator wavelength to be matched to the PMT sensitivity.

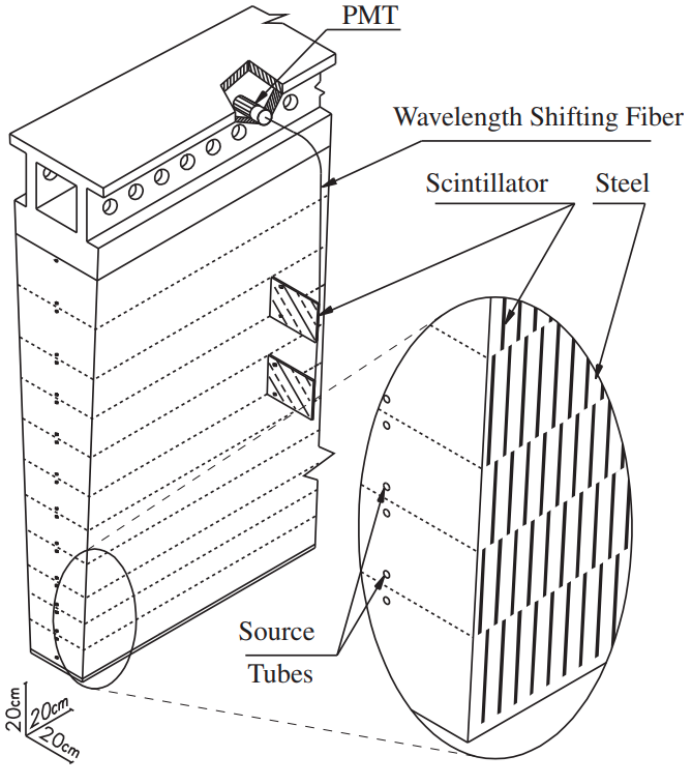


Figure 5.11: Mechanical structure of a TileCal module, showing the slots in the iron for scintillating tiles and the method of light collection by WLS fibers to PMTs. The holes for radioactive calibration source tubes that traverse the module parallel to the colliding beams are also shown [327].

5.6.3 Muon Spectrometer

The Muon Spectrometer (MS) [301], shown in Figure 5.12, is immersed in magnetic field like the ID and is the outermost portion of the ATLAS detector subsystems. The air-core barrel and endcap toroid magnets [328, 329] bend the muons in the longitudinal direction with a $\mathcal{B} \cdot dl$ between 2–6 T · m [330]. The MS is designed to provide precision momentum measurements of muons up to the highest expected energies. The MS consists of four subdetectors specially designed to measure muons, which do not get absorbed by the calorimeters. The four different detectors consist of the Monitored Drift Tube Chamber (MDT), for precision tracking in the MS bending plane with a single-tube resolution of 80 µm [331]; Resistive Plate Chambers (RPC) and Thin Gap Chambers (TGC) for triggering in the barrel and endcap, respectively; and Cathode Strip Chambers (CSC) for precision measurements in the endcap inner layer, which has occupancy levels too high for the MDTs to handle alone [330]. Around the barrel region, three concentric cylindrical shells composed of MDTs and RPCs are arranged at 5 m, 7.5 m, and 10 m. These large distances from the beamline, along with bending power of the magnets, allow for excellent momentum and position resolution for muons, which are often used to trigger on interesting events (more on triggering in Section 6.1). The momentum resolution of the detector scales with the momentum, maintaining $\Delta p_{\mathrm{T}}/p_{\mathrm{T}} = 10\%$, even up to $p_{\mathrm{T}} = 1\,\mathrm{TeV}$ [328, 332].

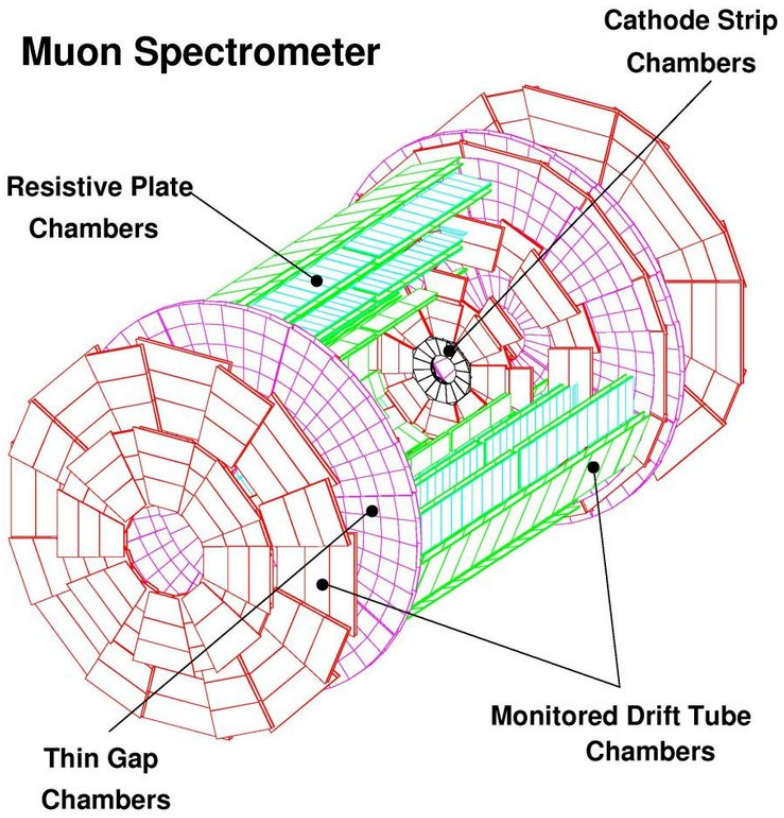


Figure 5.12: ATLAS Muon Spectrometer [333].

Chapter 6

Data and Reconstruction

The collection of hits in the various detector subsystems discussed in Section 5.6, taken in 25 ns intervals, comprise a digital snapshot of each event. The data need to be quickly and efficiently sifted in order to capture interesting events at a rate within the limits of our ability to store it. In the following section, we will discuss how we cope with such a high data flux through the use of hardware- and software-based triggering systems.

6.1 Data Acquisition and Triggering

The raw data from a typical event is approximately 1.6 MB [334] in size. Considering an event rate of 40 MHz, this would require upwards of 65 TB s^{-1} to be written to disk. This is well outside our current capabilities and budget[1]. Additionally, we would be writing many[2] uninteresting inelastic collisions to disk, requiring considerable offline processing. To accomplish the goal of reducing the data inflow, a

[1] Especially because we can't sell an event's personal info to advertisers.

[2] Recall that the total Higgs boson cross section is approximately 55 pb [54] and the total inelastic cross section is approximately 80 mb [286–288]. So the ratio of uninteresting to interesting events (those containing Higgs bosons in our case) is nearly about a billion to one.

dedicated trigger system is employed to preferentially select promising events while reducing the initial rate from 40 MHz to a more manageable 1 kHz. ATLAS has achieved this reduction rate via a two-level trigger system [335, 336]; the first is hardware-based, and the second software-based.

6.1.1 Level-1 Trigger

The first stage of the ATLAS triggering system is the Level-1 (L1) trigger [337, 338]. Operating at the LHC's bunch-crossing frequency, this hardware-level trigger uses dedicated trigger hardware in the calorimeter (L1Calo) and muon (L1Muon) detector systems to reduce the data rate from 40 MHz to at most 100 kHz. The decision [339–341] to keep an event is based on factors such as the multiplicities and energy thresholds from objects such as leptons, jets, and photons, and even takes into account missing transverse momentum. This trigger has a total latency of approximately 2.5 µs so that it reaches a decision before the L1 pipelines of the detector readout systems overflow [342].

6.1.2 High Level Trigger

The second stage of the trigger system is the software-based High-Level Trigger (HLT) [335, 338, 343], which further reduces the data rate to approximately 1 kHz. The HLT uses offline reconstruction algorithms that utilize the full detector output in the regions of interest to make a decision as to whether or not to keep an event within a 200 ms window. The HLT revises the decisions [344] made by the L1 trigger to make more sophisticated selections on muons, electrons, photons, jets, b-jets, missing transverse energy, tau leptons, and b-hadrons.

6.1.3 Trigger Menu

The decisions from the L1 and HLT can be chained together in different combinations in what is known as the trigger menu [345–348], which allows for high acceptance of physics events for various analyses studying the SM, BSM, SUSY, top-physics, etc., and for dedicated detector calibration and monitoring. Due to changing LHC conditions and continuous improvements in the trigger algorithms, the menu changes from year to year. A full list of the trigger menu is outside the scope of this dissertation, but we can outline the basics of the trigger menu strategy [348]:

- Primary triggers: used for physics measurements, usually unprescaled.
- Support triggers: used for efficiency and performance measurements, background estimates or monitoring, and typically run at a small rate (approximately 0.5 Hz each with prescaling). Roughly 15% of the HLT bandwidth is dedicated to support triggers.
- Alternative triggers: used for cases where the CPU usage or output rate of the primary trigger becomes too high. These triggers employ tighter selections and lower processing rates than the primary triggers.
- Calibration triggers: used for detector calibrations, which run at high rates compensated for by storing only the necessary detector information for each event to reduce the event size and bandwidth.

On top of this trigger menu, and of importance for correct luminosity calculations, prescale factors are added for the L1 and HLT triggers. These prescale factors reduce the number of events accepted due to thresholds that are low enough to pro-

duce a data rate which is still too high. For a prescale factor of N, only one out of N events which fulfill the trigger requirements are accepted.

6.1.4 Jet and Muon Triggers

Since the Boosted $H \to b\bar{b}$ Analysis relies on the use of large-radius (large-R) jets, we utilize several jet triggers which have been implemented over the data-taking years of Run 2. The ATLAS trigger system includes several chains aimed at the presence of a large-R jet. The performance of the unprescaled large-R jet triggers is summarized in Table B.4 of Appendix B.4. The large-R jet-triggers can trigger on either the p_{T} or both the p_{T} and mass of the large-R jet, which is reconstructed in the High Level Trigger using the anti-k_t algorithm [349, 350]. Putting a requirement on the minimum mass of a large-R jet reduces the rate by rejecting events coming from QCD jets, allowing for a lower p_{T} threshold. The jet triggers come in three flavors, based on the type of large-R jet used:

- **a10**: ungroomed anti-k_t $R = 1.0$ jets,
- **a10r**: anti-k_t $R = 0.4$ jets reconstructed with anti-k_t $R = 1.0$,
- **a10t**: trimmed[3] anti-k_t $R = 1.0$ jets with trimming parameters $\rho = 0.2$, $f_{\mathrm{cut}} = 0.05$.

The efficiency of the triggers was studied (Appendix B.4) for the analysis in data and Pythia 8 QCD Monte Carlo using the leading large-R jet in the event assuming the trigger efficiency is a property of kinematics only. The least restrictive (in p_{T} for both p_{T} and p_{T}-mass triggers) prescaled single large-R-jet triggers were used

[3]Trimming is defined in Section 6.5.3

to select the events used in the analysis. The specific triggers used in the Boosted $H \to b\bar{b}$ Analysis can be found in Section 7.2. In order to retain the majority of the gains provided by the new triggers, while also keeping the number of categories used as input to the final fits manageable, we use offline trigger cuts unified as $p_{\mathrm{T}} > 450\,\mathrm{GeV}$ and $m > 60\,\mathrm{GeV}$.

The Boosted $H \to b\bar{b}$ Analysis also relies on two control regions which trigger on high- and low-p_{T} muons, with appropriate offline cuts, to select semileptonic $t\bar{t}$ events. For muon triggering [351], a sequence of reconstruction and selection steps are employed for specific muon objects in the L1 and HLT trigger chains [352]. The level-1 muon triggers can trigger on $1 - 4$ muons and typically follow a `L1_MUX` naming convention[4], where `X` refers to the minimum p_{T} of the muon in GeV. The HLT triggers use additional information such as the muon p_{T}, muon isolation (see Table B.6 of Appendix B.4), and the muon type (which is related to how the muon was reconstructed). More will be discussed on muons in Section 6.3 and the muon specific triggers for the control regions will be discussed in Sections 8.3 and 9.5. The single-muon and multi-muon trigger chains can be found in Table B.5 in Appendix B.4

6.1.5 Data Storage and Processing

After the data rate is brought down to a more manageable 1 $\mathrm{GB\,s^{-1}}$, the bytestream files containing the raw data from the readout electronics are decoded and processed to produce the reconstructed physics objects for use in analysis. Raw files are converted to Event Summary Data (ESD) – containing the full out-

[4]The `L1_MU11` actually applies a 10 GeV threshold, but contrary to `L1_MU10`, a 3-station coincidence is required for the RPCs [352].

put of reconstructed objects (tracks and their hits, calorimeter clusters, etc.), and Analysis Object Data (AOD) – with a summary of the event reconstruction into POOL/ROOT files [353]. These files are further customized into smaller data files to meet the needs of various analyses and performance groups into a Derived AOD (DAOD or xAOD) and undergo several types of filtering to reduce the size of the data files to be analyzed:

- Skimming: removes whole events based on selection criteria
- Thinning: removes objects within events based on selection criteria
- Slimming: removes object variables across all events

Event reconstruction occurs primarily at CERN's Tier-0 computing facility [354], where one copy of the raw data is archived. Replicas of the raw data are used to produce the xAOD data, distributed to several Tier-1 computing farms. Each Tier-1 facility serves multiple Tier-2 facilities, typically hosted by participating universities or laboratories. Many smaller institutions maintain Tier-3 computing facilities, which aid in the physics analysis. In the following sections, we will discuss the types of reconstructed physics objects used in many analyses, including the Boosted $H \to b\bar{b}$ Analysis.

6.2 Tracks and Vertices

One of the most vital steps in all analyses is the identification of b-jets, which requires effective track and vertex reconstruction. Using the collection of hits from the Inner Detector allows us to precisely determine the trajectories of particles, which is necessary for an accurate determination of the momentum and mass/energy

for the particles in a given event. Accurate measurements of the vertex and track parameters aid in combating the effects of pileup and provide us precise information on the topology of jet structure.

6.2.1 Track Reconstruction

Track reconstruction relies on fitting a trajectory to the charged particles which pass through the Inner Detector. This process begins by forming track seeds, which consist of three measurements in the different pixel and SCT layers. From the seed, the track trajectory is propagated using an *inside-out* method from the pixel and SCT layers to the TRT using a combinatorial Kalman filter [355]. As the algorithm moves out, additional silicon hits are added to the seed. To remove track candidates which receive incorrectly assigned hits, an ambiguity solving step is employed to remove the candidates. Each candidate track is scored relative to others, with the score increasing when an additional measurement is added [356]. The measurements coming from the three components of the Inner Detector are weighted differently, favoring the higher precision pixel detector and suppressing the weights from lower precision detectors such as the TRT. In many instances, the track trajectory predicts a hit in the detector where none is found; this is referred to as a hole. These holes help provide a more realistic description of the detector acceptance and efficiency [357], and allow the algorithm to penalize the track candidate weight and reduce its overall score. To promote tracks coming from hard scattering processes and reduce the softer interactions from pileup, the logarithm of the transverse momentum $\ln(p_{\mathrm{T}})$ is used to enhance the scores of energetic tracks and suppress the more numerous tracks formed from incorrect combinations of clusters, which tend to have a low transverse momentum. Figures 6.1 and 6.2 show the resolution for the transverse

and longitudinal impact parameters as a function of the transverse momentum. Excellent impact parameter resolution is necessary for distinguishing tracks coming from the primary vertex or from b- or c-quark decays.

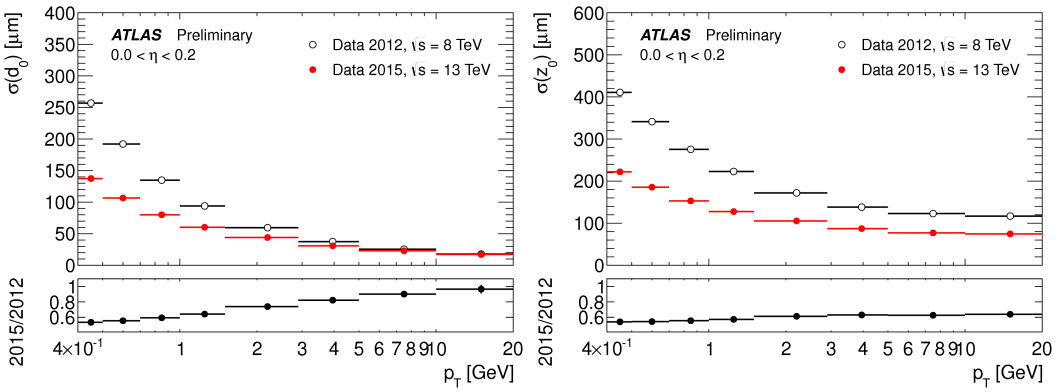


Figure 6.1: The unfolded transverse (left) and longitudinal (right) impact parameter resolutions measured from data in 2015, $\sqrt{s} = 13\,\text{TeV}$, with the Inner Detector including IBL as a function of p_{T}, for values $0.0 < \eta < 0.2$ compared to that measured from data in 2012, $\sqrt{s} = 8\,\text{TeV}$. The data in 2015 is collected with a minimum bias trigger. The data in 2012 is derived from a mixture of jet, tau, and missing E_{T} triggers [358].

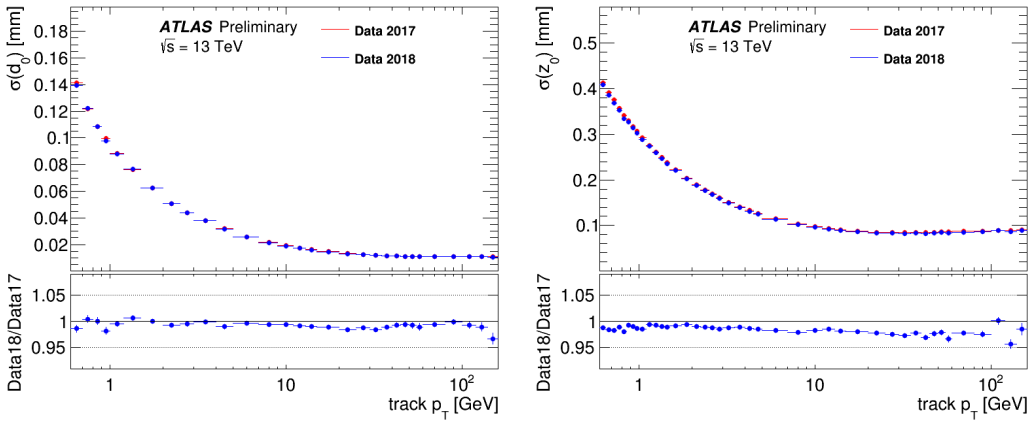


Figure 6.2: The intrinsic transverse (left) and longitudinal (right) impact resolutions for *TightPrimary* tracks associated to jets with $p_{\text{T}} > 20\,\text{GeV}$ measured in dijet-triggered pp collision data collected in 2017 (red) and 2018 (blue). The transverse impact parameter resolution is extracted in bins of p_{T} of tracks using iterative Gaussian fits [359].

6.2.2 Vertex Reconstruction

After the tracks have been reconstructed, the primary vertex candidates are constructed, which first involves finding the vertices and then fitting. Finding the vertices is a pattern recognition process where reconstructed tracks are associated to vertex candidates and the actual position and covariance matrix are produced by a fit [357]. To perform this, a set of tracks satisfying the track selection criteria are defined and the seed position for the first vertex is selected. This seed position of the vertex fit receives its x- and y-coordinates from the center of the beam spot. The z-coordinate is calculated from the mode of the z-coordinates of tracks given their respective points of closest approach to the center of the beam spot using the Half-Sample Mode algorithm [360]. The seeds are then subjected to an iterative primary-vertex finding procedure where the positions are determined using an adaptive vertex-fitting algorithm with an annealing procedure [357, 361]. Starting with a seed position, the parameters of the reconstructed tracks are used as inputs and the algorithm performs an iterative χ^2 minimization to optimize the vertex position. Each track is assigned a weight to reflect its compatibility with the vertex estimate and the position is iteratively recalculated using the weighted tracks. These track weights are calculated from the χ^2 values as:

$$\omega(\hat{\chi}^2) = \frac{1}{1 + \exp(\frac{\hat{\chi}^2 - \chi^2_{\text{cutoff}}}{2T})}, \tag{6.1}$$

where $\hat{\chi}^2$ is calculated from the position between the last estimated vertex and the point of closest approach of the track. Tracks with a lower weight have less influence on the position to reflect their incompatibility with the vertex position. The χ^2_{cutoff}, the threshold where the weight is equal to 0.5, is set to 9, corresponding to 3

standard deviations. T is the temperature parameter that controls the smoothness of the weighting procedure. When T is large, the function is fairly flat in $\hat{\chi}^2$; when small, it approaches a step function. In order to converge on a local minima, T is progressively reduced with each iteration in a predefined sequence until $T = 1$ [357]. Lastly, the final weight for each track is used to evaluate the vertex fit. Tracks incompatible with the vertex by more than 7 standard deviations are removed from the vertex candidate and returned to a pool of unused tracks for consideration in other primary vertex candidates. This process is repeated until all unused tracks have been exhausted or no vertex can be found with at least two associated tracks. From here, the primary vertex of the hard scatter is ready to be identified. The primary vertex candidate with the highest sum of the squared transverse momenta of all contributing tracks, $\sum p_{\mathrm{T}}^2$, is chosen as the primary vertex. Figure 6.3 shows the average number of reconstructed vertices as a function of the average number of interactions per bunch crossing.

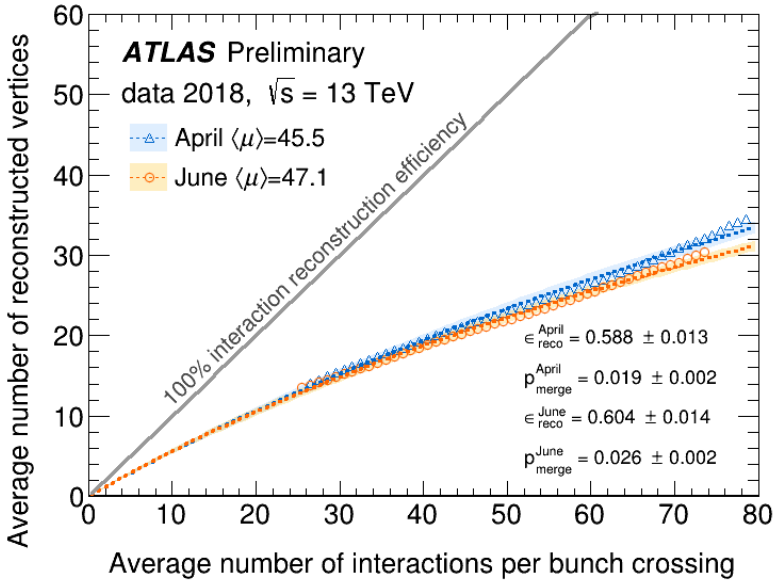


Figure 6.3: The number of vertices reconstructed as a function of the average number of interactions per bunch crossing, in two fills with different average μ taken at different times in 2018. The distributions are fit with an analytical function introduced in [357] that allows the extraction of the vertex reconstruction efficiency and vertex merging probabilities. The error bars on the data show the statistical uncertainties. The shaded bands show the 1 sigma variation of the fit parameters. The data slightly overshoot the parameterization for large μ values. A possible source for this discrepancy is the contribution of fake vertices due to combinatorics, which are neglected in the fitted parameterization [362].

6.3 Muon Reconstruction

Among all reconstructed objects, muons are some of the most straightforward, with high efficiency and low fake rates. Muons are long-lived, with a proper lifetime of 2 µs [363] ($\langle c\tau \rangle \approx 600m$), so they make it well outside of the detector. Muons are also the only charged particles that produce neither electromagnetic nor hadronic showers, so they pass beyond the calorimeters, potentially giving us measurements in both the ID and MS.

A muon first passes through the ID, bending in the transverse plane due to the 2 T solenoidal magnetic field present in that region. The hits in this portion of the ATLAS detector provide precise hits close to the IP, yielding tight constraints on

the transverse and longitudinal impact parameters, d_0 and z_0, respectively. After traversing the calorimeters with little energy loss, the muon passes into the MS and is bent along the polar direction θ by the toroidal magnets, providing excellent resolution for the charge-to-momentum ratio q/p. The various regions through which muons pass in the detector allow for four momentum reconstruction algorithms [364]:

- Combined (CB): obtained from a global re-fit of the ID and MS subdetector hits, where MS hits may be added or removed from the track to maximize the fit quality.
- Segment-tagged (ST): consist of fitted ID tracks and MS segments (MDT or CSC). This is used mostly with lower p_{T} muons that do not pass through the entire MS.
- Calorimeter-tagged (CT): consist of a fitted ID track along with an energy deposit in the calorimeters.
- Extrapolated (ME): these muons consist of an MS track only, containing at least two layers, ensuring the track is consistent with originating from the IP.

A representation of these reconstruction algorithms can be found in Figure 6.4.

Collections of muon selection criteria, often referred to as Working Points (WPs), are defined to suit the needs of various ATLAS analyses. Five WPs are currently supported: the *Loose* WP maximizes the reconstruction efficiency, taking advantage of all types of muons defined above; the *Medium* WP, an ATLAS standard, uses only CB and ME muons; the *Tight* WP maximizes selection purity through the use of CB and ME muons by using tighter restrictions on the muon isolation and number of hits required in the fit; the *Low-*p_T WP maximizes selection efficiency and

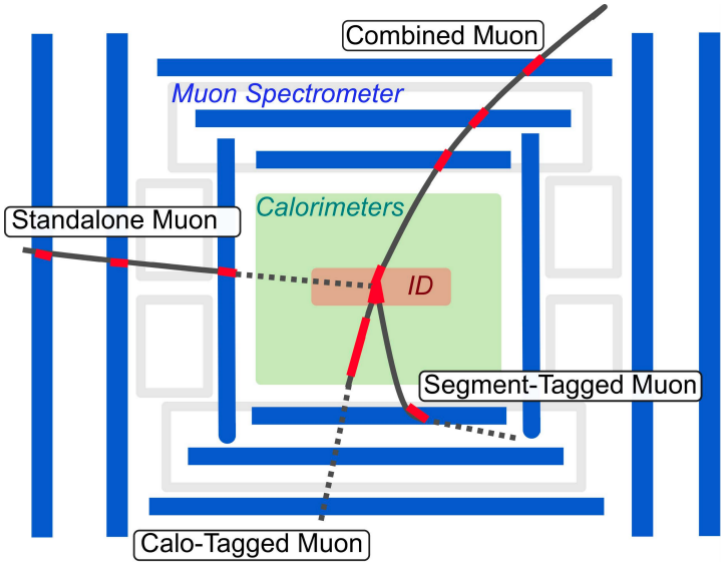


Figure 6.4: Reconstruction algorithms available in the ATLAS detector [364].

Table 6.1: Selection efficiencies for several muon selection WPs [365].

	$4 < p_{\mathrm{T}} < 20\,\mathrm{GeV}$		$20 < p_{\mathrm{T}} < 100\,\mathrm{GeV}$	
Selection	$\epsilon_{\mu}^{\mathrm{MC}}$ [%]	$\epsilon_{\mathrm{Hadrons}}^{\mathrm{MC}}$ [%]	$\epsilon_{\mu}^{\mathrm{MC}}$ [%]	$\epsilon_{\mathrm{Hadrons}}^{\mathrm{MC}}$ [%]
Loose	96.7	0.53	98.1	0.76
Medium	95.5	0.38	96.1	0.1
Tight	89.9	0.19	91.8	0.11
High-p_{T}	78.1	0.26	80.4	0.13

fake rejection for muons with $p_{\mathrm{T}} < 5\,\mathrm{GeV}$; and the *High-$p_T$* WP maximizes momentum resolution for $p_{\mathrm{T}} > 100\,\mathrm{GeV}$ muons. Table 6.1 shows the selection efficiencies for several muon WPs.

To reduce background contributions from light and heavy hadron decays from within jets, muon isolation is used. This isolation is calculated by measuring the detector activity in the vicinity of the muon candidate, using track- and calorimeter-based variables to quantify the level of isolation [365].

6.4 Flavor Tagging

In Section 6.4, we discussed the importance of identifying jets originating from b-hadrons. As mentioned previously, the properties of b-hadrons such as their relatively long lifetimes, high mass, and high decay multiplicities aid in their identification. To take advantage of these properties, b-tagging algorithms heavily rely on reconstructed tracks to identify the secondary vertex (SV) from the b-decay as well as the tertiary vertex from the subsequent decay of hadrons formed from c-quarks as seen in Figure 6.5. To get a flavor for how we tag jets, we will discuss several of the basic b-tagging algorithms used in ATLAS in the following sections. These algorithms can be quite simple, based on the impact parameters and secondary vertex, to more complicated algorithms which take the simpler ones in as input, exploiting a host of properties. Typically, jets are tagged with the light, c, or b flavor label.

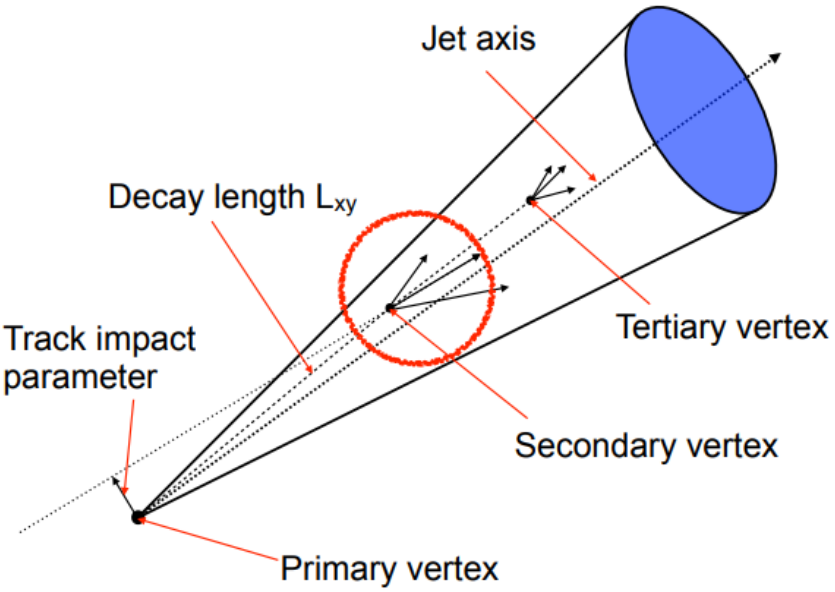


Figure 6.5: Diagram showing the decay of a b-hadron with a secondary vertex (circled) displaced from the primary vertex by distance L_{xy}, a tertiary vertex formed from the subsequent c-hadron decay, and a non-zero impact parameter [366].

6.4.1 Physics Taggers

Impact Parameter Tagging Algorithm

Impact-parameter based algorithms (IP2D and IP3D) [246, 367] exploit the long lifetimes of hadrons containing b-quarks. This long lifetime results in a topology characterized by at least one vertex displaced from the point where the hard-scatter collision occurred. The tracks generated from the b-hadrons tend to have large impact parameters (IPs)[5], enabling us to separate those tracks from the contribution of tracks from the primary vertex (PV). We typically define two impact parameters: the transverse and longitudinal IPs. The transverse impact parameter d_0 is defined as the distance of closest approach in the $r - \phi$ plane of the track to the PV, and

[5]IP was just used for interaction point. For the remainder of this section, IP is the impact parameter and interaction point will be fully written out or replaced with primary vertex (PV).

the longitudinal impact parameter $z_0 \sin\theta$ is defined as the distance of the track to the PV in the longitudinal plane at the point of closest approach in $r - \phi$. A sign, relative to the jet direction, is given to the impact parameter to designate whether the primary vertex is in front of the secondary vertex or behind it. An SV behind the PV typically comes from background and aids in calibrating the light-quark tag rate [368]. The track selection applied for the IP-algorithm is based on the following requirements:

- track p_{T} above 1 GeV
- $|d_0| < 1\,\mathrm{mm}$ and $|z_0 \sin\theta| < 1.5\,\mathrm{mm}$
- seven or more silicon hits, with at most two silicon holes, at most one of which is in the pixel detector.

The IP2D tagger utilizes the transverse impact parameter significance (d_0/σ_{d_0}) as its discriminating variable and IP3D uses both the transverse and longitudinal impact parameter significance ($z_0 \sin\theta/\sigma_{z_0 \sin\theta}$), accounting for their correlation. Monte Carlo (MC) simulation is used to construct probability density functions (PDFs) for the IP significances, which are divided into several categories depending on the characteristics of the hits in the silicon layers [246]. A log-likelihood ratio (LLR) discriminant is constructed as the sum of the per-track contributions:

$$\sum_{i=1}^{N} \log\left(\frac{p_b}{p_u}\right), \tag{6.2}$$

where N is the number of tracks for a given jet and p_b/p_u are the template PDFs for the b-/light-quark-flavor hypotheses, respectively. The distributions for the transverse and longitudinal IP significances are shown in Figure 6.6 for tracks associated

to light-, c-, and b-jets.

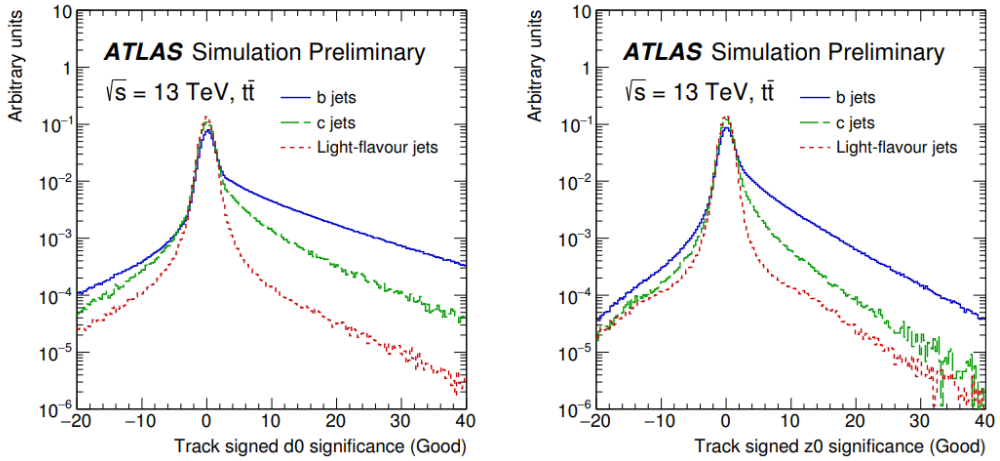


Figure 6.6: The transverse (left) and longitudinal (right) signed impact-parameter significances of tracks in simulated $t\bar{t}$ events for b- (solid blue), c- (dashed green), and light-flavor (dotted red) jets for the *Good* category [246]. See Reference [246] for details on the *Good* category description.

Secondary Vertex Finding Algorithm

The secondary vertex finding algorithm [369] reconstructs the secondary vertex within a jet, where all track pairs associated to the jet are tested for a two-track vertex hypothesis. If a two-track vertex is likely to have originated from the decay of a long-lived particle, photon conversion, or hadronic interaction with detector material, it is rejected. If the vertex is not rejected, it is then fitted with all tracks from the accepted two-track vertices, and outliers are iteratively removed from the set. Each selected track is required to have at least seven silicon hits from the pixel and SCT, with at most one shared hit in the two detectors. Because highly energetic jets tend to have a higher multiplicity of tracks as the result of increased jet fragmentation, there is an increased risk of reconstructing fake vertices. To reduce the fake rate, tracks are p_{T}-ordered and at most 25 of the highest-p_{T} tracks can be used to reconstruct the secondary vertex. In the forward regions ($|\eta| > 1.5$), an

increase in detector material leads to more interactions and a poorer resolution on the impact parameter. To improve the resolution, track cleaning is applied and the minimal number of silicon hits is increased by one to improve the quality of selected tracks. To reduce the fake rate from pileup, tracks with a low transverse IP significance (< 2) and high longitudinal IP significance (> 6) are removed. Further details regarding track selection can be found in Reference [370].

JetFitter

JetFitter [371] is a multi-vertex decay-chain reconstruction algorithm that exploits the topological structure of b- and c-decays. By reconstructing the full PV$\rightarrow$ b $\rightarrow$ c-hadron decay chain, a Kalman filter [355] algorithm finds the common line connecting the positions of the primary vertex with the bottom and charm vertices, approximating the b-hadron flight path. Provided we have a good vertex resolution, the b- and c-hadron vertices can be resolved even when only a single track is attached to each of the them [372]. An example of the generated and reconstructed track and vertex topologies for a simulated B^0 meson decay is shown in Figure 6.7.

The JetFitter algorithm makes the assumption that the b- and c-hadron decay vertices are aligned on the line defined by the b-hadron flight path, which will intersect this axis. This permits us to reconstruct incomplete topologies, including topologies with a single track from both the b- and subsequent c-hadron decays. The compatibility of the set of tracks undergoing a $b \rightarrow c$-like cascade topology is evaluated using a fit, increasing the discriminating power against c- and light-flavored jets. By constraining the tracks to lie on the b-hadron flight axis, the number of degrees of freedom of the fit are reduced, increasing the probability of resolving the b- and c-hadron vertices. Further details on the JetFitter algorithm can be found in

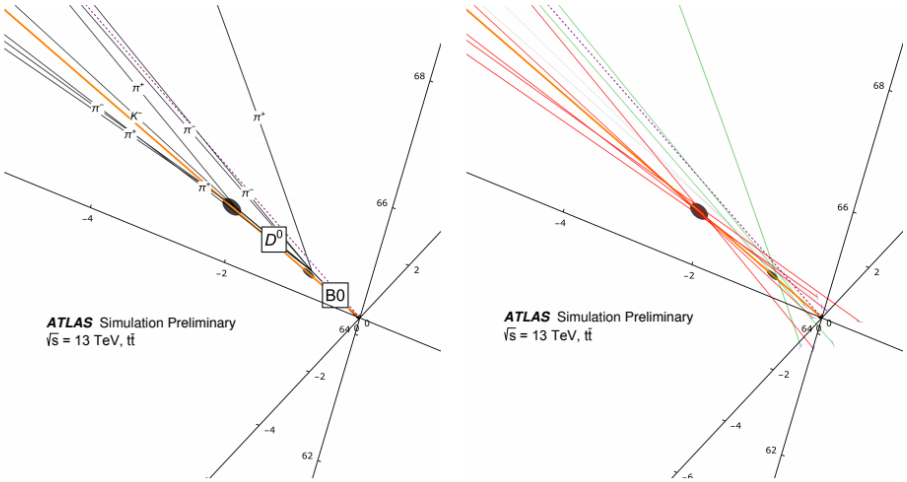


Figure 6.7: Event display of a simulated B^0 decay showing the generated (left) and reconstructed (right) particle trajectories. Only charged particles originating from the B^0 meson decay are shown. On the right, reconstructed particles (green and red for those originating from the secondary and tertiary vertices, respectively) are extrapolated to the point of closest approach to the beam-collision point. The coordinates on the axes are given in mm. JetFitter performs a multi-vertex fit constraining all vertices to lie on a single flight axis, indicated by an orange line in both figures [373].

Reference [373].

6.4.2 MV2 Flavor Tagging Discriminant

The Boosted $H \to b\bar{b}$ Analysis relies on the MV2c10 algorithm for identifying b-jets. The MV2 algorithm [372] is a multivariate classifier that takes the inputs from the three basic b-tagging algorithms discussed above and discriminates b-jets from light- and c-jets using a boosted decision tree (BDT). The BDT algorithm is trained using the ROOT Toolkit for Multivariate Data Analysis (TMVA) [374]. There are three MV2cXX (XX = 00, 10, 20) versions, where the value of XX depends upon the approximate fraction of charm-flavor jets used in the background for training. For example, MV2c10 was trained using a mixture of $\sim 10\%$ c-jets and $\sim 90\%$ light-jets. The training of this multivariate classifier was performed on jets in simulated $t\bar{t}$ events. To take advantage of correlations between the input variables, the kinematical properties are included in the training; however, the signal jets are reweighted in p_{T} and η to match the spectrum of the light-flavor background to avoid differences between the kinematical distributions of signal and background causing a bias in the output. The inputs to the MV2 algorithm can be found in Table 6.2.

Figure 6.8 shows the light-jet and c-jet rejections as a function of the b-jet efficiency for the MV2 b-tagging algorithm. The choice in the Boosted $H \to b\bar{b}$ Analysis of MV2c10 over the others is a result of the need for a higher c-jet rejection, as c-jets are more likely to be mistagged as a b-jet compared to a light-jet. MV2c10 has a higher light-jet rejection than MV2c20 (which is trained on a smaller portion of light-jet background), but a lower light-jet rejection compared to MV2c00 (trained only on a light-jet background). MV2c10 has considerably better c-jet rejection than MV2c00, and is comparable to MV2c20. The various working points (WPs), given as the b-jet efficiency, along with the rejection factors for c-, light-, and τ-jets can

Table 6.2: The input variables used by the MV2 b-tagging algorithms [372].

Input	Variable	Description
Kinematics	p_{T}(jet)	Jet transverse momentum
	η(jet)	Jet pseudorapidity
IP2D, IP3D	$\log\left(P_b/P_{\mathrm{light}}\right)$	Likelihood ratio between the b- and light-jet hypotheses
	$\log\left(P_b/P_c\right)$	Likelihood ratio between the b- and c-jet hypotheses
	$\log\left(P_c/P_{\mathrm{light}}\right)$	Likelihood ratio between the c- and light-jet hypotheses
SV	m(SV)	Invariant mass of tracks at the secondary vertex assuming pion masses
	f_E(SV)	Fraction of the charged jet energy in the secondary vertex
	N_{TrkAtVtx}(SV)	Number of tracks used in the secondary vertex
	N_{2TrkVtx}(SV)	Number of two-track vertex candidates
	L_{xy}(SV)	Transverse distance between the primary and secondary vertices
	L_{xyz}(SV)	Distance between the primary and secondary vertices
	S_{xyz}(SV)	Distance between the primary and secondary vertices divided by its uncertainty
	ΔR(jet, SV)	ΔR between the jet axis and the direction of the secondary vertex relative to the primary vertex
Jet Fitter	N_{2TrkVtx}(JF)	Number of two-track vertex candidates (prior to decay chain fit)
	m(JF)	Invariant mass of tracks at the secondary vertex assuming pion masses
	S_{xyz}(JF)	Significance of the average distance between the primary and secondary vertices
	f_E(JF)	Fraction of the charged jet energy in the secondary vertices
	$N_{1-\mathrm{trk\ vertices}}$(JF)	Number of displaced vertices with one track
	$N_{\geq 2-\mathrm{trk\ vertices}}$(JF)	Number of displaced vertices with more than one track
	N_{TrkAtVtx}(JF)	Number of tracks from displaced vertices with at least two tracks
	$\Delta R(\vec{p}_{\mathrm{jet}}, \vec{p}_{\mathrm{vtx}})$	ΔR between the jet axis and the vectorial sum of the momenta of all tracks attached to displaced vertices

be found in Table 6.3. The rejection is simply one divided by the efficiency for selecting a background jet. So, for example, a rejection factor of 1000 would mean that the background rate is diminished by a factor of 1000 when the tagger is applied. Figure 6.9 shows the b-tagging efficiency and data-to-simulation scale factors for the working point used in the Boosted $H \to b\bar{b}$ Analysis (77%) as a function of the jet p_{T}.

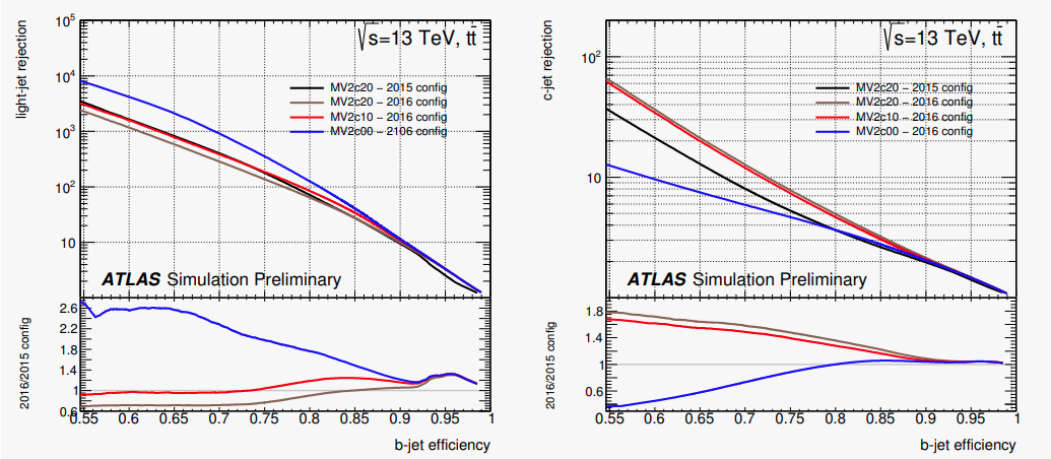


Figure 6.8: Light-flavor jet (left) and c-jet (b) rejection vs. b-jet efficiency for the previous (2015 config) and the current configuration (2016 config) of the MV2 b-tagging algorithm evaluated on $t\bar{t}$ events. MV2c10 (MV2c20) denote the MV2 outputs where 7% (15%) c-jet fractions were present in the background sample (for the 2016 configuration [246])

Table 6.3: Operating points for the MV2c10 b-tagging algorithm, including benchmark numbers for the efficiency and rejection rates. These values have been extracted from $t\bar{t}$ events, the main requirement being jet p_{T} above 20 GeV [246].

b-jet Efficiency [%]	c-jet Rejection	Light-jet Rejection	τ Rejection	BDT Cut Value
60	34	1538	184	0.9349
70	12	381	55	0.8244
77	6	134	22	0.6459
85	3.1	33	8.2	0.1758

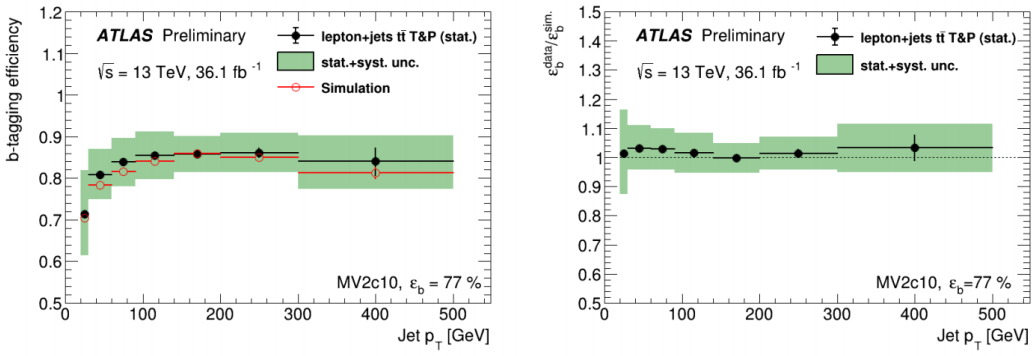


Figure 6.9: (left) The b-tagging efficiencies for the MV2c10 algorithm at the 77% WP as a function of the transverse momentum of the probe jet obtained by selecting $t\bar{t}$ single lepton (SL) events. The b-tagging efficiencies for the predictions extracted from the simulation are shown as a red line while the efficiencies measured in data by using the tag-and-probe (T&P) method are shown as black dots. The vertical error bars represent the statistical uncertainty on the measurement. The green bands indicate the total statistical and systematic uncertainties on the measurement. (right) The corresponding data-to-simulation scale factors are presented as a function of the jet p_{T} [375].

6.5 Jets

At the detector level, the signals in the calorimeter are collected into clusters of topologically connected cell signals, known as topo-clusters, allowing us to extract the significant signal from the background of electronic noise and other fluctuations such as pileup [376]. These clusters represent the full or fractional response to a single particle, the merged response for several particles, or a combination of merged full and partial showers. The observable used to determine cluster formation is the cell-signal significance $\varsigma^{\mathrm{EM}}_{\mathrm{cell}}$, which is the ratio of the cell signal to the average (expected) noise $\sigma^{\mathrm{EM}}_{\mathrm{noise,cell}}$, which is the quadratic sum of the electronic and pileup noise:

$$\varsigma^{\mathrm{EM}}_{\mathrm{cell}} = \frac{E^{\mathrm{EM}}_{\mathrm{cell}}}{\sigma^{\mathrm{EM}}_{\mathrm{noise,cell}}}. \tag{6.3}$$

Topo-clusters are formed by a growing-volume algorithm, starting with a primary seed formed by a calorimeter cell when the cell-signal significance is above a particular threshold. The seeding, growth, and boundaries are controlled by three parameters $\{S, N, P\}$, respectively, which define the signal thresholds in terms of the noise mentioned above.

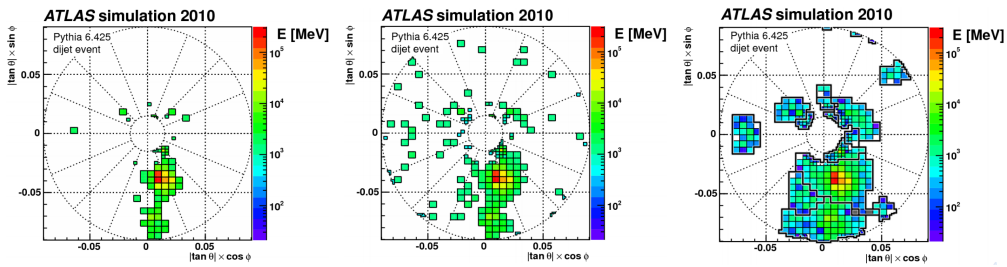
$$|\varsigma^{\mathrm{EM}}_{\mathrm{cell}}| > S \qquad \text{(primary seed threshold, default } S = 4), \tag{6.4}$$

$$|\varsigma^{\mathrm{EM}}_{\mathrm{cell}}| > N \qquad \text{(threshold for growth control, default } N = 2), \tag{6.5}$$

$$|\varsigma^{\mathrm{EM}}_{\mathrm{cell}}| > P \qquad \text{(principal cell filter, default } P = 0). \tag{6.6}$$

The formation of the clusters proceeds with a seed-and-collect process, repeated

until all topologically connected cells passing the criteria for seeding (S), growing (N), and the boundary condition (P) are met. This process can be seen visually in Figure 6.10 [376].



(a) (b) (c)

Figure 6.10: Cells passing primary-seed threshold (a), cells passing threshold for growth (b), and cells passing cell-filter threshold (c) [376].

6.5.1 Jet-Clustering Algorithms

Once the topo-clusters have been identified, they can be grouped into the same physics object – a jet. Depending on the needs of the analysis, several types of jets are formed based on the angular size (R) used in the clustering algorithm. We can have small-radius ($R = 0.2, 0.4$), variable-radius ($R = \rho/p_{\mathrm{T}}$), and large-R ($R = 1.0$) jets. Variable-radius jets allow for a more realistic description for the size of small-R jets, as the opening angle decreases with increasing transverse momentum. The most common algorithm to construct jets is a sequential recombination procedure, which is based on the distance metrics d_{ij}, the distance between jet candidates, and d_{iB}, the distance between the jet candidate and the beam. They are defined:

$$d_{ij} = \min(k_{t,i}^{2p}, k_{t,j}^{2p}) \frac{\Delta_{ij}^2}{R^2}, \tag{6.7}$$

$$d_{iB} = k_{t,i}^{2p}, \tag{6.8}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, and $k_{t,i}$, y_i, and ϕ_i are the transverse momentum, rapidity, and azimuthal angle of particle i, respectively. For each combination i, j in the list, the minimum of d_{ij} and d_{iB} is iteratively computed to merge two candidates into the same jet by summing the four-momenta. In the case that d_{iB} is the minimum, the algorithm stops, completing the jet object and removing it from the list. This process continues until all energy clusters have been grouped into jets.

The parameter p in the exponent of the transverse momentum allows us to tune the relative influence of the deposit's energy. For $p = 0$, we obtain the Cambridge-Aachen (C/A) algorithm [377–379], which is not discussed in this dissertation. When $p = 1$, the algorithm is referred to the k_t-algorithm [350], and its use will be referenced briefly in the case of large-R-jet trimming. In the case where $p = -1$, we have the anti-k_t algorithm [350, 375], which is implemented in the FastJet [380] package and is used ubiquitously in ATLAS. These three algorithms are collinear and infrared safe[6]. The anti-k_t algorithm favors clusterings that involve hard particles rather than clusterings that involve soft particles (k_t algorithm) or energy-independent clusterings (C/A) [236]. The anti-k_t algorithm forms circular boundaries around the hardest deposits, which can be seen in Figure 6.11, resulting in cone-shaped jets, making calibration considerably easier.

[6]Infrared and collinear (IRC) safety is the property that if one modifies an event by a collinear splitting or the addition of a soft emission, the set of hard jets that are found in the event should remain unchanged [236].

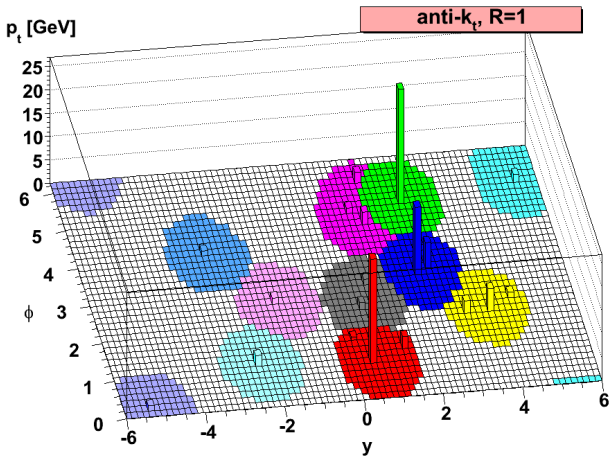


Figure 6.11: Illustration of the regularity of the jets obtained with the anti-k_t algorithm [381].

6.5.2 Variable-Radius Jets

As discussed in Section 4.4, the angle between the decay products of the Higgs boson (b-quarks in our case) decreases with increasing momentum. In the same manner, this is true for the decay products of the bottom quarks. An alternative to the fixed-radius jets (*e.g.*, $R = 0.4$ or $R = 1.0$) is to allow the jet radius to vary, permitting a cone size that more accurately reflects the true size of jets. The variable-radius (VR) jet algorithm modifies the conventional iterative recombination algorithm by allowing the jet radius to vary as a function of the jet p_{T} as [382]:

$$R \to R_{\mathrm{eff}}(p_{\mathrm{T}}) = \frac{\rho}{p_{\mathrm{T}}}. \tag{6.9}$$

The parameter ρ determines how quickly the effective radius of the jet changes with the jet p_{T}. The VR algorithm requires two additional parameters, R_{min} and R_{max}, which are the lower and upper cutoffs on the jet radius, respectively. These cutoffs prevent the jets from becoming too large at low p_{T} and from becoming smaller than the detector resolution at high p_{T}.

The optimal values for the parameters of the VR algorithm are determined by examining the truth subjet double b-labeling[7] efficiency from $H \to b\bar{b}$ decays [382]. Figure 6.12 shows the results for several scans of the parameters ρ, R_{min}, and R_{max}. In the figure, the truth subjet double b-tagging efficiency is shown for $R = 0.2$ fixed-radius track jets and different combinations of the VR algorithm parameters as a function of the Higgs jet p_{T}. VR track jets with $\rho = 30\,\text{GeV}$, $R_{\text{max}} = 0.4$, and $R_{\text{min}} = 0.02$ yields the highest truth subjet double b-labeling efficiency across the whole Higgs jet p_{T} range. These are the values used in the Boosted $H \to b\bar{b}$ Analysis.

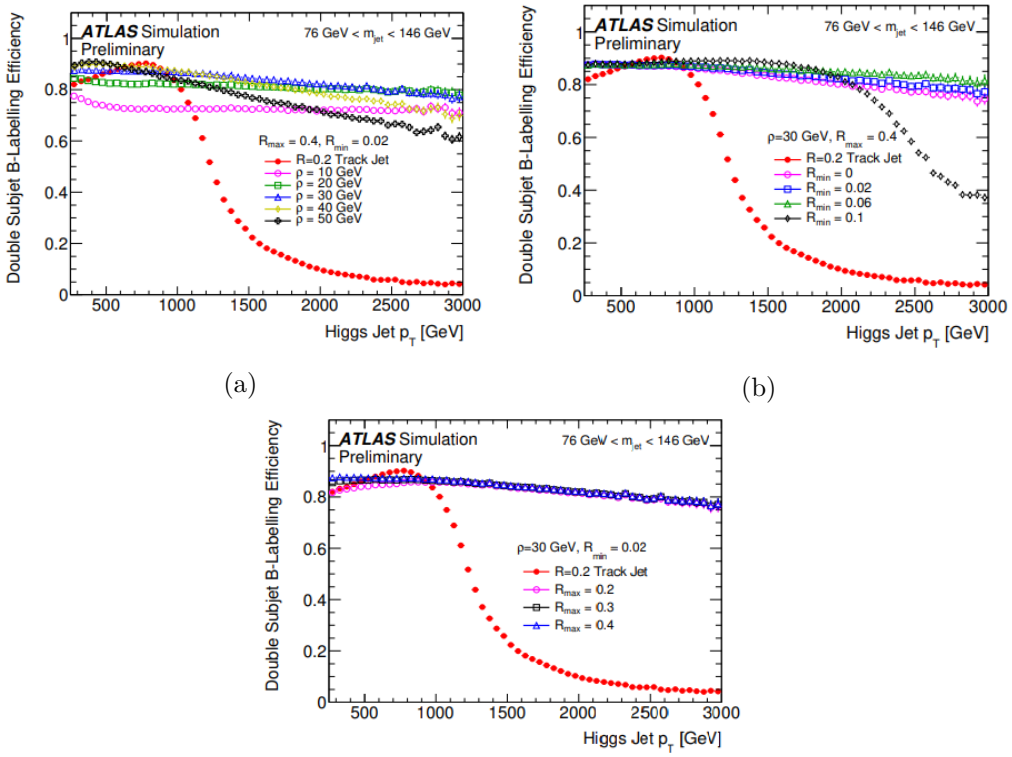


Figure 6.12: Efficiency of subjet double b-labeling at the truth level of a Higgs jet as a function of the Higgs jet p_{T}. (a) The efficiency for VR track jets with $R_{\text{min}} = 0.02$ and $R_{\text{max}} = 0.4$ for several ρ values. (b) The efficiency for VR track jets with $\rho = 30\,\text{GeV}$ and $R_{\text{max}} = 0.4$ for different values of R_{min}. (c) The efficiency for VR track jets with $\rho = 30\,\text{GeV}$ and $R_{\text{min}} = 0.02$ for varying values of R_{max}. The efficiency for $R = 0.2$ track jets is also included in all of the plots. The error bars include statistical uncertainties only [382].

[7] Double b-labeling refers to b-tagging.

6.5.3 Jet Calibration

Because the sensitivity of any analysis relying on the use of jets depends on an accurate knowledge of the mass and transverse momentum responses in the detector, it is crucial that we perform appropriate calibrations [383, 384]. For this reason, jets undergo a series of calibration steps to address effects from signal losses, detector effects, and mismodeling in simulation. The calibrations performed differ slightly between different jet definitions, but they share the same underlying principles. In the following sections, we'll look at the steps for small radius (small-R) and large radius (large-R) jets. The steps for small-R and large-R jets are briefly outlined in the flow charts of Figures 6.13 and 6.14, respectively.

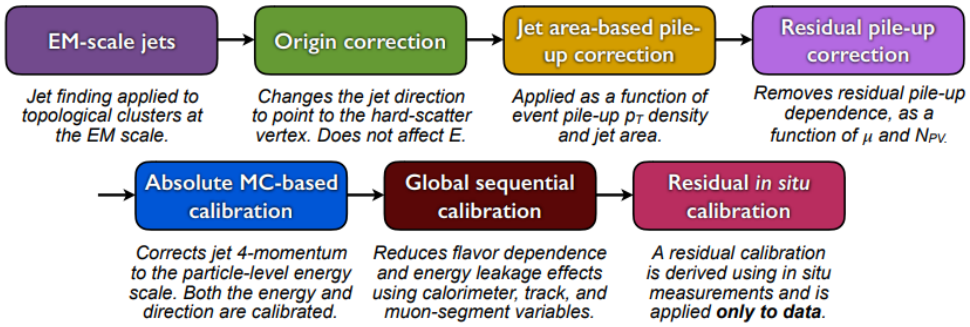


Figure 6.13: Calibration steps for EM-Scale small-R jets. Other than the origin correction, each step is applied to the four-momentum of the jet [385].

As discussed at the beginning of this section, calorimeter jets are constructed from topologically connected clusters, which enter the jet-reconstruction algorithms as massless particles. These topo-clusters are initially reconstructed at the electromagnetic (EM) scale [386–389], which measures the energy deposited in the calorimeter by particles produced by electromagnetic showers. A second topo-cluster collection is formed by calibrating the calorimeter cells such that the response of the calorimeter to hadrons is correctly reconstructed. This calibration uses the local-

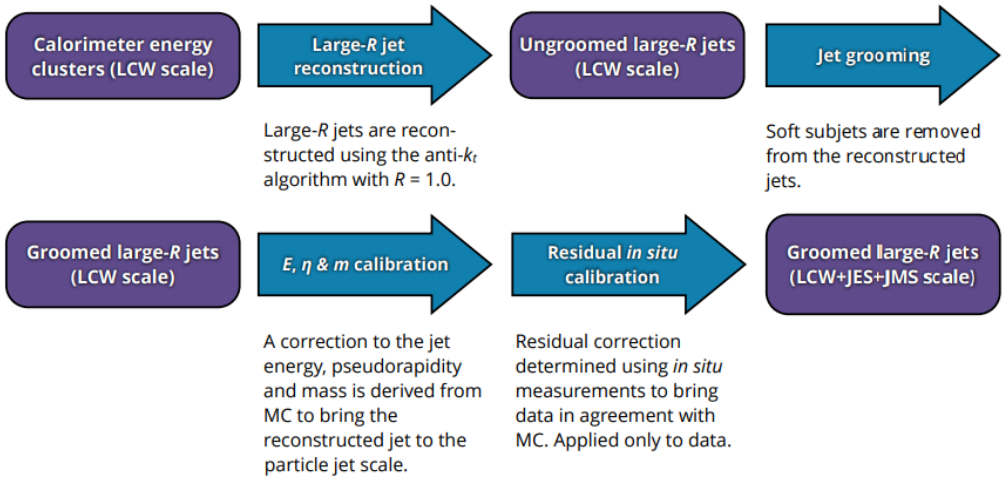


Figure 6.14: Calibration steps for large-R jets. Large-R-jet constituents formed from calorimeter energy clusters have already been adjusted to point at the event's primary hard-scatter vertex [383].

cell-signal weighting (LCW) method [376], which gives us an improved resolution compared to the EM scale by correcting the signals from hadronic deposits, reducing the fluctuations caused by the non-compensating nature of the ATLAS calorimeter systems. In the LCW method, we begin by classifying topo-clusters as either electromagnetic or hadronic, primarily using the measured energy density and longitudinal shower depth. Energy corrections are derived according to this classification from single charged and neutral pion MC simulations. Dedicated corrections address effects of calorimeter non-compensation, signal loses due to noise threshold effects, and energy lost in non-instrumented regions close to the cluster [390]. Typically, small-R jets are calibrated to the EM scale, while large-R jets are calibrated using the LCW scale.

Origin Correction

Nominally, the angular directions (η, ϕ) of the topo-cluster constituents for a jet point towards the center of the ATLAS detector. This is adjusted in the jet-

origin correction such that they point back to the hard-scatter vertex of the event. The origin-corrected calorimeter jet four-momentum is then the sum of the updated topo-cluster four-momenta.

Pileup Correction

Although the high luminosity opens the path for searches on rare SM and new physics processes, it complicates analysis by overlaying multiple soft interactions on top of the more interesting hard events we care about. This excess from the soft interactions is known as pileup (PU). While the Inner Detector can aid in sorting out uninteresting charged particle interactions by associating them to distinct vertices, neutral particles cannot be tracked here and we have to rely on the poorer spatial resolution of the calorimeters. Pileup corrections differ between small-R and large-R jets.

For small-R jets, the PU is corrected using a jet area-based PU correction and then a residual PU correction. The area-based PU correction is performed by subtracting the expected pileup contribution for a given jet by estimating the event-by-event pileup density (ρ) using $R = 0.4$ central jets ($|\eta| < 2.0$) from the k_T algorithm, which are sensitive to soft radiation and the transverse momentum. The area of the anti-k_t jets is calculated using an approach known as jet ghost association [350], where in addition to the real particles, *ghost particles* with an infinitesimal p_{T} are introduced, leaving the angular coordinates unchanged. The density is defined as:

$$\rho = \text{median} \left(\frac{p_{\mathrm{T}}^i}{A^i} \right) \Bigg|_{k_T \text{ jets}} \tag{6.10}$$

From there, using the pileup density and the jet area, the expected pileup contribu-

tion is subtracted from the jet. Additionally, since MC is generated with a set-value of PU, which will not necessarily be the same as that in data, the MC PU is reweighted to that of the data. The correction factors are functions of the pseudorapidity of the jet, the number of primary vertices N_{PV}, and the average number of interactions per bunch crossing μ. The correction factors are determined from simulation and allow for in-time and out-of-time PU effects. These two corrections for pileup are used to correct the transverse momentum:

$$p_{\mathrm{T}}^{\mathrm{corr}} = p_{\mathrm{T}}^{\mathrm{EM}} - \rho \times A - \alpha(\eta) \times (N_{\mathrm{PV}} - 1) - \beta(\eta) \times \mu. \tag{6.11}$$

Because large-R jets cover a larger area than small-R jets, they are more susceptible to pileup. To remove soft subjets created by PU, a trimming (grooming) procedure is used [391]. This algorithm takes advantage of the fact that pileup contamination resulting from multiple parton interactions (MPI) and initial-state radiation (ISR) is often much softer than the out-going partons associated with the hard-scatter and their final-state radiation (FSR). Removing these softer components from the reconstructed jet preferentially removes radiation from PU, MPI, and ISR, while only partially discarding some of the hard-scatter and FSR. The trimming procedure begins by using the k_t-algorithm to create subjets with radius R_{sub} from the constituents of the large-R jet. Subjets below some fraction of the transverse momentum of the total jet $p_{\mathrm{T}}^{i}/pt^{\mathrm{Jet}} < f_{\mathrm{cut}}$ (*e.g.* $f_{\mathrm{cut}} = 0.05$) are removed. This tends to remove a greater fraction from low-mass jets produced from light-quarks or gluons, but less from the decay products of boosted high-mass particles. The trimming procedure can be seen diagrammatically in Figure 6.15.

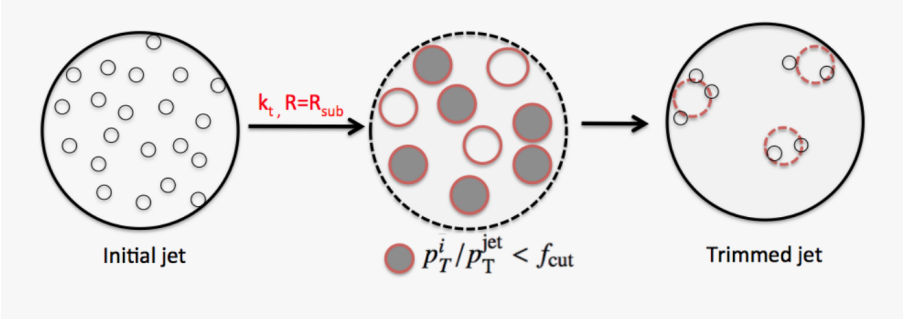


Figure 6.15: Diagram depicting the jet trimming procedure [392].

Energy and Mass Scale Calibration

Due to losses in the calorimeters which are not compensated for, along with multiple other detector effects, the EM-scale energy of a jet will tend to be lower than that of the true, or particle-level energy. These corrections are known as the jet energy scale (JES) [393] and jet mass scale (JMS) [394] corrections. To account for this using Monte Carlo, the truth and reconstructed jets are matched on a jet-by-jet basis using their angular distance (ΔR) to construct the response. Those jets without a corresponding truth jet are thrown out. The jet-energy and jet-mass responses are formed by the ratios of the reconstructed to the truth-level quantities:

$$\mathcal{R}_m = \frac{m_{\text{reco}}}{m_{\text{truth}}} \tag{6.12}$$

$$\mathcal{R}_E = \frac{E_{\text{reco}}}{E_{\text{truth}}}, \tag{6.13}$$

where $m_{\text{reco}}/E_{\text{reco}}$ are the reconstructed mass/energy after the pileup correction and $m_{\text{truth}}/E_{\text{truth}}$ are the truth-level jet quantities. These response variables are binned in η and E/p_{T} to account for effects that depend upon the location of the jet in the detector and the energy/momentum, respectively, which causes particles to deposit

energy in close-by, or the same, calorimeter clusters, degrading the reconstructed quantities. Using the response variables, we can derive a calibration factor for the jet-energy and jet-mass scales, which is the inverse of the average response, *i.e.*,

$$c_{\mathrm{JES/JMS}} = \frac{1}{\left\langle \mathcal{R}_{\mathrm{JES/JMS}} \right\rangle}. \tag{6.14}$$

Global Sequential Calibration

The numerical-inversion calibration discussed above helps to restore the average reconstructed jet energy/mass to the mean truth-jet energy/mass, but other jet characteristics, such as the parton flavor, affect the jet fragmentation and hence produce further deviations in the jet response. To improve the precision of the jet calibration for small-R jets, the Global Sequential Calibration (GSC) [395] uses information of the jet shower shape. For example, quarks and gluons hadronize differently, resulting in different responses within the detector. Gluon-initiated jets tend to have a high multiplicity of soft jets. On the other hand, quark-initiated jets have hadrons which carry a higher fraction of the momentum. The GSC uses five variables as its input to improve the resolution by using subsets of the detector:

- f_{Tile0}, $|\eta_{\mathrm{det}}| < 1.7$: the fraction of jet energy in the first layer of the Tile Calorimeter
- f_{LAr3}, $|\eta_{\mathrm{det}}| < 3.5$: the fraction of jet energy in the third layer of the EM Calorimeter
- n_{trk}, $|\eta_{\mathrm{det}}| < 1.7$: the number of tracks with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$ associated to the jet
- $\mathcal{W}_{\mathrm{trk}}$, $|\eta_{\mathrm{det}}| < 2.5$: the width of the tracks associated to the jet, weighted by

their p_{T}.

- n_{segments}, $|\eta_{\mathrm{det}}| < 2.7$: the number of muon tracks associated to the jet

which are applied sequentially [395]:

$$c_{\mathrm{GSC}} = \prod_{i=1}^{N} c_i(p_{\mathrm{T}}^{i-1}, x_i). \tag{6.15}$$

These variables combined characterize the longitudinal and transverse distribution of the energy deposited into the calorimeters by the jet.

In Situ Calibration

Even after the MC-based corrections are used to correct for the EM-scale response, further biases from mismodeling in MC simulation exist and are corrected for using data. These mismodeling effects are determined using events with topologies for which jets are expected to have a balanced momentum with respect to a well-measured reference object. To account for the detector effects not captured in simulation, a residual *in situ* correction [383] is derived as the p_{T} response ratio between data and MC and is applied to the data:

$$\frac{1}{c_{in\ situ}} = \frac{\mathcal{R}_{\mathrm{data}}}{\mathcal{R}_{MC}} = \frac{\left\langle p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\mathrm{ref}} \right\rangle_{\mathrm{data}}}{\left\langle p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\mathrm{ref}} \right\rangle_{\mathrm{MC}}}. \tag{6.16}$$

The *in situ* calibration first measures the JES using the same method to calibrate small-R jets [385], which relies on the p_{T}-balance in several final states. These reference events (dijet, Z-jet or γ-jet, and multijet events) are shown in Figure 6.16.

The JES calibration factor relies on an absolute and relative calibration. The absolute calibration is formed from a statistical combination of three measurements

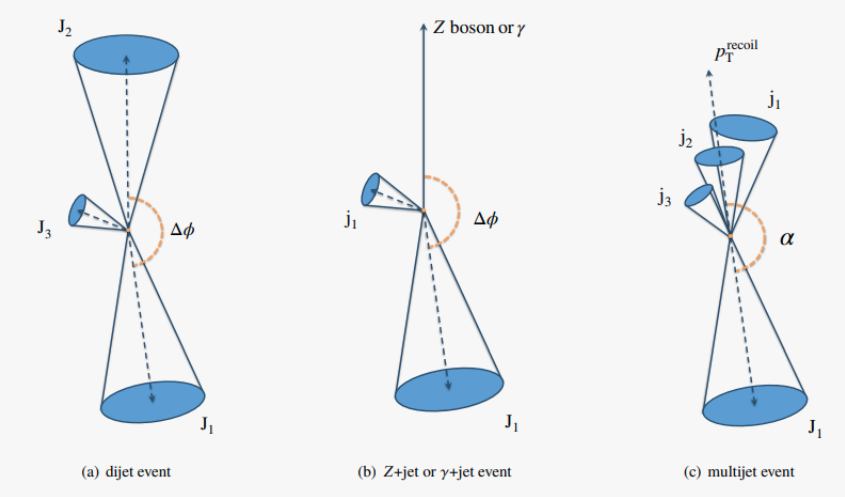


Figure 6.16: Schematic representation of the events used to measure the JES and jet mass resolution (JMR): (a) dijet event, (b) Z + jet or γ + jet event, and (c) a multijet event with several jets recoiling against the leading large-R jet. J_i refers to the ith leading large-R jet, j_i is the ith leading small-R jet fulfilling $\Delta R(J_1, j) > 1.4$. $\Delta\phi$ is the difference between the azimuthal angle of the jet and the reference object, while $\Delta\alpha$ is the difference between the azimuthal angle of the jet and the vectorial sum of the recoil system momenta [383].

from Z+jet, γ+jet, and multijet events in the central region of the detector. For the relative intercalibration, multijet events are used to propagate the well-measured central JES into the forward region of the detector. Because this JES correction is applied to the jet four-momenta, it also affects the jet mass calibration. So as a second step, the jet mass response is measured using lepton+jets in top-quark pair production by fitting the peaks in the jet mass distribution formed by W bosons and top quarks decaying to fully hadronic final states at high transverse momentum. An additional measurement, called the R_{trk} method [392], takes advantage of the independent measurements by the calorimeters and the Inner Detector to provide a calibration for the calorimeter jet-mass measurement over a broad range in p_{T}.

Additionally, the jet energy resolution (JER) and jet mass resolution (JMR) are measured *in situ* and compared with the prediction from simulation. The JER is extracted by taking advantage of the transverse momentum balance in dijet events.

JMR is obtained from fits to the top-quark and W-boson mass peaks in high-p_{T} lepton+jet $t\bar{t}$ events. We will see an example of extracting an independent JMR measurement in a fit between data and MC using a dedicated control region in the Boosted $H \to b\bar{b}$ Analysis in Section 9.5.

6.5.4 Combined Large-R Jet Mass

The large-R jet mass can be obtained from both the ID and the calorimeters. These systems have different resolutions depending on whether the particles are neutral or charged. For example, the ID is excellent for tracking charged particles, but neutral ones pass through unobstructed, only to be picked up in the calorimeters. While the calorimeters are able to give us an accurate determination of the energy (after appropriate calibrations) of particles by absorbing them, their spatial resolution is poorer compared to that of the silicon tracking system. As the goal is to obtain the highest possible precision of the large-R jet mass, we used the combined output of the tracking system and calorimeters to determine the mass and resolution of large-R jets. This combination is referred to as the *combined mass* [396].

The calorimeter mass m^{calo} for a large-R jet J with cell-cluster constituents possessing energy E_i and momentum $\vec{p}_i$ is defined as:

$$m^{\mathrm{calo}} = \sqrt{\left(\sum_{i \in J} E_i\right)^2 - \left(\sum_{i \in J} \vec{p}_i\right)^2}. \tag{6.17}$$

The track-assisted jet mass m^{TA} is the mass calculated in the tracker, scaled by the ratios of the momentum of the calorimeter and the tracker, defined as:

$$m^{\mathrm{TA}} = \frac{p_{\mathrm{T}}^{\mathrm{calo}}}{p_{\mathrm{T}}^{\mathrm{track}}} m^{\mathrm{track}}. \tag{6.18}$$

The combined mass is a linear combination of the calorimeter and track-assisted masses:

$$m^{\text{comb}} = w_{\text{calo}} m^{\text{calo}} + w_{\text{TA}} m^{\text{TA}}, \tag{6.19}$$

with the weights based on the resolutions of the track-assisted and calorimeter masses, constrained to sum to unity:

$$w_{\text{calo}} = \frac{\sigma_{\text{calo}}^{-2}}{\sigma_{\text{calo}}^{-2} + \sigma_{\text{TA}}^{-2}} \tag{6.20}$$

$$w_{\text{TA}} = \frac{\sigma_{\text{TA}}^{-2}}{\sigma_{\text{calo}}^{-2} + \sigma_{\text{TA}}^{-2}}, \tag{6.21}$$

which can also be expressed directly in terms of the combined, track-assisted, and calorimeter masses:

$$w_{\text{calo}} = \frac{m^{\text{comb}} - m^{\text{TA}}}{m^{\text{calo}} - m^{\text{TA}}} \tag{6.22}$$

$$w_{\text{TA}} = \frac{m^{\text{calo}} - m^{\text{comb}}}{m^{\text{calo}} - m^{\text{TA}}}. \tag{6.23}$$

Using the combined mass greatly improves the jet mass resolution, especially at high transverse momentum when compared to the calorimeter mass alone. This can be seen in Figure 6.17.

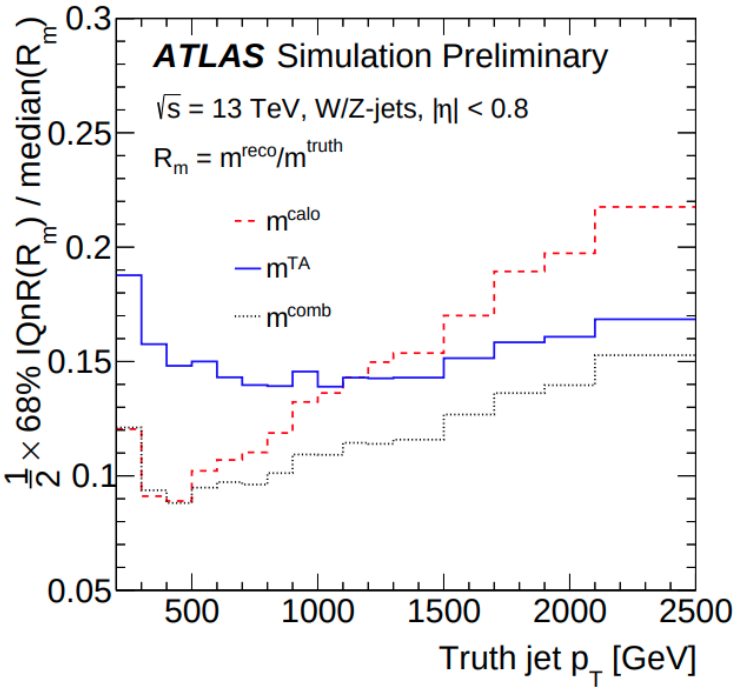


Figure 6.17: The resolution of the jet mass response as a function of the truth-jet p_{T} for W/Z boson jets for calorimeter-based (red dashed line) and track-assisted (solid blue line) jet mass compared to the combined (gray dashed line) jet mass [396].

6.6 Large-R Jet Mass Resolution

6.6.1 Semileptonic Decays

When a bottom quark decays, it primarily does so via a charm quark and a W boson. The W boson then decays either hadronically (into $q\bar{q}'$ pairs) or leptonically (into $\ell\nu$ pairs) as seen in Figure 6.18. When the W boson decays into a muon and its associated neutrino, the energy from both particles is not captured in either calorimeter[8]. This results in a shift towards lower values in the reconstructed mass, energy, and p_{T} of the jet. Standard jet-energy calibrations do not take the effects from semileptonic decays into account, which leads to a degradation in mass resolution. In what follows, we will be referring to semileptonic decays involving a muon,

[8] A small fraction of the muon's energy is deposited in the calorimeters.

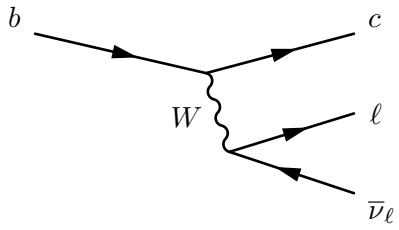


Figure 6.18: Bottom quark decaying semileptonically.

but the principle is similar for electrons. Figure 6.19 shows an example, comparing the reconstructed mass between hadronically and semileptonically decaying bottom quarks in large-R jets truth-matched to the Higgs boson. Here, the large-R jet is considered to have a semileptonic decay when at least one correction is applied to the jet (*i.e.* when at least one of the b-quarks decays semileptonically.)

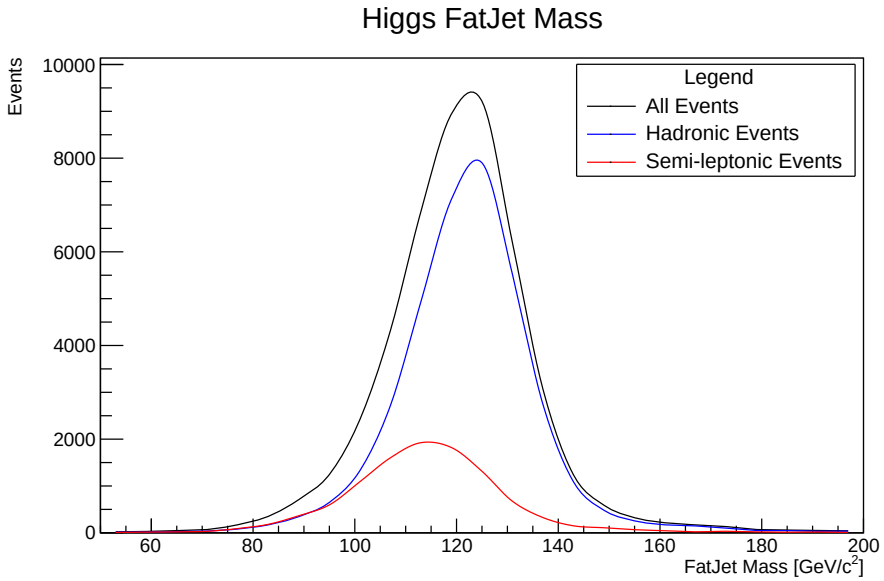


Figure 6.19: Higgs mass distribution divided into the semileptonically and hadronically decaying bottom quarks from the Higgs boson. Semileptonic decays reveal a wider mass distribution than hadronic decays, with a peak shifted towards lower mass.

A partial remedy to this problem is to add the 4-momentum of the lepton back into the candidate large-R jet. A neutrino correction has been considered in other

analyses within ATLAS, but no satisfactory solution has been found. To counter the loss caused by semileptonic decays from b-quarks, muons can be identified and their 4-momenta added to that of the large-R jet. Using electrons as well would be ideal, but due to their higher fake rates and larger energy deposits in the calorimeters compared to muons, more care must be taken. The muon-in-jet (MIJ) correction to large-R jets can improve the resolution of the reconstructed mass for heavy resonances. The Boosted $H \to b\bar{b}$ Analysis follows the prescription of the $VH(bb)$ analysis [23] regarding the MIJ correction[9] and I implemented and optimized it in the xAODAnaHelpers [397] package used to perform our analysis.

6.6.2 Muon-in-Jet Correction: Implementation

Muons used in the correction are required to be of *medium* quality and pass the *loose* isolation criteria [365]. Only muons with transverse momentum above 10 GeV and $|\eta| < 2.4$ are considered for use in a correction. Muons are matched to the nearest b-tagged variable-radius (VR) jet associated to a large-R jet and are required to satisfy the following ΔR requirement:

$$\Delta R_{\mu,\text{VR jet}} < \min(0.4, 0.04 + 10\,\text{GeV}/p_{\text{T}}^{\mu}), \tag{6.24}$$

where p_{T}^{μ} is the transverse momentum of the muon. The match between b-tagged VR jets and muons is one-to-one, but multiple b-tagged VR jets in a given large-R jet can have matched muons for the correction. The rate of multiple muons used for correction is quite small, accounting for only $\sim$ 4% of events, and does not significantly depend on the process (QCD dijet vs ggF events for example). In the

[9] Our analysis thanks Chikuma Kato for his help and instruction for implementing the MIJ correction.

Boosted $H \to b\bar{b}$ Analysis, the large-R-jet combined mass [396] is used. Here we will see how the MIJ correction is performed in the *combined* scheme because it uses the schemes for the track-assisted and calo jet masses. In the *combined* muon-in-jet correction scheme, the 4-momenta of the calorimeter-calibrated large-R jet mass is used. Before the muon 4-momenta is added to the large-R jet, the amount of energy deposited in the calorimeters is subtracted from the muon to avoid double-counting. In other words,

$$p^{\mu}_{\text{J, corr}} = p^{\mu}_{\text{J}} + p^{\mu}_{\text{muon}} \ - p^{\mu}_{\text{muon, eLoss}}, \tag{6.25}$$

where J represents the large-R jet, the μ is used to represent indices of the 4-momentum vector and *not* the muon, and the eLoss is the fraction of energy deposited into the calorimeter by the muon. The $p^{\mu}_{\text{muon, eLoss}}$ vector is anti-parallel to p^{μ}_{muon}, but of course of a smaller magnitude. The track-assisted mass is then corrected using the ratio of the corrected to the uncorrected calorimeter transverse momenta:

$$\begin{aligned} m^{\text{TA}}_{\text{corr}} &= \frac{p^{\text{calo,corr}}_{\text{T}}}{p^{\text{calo}}_{\text{T}}} m^{\text{TA}}, \\ &= \frac{p^{\text{calo,corr}}_{\text{T}}}{p^{\text{track}}_{\text{T}}} m^{\text{track}}. \end{aligned} \tag{6.26}$$

The transverse momentum is then calculated using the energy from the corrected calorimeter 4-vector and mass from the corrected track-assisted 4-vector, while main-

taining the original η of the uncorrected track-assisted large-R jet.

$$p_{\mathrm{T}}^{\mathrm{TA,corr}} = \frac{\sqrt{\left(E_{\mathrm{corr}}^{\mathrm{calo}}\right)^2 - \left(m_{\mathrm{corr}}^{\mathrm{TA}}\right)^2}}{\cosh(\eta^{\mathrm{TA}})} = |\vec{p}_{\mathrm{TA,corr}}| \sin(\theta^{\mathrm{TA}}). \tag{6.27}$$

The corrected track-assisted mass and transverse momentum, together with the calorimeter η and ϕ are used to construct the corrected track-assisted 4-vector. The weights defined in Equations 6.20 and 6.21 are then used to produce the corrected combined mass:

$$m^{\mathrm{comb,corr}} = w_{\mathrm{calo}} m^{\mathrm{calo,corr}} + w_{\mathrm{TA}} m^{\mathrm{TA,corr}}, \tag{6.28}$$

and the corrected combined transverse momentum:

$$p_{\mathrm{T}}^{\mathrm{comb,corr}} = \frac{\sqrt{\left(E_{\mathrm{corr}}^{\mathrm{calo}}\right)^2 - \left(m_{\mathrm{corr}}^{\mathrm{comb}}\right)^2}}{\cosh(\eta^{\mathrm{calo,corr}})} = |\vec{p}_{\mathrm{comb,corr}}| \sin(\theta^{\mathrm{calo,corr}}). \tag{6.29}$$

The final 4-momentum uses the corrected combined mass and momentum above with the corrected calorimeter η and ϕ.

$$p_{\mathrm{T}}^{\mathrm{TA,corr}} = \frac{\sqrt{\left(E_{\mathrm{corr}}^{\mathrm{calo}}\right)^2 - \left(m_{\mathrm{corr}}^{\mathrm{TA}}\right)^2}}{\cosh(\eta^{\mathrm{TA}})} = |\vec{p}_{\mathrm{TA,corr}}| \sin(\theta^{\mathrm{TA}}). \tag{6.30}$$

6.6.3 Muon-in-Jet Correction: Optimization

When implementing the MIJ correction, it is important to maximize the number of instances in which a genuine correction should occur (the efficiency), while simultaneously minimizing the instances in which a muon is inappropriately used to make a correction (the mistag rate). For this reason, I performed several optimiza-

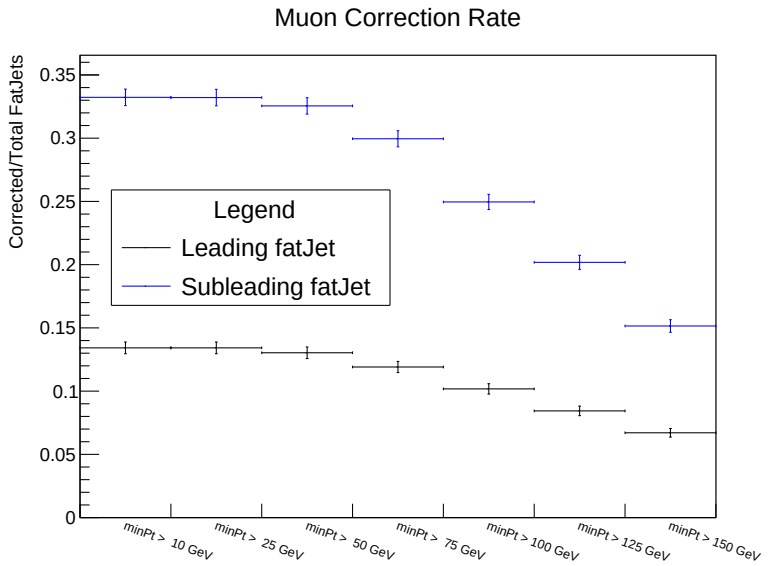
tion studies using a ggF Higgs MC sample (see Section 9.2 for details on samples used in the analysis). By optimize, we mean that we wish to determine the mass and p_{T} cuts on the muons and VR jets which yield an acceptable compromise between mistag rates (muons used to correct non-semileptonically decaying b-tagged VR jets) and corrections (which improve the resolution of the large-R jet resonances).

The minimum muon-p_{T} cut of 10 GeV was chosen after an optimization study described in this section. The ggF Higgs signal MC samples were used and the candidate was jet truth-matched to the Higgs for this study. Events were required to have a leading candidate jet with $p_{\mathrm{T}} > 410\,\mathrm{GeV}$ and two b-tagged VR track jets which each have $p_{\mathrm{T}} > 10\,\mathrm{GeV}$. In the initial studies performed, the minimum p_{T} of the VR jets (used for muon association) and the p_{T} of the muons used for the correction were varied to determine the mistag and correction rates on the large-R jets. The correction rate is defined as:

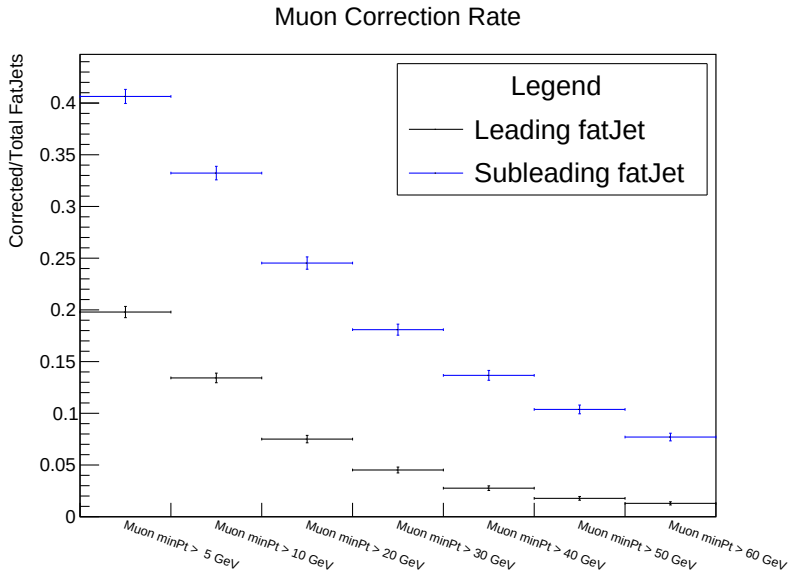
$$\text{Correction Rate} = \frac{\text{Corrected large-}R\text{ Jets}}{\text{All large-}R\text{ Jets}}, \tag{6.31}$$

and the mistag rate is computed by counting how often a muon correction occurs, but the b-hadron decay did not actually include a muon. Note, a cascade decay with $(b \to c \to \mu X)$ is *not* a mistag.

The mistag rate is relatively flat for variations on the minimum p_{T} of the VR jet up to 50 GeV and the rate drops dramatically for variations on the minimum p_{T} of the muon from 5 to 10 GeV; therefore, the default cut value on the muon minimum p_{T} of 10 GeV was maintained to provide a good compromise between correction and mistag rates (see Figures 6.20 and 6.21).

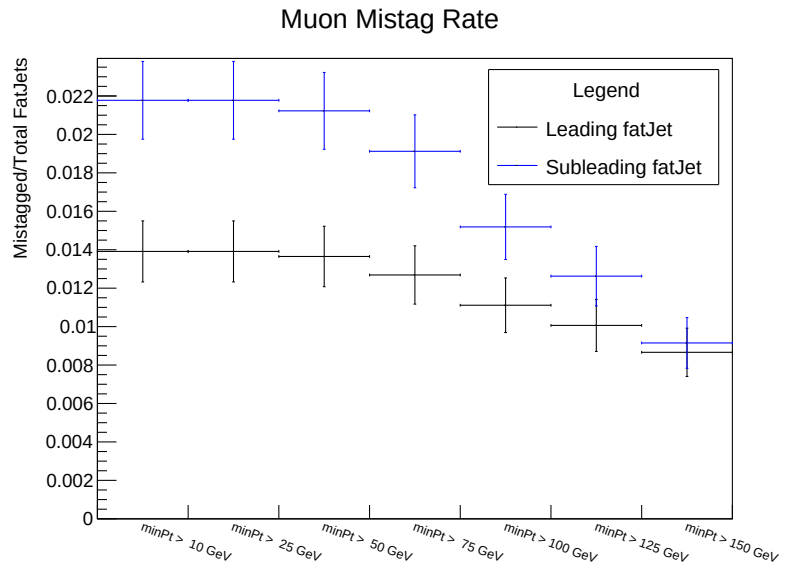


(a)

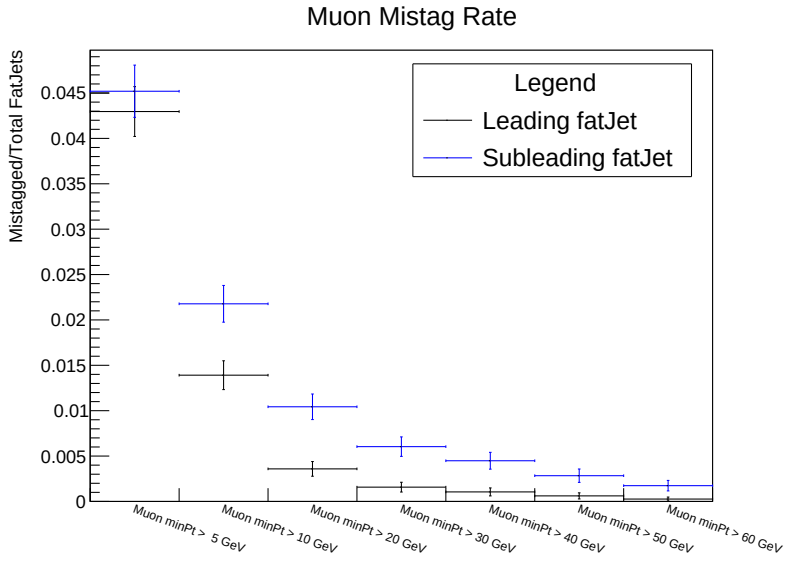


(b)

Figure 6.20: Muon correction rates for variations on the VR jet minimum p_{T} (top) and the muon minimum p_{T} (bottom) indicated with the minPt.



(a)



(b)

Figure 6.21: Muon mistag rates for variations on the VR jet minimum p_{T} (top) and the muon minimum p_{T} (bottom) indicated with the minPt.

6.6.4 Containment

The opening angle of the decay products coming from a massive particle is roughly $2m/p_{\mathrm{T}}$ [236, 252]. The mass of a particle is essentially a fixed quantity, so the opening angle is dependent upon the transverse momentum of the particle. At low transverse momentum, the decay products from the massive particle have a large opening angle which may be larger than the radius of the large-R jet. This can be seen pictorially in Figure 6.22. This loss in containment will cause some of

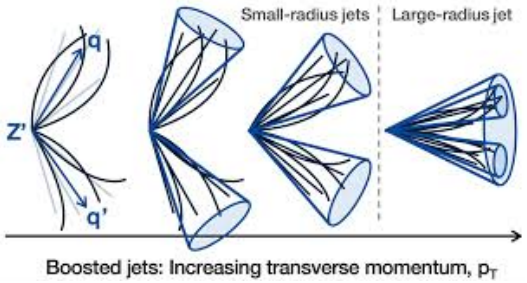


Figure 6.22: At low transverse momentum, the decay products form jets back-to-back (left). As momentum increases (moving right), the angle between the jets decreases until they are close enough to be contained within the same large-R jet (right) [398].

the mass/energy to not be accounted for, leading to a lower reconstructed mass with a wider distribution. As the p_{T} increases, the decay products become more collimated, allowing the energy to be deposited fully within the jet radius. So we expect that, as a function of p_{T}, the resolution will improve, and the reconstructed mass peak will be higher. This is indeed what we see when looking at the ggF MC sample in Figure 6.23. The results in the plot should be taken as more qualitative than quantitative, however, as no consideration of systematics nor scale factors have been taken into account. The containment is important to consider as the Boosted $H \to b\bar{b}$ Analysis reconstructs the Higgs mass in several p_{T} bins.

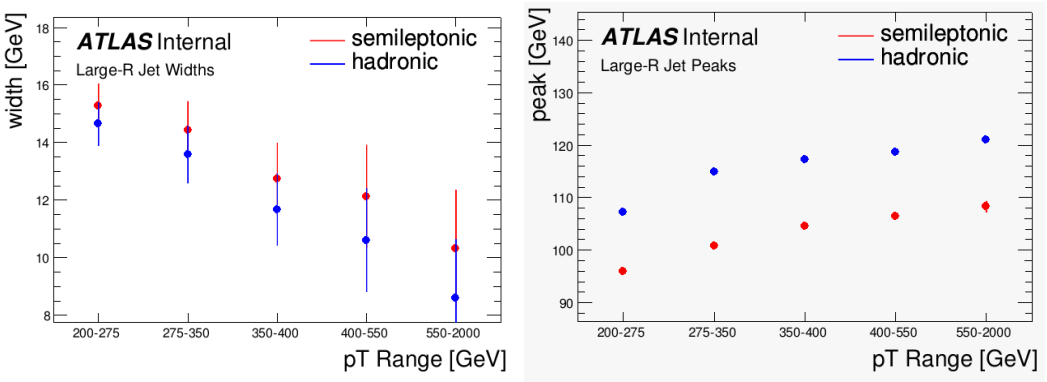


Figure 6.23: The width (left) and peak (right) of the reconstructed jet mass for jets truth-matched to the Higgs as a function of p_{T}.

Chapter 7

Boosted Hbb Analysis

We now turn to the heart of this dissertation, the Boosted $H \to b\bar{b}$ Analysis. Previously [218], ATLAS measured the inclusive (in p_{T} and production mode) cross section for $H \to b\bar{b}$ at high transverse momentum using the roughly $80\,\mathrm{fb}^{-1}$ of Run 2 data collected up to that point. In addition to the Higgs boson measurement, a search for new resonances decaying to a pair of bottom quarks was performed. For the Standard Model Higgs boson, the observed signal strength was found to be $\mu_H = 5.8 \pm 3.1\,(\mathrm{stat.}) \pm 1.9\,(\mathrm{syst.}) \pm 1.7\,(\mathrm{theo.})$. In the latest incarnation of the Boosted analysis, we shift from the static properties of the Higgs boson to its dynamic properties at high p_{T}, measuring the Higgs signal strength in bins of p_{T}. This search gives us indirect sensitivity to BSM effects at large momentum transfers. In addition to a measurement of the Higgs signal strength, the the Z boson signal strength is measured to validate the background fit model. Two similar searches were performed by CMS, one using the partial Run 2 data [219], and the other using full $137\,\mathrm{fb}^{-1}$ of data [222] collected by CMS. We will discuss the CMS results in Chapter 12 after revealing the results from the ATLAS full Run 2 Boosted $H \to b\bar{b}$ [220, 221] Analysis.

With Run 2 complete, we now have nearly twice the data collected previously and have made several improvements in the analysis techniques that have increased the sensitivity of our search. For example, the consideration of the subleading (in p_{T}) large-R jet has extended the search to lower transverse momentum, increasing the analysis sensitivity for $p_{\mathrm{T}}^{H} > 450\,\mathrm{GeV}$ by 11%, and the additional statistics have allowed us to develop the analysis techniques for a p_{T}-differential study of the Higgs cross section.

Event classification sorts the large-R jets by their transverse momentum into signal regions (SRs) for the leading (SRL) or subleading (SRS) jets in the event, which are candidates for the Higgs boson. Validation regions (VRs) for leading (VRL) and subleading (VRS) jets are constructed from anti-b-tagged large-R jets and are used in background modeling validation. A $t\bar{t}$ control region, the $\mathrm{CR}_{t\bar{t}}$ (Section 8.3), uses data to correct the normalization of the MC large-R jet mass template for $t\bar{t}$, triggering on high-p_{T} muons to select for semileptonic $t\bar{t}$ decays. Another control region, the $\mathrm{WCR}_{t\bar{t}}$, is designed to isolate W bosons from semileptonic $t\bar{t}$ decays (Section 9.5) to place constraints on the jet mass resolution in $V+\mathrm{jets}$, and triggers on lower-p_{T} muons. Additionally, muons are used as the input to the muon-in-jet (MIJ) correction (Section 7.3), which partially corrects the large-R jet four-momentum lost through leptonic decays of heavy hadrons. See Section 6.6.1 for details on the MIJ implementation and Section 7.3 for its impact on the boosted Higgs analysis.

The major Higgs production modes are those discussed in Section 3.1: gluon-gluon fusion (ggF), vector boson fusion (VBF), associated vector boson production (WH and ZH, or VH collectively), and associated top-quark production (tH and ttH). The major background contributions consist of QCD multijet, $V+\mathrm{jets}$, and

top-quark pair production ($t\bar{t}$). $V+$jets ($W+$jets and $Z+$jets) and $t\bar{t}$, along with the Higgs signal, are modeled using Monte Carlo. The QCD multijet background, by far the largest contribution, is modeled using a polynomial exponential functional form for the jet-mass distribution.

A Higgs signal strength is extracted in each p_{T} region of the analysis, shown in Table 7.1, in a simultaneous profile likelihood fit (Chapter 10) to the jet-mass distributions in the signal regions (SRL and SRS) and a separate control region ($\mathrm{CR}_{t\bar{t}}$), which uses data to correct the normalization of the $t\bar{t}$ background. Systematic uncertainties, also known as nuisance parameters, are incorporated into the likelihood as Gaussian penalty terms. An external constraint for the jet mass resolution (JMR) nuisance parameter in the $V+$jets process is applied to the fit to compensate for the mismatch between data and MC in the width of the reconstructed jet-mass distribution, and is derived from a separate control region ($\mathrm{WCR}_{t\bar{t}}$). This constraint is added to the fit and the control region itself is not included in the profile likelihood. The processes modeled with MC enter the fit as binned jet-mass distributions for large-R jets which pass the signal region criteria. The expected number of events, in each mass bin in the fit range, for each process comes from the MC distributions (H, $V+$jets, and $t\bar{t}$) or the background fit function (QCD multijet) and compared with data.

7.1 Reconstructed Physics Objects

7.1.1 Large-Radius Jets

The Higgs boson decay products are reconstructed into trimmed [383, 391] ($R_{\mathrm{sub}} = 0.2$ and $f_{\mathrm{cut}} = 0.05$) large-R ($R = 1.0$) jets, created from topological

clusters [376] in the calorimeters. The large-R jets are often colloquially referred to as *fatjets* or *fat jets*[1] (FJ), as will be seen in several of the figures in what follows. The jet clustering is performed by the anti-k_t algorithm [349, 350] implemented in FastJet [380]. These jets are calibrated using the p_{T} and η calibrations derived from simulation, as laid out in Section 6.5.3. To improve the jet mass resolution, especially at high transverse momentum, the combined jet mass is used; details can be found in Section 6.5.4. At least one large-R jet with $p_{\mathrm{T}} > 450\,\mathrm{GeV}$ and $m > 60\,\mathrm{GeV}$ is required to satisfy the trigger conditions. Triggering is covered generally in Section 6.1 and the analysis-specific triggers are discussed in Section 7.2.

For this analysis, a dedicated large-R jet energy correction, the muon-in-jet correction, is used to account for semileptonic b-hadron decays. Studies performed to optimize the correction algorithm for high correction rates with low fake rates are outlined in Section 6.6.1. Jets are p_{T}-ordered for event selection and categorization *before* the muon-in-jet correction is applied and are *not* re-ordered afterwards; the corrected jet mass is used in the final fit only, with the analysis binned in the corrected large-R jet p_{T}, providing the closest proxy for the truth Higgs p_{T} (p_{T}^{H}). Truth-matching reconstructed jets to the Higgs boson in Monte Carlo is performed by the ghost-association technique [350, 399]. Four p_{T} bins are used: 250–450, 450–650, and 650–1000, and $> 1000\,\mathrm{GeV}$, defined by the LHC Higgs Cross Section Working Group (LHCHXSWG) [104] to help theorists and experimentalists align their results. Table 7.1 shows the p_{T} ranges for the inclusive, fiducial, and differential signal regions in the analysis.

[1] As this analysis was the work of several people, there are multiple terms used to describe large-R jets.

Table 7.1: Summary of the signal regions and the corresponding large-R jet p_{T} ranges.

Region	Jet p_{T} [GeV] SRL	SRS
Inclusive	> 450	> 250
Fiducial	> 450	> 450
Differential	–	250–450,
	450–650,	450–650,
	650–1000,	650–1000
	> 1000	–

7.1.2 Variable-Radius Track Jets

Track jets are clustered using the ATLAS standard version of the variable-radius (VR) jet algorithm [382, 400], with $\rho = 30\,\mathrm{GeV}$, $R_{\mathrm{min}} = 0.02$, and $R_{\mathrm{max}} = 0.4$ on a subset of the reconstructed tracks. These track jets are ghost-associated to large-R jets *before* trimming. Only track jets supported by the Flavor Tagging Group with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$, $|\eta| < 2.5$, and at least two track constituents are considered in this analysis. Simulated track jets are labeled as b-, c-, or light-flavor depending on which hadrons ($p_{\mathrm{T}} > 5\,\mathrm{GeV}$) are located within $\Delta R < 0.3$ of the jet axis [244].

Variable Radius Track Jet Overlap Cleaning

The analysis follows the recommendations of the Flavor Tagging group [401] with regards to overlap cleaning [402]. If one of the two leading VR track jets ($p_{\mathrm{T}} > 10\,\mathrm{GeV}$) considered for b-tagging is found to overlap with another VR track jet ($p_{\mathrm{T}} > 5\,\mathrm{GeV}$), the large-$R$ jet is discarded without further consideration. Jets are considered to overlap when the distance (in ΔR) between two VR jets is less than the radius of the smaller of the two VR jets. In the signal sample, this constitutes a

7% loss in acceptance. The overlap between two VR track jets is defined as:

$$\Delta R_{\mathrm{VR,VR}i} < \min\left(R_{\mathrm{VR}}, R_{\mathrm{VR}i}\right), \tag{7.1}$$

where VR is the track jet under consideration, VRi loops over all track jets with $p_{\mathrm{T}} > 5\,\mathrm{GeV}$, and $R_{\mathrm{VR}i}$ is the variable radius of the corresponding track jet.

7.1.3 Flavor Tagging

Flavor tagging is based on the response of the MV2c10 Boosted Decision Tree (BDT) discriminant [244, 372], discussed in Section 6.4, which uses the output of low-level taggers to exploit the particle-track impact parameters, reconstructed secondary-vertex topology, and kinematics. To tag the two leading p_{T} VR track jets, the 77% fixed efficiency working point (WP) obtained for VR track jets in a $t\bar{t}$ sample is used. The light-flavor (u-, d-, s-quark, and gluon) and c-jet misidentification efficiencies are 0.9% and 25%, respectively. Only the flavor tag of the two leading (in p_{T}) VR track jets is considered. Table 7.2 shows the scale factors derived from comparisons of data and MC, as well as statistical, systematic, and total uncertainties in bins of jet p_{T} for the 77% MV2c10 WP [375].

Table 7.2: Measured data-to-simulation scale factors in the various jet p_{T} regions with statistical and systematic uncertainties (in %). The results are shown for a representative working point that corresponds to an overall b-tagging efficiency of 77% [375].

Jet p_{T} [GeV]	20–30	30–60	60–90	90–140	140–200	200–300	300–500
Scale Factor	1.014	1.032	1.029	1.016	0.998	1.014	1.034
Stat. Uncert. [%]	±1.4	±0.6	±0.6	±0.6	±1.1	±2.1	±6.2
Syst. Uncert. [%]	±14.3	±7.4	±6.9	±6.7	±4.9	±5.1	±5.2
Tot. Uncert. [%]	±14.3	±7.4	±6.9	±6.7	±5.1	±5.5	±8.1

There are additional effects on the jet b-tagging efficiency related to the angular distance to the closest jet [367] within the same large-R jet. This angular dependence is mostly due to the change in angular resolution with which the jet reproduces the heavy hadron flight direction. The jet angular resolution therefore may differ in data and simulation, resulting in a mismatch in the b-tagging efficiency with the angular distance to the closest jet. This will induce a discrepancy in the mass turn-on shape between data and MC because, at a given mass, the high-p_{T} variable-radius jets have a smaller opening angle. We unblinded data in the mass region below 70 GeV to test the mass turn-on effect due to the double-b-tagging requirement in the signal region. The correction factors obtained in the study were *not* applied to any sample used to extract the signal. They were only derived to demonstrate that the turn-on effect from b-tagging was understood and a fit range could be reliably fit. This study is covered briefly in Appendix C.11. See Appendix A of Reference [403] for the full study.

7.1.4 Muons

Muons are reconstructed using the Inner Detector and the Muon Spectrometer. Isolated muon selection criteria are:

- Combined muon,
- $p_{\mathrm{T}} > 10\,\mathrm{GeV}$,
- $|\eta| < 2.4$,
- *Medium* quality criteria [365],
- *FCLoose* muon isolation working point [365].

The various muon isolation points can be found in Table B.6 in Appendix C.2. Non-isolated muons with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ are used for the muon-in-jet correction with the same selection criteria outlined above, but without the muon isolation requirement. The $t\bar{t}$ Control Region ($\mathrm{CR}_{t\bar{t}}$), used to correct the normalization in MC to that of data, uses isolated muons with the requirements stated above and $p_{\mathrm{T}} > 52.5\,\mathrm{GeV}$. The other $t\bar{t}$ Control Region ($\mathrm{WCR}_{t\bar{t}}$) which is discussed in Section 9.5, is used to extract the jet mass resolution of large-R jets containing a resolved Wb from $t\bar{t}$ events in data and utilizes a lower-p_{T} trigger threshold of 27 GeV.

7.2 Trigger Strategy

The trigger strategy uses single large-R jet triggers. The trigger used evolves over the four years of data taking to adjust to machine conditions by increasing the p_{T} threshold and introducing a jet-mass requirement. This additional requirement on the jet mass allows for a reduction on the cut for the reconstructed large-R jet p_{T}. The relaxation on the p_{T} cut has a substantial impact on the analysis sensitivity due to the sharply falling p_{T} spectrum. A summary of the utilized triggers, along with the 99% efficiency points obtained via a fit to the Fermi function, is shown in Table 7.3. In order to maintain the gains provided by the newer triggers, while also keeping the number of categories used as input to the final fits manageable, the offline trigger cuts have been unified as $p_{\mathrm{T}} > 450\,\mathrm{GeV}$ and $m > 60\,\mathrm{GeV}$. Figures 7.1, 7.2, and 7.3 show the trigger efficiency in data and MC as a function of jet p_{T} for 2015/2016, and as a function of jet p_{T} and mass for 2017, and 2018, respectively. More plots from studies performed regarding the jet triggers are in Appendix B.4.

Table 7.3: Summary of the jet triggers (p_{T} and mass of the large-R jet, $p_{\mathrm{T,J}}$ and m_{J}, respectively) used in the analysis. They are applied as an OR and all are required to be active. The offline threshold is the offline jet-cut above which the triggers are 99% efficient.

Year	**Trigger**	**Offline Threshold** [GeV]	**Lum.** [fb^{-1}]
2015	`HLT_j360_a10_lcw_sub_L1J100`	$p_{\mathrm{T,J}} > 410$	3.2
2016	`HLT_j420_a10_lcw_L1J100`	$p_{\mathrm{T,J}} > 450$	33.0
2017	`HLT_j440_a10t_lcw_jes_L1J100`	$p_{\mathrm{T,J}} > 470$	41.2
	`HLT_j390_a10t_lcw_jes_30smcINF_L1J100`	$p_{\mathrm{T,J}} > 420$, $m_{\mathrm{J}} > 50$	41.0
2018	`HLT_j460_a10t_lcw_jes_L1J100`	$p_{\mathrm{T,J}} > 490$ m_{J}	58.5
	`HLT_j420_a10t_lcw_jes_35smcINF_L1J100`	$p_{\mathrm{T,J}} > 450$, $m_{\mathrm{J}} > 60$	58.5
	`HLT_j420_a10t_lcw_jes_35smcINF_L1SC111`	$p_{\mathrm{T,J}} > 450$, $m_{\mathrm{J}} > 60$	55.4

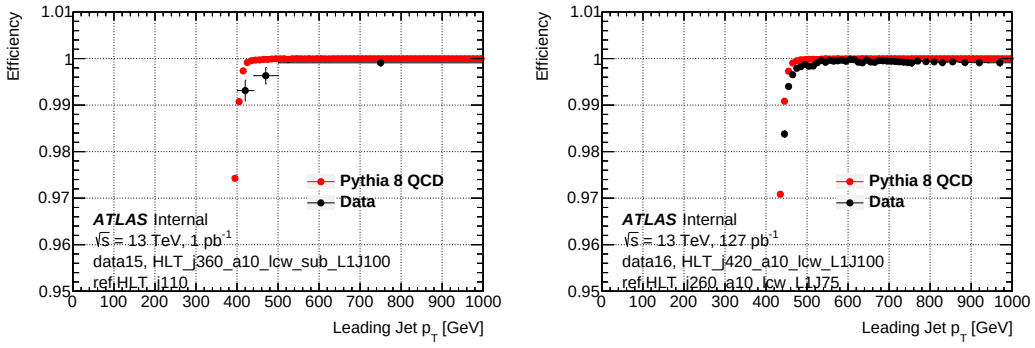
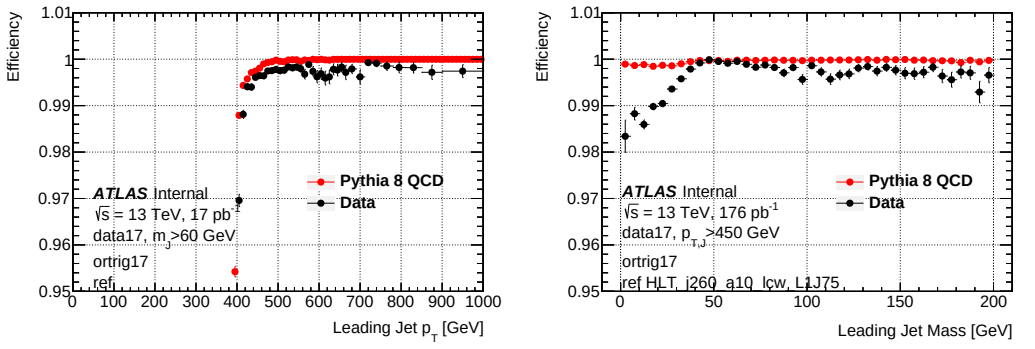
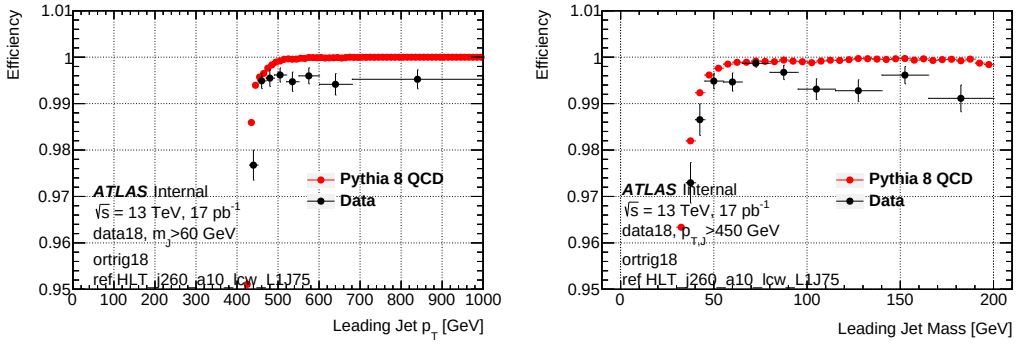


Figure 7.1: The efficiency of the trigger `HLT_j360_a10_lcw_sub_L1J100`, active in 2015 data-taking period (left) and the efficiency of the trigger `HLT_j420_a10_lcw_jes_L1J100`, active in 2016 data-taking period (right), as a function of the leading large-R jet p_{T} for data (black) and QCD Monte Carlo (red).



(a) (b)

Figure 7.2: The efficiency of the OR of the triggers `HLT_j390_a10t_lcw_jes_30smcINF_L1J100` and `HLT_j440_a10t_lcw_jes_L1J100`, active in 2017 data-taking period, as a function of the leading large-R jet p_{T} (left) and mass (right) for data (black) and QCD Monte Carlo (red).



(a) (b)

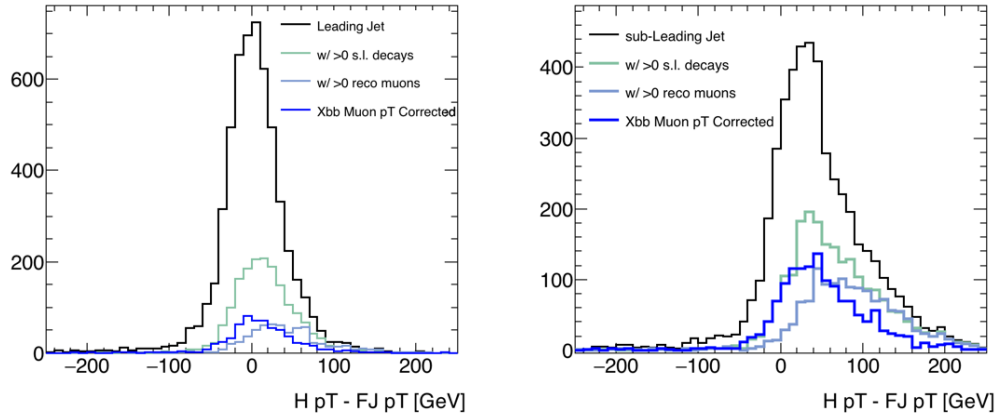
Figure 7.3: The efficiency of the OR of the triggers `HLT_j420_a10t_lcw_jes_35smcINF_L1J100`, `HLT_j420_a10t_lcw_jes_35smcINF_L1SC111` and `HLT_j460_a10t_lcw_jes_L1J100`, that were active in the 2018 data-taking period, as a function of the leading large-R jet p_{T} (left) and mass (right) for data (black) and QCD Monte Carlo (red).

7.3 Muon-in-Jet Correction

The resolution of the reconstructed mass peak of the Higgs boson is fundamental in this analysis. Besides the Higgs signal resonance, there are several background resonances with heavy hadrons capable of semileptonic decays in the mass range of 60–200 GeV. The muon-in-jet correction improves the sharpness of, and hence the separation between, these resonances. The details for the muon-in-jet correction are discussed in Sections 6.6.2 and 6.6.3. Here we will discuss its impact from studies I performed on the Higgs signal and QCD multijet background of the analysis. The studies were completed before the decision was made to add the highest p_{T} bin ($p_{\mathrm{T}} >$ 1 TeV), but the p_{T} range for the (then) last bin was extended to $p_{\mathrm{T}} = 1500$ GeV to observe the effects of the correction at high transverse momentum.

The difference between the p_{T} of the large-R jet truth-matched with the Higgs boson and its recoil jet is shown in Figure 7.4. In this analysis, the Higgs-candidate large-R jet can be the leading (Figure 7.4a) or subleading (Figure 7.4b) jet in the event based on the transverse momentum. The subleading jet has a wider, less sym-

metric, mass distribution because a semileptonic decay often causes a jet to lose enough transverse momentum for it to become subleading in the event.



(a) (b)

Figure 7.4: The difference between the generated (H p_{T}) and reconstructed (FJ p_{T}) Higgs boson p_{T} in ggF signal events for leading (left) and subleading (right) jets. The subset of jets with a semileptonic b-hadron decay (w/$>$ 0 s.l. decays) is shown in green, with a reconstructed muon (w/$>$ 0 reco muons) in light blue, and with the muon-in-jet correction applied (Xbb Muon p_{T} Corrected) in dark blue. The semileptonic decays create a broader distribution with a mean displaced from zero and the muon-in-jet correction partially compensates.

While electrons are not used to improve the jet-mass resolution in this analysis, they may likely be considered in the future. The correction used in this analysis follows the prescription of the $VH(bb)$ analysis [23] and has been implemented in the xAODAnaHelpers package [397].

7.3.1 Impact on the Higgs

It is important to reiterate that the muon-in-jet correction is only applied to the analysis objects requiring the best resolution: the large-R jet mass used in the signal extraction fit and the p_{T} used in the differential cross-section determination. The corrected quantities are *not* used in defining the offline trigger thresholds, nor the p_{T} ordering which defines the leading and subleading signal and validation region

classifications. This choice ensures that semileptonic decays predominantly populate the subleading categories. The mass resolution is therefore narrower in the leading jet category and wider in the subleading, both of which are fit simultaneously with their respective templates.

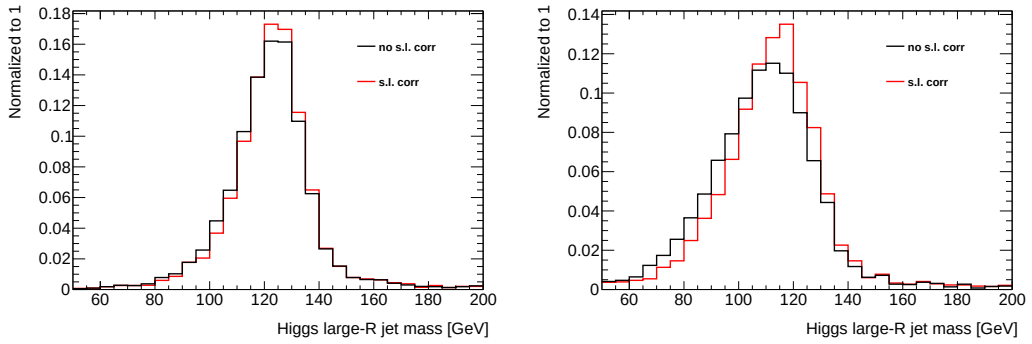
The correction has a significant impact on the inclusive mass resolution, especially in the subleading Higgs-matched large-R jets as seen in Figure 7.5. Table 7.4 shows the fraction of corrected large-R jets, the change in mass peak, and relative reduction in width of the mass distribution for the leading and subleading large-R jets truth-matched to the Higgs boson. The correction fraction in the table refers to the ratio of large-R jets which receive a muon-in-jet correction to the total. The subleading jets are more susceptible to semileptonic decays, and hence receive more corrections. The peak of the leading jet only obtains a minor improvement, shifting up by $\sim 0.5\,\text{GeV}$, and the width is decreased by 5%. The relative improvements on the peak and resolution for the subleading jets are approximately three times greater than the improvements for the leading jets. The effects of the correction were also probed in the analysis p_{T} bins, which can be found in Figure 7.6 for the leading jets and Figure 7.7 for the subleading jets. The plots are separated into inclusive (hadronic and leptonic decays combined), muonic (where at least one b-quark decays leptonically to a muon), and non-muonic (b-quarks all decay hadronically). The results are summarized in Table 7.5, where we see that the greatest improvements on the resolution are obtained at low p_{T}.

Table 7.4: Impact of the muon-in-jet correction on the leading and subleading jet truth-matched with the Higgs boson in ggF events, in terms of the fraction of corrected jets, the peak mass shift and shrinkage of the mass distribution width obtained with a Bukin function fit [404].

Higgs-matched large-R jet	Correction frac.	ΔM [GeV]	$\Delta\sigma/\sigma$
Leading large-R jet	0.13	0.45	-0.046
Subleading large-R jet	0.33	1.52	-0.116

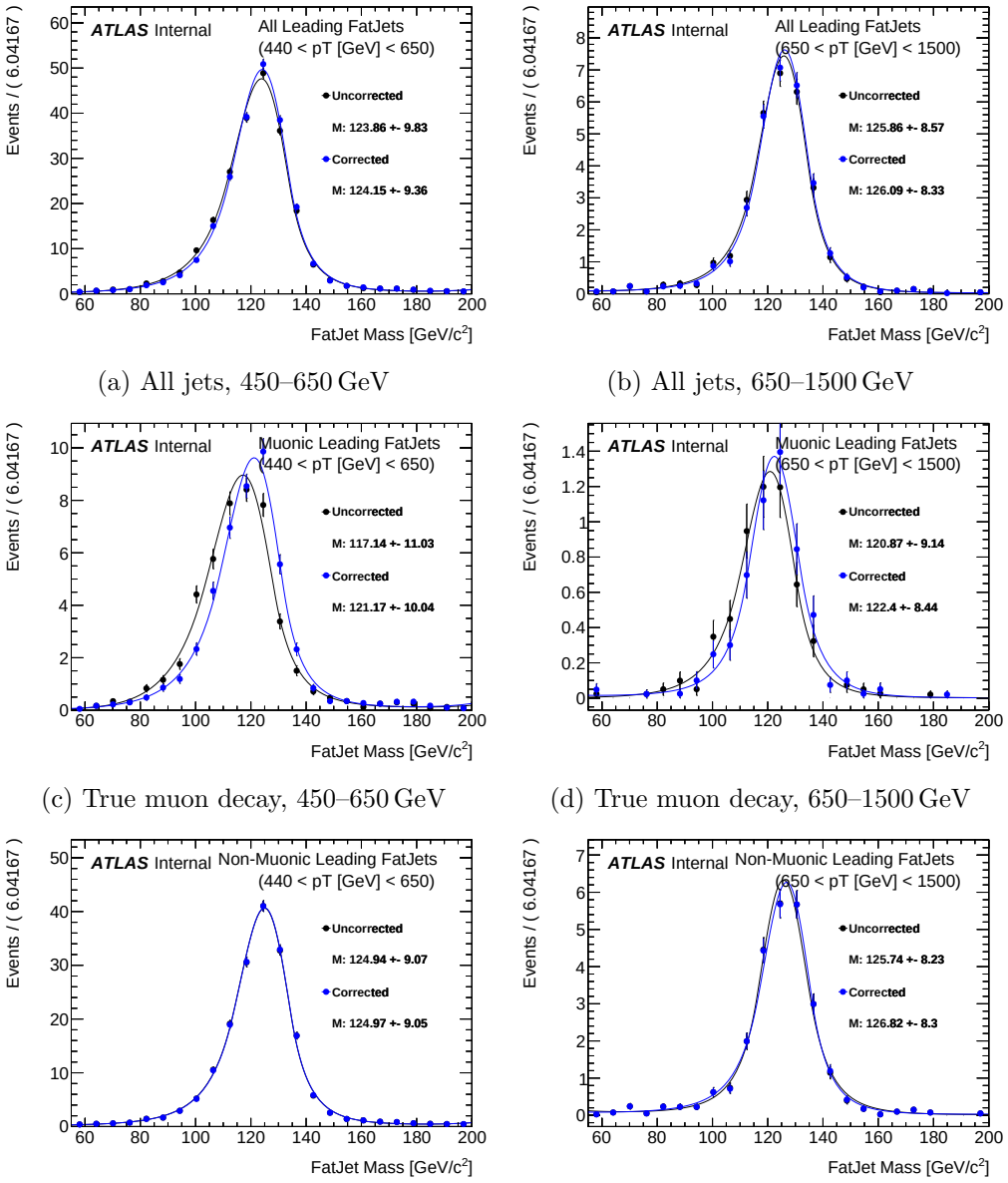
Table 7.5: Impact of the muon-in-jet correction on the leading and subleading jets, truth-matched to the Higgs boson, in simulated ggF events. Three sets of numbers are shown with the first corresponding to the actual analysis improvement for this signal. The second two sets show the improvement when truth information is used to identify Higgs jets with (middle) and without (bottom) a semileptonic b-hadron decay. All masses and widths have been obtained with Bukin function fits [404]. The small, but non-zero changes for the non-muonic category result from corrections made to the large-R jet with muons misidentified as coming from a heavy hadron. This happens rarely. The more significant changes seen in the highest p_{T} bins are the result of low statistics.

jet p_{T} [GeV]	250–450		450–650		650–1500	
jet category	ΔM [GeV]	$\Delta\sigma/\sigma$ [%]	ΔM [GeV]	$\Delta\sigma/\sigma$ [%]	ΔM [GeV]	$\Delta\sigma/\sigma$ [%]
All large-R jets						
leading	–	–	+0.29	-4.78	+0.23	-2.80
subleading	+3.44	-21.4	+1.64	-12.7	+1.33	-10.1
Muonic decay						
leading	–	–	+4.02	-8.98	+1.53	-7.66
subleading	+13.69	-9.05	+6.17	-17.2	+5.15	-14.9
Non-muonic decay						
leading	–	–	+0.03	-0.22	+1.08	0.0
subleading	-0.21	-2.0	0.0	-0.74	+0.19	-0.37



(a) (b)

Figure 7.5: Mass distribution of the leading (left) and subleading (right) large-R jets truth-matched with the Higgs boson, with and without muon-in-jet correction.



(a) All jets, 450–650 GeV

(b) All jets, 650–1500 GeV

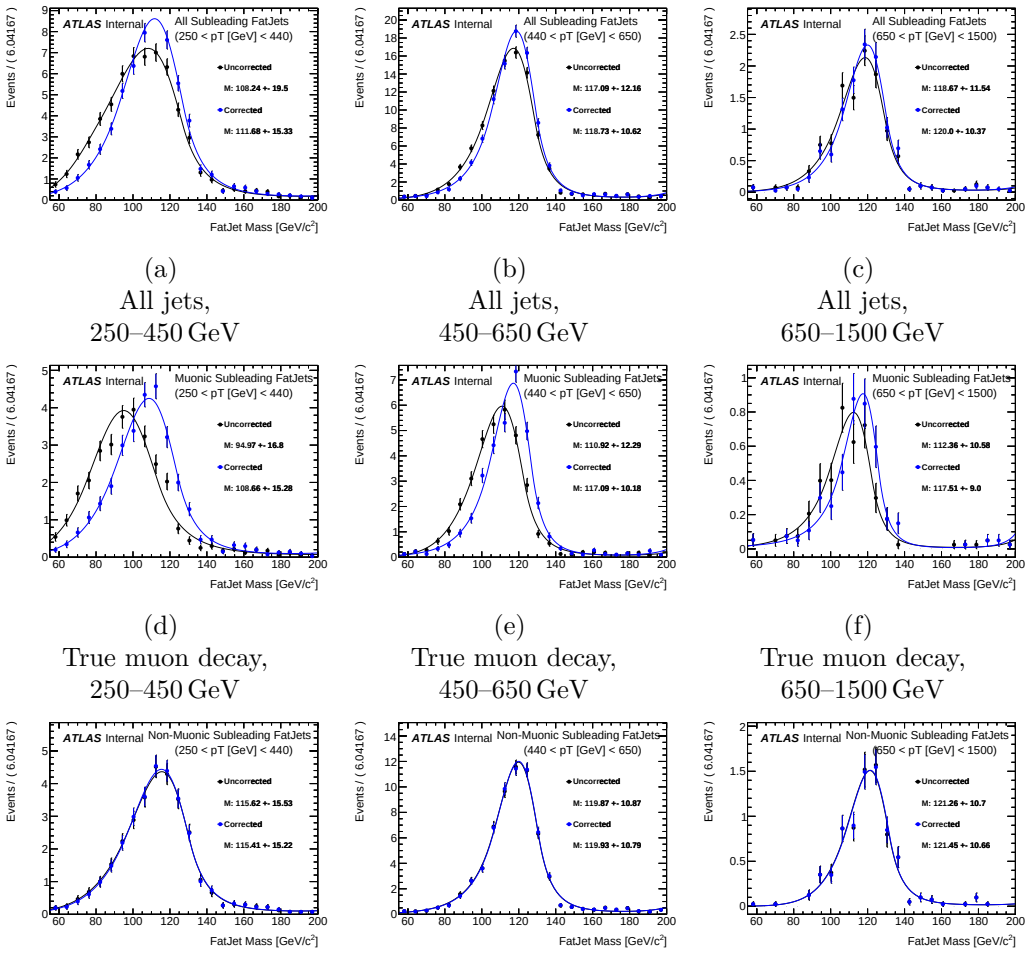
(c) True muon decay, 450–650 GeV

(d) True muon decay, 650–1500 GeV

(e) Non-muonic decays, 450–650 GeV

(f) Non-muonic decays, 650–1500 GeV

Figure 7.6: The impact of the muon-in-jet correction on the leading jet truth-matched to the Higgs boson in simulated ggF events. Three sets of plots are shown where the first corresponds to the actual analysis improvement for this signal. The second two sets show the improvement when truth information is used to identify Higgs jets with (middle) and without (bottom) a semileptonic b-hadron decay.



(a) All jets, 250–450 GeV

(b) All jets, 450–650 GeV

(c) All jets, 650–1500 GeV

(d) True muon decay, 250–450 GeV

(e) True muon decay, 450–650 GeV

(f) True muon decay, 650–1500 GeV

(g) Non-muonic decay, 250–450 GeV

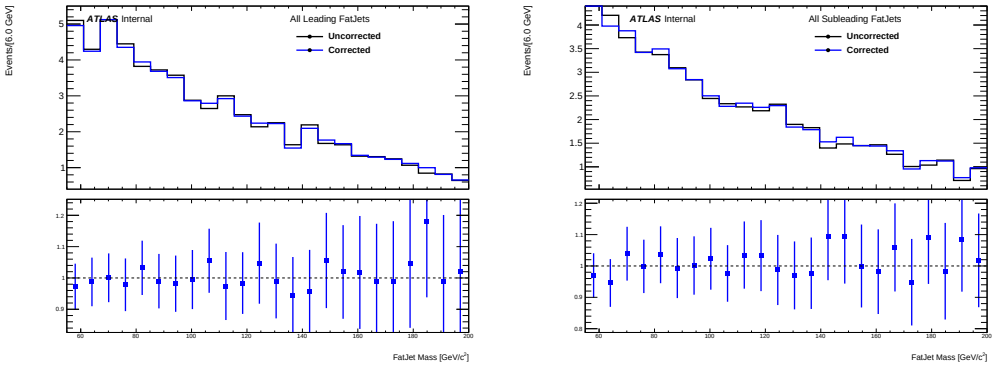
(h) Non-muonic decay, 450–650 GeV

(i) Non-muonic decay, 650–1500 GeV

Figure 7.7: The impact of the muon-in-jet correction on the subleading jet truth-matched to the Higgs boson in simulated ggF events. The inclusive figure (top) demonstrates the improvement to the analysis for this signal. The other separate the events into Higgs jets with (middle) and without (bottom) a semileptonic b-hadron decay.

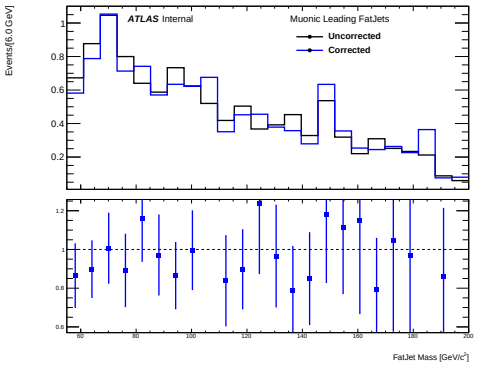
7.3.2 Impact on QCD Mass distribution

Using the JZ4W slice of the PYTHIA QCD dijet MC sample (leading truth-jet p_{T} range [400,800] GeV), the impact on the QCD background was tested. With this slice, only the lowest analysis p_{T} bin (250–450 GeV) was sufficiently populated. Muonic decays were found in 18% and 32% of the leading and subleading jet signal regions, respectively. As shown in Figure 7.8, the QCD mass distribution remains monotonically decreasing within the statistics available. This is an important aspect to point out, as it shows the MIJ correction does not significantly sculpt the QCD background, which could pose a significant challenge if it altered the low-mass region where the fit begins.

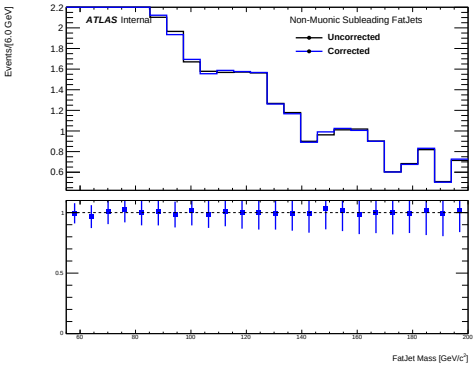


(a) all jets, lead-jet SR

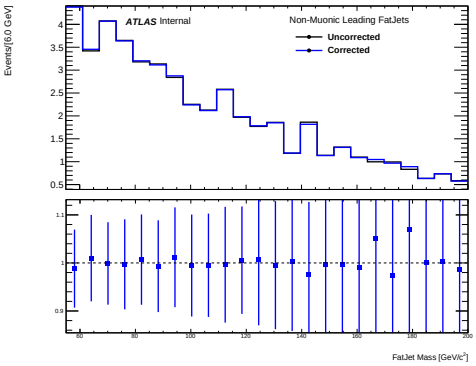
(b) all jets, sublead-jet SR



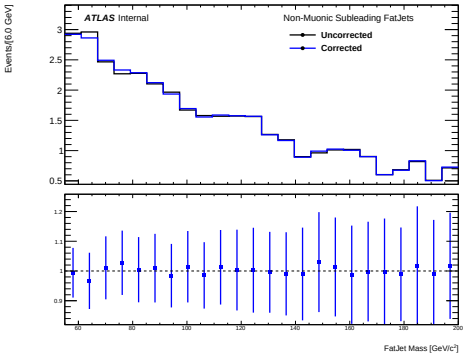
(c) true-muon decay, lead-jet SR



(d) true-muon decay, sublead-jet SR



(e) Non-muonic decay, lead-jet SR



(f) Non-muonic decay, sublead-jet SR

Figure 7.8: The impact of the muon-in-jet correction on the leading and subleading jet SRs in simulated QCD events using the JZ4W slice and populating only the leading analysis p_{T} bin (250–450 GeV). Three sets of plots are show where the top panels corresponds to the actual analysis impact. The second two sets show the improvement with (middle panels) and without (bottom panels) a semileptonic heavy hadron decay.

7.3.3 Muon-in-Jet Induced Mass Effects

In addition to effects from the angular distance between b-jets (Section 7.1.3) and the trigger threshold cut[2], the muon-in-jet correction induces a shaping of the mass spectrum near the mass threshold. This is due to the upward migration of the muon-in-jet-corrected masses that is not compensated for by candidate jets with masses below the threshold. More plainly, events below the mass/p_{T} thresholds before the correction are *not* permitted to migrate into the signal or validation regions after the correction. Figure 7.9 shows a comparison between data and Monte Carlo of the corrected and non-corrected large-R jet mass in the analysis p_{T} bins for the leading and subleading signal regions. The effect (see the ratio in the bottom panel of the plots) is minor for the leading jets due to the smaller rate of MIJ corrections. For the subleading jets, the effect decreases with the jet p_{T} because the mass contribution of muons with a constant p_{T} threshold is a smaller fraction for jets with a higher p_{T}. The results of the studies provided a lower limit on the values for the QCD fit ranges. The b-tagging and MIJ correction induced turn-ons impact the leading and subleading regions differently. Coincidentally, these effects conspire to force the QCD fit to start at nearly the same mass in each p_{T} region, which helps fix the lowest mass considered in the analysis.

[2]Recall that the offline trigger requirements of $p_{\mathrm{T}} > 450\,\mathrm{GeV}$ and $m > 60\,\mathrm{GeV}$ are applied to *uncorrected* jet quantities.

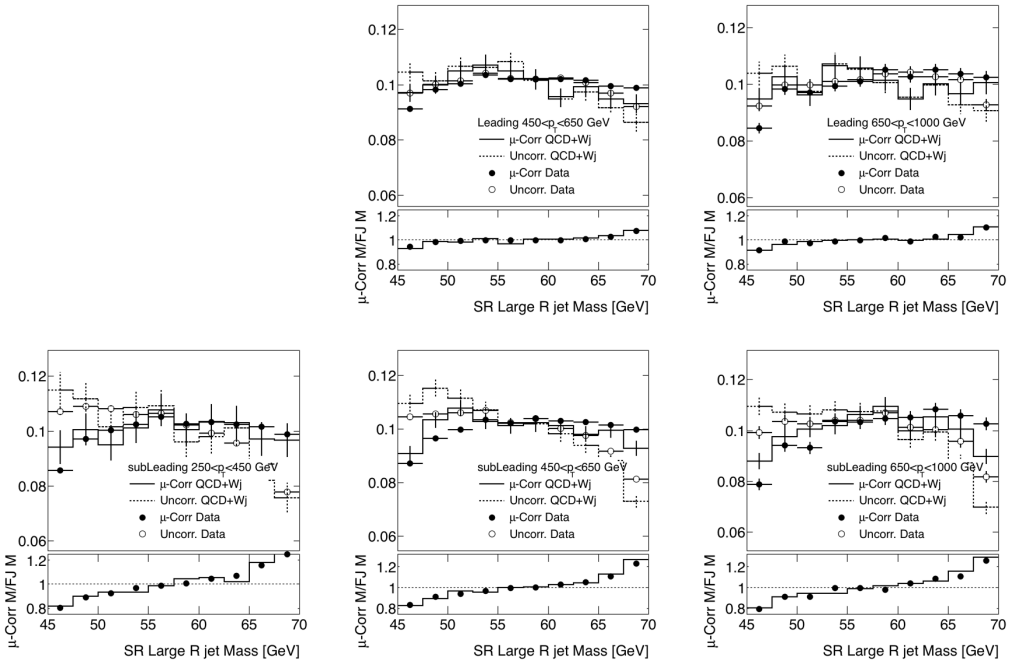


Figure 7.9: Distribution of the candidate large-R jet mass for the leading (upper panels) and subleading (lower panel) categories in data (points with error bars) and simulation (histogram) before and after applying the muon-in-jet mass correction in the three p_{T} intervals. The upper left panel is intentionally empty as the leading signal region does not populate the lowest p_{T} range. The bottom panel in each plot shows the ratio of the muon-corrected large-R jet mass (μ-Corr M) to the uncorrected mass (FJ M).

Chapter 8

Event Selection and Classification

8.1 Event Selection

Events in the Boosted $H \to b\bar{b}$ Analysis are classified into three mutually exclusive regions: a signal region (SR), a $t\bar{t}$ control region ($\mathrm{CR}_{t\bar{t}}$) used to control the normalization of the $t\bar{t}$ background, or validation region (VR), which is used to study the multijet and $V+$jets background models. In all instances, the data are divided into kinematical regions based on the candidate-jet p_{T}. Events in the signal and validation regions are required to pass the following selection criteria:

- pass detector-quality criteria enforced by the Good Runs List (data only),
- satisfy the large-R jet triggers described in Section 7.2,
- have at least one large-R jet with $p_{\mathrm{T}} > 450\,\mathrm{GeV}$ and $m > 60\,\mathrm{GeV}$,
- pass the *looseBad* calorimeter noise cleaning criteria,
- have a second large-R jet with $p_{\mathrm{T}} > 200\,\mathrm{GeV}$ to ensure dijet topology,
- events sorted into a signal leading/subleading (SRL/SRS) or a validation lead-

ing/subleading (VRL/VRS) region. Otherwise it is discarded.

The uniform offline requirement for both the SR and VR in all data-taking years of at least one jet with $p_{\mathrm{T}} > 450\,\mathrm{GeV}$ and $m > 60\,\mathrm{GeV}$ removes the trigger-induced bias in this kinematical regime of the analysis. The selection requirements of the $\mathrm{CR}_{t\bar{t}}$ are discussed in Section 8.3.

8.1.1 Event Classification

Event categorization depends, in part, on the flavor-tagging information for the jet. As mentioned, the jet p_{T} and mass *before* muon-in-jet correction are used to order the jets. The correction is only applied to the final mass used to extract the signal yield. The uncorrected quantities are used to determine which of the analysis jet-p_{T} bins (discussed below) the event populates. Unless specified, all plots of the jet variables for any region are muon-corrected. The category definitions are summarized in Figure 8.1.

First, events are tested against the criteria for the signal regions (SRs), populating the first one satisfied. When considering the b-tagging of the subjets, *only* the leading and subleading (in p_{T}) variable-radius jets are considered. The ordered signal regions and criteria are:

- **Signal Region Leading (SRL) Jet**: the leading-p_{T} jet is a candidate jet with two b-tags,
- **Signal Region Subleading (SRS) Jet**: the subleading-p_{T} jet is a candidate jet with two b-tags (only in events where the leading jet fails the selection as SRL).

The motivation for including the subleading jet in the analysis is laid out in the next section. Events which do not satisfy the signal-region criteria can be used in the two validation regions (VRs). If either candidate jet has a single b-tag, the event is discarded as to not inflate the $t\bar{t}$ contribution in the validation regions. Unlike the signal regions, an event can contribute to *both* the leading- and subleading-jet validation regions:

- **Validation Region Leading (VRL) Jet**: the leading-p_{T} jet is a candidate jet with zero b-tags,
- **Validation Region Subleading (VRS) Jet**: the subleading-p_{T} jet is a candidate jet with zero b-tags.

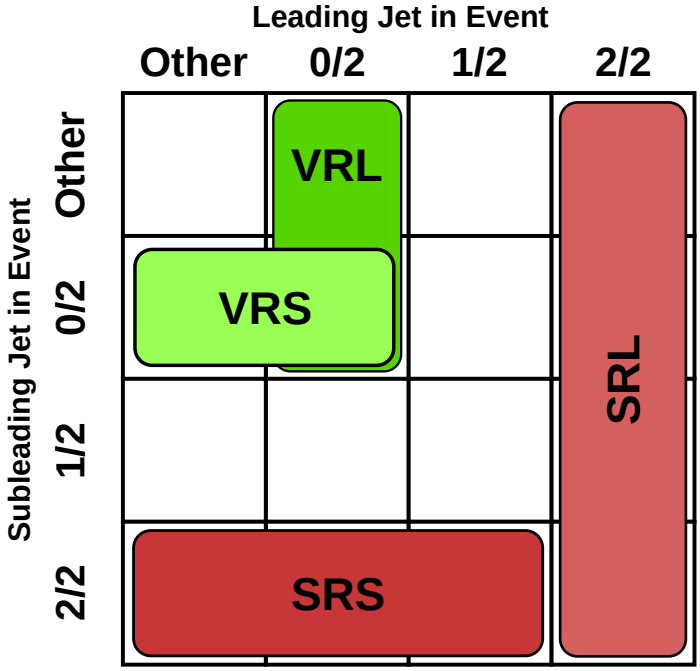


Figure 8.1: Diagram showing the event categorization criteria. The columns (rows) are divided into 4 categories: "other" when the leading (subleading) jet is not a candidate jet, "0/2" when neither of the first two p_{T}-ordered track jets is b-tagged, "1/2" when one of the track jets is b-tagged and "2/2" when both track jets are b-tagged.

All four regions are split into analysis p_{T} bins based on the muon-corrected can-

didate large-R-jet p_{T}: $250-450\,\mathrm{GeV}$ (SRS0/VRS0), $450-650\,\mathrm{GeV}$ (SRx1/VRx1), $650-1000\,\mathrm{GeV}$ (SRx2/VRx2), and $>1000\,\mathrm{GeV}$ (SRL3/VRL3), where x = L for the leading and x = S for the subleading jets. The lowest analysis p_{T} bin is only populated by the subleading regions due to trigger cuts requiring a jet p_{T} above 450 GeV, and the highest p_{T} bin is only populated by the leading regions due to limited statistics in the subleading jet categories. The upper jet-p_{T} cut of 1000 GeV is extended to 3000 GeV when using the inclusive signal regions to extract the signal strength inclusively in p_{T}. Table 8.1 shows a summary of the signal regions and corresponding measurement fiducial volumes. The VRs are used to validate the modeling of the QCD multijet and $V+$jets backgrounds, which will be discussed in more detail in Section 9.4. Events with an isolated muon recoiling against a large-R jet are used to construct a control region for the rate of the $t\bar{t}$ background, which is discussed in Section 8.3.

Table 8.1: A summary of the signal regions and the corresponding measurement fiducial volumes. Left of the break, the large-R jet p_{T} requirements are given for the different analysis configurations. Each SR has an associated $\mathrm{CR}_{t\bar{t}}$ and VR in the same kinematic region. Right of the break, the p_{T}^{H} (Higgs boson transverse momentum) and $|y_H| < 2$ requirements used to determine which signal events are counted in the measurement within the fiducial and differential signal regions are given.

Region	Jet p_{T} [GeV]		p_{T}^{H} [GeV]	$\lvert y_H\rvert$
	SRL	SRS		
Inclusive	>450	>250	–	–
Fiducial	>450	>450	>450	<2
Differential	–	250–450,	300–450,	<2
	450–650,	450–650,	450–650,	
	650–1000,	650–1000	650–1000,	
	>1000	–	>1000	

8.2 Subleading Jet Motivation

In the previous iteration of the analysis [218], only the leading candidate jet was considered. The boosted topology in the ggF channel consists of two back-to-back large-R jets corresponding to the Higgs decay and a recoil jet. At the parton level, the Higgs and QCD recoil have equal transverse momenta; however, at the level of hadronic jets, there are several sources that impact the jet-p_{T} ordering: Final state radiation (FSR), jet mass resolution (JMR) (Sections 6.5.3), out-of-jet-cone activity (Section 6.6.4), and semileptonic decays of heavy quarks (Section 6.6.1) can all result in the Higgs jet becoming subleading in jet p_{T}. While ISR/FSR and jet-resolution effects from the detector are similar between the Higgs and QCD jets, containment and semileptonic decays affect the $H \to b\bar{b}$ system more severely. Higgs production processes other than ggF can also contribute significantly to the subleading jet, in particular the ttH production mode because a hadronically decaying top can satisfy the trigger requirements without the Higgs boson having a high transverse momentum. Due to this possibility, this analysis allows for a subleading Higgs jet in the event. Figure 8.2 shows the relative fractions when the Higgs is leading/subleading in the event, as well as the relative difference in p_{T} between the Higgs and recoil jets.

By using the subleading jets as signal candidates, we ensure a higher signal efficiency while also improving the sensitivity to lower-p_{T} Higgs bosons below the offline trigger threshold. The Higgs p_{T} bin from 250–450 GeV is solely populated by the subleading signal region. The jet mass resolution of the subleading jets is worse than that of the leading, due to issues with containment at lower p_{T}, as well as semileptonic b-decays. The jet ordering in p_{T} is performed *before* the muon-in-

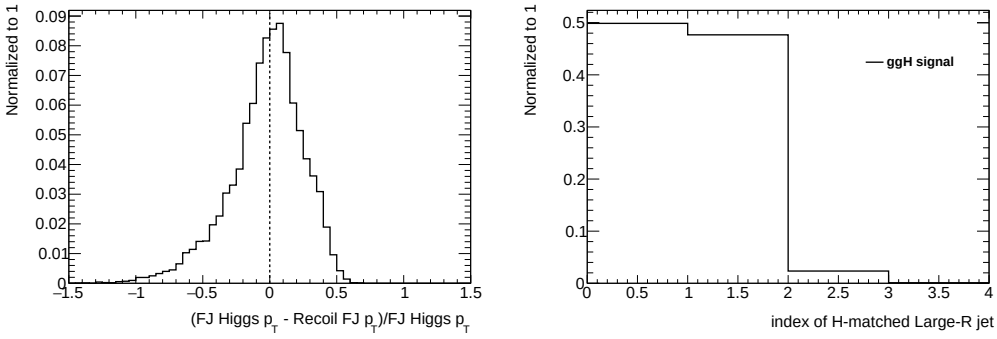


Figure 8.2: (Left) shows the relative difference between the transverse momentum of the large-R jet matched with the Higgs (FJ Higgs p_{T}) and the recoil jet (Recoil FJ p_{T}) in ggF MC. (Right) shows the index of the Higgs (0-leading, 1-subleading, etc) matched large-R jet, ordered in p_{T}, shown for the ggF signal sample.

jet correction (Sections 6.6.1 and 7.3) to maintain this separation in the resolution between signal categories. When considering jet resolution or scale systematics, and the correlation among categories, it is convenient to have one region more impacted by resolution effects when an applied correction targets one of those effects.

8.3 $t\bar{t}$ Control Region

It has been shown in previous analyses [405] that the Powheg + Pythia 8 $t\bar{t}$ MC over-predicts the cross section, especially in a boosted topology such as ours, as seen in Figure 8.3. To constrain the normalization of $t\bar{t}$ in the signal regions (inclusively and in p_{T} bins), a $t\bar{t}$ control region was created to identify boosted semileptonic $t\bar{t}$ events and compare MC with data. I designed the $t\bar{t}$ control region selection and implemented it into the analysis framework. The studies were completed before the decision was made to extend the differential measurement with large-R jets with $p_{\mathrm{T}} > 1000\,\mathrm{GeV}$, but the normalization value in each region is presented in Chapter 11. The $\mathrm{CR}_{t\bar{t}}$ is defined as having a boosted leptonically decaying top quark in one hemisphere and a boosted hadronically decaying top in the other. Using a vari-

ation of the tag-and-probe approach often used for measuring efficiencies, objects on the tag side are identified as a semileptonic decay of a top quark, and objects on the probe side are identified as the hadronic decay of a top quark. This allows us to select an unbiased sample of large-R jets containing the hadronic decay products of top quarks. Events are triggered using the `HLT_mu50` trigger. To match the effects of the jet trigger in the first p_{T} bin of the subleading jet category, a p_{T} reweighting scheme is implemented.

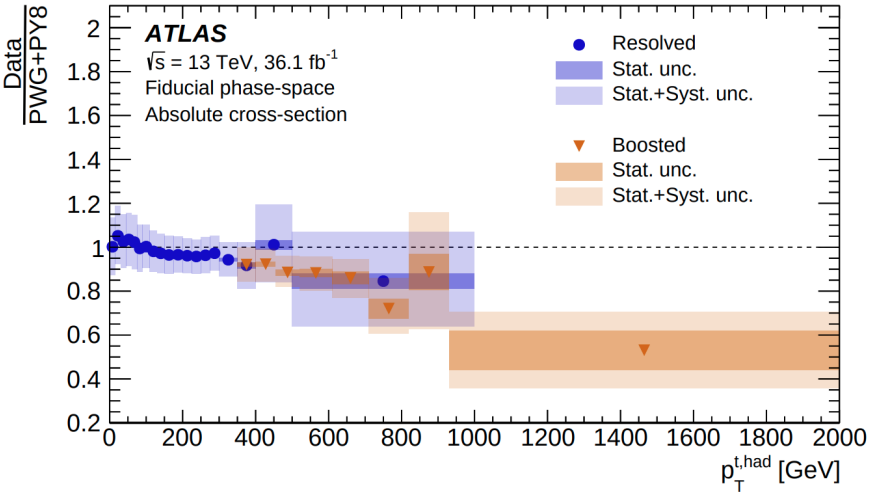


Figure 8.3: The ratios of the measured fiducial phase-space absolute differential cross sections in $t\bar{t}$ production to the predictions obtained with the Powheg + Pythia 8 MC generator in the resolved and boosted topologies as a function of the transverse momentum of the hadronic top quark. The bands indicate the statistical and total uncertainties of the data in each bin [405].

8.3.1 CR$_{t\bar{t}}$ Selection Criteria

First, the leading muon in the event is identified as having $p_{\mathrm{T}} > 52.5\,\mathrm{GeV}$ (5% above the trigger threshold for full efficiency). This muon must be of medium quality and isolated with the loose WP (see Section 6.3). A tag-jet, which is the large-R jet with a semileptonic decay, is identified if one and only one of its two leading (in p_{T}) VR jets is b-tagged at the MV2c10 77% WP and satisfies

$0.04 + 10/p_{\mathrm{T}}^{\mathrm{muon}} < \Delta R_{\mathrm{muon,VR\ jet}} < 1.5$. These VR jets must have $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and pass overlap cleaning (Section 7.1.2). The ΔR cut reduces the $V+\mathrm{jets}$ and VV contamination. The probe-jet, or large-R jet produced by a hadronically decaying top quark, is identified as being in the opposite hemisphere of the tag-jet using an angular requirement of $\Delta\phi_{\mathrm{tag,probe}} > 2\pi/3$. The probe-jet is further required to have at least 3 VR jets with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ with exactly one of the leading three b-tagged. The $\Delta\phi$ requirement reduces the contamination from QCD and $V+\mathrm{jets}$, and requiring 3 VR jets in the large-R jet ensures a fully contained top-quark decay. Considering up to 2 (3) VR jets in the tag- (probe-) jet boosts the acceptance by 7.5% (10%) without altering the shape of the mass distribution. To make the $\mathrm{CR}_{t\bar{t}}$ events more compatible with those in the signal region, the mass range $140 < m_{\mathrm{J}} < 200\,\mathrm{GeV}$ is used in the fit. The selection is visually represented in Figure 8.4.

The $\mathrm{CR}_{t\bar{t}}$ is binned in the p_{T} of the probe-jet, following the same binning used in the signal region. Reasonable modeling is seen, and the $t\bar{t}$ purity is greater than 95% in all three p_{T} bins (Figures 8.7 and 8.8).

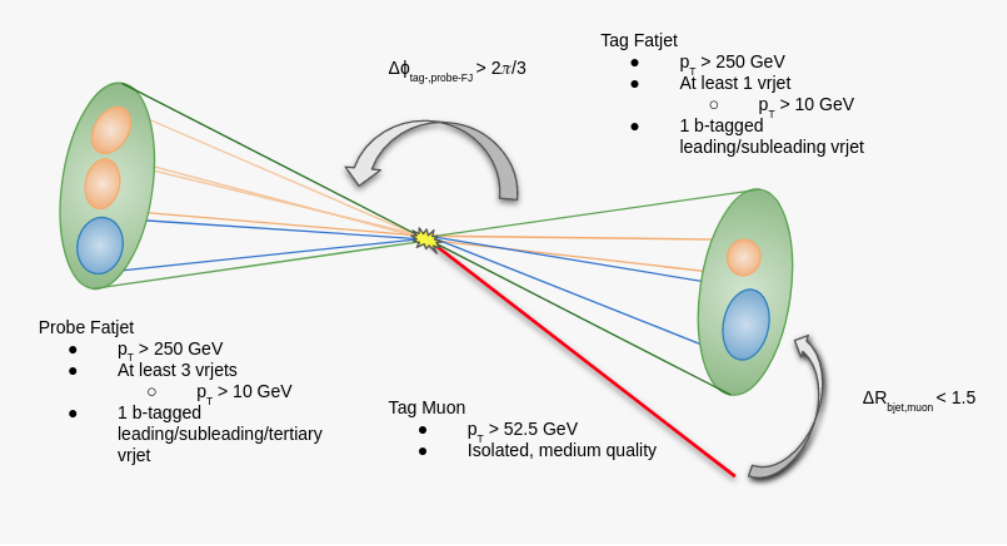
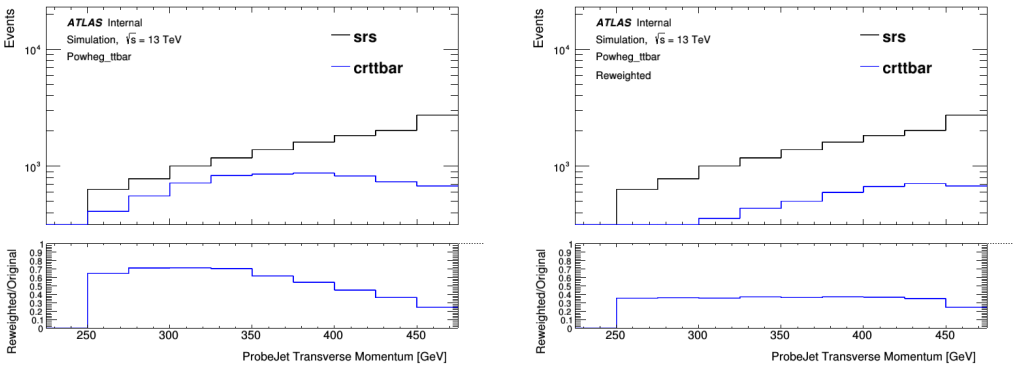


Figure 8.4: Graphic displaying the topology of $\mathrm{CR}_{t\bar{t}}$ events.

8.3.2 p_{T} Reweighting of the $\mathrm{CR}_{t\bar{t}}$

The SR and $\mathrm{CR}_{t\bar{t}}$ rely on different triggers for their event selections. The $\mathrm{CR}_{t\bar{t}}$ triggers on a muon, whereas the SR triggers on the mass and p_{T} of the large-R jet. This difference in triggering sculpts the low end of the p_{T} distribution in different ways. This is noticeable in the first p_{T} bin ($250 < p_{\mathrm{T}}^{\mathrm{jet}} < 450\,\mathrm{GeV}$), where the p_{T} of the SRS is rising over the full range, but the $\mathrm{CR}_{t\bar{t}}$ peaks at $\sim$ 350 GeV and begins to fall, as can be seen in Figure 8.5a.



(a) (b)

Figure 8.5: Comparison between p_{T} distributions between the unweighted (a) and reweighted (b) POWHEG $t\bar{t}$ MC samples in the SRS and $\mathrm{CR}_{t\bar{t}}$. After reweighting, the ratio between SRS and $\mathrm{CR}_{t\bar{t}}$ is flat in p_{T}.

To correct for the mismatch in shape, the $\mathrm{CR}_{t\bar{t}}$ p_{T} spectrum needed to be reweighted to the lowest p_{T} bin so that its shape matches that of the SRS. To accomplish this, a second-order polynomial function was fit to the normalized ratio of the SRS-to-$\mathrm{CR}_{t\bar{t}}$ p_{T} spectrum such that the weight does not exceed 1. The reweighting was performed in similar, but different, ways between data and MC. For MC, the weight is multiplied by the value of the polynomial function evaluated for a particular value of the transverse momentum. In data, this would not be acceptable as it would falsely increase the statistical precision in the fit after reweighting. To solve this problem, a "random toss" is performed in data using the acceptance-rejection

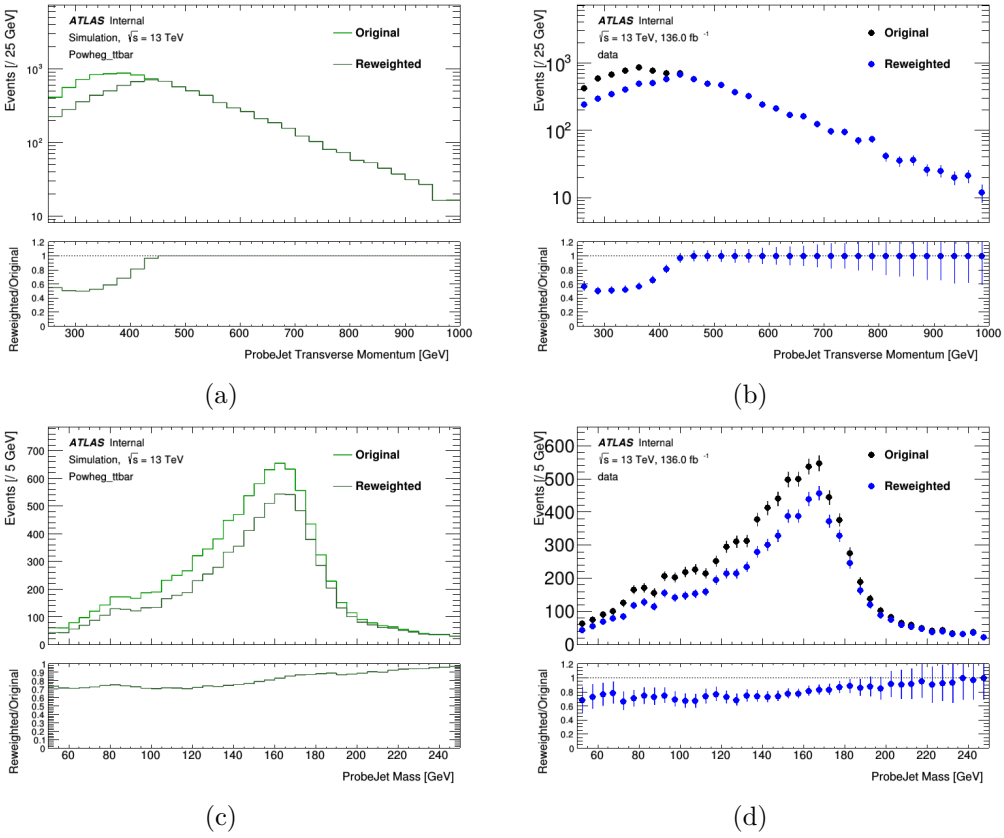


Figure 8.6: Comparison between the original and reweighted p_{T} (top) and mass (bottom) distributions in Powheg + Pythia 8 $t\bar{t}$ MC (left) and data (right).

method, with the polynomial used as the probability distribution.

The results of the reweighting in MC and random toss in data for the p_{T} and mass distributions are shown in Figure 8.6. The effect in data is consistent with that seen in MC. Approximately 1/3 of the data events from the lowest p_{T} bin are "tossed."

8.3.3 CR$_{t\bar{t}}$ Data/MC Comparisons

The reweighting procedure described in the previous section provides the inclusive distributions shown in Figure 8.7. For distributions in large-R-jet p_{T} bins, see Figure 8.8a for $250 < p_{\mathrm{T}}^{\text{probe-jet}} < 450\,\mathrm{GeV}$, Figure 8.8b for $450 < p_{\mathrm{T}}^{\text{probe-jet}} < 650\,\mathrm{GeV}$, and Figure 8.8c for $650 < p_{\mathrm{T}}^{\text{probe-jet}} < 1000\,\mathrm{GeV}$. The simulation is scaled

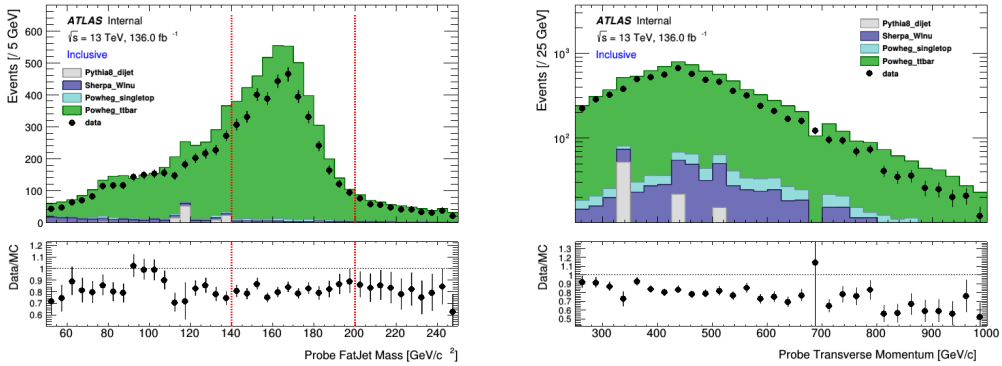
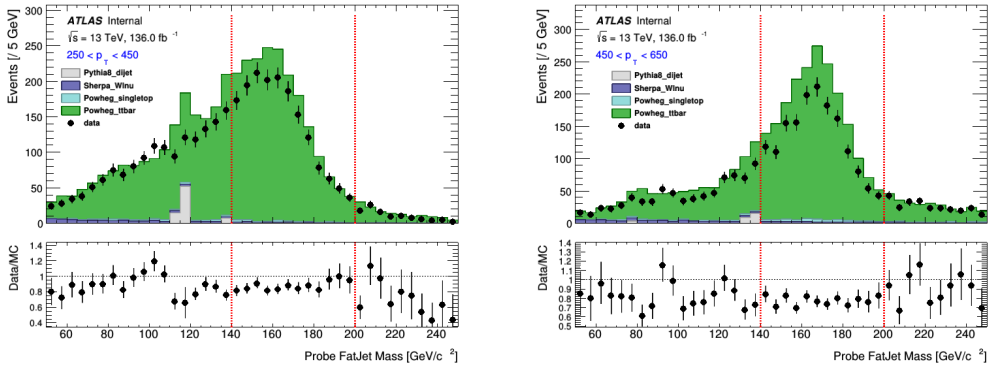
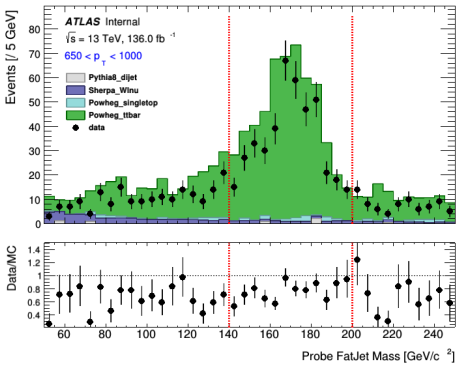


Figure 8.7: Pre-fit Inclusive distributions of jet mass (left) and p_{T} (right) for the large-R jets (Probe FatJet) in the CR$_{t\bar{t}}$ region after the p_{T} reweighting procedure. The large dijet contribution near 120 GeV is due to a single event with a large weight.



(a) $250 < p_{\mathrm{T}}^{\text{probe-jet}} < 450\,\mathrm{GeV}$ (b) $450 < p_{\mathrm{T}}^{\text{probe-jet}} < 650\,\mathrm{GeV}$



(c) $650 < p_{\mathrm{T}}^{\text{probe-jet}} < 1000\,\mathrm{GeV}$

Figure 8.8: Jet mass distributions for events in different p_{T} ranges of the CR$_{t\bar{t}}$ selection.

to agree with data, and the normalization factors are given in Table 8.2, which agree with those in [405] (Figure 8.3) and are only given as a reference to the precision of the CR. The normalization factors extracted from fits to data will be presented

in Chapter 11. The yields before any normalization factors from data were applied are found in Table 8.3. The corresponding post-fit plots are shown in Figure 9.6 of Section 9.4.2.

Table 8.2: A simple scale factor for the $t\bar{t}$ process determined in the $\mathrm{CR}_{t\bar{t}}$ for the large-R-jet mass range 140 to 200 GeV. The quoted uncertainties are statistical only. Note, these estimates were made before the decision to use jets with $p_{\mathrm{T}} > 1000\,\mathrm{GeV}$ was considered.

Range [GeV]	Scale Factor	Statistical Error
$250 < p_{\mathrm{T}}^{\text{probe-jet}} < 450$	0.86	0.02
$450 < p_{\mathrm{T}}^{\text{probe-jet}} < 650$	0.77	0.03
$650 < p_{\mathrm{T}}^{\text{probe-jet}} < 1000$	0.77	0.05

Table 8.3: The yields for all simulated events and data in $\mathrm{CR}_{t\bar{t}}$ in probe-jet p_{T} bins. The scale factors discussed above are not applied here.

Process	$p_{\mathrm{T}} > 250\,\mathrm{GeV}$	$250 < p_{\mathrm{T}} < 450\,\mathrm{GeV}$	$450 < p_{\mathrm{T}} < 650\,\mathrm{GeV}$	$650 < p_{\mathrm{T}} < 1000\,\mathrm{GeV}$
$t\bar{t}$	5354.93 ± 73.18	2809.28 ± 53.00	1981.07 ± 44.51	520.79 ± 22.82
single t	82.26 ± 9.07	29.98 ± 5.48	38.46 ± 6.20	11.27 ± 3.36
$W+$jets	58.80 ± 7.67	17.64 ± 4.20	23.39 ± 4.84	12.86 ± 3.59
dijet	3.05 ± 1.75	0.0 ± 0.0	0.0 ± 0.0	2.96 ± 1.72
data	4492.00 ± 67.02	2458.00 ± 49.58	1580.00 ± 39.75	421.00 ± 20.52

Chapter 9

Event Samples and Process Modeling

9.1 Data

The dataset consists of Run 2 data gathered by ATLAS during 2015 – 2018. After a Good Runs List (GRL) to meet data quality requirements [406] is used to remove collisions with bad detector performance, the remaining total integrated luminosities are $36.2\,\text{fb}^{-1}$, $41.0\,\text{fb}^{-1}$, and $58.5\,\text{fb}^{-1}$ for jet triggers used in the 2015 + 2016, 2017, and 2018 data-taking periods, respectively, yielding a total integrated luminosity of $136\,\text{fb}^{-1}$. The GRL used can be found in Appendix C.1.

9.2 Simulated Signal and Background Samples

The resonant backgrounds ($W+$jets, $Z+$jets, and top-quark production) and signal (Higgs: ggF, VBF, VH, and ttH) are modeled with Monte Carlo simulation [218]. The mass distributions of the data and MC form template histograms, which serve as the inputs to the global likelihood fit. The QCD multijet background is fit using a data-driven method, which we will discuss in Section 9.4.3. Pythia 8 is

used for parton showing and hadronization for all simulated samples except $V+$jets, which uses SHERPA 2.2.8.

Higgs Boson Production

Gluon-gluon fusion (ggF) is simulated at next-to-leading order (NLO) accuracy in QCD with finite mass effects using the HJ-MiNLO prescription [407–409] with POWHEG [410–412] as laid out in Reference [413]. The VBF and ttH processes are calculated at NLO accuracy in QCD and $gg \to VH$ production at LO accuracy using POWHEG-BOX V2 [410–412, 414, 415]. The $qq \to VH$ process uses the POWHEG-BOX V2 program with the improved MiNLO calculation [416] and the GOSAM [417] program for NLO accuracy in QCD. Electroweak corrections are applied using HAWK [418, 419] as a function of the generated Higgs boson transverse momentum p_{T}^{H} for VBF, VH, and ttH production. The Higgs production cross sections are compatible with those presented in Reference [420] with the exception of ttH, which receives an overall scale factor of 0.71 to recover the difference[1]. The Higgs boson branching fractions are calculated with HDECAY [421–423] and PROPHECY4F [424–426].

$V+$jets Production

$V+$jets production with hadronically decaying bosons is simulated with SHERPA 2.2.8 to NLO QCD accuracy for one additional parton and LO QCD for up to four additional partons. The approximate NLO EW corrections are applied as a function of the simulated vector boson transverse momentum p_{T}^{V}. The corrections

[1] This scale factor results from the higher order corrections in the state-of-the-art calculations [104] not included in the generator (POWHEG) used in the analysis.

have an appreciable effect on the differential production cross section, decreasing the predicted yield by $\sim 10\%$ at $p_{\mathrm{T}}^{V} \approx 500\,\mathrm{GeV}$ and $\sim 20\%$ above 1 TeV. Calculations of next-to-next-to-leading order (NNLO) QCD corrections to $V+\mathrm{jets}$ productions are available [104]. The NNLOJET Group [427] performed the calculation for $\sqrt{s} = 8\,\mathrm{TeV}$ [428, 429], and has provided custom corrections for the analysis fiducial region at $\sqrt{s} = 13\,\mathrm{TeV}$ [430] as a function of p_{T}^{V}. They vary from 1.013 to 1.081 and are applied as a multiplicative factor on top of the NLO EW corrections.

Top-Quark Production

The production of top-quark pairs, associated production of top quarks with W bosons Wt, and single-top t- and s-channel production are modeled with the Powheg-Box V2 [410–412, 431–434] generator at NLO in QCD. The diagram removal scheme [435] is used in Wt events to account for interference and overlap with $t\bar{t}$ production.

QCD Multijet Production

The jet-mass distribution of the non-resonant multijet events is modeled with an analytic function discussed in Section 9.4.3. Simulated events used to study the multijet model are generated using Pythia 8.230 [436]. This program includes both LO matrix elements for dijet production and a p_{T}-ordered parton shower.

Pileup

Simulated collision particles are processed through the ATLAS detector simulation [437] based on Geant 4 [438]. Pileup is modeled by overlaying simulated inelastic pp events generated with Pythia 8.186 [436] using the NNPDF2.3LO set

Table 9.1: The generators used for the simulation of the signal and background processes. Matrix element, parton shower and underlying event are abbreviated as ME, PS, and UE respectively.

Process	ME generator	ME PDF	PS and Hadronization	UE model tune	Cross-section order
Higgs Boson					
$gg \to H \to b\bar{b}$	Powheg-Box v2*[410–412] + MiNLO [407–409]	NNPDF3.0NNLO [447]	Pythia 8.212 [436]	AZNLO [448]	NLO(QCD) + LO(EW)
$qq \to H \to q'q'b\bar{b}$	Powheg-Box v2 [410–412, 414]	NNPDF3.0NLO [447]	Pythia 8.230	AZNLO	NLO(QCD) + NLO(EW)•
$qq \to WH$ $\to qq'b\bar{b}$ $\to \ell\nu b\bar{b}$	Powheg-Box v2 + GoSam [417] + MiNLO [416]	NNPDF3.0NLO	Pythia 8.240 Pythia 8.212	AZNLO	NNLO(QCD) + NLO(EW)•
$qq \to ZH$ $\to q\bar{q}b\bar{b}$ $\to \nu\nu b\bar{b}$ $\to \ell\ell b\bar{b}$	Powheg-Box v2 + GoSam+ MiNLO	NNPDF3.0NLO	Pythia 8.240 Pythia 8.212	AZNLO	NNLO(QCD) + NLO(EW)•
$gg \to ZH$ $\to q\bar{q}b\bar{b}$ $\to \nu\nu b\bar{b}$ $\to \ell\ell b\bar{b}$	Powheg-Box v2	NNPDF3.0NLO	Pythia 8.240 Pythia 8.212	AZNLO	LO + NLL(QCD)
$gg \to ttH$ $tt \to$ all $H \to$ all	Powheg-Box v2 [415]	NNPDF3.0NLO	Pythia 8.230	AZNLO	NLO(QCD) + NLO(EW)°
Vector boson + jets					
$W \to qq$ $Z \to qq$	Sherpa 2.2.8 [449–451]	NNPDF3.0NNLO	Sherpa 2.2.8 [452, 453]	Default	NNLO(QCD)†[428, 429, 454] approx NLO(EW) [455, 456]
Top quark, mass set to 172.5 GeV					
$t\bar{t} \to$ all	Powheg-Box v2 [410–412, 431]	NNPDF3.0NLO	Pythia 8.230	A14 [457]	NNLO+NNLL [458]
tW	Powheg-Box v2 [410–412, 432]	NNPDF3.0NLO	Pythia 8.230	A14	NLO
t t-channel	Powheg-Box v2 [410–412, 433]	NNPDF3.0NLO	Pythia 8.230	A14	NLO
t s-channel	Powheg-Box v2 [410–412, 434]	NNPDF3.0NLO	Pythia 8.230	A14	NLO
QCD Multijet					
Dijets	Pythia 8.230	NNPDF2.3LO [439]	Pythia 8.230	A14	LO

* Powheg was configured to output events with Born k_T above 200 GeV using the *bornktmin* setting.
• Corrections for NLO EW effects computed with Hawk [418, 419] are applied as a function of the generated Higgs boson transverse momentum.
° Corrections for NLO EW effects computed with Sherpa+OpenLoops [450, 459, 460] are applied as a function of the generated Higgs boson transverse momentum and were provided by Ref. [420].
† Sherpa provides 1 additional parton at NLO accuracy and up to 4 additional partons at LO in QCD and custom NNLO QCD corrections were provided by the NNLOJET group.

of parton distribution functions (PDFs) [439] and A3 tune [440] over the original hard-scatter event. For Higgs boson and top-quark production, the EvtGen v1.2.0 program [441] models the decays of bottom and charm hadrons.

Sample Summary

For each sample, Table 9.1 summarizes the MC generators, PDFs, and underlying event tunes used, as well as the order of perturbative QCD computations and EW corrections obtained for the cross section. For additional information, see Reference [442] for V + jets events, References [443–445] for $t\bar{t}$ events, and Reference [446] for multijet events. The systematic uncertainties are discussed in Section 10.2.

9.3 Signal Modeling

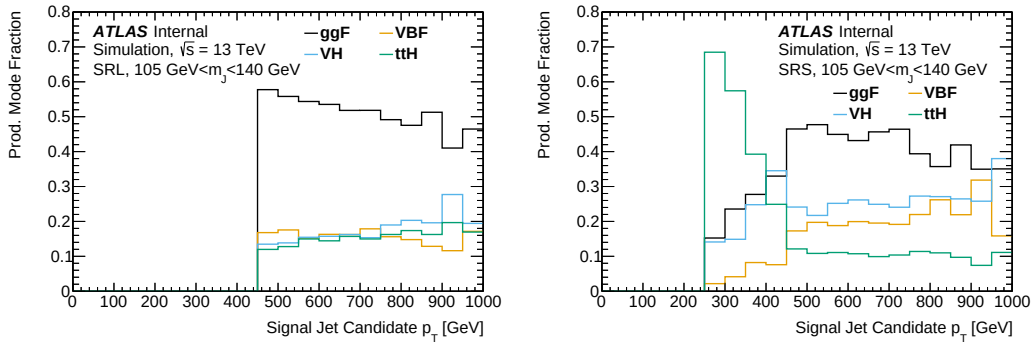
As discussed in Section 3.1, there are four primary production modes for the Higgs boson: ggF, VH, ttH, and VBF. In this section, we will discuss the signal production modes and their theoretical corrections to determine their expected signal rates.

9.3.1 Higgs Production Modes

Because the analysis selection does not investigate the properties of the recoil system, it is inclusive in terms of the Higgs production modes, with varying contributions from each mode as a function of the transverse momentum. The relative contributions of the four main production modes as a function of p_{T} can be found in Figure 9.1, where only the events in the mass window $105 < m_{\mathrm{J}} < 140\,\mathrm{GeV}$ are considered. A summary of this breakdown can be found in Table 9.2.

In the SRS, ttH comprises nearly half of the events in the $250 < p_{\mathrm{T}}^{\mathrm{jet}} < 450\,\mathrm{GeV}$ bin in the Higgs mass window. ttH events form a three-body system, where a hadronic top can satisfy the trigger mass and p_{T} requirements without providing a significant boost to the Higgs boson. This results in a large fraction of ttH events with low Higgs p_{T} in the SRS.

The mass distributions, subdivided into the four Higgs production modes and broken down into the four inclusive analysis regions (SRL, SRS, VRL, VRS), are shown in Figure 9.2.



(a) (b)

Figure 9.1: The fractional contribution of each production mode as a function of the signal candidate jet p_{T} to the inclusive leading (left) and subleading (right) signal regions. Only events contributing to the Higgs peak, defined by $105 < m_{\mathrm{J}} < 140\,\mathrm{GeV}$, are considered.

Table 9.2: The fractional contribution of each production mode to a given analysis bin in the Higgs mass window ($105 < m_{\mathrm{J}} < 140\,\mathrm{GeV}$). The fraction is given with respect to the total yield in the analysis bin in question.

	Jet p_{T} Range [GeV]				
Process	Inclusive	250–450	450–650	650–1000	> 1000
			SRL		
ggF	0.55	–	0.56	0.50	0.39
VBF	0.17	–	0.17	0.16	0.17
VH	0.14	–	0.14	0.18	0.25
ttH	0.14	–	0.13	0.16	0.19
			SRS		
ggF	0.40	0.28	0.46	0.43	–
VBF	0.15	0.07	0.19	0.21	–
VH	0.25	0.26	0.24	0.26	–
ttH	0.21	0.39	0.11	0.10	–

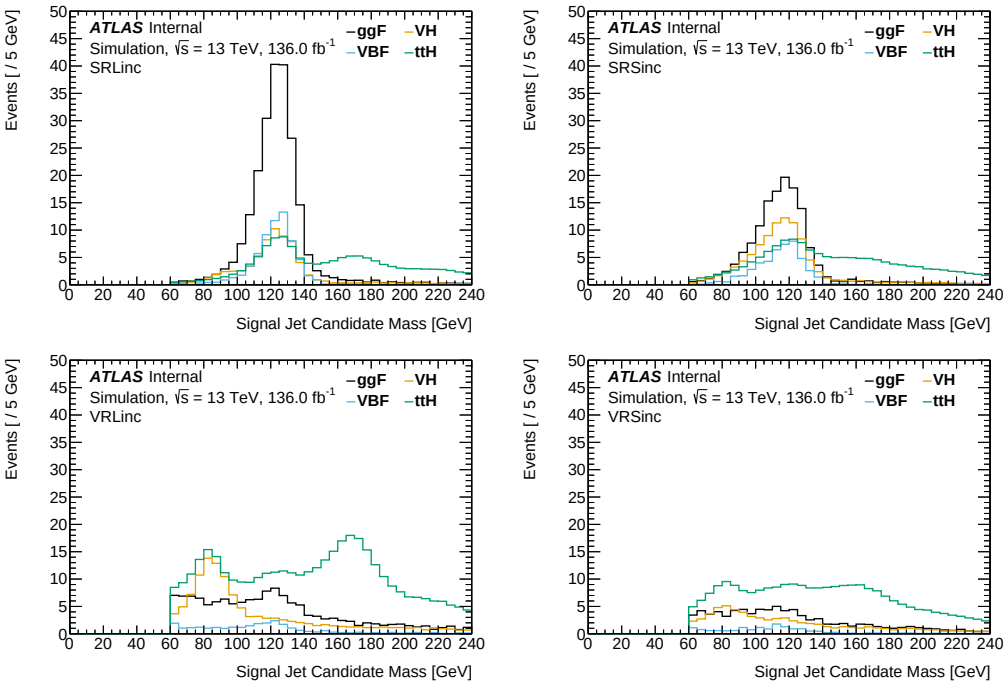


Figure 9.2: Breakdown of the Higgs contribution to the signal jet candidate mass peak into the different Higgs production modes in the inclusive leading (left) and subleading (right) signal (top) and validation (bottom) regions.

9.4 Background Modeling

The dominant background to this analysis is by far the monotonically falling non-resonant QCD spectrum, with significant additional contributions from $V+$jets and $t\bar{t}$ production. The Higgs signal sits atop the enormous QCD background, which dominates in all regions, and is situated between the Z and $t\bar{t}$ contributions. Figure 9.3 shows the background shapes and the nominal normalization from MC for the inclusive signal regions. In the signal regions, the $V+$jets contribution is dominated by $Z+$jets ($Z \to b\bar{b}$), and the ratio of Z to W events varies from 5:1 to 10:1. The $t\bar{t}$ background is most prevalent at high mass, but has large low-mass tails due to uncontained top quarks. Comparisons of the Higgs mass distribution to only the minor backgrounds[2] can be found in Figure C.2 of Appendix C.3.

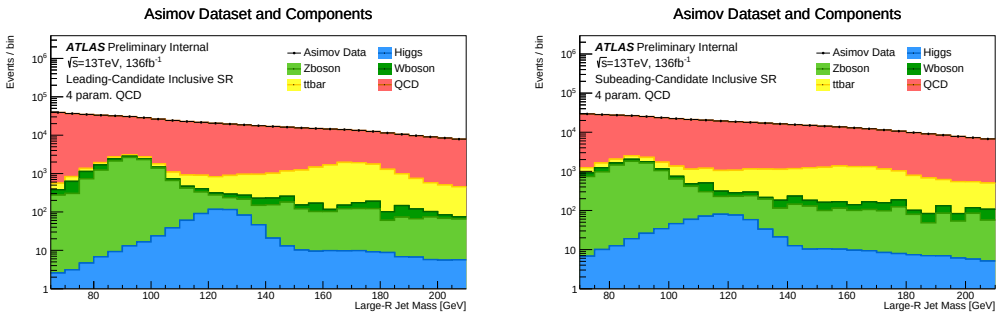


Figure 9.3: Comparison of the large-R jet mass distributions of the expected nominal MC estimates (the Asimov dataset) for the major backgrounds in the inclusive SRL (left) and SRS (right) regions. The QCD background dominates, but the Z, $t\bar{t}$, and W resonances all have significant contributions with either high- or low-mass tails which overlap in the Higgs mass window. The Asimov dataset[3]is the most likely dataset distribution, constructed from the medians of the various contributions.

[2] All backgrounds except the QCD multijet background.

[3] Inspired by Isaac Asimov's short story *Franchise* [461], where a representative individual is selected to represent the entire population in the presidential election. In experiments, the Asimov dataset is synonymous with the representative dataset derived from simulation [462].

9.4.1 $V+$jets

The $W+$jets and $Z+$jets processes, collectively referred to as $V+$jets, populate the large-R jet mass region between approximately 70 to 140 GeV in the signal region, with high-mass tails covering the Higgs mass region. Each process is modeled with large-R jet mass from Monte Carlo, and each receives a separate normalization factor in fit to signal and control regions. Over this mass range, the $W+$jets process adds to the QCD background near the start of the mass fit and the upper tail of the $Z+$jets process intrudes into the signal region. The decay structure and relative experimental resolution of the $Z+$jets mass peak is similar to that of the Higgs boson. Additionally, the results of the QCD multijet background fit depend on the modeling of the simulated backgrounds, in particular $V+$jets. Therefore, the modeling of the $V+$jets backgrounds is of great importance and offers a unique opportunity to evaluate the experimental performance of the analysis. The prominent decay mode for Z boson decays in the SR is, by construction, $Z \to b\bar{b}$. Figure 9.4 shows the fraction of $Z+$jets events coming from $Z \to b\bar{b}$. For $W+$jets, an appreciable fraction of events come not from W boson decays, but from b-tagging the recoil jet. Events from $V+$jets outnumber Higgs bosons by more than a factor of 20, so experimental effects which are more challenging to discern in the statistically limited H production measurements will be more evident in the Z observation. This forces us to need a well-understood Z measurement to produce a reliable measurement of the Higgs boson.

Events from $W+$jets outnumber $Z+$jets nearly three to one in the VR in part due to the sizable $Z \to b\bar{b}$ branching fraction and flavor-tagging requirements (*i.e.,* two anti-b-tagged subjets). The decay products of the vector bosons are recon-

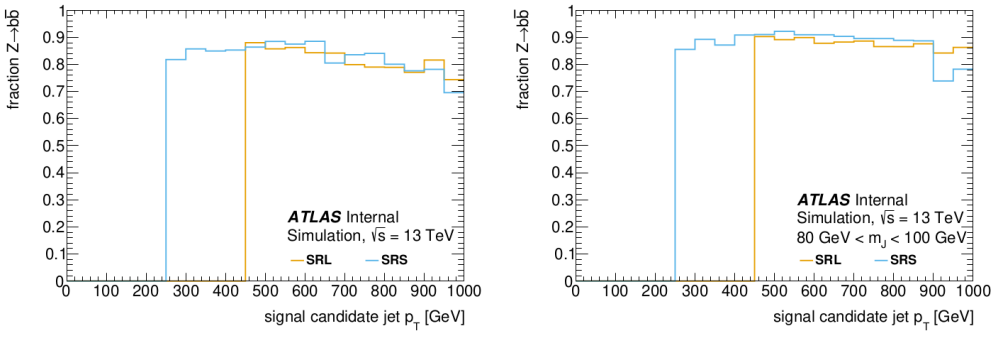


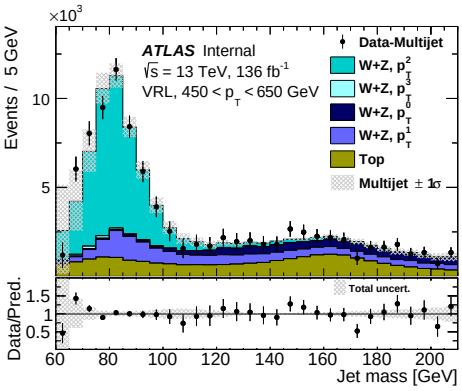
Figure 9.4: Contribution of $Z+$jets events in the signal regions that come from $Z \to b\bar{b}$ decays as a function of signal-candidate jet p_{T}. The fraction of all $Z+$jets events (left) and events with the signal candidate jet mass within a ± 10 GeV window around m_Z (right).

structed within the selected candidate large-R jets in approximately 40% and 90% of the W and Z events in the signal region, respectively. Within the validation region, over 60% of candidate jets contain some or all of the decay products of the vector boson, with the remainder created by the recoil jets, resulting in a non-resonant mass distribution similar in shape to the multijet background, enhancing the high-mass tail.

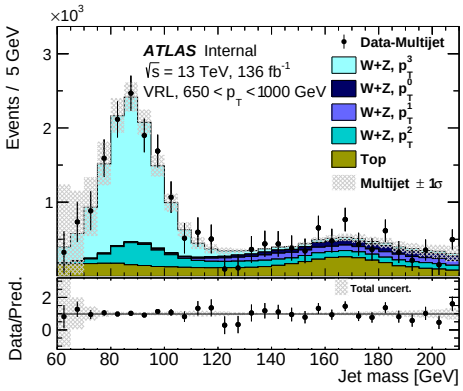
The impact of modeling systematic uncertainties is limited to changes in acceptance because the $Z+$jets normalizations are determined directly from data. The $W+$jets cross section has a 10% uncertainty [463]. The maximum variation from seven independent pairs of renormalization μ_R and factorization μ_F scale variations from 0.5 to 2 corresponds to a 3–20% error on the acceptance. An alternative PDF set (`MMHT2014nlo`), α_s variations within the nominal PDF set, and changing the cluster fragmentation model to the Lund string model [464] did not lead to significant differences in the acceptance estimate with respect to the nominal model. The maximum variations in μ_R/μ_F and fragmentation modeling are in Figures C.3 and C.4 of Appendix C.4, respectively. The fractional changes in the area around

the Higgs mass for the alternative PDF set, scale variations, and α_s variations can be found in Table C.1 of Appendix C.4.

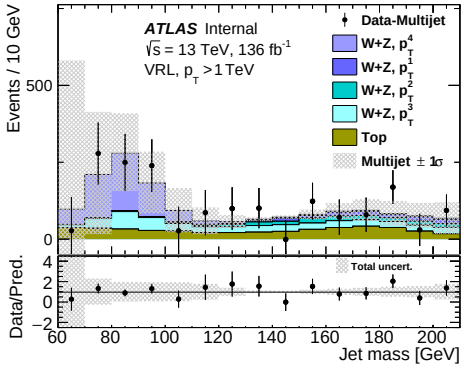
The experimental uncertainties discussed in Section 10.2 are applied to the $V+$jets samples in the fit. Uncertainties on the jet mass resolution (JMR) and jet mass scale (JMS) have the largest impact on the $V+$jets normalization. The external constraint of the $V+$jets JMR will be discussed in Section 9.5. Using the QCD multijet model described in Section 9.4.3 and the likelihood description in Section 10.1, the leading-jet invariant-mass distribution in the validation region shows good agreement between simulation and data within uncertainties, as shown in Figure 9.5.



(a) VRL: $450 < p_{\mathrm{T}} < 650\,\mathrm{GeV}$.



(b) VRL: $650 < p_{\mathrm{T}} < 1000\,\mathrm{GeV}$.



(c) VRL: $p_{\mathrm{T}} > 1\,\mathrm{TeV}$.

Figure 9.5: Post-fit leading-jet invariant-mass distributions after the multijet background subtraction in the validation region for data (points with error bars), and $V+$jets (W and Z) and Top components (histogram templates) for $450 < p_{\mathrm{T}} < 650\,\mathrm{GeV}$ (a), $650 < p_{\mathrm{T}} < 1000\,\mathrm{GeV}$ (b), and $p_{\mathrm{T}} > 1\,\mathrm{TeV}$ (c). The five fiducial volumes are labeled $p_{\mathrm{T}}^0 - p_{\mathrm{T}}^4$, corresponding to $p_{\mathrm{T}}^V < 300\,\mathrm{GeV}$, 300–450 GeV, 450–650 GeV, 650–1000 GeV, and $> 1000\,\mathrm{GeV}$, respectively. The p_{T}^0 event yield is constrained to its SM value within the theoretical and experimental uncertainties, and the free parameters act independently on the remaining four volumes. The $t\bar{t}$ normalization and its uncertainty are set to the corresponding values from the $\mathrm{CR}_{t\bar{t}}$.

9.4.2 Top Quark

The top-quark ($t\bar{t}$ and tW, jointly referred to as *Top*[4]) background populates the mass region above m_H. We designed the $\mathrm{CR}_{t\bar{t}}$, discussed in detail in Section 8.3, to ensure that the same underlying physics processes responsible for top-quark events

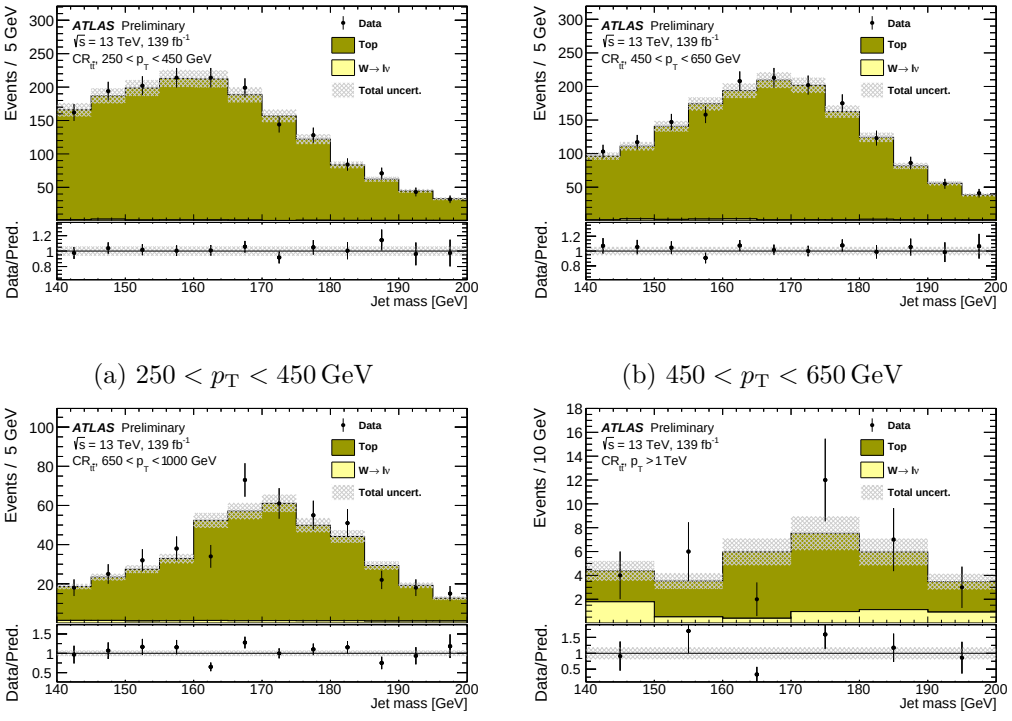
[4]The tW process is a minor contribution compared to $t\bar{t}$. It shares the same shape in the mass distribution. Both processes are combined with the normalization extracted from the $\mathrm{CR}_{t\bar{t}}$ control region.

in the SR also populate the control region (CR). The simulated jet mass distributions in both regions are similar in shape near the top-quark mass because they both probe the same top-quark transverse momentum. Therefore, we can apply the normalization scale factor which improves the agreement between MC and data in the CR directly to the SR. This is achieved by including the $\mathrm{CR}_{t\bar{t}}$ in the global likelihood fit described in Section 10.1. The purity of the $\mathrm{CR}_{t\bar{t}}$, as mentioned in Section 8.3, is nearly 97% inclusively, with similar levels found in the differential analysis regions, with lower purity above 1 TeV, which is statistically limited. This allows us to determine the Top normalization directly from data with better than 10% precision in most regions.

We performed comparisons[5] between the nominal and alternative samples to produce modeling systematic uncertainty estimates for variations in the parton shower model (Herwig 7 vs. Pythia 8) and the matrix element calculation (MadGraph5_aMC@NLO vs. Powheg-Box V2), resulting in a 6–20% and 1–19% difference in yield in the analysis regions, respectively. Within the nominal sample, variations of internal weights are used to extract the systematic uncertainties associated with ISR and FSR (1–7%), as well as μ_F and μ_R scale variations, which were negligible. Plots showing the effects of the variations on the invariant-mass distributions can be found in Appendix C.5.

We used the experimental uncertainties discussed in Section 10.2 in the fit. The uncertainties on b-tagging efficiency and JMS have the largest impact on the Top normalization. The post-fit jet-mass distribution for the $\mathrm{CR}_{t\bar{t}}$ MC shows good agreement with data, as shown in Figure 9.6.

[5]Following the recommendations from the Top Physics Group and Physics Modeling Group [465].



(a) $250 < p_{\mathrm{T}} < 450\,\mathrm{GeV}$

(b) $450 < p_{\mathrm{T}} < 650\,\mathrm{GeV}$

(c) $650 < p_{\mathrm{T}} < 1000\,\mathrm{GeV}$

(d) $p_{\mathrm{T}} > 1000\,\mathrm{GeV}$

Figure 9.6: The post-fit $\mathrm{CR}_{t\bar{t}}$ jet mass distribution in the different p_{T} regions.

9.4.3 QCD Multijet Background

The simulated QCD dijet event samples have an effective luminosity of only $\sim 5\%$ that of the data. This low level of statistics for such a large background makes using template from the Monte Carlo somewhat unreliable, as statistical fluctuations larger than the expected Higgs signal contribution are possible. Therefore, an exponential of polynomial functions is used to model the QCD background:

$$f_N(x|\vec{\theta}) = \theta_0 e^{\sum_{i=1}^{N} \theta_i x^i}, \tag{9.1}$$

where $x = (m_{\mathrm{J}} - 140\,\mathrm{GeV})/70\,\mathrm{GeV}$ and θ_i are the parameters extracted from a fit. The parameter values are determined independently in each region in a simultaneous fit, and the optimal degree for the polynomial N in Equation 9.1 depends on the mass

shape and number of events in each analysis region. We performed tests to determine the polynomial degree using the validation region data discussed in Section 8.1. The VRL (VRS) contain 58 (51) times the data in the SRL (SRS), allowing us to perform *ensemble tests* using modified VR subsets with approximately the same number of events as the corresponding signal region to determine N. Events are randomly assigned into the various subsets, which are referred to as hybrid validation region $\mathrm{VR_{hyb}}$ *slices*.

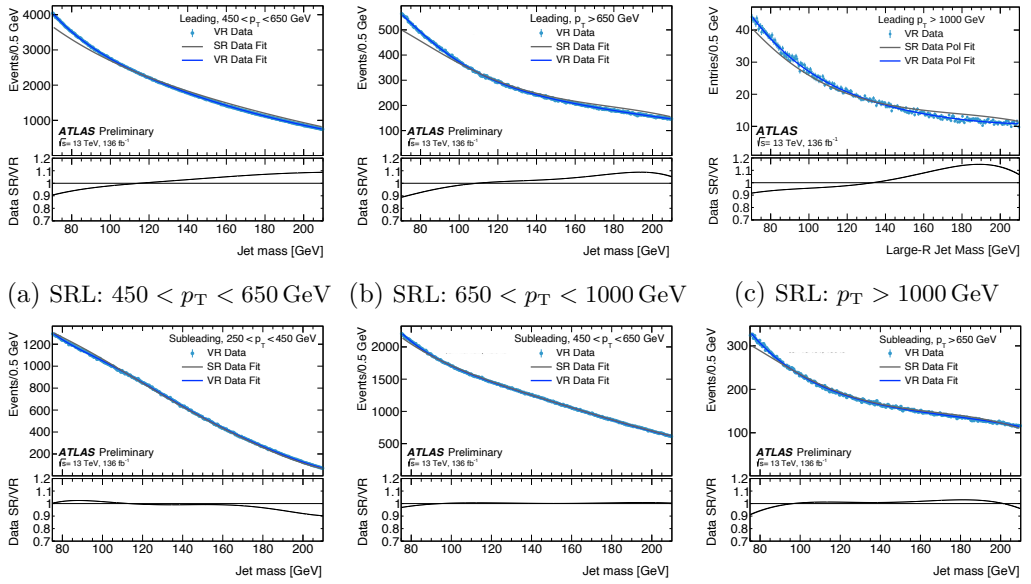
The hybrid VRs are the best proxies in the blinded mass region for the SRs. The hybrid VR is defined by replacing the VR resonance peaks with the SM prediction from the corresponding SR and correcting the underlying multijet shape for SR-acceptance effects. To determine the resonance and multijet shape, we created a VR multijet estimate ($\mathrm{MJ_{VR}}$) from the average of the multijet model parameter values obtained from likelihood fits to ten[6] random, mutually exclusive subsets of the VR including all experimental modeling uncertainties. The $\mathrm{CR}_{t\bar{t}}$ provides the $t\bar{t}$ normalization and associated uncertainty. The VR $V+$jets and Top estimates[7] ($\mathrm{V_{VR}}$ and $\mathrm{Top_{VR}}$, respectively) are calculated as the average post-fit $V+$jets and Top contributions from the fits to the VR slices. The SR QCD multijet estimate $\mathrm{MJ_{SR}}$ is obtained from a global likelihood fit to the SR. Each $\mathrm{VR^{\mathit{i}}_{hyb}}$ is:

$$\mathrm{VR^{\mathit{i}}_{hyb}} = \left(\frac{\mathrm{MJ_{SR}}}{\mathrm{MJ_{VR}}}\right)(\mathrm{VR}^i - \mathrm{V_{VR}} - \mathrm{Top_{VR}}) + \mathrm{V_{SR}} + \mathrm{Top_{SR}} + \mathrm{H_{SR}}, \tag{9.2}$$

where VR^i is the full mass distribution in a VR slice and $\mathrm{V_{SR}}$, $\mathrm{Top_{SR}}$, and $\mathrm{H_{SR}}$

[6] Ten slices balance the statistical precision with the need to use larger values of N to model a larger dataset.

[7] *Top* is the combined template of $t\bar{t}$ and single top.



(a) SRL: $450 < p_{\mathrm{T}} < 650\,\mathrm{GeV}$ (b) SRL: $650 < p_{\mathrm{T}} < 1000\,\mathrm{GeV}$ (c) SRL: $p_{\mathrm{T}} > 1000\,\mathrm{GeV}$

(d) SRS: $250 < p_{\mathrm{T}} < 450\,\mathrm{GeV}$ (e) SRS: $450 < p_{\mathrm{T}} < 650\,\mathrm{GeV}$ (f) SRS: $650 < p_{\mathrm{T}} < 1000\,\mathrm{GeV}$

Figure 9.7: Comparison of the QCD multijet shapes from the different p_{T}-binned analysis regions. The solid lines show the multijet function shape after a fit to the SR (gray) and VR (blue). The solid points are the averaged data from the VR slices with the same number of events as the SR. The $250 < p_{\mathrm{T}} < 450\,\mathrm{GeV}$ region is only populated by the subleading jet SR and the $p_{\mathrm{T}} > 1\,\mathrm{TeV}$ region is only populated by the leading jet SR.

are the nominal MC predictions for $V+$jets, Top, and H production, respectively. The ratio $\mathrm{MJ_{SR}}/\mathrm{MJ_{VR}}$ defines the acceptance differences between the two regions as shown in Figure 9.7.

We used three tests to choose the optimal N for each region. All three tests were performed using the hybrid VR data. A full description of the procedures can be found in Section 9.2 and Appendix B of Reference [403].

First, a log-likelihood ratio (LLR) test compared the result of an N-parameter fit (null hypothesis) to an $N+1$-parameter fit (alternative hypothesis) in each $\mathrm{VR_{hyb}}$ slice without any injected SM resonances. By Wilks' theorem [466], the likelihood ratio follows an asymptotic χ^2 distribution when the data correspond to the null hypothesis (*i.e.* the corresponding distribution of p-values is flat). The smallest N

yielding a uniform p-value distribution was chosen. Figure 9.8 shows the cumulative distribution function vs p-value distributions for several N vs. $N+1$ hypotheses.

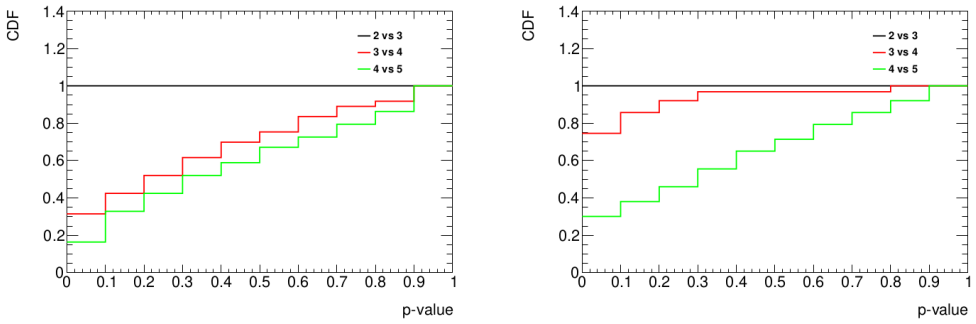


Figure 9.8: Cumulative distribution function (CDF) distributions of the probability for the difference in the log-likelihood obtained for exponential polynomial functions of order N and $N+1$. The results from the VRL and VRS jet categories are shown on the left and right panels, respectively. A flat probability distribution results in a cumulative distribution function that increases evenly as the 3 vs. 4 (red) and 4 vs. 5 (green) do for the VRL and VRS.

The LLR tests ensure a good global description of data over the full mass range, but resonances make us sensitive to local effects. We employed two rate tests using $\mathrm{VR}_{\mathrm{hyb}}$ slices with SR resonances injected at SM rates. These tests are sensitive to local effects and rely on the fit result of a free normalization parameter and its associated error (generalized as $\mu^{\mathrm{VR}} \pm \sigma^{\mathrm{VR}}_{\mathrm{stat}}$) on either the $Z+$jets process or the Higgs boson process. We estimate the probability that the multijet model allows for a substantial artificial excess with $F_{2\sigma}$, the fraction of results where $|\mu^{\mathrm{VR}} - 1|$ is beyond twice its error $\sigma^{\mathrm{VR}}_{\mathrm{stat}}$. A 2σ threshold ensures that some results from the full set of VR slices cross the boundary. The average value for the $(\mu^{\mathrm{VR}} - 1)/\sigma^{\mathrm{VR}}_{\mathrm{stat}}$, $\overline{\mu/\sigma}$, indicates if there is a bias in the signal strength determination. The N chosen by the LLR test is incremented until $F_{2\sigma}$ is compatible with 0.05 and $\overline{\mu/\sigma}$ stabilizes for both $Z+$jets and H production. The $F_{2\sigma}$ values vs. N for the inclusive SRL and SRS can be seen in Figure 9.1b

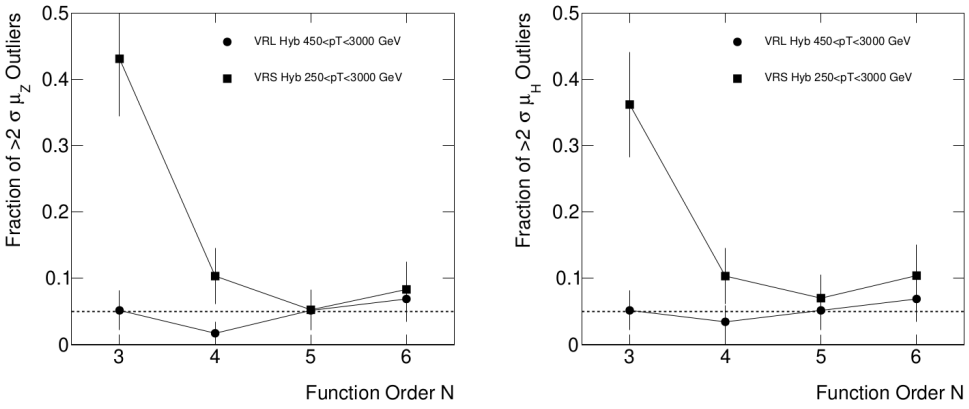


Figure 9.9: Spurious signal test on the leading and subleading inclusive VR hybrid data slices using $F_{2\sigma}$ for the Z (left) and H (right). The fraction of fitted signal values in excess of $2\sigma_{\text{stat}}^{\text{VR}}$ is shown as a function of the polynomial order N.

When including systematic uncertainties and when not injecting $Z+$jets and H processes into the hybrid slices (with appropriate changes to the definition of $F_{2\sigma}$ and $\overline{\mu/\sigma}$) the resultant N agrees. A fifth degree polynomial is used in both inclusive regions and either a fourth or fifth degree polynomial is used in the analysis p_{T} bins. The $\overline{\mu/\sigma}$ for the chosen N in each region indicates if there is a bias in the background model and defines the spurious signal systematic uncertainty and ranges from 0.01–0.33 for H and 0.15–0.65 for $Z+$jets production. In both cases, it has an insignificant impact on the sensitivity. We also considered the use of a Laurent series as an alternative fit function. It provides a similar level of agreement with the QCD multijet background in the hybrid VR dataset. The differences between the two models is significantly smaller than the statistical uncertainty of the Higgs signal and thus did not necessitate an additional systematic uncertainty.

9.5 Jet Mass Resolution and W-$t\bar{t}$ Control Region

9.5.1 Jet Mass Resolution

In Section 9.4.3, we saw that the modeling of the QCD multijet background makes use of an exponential polynomial fit rather than template histograms produced from MC, which are used for the signal and other backgrounds. The use of a multi-parameter functional form to describe the QCD background trades the modeling uncertainties for a statistical uncertainty and enhances our sensitivity to the modeling of the broad resonances from the $V+$jets and $t\bar{t}$ processes, as well as making us sensitive to the concavity of the function. Within the fitting framework, discussed in Chapter 10, we have several fit parameters and only ~ 30 mass bins. The heavy resonances (H, W, and Z bosons and top quarks) are each largely concentrated within a third of the fitted mass range. Many of the mass bins are affected by the same parameters, such as the modeling uncertainties, and the large correlations make it imperative that we control the systematic uncertainties carefully in particular the jet mass resolution (JMR). The interplay between the QCD fit and the broad resonances can significantly alter the estimate of the multijet background below the resonances and therefore the evolution of the resonance's rate μ. The contribution of the Higgs signal is within the statistical error of the data and hence the QCD model is not expected to be able to deform itself around the Higgs template in a way that minimizes the likelihood.

The fitted Z normalization in the signal region shows a significant correlation with the reconstructed jet mass resolution uncertainty due to the interaction between the $Z+$jets component in the mass spectrum and the QCD multijet model's flexibility, opening up a local minimum. This feature was revealed using subsets of

the hybrid validation region $\mathrm{VR_{hyb}}$, constructed to have a known amount of each process. Using the $\mathrm{VR_{hyb}}$, the μ_V-JMR correlation can be studied in detail. Figure 9.10 shows QCD, μ_Z, and JMR in the 58 fits to the hybrid VRL slices. The μ_V value becomes inflated when there are large ($> 2\,\sigma$) upward pulls from the JMR, thus broadening the $V+$jets peak, which correlates with a downward distortion in the QCD for that mass range. While the jet mass scale systematics, which control the peak position, do not show significant correlations with the fitted μ_Z values in the SR/VR, the high correlation between the $V+$jets JMR systematic and μ_Z negatively affects the reliability of the μ_Z extracted from the signal regions.

For some $\mathrm{VR_{hyb}}$ subsets, the JMR parameter broadens the Z-peak, corresponding to an increase in the $Z+$jets normalization and a decrease in the QCD multijet contribution from the nominal expectation. Although the parameter represents a JMR uncertainty, or a lack of knowledge of the spread of the detector response in the jet mass, in this situation it is more of an effective parameter which covers all systematic effects related to the jet-mass width (as these effects are indistinguishable with the available information) and is sensitive to the reconstructed jet-mass width. The analysis follows the recommendations of the Jet/ETMiss Group for the jet mass scale and jet energy scale uncertainties [467]. The JMR recommendation of a 20% uncertainty was obtained from differences seen in cross-calibration samples [468]. The recommendations from recent measurements [469] for the JMR (X = heavy resonance) uncertainties are applied to $H \to b\bar{b}$, top, and W/Z jets separately. Figure 9.11 shows the high level of correlation between the Z yield and QCD estimate through the JMR uncertainty. The JMR uncertainty is also obtained by the ATLAS Jet/ETMiss Group [470] using the official JMR smearing tool using a prescribed

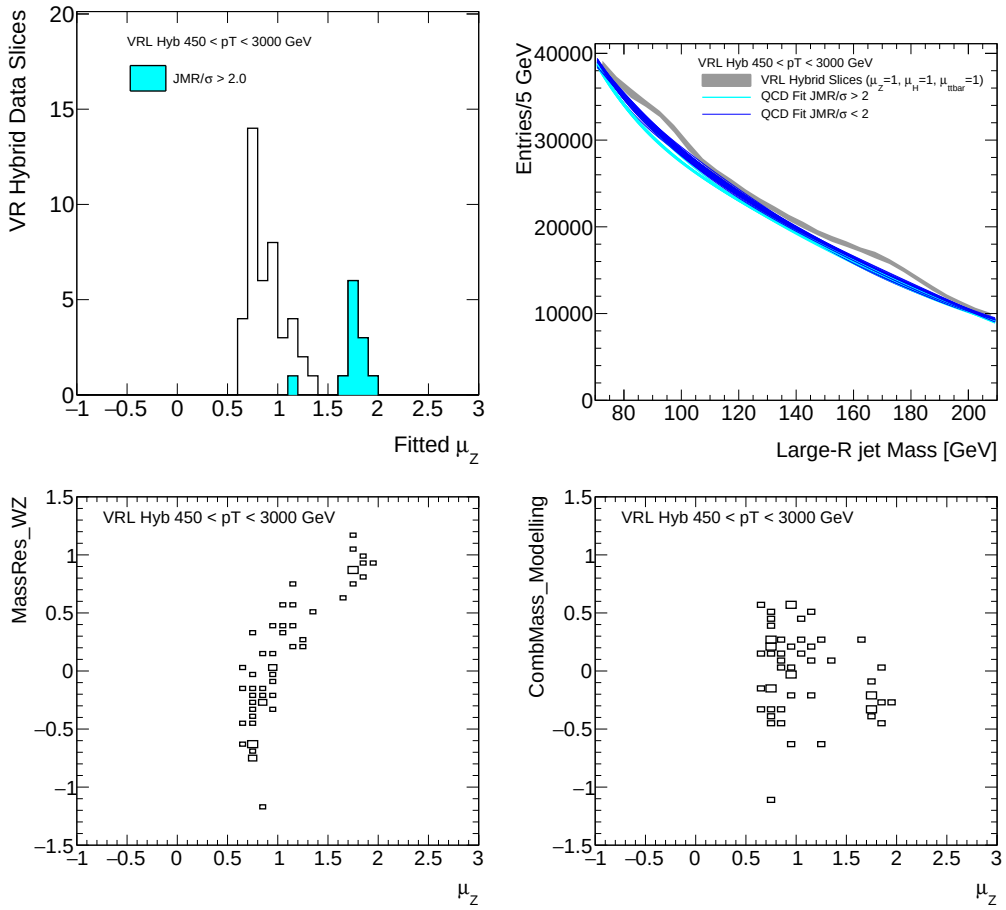


Figure 9.10: From a fit to the hybrid slices from the inclusive VRL: the upper left panel shows the fitted value of μ_Z and the upper right panel the results of the QCD fits overlaid with the hybrid spectra with those resulting in a significant JMR upward pull highlighted in cyan. The lower left and right panels show the correlation between JMR and JMS, respectively, with the fitted value of μ_Z.

20% uncertainty. The correlation between the JMR uncertainty and the fitted μ_Z parameter is nearly 90% before the introduction of the external constraint, and when applied is reduced to nearly 30%.

To stabilize our fit response, we used two data samples to measure the jet-mass widths of the W and Z resonances to place an additional constraint in the global likelihood. The two data samples are an alternate $t\bar{t}$ control region ($\text{WCR}_{t\bar{t}}$) and the VRL. I designed the $\text{WCR}_{t\bar{t}}$, implemented it into the analysis framework, and performed fits to evaluate the region. The $\text{WCR}_{t\bar{t}}$, like the $\text{CR}_{t\bar{t}}$ discussed in Section 8.3,

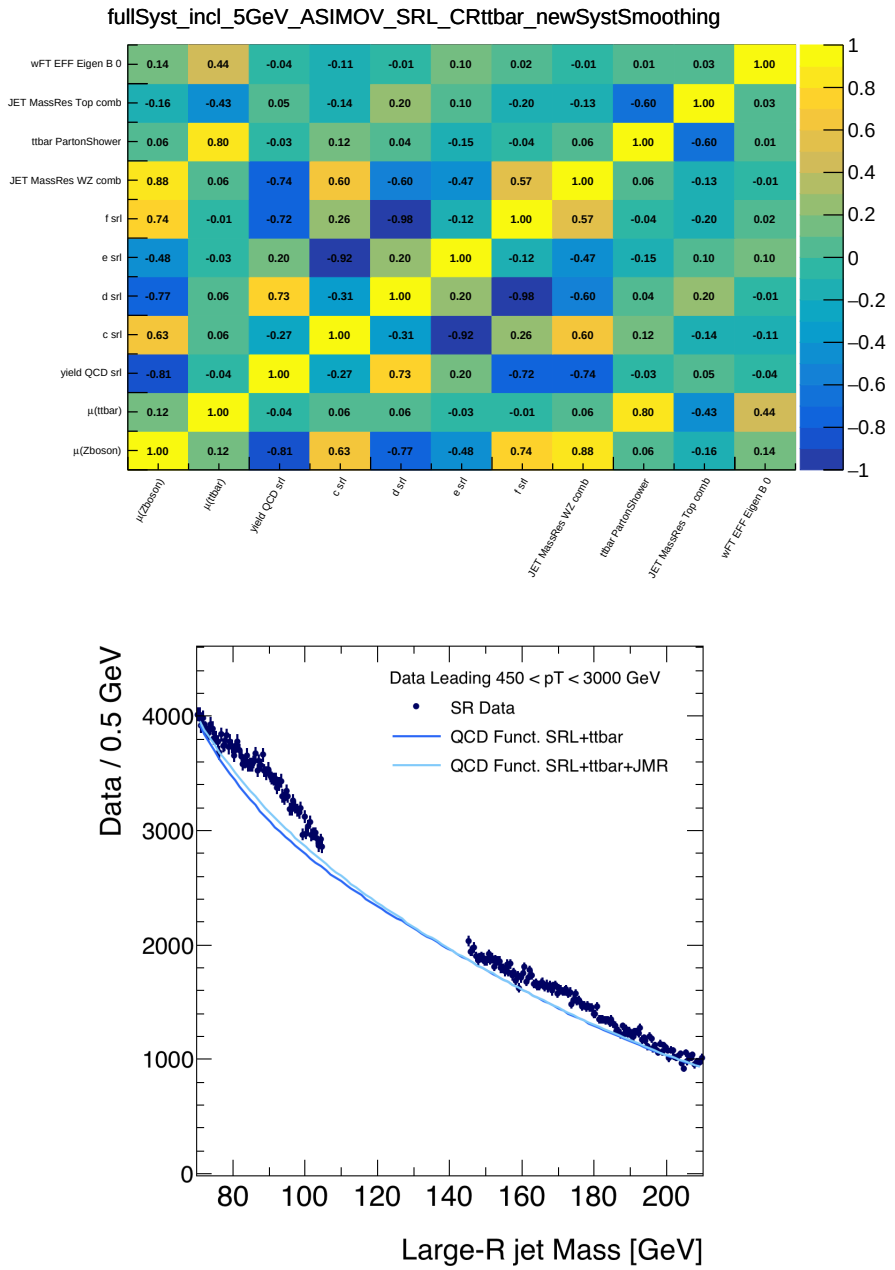


Figure 9.11: From a fit to the inclusive SRL region, the upper panel shows the high level of correlation between μ_Z, the JMR uncertainty (`JETMassResWZcomb`), and the QCD function parameters while the lower panel shows the large change in the QCD estimate in a fit without the WZ JMR uncertainty (dark blue) compared with the JMR uncertainty (light blue).

takes advantage of the semileptonic decay signature of $t\bar{t}$ events. In contrast with the $\mathrm{CR}_{t\bar{t}}$, the $\mathrm{WCR}_{t\bar{t}}$ selects for a resolved Wb pair from the hadronically decaying top quark, where the decay products of the W boson are captured within a large-R jet and a variable-radius b-tagged jet is located outside the W-jet. The region provides a high-purity reconstructed W-peak in the jet-p_{T} range of 200–600 GeV. As the transverse momentum increases, the opening angle between the W boson and b-quark decreases, causing the event to fail selection; so, by construction, this region does not have the full kinematical reach that the $\mathrm{CR}_{t\bar{t}}$ has. The VRL provides a clear W-peak and covers the entire jet-p_{T} range of the analysis, but has a considerably larger multijet contribution compared to the $\mathrm{WCR}_{t\bar{t}}$. The details of the $\mathrm{WCR}_{t\bar{t}}$, its selection criteria, studies, and implementation will be discussed in the following sections. The use of the CR to extract the external constraint [471] for the fit is not an obvious procedure, so we will cover the topic in detail. I should note that the work concerning the constraint transfer was not performed by myself, but is an important component to understanding how the JMR external constraint, extracted from a sample rich in $W \to jj$ large-R jets ($jj \to$ light-quark jets), is applied to a sample comprised of $Z \to b\bar{b}$ large-R jets.

9.5.2 $\mathrm{WCR}_{t\bar{t}}$ Selection Criteria

A dedicated control region ($\mathrm{WCR}_{t\bar{t}}$) – rich in large-R jets containing the decay products of the W boson from semileptonic $t\bar{t}$ decays – was designed to extract information on the jet mass resolution needed to constrain the JMR systematic uncertainty for $V+$jets. The Sherpa 2.2.1 $t\bar{t}$ MC sample is used because it is the available sample closest to that used for the $V+$jets process (Sherpa 2.2.8). Differences between the two versions only have minor effects on our results. In par-

ticular, SHERPA 2.2.1 is known to have an abnormally large rate of heavy-flavored baryons [444]. However, the effect of heavy-flavored baryons on the boson mass-peak width has been studied for b-baryons in $Z \to b\bar{b}$ events and found to be on the order of a few percent. We will see comparisons between SHERPA versions and POWHEG + PYTHIA 8 in the following sections.

The $\text{WCR}_{t\bar{t}}$ is defined, similar to the $\text{CR}_{t\bar{t}}$, as having a leptonically decaying top quark in one hemisphere and a hadronically decaying top quark in the other, with the goal of isolating the decay products of the W boson coming from the hadronically decaying top in a large-R jet. Events are triggered using either the `HLT_mu26_ivarmedium` or `HLT_mu20_iloose_L1MU15` triggers, depending upon the year.

Events are identified first by finding an isolated trigger-matched muon with $p_{\text{T}} > 27\,\text{GeV}$ to be efficient across all years. A b-tagged variable-radius jet at the 77% WP, with $p_{\text{T}} > 25\,\text{GeV}$, is then required within $\Delta R < 1.5$ of the muon. This b-jet is required to have a $\Delta R > 2.0$ of the leading large-R jet in the event, which is identified as the W-candidate. This ensures a back-to-back topology in the $t\bar{t}$ decay. To contain the W boson only, a second b-tagged variable radius jet, with $p_{\text{T}} > 10\,\text{GeV}$, is required within $1.0 < \Delta R < 1.5$ of the W-candidate large-R jet to ensure they are resolved. The selection is visually represented in Figure 9.12 and summarized below:

- At least 1 large-R Jet (W-candidate), $p_{\text{T}} > 200\,\text{GeV}$,
- Trigger on isolated medium-quality muon, $p_{\text{T}} > 27\,\text{GeV}$,
- b-tagged variable radius jet (bjet1)

- $p_{\mathrm{T}} > 25\,\mathrm{GeV}$
- $0.04 + 10/p_{\mathrm{T}}^{\mu} < \Delta R_{\mathrm{jet1,muon}} < 1.5$,
- $\Delta R_{b\mathrm{jet1,Jet}} > 2.0$

- Second b-tagged variable radius jet (bjet2)
  - $p_{\mathrm{T}} > 10\,\mathrm{GeV}$
  - $1.0 < \Delta R_{\mathrm{bjet2,Jet}} < 1.5$

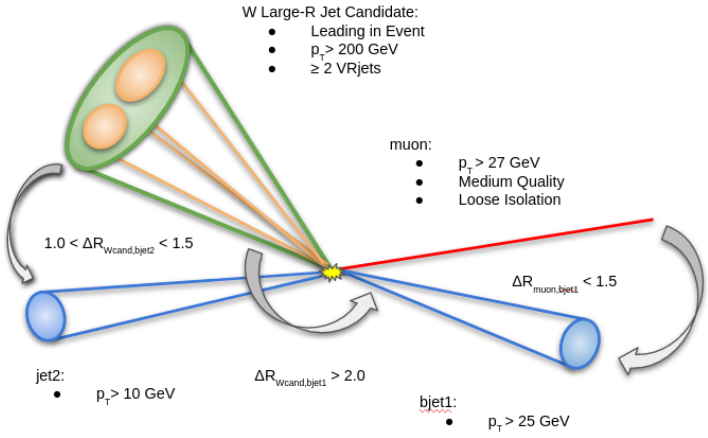


Figure 9.12: Topology of $\mathrm{WCR}_{t\bar{t}}$ events.

The region is analyzed inclusively and in W-jet p_{T} bins of $200 < p_{\mathrm{T}} < 275\,\mathrm{GeV}$, $275 < p_{\mathrm{T}} < 375\,\mathrm{GeV}$, and $375 < p_{\mathrm{T}} < 1500\,\mathrm{GeV}$. The p_{T} bins used in the $\mathrm{WCR}_{t\bar{t}}$ are not the same as the analysis due to the large-R jet containing the b-quark at high transverse momentum as mentioned in the previous section. They were chosen to provide adequate statistics in the jet mass distribution across several bins in p_{T}. The kinematics of the selection do not allow for good statistics at the upper end of the signal region for the jet-p_{T} ranges probed. However, this can still provide a good confirmation of effects seen in the lower end of the validation region. The selection

criteria described produce an exceptionally pure sample of large-R jets rich in W bosons, as can be seen in Table 9.3.

9.5.3 WCR$_{t\bar{t}}$ Data/MC Comparisons

Table 9.3: The pre-fit yields for all simulated events and data in WCR$_{t\bar{t}}$.

Process	$p_{\mathrm{T}} > 200\,\mathrm{GeV}$	$200 < p_{\mathrm{T}} < 275\,\mathrm{GeV}$	$275 < p_{\mathrm{T}} < 375\,\mathrm{GeV}$	$375 < p_{\mathrm{T}} < 1500\,\mathrm{GeV}$
$W\ell\nu$	151.8 ± 12.3	66.5 ± 8.2	44.9 ± 6.7	40.3 ± 6.3
Zqq	0.3 ± 0.6	0.3 ± 0.6	0.0 ± 0.0	0.0 ± 0.2
single t	495.4 ± 22.3	247.9 ± 15.7	151.5 ± 12.3	95.8 ± 9.8
dijet	15.0 ± 3.9	0.0 ± 0.0	14.6 ± 3.8	0.5 ± 0.7
$t\bar{t}$	$26{,}327.0 \pm 162.3$	$16{,}247.7 \pm 127.5$	7628.4 ± 87.3	2450.9 ± 49.5
data	$25{,}020.0 \pm 158.2$	$15{,}531.0 \pm 124.6$	7168.0 ± 84.7	2320.0 ± 48.2

Pre-fit plots are shown for the inclusive sample in Figure 9.13 and in jet-p_{T} bins in Figure 9.14.

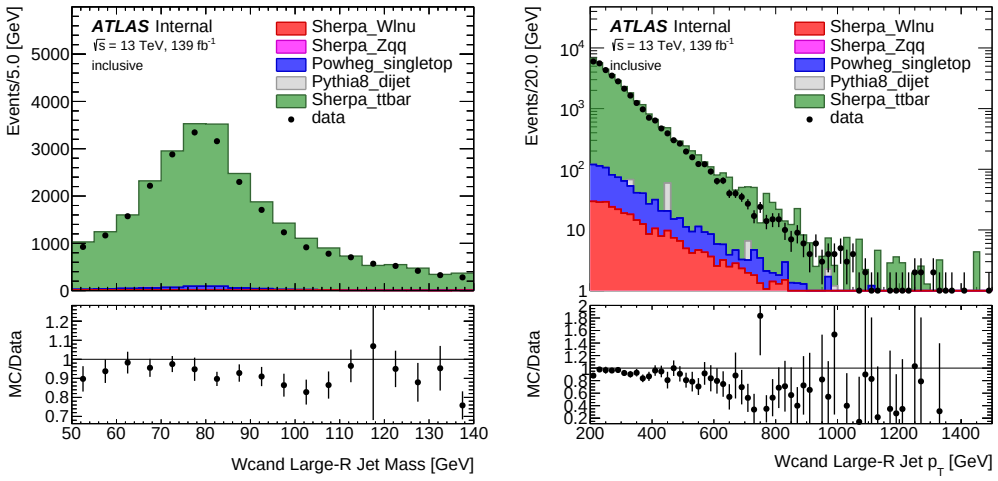
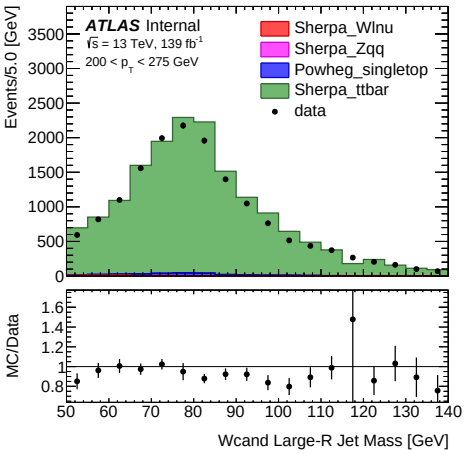
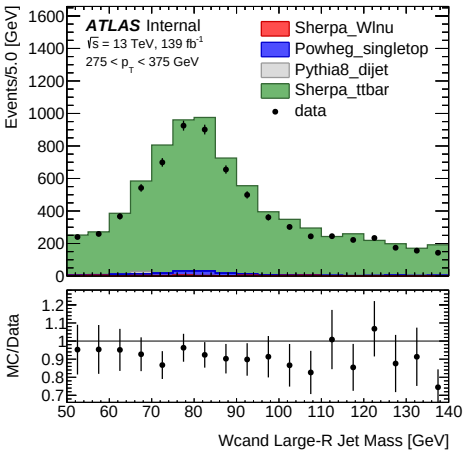


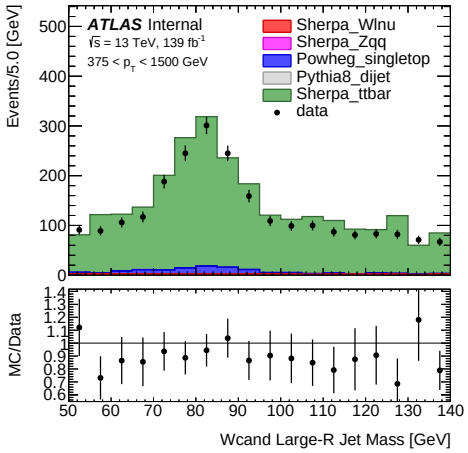
Figure 9.13: Pre-fit inclusive WCR$_{t\bar{t}}$ distributions of the W-from-top large-R jet mass (left) and p_{T} (right).



(a) $200 < p_{\mathrm{T}} < 275\,\mathrm{GeV}$



(b) $275 < p_{\mathrm{T}} < 375\,\mathrm{GeV}$



(c) $375 < p_{\mathrm{T}} < 1500\,\mathrm{GeV}$

Figure 9.14: Pre-fit WCR$_{t\bar{t}}$ distributions of the W-from-top large-R jet mass in three jet p_{T} bins.

9.5.3.1 Post-Fit Plots

After a fit including a full set of systematic uncertainties (see Section 10.2), the inclusive post-fit mass distribution can be found in Figure 9.15. Figure 9.16 displays the post-fit mass distribution in for the p_{T}-binned result. Initial tests including other contributions in the WCR$_{t\bar{t}}$ showed negligible effects. Hence, only the $t\bar{t}$ contribution is considered in the final fit.

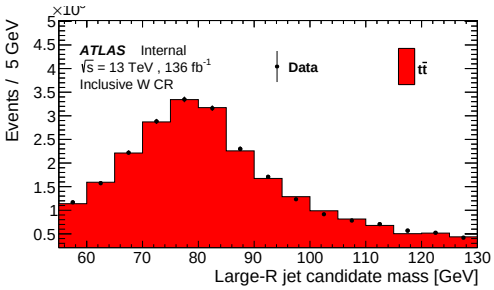
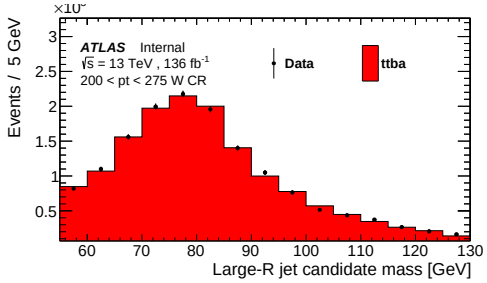
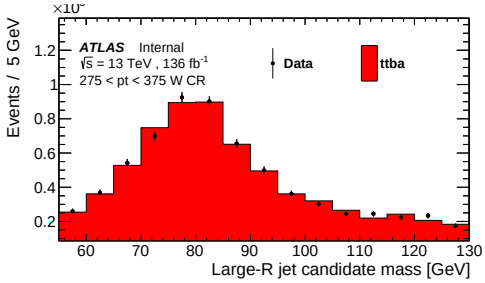


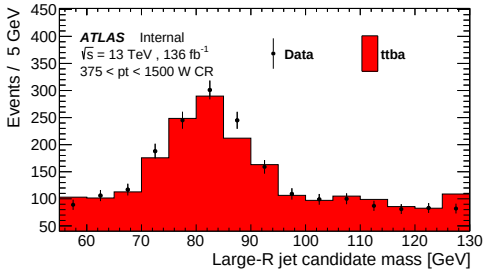
Figure 9.15: Post-fit data and $t\bar{t}$ MC large-R jet-mass comparison in WCR$_{t\bar{t}}$. The cut off power of ten for the y-axis is 10^3.



(a) $200 < p_{\mathrm{T}} < 275\,\mathrm{GeV}$



(b) $275 < p_{\mathrm{T}} < 375\,\mathrm{GeV}$



(c) $375 < p_{\mathrm{T}} < 1500\,\mathrm{GeV}$

Figure 9.16: Post-fit data and $t\bar{t}$ MC large-R jet-mass comparison in WCR$_{t\bar{t}}$ in the three p_{T} bins available. The cut off power of ten for the lower-p_{T} bins for the y-axis is 10^3.

9.5.4 Comparisons (Generator and MC/Data)

We compared the widths, extracted from a Gaussian fit, of the truth-matched large-R-jet mass distributions in the SHERPA 2.2.1 $t\bar{t}$ $\text{WCR}_{t\bar{t}}$ and SHERPA 2.2.8 $W+$jets VR samples, where the p_{T} distribution of the matched jets have been restricted to 250–600 GeV and the VR jets have been reweighted to the $\text{WCR}_{t\bar{t}}$ p_{T} spectrum. The resulting mass spectra are in Figure 9.17. The Gaussian widths fitted from the core of the mass peaks are 9.30 ± 0.03 GeV and 9.42 ± 0.41 GeV for the SHERPA 2.2.8 VR and SHERPA 2.2.1 $\text{WCR}_{t\bar{t}}$ samples, respectively. To compare with the results presented in Reference [469], we included the POWHEG + PYTHIA 8 $\text{WCR}_{t\bar{t}}$ in the fit. The mass distributions between the two generators for the $\text{WCR}_{t\bar{t}}$, binned in p_{T} and pileup, can be found in Appendix C.7. Only the results obtained in the SHERPA simulation are used for the constraints in the analysis. The fitted JMR values, which define the nominal mass-peak width, from the two samples are not expected to agree numerically, since they are relative to the predictions of each MC. However, trends in the JMR parameter vs. p_{T} are comparable, as can be seen in Figure 9.18.

To perform a direct comparison of the results obtained from the SHERPA and POWHEG + PYTHIA 8 generators, we translated the fitted JMR values into the corresponding widths of the W and Z resonances. By measuring the widths corresponding to the nominal and $1\,\sigma$ variations of the JMR for each Monte Carlo and p_{T} range, we obtain a calibration for the JMR variations in terms of the variations of the mass-peak width. This allows us to extract the widths and errors from the fit result. Figure 9.19 shows the resulting mass widths as a function of the p_{T}. We see good agreement between the W mass-peak widths in the SHERPA and POWHEG

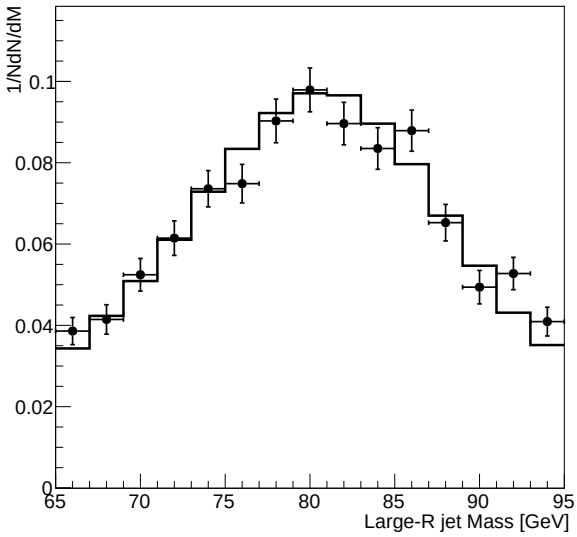


Figure 9.17: Comparison of the shape of truth-matched and p_{T}-reweighted W boson reconstructed spectra for the SHERPA 2.2.8 $W+$jets VR (continuous histogram) and SHERPA 2.2.1 $t\bar{t}$ WCR$_{t\bar{t}}$ (points with error bars) samples.

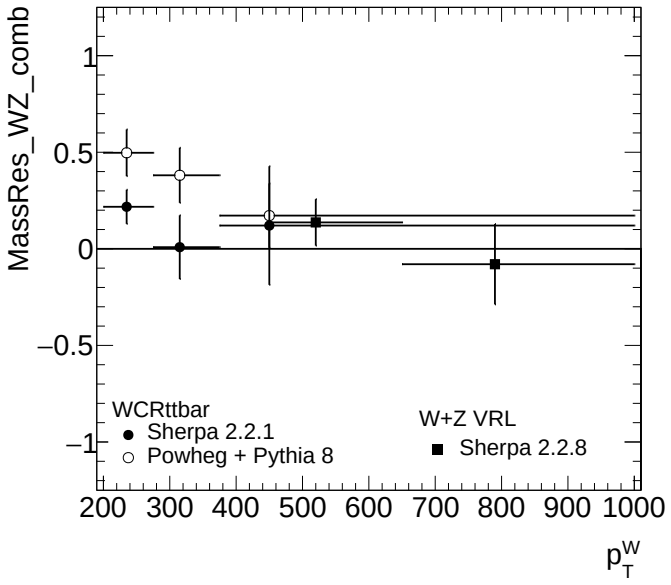


Figure 9.18: A summary of the results of the fits of the JMR values as a function of the large-R jet p_{T} using the resolved W in top decays in the WCR$_{t\bar{t}}$ region and the $W+Z$ in the anti-b-tagged validation region of the analysis. Results obtained with SHERPA 2.2.1 and POWHEG + PYTHIA 8 are compared for the WCR$_{t\bar{t}}$. Given that the JMR value is referenced to each MC generator, the results from the two MC samples do not need to be numerically compatible. The length of the horizontal bars for each point corresponds to the size of the p_{T} bin used in the fits while the marker position corresponds to the average p_{T} of the bin in the corresponding region.

+ Pythia 8 MC, despite the tension between fitted JMR values. It shows that the JMR uncertainty is generous enough to cover the differences between the generator's starting point and the data. The posterior constraint is driven by the data, so an increase in the JMR, in its capacity as an effective prior, should have little impact on the result. Additionally, we observe that the variations of the vector boson peak widths are expected to have a minimum mass resolution for the combined large-R jet around $p_{\mathrm{T}} = 800\,\mathrm{GeV}$, which corresponds to our highest fitted point [396] (see Figure 6.17 in Section 6.5.4). Lastly, the fact that we have a mixture of the W and Z bosons necessitates an additional uncertainty to account for possible differences in the rate of the two bosons, which distorts the shape of the fitted mass spectrum. As with the SRs, a $\pm 10\%$ uncertainty is associated to the relative W rate, which is summed in quadrature with the uncertainty arising from differences in the apparent mass width from the corresponding JMR fit values.

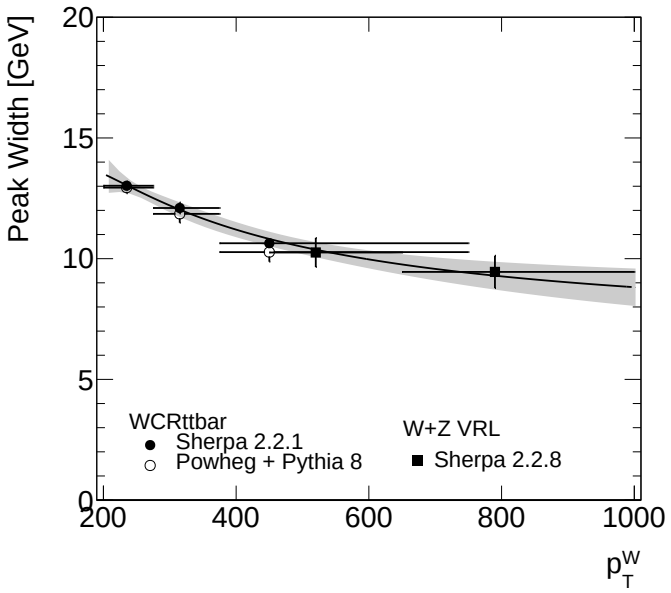


Figure 9.19: A summary of the results of the JMR fits expressed in terms of W and Z mass resonance widths as a function of the large-R jet p_{T} using the resolved W in top decays in the $\mathrm{WCR}_{t\bar{t}}$ region and the $W+Z$ in the anti-b-tagged validation region of the analysis. Results obtained with Sherpa 2.2.1 and Powheg + Pythia 8 have compatible values of mass width.

The fitted values of the JMR, as a function of the candidate W/Z large-R-jet p_{T} are interpolated using a function of the form:

$$\mathrm{JMR}(p_{\mathrm{T}}) = a + b \cdot p_{\mathrm{T}} + \frac{c}{p_{\mathrm{T}}}. \tag{9.3}$$

The 68% CL band of the resulting fit for the constraint in each region, accounting for correlations, is used as a statistical uncertainty. For the final result, the constraint terms are derived in two ways using the fit to JMR parameters. The first is simply to use the function value and associated error for the average p_{T} of the region being examined. The second method incorporates the increase in statistical uncertainty vs. p_{T} by averaging, or folding, the function value and error over the p_{T} spectrum for the region. This is achieved by sampling the Z p_{T} spectrum in the corresponding signal region 250,000 times, noting the function value and error at the sampled p_{T}, then averaging the values. To transfer this constraint derived from $W \to jj$ to $Z \to b\bar{b}$ large-R jets, we must take into account differences in the detector response between light-quark jets and b-jets, as well as the characteristics of b-decays compared to light-quark fragmentation. We will cover the constraint transfer procedure in the following section.

9.5.5 Constraint Transfer

As we've discussed, the constraints derived from the validation region and the $\mathrm{WCR}_{t\bar{t}}$ come from decays of the vector bosons to light quarks. We then need to transfer this constraint to the $Z \to b\bar{b}$-dominated 2-tag region. After accounting for the mass differences between the two bosons, the difference in the jet-mass width, both at truth level and simulated detector level, is driven by the differences in the

decay patterns of light and bottom quarks. We thus consider additional systematic uncertainties resulting from differences in the detector response and fragmentation properties between light- and bottom-flavored jets.

Detector Response

Following the methodology of the studies performed in Reference [469], the detector response for $b\bar{b}$ and $q\bar{q}$ decays was compared using the distribution of the ratio $m_{\text{reco}}/m_{\text{truth}}$ fitted with a Gaussian in a $\pm 0.75 * \text{RMS}$ window centered on the maximum bin. A p_{T}-dependent cut on m/p_{T} was used to separate out the high-mass-tail fluctuations from the bulk of the resonance. Figure 9.20 shows the mean and width from the fits. Only the width is used. The results from these studies show that the detector response is largely independent of the flavor of the quarks and that transferring the constraint from $W \to jj$ to $Z \to b\bar{b}$ jets is valid up to potential differences between simulation and data unique to b-quarks.

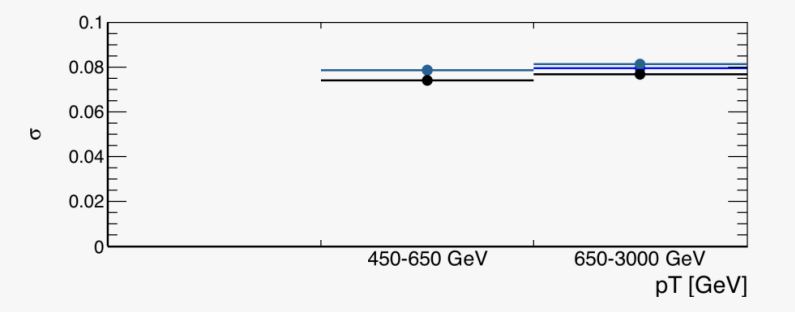
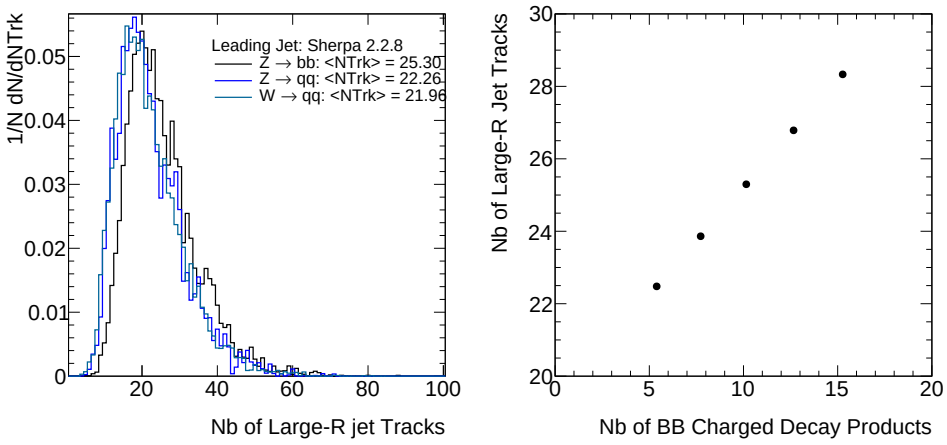


Figure 9.20: From fitting a Gaussian to $m_{\text{reco}}/m_{\text{truth}}$ in the signal region, the resolution of the detector response is comparable between $W \to jj$ (gray), $Z \to jj$ (blue), and $Z \to b\bar{b}$ (black). Differences in the mean would inform decisions on the jet mass scale which is beyond the scope of this work.

b-hadron Modeling

The particle multiplicity in b-decay products is largely determined by the b-quark decay chain rather than the quark fragmentation, resulting in an increase in the number of charged particles in the $Z \to b\bar{b}$ large-R jets (increase of approximately 12.5%) compared to $Z \to jj$ jets, which are similar to $W \to jj$ jets. A *difference* of 3.04 tracks is predicted by the SHERPA simulation, as seen in Figure 9.21, in good agreement with the SLD [472, 473] measurement of 2.93 ± 0.33 [474]. The charged-particle multiplicity of b-hadron decays has been measured precisely at e^+e^- colliders. The results obtained at LEP [95], an experiment which shared a similar mixture of b-hadron species present at the LHC, of 4.97 ± 0.07 [475] is propagated along with the uncertainty on the Z mass resolution to obtain an uncertainty on the $Z \to b\bar{b}$ jet-mass width of 0.3 GeV. The effect is estimated by reweighting events according to the number of semileptonic decays from the heavy-flavored quarks contained in the large-R jet. From the variation motivated by a $\pm 0.4\%$ uncertainty on the semileptonic branching fraction [476], the $Z \to b\bar{b}$ jet-mass width uncertainty is estimated to be ± 0.02 GeV.

Figure 9.22 shows the width of the jet mass as a function of the number of decay products, where a clear dependence on the track multiplicity is shown for $Z \to b\bar{b}$ jets. The slope of this difference between $Z \to b\bar{b}$ decays with differing numbers of semileptonic decays present agrees with $W \to jj$ and $Z \to jj$ jets. Additionally, the b-tagging efficiency of the MV2c10 BDT discriminant is track-multiplicity dependent. As shown in Table 6.2 of Section 6.4.2, the impact parameter and vertexing physics taggers (inputs to the BDT) rely on the track multiplicity. The $Z \to b\bar{b}$ tagging efficiency for the standard working points (WPs: 60%, 70%, 77%,



(a) (b)

Figure 9.21: (a) The number of reconstructed tracks in leading large-R jets within the analysis phase space. The b-decay products explain the differences seen in $Z \to b\bar{b}$ jets. In (b), the number of tracks within a $Z \to b\bar{b}$ jet is shown vs. the number of b-decay products.

and 85%) vs. the number of b-decay products and the change in the $Z \to b\bar{b}$ mass width for the respective WPs is shown in Figure 9.23. The systematic uncertainty on the b-tagging efficiency of 77% WP from the calibration (see Appendix M of Reference [403]) is propagated to the mass resolution with an effect estimated at $\pm 0.05\,\text{GeV}$.

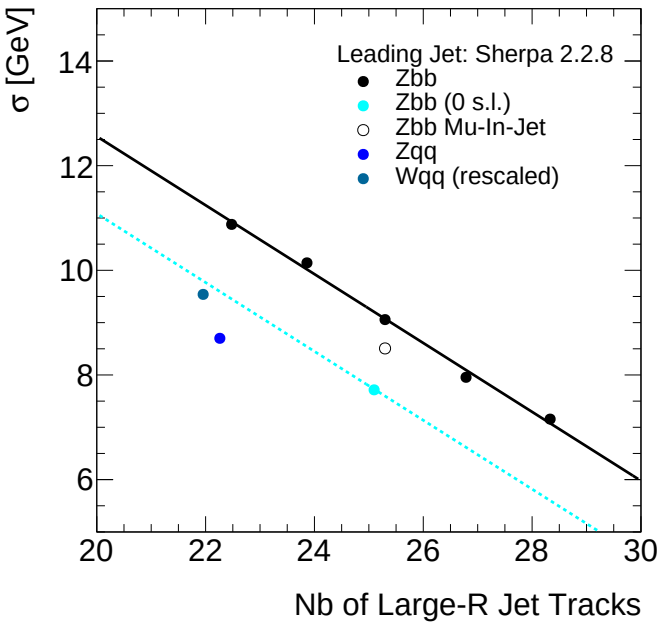


Figure 9.22: Effects of b-decay multiplicity drives the difference in large-R track multiplicity and therefore the mass resolution. $W \to jj$ events are rescaled by the W/Z mass difference. The dotted cyan line follows the same slope as the black points ($Z \to b\bar{b}$) showing good agreement between $Z \to b\bar{b}$, $W \to jj$ and $Z \to jj$ when all decay products are visible to the detector. The degraded resolution resolution due to energy losses in decay chain (s.l. decays) is partially recovered through muon-in-jet correction.

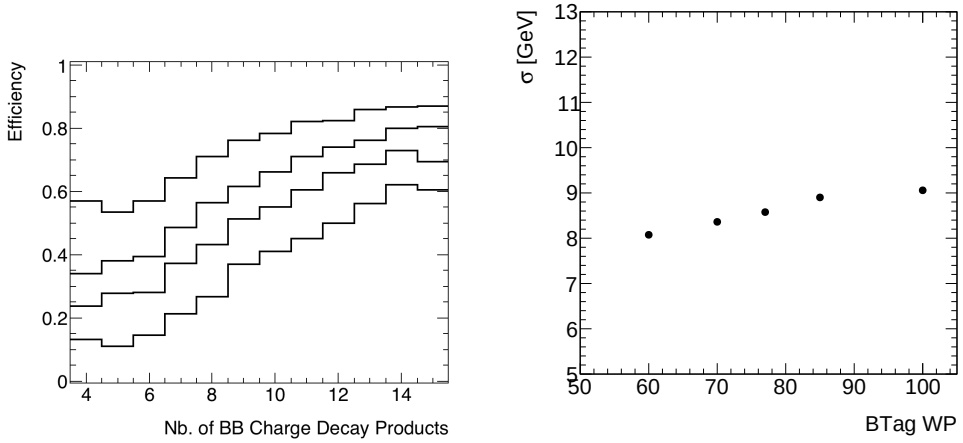


Figure 9.23: $Z \to b\bar{b}\ bb$ tagging efficiency for four standard flavor-tagging working points (ascending from 60% to 85% in the plot) vs. the number of b-decay products (left) and $Z \to b\bar{b}$ mass-width for the same working points (right).

b-Fragmentation Modeling

The b-fragmentation function [477–480] was measured at LEP with a reported average value of $\langle x_B \rangle = 0.7092 \pm 0.0025$ [481]. Varying the fraction of the primordial b-quarks carried by the b-hadrons associated to the large-R jets to reproduce the LEP measurement error in the fragmentation function average value corresponds to a negligible additional systematic on the jet mass resolution. A recent ATLAS Conference Note [482] does not show any significant disagreement between the SHERPA predictions and data in the related quantities.

9.5.6 Total Systematic

The estimates for the discussed systematic uncertainties associated with the transfer of the constraint derived from the light-quark pairs in the control and validation regions and used for the JMR extraction to the $Z \to b\bar{b}$-dominated signal region are given in Table 9.4, given in units of the JMR nuisance parameter. The total systematic uncertainty obtained is ± 0.06 JMR units, but to be conservative, we inflate the error to ± 0.10 in the analysis. We add this uncertainty in quadrature to the statistical and extrapolation uncertainty, which can be seen in Figure 9.24. The values used in the signal region are summarized in Table C.14 of Appendix C.8. In the inclusive fit, the constraint is determined using the p_{T} of the VRL. Using the SRS instead has a $< 1\%$ effect on the value of μ_Z. In the systematic uncertainty pull plots, the external constraint is seen directly (just as if the CR were included in the fit). Table C.13 of Appendix C.8 shows the JMR NP post-fit mean and uncertainty from the SR. In the following chapter, we will discuss the global likelihood fit, and how systematic uncertainties enter into the signal determination.

Table 9.4: Sources of systematic uncertainty for the JMR constraint transfer given in units of the JMR nuisance parameter. To be conservative, the systematic is inflated to ±0.10 for the analysis.

Source	Transfer Uncertainty
b-Decay Track Multiplicity	±0.03
Semileptonic BR	±0.02
b-Tagging Efficiency	±0.05
Total	±0.06

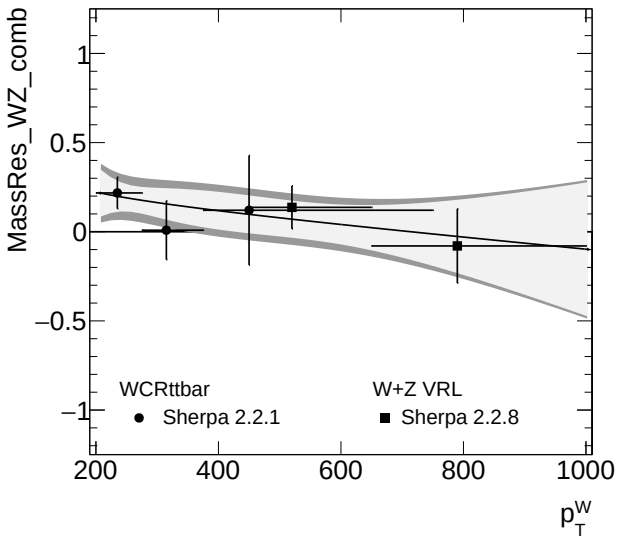


Figure 9.24: A summary of the constraints on the JMR values fit from the CR and VR as a function of the large-R jet p_{T}. The points are the measurements and the gray area shows the 68% C.L. from a fit with a function $a + bx + c/x$. The length of the horizontal bars for each point corresponds to the size of the p_{T} bin used in the fits while the marker position corresponds to the average p_{T} of the bin in the corresponding region.

Chapter 10

Likelihood Fit

As we've discussed in the previous chapters, the simulated signal and backgrounds (H, $Z+\text{jets}$, $W+\text{jets}$, and $t\bar{t}$) enter the fit as template histograms. Each of the simulated template histogram normalizations are unconstrained, or free parameters within each jet p_{T} region, or within a given fiducial volume common to the SRL, SRS, and $\text{CR}_{t\bar{t}}$. The normalization of the $t\bar{t}$ background template is determined in the $\text{CR}_{t\bar{t}}$, which is fit around the peak of the top-quark in the mass range of 140–200 GeV. The observable used is the distribution of large-R jet masses passing the b-tagging selection criteria for the respective regions. The fit range in the signal region is 70–210 GeV, with 5 GeV binning for a total of 28 mass bins. The QCD jet-mass distribution uses a polynomial exponential functional form, whose coefficients are free parameters and independent for each region. The parameters of interest in the fit are the signal-strength parameters μ ($\mu = \sigma_{\text{obs}}/\sigma_{\text{SM}}$) for the H+jet and $Z+\text{jets}$ ($H/Z \to b\bar{b}$) signal processes, in each of the transverse momentum ranges defined in Table 7.1. These signal strengths are obtained by normalizing the fitted number of signal events to the corresponding SM predictions. Table 10.1 shows the μ values, the floating normalizations of the H, Z, and $t\bar{t}$, that are extracted from the

fit in their respective regions. The number of parameters in the QCD background fit function was determined according to the studies presented in Section 9.4.3, and the number in each analysis region is shown in Table 10.2.

Table 10.1: A summary of the floating normalizations included in the statistical model for the Boosted $H \to b\bar{b}$ Fit. The QCD background has a purely data-driven model with free parameters not included in this table. $p_{\mathrm{T}}^{\mathrm{jet}}$ refers to the reconstructed jet p_{T} while p_{T}^{H} refers to the **truth** Higgs p_{T}. Inclusive and fiducial fits include all the p_{T} regions summed into one jet-mass distribution while maintaining the separation of the leading- and subleading-jet regions but with the upper bound on the $p_{\mathrm{T}}^{\mathrm{jet}}$ extended to 3 TeV. The parameters listed are correlated in different jet category regions (leading, subleading, and probe where applicable). The p_{T} for all $V+$jets and $t\bar{t}$ backgrounds is that of the large-R jet ($p_{\mathrm{T}}^{\mathrm{jet}}$). For the Higgs signal, the inclusive p_{T} region uses the large-R jet p_{T} ($p_{\mathrm{T}}^{\mathrm{jet}}$), and the fiducial and differential regions use the truth Higgs boson p_{T} (p_{T}^{H}).

Process	Inclusive	Fiducial	Differential			
			Backgrounds			
	$p_{\mathrm{T}}^{\mathrm{jet}} > 250/450$	> 450	250–450	450–650	650–1000	> 1000
$Z+$jets	μ_Z	μ_Z	$\mu_Z 0$	$\mu_Z 1$	$\mu_Z 2$	$\mu_Z 3$
$t\bar{t}$	$\mu_{t\bar{t}}$	$\mu_{t\bar{t}}$	$\mu_{t\bar{t}}^{0}$	$\mu_{t\bar{t}}^{1}$	$\mu_{t\bar{t}}^{2}$	$\mu_{t\bar{t}}^{3}$
			Signal			
	$p_{\mathrm{T}}^{\mathrm{jet}} > 250/450$	$p_{\mathrm{T}}^{H} > 450$	300–450*	450–650	650–1000†	> 1000†
H (all)	μ_H	μ_H	μ_H^{pT1}	μ_H^{pT2}	μ_H^{pT3}	μ_H^{pT4}

* Note that $pT0$ is the category of events with $p_{\mathrm{T}}^{H} < 300$ GeV which are held at the SM expectation within theoretical and experimental errors.

† The upper truth p_{T}^{H} cut is released to allow the small fraction of events with truth $p_{\mathrm{T}}^{H} > 1000$ GeV and reconstructed jet $p_{\mathrm{T}} < 1000$ GeV to be included.

Table 10.2: The number of free parameters for the QCD model in each region. Note, in Section 9.4.3, *number of parameters* does not include the normalization factor. To avoid confusion, the number of free parameters is shown as $N+1$ where N, the order of the polynomial, is determined from the LLR and spurious signal tests as described in Section 9.4.3.

	Inclusive	Fiducial	Differential			
Candidate Jet	$> 250/450$	> 450	250–450	450–650	650–1000	> 1000
Leading	5+1	5+1	–	5+1	4+1	4+1
Subleading	5+1	5+1	5+1	4+1	5+1	–

In an idealized world, we would be able to calculate our cross section to all orders with infinite detector precision; no electronic noise, no pileup, and unlimited statistics for data and MC. Unfortunately we must contend with the fact that we have limited knowledge and statistics. Statistical and systematic uncertainties allow us to incorporate our ignorance into the fit to data to allow for a meaningful interpretation of the results. Generally, any attempt to measure some parameter will be affected by these uncertainties.

Statistical effects are quite straightforward; if we were to repeat our experiment (or simulation), random effects would change the obtained result. This can come from the inherent Poisson variability for observing a number of counts n, or effects from the limited accuracy and precision of our measurement devices. An increase in the data or simulation (better yet, both) will decrease the statistical uncertainty.

On the other hand, systematic uncertainties come from a variety of sources and require extreme care and consideration. Systematic effects can shift the measurements from their true value, altering the yields and shapes of the observable, and need to be accounted for [483]. We cannot calculate the cross section to infinite[1] order, some parameters are inherently not well known, etc., and thus we face theoretical uncertainties. The detector components are inherently noisy, and they have limited spatial resolution and response, etc., further obscuring our knowledge. In the fit, these systematic effects are translated into nuisance parameters (NPs), which still need to be determined, but are not the parameters of interest. The magnitude of the systematic effects in our parameter determination is assessed by fitting the data with different values of the NPs, and observing how the result changes when the NP parameters are varied by their $\pm 1\,\sigma$ uncertainty [483].

[1] Not even close

All of the systematic uncertainties, or nuisance parameters (NP), enter the fit as template histograms, which can vary from the nominal Monte Carlo template in shape and normalization. There are many systematic uncertainties, related to jets, b-tagging, and modeling in simulation, which are discussed in more detail in Section 10.2. To give a sense of the affect from these systematic uncertainties, let's take two relevant examples from the analysis of a jet systematic, the jet mass resolution (JMR) and jet mass scale uncertainties. The JMR NP incorporates our lack of knowledge about the width of a resonance in the Monte Carlo, altering the shape of the large-R jet-mass distribution and affecting the agreement between the model and data by varying the JMR from the nominal value by its uncertainty. Similarly, the jet mass scale (JMS) uncertainty incorporates our lack of knowledge about the peak position of the resonance. The effects of both uncertainties can be seen in Figure 10.1, which shows the MC templates for the JMR and JMS variations on the $\text{WCR}_{t\bar{t}}$ region. These effects can range in severity for each process and region, but the principle remains the same. The various systematic uncertainties must be incorporated into the fit simultaneously so that correlations, which are bound to exist, can be understood. To do this, we need some statistical procedure.

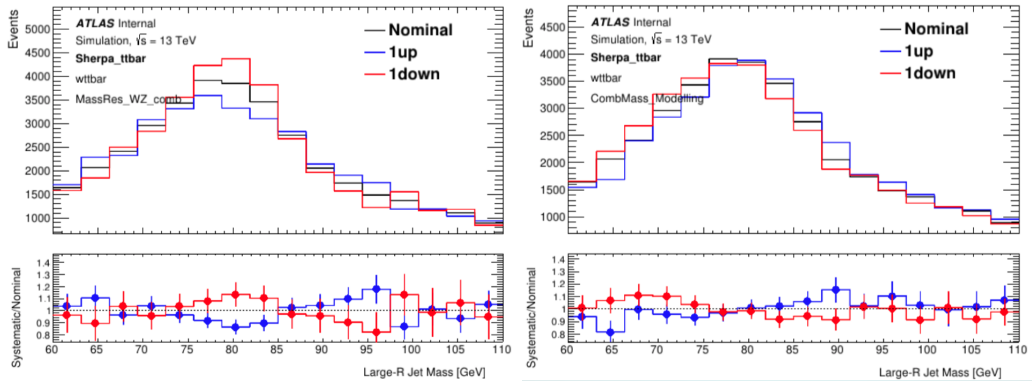


Figure 10.1: The JMR (left) and JMS (right) $\pm 1\,\sigma$ (*1up* and *1down*) m_J distributions (red and blue) compared to the nominal (black).

10.1 Likelihoood Definition

There are multiple statistical tests available [257, 258, 483–486] to test a background-only and signal + background hypotheses to the data. There is considerable discussion regarding the advantages and disadvantages for various methods [487], but we will avoid assigning one as better than another. Among the most widely used methods in statistics [488], as well as in ATLAS and CMS [489], is a maximum likelihood ratio test. Wilks' theorem [466] states that the asymptotic null distributions of the likelihood ratio belong to the χ^2 family of functions [490]. In our analysis, we use the profile likelihood ratio [462], which profiles the signal strengths μ.

The likelihood is the function that returns the value of the probability density function (PDF) evaluated on the observed data sample for a given value of the unknown parameters. If we have the measured values of n random variables $\boldsymbol{x} = x_1, \ldots, x_n$ (*i.e.,* our observable values in the different regions) and our PDF depends on the unknown parameters μ (the parameters of interest, of which there can be more than one.) and $\boldsymbol{\theta} = \theta_1, \ldots, \theta_m$ (the nuisance parameters), the likelihood function has the general form [258]:

$$L(\boldsymbol{x}; \mu, \boldsymbol{\theta}) = f(\boldsymbol{x}; \mu, \boldsymbol{\theta}), \tag{10.1}$$

where f is the joint PDF of the random variables $\boldsymbol{x}$. The *maximum likelihood estimator* of the various parameters μ and $\boldsymbol{\theta}$ is the function that returns the values of the unknown parameters for which the likelihood function, given the data $\boldsymbol{x}$, is a maximum. For a counting experiment such as ours, the likelihood takes the form

of a Poisson distribution. The systematics are included as Gaussian PDFs which constitute an effective penalty term, penalizing (lowering) the likelihood for values of the NPs as they deviate from the nominal, or central value.

In our analysis, we use a binned likelihood function whose maximum corresponds to the best description of data. It is defined as the product over all mass bins of the Poisson probability to observe N_b^{obs} data events given a prediction of $N_b^{\text{exp}}(\mu, \mathbf{k}, \boldsymbol{\theta})$ events in a certain mass bin i [491]:

$$L(\mu, \boldsymbol{k}, \boldsymbol{\theta}) = \prod_{i \in \text{ bins}} \frac{\left(N_i^{\text{exp}}(\mu, \boldsymbol{k}, \boldsymbol{\theta})\right)^{N_i^{\text{data}}}}{N_i^{\text{data}}\,!} \cdot e^{-N_i^{\text{exp}}(\mu, \boldsymbol{k}, \boldsymbol{\theta})}. \tag{10.2}$$

Bin widths are set to 5 GeV, necessitating technical advancements to fit an analytic function (for the multijet background) to a wide-binned dataset [492–494]. In this likelihood, the number of predicted events is made dependent on three sets of parameters: the signal strength μ, the background scale factors $\mathbf{k} = \{k_1, ..., k_j\}$, and the nuisance parameters $\boldsymbol{\theta} = \{\theta_1, \ldots, \theta_l\}$, as follows

$$N_i^{\text{exp}}(\mu, \mathbf{k}, \boldsymbol{\theta}) = \mu \cdot N_{i,\text{sig}}^{\text{exp}}(\boldsymbol{\theta}) + \sum_{b \in \text{ bkg}} k_b \cdot N_{i,b}^{\text{exp}}(\boldsymbol{\theta}). \tag{10.3}$$

The μ parameter here is the signal strength for the Higgs boson. The k_b scale factors are the normalizations for the various backgrounds to the Higgs signal, and the normalizations for the $Z+$jets and $t\bar{t}$ processes are designated with a μ value (*i.e.* $k_Z \equiv \mu_Z$ and $k_{t\bar{t}} \equiv \mu_{t\bar{t}}$, respectively). Signal yields are extracted by minimizing the negative logarithm of the likelihood function in Equation 10.2 using the RooStats framework [495]. The parameter of interest $\mu = \sigma/\sigma_{\text{SM}}$ is the ratio between the measured and the expected signal cross sections. As each scale factor does for its

associated background component, the signal strength scales the amount of signal linearly without any prior constraint or penalty in the likelihood function.

The nuisance parameters θ_i encode the dependence of the prediction on systematic uncertainties into continuous parameters in the likelihood. The prior knowledge of these parameters is reflected in the Gaussian penalty term $\text{Gauss}(0|\theta_i, 1)$, which is added to the log likelihood for each nuisance parameter to disfavor large displacements from the nominal value. The JMR constraints obtained from the $\text{WCR}_{t\bar{t}}$ and the VRL regions are also included as Gaussian PDF priors. The unconstrained or *free* parameters μ control the normalizations of the MC templates within each jet-p_{T} region or within a given fiducial volume and is common to the SRL, SRS, and $\text{CR}_{t\bar{t}}$.

The nominal fit result, in terms of μ and σ_μ, is obtained by maximizing the likelihood function with respect to all parameters in what is known as the maximized log-likelihood value, or MLL. The test statistic for the profile likelihood, q_μ, is defined as:

$$q_\mu = -2 \ln \left(\frac{L(\mu, \hat{\hat{\boldsymbol{k}}}, \hat{\hat{\boldsymbol{\theta}}}_\mu)}{L(\hat{\mu}, \hat{\boldsymbol{k}}, \hat{\boldsymbol{\theta}})} \right). \tag{10.4}$$

where $\hat{\mu}$ and $\hat{\theta}$ are the parameters that maximize the likelihood (with the constraint $0 \leq \hat{\mu} \leq \mu$), and $\hat{\hat{\theta}}_\mu$ are the nuisance parameter values which maximize the likelihood for a given value of μ. The test statistic is used to measure the compatibility of the background-only model with the observed data, extract the local p_0 value, and, if no hint of signal is found in this procedure, to derive the exclusion intervals using the CL_s method [462, 496].

10.2 Systematic Uncertainties

The major uncertainties in this analysis result from the jet mass resolution (JMR) modeling and jet mass scale (JMS) calibration. We used the most recent recommendations [470] available for the large-R-jet jet energy resolution (JER), JMS, and JMR systematic uncertainties. We chose the category reduction scheme [467] configuration following the updated large-R-jet jet energy scale (JES) recommendations. This configuration contains roughly 30 nuisance parameters and is aimed at analyses which intend to perform combinations. The JES and JMS uncertainties are defined by a ratio of calorimeter-based and track-based measurements [383] in dijet data, and the two systematics are divided into 23 and 6 separate components, respectively, to account for different sources of uncertainties. The JMS agreement between data and Monte Carlo is within the systematic uncertainties, but it displays a clear process and p_{T} dependence. During the blinded phase of the analysis, the $t\bar{t}$ JMS uncertainty was decorrelated from the other processes per the request of the Hbb Group [497] following observations in the Boosted $VH(b\bar{b})$ analysis [129, 498]. The dominant component of the reconstructed mass scale uncertainty is further separated to act independently on all processes ($t\bar{t}$, $V+$jets, and H) and in all analysis p_{T} bins. The jet-kinematic observables are smeared in simulation to asses the impact of JES and JMR uncertainties. The JER has an 2% absolute uncertainty, consistent with previous studies for trimmed jets [396, 499]. The JMR systematic carries a 20% *relative* uncertainty. The JetETMiss Group recommends that the JMR systematic be decorrelated between the top, W/Z, and Higgs processes. As discussed in Section 9.5, we derive an external JMR constraint for $V+$jets using the $\mathrm{WCR}_{t\bar{t}}$ and VRL regions in each p_{T} bin of the analysis, which is included as a Gaussian PDF

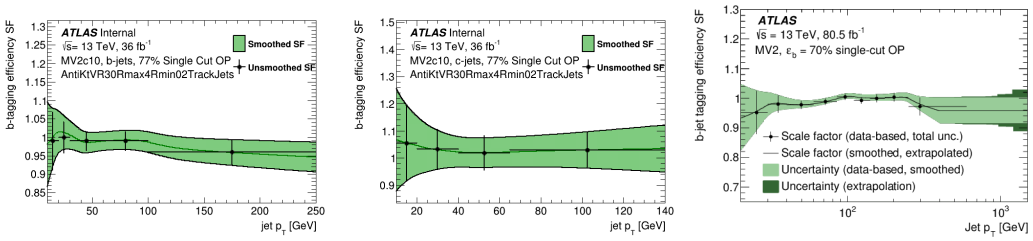


Figure 10.2: The left two plots show the FTag scale factors vs. VR track jet p_{T}. The right-most plot shows the extrapolation uncertainty taken from Ref [503] for calo jets to give the reader a sense of the expected impact.

prior in the global likelihood. Therefore, the JMR systematic is decorrelated among processes and p_{T} regions.

The impact of uncertainties on b-tagging rates for b-, c-, and light-flavor jets are determined separately in various kinematic regions [244, 500, 501]. Each flavor category uncertainty is decomposed independently in the eigenvector scheme. Figure 10.2 shows the calibration scale factors for the b- and c-jets binned in jet p_{T}; 100–250 GeV for b-jets and 65–140 GeV for c-jets. The extrapolation uncertainty uses the MC to translate the calibration error above the kinematic reach of the data [502] and has a sizable impact above 400 GeV.

The uncertainties on b-tagging efficiency scale factors and JES are a minor contribution to the total systematic uncertainty. The remaining uncertainties, including the modeling uncertainties for triggering, reconstruction, identification, and isolation-range for muons [365] are negligible. The uncertainty concerning the total integrated luminosity of 136 fb^{-1} and 139 fb^{-1} for the jet- and muon-triggered data, respectively, is 1.7% [283, 504]. A more complete list of the systematic uncertainties studied in the analysis is available in Tables C.15–C.19 of Appendix C.9. Table 10.3 summarizes the systematic uncertainties considered in the likelihood fit. In addition, uncertainties due to the limited number of events in the simulated samples used

for the background predictions are parameterized using the Beestow-Barlow technique [505]. Systematic variations yielding large statistical fluctuations are smoothed using custom algorithms which also remove variations resulting in effects below 2% within each region.

In the following chapter, we will discuss the results of the ATLAS Boosted $H \to b\bar{b}$ Analysis in the inclusive, fiducial, and differential regions. The main contributions to the total systematic uncertainties in each p_{T} region will also be shown. Lastly, we will discuss the results from the CMS analysis in the same channel to compare the differential results.

Table 10.3: A summary of the systematic uncertainties included within the profile likelihood for the H and Z signal strength extraction. For a given uncertainty, the second column lists each process which has independent nuisance parameters within the likelihood. The third column describes how the systematic uncertainty is correlated across regions: *all* indicates a fully correlated parameter, p_T *bins* indicates a decorrelation between the analysis p_{T} bins, and *LS* means it is decorrelated between the SRL and SRS. For the inclusive analysis, *bins* does not apply, and should be understood to mean the same as *all*. The fourth column describes the change induced by the parameter. S means the m_{J} shape will change while N denotes parameters which change the normalization and can result in a migration of events between regions.

Description	Processes	Category	Effect
Experimental Systematic Uncertainties			
JMR	$t\bar{t}$, $V+$jets, H	p_{T} bins	N+S
JMS (dominant)	$t\bar{t}$, $V+$jets, H	p_{T} bins	N+S
JMS (rest)	$t\bar{t}$, $V+$jets $+H$	all	N+S
Jet energy scale	all*	all	N+S
Jet energy resolution	all	all	N+S
b-tag efficiency b-jets	all	all	N+S
b-tag efficiency c-jets	all	all	N+S
b-tag efficiency l-jets	all	all	N+S
Modeling Systematic Uncertainties			
Cross section and acceptance	$W+$jets, $W(\ell\nu)$, t	all	N
Renormalization and factorization scale	$V+$jets	all	N+S
Parton shower model	$t\bar{t}$	all	N+S
Matrix element calculation	$t\bar{t}$	all	N+S
Initial-/Final-state radiation	$t\bar{t}$	all	N+S
Acceptance	H	all	N
NLO EW corrections	VBF $+VH$ $+ttH$	all	N
	$V+$jets	all	N
Spurious signal	H	p_{T}^{H} bins $\times$ LS	N
	$Z+$jets †	p_{T}^{Z} bins $\times$ LS	N

* Two minor components separately apply to $t\bar{t}$ and $V+$jets events.

† The spurious signal uncertainty is only applied to $Z+$jets when the procedure to extract signal strengths in fiducial volumes is tested using $Z+$jets events in the SR.

Chapter 11

Results

We designed the analysis regions, as summarized in Table 8.1, to probe Higgs boson production with considerable transverse momentum. They include one all-encompassing region to measure the H signal strength, one region for a fiducial cross-section measurement, and four bins for a differential measurement. All of the Higgs production modes are considered for the signal-strength extraction. For Higgs signal strength measurements in the fiducial and differential regions, the signal events are defined by requirements on the generator truth record for the Higgs boson rapidity y_H and transverse momentum p_{T}^H. The analysis jet-p_{T} bins align well with the p_{T}^H-defined fiducial volumes and the yield of signal events outside the targeted fiducial volume(s) are constrained to their SM prediction within the theoretical and experimental uncertainties. The cross sections are derived from the fitted signal yields divided by the integrated luminosity, corrected by the product of the estimated selection efficiency and fiducial acceptance. Using the same category definitions, differential cross-section measurements of $V+\mathrm{jets}$ production in the VRL and $Z+\mathrm{jets}$ production in the SR validate the method.

Table 11.1: Expected and observed values of the signal strengths for the H, Z and $t\bar{t}$ components in the inclusive fit. The value of $\mu_{t\bar{t}}$ obtained is in agreement with unfolded measurements of $t\bar{t}$ events in a similar kinematic phase space [405].

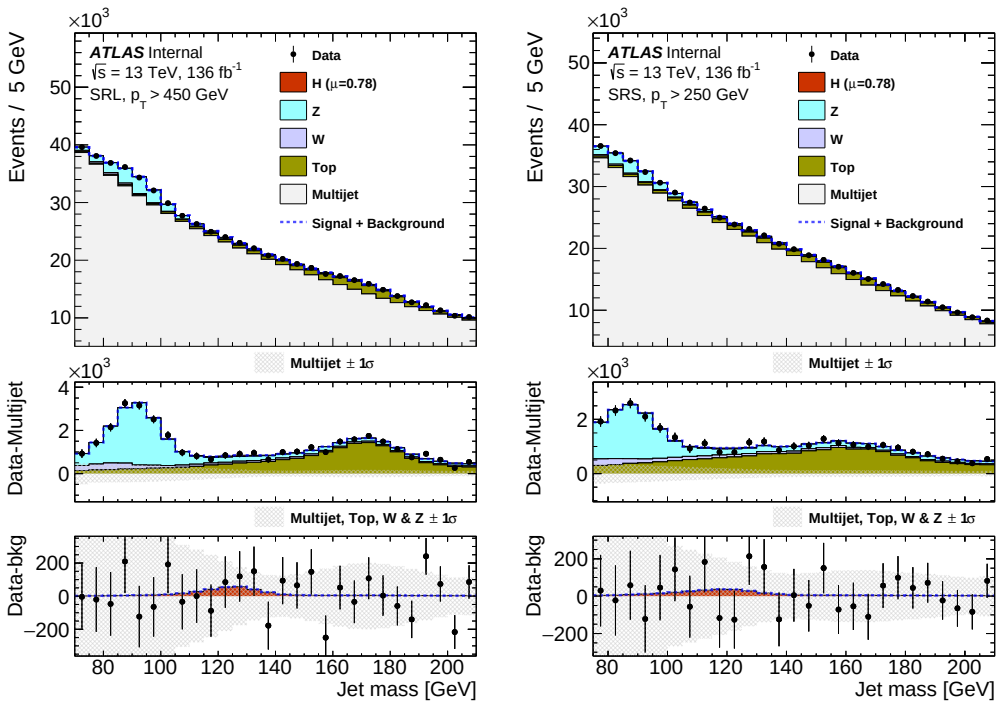
Result	μ_H	μ_Z	$\mu_{t\bar{t}}$
Expected	1.0 ± 3.2	1.00 ± 0.17	1.00 ± 0.07
Observed	0.8 ± 3.2	1.29 ± 0.22	0.80 ± 0.06

11.1 Inclusive Region

The fit to the inclusive region yields a Higgs boson signal strength for the combination of the SRL, SRS, and $\text{CR}_{t\bar{t}}$ is $\mu_H = 0.8 \pm 3.2$. The fit χ^2 probability values are 0.19 and 0.67 for the SRL and SRS, respectively. The post-fit SRL and SRS jet mass distributions are shown in Figure 11.1 Signal strength results are summarized in Table 11.1 and event yields in Table 11.2. The Higgs boson signal strength uncertainty is statistically dominated, with the leading source of systematic uncertainties being the jet mass resolution and jet mass scale. Figure 11.2 shows the ranking plot. JMR is the leading systematic due to its correlation with $V+$jets (see Section 9.5). The next few leading systematics relate to the Top background, followed by the Higgs JMR systematic nuisance parameter, but the small Higgs signal yield causes it to be too small to play a large role in the QCD minimum in the fit.

Table 11.2: Event yields and associated uncertainties after the global likelihood fit in the inclusive region. Entries with "–" indicate that the process does not contribute to the region.

Process	SRL	SRS	$\mathrm{CR}_{t\bar{t}}$
Multijet	$590{,}800 \pm 4200$	$529{,}400 \pm 3500$	–
$Z+\text{jets}$	$16{,}100 \pm 2800$	$12{,}000 \pm 2100$	–
$W+\text{jets}$	3050 ± 720	2510 ± 500	–
Top	$16{,}200 \pm 1900$	$15{,}900 \pm 2000$	3737 ± 68
$W(\ell\nu)$	–	–	53 ± 16
H	400 ± 1500	300 ± 1300	–
Total	$626{,}530 \pm 810$	$560{,}090 \pm 770$	3790 ± 66
Data	626,532	560,083	3791



(a) SRL: $p_{\mathrm{T}} > 450\,\mathrm{GeV}$ (b) SRS: $p_{\mathrm{T}} > 250\,\mathrm{GeV}$

Figure 11.1: Post-fit signal jet mass distributions with the various components for the inclusive SRL (a) and SRS (b) regions. The middle panel for each plot shows the distributions after subtracting the QCD multijet distribution. The shaded areas indicate the 68% CL for the multijet background from the fitted parameters and normalizations of the exponential polynomials. The lower panels in each plot shows the distributions after subtracting the multijet, $V+\text{jets}$, and Top background processes. The shaded areas show the 68% CL for all background processes.

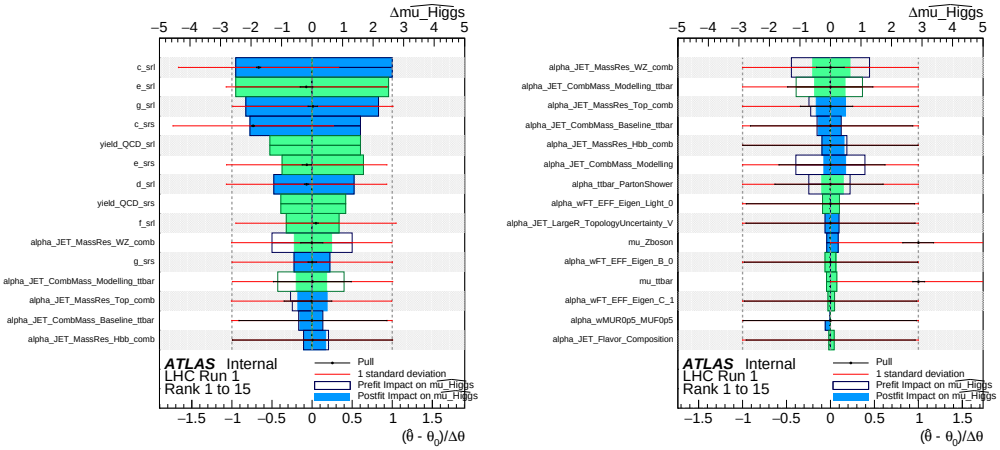


Figure 11.2: Ranking plot for all NPs (left) and ranking plot for all NPs excluding the QCD background function parameters (right) for the inclusive Higgs fit. The QCD fit parameters are labeled starting with *c-h* The precision of the result is largely limited by the statistics of the sample.

11.2 Fiducial Region

The fiducial region determines the Higgs boson yield and cross section in the phase space defined by the Higgs boson rapidity range $|y_H| < 2.0$ and transverse momentum $p_{\mathrm{T}}^{H} > 450\,\mathrm{GeV}$. Compared to the inclusive measurement discussed above, the fiducial region does not include the SRS region below 450 GeV. The acceptance times efficiency values for the different SM Higgs boson production modes are given in Table 11.3.

Two Higgs boson mass templates are used in the fiducial fit for the signal region. The first describes the mass distribution of events with signal jet and Higgs boson p_{T} ($p_{\mathrm{T}}^{\mathrm{jet}}$ and p_{T}^{H}, respectively) above the chosen cut; the second, those with signal jet p_{T} above the cut, but p_{T}^{H} below said cut. The first component accounts for more than 80% of the Higgs boson signal selected and the yield is left free in the fit. The second is constrained to the SM value within theoretical and experimental uncertainties. The latter contribution tends to have a broader mass spectrum, which is shifted to higher values.

This fit procedure is first tested with $W \to q\bar{q}'$ and $Z \to q\bar{q}$ in the VR and $Z \to b\bar{b}$ in the SR. For these tests, the W and Z mass templates are structured similarly to those of the Higgs boson described above. The Higgs boson yields are

Table 11.3: Signal acceptance times efficiency for the signal regions in the fiducial measurements.

Process	$p_{\mathrm{T}}^{H} > 450\,\mathrm{GeV}$, $\lvert y_H \rvert < 2$
All	0.24
ggF	0.26
VH	0.27
VBF	0.22
ttH	0.20

Table 11.4: Expected and observed values of the signal strengths for the H, Z and $t\bar{t}$ components in the fiducial fit. The value for μ_H^r efers to the fiducial volume $|y_H| < 2.0$ and $p_{\mathrm{T}}^H > 450\,\mathrm{GeV}$, while those for μ_Z and $\mu_{t\bar{t}}$ refer to regions of jet p_{T} ($p_{\mathrm{T,J}}$) above 450 GeV. The values of $\mu_{t\bar{t}}$ obtained are in agreement with unfolded measurements of $t\bar{t}$ events in a similar kinematic phase space [405].

Result	μ_H	μ_Z	$\mu_{t\bar{t}}$
Expected	1.0 ± 3.4	1.00 ± 0.18	1.00 ± 0.08
Observed	-0.1 ± 3.5	1.30 ± 0.22	0.75 ± 0.06

kept fixed to the SM expectations in the fit. In the VR, the fitted signal strength for $V+$jets with $p_{\mathrm{T}}^V > 450\,\mathrm{GeV}$ is $\mu_V = 1.01 \pm 0.09$. In the SR, the μ_Z values for Z events with $p_{\mathrm{T}}^Z > 450\,\mathrm{GeV}$ is 1.35 ± 0.23. These results are in agreement with the SM.

When extracting the Higgs boson signal strength, the likelihood fit for $p_{\mathrm{T}}^H > 450\,\mathrm{GeV}$ results in a signal strength of $\mu_H = -.01 \pm 3.5$. These yields correspond to Higgs boson production cross-section values in the fiducial region at the 95% confidence level (CL) upper limits of

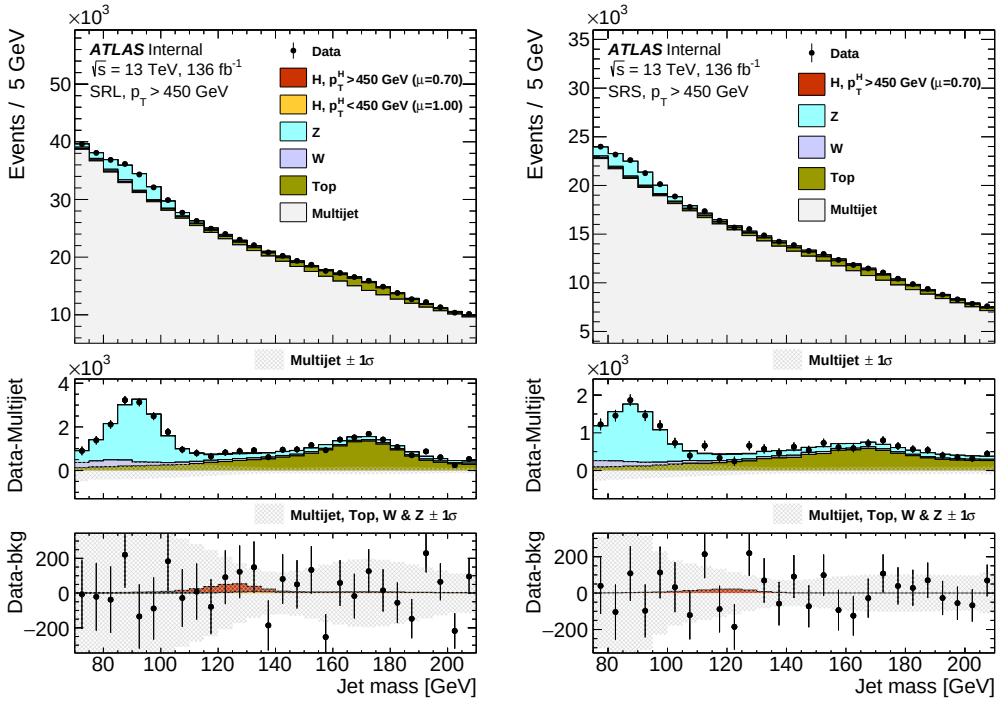
$$\sigma_H(p_{\mathrm{T}}^H > 450\,\mathrm{GeV}) < 115\,\mathrm{fb}$$

The extracted signal strengths are summarized in Table 11.4.

The post-fit SRL and SRS jet-mass distributions are shown in Figure 11.3. Results are summarized in Table 11.4. The uncertainties are statistically dominated and the contributions of the main categories of systematic uncertainties are given in Table 11.5. The jet uncertainties give the largest contribution of the systematic

uncertainties, driven by the JMS effects at lower p_{T} and by JMR above 1 TeV where the constraints are looser due to the extrapolation uncertainties (see Figure 9.24).

Similarly, the flavor-tagging uncertainties increase above 1 TeV due to the extrapolation from the p_{T} range of the calibration dataset.



(a) SRL: $p_{\mathrm{T}}^{H} > 450\,\mathrm{GeV}$ (b) SRS: $p_{\mathrm{T}}^{H} > 450\,\mathrm{GeV}$

Figure 11.3: Post-fit signal jet-mass distributions with the various components for the fiducial regions with $p_{\mathrm{T}}^{H} > 450\,\mathrm{GeV}$ in the SRL (a) and SRS (b). The middle panels for each plot show the distributions after subtraction of the multijet distribution. The shaded areas indicate the 68% CL for the multijet background from the fitted parameters and normalizations of the exponential polynomials. The lower panels in each plot show the distributions after subtraction of the fitted background processes: multijet, V + jets and Top. The shaded areas indicate the 68% CL for all background processes. Contributions below 0.5 per mil of the total yield are not shown.

Table 11.5: Contributions to the systematic uncertainties for the fiducial signal strength measurements. The total uncertainty is also given for comparison.

Uncertainty Contribution	$p_{\mathrm{T}}^{H} > 450\,\mathrm{GeV}$
Total	3.5
Statistical	2.6
Systematic	2.3
Jet Systematics	1.8
Modeling and Theory Systs.	0.8
Flavor Tagging Systs.	0.2

11.3 Differential Regions

The differential regions, defined in Table 8.1 of Section 8.1.1, are used to measure the differential spectrum of the Higgs boson transverse momentum, where possible deviations from the SM predictions could manifest with an amplitude increasing with p_{T}^{H}. In an extension of the procedure adopted for the fiducial measurements, several Higgs boson mass templates corresponding to the same jet p_{T}, but different p_{T}^{H} ranges, are used in the fits. Figure 11.4 presents the signal yield in each reconstructed event category for each fiducial volume and the corresponding fraction of signal events.

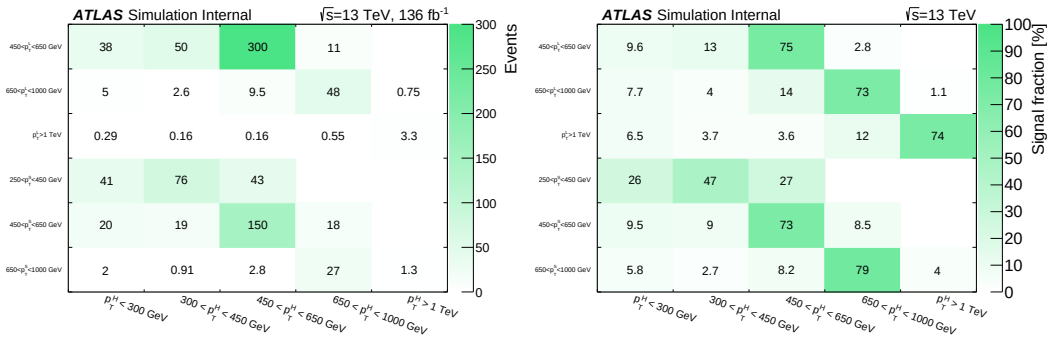
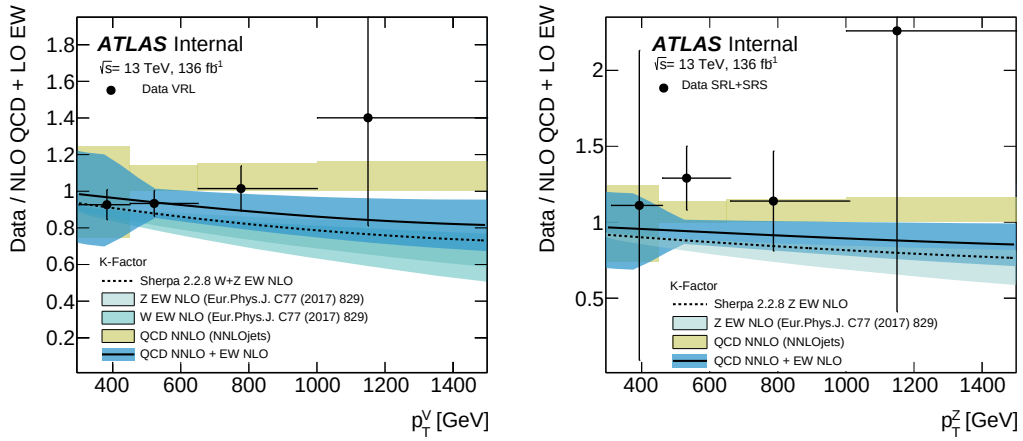


Figure 11.4: For each of the p_{T}^{H} differential regions (horizontal axis), the expected signal event yield for all Higgs boson events (left) and the fraction of signal in percent (right) in each jet-p_{T} region (vertical axis) is shown. The leading jet p_{T} of the SRL is denoted $p_{\mathrm{T}}^{\mathrm{L}}$ and the subleading jet p_{T} of the SRS is denoted as $p_{\mathrm{T}}^{\mathrm{S}}$.

The procedure is tested with $W \to q\bar{q}'$ and $Z \to q\bar{q}$ in the VR, and tested with $Z \to b\bar{b}$ in the SR. For the differential $V+$jets analysis, the larger statistics of the VRL allow the dataset to be divided into five slices. The fit is performed independently on each slice, and the results combined. In the SR, the differential Z fit is performed in the SRL, SRS, and $\mathrm{CR}_{t\bar{t}}$ regions with the Higgs boson contribution fixed to the SM prediction. Results of the two fits are shown in Figure 11.5, where they are compared to the predictions for the EW NLO and QCD NNLO corrections as a function of the reconstructed p_{T}^{V}. Both results agree with SM expectations.



(a) $V+$jets in the VRL region. (b) $Z+$jets in the SR region.

Figure 11.5: Comparison of differential fit signal strengths for $V+$jets in the VRL (a) and $Z+$jets in the SR (b). The signal strength is calculated with respect to the prediction at NLO QCD and LO EW accuracy. They are compared to the NLO EW correction provided by Sherpa, the NNLO QCD correction provided by the NNLOJET group, and their product. The Higgs boson yields are kept fixed to the SM expectation when extracting the $Z+$jets signal strength within the fiducial volumes.

To extract the Higgs boson signal strength, the ten differential SR and CR regions defined in Table 8.1 are simultaneously fitted with the corresponding systematic uncertainties shown in Table 10.3. Results are summarized in Table 11.6 and Figure 11.6. The four Higgs boson signal strengths are compatible with a p-value of 0.53. The acceptance times efficiency values for the different Higgs boson production

Table 11.6: Expected and observed values of the signal strengths for the H, Z, and $t\bar{t}$ components in the differential fits. The value for μ_H refers to a fiducial p_{T}^H volume, while those for μ_Z and $\mu_{t\bar{t}}$ pertain to the corresponding jet p_{T} regions. The values of $\mu_{t\bar{t}}$ obtained are in agreement with unfolded measurements of $t\bar{t}$ events in a similar kinematic phase space [405].

p_{T}^H [GeV]	μ_H	
	Exp.	Obs.
300–450	1.0 ±18	-6.0 ±18
450–650	1.0 ±3.3	-3.0 ±5.0
650–1000	1.0 ±6.2	5.2 ±6.5
> 1000	1.0 ±30	18 ±32

Jet p_{T} [GeV]	μ_Z		$\mu_{t\bar{t}}$	
	Exp.	Obs.	Exp.	Obs.
250–450	1.00 ±1.05	1.76 ±1.13	1.0 ±0.07	0.85 ±0.06
450–650	1.00 ±0.17	1.28 ±0.22	1.0 ±0.07	0.76 ±0.06
650–1000	1.00 ±0.33	1.35 ±0.42	1.0 ±0.09	0.74 ±0.08
> 1000	1.00 ±1.62	2.35 ±1.70	1.0 ±0.22	0.57 ±0.18

Table 11.7: Signal acceptance times efficiency for the signal regions in the differential measurements. The ranges correspond to the Higgs boson transverse momentum p_{T}^H. For events with $p_{\mathrm{T}}^H < 300\,\mathrm{GeV}$, the acceptance times efficiency is less than 0.1×10^{-2}.

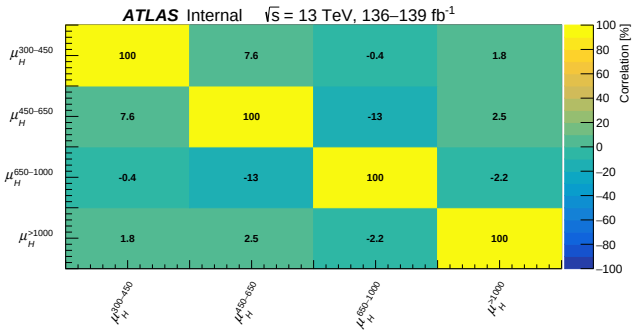
Process	300–450	450–650	650–1000	> 1000
All	1.3×10^{-2}	0.23	0.31	0.23
ggF	0.7×10^{-2}	0.25	0.35	0.28
VBF	0.4×10^{-2}	0.21	0.32	0.25
VH	1.7×10^{-2}	0.26	0.30	0.20
ttH	4.7×10^{-2}	0.19	0.24	0.19

processes are given in Table 11.7. The post-fit jet-mass distribution from the most sensitive category in each jet p_{T} bin is shown in Figure 11.7.

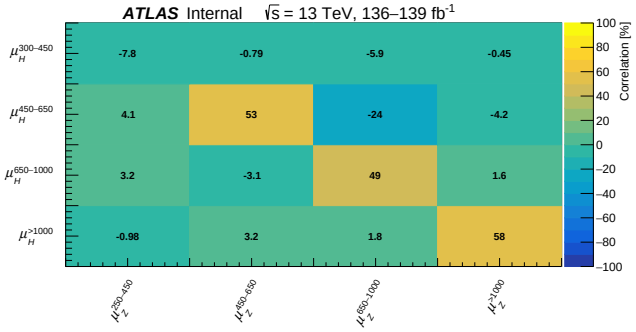
The resulting Higgs boson production cross section for $p_{\mathrm{T}}^H > 1000\,\mathrm{GeV}$ is

$$\sigma_H(p_{\mathrm{T}}^H > 1\,\mathrm{TeV}) = 2.3 \pm 3.9\,(\mathrm{stat.}) \pm 1.3\,(\mathrm{syst.}) \pm 0.5\,(\mathrm{theory})\,\mathrm{fb}.$$

The differential results corresponding to the following 95% CL upper limits on the



(a) μ_H



(b) μ_Z

Figure 11.6: Correlations among the four Higgs boson signal strengths, and between the four Higgs boson and $Z+$jets signal strengths. The Higgs boson signal strengths μ_H are labeled with the corresponding p_{T}^H range as a superscript. The $Z+$jets signal strengths μ_Z are labeled with the corresponding jet p_{T} range as a superscript.

Higgs boson differential production cross sections:

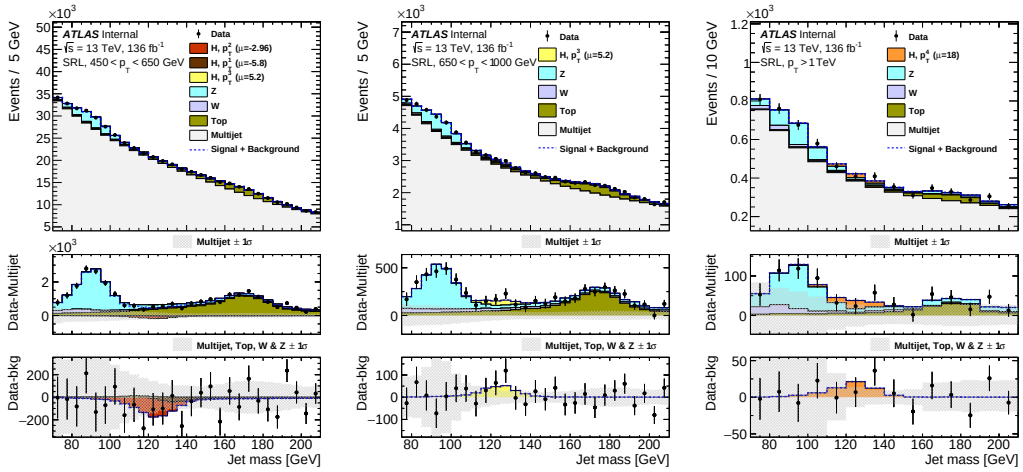
$$\sigma_H(300 < p_{\mathrm{T}}^H < 450\,\mathrm{GeV}) < 2.9\,\mathrm{pb},$$

$$\sigma_H(450 < p_{\mathrm{T}}^H < 650\,\mathrm{GeV}) < 89\,\mathrm{fb},$$

$$\sigma_H(650 < p_{\mathrm{T}}^H < 1000\,\mathrm{GeV}) < 39\,\mathrm{fb},$$

$$\sigma_H(p_{\mathrm{T}}^H > 1000\,\mathrm{GeV}) < 9.6\,\mathrm{fb}.$$

The uncertainties are statistically dominated and contributions of the main categories of systematic uncertainties are summarized in Table 11.8. The dominant source of systematic uncertainty results from the jet uncertainties, driven largely by



(a) SRL: $450 < p_{\mathrm{T}} < 650\,\mathrm{GeV}$ (b) SRL: $650 < p_{\mathrm{T}} < 1000\,\mathrm{GeV}$ (c) SRL: $p_{\mathrm{T}} > 1000\,\mathrm{GeV}$

Figure 11.7: Post-fit signal jet mass distributions for the differential signal region defined by the leading jet with $450 < p_{\mathrm{T}} < 650\,\mathrm{GeV}$ (a), $650 < p_{\mathrm{T}} < 1000\,\mathrm{GeV}$ (b), and $p_{\mathrm{T}} > 1000\,\mathrm{GeV}$ (c) with the various components. The middle panels in each plot show the distributions after subtraction of the multijet distribution. The shaded areas indicate the 68% CL for the multijet background from the fitted parameters and normalizations of the exponential polynomials. The lower panels in each plot show the distributions after subtraction of the fitted background processes: multijet, $V+$jets and Top. The shaded areas indicate the 68% CL for all background processes. The four fiducial volumes are labeled p_{T}^{0}–p_{T}^{4} corresponding to $p_{\mathrm{T}}^{H} < 300\,\mathrm{GeV}$, 300–450 GeV, 450–650 GeV, 650–1000 GeV and $> 1000\,\mathrm{GeV}$, respectively. The p_{T}^{0} event yield is constrained to its SM value within the theoretical and experimental uncertainties, and free parameters act independently on the remaining three volumes. Contributions below 0.5 per mil of the total yield are not shown.

JMS effects. The impact of the JMR uncertainty increases above 1 TeV because the JMR measurement constraints are looser due to the extrapolation uncertainties (see Figure 9.24). In the same kinematical region, the flavor-tagging uncertainties increase due to the extrapolation from the p_{T} range of the calibration regions. The $t\bar{t}$ modeling systematic uncertainties are more relevant in the first jet-p_{T} bin and decrease above 450 GeV, where the top-quark decay products become more collimated, thus reducing the contamination around the Higgs boson mass peak. Figure 11.8 shows the nuisance parameter ranking plot in the differential regions. When focusing on the constrained systematic uncertainties, the $t\bar{t}$-related systematics are more prevalent

Table 11.8: Contributions to the systematic uncertainties for the differential signal strength measurements. The ranges correspond to the Higgs boson transverse momentum p_{T}^{H}. The total uncertainty reported in the top portion of Table 11.6 is also given for comparison.

Uncertainty Contribution	300–450	450–650	650–1000	> 1000
Total	18	4.9	6.5	32
Statistical	16	3.0	5.5	30
Systematic	7	3.9	3.4	10
Jet Systematics	6	4.3	3.3	10
Modeling and Theory Systs.	4	0.7	0.7	2
Flavor Tagging Systs.	0.2	0.4	0.4	2

in pT1, but drop in the rankings as p_{T}^{H} increases where the $V+\mathrm{jets}$ errors become relatively important.

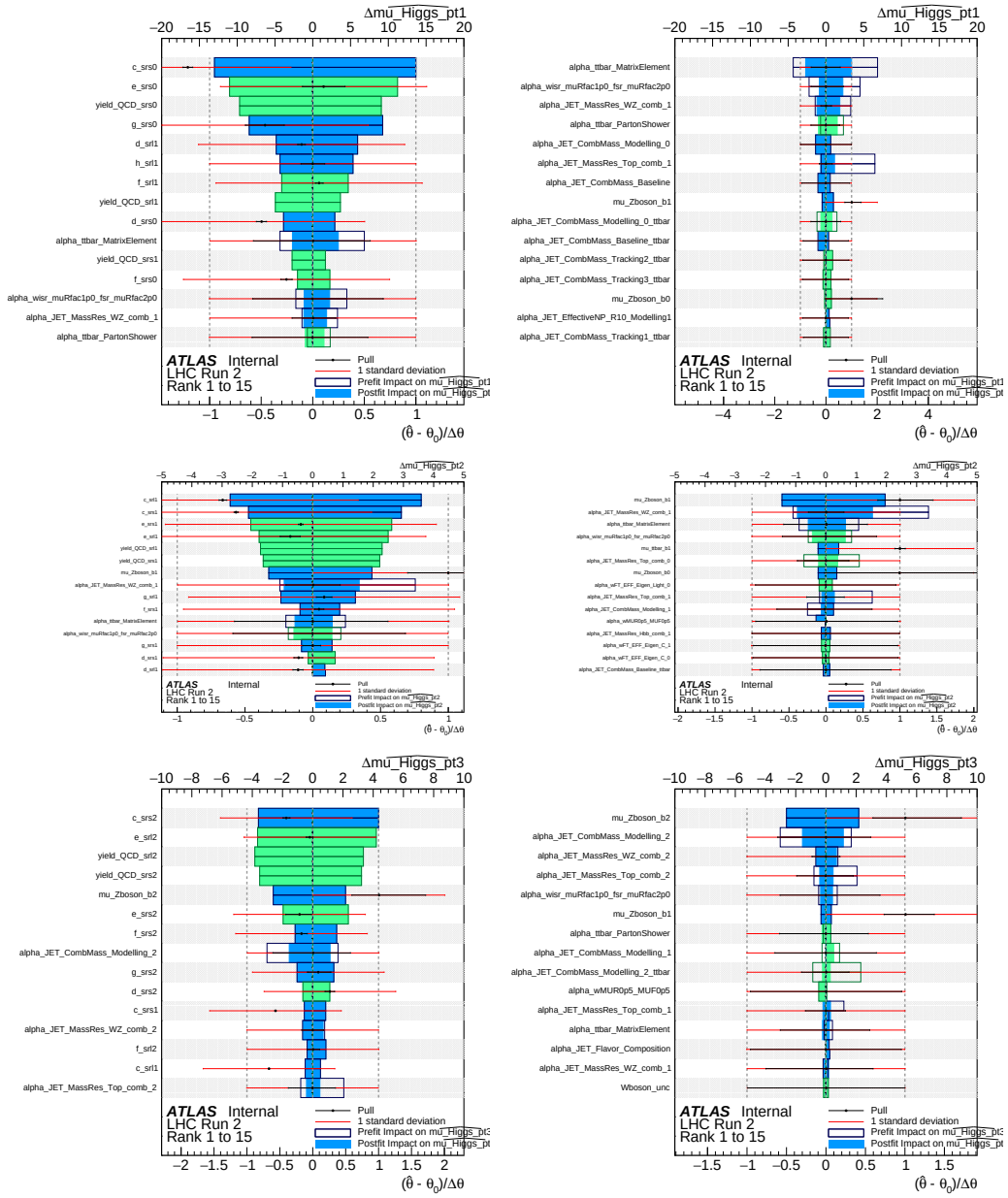


Figure 11.8: The ranking plot for pT1 (top panel), pT2 (middle panel), pT3 (bottom panel) both with (left) and without (right) QCD parameters shown.

Chapter 12

Conclusion

12.1 ATLAS Run 2 Boosted $H \to b\bar{b}$ Results

We have presented an inclusive measurement [220, 221] of the Higgs boson produced at high transverse momentum and decaying to a pair of bottom quarks. The results are based on the Run 2 dataset of pp collision data, collected at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector at the LHC, corresponding to an integrated luminosity of $136\,\text{fb}^{-1}$. The Higgs boson is reconstructed using a single large-R ($R = 1.0$) jet and identified using b-tagging techniques. We extract the Higgs production cross section in several regions of increasing p_{T}. We also perform cross-section measurements of $Z \to b\bar{b}$ in the same kinematical regimes. The results agree with the Standard Model predictions, validating the methods.

Summarizing the results from the previous chapter, the fit to the inclusive region yields a Higgs boson signal strength for the combination of the SRL, SRS, and $\text{CR}_{t\bar{t}}$ is $\mu_H = 0.8 \pm 3.2$. The Higgs boson production cross section in the fiducial and differential regions is presented below.

Fiducial

The upper limits on Higgs boson cross-section production at 95% CL for the fiducial measurement, defined by the phase space with $p_{\mathrm{T}}^{H} > 450\,\mathrm{GeV}$ and $|y_H| < 2$:

$$\sigma_H(p_{\mathrm{T}}^{H} > 450\,\mathrm{GeV}) < 115\,\mathrm{fb}.$$

Differential

The upper limits on Higgs boson cross-section production at 95% CL, obtained in the four differential volumes with $|y_H| < 2$:

$$\sigma_H(300 < p_{\mathrm{T}}^{H} < 450\,\mathrm{GeV}) < 2.9\,\mathrm{pb},$$
$$\sigma_H(450 < p_{\mathrm{T}}^{H} < 650\,\mathrm{GeV}) < 89\,\mathrm{fb},$$
$$\sigma_H(650 < p_{\mathrm{T}}^{H} < 1000\,\mathrm{GeV}) < 39\,\mathrm{fb},$$
$$\sigma_H(p_{\mathrm{T}}^{H} > 1\,\mathrm{TeV}) < 9.6\,\mathrm{fb}.$$

The Higgs boson production cross section for $p_{\mathrm{T}}^{H} > 1\,\mathrm{TeV}$ is:

$$\sigma_H(p_{\mathrm{T}}^{H} > 1\,\mathrm{TeV}) = 2.3 \pm 3.9\,(\mathrm{stat.}) \pm 1.3\,(\mathrm{syst.}) \pm 0.5\,(\mathrm{theory})\,\mathrm{fb}.$$

All of the Higgs boson results are consistent with their Standard Model predictions. The fitted values and upper limits at 95% CL of the Higgs boson signal strengths in three p_{T}^{H} regions, are shown in Figure 12.1.

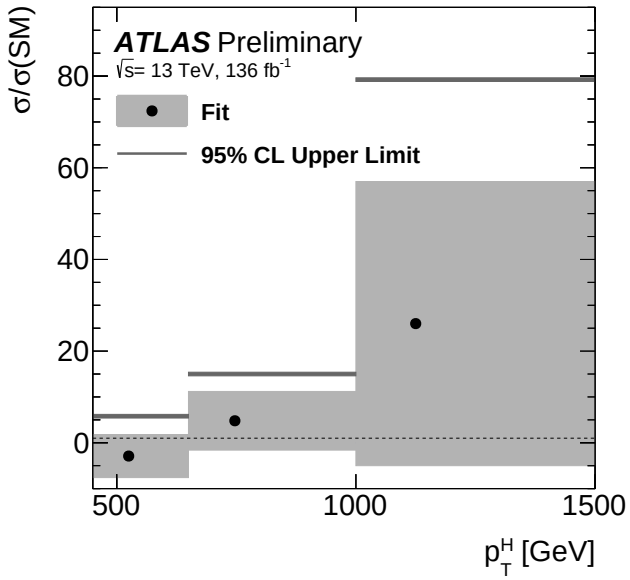


Figure 12.1: A summary of the Higgs boson signal strengths measured in three p_{T}^{H} differential volumes, defined by the transverse momentum of the Higgs boson p_{T}^{H}.

12.2 CMS Run 2 Boosted $H \to b\bar{b}$ Results

CMS also performed an inclusive search [222] for the Higgs boson production at high transverse momentum, decaying to a pair of bottom quarks. The search used the LHC Run 2 data collected by CMS at $\sqrt{s} = 13\,\text{TeV}$, corresponding to an integrated luminosity of $137\,\text{fb}^{-1}$. The Higgs boson is reconstructed into a single large-R ($R = 0.8$) jet with $p_\text{T} > 450\,\text{GeV}$. Similar to the ATLAS analysis, CMS measured the $Z+\text{jets}$ ($Z \to b\bar{b}$) process to validate their method, finding it consistent with the Standard Model prediction. The Higgs boson signal strength is measured to be $\mu_H = 3.68 \pm 1.20\,(\text{stat.})^{+0.63}_{-0.66}\,(\text{syst.})^{+0.81}_{-0.46}\,(\text{theo.}) = 3.68^{+1.58}_{-1.46}$ based on the theoretical predictions from the HJ-MiNLO [409, 412, 416, 506], corresponding to an observed (expected) significance of $2.54\,\sigma$ ($0.71\,\sigma$) with respect to the background-only hypothesis. The differential cross section in the three differential transverse momentum regions is:

$$\sigma_H(300 < p_\text{T}^H < 450\,\text{GeV}) = 600 \pm 700\,(\text{stat.}) \pm 350\,(\text{syst.})\,\text{fb}$$

$$\sigma_H(450 < p_\text{T}^H < 650\,\text{GeV}) = 5 \pm 37\,(\text{stat.}) \pm 22\,(\text{syst.})\,\text{fb}$$

$$\sigma_H(p_\text{T}^H > 650\,\text{GeV}) = 29 \pm 9\,(\text{stat.}) \pm 7\,(\text{syst.})\,\text{fb}.$$

Figure 12.2 shows the signal-strength parameters for the Higgs boson production, in several p_T bins, from their fit to the data.

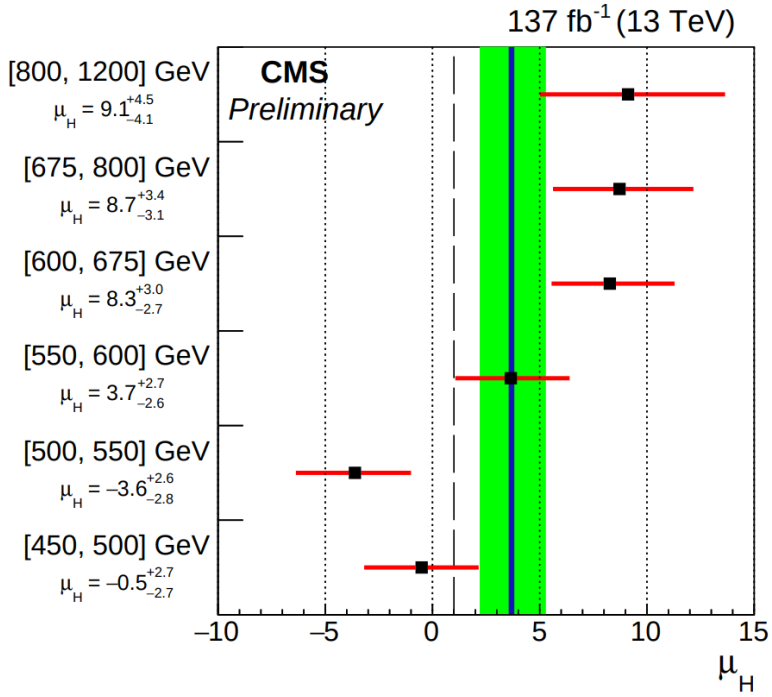


Figure 12.2: The best-fit signal strength μ_H (black squares) and uncertainty (red lines) per p_{T} category based on the HJ-MiNLO [409, 506] prediction. The dashed black line indicates the SM expectation. The solid blue line and green bands represent the combined best-fit signal strength and uncertainty, respectively: $\mu_H = 3.68^{+1.58}_{-1.46}$ extracted from a simultaneous fit of all channels [222].

12.3 Comparisons

We performed a comparison between the ATLAS and CMS results reported above, both extracted from the full Run 2 dataset. Results from the latest ATLAS $H \to \gamma\gamma/ZZ^* \to 4\ell$ combination [507], which covers the lower-end of the p_{T}^{H} spectrum, are also compared. The values from the studies are shown in Figure 12.3, where we see a similar trend between the ATLAS and CMS results in the highest p_{T} regions probed by the respective analyses. The result from CMS is more sensitive than that of ATLAS, in large part due to a reduction the QCD background to roughly half that of ATLAS through the use of jet substructure variables to better discriminate $g \to b\bar{b}$ vs $H \to b\bar{b}$ jets.

We did not perform an EFT interpretation of the results, but we intend to pursue this study in the future. Loop-induced processes such as ggF are quite sensitive to BSM physics and could produce the largest changes in some of the EFT parameters. In the next section, we will briefly discuss the experimental results and how they may be compared to different values of a few of the EFT parameters, specifically a subset involving the Higgs couplings.

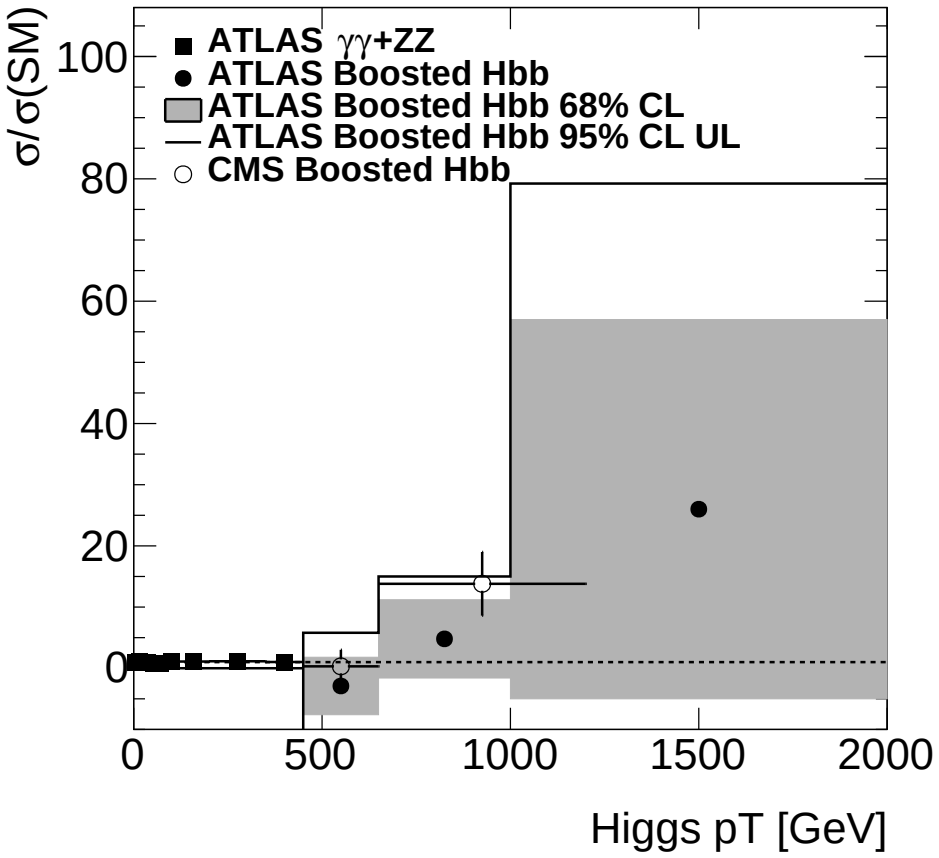


Figure 12.3: Comparison of the latest ATLAS $H \to \gamma\gamma/ZZ^* \to 4\ell$ combination [507] combination results for $p_{\mathrm{T}}^H < 350\,\mathrm{GeV}$ and the ATLAS [220, 221] and CMS [222] Boosted $H \to b\bar{b}$ results for $p_{\mathrm{T}}^H > 350\,\mathrm{GeV}$.

12.4 Effective Field Theory: Redux

We now pick up our discussion regarding effective field theories (EFTs) [194, 195, 508, 509] from Section 3.4, where we introduced the following EFT Lagrangian:

$$\mathcal{L}_{\mathrm{EFT}} = \mathcal{L}_{\mathrm{SM}} + \sum_i \frac{c_i}{\Lambda^2}\mathcal{O}_i, \tag{12.1}$$

where Λ is the scale of new physics, $\mathcal{O}_i$ are dimension-6 EFT operators, and c_i are the Wilson coefficients, which parameterize the virtual effects of heavy new physics into the lower-energy observables (*i.e.,* our couplings to the Higgs boson).

The odd-dimension operators are removed from the EFT Lagrangian to conserve baryon and lepton numbers, and higher order operators have been truncated as they are suppressed by higher and higher powers of Λ [181, 510–512]. Many groups have already used the LHC data and data from earlier experiments to place bounds on the Wilson coefficients for the dimension-6 operators [513–518]. For our discussion, we'll refer to the strongly-interacting light Higgs (SILH) basis [182], where the Wilson coefficients are expressed as factors in the canonically normalized Lagrangian. The SILH basis contains 59 operators of dimension-6, assuming flavor-universal couplings, with an additional seventeen operators for the hermitian conjugates. The majority of these operators either do not affect Higgs physics or have coefficients that are already tightly constrained by precision electroweak data at leading order [519]. For a (more) complete review of the connection between EFTs and Higgs physics, see References [517, 520–524].

By reducing this set of operators, we can focus on a promising subset of operators, related to the effective couplings for gluon-gluon fusion involving the Higgs boson, top and bottom quarks, and gluons, as seen in Equation 12.2 [199]:

$$\begin{aligned}
\mathcal{O}_1 &= |H|^2 G^a_{\mu\nu} G^{a,\mu\nu}, \\
\mathcal{O}_2 &= |H|^2 \overline{Q}_L H^c u_R + h.c., \\
\mathcal{O}_3 &= |H|^2 \overline{Q}_L H d_R + h.c \\
\mathcal{O}_4 &= \overline{Q}_L H \sigma^{\mu\nu} T^a u_R G^a_{\mu\nu} + h.c.,
\end{aligned} \tag{12.2}$$

where $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$. These operators are interesting in that $\mathcal{O}_1$ yields a contact

interaction between gluons and the Higgs, and $\mathcal{O}_2$ and $\mathcal{O}_3$ describe modifications to the top and bottom Yukawa couplings respectively. $\mathcal{O}_4$ is known as the chromomagnetic dipole-moment operator, which modifies the coupling between gluons and the top quark. With careful work, it is possible to constrain the values of the effective couplings, giving us hints at the structure of BSM physics. We can represent these operators as:

$$\begin{aligned}
\frac{c_1}{\Lambda^2}\mathcal{O}_1 &\rightarrow \frac{\alpha_s}{\pi v} c_g h G^a_{\mu\nu} G^{a,\mu\nu} \\
\frac{c_2}{\Lambda^2}\mathcal{O}_2 &\rightarrow \frac{m_t}{v} c_t h \bar{t}t \\
\frac{c_3}{\Lambda^2}\mathcal{O}_3 &\rightarrow \frac{m_b}{v} c_b h \bar{b}b \\
\frac{c_4}{\Lambda^2}\mathcal{O}_4 &\rightarrow c_{tg}\frac{g_s m_t}{2v^3}(v+h) G^a_{\mu\nu}(\bar{t}_L \sigma^{\mu\nu} T^a t_R + h.c)
\end{aligned} \tag{12.3}$$

The couplings c_g, c_t, and c_b can be probed in Higgs boson processes. The coupling c_{tg} is constrained by top pair production measurements [525], but could still contribute significantly, within the current bounds, to the gluon fusion cross section [199]. The couplings c_t and c_b, modifiers to the y_t and y_b Yukawa couplings, respectively, may be measured directly in ttH and bbH production modes, with c_b also directly accessible in the $H \rightarrow b\bar{b}$ decay channel. The vertices where these new couplings could appear in the gluon fusion diagrams are shown in Figure 12.4.

The results from Reference [199] report interesting consequences regarding the shape of the Higgs boson transverse momentum distribution that are worth reiterating here to give a more intuitive feel for the impact of the EFT couplings. The variations of the Wilson coefficients manifest themselves in *different* regions of the Higgs

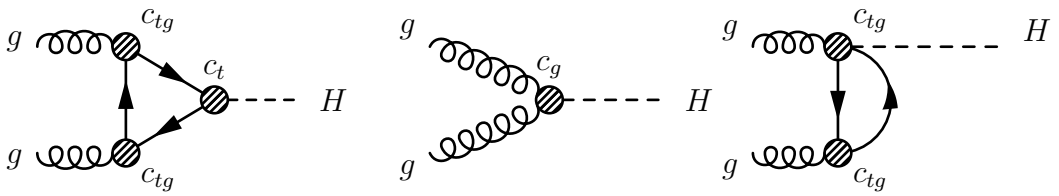


Figure 12.4: Feynman diagrams contributing to the LO $gg \to H$ production. The blobs indicate possible interactions formed by the dimension-6 operators.

p_{T} spectrum. Modifying the Higgs-bottom Yukawa coupling (c_b in Equation 12.3) almost exclusively induces effects at low transverse momenta. On the other hand, the effective coupling of the Higgs boson to gluons (c_g in Equation 12.3) changes the shape of the distribution primarily in the high-p_{T} tail. Additionally, the shape depends on the mass of the particle that mediates the Higgs-gluon coupling; the higher the mass of the particle, the harder the resulting spectrum. Changes to the top-quark Yukawa coupling (c_t in Equation 12.3) primarily affect the overall normalization and approximately correspond to a simple rescaling of the transverse momentum distribution. Interestingly, while individual variations of the Wilson coefficients produce effects roughly on par with the NLL+NLO [526, 527] perturbative uncertainties, variations of two or more coefficients are capable of significantly distorting the spectrum beyond said uncertainties [199]. An example from Reference [199], in Figure 12.5, shows the possible distortions in the Higgs p_{T} spectrum for different values of c_t and c_g. Notice that at low transverse momentum, the changes to the p_{T} spectrum are within the uncertainties formed by scale variations on μ_F and μ_R, but with $p_{\mathrm{T}} > 350\,\mathrm{GeV}$, the deformations become larger than the uncertainties.

The operators we've focused on are primarily important for the ggF production mode; different production modes are sensitive to the various EFT couplings to

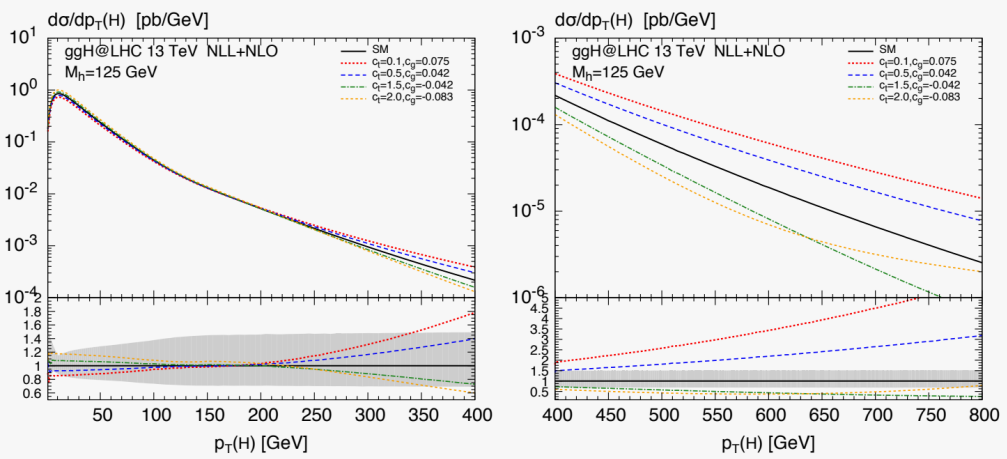


Figure 12.5: The Higgs transverse-momentum spectrum in the SM (solid black line) compared to simultaneous variations of c_t and c_g for (left) $0 < p_{\mathrm{T}} < 400\,\mathrm{GeV}$ and (right) $400 < p_{\mathrm{T}} < 800\,\mathrm{GeV}$. The lower frames show the ratio with respect to the SM prediction. The shaded band in the ratio indicates the uncertainty due to scale variations [199].

varying degrees. Some modes are sensitive to operators not specifically mentioned here. This motivates at least some minimal definition of the recoil system in future measurements to provide some disentanglement of the main Higgs production modes. For example, some of the EFT operators in the SILH basis [528], in particular those with the c_g and c_{tg} coupling modifiers, motivate measuring ttH and ggF production. The possible high-p_{T} enhancements could lie within the sensitivity of the Boosted $H \to b\bar{b}$ Analysis without creating significant deviations from the SM prediction at low transverse momentum of the Higgs boson. However, the effect is not uniform in both production modes, with ttH being less sensitive. Concerning other production modes, recent results on VBF production where the Higgs boson decays to photons or leptons also include high-p_{T}^{H} event categories, but they have limited reach [26, 125, 126, 207, 529, 530]. On the other hand, the analyses of VH production with leptonic V-decays have achieved considerable sensitivity in the high-p_{T} regime for a specific production mode [25, 531, 532]. Although we did not perform an EFT interpretation of the results, we did perform a 2D scan of the c_g and c_{tg} Wilson

coefficients for specific phase-space points to demonstrate the sensitivity potential of highly boosted Higgs production to new physics. The results of the scan, along with the combined results shown in the previous section, are shown in Figure 12.6.

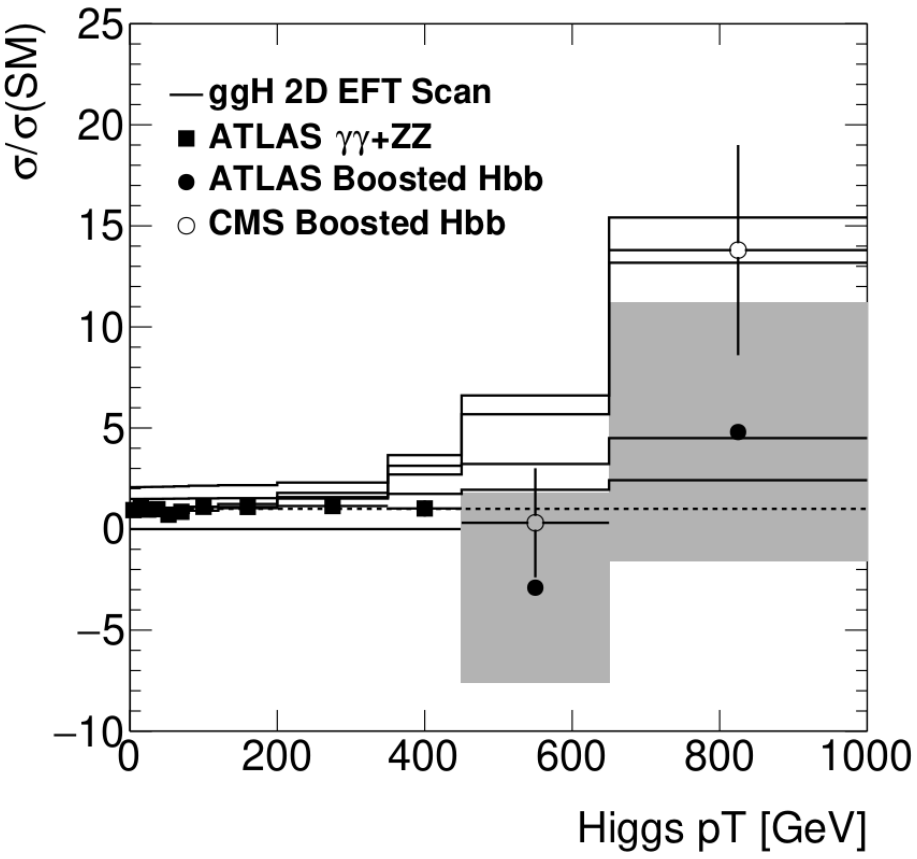


Figure 12.6: Comparison of the latest ATLAS $H \to \gamma\gamma/ZZ^* \to 4\ell$ combination [507] combination results for $p_{\mathrm{T}}^H < 350\,\mathrm{GeV}$ and the ATLAS [220, 221] and CMS [222] Boosted $H \to b\bar{b}$ results for $p_{\mathrm{T}}^H > 350\,\mathrm{GeV}$. Expected EFT-to-SM cross-section ratios for points in the (c_g, c_{tg}) phase space are representative of observed p_{T}^H patterns. The black lines are from multiple scans of the EFT parameters c_g and c_{tg}.

12.5 Future Prospects

At the time of writing, the LHC is in its second long shutdown (LS2) phase. The next phase of data taking, Run 3, is expected to begin in early 2022 and continue until about 2025, with plans to deliver a total integrated luminosity of $350\,\mathrm{fb}^{-1}$. Af-

terwards, the LHC will shut down again (LS3) and prepare for the High Luminosity LHC upgrade (HL-LHC) [533–535], which is expected to deliver a total integrated luminosity of $3000\,\mathrm{fb}^{-1}$ ($3\,\mathrm{ab}^{-1}$) [536] at center-of-mass energy $\sqrt{s} = 14\,\mathrm{TeV}$. The plans for the LHC from Run 1 until Runs 4 and 5 are shown in Figure 12.7.

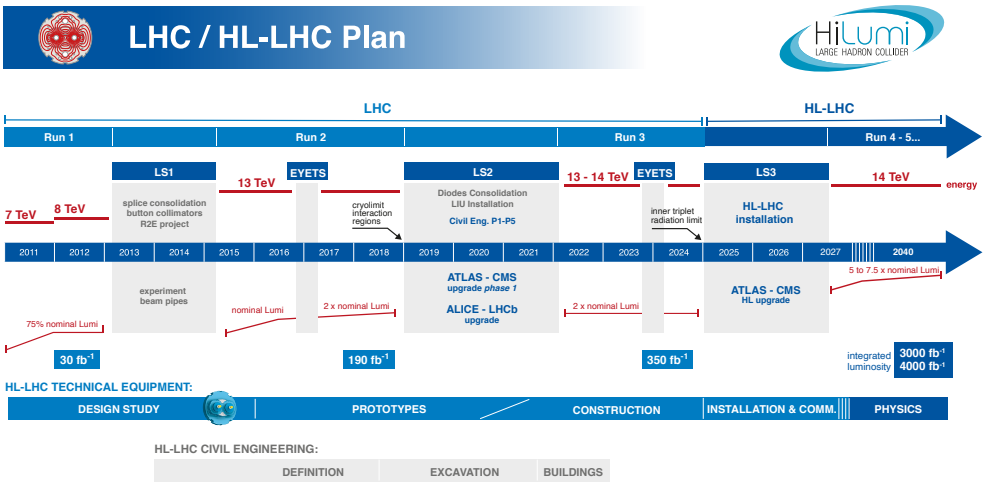


Figure 12.7: LHC and High Luminosity LHC plan from 2011 until 2037 and beyond [535].

Both ATLAS [537–544] and CMS [545–557] will be upgrading their detectors[1] with new components to reflect the changing state of technology, which are expected to withstand a much higher level of radiation from the increased luminosity than in previous Runs of the LHC [558].

Having occupied the particle physics community for half a century, the Higgs boson certainly has shaped the physics program of the LHC and of prospective future colliders [157]. With its discovery secured, the task remains to obtain higher precision measurements to more accurately describe the properties of the Higgs boson, as well as search for deviations from the Standard Model predictions. The increase in our

[1] Early in the PhD program at UCSC, before joining the Boosted $H \to b\bar{b}$ Analysis, I worked on the upgrade of the strip detector (SCT replacement) [540] for the Inner Tracker (ITk) [558]. See References [559–566] for my contributions to sensor performance studies, glue qualification studies, etc.

sensitivity is expected to scale with the square root of the luminosity [567], and the HL-LHC will bring a new precision era in Higgs boson measurements, improving the current precision by a factor of 5–10 for nearly all observables [79]. For the dimension-6 Wilson coefficients, we expect improvements in the sensitivity to scale similarly to the SM Higgs observables [568]. As an example, in Figure 12.8, we see the expected exclusion limits as a function of the luminosity for the Higgs-gluon effective coupling c_g. Personally, I feel that the physics program of the next decade at the LHC is incredibly exciting. The data have thus far been in frustratingly good agreement with the Standard Model, and I look forward to seeing, and participating in, what comes next.

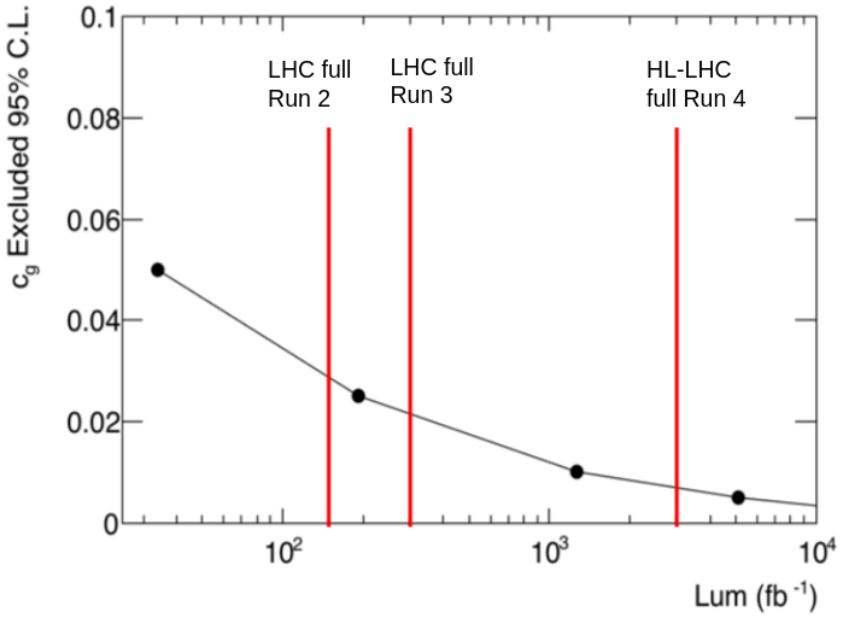


Figure 12.8: Expected c_g exclusion limits at the 95% CL as a function of luminosity. The red lines show the approximate total integrated luminosities for the LHC Runs 2, 3, and 4.

Appendix A

Theory Appendix

A.1 Particle Discovery

Table A.1: Timeline of particle discoveries from Ref. [320]. Additions to the table were made by me.

Particle	Symbol	Discovery	Person/Collaboration	Paper
Photon*	γ	1895	Röntgen, Einstein	[569–572]
Electron	e	1897	J.J. Thomson	[573]
Proton	p	1919	Ernest Rutherford	[574]
Neutron	n	1932	James Chadwick	[575]
Muon	μ	1937	Neddermeyer, Anderson	[576]
Electron Neutrino	ν_e	1956	Cowan, Reines	[577]
Muon Neutrino	ν_μ	1962	BNL (AGS)	[578]
Up Quark†	u	1969	SLAC	[579]
Down Quark†	d	1969	SLAC	[579, 580]
Strange Quark†	s	1969	Rochester and Butler, SLAC	[579, 581]
Charm Quark	c	1974	SLAC, MIT	[582, 583]
Tau	τ	1975	SLAC, LBL	[584]
Bottom Quark	b	1977	Fermilab (E288)	[49]
Gluon	g	1979	DESY (PETRA)	[585]
W Boson	W	1983	CERN (UA1, UA2)	[42, 43]
Z Boson	Z	1983	CERN (UA1, UA2)	[44, 45]
Top Quark	t	1995	Fermilab (DØ, CDF)	[52, 53]
Tau Neutrino	ν_τ	2000	Fermilab (DONUT)	[586]
Higgs Boson	H	2012	CERN LHC (ATLAS, CMS)	[21, 22]

* Who discovered the photon is actually a bit tricky. Röntgen discovered gamma rays in 1895 [569], but the concept of a photon as a quanta of light wasn't proposed until 1905 by Einstein in his famous paper on the photo-electric effect [570]. The term photon wasn't coined until 1926 by Gilbert Lewis [571]. Its absolute verification may be said to have occurred as late as 1977 [572] by photon-correlation experiments, but we humans have been detecting photons with our eyes for much longer.

† The baryons and mesons containing the up, down, and strange quarks (protons, neutrons, pions, kaons, ect. were discovered much earlier, but the experiment at SLAC verifying the existence of quarks [579] was not performed until 1969. Kaons, the first hadrons containing strange quarks, were discovered in 1947 [581].

A.2 Noether's Theorem

Recall the Lagrangian density is the difference between the kinetic and potential energy (densities) of a system:

$$\mathcal{L} = \mathcal{T} - \mathcal{V} \tag{A.1}$$

where $\mathcal{L}$ is the Lagrangian density, $\mathcal{T}$ is the kinetic (or dynamic/propagating) term, and $\mathcal{V}$ is the potential (or interaction) term. The term "density" will be dropped for simplicity, and terms using the calligraphic script will be used to represent the

quantity's density. Consider as an example a simple Lagrangian as a function of a field ϕ, which is itself a function of space-time. Just like with classical mechanics, the dynamics for a Lagrangian system in QFT are determined by the *principle of least action* [587], where the action is the integral of the Lagrangian over space-time:

$$S = \int d^4x \, \mathcal{L}(\phi, \partial_\mu \phi), \tag{A.2}$$

where our Lagrangian is a function of a field ϕ and its first derivatives. We can now imagine applying an infinitesimal variation on the field, allowing $\phi \to \phi + \delta\phi$. We find with a little calculus of variations, the variation of the action becomes:

$$\begin{aligned} \delta S &= \int d^4x \left[\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta(\partial_\mu \phi) \right] \\ &= \int d^4x \left(\left[\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right] \delta\phi + \cancelto{0}{\partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi \right]} \right). \end{aligned} \tag{A.3}$$

The last term is a total derivative, so using the physical assumption that our fields vanish at the boundaries, we drop it. If the fields satisfy the equations of motion, which are themselves determined by the principle of least action, the action will be stationary when $\frac{\delta S}{\delta \phi} = 0$, so that the integrand of Equation A.3 yields the Euler-Lagrange equations:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = 0. \tag{A.4}$$

Any continuous transformation which leaves the equations of motion unaffected defines a symmetry of the Lagrangian. We can generalize our earlier example of an

infinitesimal variation on a complex field ($\phi \to \phi + \delta\phi$), which could be written as $\phi = \phi_1 + i\phi_2$, to a continuous symmetry:

$$\phi \to e^{-i\alpha}\phi \quad \phi^* \to e^{i\alpha}\phi^*. \tag{A.5}$$

Given a symmetry that depends upon some parameter α, which can be taken to be small (*i.e.* we have a *continuous* symmetry) we find that:

$$\frac{\delta\mathcal{L}}{\delta\alpha} = \sum_n \left(\cancelto{0}{\left[\frac{\partial\mathcal{L}}{\partial\phi_n} - \partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_n)}\right]} \frac{\delta\phi_n}{\delta\alpha} + \partial_\mu \left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_n)} \frac{\delta\phi_n}{\delta\alpha}\right] \right) = 0, \tag{A.6}$$

where ϕ_n could be any number of fields the Lagrangian depends on. If the equations of motion are satisfied, the first term in brackets is zero as we saw in Equation A.4. The second term in brackets is the Noether current J_μ,

$$J_\mu = \sum_n \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_n)} \frac{\delta\phi_n}{\delta\alpha}, \tag{A.7}$$

which we see from Equation A.6 that this current is conserved, *i.e.* that

$$\partial_\mu J_\mu = 0. \tag{A.8}$$

Given a vector field J_μ that satisfies the above relationship, we can show that the total charge Q associated with this symmetry is a conserved quantity. Q is defined:

$$Q = \int d^3x\, J_0, \tag{A.9}$$

and assuming $\vec{J}$ vanishes at the spatial boundary, satisfies

$$\partial_t Q = \int d^3x \, \partial_t J_0 = \int d^3x \, \vec{\nabla} \cdot \vec{J} = 0. \tag{A.10}$$

A.3 A Little SM History

The history behind the quantum field theory formulation of the Standard Model, especially for the weak and strong nuclear forces is a long one – full of many twists, turns, and a few dead-ends. But these historical missteps are insightful (and perhaps, at times, inciteful), as they provided a guiding light by which to press forward. With that in mind, I want to take a step back to motivate some of the directions taken by particle physicists in the decades preceding the current form of the SM.

In the 1930s, protons and neutrons, the latter had just been discovered by James Chadwick [575], were thought to be fundamental particles. Electromagnetism was being formulated as a quantum field theory in QED, though without the procedure of renormalization, it was only useful at leading order in perturbation theory. Two other interactions had also been identified; the weak nuclear force, responsible for beta decay, and the strong nuclear force, which was responsible for holding protons and neutrons together. Given that the neutron and proton have nearly the same (degenerate) masses, it was suggested in 1932 by Heisenberg [588] that, if we ignore the different electric charges, the two were degenerate states (the same particle) with regards to the strong nuclear force. This approximate degeneracy led to the suspicion that it was an approximate symmetry in nature, but an exact symmetry for nuclear interactions. Viewing the neutron and proton as two linearly independent states of the same particle allows us to represent them as a two-component vector, similar to

the spin-up and spin-down states for fermions.

$$\Psi = \begin{pmatrix} \Psi_{\text{Proton}} \\ \\ \Psi_{\text{Neutron}} \end{pmatrix} \tag{A.11}$$

Much like the case with spin, the choice of which of these nucleons is the "up" state (proton) and which is the "down" state (neutron) is arbitrary, but once this choice is made, we need to stick to it. Transformations under this symmetry can be handled analogously to that of spin; that is, using the unitary transformation under the $SU(2)$ Lie group:

$$\Psi \rightarrow e^{i\alpha_a \frac{\sigma^a}{2}} \Psi \tag{A.12}$$

where $\sigma^a/2$ are the generators of the group, responsible for "rotating" between the two particle types and σ^a are the Pauli spin matrices,

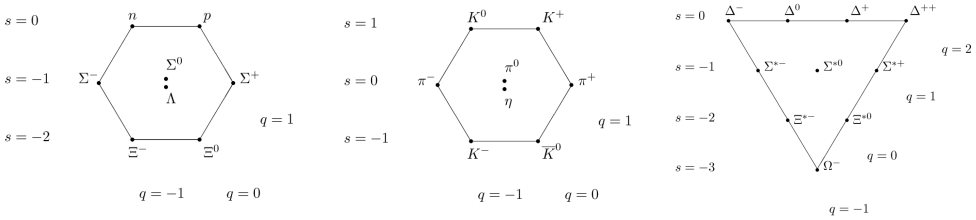
$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ \\ 0 & -1 \end{pmatrix}. \tag{A.13}$$

The term used for this spin-like property was introduced by Eugene Wigner [589] as isotopic spin, or isospin, due to its close analogy to quantum mechanical spin. In this context, the proton has isospin $+1/2$ and the neutron isospin $-1/2$. This grouping was also used for the newly discovered pions, which also have near degenerate masses; The π^+, π^0, and π^- were grouped into an isospin triplet with values $+1$, 0, and -1, respectively. It's important to note that these generators do not commute, which implies that the gauge fields can self-interact if they are charged under the symmetry group. This is an important property of $SU(2)$ as well as the general case of $SU(N)$,

which led physicists to devise a method to generalize $SU(N)$ into what we now refer to as a Yang-Mills theory, discussed in Section A.5. Before that though, we shall look at how the strong and weak nuclear forces came to have the symmetry groups they have in the SM.

A.3.1 The Strong Nuclear Force

The invention of the bubble and spark chambers of the 1940s and 1950s produced a rapid progression in the discovery of new, possibly fundamental, particles we generically call hadrons. The hadrons come in two varieties: baryons, which are fermions with half-integer spins, and mesons, which are bosons with integer spin. With the addition of so many new "fundamental" particles, patterns began to emerge as new properties were seen. As more and more particles were being discovered, particles with a peculiar property were noticed, one called *strangeness*; these particles appeared to be easily produced in cosmic-ray and accelerator reactions, but decayed slowly, as if something were hindering their decays [590]. In 1961, Murray Gell-Mann and Yuval Ne'eman [591] collected these baryons and mesons into groups with similar properties, in something called the Eight-Fold Way (Figure A.1).



(a) (b) (c)

Figure A.1: Baryon octet (a), meson octet (b), and baryon decuplet (c) [592].

The Eight-Fold Way organizes eight spin-0 mesons into an octet (Figure A.1b), eight spin-1/2 baryons into another octet (Figure A.1a), and ten spin-3/2 baryons

into a decuplet (Figure A.1c), based on electric charge and strangeness. It was Gell-Mann [593] and George Zweig [594] who, independently, realized that the variety they saw in the hadrons could be explained if they were composed of new particles called quarks[1] (by Gell-Mann) or *Aces* (by Zweig) with fractional electromagnetic charges (either $+2/3$ or $-1/3$). Because we never observe fractional charges in nature, this suggests that quarks are grouped in such a way as to have only integer units of electrical charge. This implies we can see baryons as triplets of either three quarks or antiquarks and mesons as quark-antiquark pairs. The proton is no longer a fundamental particle, but a collection of 2 "up" quarks ($+2/3$ charge) and 1 "down" quark ($-1/3$ charge) and the neutron 1 up and 2 down. The observed strangeness was in fact related to the number of strange quarks (electric charge $-1/3$) present in the particle. The strong interaction was thought to conserve strangeness, at least approximately, while the weak violated it. This meant that strangeness changing strong decays of particles such as the K^+ (a kaon) were forbidden, giving it a higher than usual lifetime (a whopping 1.24×10^{-8} seconds).

With 3 quarks (u, d, and s), it seemed the strong interaction didn't obey the $SU(2)$ symmetry of nucleons (Section A.3), but rather an $SU(3)$ *flavor* symmetry. This turned out to be an approximate symmetry, one that was less useful as more quarks were discovered. Instead, it turned out that these quarks possessed a new quantum number, called *color*, whose invention was precipitated by the Δ^{++} baryon (3 up-quarks with spin up), whose wave function *must* be antisymmetric because it is a spin-3/2 fermion. Before the introduction of color however, the total wave-function

[1] Coming from the phrase "three quarks for Muster Mark" in James Joyce's *Finnegan's Wake*.

(a product of the spacial, spin, and flavor wave-functions) is symmetric:

$$\psi_{\text{Total}} = \psi_{\text{Spacial}}\psi_{\text{Spin}}\psi_{\text{Flavor}}. \tag{A.14}$$

In the ground state, the spatial wave-function is symmetric. The spin and flavor matrices are manifestly symmetric as well. This suggested an additional internal symmetry, consisting of three colors, which provides us with an anti-symmetric configuration:

$$\psi_{\text{Color}} = \frac{1}{\sqrt{6}}(RGB - RBG + GBR - GRB + BRG - BGR). \tag{A.15}$$

The theory that explains the interactions under this color, or strong interaction is known as quantum chromodynamics (QCD) with the symmetry group $SU(3)_{\text{color}}$, which will be formally laid out in Section A.7. Although the colors are not at all related to our ordinary notion, the naming scheme is mildly clever. The name arises from the fact that quarks are typically bound in quark triplets of particles (with colors red, green, and blue) or antiparticles (with anti-red, -green, -blue)[2], or bound in quark anti-quark pairs (red and anti-red, etc). While three colors are not necessarily predicted by the theory, experimental evidence supports the existence of only three [595].

[2]Or more fun; cyan, magenta, and yellow. Coming from properties of the real colors in which I am in no way qualified to talk about.

A.3.2 The Weak Nuclear Force

The weak interaction was first introduced by Enrico Fermi [596] in 1934 to explain beta decay which involves the decay of a neutron (along with a neutrino) into a proton and electron[3]. The way interactions associated with the weak nuclear force change up- and down- type quarks and swap neutrinos and leptons hinted early on that the weak interaction obeyed an $SU(2)$ gauge theory, much like the case for nucleons discussed earlier, which conserves the quantum number known as weak isospin.

An $SU(2)$ gauge theory such as this however not only predicts massless fermions, but massless bosons, which is inconsistent with expectations. The short-ranged nature of the weak force implies that the force carriers should be quite massive. In the massless limit, left- and right-handed fermions would decouple in the Dirac equation, so this would not necessarily be a problem, but electrons *do* have mass. We will need to resolve this issue to see how to mix these different helicity (handedness) states to generate mass. (Section 2.3). Further issues were discovered in the 1950s following the experiments carried out by Chien-Shiung Wu [599] on the decay of ^{60}Co to determine if the weak force conserved parity. In this decay, a neutron decays to a proton, emitting an electron and an anti-neutrino in a magnetic field. If parity were conserved, the electrons would not have a preferred direction of decay relative to the nuclear spin. But indeed the electrons *did* have a preferred direction which was opposite to the nuclear spin. The results from this experiment showed that the weak force not only violates parity, but it only interacts with left-handed (right-handed)

[3] Amazingly, when Fermi submitted his theory of beta decay to Nature, it was outright rejected "because it contained speculations too remote from reality to be of interest to the reader." [597] While it was later published in German and Italian journals, it was not published in English until it appeared in the American Journal of Physics in 1968 [598]!

particles (anti-particles).

So where does this leave us? Massive bosons and fermions will not manifest from a gauge invariant $SU(2)$ as we will see in Section 2.3 and experimental evidence made it clear that the weak interaction violated parity, which had been firmly believed to be conserved, leading Pauli to say "*I cannot believe that God is a weak left-hander.*" But still, it seemed the weak force observed some kind of $SU(2)_L$ (L = Left) structure, where left-handed particles could be treated as doublet wave functions and right-handed as singlets. These left- and right-handed fields are expressed as

$$\psi_L = \frac{1}{2}(1-\gamma_5)\psi, \quad \psi_R = \frac{1}{2}(1+\gamma_5)\psi. \tag{A.16}$$

Where $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$ and $\frac{1}{2}(1 \pm \gamma_5)$ are the right/left helicity projectors P_R/P_L, satisfying $P_R + P_L = 1$. The left-handed doublet fields have isospin 1/2 and right-handed singlet fields have isospin 0:

$$\psi_L = \begin{pmatrix} q_u \\ \\ q_d \end{pmatrix}_L, \quad \begin{pmatrix} \nu_\ell \\ \\ \ell \end{pmatrix}_L, \quad \psi_R = (q_u)_R, (q_d)_R, (\ell)_R, \quad \begin{array}{l} q_u = u, c, t, \\ q_d = d, s, b, \\ \ell = e, \mu, \tau \end{array} \tag{A.17}$$

Notice the omission of any right-handed neutrinos. While not forbidden, experimental evidence has yet to observe any. This is not terribly surprising since they are not only left out of the weak interaction, but the electromagnetic interaction as well due to their neutral charge.

A.4 Definition of a Group

A symmetry group G consists of a set of distinct elements g_a and group multiplication satisfies the following:

- Closure: The product of two group elements is also a group element.
- Associativity: The order of multiplication of three or more group elements does not affect the product.
- Identity: There exists a unique identity element which when multiplied by another group element, leaves that element unchanged.
- Inverse: Each group element has a unique inverse, such that multiplication of the element with its inverse yields the identity element.

A.5 Yang-Mills Theories

Unlike the relative simplicity of QED's $U(1)$ symmetry, the Standard Model relies on a more rich and complex set of symmetries described in Yang-Mills theories. Because we will need to make use of them to understand quantum chromodynamics (Section A.7) and the Electroweak (Section 2.3) interaction, it's worth taking a moment to understand some of the key structure involved in a Yang-Mills theory. In 1954, Chen Ning Yang and Robert Mills [600] extended gauge theory for Abelian groups such as QED's $U(1)$ to non-Abelian groups. In Abelian groups, all pairs of group elements commute. In contrast, for the non-Abelian case, there exists at least one pair of generators that do not commute. Yang-Mills theories offer a unique generalization of QED which involve multiple, renormalizable, self-interacting spin-1 particles. An $SU(N)$ group has $N^2 - 1$ dimensions and hence the same number

of generators, or gauge fields, of the group. Under this symmetry, ϕ transforms as $\phi_i \to U_{ij}\phi_j$ with U an $N \times N$ unitary matrix. While we didn't have to suffer knowledge of Lie groups when discussing QED, it behooves us to introduce them briefly. A Lie group is a group which contains an infinite number of elements that are also differential manifolds. These groups all contain the identity matrix $\mathbb{1}$ and any group element continuously connected to the identity can be expressed as

$$U = e^{i\theta_a T^a}\mathbb{1}, \tag{A.18}$$

where θ_a are the numbers which parameterize the group elements and T^a are generators of the group. In the fundamental representation, these generators are a set of $N \times N$ traceless Hermitian matrices with determinant 1. The normalization convention for the generators is

$$\mathrm{Tr}(T^a T^b) = \frac{1}{2}\delta_{ab}. \tag{A.19}$$

If you are given an explicit form of the elements U of a Lie group, you can determine the generators by expanding infinitesimally about $\mathbb{1}$. What we are most interested in currently is the Lie algebra for the generators, defined using the commutation relations:

$$[T^a, T^b] = if^{abc}T^c, \tag{A.20}$$

where f^{abc} are known as the structure constants, which are totally antisymmetric with respect to the interchange of any indices. For the Abelian case, $f^{abc} = 0$ (*i.e.* they commute). In the case of $SU(2)$, which we'll see in Sections 2.3 and 2.4, $f^{abc} =$

ϵ^{abc}, the familiar Levi-Cevita symbol.

A.6 Quantum Electrodynamics

Quantum electrodynamics (QED) was QFT's first successful implementation, able to make precise predictions such as the calculation of the leading radiative correction to the electron magnetic moment ($g_e - 2 = \alpha/\pi$) [64, 601, 602]. We begin with the Dirac Lagrangian describing the dynamics of the free spin-1/2 fermion $\psi(x)$ with mass m,

$$\mathcal{L} = i\overline{\psi}(x)\gamma^\mu\partial_\mu\psi(x) - m\overline{\psi}(x)\psi(x), \tag{A.21}$$

where γ^μ ($\mu = 0, 1, 2, 3$) are the 4×4 γ matrices satisfying $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$ and $\overline{\psi}(x) = \psi^\dagger(x)\gamma_0$. We can condense this Lagrangian into a neater, more compact form:

$$\mathcal{L} = \overline{\psi}(x)(i\partial\!\!\!/ - m)\psi(x), \tag{A.22}$$

where $\partial\!\!\!/ = \gamma^\mu\partial_\mu$. This Lagrangian is invariant under the $U(1)$ local gauge transformation

$$\psi(x) \to e^{i\alpha(x)}\psi, \tag{A.23}$$

or at least it will be with a few other considerations. With $\alpha(x)$ a function of space-time, the space-time derivative does not transform simply under this local $U(1)$

transformation. We need to define the gauge field A_μ, which transforms as

$$A_\mu \to A_\mu + \frac{1}{e}\partial_\mu \alpha(x), \tag{A.24}$$

where e, the conserved electromagnetic charge, is the coupling to the electromagnetic field. Then, to make the kinetic terms invariant under the symmetry, by adding a correction term that compensates for the choice of gauge, we replace the ordinary derivative ∂_μ by a covariant derivative D_μ:

$$D_\mu = \partial_\mu - ieA_\mu. \tag{A.25}$$

And hence we obtain the proper Lagrangian which obeys our locally imposed symmetry,

$$\mathcal{L} = \overline{\psi}(i\not{D} - m)\psi. \tag{A.26}$$

These new stipulations on the invariance under a local gauge transformation work out wonderfully, as the electromagnetic coupling between the photon (the gauge field A_μ) and the fermion (ψ) pops out as a result! Typically, we re-express the EM coupling in terms of the fine-structure constant $\alpha = e^2/4\pi$. So far though, we do not have the full Lagrangian for QED. To achieve this, we add the gauge-invariant kinetic term of the photon,

$$\mathcal{L}_{\text{kinetic},\gamma} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{A.27}$$

where $F_{\mu\nu}$ is the familiar EM field strength tensor, which also employs the covariant derivative from Equation A.25 in its definition:

$$F_{\mu\nu} = -\frac{i}{e}[D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{A.28}$$

We should not just note, but *stress*, the fact that we do not have a mass term for the photon (*i.e.* $m^2 A_\mu A^\nu$) because it manifestly breaks gauge invariance. This should sit well with us however, as experimental evidence has long shown the photon is massless. Thus, our full Lagrangian for QED is:

$$\mathcal{L}_{\text{QED}} = \overline{\psi}(x)(i\not{D} - m)\psi(x) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{A.29}$$

The lessons learned in the simple construction of QED translate well to the other fundamental forces, albeit with more complications. The strong and weak nuclear forces are couched in the same formalism, as we will see in the following sections.

A.7 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the theory describing the interactions of quarks and gluons under the strong interaction. The force-mediating bosons are eight gluons, which "glue" quarks together within the nuclei of atoms. The underlying symmetry group is $SU(3)_c$ color symmetry, where our quark fields ψ with color charges (r, g, b):

$$\psi = \begin{pmatrix} \psi_r \\ \psi_g \\ \psi_b \end{pmatrix}, \tag{A.30}$$

under a local transformation, transform as

$$\psi(x) \rightarrow e^{i\eta_a(x)T^a}\psi(x). \tag{A.31}$$

Again, $\eta_a(x)\,(a = 1, 2, ..., 8)$ parameterize the group elements and $T^a = \lambda^a/2$, where λ^a are the generators of $SU(3)_c$, the eight traceless Hermitian Gell-Mann matrices:

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \tag{A.32}$$

$$\lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \lambda^8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Just as in QED, to preserve gauge invariance, we define how the gauge fields transform using Equation 2.3:

$$G^a_\mu \rightarrow G^a_\mu + \frac{1}{g_s}\partial_\mu \eta^a(x) - f^{abc}\eta^b(x)G^c_\mu, \tag{A.33}$$

where g_s is the strong coupling constant whose value depends upon the scale (or energy) of the interaction due to quantum fluctuations. The coupling g_s gets larger for low-momentum transfers (long-distance) and weaker for high-momentum transfers (short-distance) in a process known as asymptotic freedom. Typically, g_s is re-expressed in terms of α_s (similar to the fine structure constant α from QED)

where $\alpha_s = g_s^2/4\pi$. With the field transformation defined, we introduce the covariant derivative as expressed in Equation 2.5:

$$D_\mu = \partial_\mu - i\frac{g_s}{2}\lambda_a G_\mu^a. \tag{A.34}$$

And finally, the field strength tensor $G_{\mu\nu}^a$, as defined in Equation 2.6:

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g_s f^{abc} G_\mu^b G_\nu^c. \tag{A.35}$$

The last terms in Equations A.33 and A.35 contain the structure constants f^{abc} of $SU(3)$ which are defined in Equation A.20 in Appendix A.5 in a discussion of the non-Abelian gauge theories of $SU(N)$ in Yang-Mills theories. And thus, by using gauge-invariant terms involving the quark fields and gauge fields, our QCD Lagrangian becomes:

$$\mathcal{L}_{\text{QCD}} = \overline{\psi}(i\not{D} - m)\psi - \frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu}. \tag{A.36}$$

We should take note that we have no mass terms for the gluon fields, *i.e.*

$$G_{\mu\nu}^a G_a^{\mu\nu} \sim (\partial_\mu G_\nu - \partial_\nu G_\mu)^2 + g_s(\partial_\mu G_\nu - \partial_\nu G_\mu)G^\mu G^\nu + g_s^2 G_\mu G_\nu G^\mu G^\nu \tag{A.37}$$

contains no term with the form $m^2 G_\mu^a G_a^\mu$, implying massless gluons. The first term in this expression describes the gluon propagator, the second the triple gluon self-interactions, and the third the quartic interactions.

A.8 The Stability of the Higgs Potential

The measured value for the Higgs boson mass of $m_H \approx 125\,\text{GeV}$ has interesting consequences within the SM. The value of the quartic coupling λ in the Higgs potential is precisely determined by m_H ($\lambda = m_H^2/(2v^2)$). A larger mass for the Higgs boson ($m_H \gtrsim 170\,\text{GeV}$ [603]), λ would become negative at some lower energy scale, potentially well below the Plank scale ($\sim 10^{18}$) [604, 605]. For a smaller Higgs mass ($m_H \lesssim 115\,\text{GeV}$ [603]), λ would become negative at some lower energy scale and become unstable, with a lifetime shorter than the age of the Universe [606–608]. This happens because the SM couplings and fields experience an evolution once radiative corrections are included. This evolution is called the renormalization group (RG) evolution [609], and it can cause our theory to become theoretically inconsistent. In the SM, the renormalization group evolution of λ is dependent on the logarithm of the square of the energy scale μ. As the coupling of the Higgs bosons to fermions and vector bosons is determined by their respective masses, the RG evolution is dominated by contributions from the massive vector bosons and the top quark. The SM RG evolutions of the gauge couplings and the quartic coupling λ are in Figure A.2. Interestingly, for the experimentally measured value for the Higgs boson mass, λ is close to zero at the Plank scale and the electroweak vacuum of the Higgs potential is likely meta-stable [144, 610, 611]. The high-energy evolution of λ shows it becomes negative at energies $\Lambda \sim \mathcal{O}(10^{18}\,GeV)$, calling the long-term existence of the EW vacuum into question. The stability of the vacuum, as a function of the top quark and Higgs masses, is shown in Figure A.3. The accepted Higgs boson mass implies that the lifetime of the SM vacuum will likely exceed the current age of the Universe by several orders of magnitude, but the behavior of the potential at high energy

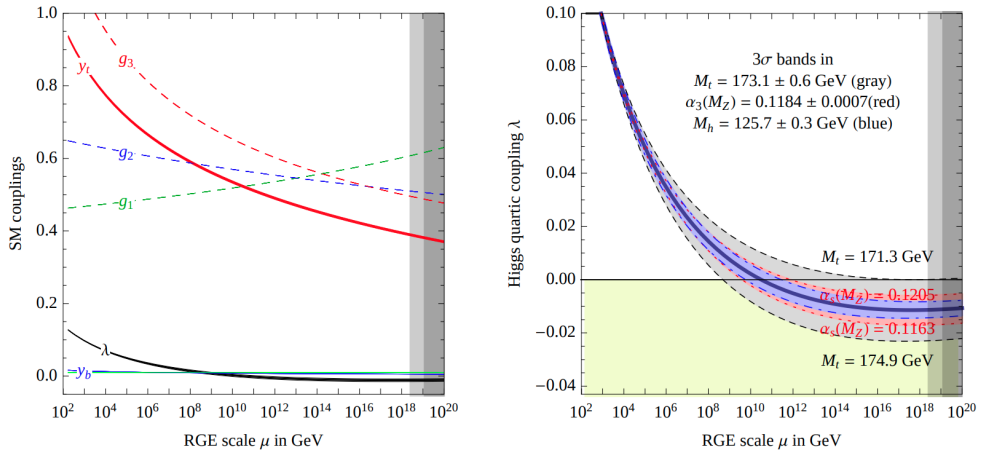


Figure A.2: (left) SM RG evolution of the gauge couplings $g_1 = \sqrt{5/3}g'$, $g_2 = g$, $g_3 = g_s$, of the top and bottom Yukawa couplings (y_t, y_b), and of the Higgs quartic coupling λ. All couplings are defined in the $\overline{\text{MS}}$ scheme. The line thickness indicates the $\pm\sigma$ uncertainty. (right) RG evolution of λ varying M_t, M_H, and α_s by $\pm 3\,\sigma$ [144].

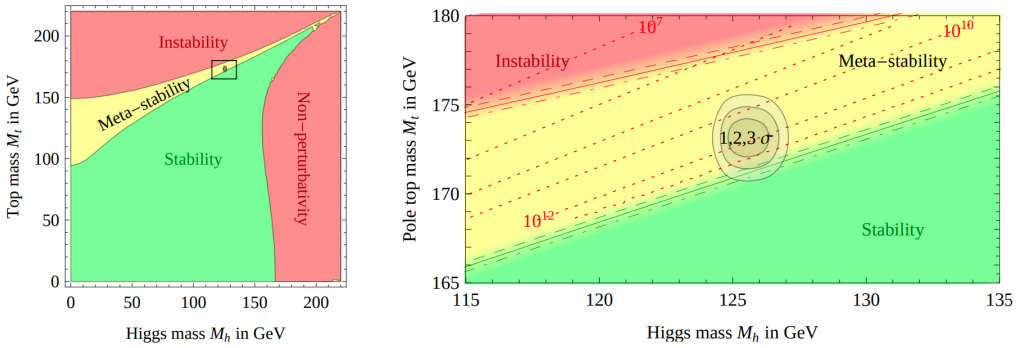


Figure A.3: Regions of absolute stability, meta-stability, and instability of the SM vacuum in the $M_t - M_H$ plane (upper left) and in the $\lambda - y_t$ plane, in terms of parameter renormalized at the Plank scale (upper right). **Right**: Zoom in of the region of preferred experimental range of M_H and M_t (the gray areas denote the allowed region at 1, 2, and 3 σ). The three boundary lines correspond to $\alpha_s(M_Z) = 0.1184 \pm 0.0007$, and the grading of the colors indicates the size of the theoretical error. The dotted contour-lines show the instability scale Λ in GeV assuming $\alpha_s(M_Z) = 0.1184$ [144].

scales could call for new physics at an intermediate scale, below the Planck scale, before the instability develops. A better measurement of the masses and couplings tells us if we are truly in a meta-stable point. For more on implications on the electroweak vacuum stability, see Reference [612]. The consequences of the instability of the EW vacuum on inflation in reference [613, 614].

A.9 The Hierarchy Problem

An interesting unanswered question in the SM is that of the Hierarchy Problem [145, 146]; an issue in which the Higgs is the central figure. The question to ask is why the weak scale ($\mathcal{O} \simeq 100\,\text{GeV}$) is so much smaller than the Plank Scale ($\mathcal{O} \sim 10^{19}\,\text{GeV}$)? As discussed in Section 2.4, the weak scale is given by the vev of the Higgs ($v \approx 246\,\text{GeV}$). Does this imply that the SM is a good effective theory at the weak scale, requiring New Physics at some higher scale? We can see the problem more clearly by examining the corrections to the Higgs mass. At loop level, the Higgs mass receives corrections from fermion loops, gauge loops, and self interactions as depicted diagrammatically in Figure A.4. For some cutoff scale for new physics Λ,

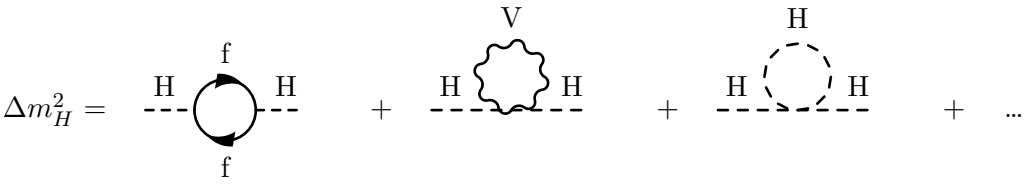


Figure A.4: Feynman diagrams showing the loop contributions to the Higgs mass corrections for femions f, vector bosons V, and Higgs bosons H.

these loops diverge quadratically as $\int d^4\,k(k^2 - m^2)^{-1} \sim \Lambda^2$ [145]: With the Higgs mass sensitive to any high scale in the theory, we must either accept that there is considerable fine-tuning order-by-order in perturbation theory – which implies spectacular cancellations to one part in 10^{28} – or, that we have new physics at the TeV scale which can eliminate these large loop corrections [146]. Given that the LHC allows us to glimpse into the TeV scale, it behooves us to take the opportunity to probe this high scale in the search for the missing pieces.

Appendix B

Detector Appendix

B.1 ATLAS Inner Detector

The original pixel detector consists of three barrel layers (Layer 1, Layer 2, Layer 3) and two end-caps with three disks each. The radii of the three barrel layers are 50.5 mm, 88.5 mm, 122.5 mm, respectively. An additional layer, added for Run 2, is the Insertable B-Layer (IBL). The IBL is located closer to the beamline ($r = 33.25\,\text{mm}$) to provide higher tracking resolution. The detector is comprised of 1744 pixel-sensor modules ($10\,\text{cm}^2$) with each module containing 46,080 pixels $(r - \phi) \times z = 50\,\mu\text{m} \times 400\,\mu\text{m}$ in size, providing a resolution of $(r - \phi) \times z = 10\,\mu\text{m} \times 115\,\mu\text{m}$. In total, the IBL and pixel detector are composed of nearly 80 million pixels. Table B.1 shows some summary values of the basic parameters for the pixel barrel layers.

The Semiconductor Tracker (SCT) [296, 297] is composed of 4088 two-sided silicon micro-strip modules arranged in four barrel layers (SCT1, SCT2, SCT3, SCT4) [311] and two end-caps with nine wheels each. The barrel layers are located at the radii of 299 mm, 371 mm, 443 mm, and 514 mm, respectively. In all, the SCT

Table B.1: Basic parameters for the barrel region of the ATLAS IBL and pixel detector system [309, 615, 616].

Layer	Mean Radius [mm]	Number of Staves	Number of Modules	Number of Channels	Active Area [m^2]
B-Layer	33.25	14	280	$\sim$ 12,000,000	0.15
Layer 1	50.5	22	286	$\sim$ 13,200,000	0.28
Layer 2	88.5	38	494	$\sim$ 22,800,000	0.49
Layer 3	122.5	52	676	$\sim$ 31,100,000	0.67
Total		126	1736	$\sim$ 79,100,000	1.59

has over 6 million readout strips. Table B.2 shows summary values for the basic parameters in the SCT barrel layers.

Table B.2: Basic parameters for the barrel region of the ATLAS SCT detector system [617]. (Note: the tilt angle is the angle of the modules relative to the local tangent to the surface of their supporting cylinder)

Layer	Mean Radius [mm]	Length (full) [mm]	Tilt Angle (in ϕ) [deg]	Number of Modules
SCT1	299	1492	11	384
SCT2	371	1492	11	480
SCT3	443	1492	11.25	576
SCT4	514	1492	11.25	672
Total				2112

B.2 ATLAS Muon Spectrometer

A summary table of the detector types, their locations, coverage in η, and the number of readout channels can be found in Table B.3.

Table B.3: ATLAS Muon Spectrometer detector types, their purpose, location, coverage in η, and the number of channels of each [330].

Type	Purpose	Location	Coverage	Channels
MDT	Tracking	Barrel/EndCap	$0.0 < \|\eta\| < 2.7$	354k
CSC	Tracking	EndCap Lyr 1	$2.0 < \|\eta\| < 2.7$	30.7k
RPC	Trigger	Barrel	$0.0 < \|\eta\| < 1.0$	373k
TGC	Trigger	EndCap	$1.0 < \|\eta\| < 2.4$	318k

B.3 A Holistic View of the ATLAS Detector

Considering what we've seen with the various detector subsystems, we can step back and take in a more holistic picture of which particles can be observed in the detector – and where. Charged particles such as electrons, muons, and charged hadrons will all leave traces through ionization in the pixel, SCT, and TRT. Electrons and photons will interact with the dense layers of the ECal, being fully absorbed. Muons, because they are significantly heavier than electrons and do not interact strongly, will pass through the calorimeters with little energy loss, but they leave traces in the muon spectrometer, which allows for accurate measurements of its momentum. Protons and other heavy charged hadrons will leave tracks in the ID, and will deposit a fraction of their energy in the ECal, finally being stopped in the dense materials of the HCal. Neutrons and other neutral heavy hadrons, however, will pass through the ID and ECal undetected due to their neutral charge, depositing almost all of their energy through interactions in the HCal. Neutrinos, uncharged and only participating in weak interactions, will escape the detector without a trace, leaving us to only infer their four-momenta. These interactions can be seen in Figure B.1. The differences in how the particles are detected and interact with the various components of ATLAS have a strong effect on our ability to reconstruct physics objects.

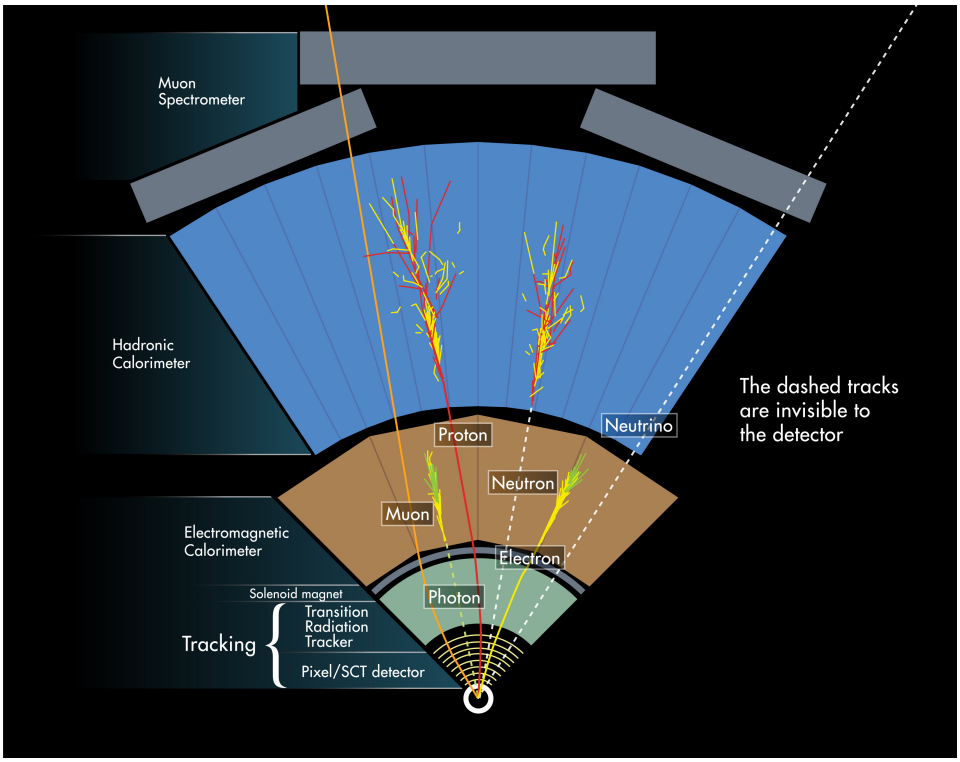


Figure B.1: The types of trails left by various particles in the ATLAS detectors [618].

B.4 Trigger

B.4.1 Jet and Muon Trigger Tables

Seven isolation criteria are defined, each optimized for the various physics analyses. A list of the isolation WPs and their discriminating variables can be found in Table B.6.

B.4.2 Jet Trigger Studies

The efficiency of the triggers is studied both in data and Pythia 8 QCD Monte Carlo. The least prescaled single large-R-jet trigger is used to select the events used in the analysis. The trigger plateau is studied using the leading large-R jet in the event. This assumes that the trigger efficiency is a property of kinematics only.

The 2015 data used single large-R p_{T} triggers of ungroomed and reclustered

Table B.4: Summary of the unprescaled large-R jet triggers [403]. The triggers used in the analysis are highlighted in blue.

Year	Trigger	Offline Threshold [GeV]	Luminosity [fb^{-1}]
2015	HLT_j360_a10_lcw_sub_L1J100	$p_{\mathrm{T,J}} > 410$	3.2
	HLT_j360_a10r_L1J100	$p_{\mathrm{T,J}} > 420$	3.2
2016	HLT_j420_a10_lcw_L1J100	$p_{\mathrm{T,J}} > 450$	33.0
	HLT_j420_a10r_L1J100	$p_{\mathrm{T,J}} > 470$	33.0
2017	HLT_j390_a10t_lcw_jes_30smcINF_L1J100	$p_{\mathrm{T,J}} > 420, m_{\mathrm{J}} > 50$	41.0
	HLT_j420_a10t_lcw_jes_40smcINF_L1J100	$p_{\mathrm{T,J}} > 450, m_{\mathrm{J}} > 70$	44.3
	HLT_j440_a10_lcw_subjes_L1J100	$p_{\mathrm{T,J}} > 470$	41.0
	HLT_j440_a10r_L1J100	$p_{\mathrm{T,J}} > 480$	41.0
	HLT_j440_a10t_lcw_jes_L1J100	$p_{\mathrm{T,J}} > 470$	41.2
	HLT_j460_a10_lcw_subjes_L1J100	$p_{\mathrm{T,J}} > 490$	44.3
	HLT_j460_a10_lcw_subjes_L1SC111	$p_{\mathrm{T,J}} > 490$	44.3
	HLT_j460_a10r_L1J100	$p_{\mathrm{T,J}} > 500$	44.3
	HLT_j460_a10r_L1SC111	$p_{\mathrm{T,J}} > 500$	44.3
	HLT_j460_a10t_lcw_jes_L1J100	$p_{\mathrm{T,J}} > 490$	44.3
	HLT_j460_a10t_lcw_jes_L1SC111	$p_{\mathrm{T,J}} > 490$	44.3
2018	HLT_j460_a10t_lcw_jes_L1SC111	$p_{\mathrm{T,J}} > 490$	58.5
	HLT_j460_a10r_L1J100	$p_{\mathrm{T,J}} > 500$	58.5
	HLT_j460_a10t_lcw_jes_L1J100	$p_{\mathrm{T,J}} > 490$	58.5
	HLT_j420_a10t_lcw_jes_30smcINF_L1J100	$p_{\mathrm{T,J}} > 450, m_{\mathrm{J}} > 50$	39.9
	HLT_j420_a10t_lcw_jes_35smcINF_L1J100	$p_{\mathrm{T,J}} > 450, m_{\mathrm{J}} > 60$	58.5
	HLT_j420_a10t_lcw_jes_35smcINF_L1SC111 m_{J}	$p_{\mathrm{T,J}} > 450, m_{\mathrm{J}} > 60$	55.4

Table B.5: Sequence for the muon trigger chains at L1 and HLT for pp collision data. The p_{T} and isolation cuts applied at each step of the chain are also shown. The isolation requirement was updated in 2018 by tightening the Δz selection on additional tracks to reduce inefficiencies observed in the high-pileup conditions during data-taking in 2017. The threshold on the tri-muon trigger was increased from 4 GeV to 6 GeV in 2018 [619]. See Section 6.3 for the definitions of CB, SA, and FS.

Trigger Chain	Level-1	HLT
	Single-Muon Triggers	
HLT_mu26_ivarmedium	L1_MU20	≥ 1 CB muon with $p_{\mathrm{T}} > 26$ GeV and $\Sigma_{\Delta z<6(2)\,\mathrm{mm}} p_{\mathrm{T}}^{\mathrm{trk}}/p_{\mathrm{T}} < 0.07$
HLT_mu50	L1_MU20	≥ 1 CB muon with $p_{\mathrm{T}} > 50$ GeV
HLT_mu50_0eta105_msonly	L1_MU20	≥ 1 SA muon with $p_{\mathrm{T}} > 60$ GeV in $\lvert\eta\rvert < 1.05$
	Multi-muon Triggers	
HLT_2mu14	L1_2MU10	≥ 2 CB muons with $p_{\mathrm{T}} > 14$ GeV
HLT_mu22_mu8noL1	L1_MU20	≥ 1 CB muon with $p_{\mathrm{T}} > 22$ GeV (mu22 trigger) and ≥ 2 FS muons with $p_{\mathrm{T}} > 22$ and > 8 GeV
HLT_3mu4(6)	L1_3MU4(6)	≥ 3 CB muons with $p_{\mathrm{T}} > 4(6)$ GeV
HLT_3mu6_msonly	L1_3MU6	≥ 3 SA muons with $p_{\mathrm{T}} > 6$ GeV
HLT_mu20_2mu4noL1	L1_mu20	≥ 1 CB muon with $p_{\mathrm{T}} > 20$ GeV (mu20 trigger) and ≥ 1 FS muon with $p_{\mathrm{T}} > 20$ GeV and ≥ 2 FS muons with $p_{\mathrm{T}} > 4$ GeV

Table B.6: Definition of the seven isolation working points. The discriminating variables are listed in the second column and the criteria used in the definition are in the third [365].

Isolation WP	Discriminating variable(s)	Definition
LooseTrackOnly	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu}$	99% efficiency constant in η and p_{T}
Loose	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu}$, $E_T^{\mathrm{topocone20}}/p_{\mathrm{T}}^{\mu}$	99% efficiency constant in η and p_{T}
Tight	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu}$, $E_T^{\mathrm{topocone20}}/p_{\mathrm{T}}^{\mu}$	96% efficiency constant in η and p_{T}
Gradient	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu}$, $E_T^{\mathrm{topocone20}}/p_{\mathrm{T}}^{\mu}$	$\geq 90(99)\%$ efficiency at 25 (60) GeV
GradientLoose	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu}$, $E_T^{\mathrm{topocone20}}/p_{\mathrm{T}}^{\mu}$	$\geq 95(99)\%$ efficiency at 25 (60) GeV
FixedCutTightTrackOnly	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu}$	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu} < 0.06$
FixedCutLoose	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu}$, $E_T^{\mathrm{topocone20}}/p_{\mathrm{T}}^{\mu}$	$p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu} < 0.15$, $E_T^{\mathrm{topocone20}}/p_{\mathrm{T}}^{\mu} < 0.30$

varieties. Figure B.2 shows the efficiency of the 2015 triggers in data, fitted with the Fermi function. The data vs. MC comparison is shown in Figure B.3. Unfortunately, there is not a prescaled large-R jet trigger available in 2015 for efficiency studies. Instead, the most prescaled anti-k_t $R = 0.4$ jet trigger, `HLT_110`, is used for the reference.

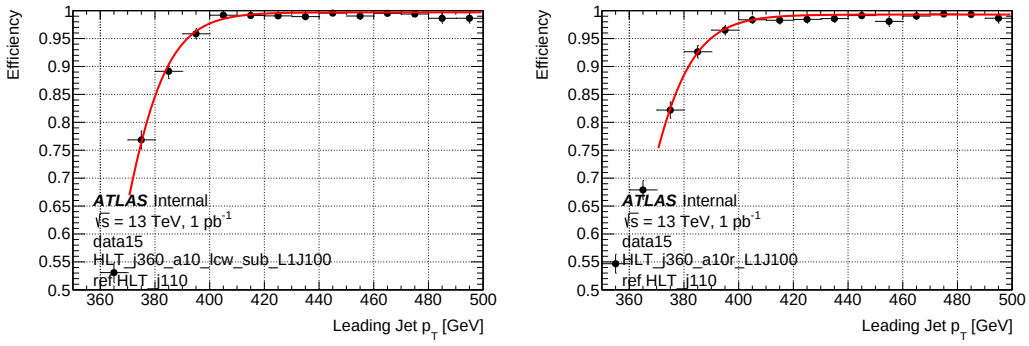


Figure B.2: Trigger efficiency for the single large-R jet-p_{T} trigger in 2015 data-taking. Triggers for ungroomed (left) and reclustered (right) versions are shown. The efficiency curves are fitted with a Fermi function to estimate the 99% efficiency point.

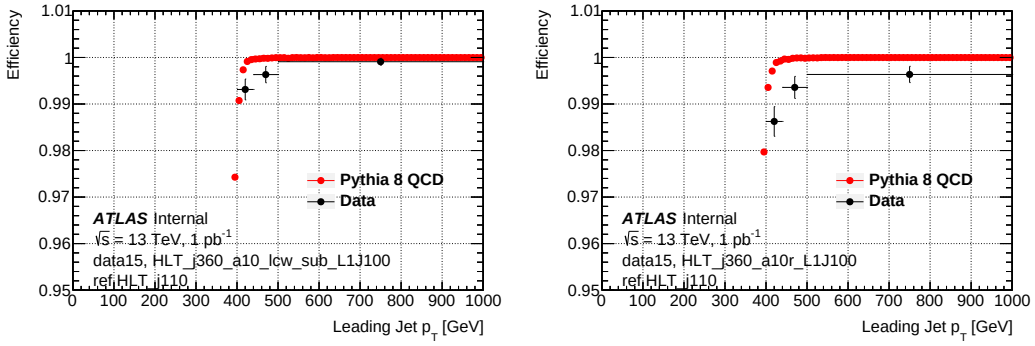


Figure B.3: Trigger efficiency comparison between data and PYTHIA 8 QCD Monte Carlo for the single large-R jet-p_{T} triggers in 2015 data-taking. Triggers for ungroomed (left) and reclustered (right) versions are shown.

The 2016 data used single large-R jet-p_{T} triggers of ungroomed and reclustered varieties. Figure B.4 shows the efficiency of the 2016 triggers in data, fitted with the Fermi function. The data vs. MC comparison is shown in Figure B.5.

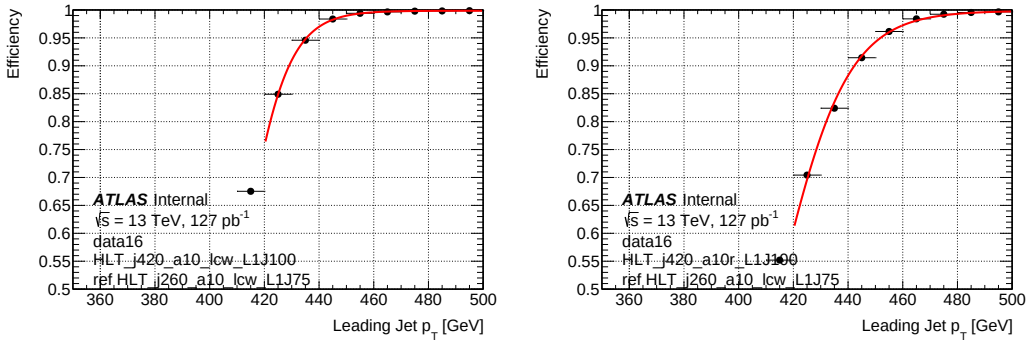


Figure B.4: Trigger efficiency for the single large-R jet-p_{T} trigger in 2016 data-taking. Triggers for ungroomed (left) and reclustered (right) versions are shown. The efficiency curves are fitted with a Fermi function to estimate the 99% efficiency point.

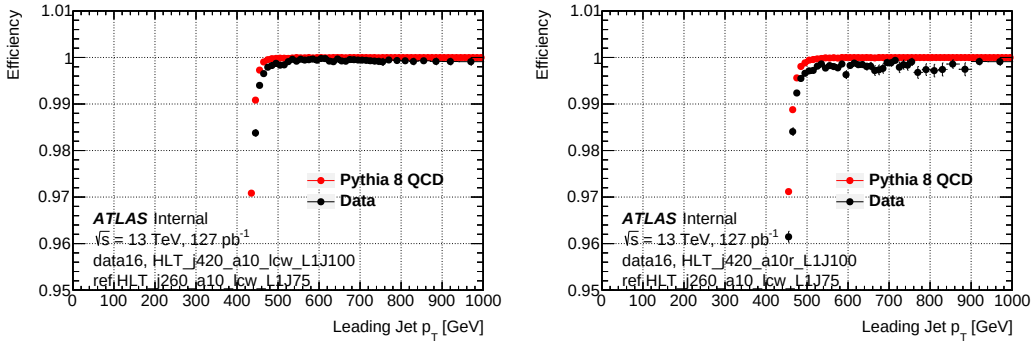


Figure B.5: Trigger efficiency comparison between data and PYTHIA 8 QCD Monte Carlo for the single large-R jet p_{T} triggers in 2016 data-taking. Triggers for ungroomed (left) and reclustered (right) versions are shown.

The 2017 data used single large-R jet-p_{T} and p_{T} + mass triggers. Both the trimming algorithm and the p_{T} + mass triggers were introduced in this year. Figures B.6 and B.8 show the efficiency of the 2017 triggers in data, fitted with the Fermi function. The data vs. MC comparison is shown in Figures B.7 and B.9.

The 2018 data used a mix of single large-R jet-p_{T} and p_{T} + mass triggers. Partway through this year, the simple cone triggers at L1 (L1SC111) were introduced to improve the L1 trigger efficiency for large-R jets. The recommendation is to use an OR of the L1J100 and L1SC111 chains. The `HLT_j260_a10t_lcw_jes_L1J75` trigger was used as the reference trigger in all cases, including the L1SC111-based chains. Figures B.10, B.12 and B.14 show the efficiency of the 2018 p_{T} and p_{T} + mass triggers in data, fitted with the Fermi function. The data vs. MC comparison is shown in figures B.11, B.13 and B.15. For the p_{T} + mass triggers, there is small decrease in efficiency for high-p_{T} jets. This is possibly due to the HLT using calorimeter mass triggers and offline using combined mass. This inefficiency at high p_{T} is recovered in the analysis by OR'ing with the p_{T}-only trigger, as seen in Figure B.16.

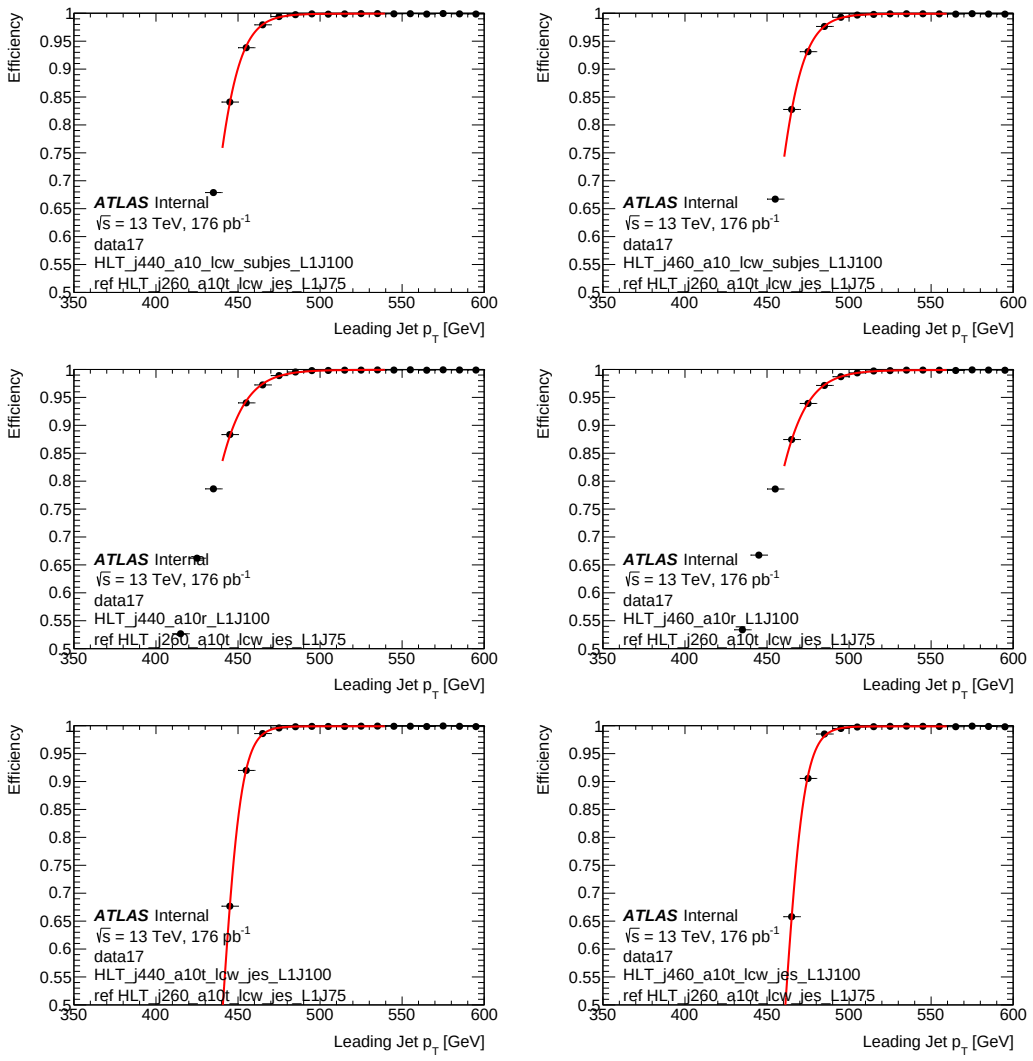


Figure B.6: Trigger efficiency for the single large-R jet-p_{T} trigger in 2017 data-taking. Triggers using online p_{T} of 440 GeV (left) and 460 GeV (right) are shown for ungroomed (top), reclustered (middle) and trimmed (bottom) versions. The efficiency curves are fitted with a Fermi function to estimate the 99% efficiency point.

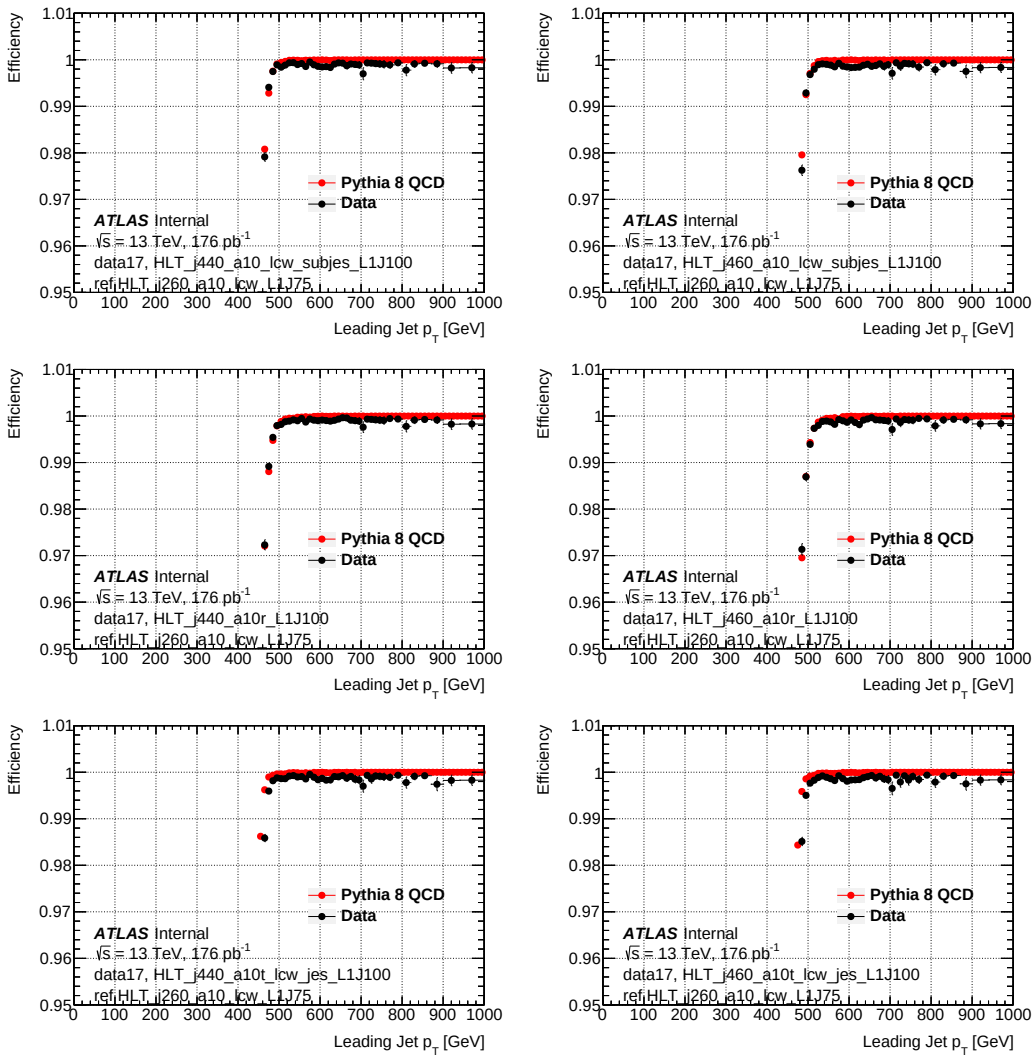


Figure B.7: Trigger efficiency comparison between data and PYTHIA 8 QCD Monte Carlo for the single large-R jet-p_{T} triggers in 2017 data-taking. Triggers using online p_{T} of 440 GeV (left) and 460 GeV (right) are shown for ungroomed (top), reclustered (middle) and trimmed (bottom) versions.

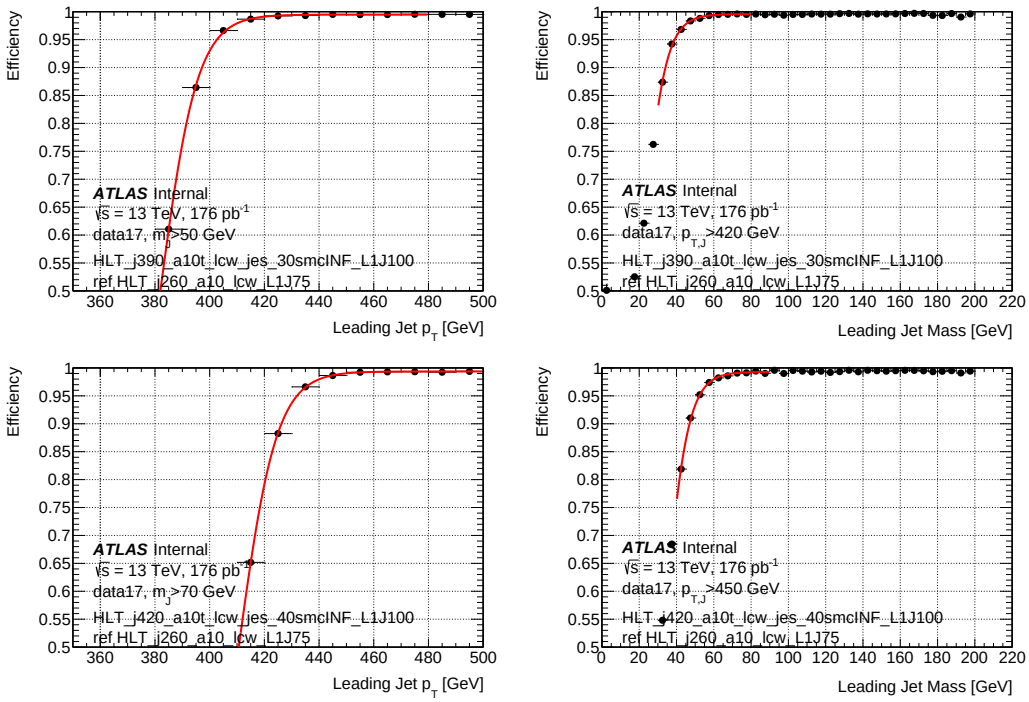


Figure B.8: Trigger efficiency for the single large-R jet-p_{T} + mass triggers in 2017 data-taking. The efficiency curves are fitted with a Fermi function to estimate the 99% efficiency point.

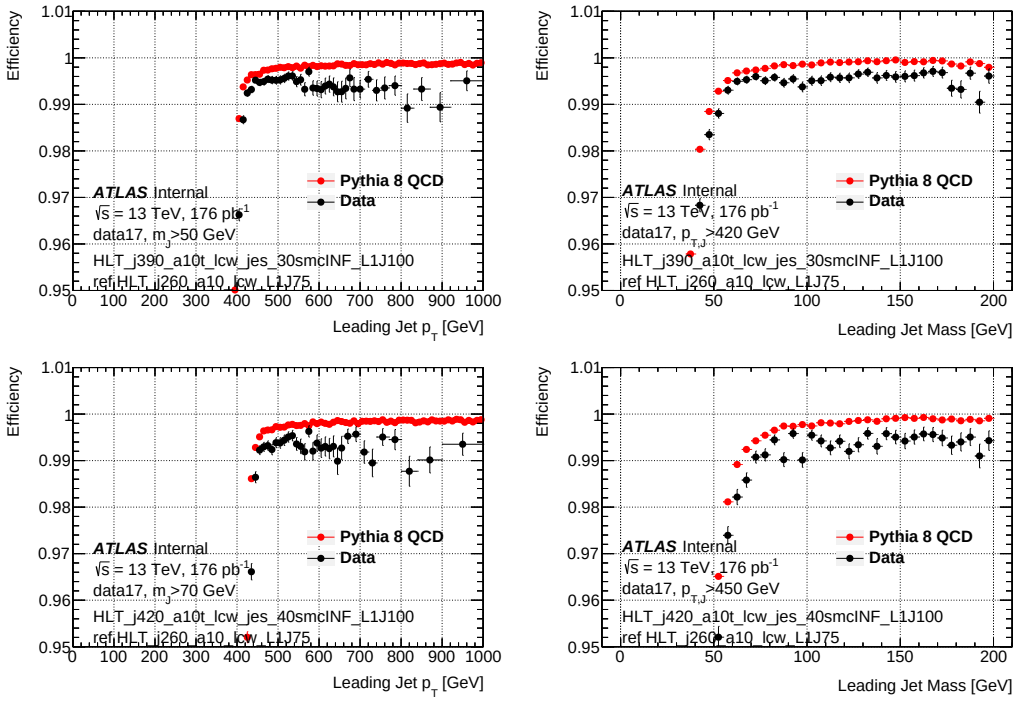


Figure B.9: Trigger efficiency comparison between data and Pythia 8 QCD Monte Carlo for the single large-R jet-p_{T} + mass triggers in 2017 data-taking.

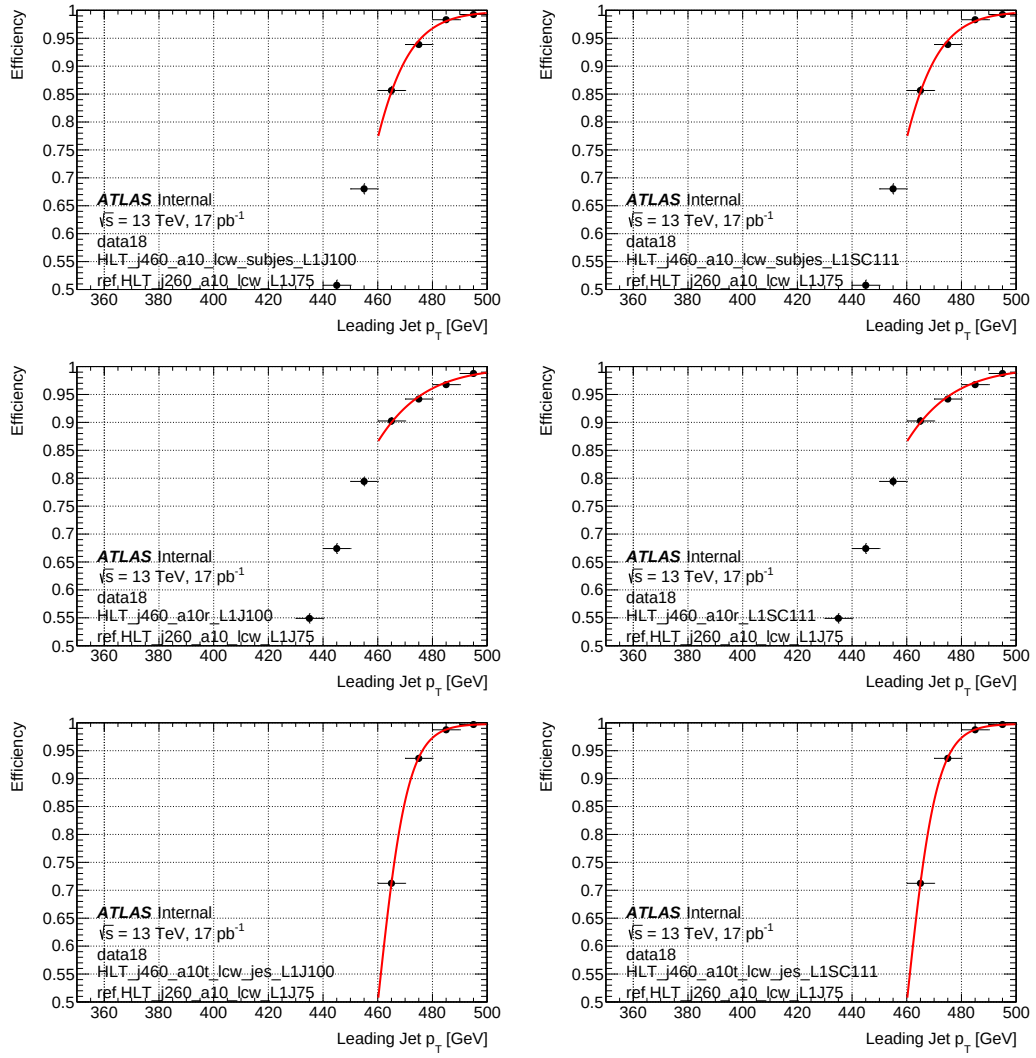


Figure B.10: Trigger efficiency for the single large-R jet p_{T} trigger in 2018 data-taking. The efficiency curves are fitted with a Fermi function to estimate the 99% efficiency point.

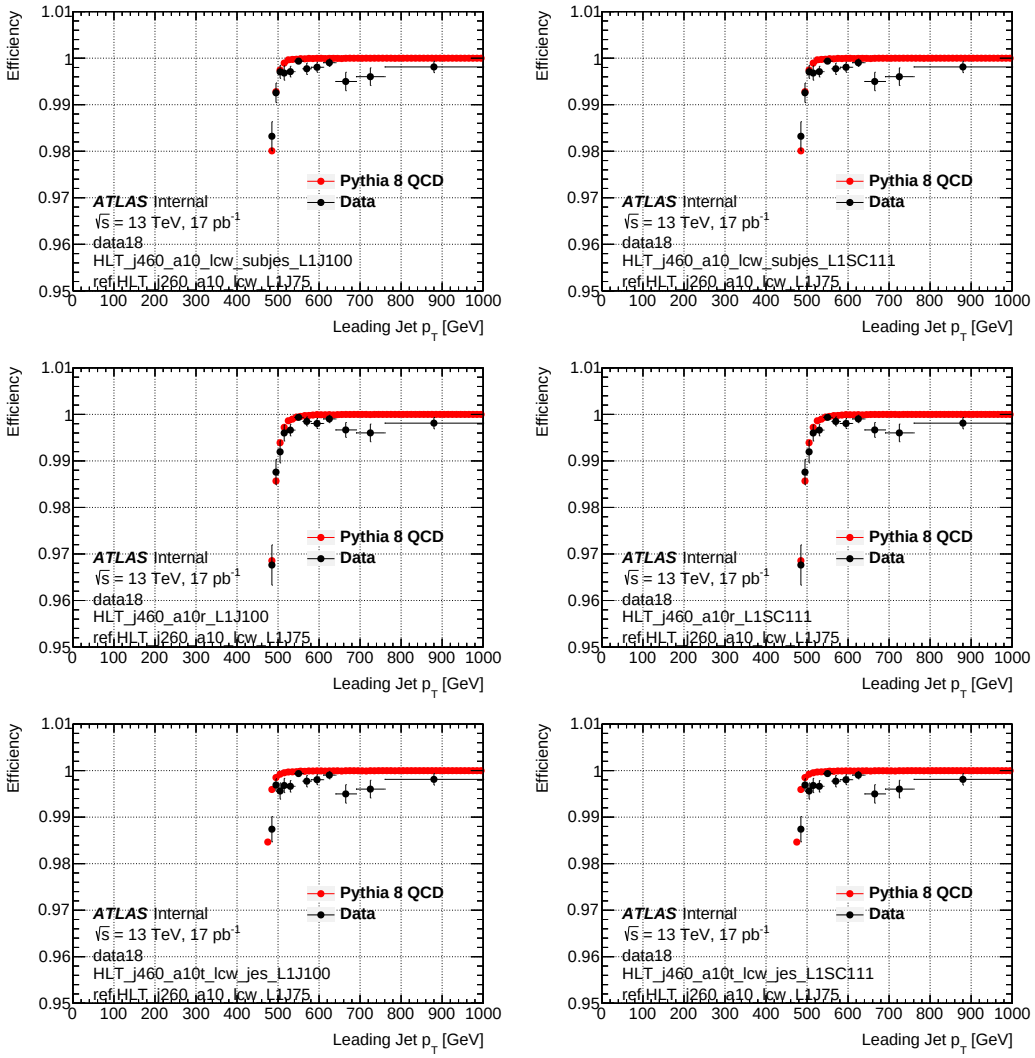


Figure B.11: Trigger efficiency comparison between data and PYTHIA 8 QCD Monte Carlo for the single large-R jet p_{T} triggers in 2018 data-taking. The L1J100 (left) and L1SC111 (right) are shown.

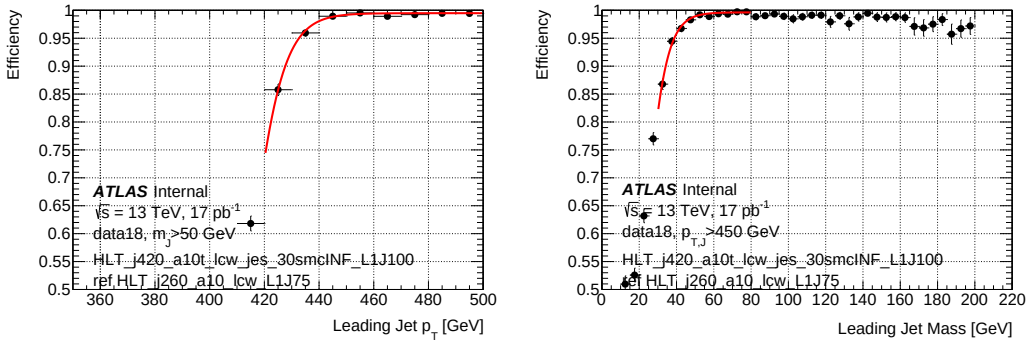


Figure B.12: Trigger efficiency for the single large-R jet p_{T} + mass trigger requiring an HLT large-R jet with online $p_{\mathrm{T}} > 420\,\mathrm{GeV}$ and $m > 30\,\mathrm{GeV}$ in 2018 data-taking. The efficiency with respect to offline p_{T} (left) and mass (right) is shown and fitted with a Fermi function to estimate the 99% efficiency point.

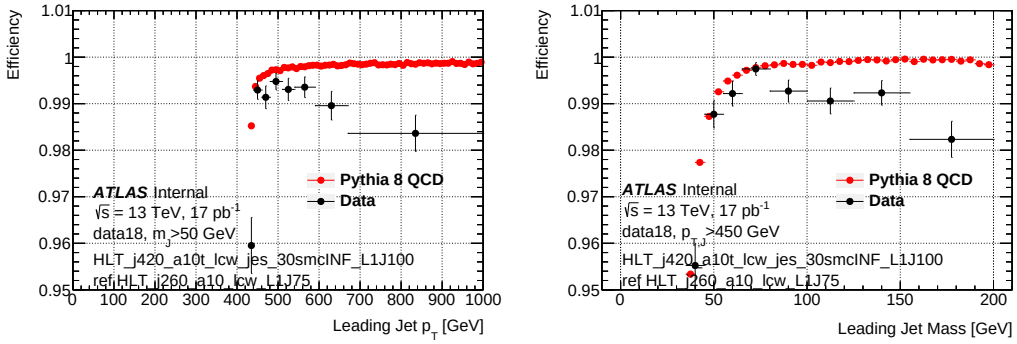


Figure B.13: Trigger efficiency comparison between data and PYTHIA 8 QCD Monte Carlo for the single large-R jet p_{T} + mass trigger requiring an HLT large-R jet with online $p_{\mathrm{T}} > 420\,\mathrm{GeV}$ and $m > 30\,\mathrm{GeV}$ in 2018 data-taking. The efficiency with respect to offline p_{T} (left) and mass (right) is shown.

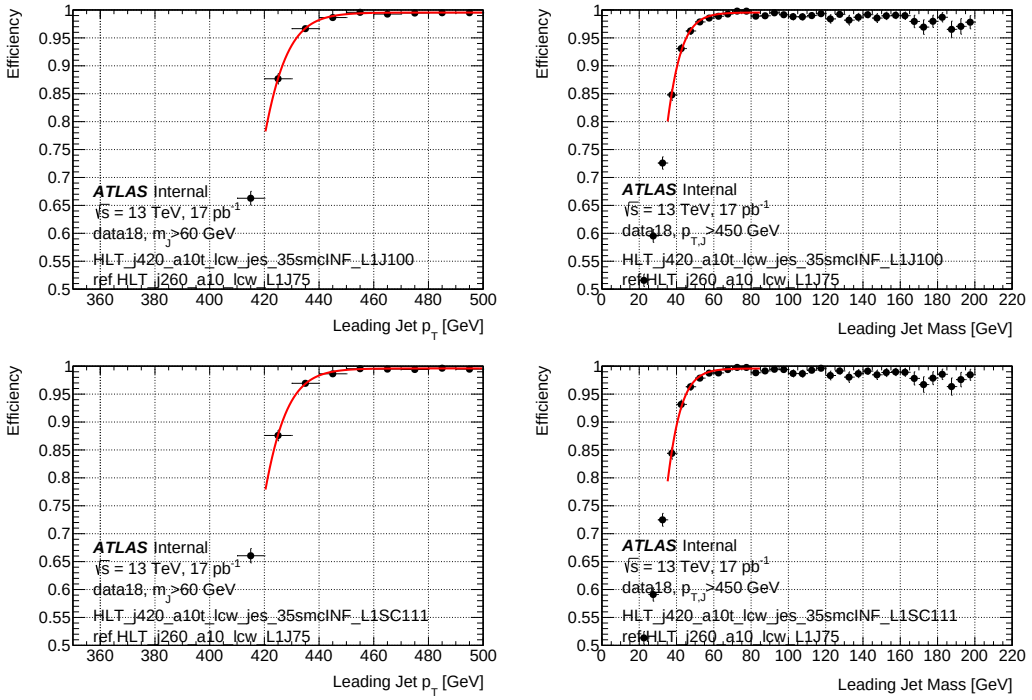


Figure B.14: Trigger efficiency for the single large-R jet p_{T} + mass trigger requiring an HLT large-R jet with online $p_{\mathrm{T}} > 420\,\mathrm{GeV}$ and $m > 35\,\mathrm{GeV}$ in 2018 data-taking. The L1J100 (top) and L1SC111 (middle) chains are shown. The L1 The efficiency with respect to offline p_{T} (left) and mass (right) is shown. The efficiency curves are fitted with a Fermi function to estimate the 99% efficiency point.

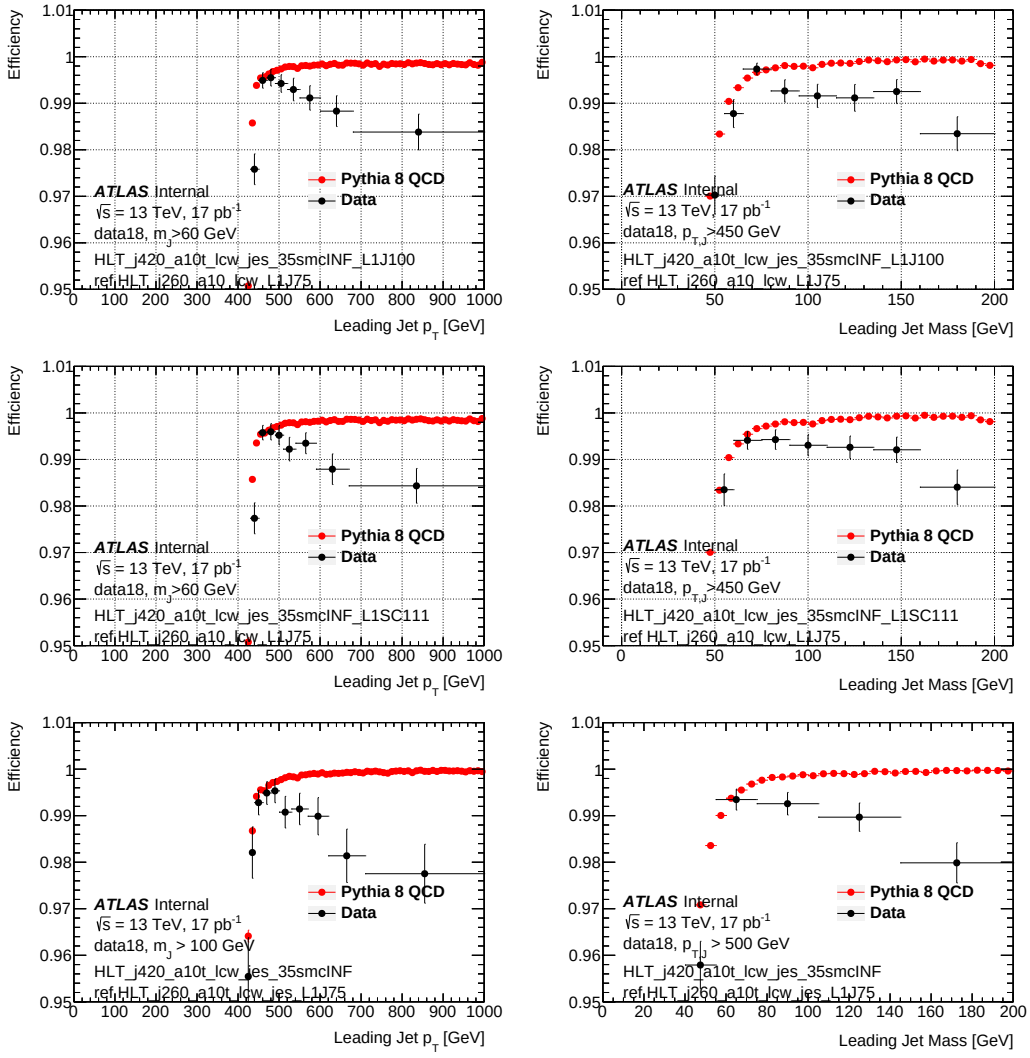


Figure B.15: Trigger efficiency comparison between data and PYTHIA 8 QCD Monte Carlo for the single large-R jet p_{T} + mass trigger requiring an HLT large-R jet with online $p_{\mathrm{T}} > 420\,\mathrm{GeV}$ and $m > 35\,\mathrm{GeV}$ in 2018 data-taking. The L1J100 (top), L1SC111 (middle) and OR of the two (bottom) chains are shown. The efficiency with respect to offline p_{T} (left) and mass (right) is shown.

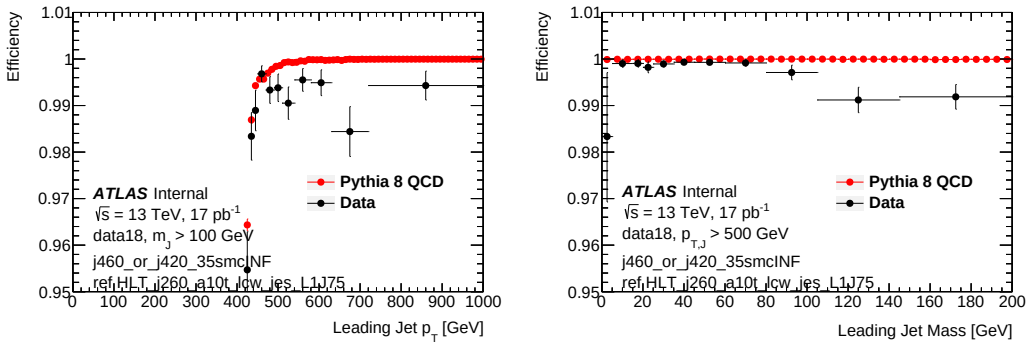


Figure B.16: Trigger efficiency comparison between data and PYTHIA 8 QCD Monte Carlo for the OR of the unprescaled single large-R jet p_{T} and p_{T} + mass triggers; `HLT_j420_a10t_lcw_jes_35smcINF` and `HLT_j460_a10t_lcw_jes`. The efficiency with respect to offline p_{T} (left) and mass (right) is shown.

Appendix C

Analysis Appendix

C.1 Good Runs List

The following GRL's, targeting data reprocessed in release 21, were used:

- data15_13TeV.periodAllYear_DetStatus-v89-pro21-02_Unknown_PHYS_StandardGRL_All_Good_25ns.xml
- data16_13TeV.periodAllYear_DetStatus-v89-pro21-01_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns.xml
- data17_13TeV.periodAllYear_DetStatus-v99-pro22-01_Unknown_PHYS_StandardGRL_All_Good_25ns_ Triggerno17e33prim.xml
- data18_13TeV.periodAllYear_DetStatus-v102-pro22-04_Unknown_PHYS_StandardGRL_All_Good_25ns_ Triggerno17e33prim.xml

C.2 Simulated Signal and Background Samples

All simulated events are run through the full detector simulation using GEANT [620] and overlayed with minimum bias events simulated using PYTHIA 8.1 [621] with the A3 set of tunes and the NNPDF23LO [440] parton distribution function to account for pileup. The mc16a (2015 + 2016 data pileup profile), mc16d (2017 data pileup profile), and mc16e (2018 data pileup profile) campaigns

are used, with each reweighted to match the observed pileup profile in data to the corresponding years.

C.2.1 Simulated Signal Samples

The analysis is optimized for a signal consisting of a single Standard Model Higgs boson decaying to two b-quarks produced in association with a jet, $pp \to H(\to b\bar{b}) + j$. Higgs events were simulated in the four main production modes discussed in Section 3.1: gluon-gluon fusion (ggF), vector boson fusion (VBF), associated vector boson production (VH), and Higgs production in association with a top-quark pair (ttH) – all of which constitute a non-negligible contribution to the signal.

The ggF + jet events are generated at NLO in QCD using the HJ+MiNLO prescription with finite mass effects using POWHEG-BOX V2 [622] and the NNPDF30 NNLO parton distribution function. The implementation discussed in Ref. [623] and the mass effects include only the top contribution, implemented through a rescaling of the EFT result by the spectrum calculated to LO accuracy at 1-loop level (missing real emissions and virtual contributions where the gluon energy is not small with respect to the top mass). Events are showered using PYTHIA 8.212 [624] with the AZNLO tune and the CTEQ6L1 [625] parton distribution function. The decay of b-hadrons is performed using EVTGEN [626]. To obtain a large sample at high p_{T}, POWHEG was configured only to output events with Born-k_T above 200 GeV using the *bornktmin* setting.

VBF Higgs boson events are generated using POWHEG-BOX V2 [627] with the NNPDF30 NLO parton distribution function. The showering is performed using PYTHIA 8.230 [624] with the AZNLO tune and CTEQ6L1 [625] parton distribution

function. The decay of b-hadrons is performed using EvtGen [626].

The Higgs boson events produced in association with a W or Z boson are generated using Powheg-Box V2 at NLO in QCD [622] and the NNPDF30 NLO parton distribution function. They also include the $gg \to ZH$ contribution calculated at leading order. The events are showered using Pythia 8.240 [624] for hadronic and 8.212 for leptonic decays of the vector boson with the AZNLO tune and the CTEQ6L1 [625] parton distribution function. The decay of b-hadrons is performed using EvtGen [626].

The Higgs boson events produced in association with two top quarks are generated using Powheg-Box V2 at NLO in QCD [628] in the five-flavor scheme with the NNPDF30 NLO parton distribution function. The samples are split into all-hadronic, semileptonic and dileptonic slices based on the decay of the two top quarks. All decay modes of the Higgs particle, with the theoretically predicted branching ratios, are included. The events are showered using Pythia 8.230 [624] with the AZNLO tune and the CTEQ6L1 parton distribution function. The decay of b-hadrons is performed using EvtGen [626]. The cross sections from the ATLAS generators differ from the state-of-the-art calculations [420] by a factor of 0.71. The ATLAS samples were therefore scaled by this factor (see Figure C.1).

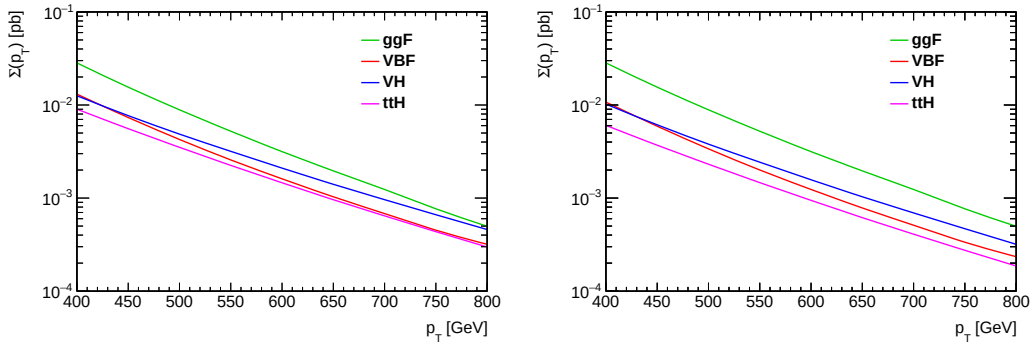


Figure C.1: Cumulative cross section as a function of the Higgs p_{T} for the different production modes as predicted by the ATLAS samples (left) and after all corrections (*i.e.* NLO EW) (right) [420].

C.2.2 Simulated Background Samples

Monte Carlo samples of the expected backgrounds are used during the development of the non-resonant background estimation procedure and to model the resonant backgrounds themselves. To develop a background estimation method for the QCD background, simulated QCD dijet events are generated using the Pythia 8.235 [624] generator with the A14 tune and the NNPDF23 LO PDF [629]. The decay of b-hadrons is performed using EvtGen [626]. The effective luminosity of these samples is approximately 5% of the actual data. To maintain a constant statistical power over a large energy range, the weighted events are generated with a flat jet p_{T} spectrum and split into several slices based on p_{T} of the leading truth jet reconstructed using the anti-k_t $R = 0.6$ jet algorithm ran on the final-state truth particles. However, even then, the samples are statistically smaller than the current Run 2 dataset.

The hadronically decaying W and Z events, with 1 additional parton at next-to leading order accuracy and up to 4 additional partons at leading order, were generated using Sherpa 2.2.8 [630] with the NNPDF30 [631] NNLO parton distribution function. They were separated into several orthogonal datasets based on the p_{T} of

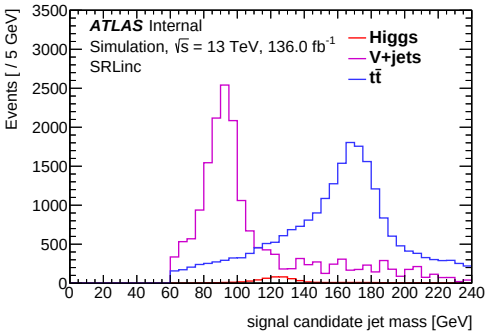
the vector boson. EW NLO corrections are available within the sample as additional weights.

The simulated $t\bar{t}$ samples were generated at tree-level using POWHEG-BOX V2 [627] with $hdamp = 1.5m_t$ and the NNPDF30 [631] NLO parton distribution function. The hadronization is performed using PYTHIA 8.230 [624] with the A14 tune and the NNPDF23 LO [629] parton distribution function. The decay of b-hadrons is performed using EvtGen. The events are split into two samples, one where both tops decay hadronically and one when one or both of the top decays leptonically.

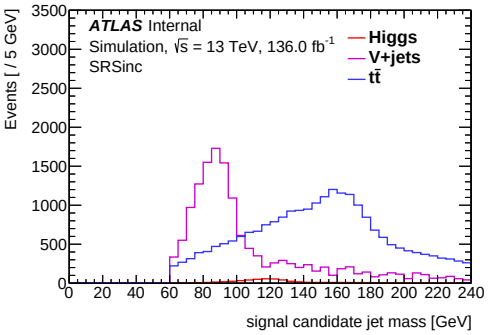
A sample containing events with single (anti-)top quark and a $W^{-(+)}$ was generated at tree-level using POWHEG-BOX V2 [632] with the NNPDF30 parton distribution function. The process is showered using PYTHIA 8.230 [624] configured for the NNPDF23 PDF set [629], the A14 tune, and b-hadrons decays using EVTGEN [626].

C.3 Comparison of Higgs Signal to Minor Backgrounds

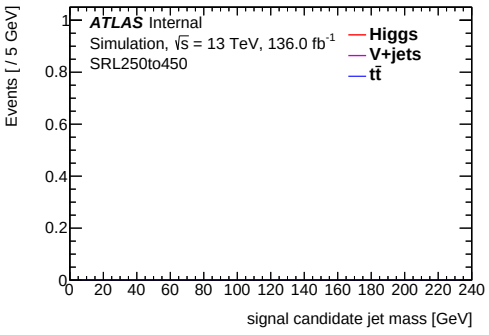
Figure C.2 compares the yield of the backgrounds modeled by MC at the SM prediction. When comparing the shape of the mass spectrum for signal region events in the SRL and SRS categories, a broader mass distribution is observed for the SRS, with an accompanying lower peak value. In the case of W and $t\bar{t}$ events, the resonance shape flattens considerably. In the analysis p_{T} bins, the relative rate of QCD events in the Higgs mass window decreases with increasing p_{T} due to the rapidly falling QCD spectrum as a function of the transverse momentum.



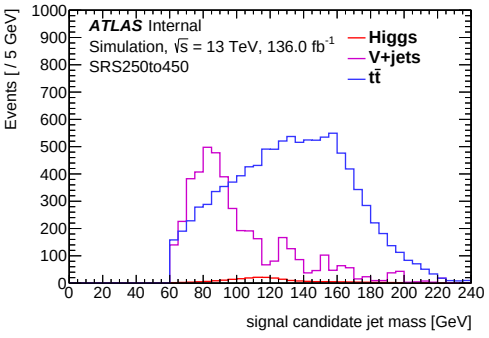
(a) Inclusive SRL



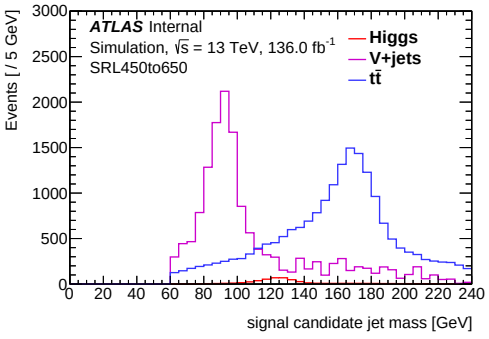
(b) Inclusive SRS



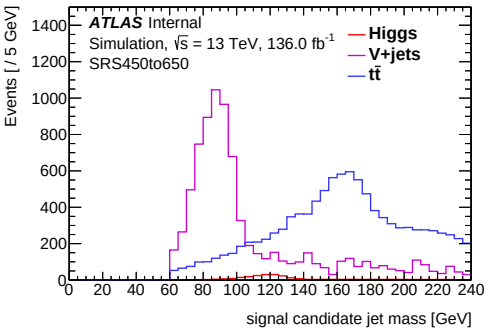
(c) SRL: $250 < p_{\mathrm{T}} < 450\,\mathrm{GeV}$



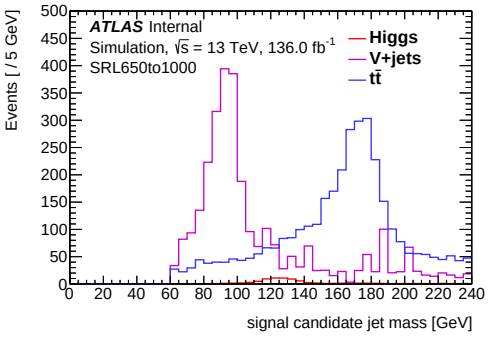
(d) SRS: $250 < p_{\mathrm{T}} < 450\,\mathrm{GeV}$



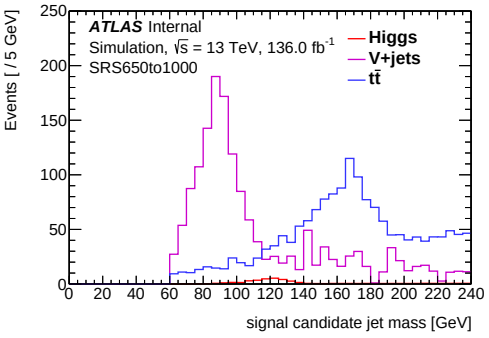
(e) SRL: $450 < p_{\mathrm{T}} < 650\,\mathrm{GeV}$



(f) SRS: $450 < p_{\mathrm{T}} < 650\,\mathrm{GeV}$



(g) SRL: $650 < p_{\mathrm{T}} < 1000\,\mathrm{GeV}$



(h) SRS: $650 < p_{\mathrm{T}} < 1000\,\mathrm{GeV}$

Figure C.2: Comparison of the nominal MC estimates for the $V+\mathrm{jets}$ and $t\bar{t}$ backgrounds compared to the signal in the inclusive (top) and p_{T}-binned SRL (left) and SRS (right) regions.

C.4 $V+$jets Background

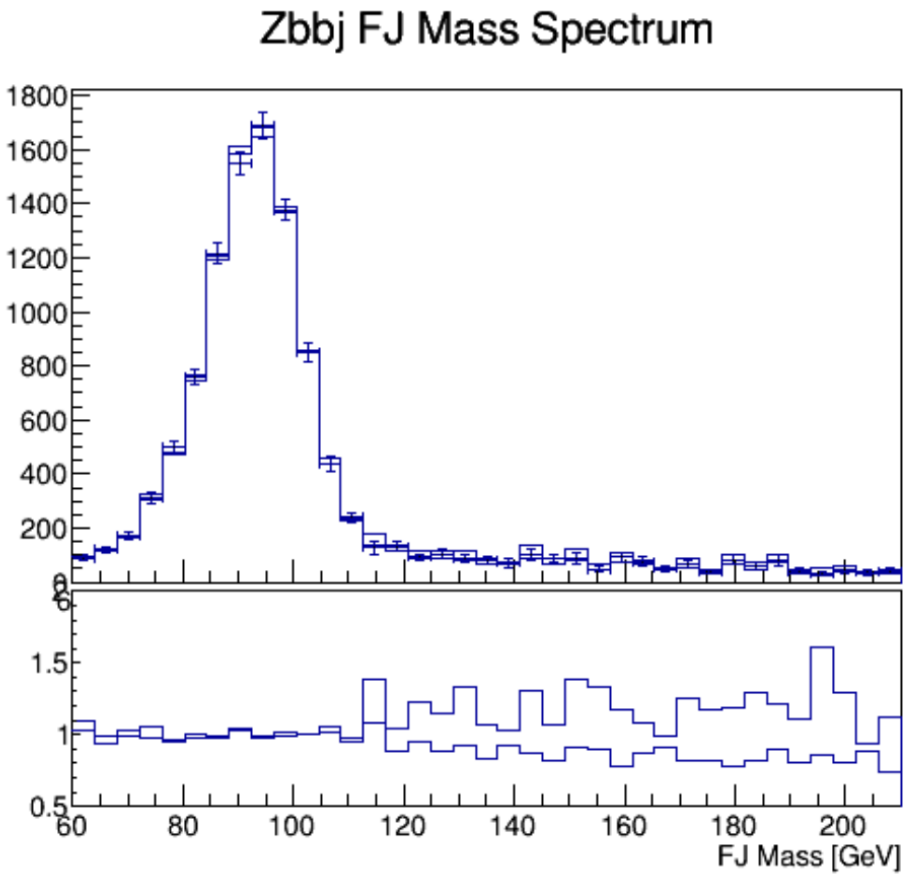


Figure C.3: The large-R jet invariant mass distribution for Z signal events generated with SHERPA 2.2.8 for the $b\bar{b}$ channel. The baseline weights and those corresponding to the up- and down-variations of the renormalization and factorization scale parameters.

Table C.1: The symmetrized fractional changes in the area around the Higgs mass for the PDF and scale variations of the $V+$jets backgrounds. Since the $V+$jets yield is left free in the final fit, the numbers in the table show the change in the SRL after overall normalization changes have been removed.

Source	Range	Symmetrized Fractional Change of Area $105 < m < 140$ GeV
Scale	$0.5 < \mu < 2$	± 0.045
PDF α_s	0.118 ± 0.001	± 0.003
PDF set	`NNPDF-MMHT2014`	± 0.001

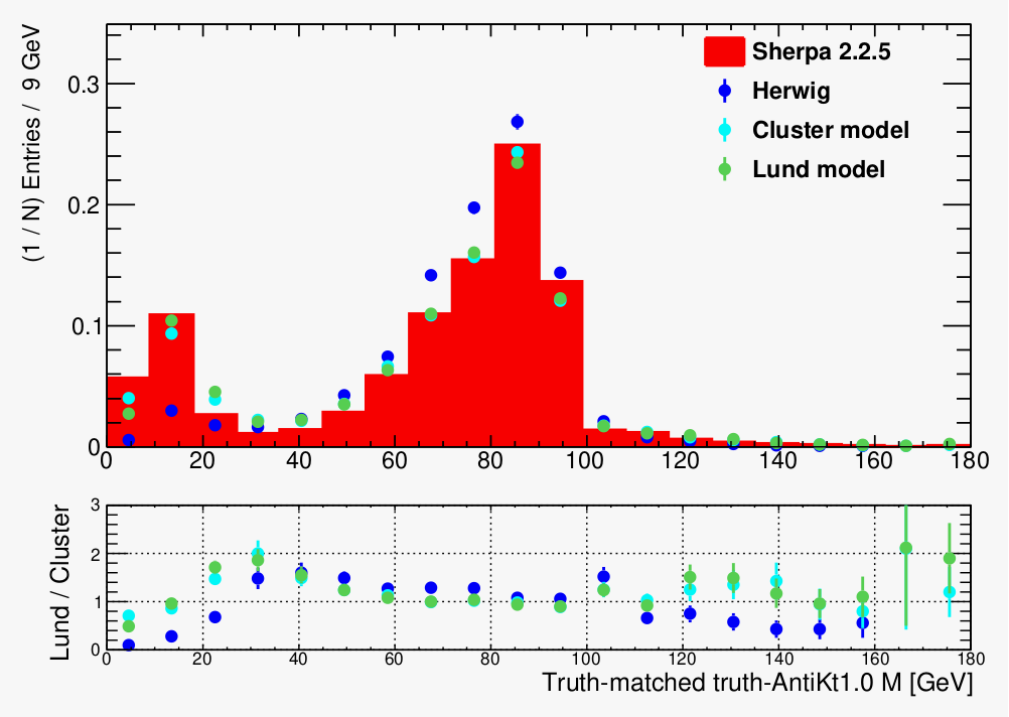


Figure C.4: Comparison of (LO) Sherpa 2.2.5 (red) to Herwig (blue), Sherpa 2.2.8 using the Cluster model (light blue), and Sherpa 2.2.8 using the Lund string model (green) is shown for the mass of the truth-jet matched to the Z boson. The comparison is performed using truth jets (Truth3 derivation) and requiring $300 < p_{T_Z}^{\text{truth}} < 500\,\text{GeV}$.

C.5 $t\bar{t}$ Background

The following two systematics are evaluated by comparing the nominal sample to an alternate sample:

- *Parton Shower* model replacing Pythia 8 with Herwig
- *Matrix Element* where MadGraph5_aMC@NLO replaces Powheg, both interfaced to Pythia 8

Both are shown in Figure C.5.

- *Renormalization and Factorization scales* (μ_R and μ_F)
- *Initial State Radiation* rate

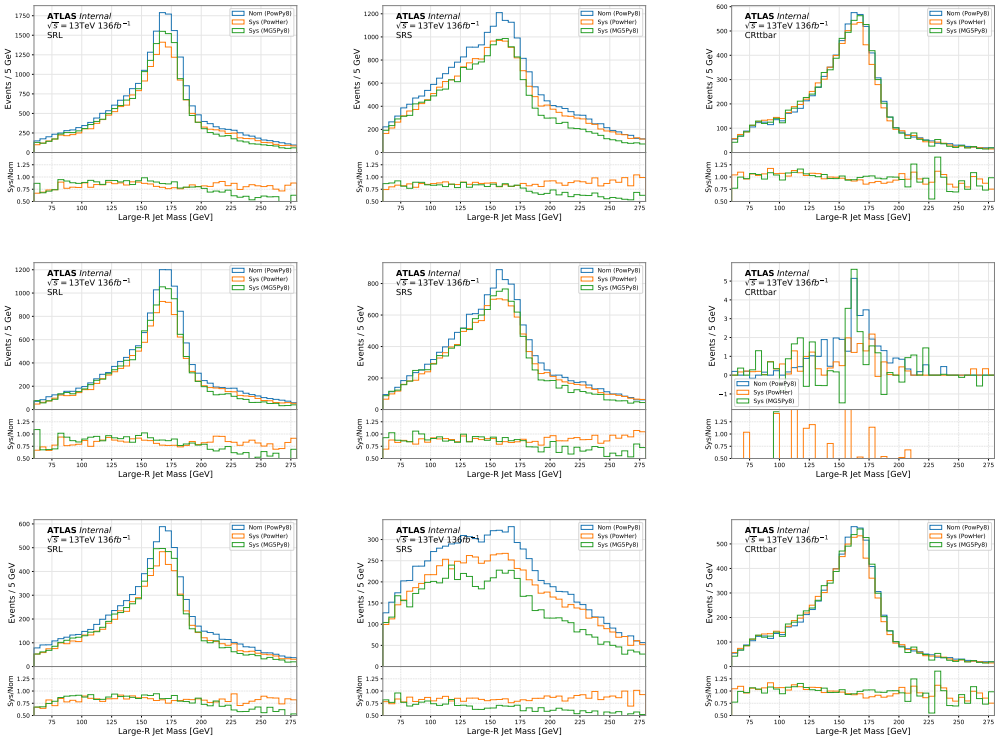


Figure C.5: Comparisons for the $t\bar{t}$ systematics derived from alternate sample comparisons. From left to right, the SRL, SRS and CR$_{t\bar{t}}$ are shown. The rows show the inclusive, all-had and non-all-had $t\bar{t}$ distributions starting from top.

- *Final State Radiation* rate

The 7-point variation of μ_R and μ_F is shown in Figure C.6 and the envelope created by the largest variation is used as the systematic uncertainty. The impact of the weight change for initial and final state radiation is shown in Figure C.7.

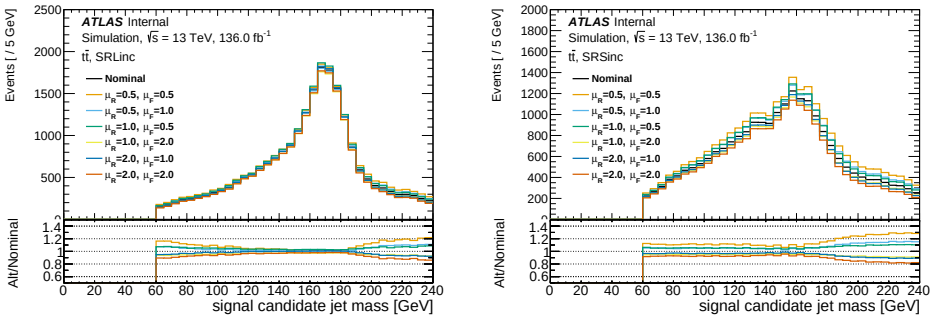


Figure C.6: Effects of the $t\bar{t}$ systematics derived from varying internal weights in the nominal sample to vary the factorization and renormalization scales μ_F and μ_R in a 7-point scheme.

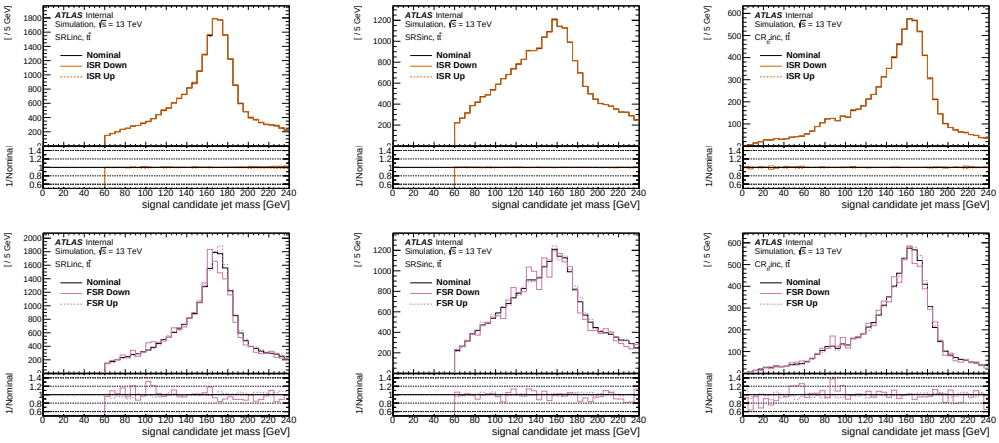


Figure C.7: Effects of the $t\bar{t}$ systematics derived from varying internal weights in the nominal sample. The leading SR, subleading SR and $\mathrm{CR}_{t\bar{t}}$ are shown from left to right. The top row shows the ISR variations and FSR variations is shown on the bottom.

C.6 Efficiencies and Yields

Table C.2: The MC prediction for the expected number of events and efficiencies surviving each analysis cut for the ggF signal process.

Requirement	MC Events	Rel. Eff. (%)	Abs. Eff. (%)
Trigger	3236	100.00	100.00
≥ 1 large-R Jet, $p_{\mathrm{T}} > 450\,\mathrm{GeV}$, $m > 60\,\mathrm{GeV}$	1018	31.48	31.48
Jet Noise Cleaning	1018	100.00	31.48
≥ 2 large-R Jets, $p_{\mathrm{T}} > 200\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.0$	954	93.70	29.50
≥ 1 Candidate Jet	895	93.78	27.66
Signal Region Leading (SRL)	226	25.33	7.01
Signal Region Subleading (SRS)	139	15.64	4.33
Validation Region Leading (VRL)	145	16.29	4.51
Validation Region Subleading (VRS)	87	9.81	2.71

Table C.3: The MC prediction for the expected number of events and efficiencies surviving each analysis cut for the VBF signal process.

Requirement	MC Events	Rel. Eff. (%)	Abs. Eff. (%)
Trigger	1050	100.00	100.00
≥ 1 large-R Jet, $p_{\mathrm{T}} > 450\,\mathrm{GeV}$, $m > 60\,\mathrm{GeV}$	311	29.64	29.64
Jet Noise Cleaning	311	100.00	29.64
≥ 2 large-R Jets, $p_{\mathrm{T}} > 200\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.0$	268	86.07	25.51
≥ 1 Candidate Jet	250	93.48	23.85
Signal Region Leading (SRL)	65	26.07	6.22
Signal Region Subleading (SRS)	47	19.11	4.56
Validation Region Leading (VRL)	26	10.56	2.52
Validation Region Subleading (VRS)	16	6.44	1.54

Table C.4: The MC prediction for the expected number of events and efficiencies surviving each analysis cut for the VH signal process.

Requirement	MC Events	Rel. Eff. (%)	Abs. Eff. (%)
Trigger	1492	100.00	100.0
≥ 1 large-R Jet, $p_{\mathrm{T}} > 450\,\mathrm{GeV}$, $m > 60\,\mathrm{GeV}$	623	41.76	41.76
Jet Noise Cleaning	623	100.00	41.76
≥ 2 large-R Jets, $p_{\mathrm{T}} > 200\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.0$	543	87.25	36.44
≥ 1 Candidate Jet	515	94.83	34.55
Signal Region Leading (SRL)	70	13.75	4.75
Signal Region Subleading (SRS)	97	18.83	6.51
Validation Region Leading (VRL)	124	24.09	8.32
Validation Region Subleading (VRS)	70	13.66	4.72

Table C.5: The MC prediction for the expected number of events and efficiencies surviving each analysis cut for the ttH signal process.

Requirement	MC Events	Rel. Eff. (%)	Abs. Eff. (%)
Trigger	4700	100.00	100.00
≥ 1 large-R Jet, $p_{\mathrm{T}} > 450\,\mathrm{GeV}$, $m > 60\,\mathrm{GeV}$	2102	44.73	44.73
Jet Noise Cleaning	2102	100.00	44.73
≥ 2 large-R Jets, $p_{\mathrm{T}} > 200\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.0$	1912	90.97	40.69
≥ 1 Candidate Jet	1791	93.65	38.11
Signal Region Leading (SRL)	150	8.38	3.19
Signal Region Subleading (SRS)	163	9.12	3.48
Validation Region Leading (VRL)	426	23.83	9.08
Validation Region Subleading (VRS)	270	15.11	5.76

Table C.6: The MC prediction for the expected number of events and efficiencies surviving each analysis cut for the $W+$jets signal process.

Requirement	MC Events	Rel. Eff. (%)	Abs. Eff. (%)
Trigger	2163284	100.00	100.00
≥ 1 large-R Jet, $p_{\mathrm{T}} > 450\,\mathrm{GeV}$, $m > 60\,\mathrm{GeV}$	852831	39.42	39.42
Jet Noise Cleaning	852831	100.00	39.42
≥ 2 large-R Jets, $p_{\mathrm{T}} > 200\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.0$	805643	94.47	37.24
≥ 1 Candidate Jet	753016	93.47	34.81
Signal Region Leading (SRL)	3652	0.49	0.17
Signal Region Subleading (SRS)	3207	0.43	0.15
Validation Region Leading (VRL)	470047	62.42	21.73
Validation Region Subleading (VRS)	335611	44.57	15.51

Table C.7: The MC prediction for the expected number of events and efficiencies surviving each analysis cut for the $t\bar{t}$ signal process.

Requirement	MC Events	Rel. Eff. (%)	Abs. Eff. (%)
Trigger	2205976	100.00	100.00
≥ 1 large-R Jet, $p_{\mathrm{T}} > 450\,\mathrm{GeV}$, $m > 60\,\mathrm{GeV}$	821649	37.25	37.25
Jet Noise Cleaning	821649	100.00	37.25
≥ 2 large-R Jets, $p_{\mathrm{T}} > 200\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.0$	725747	88.33	32.90
≥ 1 Candidate Jet	683640	94.20	30.99
Signal Region Leading (SRL)	24678	3.61	1.12
Signal Region Subleading (SRS)	25181	3.68	1.14
Validation Region Leading (VRL)	196076	28.68	8.89
Validation Region Subleading (VRS)	123586	18.08	5.60

Table C.8: The MC prediction for the expected number of events and efficiencies surviving each analysis cut for the $Z+$jets signal process.

Requirement	MC Events	Rel. Eff. (%)	Abs. Eff. (%)
Trigger	890537	100.00	100.00
≥ 1 large-R Jet, $p_{\mathrm{T}} > 450\,\mathrm{GeV}$, $m > 60\,\mathrm{GeV}$	342033	38.41	38.41
Jet Noise Cleaning	342033	100.00	38.41
≥ 2 large-R Jets, $p_{\mathrm{T}} > 200\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.0$	324088	94.75	36.39
≥ 1 Candidate Jet	302266	93.27	33.94
Signal Region Leading (SRL)	12436	4.11	1.40
Signal Region Subleading (SRS)	10514	3.48	1.18
Validation Region Leading (VRL)	164206	54.32	18.44
Validation Region Subleading (VRS)	115917	38.35	13.02

Table C.9: The MC prediction for the expected number of events and efficiencies surviving each analysis cut for the QCD signal process.

Requirement	MC Events	Rel. Eff. (%)	Abs. Eff. (%)
Trigger	569303040	100.00	100.00
≥ 1 large-R Jet, $p_{\mathrm{T}} > 450\,\mathrm{GeV}$, $m > 60\,\mathrm{GeV}$	116397064	20.45	20.45
Jet Noise Cleaning	116397064	100.00	20.45
≥ 2 large-R Jets, $p_{\mathrm{T}} > 200\,\mathrm{GeV}$, $\lvert\eta\rvert < 2.0$	107753720	92.57	18.93
≥ 1 Candidate Jet	96251352	89.33	16.91
Signal Region Leading (SRL)	812767	0.84	0.14
Signal Region Subleading (SRS)	635522	0.66	0.11
Validation Region Leading (VRL)	59744280	62.07	10.49
Validation Region Subleading (VRS)	40471288	42.05	7.11

Table C.10: Efficiencies and yields for the inclusive (all jet p_{T} bins) SRs and VRs. Efficiencies are relative to events passing the trigger.

	Inclusive SRL		**Inclusive SRS**		**Inclusive VRL**		**Inclusive VRS**	
Process	Yield	Eff. [%]	Yield	Eff. [%]	Yield	Eff. [%]	Yield	Eff. [%]
$Z+$jets	12436	1.40	10514	1.18	164206	18.44	115917	13.02
$W+$jets	3652	0.17	3207	0.15	470047	21.73	335611	15.51
$t\bar{t}$	24678	1.12	25181	1.14	196076	8.89	123586	5.60
ggF ($H \to b\bar{b}$)	226	7.01	139	4.33	145	4.51	87	2.71
ttH ($H \to b\bar{b}$)	150	3.19	163	3.48	426	9.08	270	5.76
VBF ($H \to b\bar{b}$)	65	6.22	47	4.56	26	2.52	16	1.54
VH ($H \to b\bar{b}$)	70	4.75	97	6.51	124	8.32	70	4.72
Total Higgs	5113	4.90	448	4.28	723	6.90	444	4.25
Data	605809	0.17	521212	0.15	47888436	13.73	349265520	10.02

Table C.11: Yields(Efficiencies) for the signal and validation regions in $p_{\mathrm{T}}^{\mathrm{Jet}}$ bins with respect to the inclusive SR and VR requirements.

	Jet Transverse Momentum Range [GeV]					
	250–450	450–650	650–1000	250–450	450–650	650–1000
Process		SRL			SRS	
QCD		671811 (0.83)	127765 (0.16)	187049 (0.29)	369861 (0.58)	70846 (0.11)
$Z+$jets		10152 (0.82)	2083 (0.17)	2868 (0.27)	6260 (0.60)	1257 (0.12)
$W+$jets		2620 (0.70)	916 (0.25)	1008 (0.30)	1591 (0.47)	606 (0.18)
$t\bar{t}$		19677 (0.80)	4511 (0.18)	10492 (0.42)	11517 (0.46)	2823 (0.11)
ggF ($H \to b\bar{b}$)		196 (0.87)	29 (0.13)	47 (0.34)	81 (0.58)	11 (0.08)
VBF ($H \to b\bar{b}$)		56 (0.86)	9 (0.14)	10 (0.21)	32 (0.68)	5 (0.11)
VH ($H \to b\bar{b}$)		59 (0.83)	11 (0.15)	43 (0.45)	44 (0.46)	8 (0.09)
ttH ($H \to b\bar{b}$)		117 (0.78)	30 (0.20)	82 (0.50)	62 (0.38)	17 (0.10)
Total Higgs		428 (0.83)	79 (0.15)	183 (0.41)	220 (0.49)	41 (0.09)
Data		501282 (0.83)	95393 (0.16)	144438 (0.28)	310564 (0.60)	60398 (0.12)
		VRL			VRS	
QCD		49995589 (0.84)	8988738 (0.15)	119635 (0.45)	115526 (0.43)	28772 (0.11)
$Z+$jets		126884 (0.77)	33065 (0.20)	30265 (0.26)	66340 (0.57)	16942 (0.15)
$W+$jets		363980 (0.77)	93908 (0.20)	88702 (0.27)	189798 (0.57)	48953 (0.15)
$t\bar{t}$		155553 (0.79)	36364 (0.19)	54402 (0.44)	55047 (0.45)	12692 (0.10)
ggF ($H \to b\bar{b}$)		119 (0.82)	24 (0.17)	30 (0.35)	48 (0.55)	8 (0.10)
VBF ($H \to b\bar{b}$)		21 (0.80)	5 (0.18)	4 (0.23)	10 (0.62)	2 (0.13)
VH ($H \to b\bar{b}$)		96 (0.78)	25 (0.20)	23 (0.33)	36 (0.52)	10 (0.14)
ttH ($H \to b\bar{b}$)		321 (0.75)	93 (0.22)	129 (0.48)	108 (0.40)	29 (0.11)
Total Higgs		558 (0.77)	146 (0.20)	187 (0.42)	203 (0.46)	49 (0.11)
Data		40104882 (0.84)	7177255 (0.15)	8761920 (0.25)	22017384 (0.63)	3822307 (0.11)

Table C.12: Yields for the signal and validation regions in $p_{\mathrm{T}}^{\mathrm{Jet}}$ bins and restricting to the candidate jet mass range of 105–140 GeV. The last row gives the $s/\sqrt{b}$ with s being the combined Higgs signal and b the QCD background.

	Jet Transverse Momentum Range [GeV]					
	250–450	450–650	650–1000	250–450	450–650	650–1000
Process		SRL			SRS	
QCD		147402	23005	46180	83715	12976
$Z+$jets		1236	235	344	697	129
$W+$jets		441	103	259	278	64
$t\bar{t}$		3326	501	3216	1999	263
ggF ($H \to b\bar{b}$)		160	25	23	61	10
VBF ($H \to b\bar{b}$)		48	8	5	24	5
VH ($H \to b\bar{b}$)		40	7	21	31	6
ttH ($H \to b\bar{b}$)		37	8	32	15	2
Total Higgs		285	47	81	132	24
Higgs $s/\sqrt{b}$		0.79	0.33	0.45	0.48	0.22
		VRL			VRS	
QCD		10864069	1572090	33183	23850	4626
$Z+$jets		21238	4511	6952	11144	1992
$W+$jets		53038	10700	20112	30736	5811
$t\bar{t}$		25718	4548	14687	9281	1406
ggF ($H \to b\bar{b}$)		39	7	9	15	3
VBF ($H \to b\bar{b}$)		9	2	1	5	1
VH ($H \to b\bar{b}$)		15	4	6	8	2
ttH ($H \to b\bar{b}$)		61	13	38	19	3
Total Higgs		123	26	54	47	8
Data		8800009	1254567	2348784	4830267	624080
Higgs $s/\sqrt{b}$		0.05	0.03	0.04	0.02	0.01

C.7 WCR$_{t\bar{t}}$: Powheg + Pythia 8 vs. Sherpa

Extensive studies comparing the Sherpa 2.2.1 and Powheg + Pythia 8 generators were performed in the analysis p_{T} bins, as well as in several pileup bins to ensure major differences were not seen, which would necessitate the need for an additional systematic for differences in the generator used. As can be seen in Figure C.8 for the inclusive distributions, and Figure C.9 for the p_{T}-binned distributions, Powheg + Pythia 8 benefits from higher statistics, but the Sherpa 2.2.1 samples more reliably match the resolution found in data. Comparisons were also made in three pileup bins: $[0, 30]$, $[30, 40]$, and $[40, 100]$, as seen in Figure C.10. No significant differences, nor trends as a function of the PU are seen.

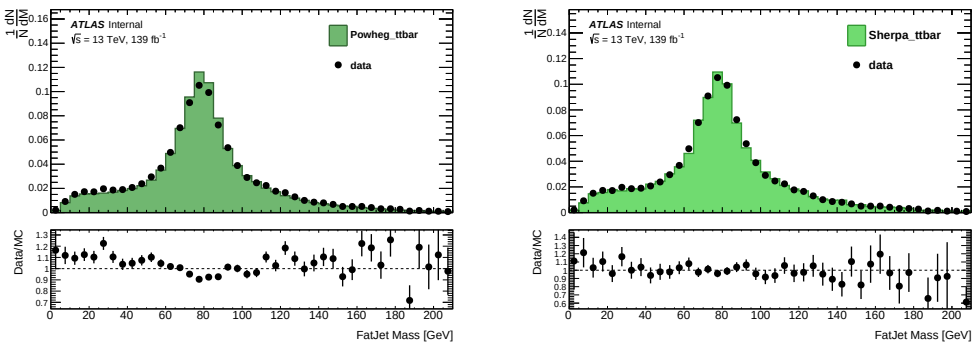


Figure C.8: Comparisons of the Powheg + Pythia 8 MC generator (left) with Sherpa 2.2.1 (right) against data.

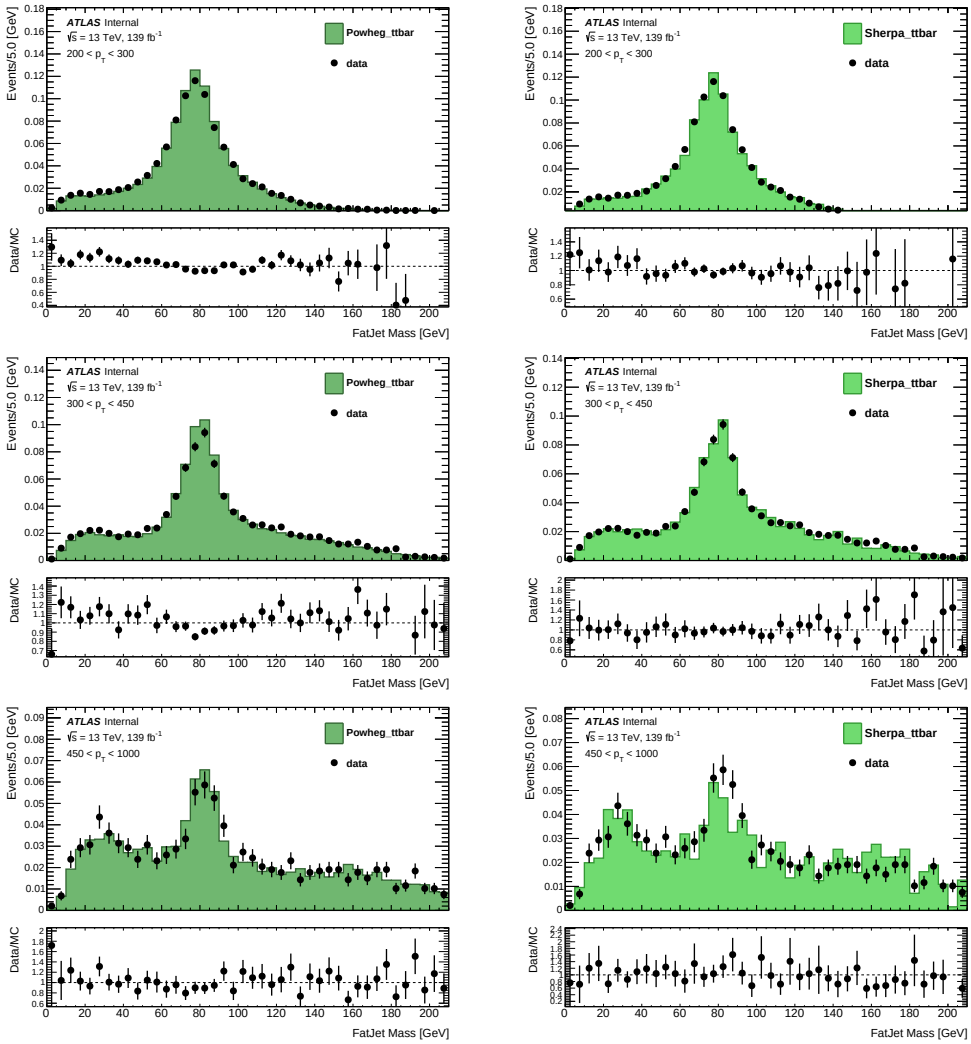


Figure C.9: Comparisons of the POWHEG + PYTHIA 8 MC generator (left) with SHERPA 2.2.1 (right) against data in the analysis p_{T} bins.

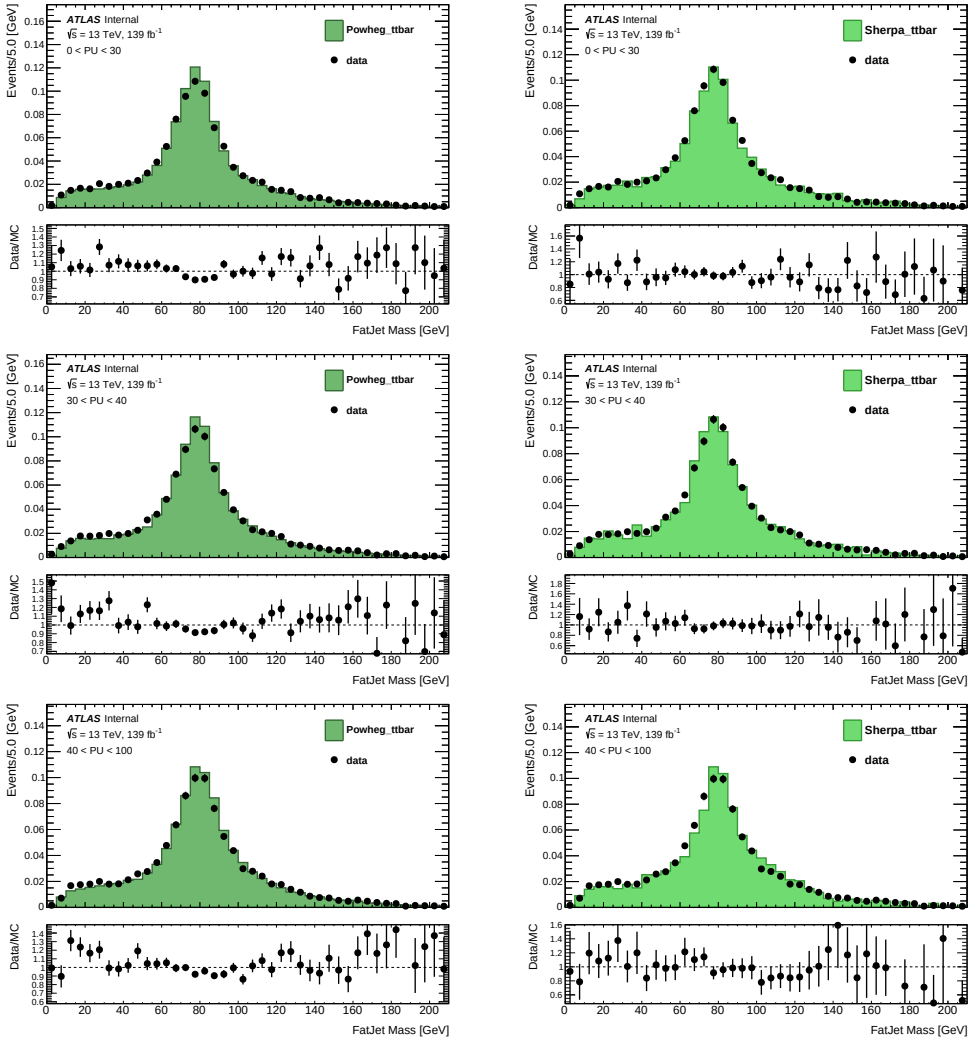


Figure C.10: Comparisons of the Powheg + Pythia 8 MC generator (left) with Sherpa 2.2.1 (right) against data, binned by pileup.

C.8 Jet Mass Resolution Nuisance Parameter Post-Fit Mean and Uncertainty

Table C.13: A fit is performed on the inclusive SRs using the JMR external constraint. Using the formula for the mean and width of two convoluted Gaussian distributions, the post-fit value of the JMR NP ("MassRes_WZ_comb") is used to solve for the NP posterior of the SR alone. For the inclusive result, an older constraint was used ($\mu = 0.056, \sigma = 0.11$). The constraints used for the differential analysis are given in Table C.14. For clarity, the equations are: $\sigma_{\text{SR}}^{-2} = \sigma_{\text{post-fit}}^{-2} - \sigma_{\text{constraint}}^{-2}$ and $\mu_{\text{SR}}/\sigma_{\text{SR}}^2 = \mu_{\text{post-fit}}/\sigma_{\text{post-fit}}^2 - \mu_{\text{external}}/\sigma_{\text{external}}^2$ where "post-fit" is the value obtained from the post-fit plot after the fit is complete.

Region	$\mu_{\text{post-fit}}$	$\sigma_{\text{post-fit}}$	μ_{SR}	σ_{SR}
Inclusive SR	0.081	0.106	0.41	0.40
SR0	0.128	0.149	0.17	1.08
SR1	0.113	0.145	0.70	0.56
SR2	0.002	0.204	0.29	0.88

Table C.14: A summary of the JMR Gaussian external constraints acting on `MassRes_WZ_comb` in the SR likelihood given for each region. The two methods discussed in Section 9.5.4 are compared and the final value using the p_{T}-folded method plus the transfer systematic is also shown. *Fit Value* refers to the value and error of the fit shown in Figure 9.24 at the average jet p_{T} of the corresponding region. *p_T-Folded* refers to the value and error of the fit shown in Figure 9.24 averaged over the p_{T} distribution of the corresponding region. The last two columns give the values used to build the Gaussian external constraint with the mean (μ_{ext}) coming directly from the p_{T}-folded method and the width (σ_{ext}) coming from p_{T}-folded method added in quadrature with the transfer systematic of 0.1.

Region	Average p_{T} [GeV]	Fit value	p_{T}-Folded	μ_{ext}	σ_{ext}
Inclusive SRL	562	0.051 ± 0.108	0.056 ± 0.122	0.056	0.158
SRL1	520	0.070 ± 0.104	0.071 ± 0.107	0.071	0.146
SRL2	775	-0.014 ± 0.185	-0.018 ± 0.198	-0.018	0.222
Inclusive SRS	516	0.070 ± 0.104	0.074 ± 0.116	0.074	0.153
SRS0	381	0.127 ± 0.105	0.127 ± 0.100	0.127	0.141
SRS1	520	0.070 ± 0.104	0.071 ± 0.107	0.071	0.146
SRS2	775	-0.014 ± 0.185	-0.018 ± 0.198	-0.018	0.212

C.9 Systematics

Table C.15: A summary of the modeling systematics. N means the parameter has a normalization effect within a given region, while S means the parameter can change the mass shape in a given region. Exceedingly small variations have not been included and are summarized in the relevant section.

Nusiance Parameter	Description	Category	Value	Effect d
	$V+$jets			
`Wboson_unc`	$W+$jets Normalization	all	10%	N
`wMUR0p5_MUF0p5`	ME+PS Scale Variations	all	3–20%	N + S
	$t\bar{t}$			
`ttbar_PartonShower`	Fragmentation/Hadronization from POWHEG + PYTHIA/HERWIG Comparisons	all	6–20%	N + S
`ttbar_MatrixElement`	POWHEG vs. aMC@NLO	all	1–19%	N + S
`wisr_muRfac1p0_ fsr_muRfac2p0`	ISR/FSR Variation	all	1–7%	N + S
	Higgs			
`Higgs_acc`	ggF Acceptance	all	20%	N
	VBF Acceptance	all	0.5%	N
	VH Acceptance	all	5%	N
	ttH Acceptance	all	13%	N
`Higgs_EW`	NLO EW corrections (VBF, VH, ttH)	all	2–6%	N

Table C.16: A summary of the 16 FTAG systematics considered. N means the parameter has a normalization effect within a given region while S means the parameter can change the mass shape in a given region. All names start with the prefix `FT_EFF_`.

Nuisance Parameter	Description	Category	Effect
`Eigen_B_[0-4]`	5 variations of b-jet scale factors	true b-jets	N + S
`extrapolation`	high p_{T} ($p_{\mathrm{T}} > 400\,\mathrm{GeV}$) b-jet scale factor uncertainty	all	N + S
`Eigen_C_[0-3]`	4 variations of c-jet scale factors	true c-jets	N + S
`Eigen_Light_[0-4]`	5 variations of light-jet scale factors	true light jets	N + S
`extrap_from_charm`	τ-jet scale factor uncertainty	true τ-jets	N + S

Table C.17: A summary of the 23 Jet Energy Scale (JES) systematics. In the process column, *all* means it applies to all simulated processes (not the QCD background). *N* means the parameter has a normalization effect within a given region while *S* means the parameter can change the mass shape in a given region. The systematics are presented in categories where the full name is found by replacing *TYPE* with the name found in the first column.

Nuisance Parameter	Description	Category	Process	Effect
`JET_EffectiveNP_R10_TYPEConstraint`				
`Detector[1-2]`	Experimental Errors	all	all	N + S
`Mixed[1-4]`	Mix	all	all	N + S
`Modelling[1-4]`	Usually Generator Comparisons	all	all	N + S
`Statistical[1-6]`	Statistical Limitations	all	all	N + S
`JET_EtaIntercalibration_TYPEConstraint`				
`Modelling`	Usually Generator Comparisons	all	all	N + S
`NonClosure_2018data`	Unknown Source	all	all	N + S
`R10_TotalStat`	Statistics	all	all	N + S
`JET_Flavor_TYPEConstraint`*				
`Composition`	Quark-Gluon Fraction	all	all	N + S
`Response`	Calorimeter Response for Quark vs. Gluon	all	all	N + S
`JET_LargeR_TopologyUncertainty_TYPEConstraint`†				
`top`	Top-jet Topology	all	$t\bar{t}$	N + S
`V`	V-jet Topology	all	V + jets	N + S

* Only applied to jets that are truth-labeled as QCD-like jet.

† The topology uncertainty variations are only applied to jets matched to the corresponding truth particle. The Higgs variation is currently being developed by the JetEtMiss group and therefore not available.

Table C.18: A summary of the 13 Jet Mass Scale (JMS) systematics derived *in situ* using the R-track procedure. In the process column, *all* means it applies to all simulated processes (not the QCD background). N means the parameter has a normalization effect within a given region while S means the parameter can change the mass shape in a given region. The systematics are presented in categories where the full name is found by replacing `TYPE` with the name found in the first column. In the row Category, *all* means a full correlation between analysis p_{T} bins for the differential analysis while *p_T-bins* means decorrelated between the three analysis p_{T} bins in the differential analysis.

Nuisance Parameter	Description	Category	Process	Effect
		`JET_CombMass_TYPEConstraint`		
`Baseline`	Difference Between Data and Pythia 8	all	H,V + jets	N + S
`Baseline_ttbar`		all	$t\bar{t}$	N + S
`Modelling`	Maximum Difference Between Pythia 8 and Herwig 7 or Pythia 8 and Sherpa	p_{T}-bins	V + jets	N + S
`Modelling_ttbar`		p_{T}-bins	$t\bar{t}$	N + S
`Modelling_H`		p_{T}-bins	H	N + S
`TotalStat`	Statistical Uncertainty on the Measurement	all	H,V + jets	N + S
`TotalStat_ttbar`		all	$t\bar{t}$	N + S
`Tracking[1-3]`	Tracking Eff., Fake Rate, q/p_{T} bias	all	H,V + jets	N + S
`Tracking[1-3]_ttbar`		all	$t\bar{t}$	N + S

† Since the flavor composition and radiation in a hadronic top jet is not the same as a $H \to b\bar{b}$ or $Z \to b\bar{b}$ jet, the `Baseline` and `Modelling` systematics are not correlated between the top background and other processes. This does not follow a JetETMiss recommendation but arose during the Hbb group review.

Table C.19: A summary of the 9 Jet Mass Resolution (JMR) systematics. N means the parameter has a normalization effect within a given region while S means the parameter can change the mass shape in a given region. The systematics are presented in categories where the full name is found by replacing `TYPE` with the name found in the first column. The decorrelation of the Higgs, V and Top jets is the default JetETMiss recommendation. See Section 9.5 for discussion of external constraint added to $V+$jets systematic.

Nuisance Parameter	Description	Category	Process	Effect
		`JET_MassRes_TYPE_comb`		
`Hbb`	"From differences seen in cross calib. samples, but is a very rough estimate!" [468]	p_{T} bins	$H \to b\bar{b}$	S
`Top`		p_{T} bins	$t\bar{t}$	S
`WZ`		p_{T} bins	$V+$jets	S

C.10 Fitting Framework

The fitting framework is based on the XML Analytic Workspace Builder (xmlAnaWSBuilder) [633], which is widely used in Higgs analyses. XmlAnaWSBuilder creates RooFit workspaces using one-dimensional observables and the workflow of the framework is summarized in Figure C.11.

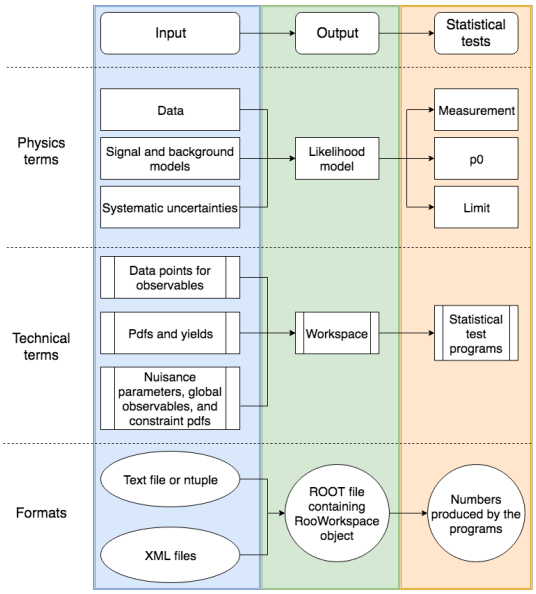


Figure C.11: The XmlAnaWSBuilder workflow. For each of the input (blue), output (green), and statistical test (orange) steps, there is a description in terms of the physics (involving the data, MC, and uncertainties), the technical implementation (in terms of the template pdfs, signal strengths, and other parameters), and the file formats.

C.11 Mass Effects from b-tagging

The mass distributions for data and background simulation in the validation and signal regions are shown in Figure C.12.

The mass turn-on effect is evident in the signal region while the validation region shows a monotonic decrease of the mass spectrum. It is evident that the data reach a plateau followed by a subsequent point of monotonic decrease at higher mass values compared to simulation, in particular for the high-p_{T} bins. In order to study the effect of the b-tagging response on the mass spectrum, the simulation is weighted by the correction function obtained from the study with boosted $t\bar{t}$ events (see Figure C.13). Details of the study are given in Appendix A of Reference [403].

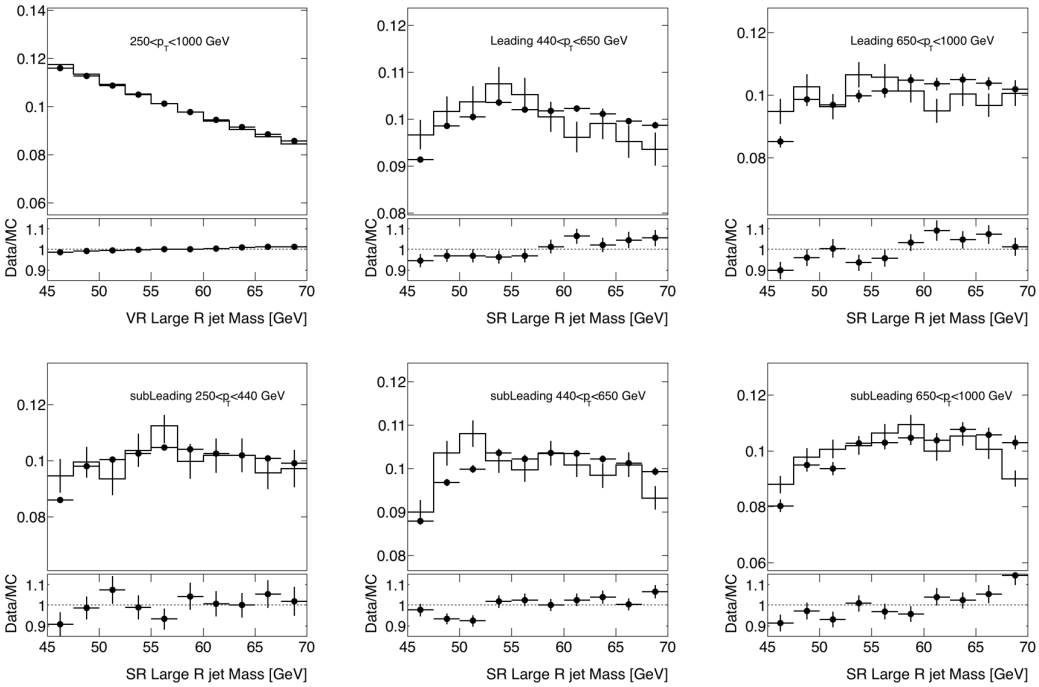


Figure C.12: Normalized distributions of the candidate large-R jet mass for the leading (upper panel) and subleading (lower panel) categories in data (points with error bars) and simulation (histogram). The upper left panel shows the validation region and the other panels the signal region in the three p_{T} intervals. The "turn-on" is attributed to a decrease in b-tagging efficiency as the angular separation between b-hadrons decreases.

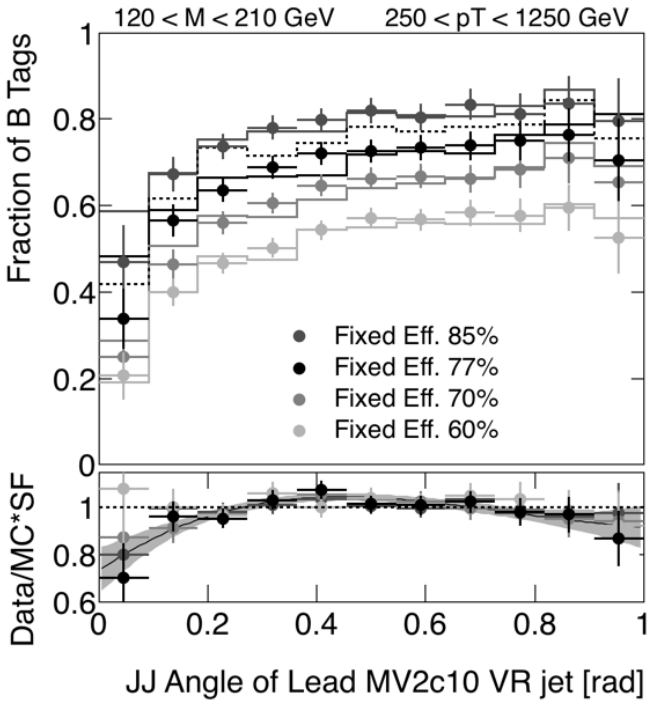


Figure C.13: Rate of b tags for different working points as a function of the angle between the b-candidate and the closest selected VR track jets in a top-dominated sample. The simulation predictions are shown by the lines and the data by the points with error bars. The dashed line represents the fraction of jets associated to a genuine b quark among the selected b-candidate jets. The ratio between the rate of b tags in data to simulation is shown in the lower pad. The fitted polynomial function is used for the re-weighting of the mass spectrum for simulated QCD events with the gray band giving the 68% C.L. of the fit. See Appendix A of Reference [403] for details.

Bibliography

[1] ATLAS Collaboration. *ATLAS detector and physics performance: Technical Design Report, 1*. Technical design report. ATLAS. Geneva: CERN, 1999. URL: https://cds.cern.ch/record/391176 (on pages 1, 75, 76, 80).

[2] ATLAS Collaboration. *ATLAS detector and physics performance: Technical Design Report, 2*. Technical design report. ATLAS. Geneva: CERN, 1999. URL: https://cds.cern.ch/record/391177 (on pages 1, 75, 76, 80).

[3] R. Raitio and W. W. Wada. "Higgs Boson Production at Large Transverse Momentum in QCD". In: *Phys. Rev. D* 19 (1979), p. 941 (on page 1).

[4] R. Ellis et al. "Higgs decay to + −A possible signature of intermediate mass Higgs bosons at high energy hadron colliders". In: *Nuclear Physics B* 297.2 (1988), pp. 221–243. ISSN: 0550-3213. URL: https://www.sciencedirect.com/science/article/pii/0550321388900193 (on page 1).

[5] J. M. Butterworth et al. "Jet Substructure as a New Higgs-Search Channel at the Large Hadron Collider". In: *Phys. Rev. Lett.* 100 (24 2008), p. 242001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.100.242001 (on page 1).

[6] S. L. Glashow. "Partial-symmetries of weak interactions". In: *Nuclear Physics* 22.4 (1961), pp. 579 –588. ISSN: 0029-5582. URL: http://www.sciencedirect.com/science/article/pii/0029558261904692 (on pages 1, 6, 17).

[7] S. Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (21 1967), pp. 1264–1266. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264 (on pages 1, 6, 17).

[8] A. Salam. "Weak and Electromagnetic Interactions". In: *Conf. Proc. C* 680519 (1968), pp. 367–377. URL: https://inspirehep.net/literature/53083 (on pages 1, 6, 17).

[9] T. P. Cheng and L. F. Li. *GAUGE THEORY OF ELEMENTARY PARTICLE PHYSICS*. 1984. ISBN: 978-0-19-851961-4. URL: https://inspirehep.net/literature/221340 (on pages 1, 6).

[10] F. Halzen and A. D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. 1984. ISBN: 978-0471887416. URL: https://inspirehep.net/literature/205394 (on pages 1, 6).

[11] M. Ogilvie. “Dynamics of the standard model, 2nd edition, John F. Donoghue, Eugene Golowich and Barry R. Holstein”. In: *Contemporary Physics* 57 (June 2015), pp. 1–3. URL: https://www.researchgate.net/publication/281189950_Dynamics_of_the_standard_model_2nd_edition_John_F_Donoghue_Eugene_Golowich_and_Barry_R_Holstein (on pages 1, 6).

[12] P. Skands. *QCD for Collider Physics*. 2012. arXiv: 1104.2863 [hep-ph] (on pages 1, 6).

[13] T. Muta. *Foundations of Quantum Chromodynamics: An Introduction to Perturbative Methods in Gauge Theories, (3rd ed.)* 3rd. Vol. 78. World scientific Lecture Notes in Physics. Hackensack, N.J.: World Scientific, 2010. ISBN: 978-981-279-353-9 (on pages 1, 6).

[14] F. J. Yndurain. *The Theory of Quark and Gluon Interactions*. Theoretical and Mathematical Physics. Berlin, Germany: Springer, 2006. ISBN: 978-3-540-33209-1, 978-3-540-33210-7 (on pages 1, 6).

[15] M. J. Herrero. “Introduction to the symmetry breaking sector”. In: *23rd International Meeting on Fundamental Physics: The Top Quark, Heavy Flavor Physics and Symmetry Breaking*. Jan. 1996. arXiv: hep-ph/9601286 (on pages 1, 6).

[16] B. Schumm. *Deep Down Things: The Breathtaking Beauty of Particle Physics*. 2004. ISBN: 978-0801879715 (on pages 1, 6).

[17] M. Herrero. “The Standard model”. In: *NATO Sci. Ser. C* 534 (1999). Ed. by T. Ferbel, pp. 1–59. arXiv: hep-ph/9812242 (on pages 1, 6).

[18] L. Evans and P. Bryant. “LHC Machine”. In: *Journal of Instrumentation* 3.08 (2008), S08001–S08001. URL: https://doi.org/10.1088/1748-0221/3/08/s08001 (on pages 1, 30, 66).

[19] O. S. Brüning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. URL: https://cds.cern.ch/record/782076 (on pages 1, 30, 66, 72).

[20] CERN. *The European Organization for Nuclear Research is born*. URL: https://timeline.web.cern.ch/european-organization-nuclear-research-born (on pages 1, 66).

[21] ATLAS Collaboration. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Physics Letters B* 716.1 (2012), pp. 1–29. ISSN: 0370-2693. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.020 (on pages 2, 10, 37, 38, 249).

[22] CMS Collaboration. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Physics Letters B* 716.1 (2012), pp. 30–61. ISSN: 0370-2693. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.021 (on pages 2, 10, 37, 38, 249).

[23] ATLAS Collaboration. “Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector”. In: *Journal of High Energy Physics* 2017.12 (2017). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP12(2017)024 (on pages 2, 131, 149).

[24] CMS Collaboration. "Observation of Higgs Boson Decay to Bottom Quarks". In: *Physical Review Letters* 121.12 (Sept. 2018). ISSN: 1079-7114. URL: http://dx.doi.org/10.1103/PhysRevLett.121.121801 (on pages 2, 39, 52).

[25] ATLAS Collaboration. "Measurement of the associated production of a Higgs boson decaying into b-quarks with a vector boson at high transverse momentum in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector". In: *Physics Letters B* 816 (May 2021), p. 136204. ISSN: 0370-2693. URL: http://dx.doi.org/10.1016/j.physletb.2021.136204 (on pages 2, 244).

[26] CMS Collaboration. "Search for the standard model Higgs boson produced through vector boson fusion and decaying to $b\bar{b}$". In: *Physical Review D* 92.3 (2015). ISSN: 1550-2368. URL: http://dx.doi.org/10.1103/PhysRevD.92.032008 (on pages 2, 53, 244).

[27] ATLAS Collaboration. "Search for the Standard Model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector". In: *Phys. Rev. D* 97.7 (2018), p. 072016. arXiv: 1712.08895 [hep-ex] (on pages 2, 53).

[28] CMS Collaboration. "Search for $t\bar{t}H$ production in the $H \to b\bar{b}$ decay channel with leptonic $t\bar{t}$ decays in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *JHEP* 03 (2019), p. 026. arXiv: 1804.03682 [hep-ex] (on page 2).

[29] A. A. Coley and D. L. Wiltshire. "What is general relativity?" In: *Physica Scripta* 92.5 (2017), p. 053001. ISSN: 1402-4896. URL: http://dx.doi.org/10.1088/1402-4896/aa6857 (on page 6).

[30] A. Arbuzov. "Quantum Field Theory and the Electroweak Standard Model". In: (Jan. 2018) (on page 7).

[31] W. Pauli. "Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren". In: *Zeitschrift für Physik* 31.1 (Feb. 1925), pp. 765–783. ISSN: 0044-3328. URL: https://doi.org/10.1007/BF02980631 (on page 7).

[32] M. Goldhaber, L. Grodzins, and A. W. Sunyar. "Helicity of Neutrinos". In: *Phys. Rev.* 109 (3 Feb. 1958), pp. 1015–1017. URL: https://link.aps.org/doi/10.1103/PhysRev.109.1015 (on page 8).

[33] J. W. F. Valle. "Physics beyond the Standard Model". In: (1990), 27 p. URL: https://cds.cern.ch/record/213349 (on page 8).

[34] R. Davis, D. S. Harmer, and K. C. Hoffman. "Search for Neutrinos from the Sun". In: *Phys. Rev. Lett.* 20 (21 May 1968), pp. 1205–1209. URL: https://link.aps.org/doi/10.1103/PhysRevLett.20.1205 (on page 8).

[35] LHCb Collaboration. "Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays". In: *Physical Review Letters* 115.7 (Aug. 2015). ISSN: 1079-7114. URL: http://dx.doi.org/10.1103/PhysRevLett.115.072001 (on page 9).

[36] LHCb collaboration. "Observation of structure in the J/ψ-pair mass spectrum". In: *Science Bulletin* 65.23 (Dec. 2020), pp. 1983–1993. ISSN: 2095-9273. URL: http://dx.doi.org/10.1016/j.scib.2020.08.032 (on page 9).

[37] TOTEM Collaboration. *Total cross-section, elastic scattering and diffraction dissociation at the Large Hadron Collider at CERN: TOTEM Technical Design Report*. Technical design report. TOTEM. Geneva: CERN, 2004. URL: https://cds.cern.ch/record/704349 (on page 9).

[38] TOTEM Collaboration. "The TOTEM Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08007. URL: https://cds.cern.ch/record/1129807 (on page 9).

[39] B. Pifer. *An Experiment at D0 to Study anti-Proton - Proton Collisions at 2-TeV: Design Report*. Tech. rep. 1983. URL: https://cds.cern.ch/record/1478625 (on page 9).

[40] E. Martynov and B. Nicolescu. "Discovery of the Odderon by TOTEM experiments and the FMO approach". In: *EPJ Web Conf.* 206 (Oct. 2018), 06001. 10 p. arXiv: 1810.08930. URL: https://cds.cern.ch/record/2646169 (on page 9).

[41] TOTEM Collaboration. *Comparison of pp and $p\bar{p}$ differential elastic cross sections and observation of the exchange of a colorless C-odd gluonic compound*. 2020. arXiv: 2012.03981 [hep-ex] (on page 9).

[42] UA1 Collaboration. "Experimental observation of isolated large transverse energy electrons with associated missing energy at $s = 540\,\text{GeV}$". In: *Physics Letters B* 122.1 (1983), pp. 103–116. ISSN: 0370-2693. URL: https://www.sciencedirect.com/science/article/pii/0370269383911772 (on pages 10, 249).

[43] UA2 Collaboration. "Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider". In: *Physics Letters B* 122.5 (1983), pp. 476–485. ISSN: 0370-2693. URL: https://www.sciencedirect.com/science/article/pii/0370269383916052 (on pages 10, 249).

[44] UA1 Collaboration. "Experimental observation of lepton pairs of invariant mass around 95 GeV/c2 at the CERN SPS collider". In: *Physics Letters B* 126.5 (1983), pp. 398–410. ISSN: 0370-2693. URL: https://www.sciencedirect.com/science/article/pii/0370269383901880 (on pages 10, 249).

[45] UA2 Collaboration. "Evidence for Z0→e+e− at the CERN pp collider". In: *Physics Letters B* 129.1 (1983), pp. 130–140. ISSN: 0370-2693. URL: https://www.sciencedirect.com/science/article/pii/037026938390744X (on pages 10, 249).

[46] J. Timmer. "The UA1 Detector". In: *3rd Moriond Workshop: Antiproton-Proton Physics*. 1983 (on page 10).

[47] B. Mansoulie. "The UA2 Apparatus at the CERN P Anti-P Collider". In: *3rd Moriond Workshop: Antiproton-Proton Physics*. Mar. 1983 (on page 10).

[48] A. W. Kolb and L. H. Hoddeson. "A New frontier in the Chicago suburbs: Settling Fermi National Accelerator Laboratory, 1963-1972". In: Dec. 1992 (on page 10).

[49] S. W. Herb et al. “Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions”. In: *Phys. Rev. Lett.* 39 (5 1977), pp. 252–255. URL: https://link.aps.org/doi/10.1103/PhysRevLett.39.252 (on pages 10, 249).

[50] V. Abazov et al. “The upgraded DØ detector”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 565.2 (2006), pp. 463–537. ISSN: 0168-9002. URL: http://dx.doi.org/10.1016/j.nima.2006.05.248 (on page 10).

[51] CDF Collaboration. *The CDF-II detector: Technical design report*. Tech. rep. 1996. URL: https://cds.cern.ch/record/1478626 (on page 10).

[52] D0 Collaboration. “Observation of the Top Quark”. In: *Phys. Rev. Lett.* 74 (14 1995), pp. 2632–2637. URL: https://link.aps.org/doi/10.1103/PhysRevLett.74.2632 (on pages 10, 249).

[53] CDF Collaboration. “Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab”. In: *Phys. Rev. Lett.* 74 (14 1995), pp. 2626–2631. URL: https://link.aps.org/doi/10.1103/PhysRevLett.74.2626 (on pages 10, 249).

[54] P. Zyla et al. “Review of Particle Physics”. In: *PTEP* 2020.8 (2020), p. 083C01. URL: https://pdg.lbl.gov/index.html (on pages 10, 31, 33, 36, 91).

[55] LIGO Scientific Collaboration and Virgo Collaboration. “Observation of Gravitational Waves from a Binary Black Hole Merger”. In: *Phys. Rev. Lett.* 116 (6 2016), p. 061102. URL: https://link.aps.org/doi/10.1103/PhysRevLett.116.061102 (on page 10).

[56] T. Guerreiro. “Quantum effects in gravity waves”. In: *Classical and Quantum Gravity* 37.15 (2020), p. 155001. URL: https://doi.org/10.1088/1361-6382/ab9d5d (on page 10).

[57] P. A. M. Dirac. “Quantum theory of emission and absorption of radiation”. In: *Proc. Roy. Soc. Lond. A* 114 (1927), p. 243 (on page 11).

[58] H. A. Bethe. “The Electromagnetic Shift of Energy Levels”. In: *Phys. Rev.* 72 (4 1947), pp. 339–341. URL: https://link.aps.org/doi/10.1103/PhysRev.72.339 (on page 11).

[59] Muon $g-2$ Collaboration. “Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm”. In: *Phys. Rev. Lett.* 126 (14 2021), p. 141801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.126.141801 (on page 11).

[60] Muon g-2 Collaboration. “Final report of the E821 muon anomalous magnetic moment measurement at BNL”. In: *Phys. Rev. D* 73 (7 2006), p. 072003. URL: https://link.aps.org/doi/10.1103/PhysRevD.73.072003 (on page 11).

[61] J. Schwinger. “On Quantum-Electrodynamics and the Magnetic Moment of the Electron”. In: *Phys. Rev.* 73 (4 1948), pp. 416–417. URL: https://link.aps.org/doi/10.1103/PhysRev.73.416 (on page 11).

[62] F. J. Dyson. “The Radiation Theories of Tomonaga, Schwinger, and Feynman”. In: *Phys. Rev.* 75 (3 Feb. 1949), pp. 486–502. URL: https://link.aps.org/doi/10.1103/PhysRev.75.486 (on page 12).

[63] R. P. Feynman. “A Relativistic Cut-Off for Classical Electrodynamics”. In: *Phys. Rev.* 74 (8 Oct. 1948), pp. 939–946. URL: https://link.aps.org/doi/10.1103/PhysRev.74.939 (on page 12).

[64] J. Schwinger. “Quantum Electrodynamics. I. A Covariant Formulation”. In: *Phys. Rev.* 74 (10 1948), pp. 1439–1461. URL: https://link.aps.org/doi/10.1103/PhysRev.74.1439 (on pages 12, 261).

[65] S. Kanesawa and S.-i. Tomonaga. “On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields V*: Case of Interacting Electromagnetic and Meson Fields”. In: *Progress of Theoretical Physics* 3.2 (June 1948), pp. 101–113. ISSN: 0033-068X. eprint: https://academic.oup.com/ptp/article-pdf/3/2/101/5260297/3-2-101.pdf. URL: https://doi.org/10.1143/ptp/3.2.101 (on page 12).

[66] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory*. Boulder, CO: Westview, 1995. URL: https://cds.cern.ch/record/257493 (on page 12).

[67] M. D. Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, Mar. 2014. ISBN: 978-1-107-03473-0, 978-1-107-03473-0 (on page 12).

[68] E. Noether. “Invariante Variationsprobleme”. ger. In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1918 (1918), pp. 235–257. URL: http://eudml.org/doc/59024 (on page 13).

[69] E. Noether. “Invariant variation problems”. In: *Transport Theory and Statistical Physics* 1.3 (1971), pp. 186–207. ISSN: 1532-2424. URL: http://dx.doi.org/10.1080/00411457108231446 (on page 13).

[70] T. Tao. *What is a gauge?* URL: https://terrytao.wordpress.com/2008/09/27/what-is-a-gauge/ (on page 13).

[71] A. Zee. *Group Theory in a Nutshell for Physicists*. USA: Princeton University Press, 2016. ISBN: 978-0-691-16269-0, 978-0-691-16269-0, 978-1-4008-8118-5 (on page 14).

[72] H. Kragh. *Physics and the Totalitarian Principle*. 2019. arXiv: 1907.04623 [physics.hist-ph] (on page 16).

[73] J. C. Maxwell. “XXV. On physical lines of force”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 21.139 (1861), pp. 161–175. eprint: https://doi.org/10.1080/14786446108643033. URL: https://doi.org/10.1080/14786446108643033 (on page 17).

[74] J. C. Maxwell. “VIII. A dynamical theory of the electromagnetic field”. In: *Philosophical Transactions of the Royal Society of London* 155 (1865), pp. 459–512. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rstl.1865.0008. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rstl.1865.0008 (on page 17).

[75] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. In: *Phys. Rev. Lett.* 13 (9 1964), pp. 321–323. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321 (on pages 20, 37).

[76] P. W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (16 1964), pp. 508–509. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508 (on pages 20, 38).

[77] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". In: *Phys. Rev. Lett.* 13 (20 1964), pp. 585–587. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.585 (on page 20).

[78] J. Ellis. "Higgs Physics". In: (2013), 117–168. 52 p. arXiv: 1312.5672. URL: https://cds.cern.ch/record/1638469 (on pages 20, 21).

[79] P. Zyla et al. "Status of Higgs Physics". In: (2020). URL: https://pdg.lbl.gov/2020/reviews/rpp2020-rev-higgs-boson.pdf (on pages 21, 22, 36, 43, 45, 52, 53, 63, 247).

[80] Y. Nambu. "Quasi-Particles and Gauge Invariance in the Theory of Superconductivity". In: *Phys. Rev.* 117 (3 1960), pp. 648–663. URL: https://link.aps.org/doi/10.1103/PhysRev.117.648 (on pages 21, 44).

[81] J. Goldstone. "Field theories with « Superconductor » solutions". In: *Il Nuovo Cimento (1955-1965)* 19.1 (1961), pp. 154–164. ISSN: 1827-6121. URL: https://doi.org/10.1007/BF02812722 (on pages 21, 22, 44).

[82] J. Goldstone, A. Salam, and S. Weinberg. "Broken Symmetries". In: *Phys. Rev.* 127 (3 1962), pp. 965–970. URL: https://link.aps.org/doi/10.1103/PhysRev.127.965 (on pages 21, 44).

[83] ATLAS Collaboration. "Measurement of the W-boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 78.2 (2018), p. 110. arXiv: 1701.07240 [hep-ex] (on page 24).

[84] S. Schael et al. "Precision electroweak measurements on the Z resonance". In: *Phys. Rept.* 427 (2006), pp. 257–454. arXiv: hep-ex/0509008 (on page 24).

[85] Particle Data Group. "14. Neutrino Masses, Mixing, and Oscillations". In: (2020) (on page 27).

[86] M. Gell-Mann, P. Ramond, and R. Slansky. "Complex Spinors and Unified Theories". In: *Conf. Proc. C* 790927 (1979), pp. 315–321. arXiv: 1306.4669 [hep-th] (on page 27).

[87] *Proceedings: Workshop on the Unified Theories and the Baryon Number in the Universe: Tsukuba, Japan, February 13-14, 1979*. Tsukuba, Japan: Natl.Lab.High Energy Phys., 1979 (on page 27).

[88] R. N. Mohapatra and G. Senjanovi. "Neutrino Mass and Spontaneous Parity Nonconservation". In: *Phys. Rev. Lett.* 44 (14 1980), pp. 912–915. URL: https://link.aps.org/doi/10.1103/PhysRevLett.44.912 (on page 27).

[89] P. Ramond. "The Family Group in Grand Unified Theories". In: *International Symposium on Fundamentals of Quantum Theory and Quantum Field Theory*. Feb. 1979. arXiv: hep-ph/9809459 (on page 27).

[90] W. Konetschny and W. Kummer. "Nonconservation of total lepton number with scalar bosons". In: *Physics Letters B* 70.4 (1977), pp. 433–435. ISSN: 0370-2693. URL: https://www.sciencedirect.com/science/article/pii/0370269377904075 (on page 27).

[91] T. Yanagida. “Horizontal gauge symmetry and masses of neutrinos”. In: *Conf. Proc. C* 7902131 (1979). Ed. by O. Sawada and A. Sugamoto, pp. 95–99 (on page 27).

[92] N. Arkani-Hamed et al. “Neutrino masses from large extra dimensions”. In: *Physical Review D* 65.2 (2001). ISSN: 1089-4918. URL: http://dx.doi.org/10.1103/PhysRevD.65.024032 (on page 27).

[93] M. Drees et al. “A Supersymmetric resolution of solar and atmospheric neutrino puzzles”. In: *Phys. Rev. D* 57 (1998), pp. 5335–5339. arXiv: hep-ph/9712392 (on page 27).

[94] “LHC Guide”. 2017. URL: https://cds.cern.ch/record/2255762 (on page 30).

[95] *LEP design report*. Geneva: CERN, 1984. URL: https://cds.cern.ch/record/102083 (on pages 30, 204).

[96] DELPHI Collaboration, ALEPH Collaboration, L3 Collaboration, OPAL Collaboration. “Search for the Standard Model Higgs Boson at LEP”. In: CERN-ALEPH-2002-024. CERN-ALEPH-CONF-2002-013. DELPHI-2002-088-CONF-621. CERN-DELPHI-2002-088-CONF-621. L3-Note-2766. OPAL-TN721. LHWG-Note-2002-01 (2002), 28 p. URL: https://cds.cern.ch/record/2310899 (on page 30).

[97] *Design Report Tevatron 1 project*. Tech. rep. 1984. URL: https://cds.cern.ch/record/1478620 (on page 30).

[98] The TEVNPH Working Group. *Combined CDF and D0 Search for Standard Model Higgs Boson Production with up to 10.0 fb-1 of Data*. 2012. arXiv: 1203.3774 [hep-ex] (on page 30).

[99] M. PERELSTEIN. “INTRODUCTION TO COLLIDER PHYSICS”. In: *Physics of the Large and the Small* (2011). URL: http://dx.doi.org/10.1142/9789814327183_0008 (on page 31).

[100] S. Dittmaier et al. “Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables”. In: (Jan. 2011). arXiv: 1101.0593 [hep-ph] (on page 31).

[101] S. Dittmaier et al. “Handbook of LHC Higgs Cross Sections: 2. Differential Distributions”. In: (Jan. 2012). arXiv: 1201.3084 [hep-ph] (on page 31).

[102] J. R. Andersen et al. “Handbook of LHC Higgs Cross Sections: 3. Higgs Properties”. In: (July 2013). Ed. by S Heinemeyer et al. arXiv: 1307.1347 [hep-ph] (on pages 31, 36, 46).

[103] D. de Florian et al. “Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector”. In: 2/2017 (Oct. 2016). arXiv: 1610.07922 [hep-ph] (on pages 31–33, 36).

[104] K. Becker et al. *Precise predictions for boosted Higgs production*. 2020. arXiv: 2005.07762 [hep-ph] (on pages 31, 142, 172, 173).

[105] M. Rauch. *Vector-Boson Fusion and Vector-Boson Scattering*. 2016. arXiv: 1610.08420 [hep-ph] (on page 32).

[106] F. Bishara, R. Contino, and J. Rojo. “Higgs pair production in vector-boson fusion at the LHC and beyond”. In: *The European Physical Journal C* 77.7 (2017), p. 481. ISSN: 1434-6052. URL: https://doi.org/10.1140/epjc/s10052-017-5037-9 (on page 33).

[107] A. Djouadi. "The anatomy of electroweak symmetry breaking". In: *Physics Reports* 457.1-4 (2008), pp. 1–216. ISSN: 0370-1573. URL: http://dx.doi.org/10.1016/j.physrep.2007.10.004 (on page 33).

[108] F. Demartin et al. *Higgs production in association with a single top quark at the LHC*. 2015. arXiv: 1504.00611 [hep-ph] (on page 34).

[109] J. Chang et al. "Probing the top-Yukawa coupling in associated Higgs production with a single top quark". In: *Journal of High Energy Physics* 2014.5 (2014). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP05(2014)062 (on page 34).

[110] V. Barger, M. McCaskey, and G. Shaughnessy. "Single top and Higgs associated production at the LHC". In: *Physical Review D* 81.3 (2010). ISSN: 1550-2368. URL: http://dx.doi.org/10.1103/PhysRevD.81.034020 (on page 34).

[111] T. M. P. Tait and C.-P. Yuan. "Single top quark production as a window to physics beyond the standard model". In: *Physical Review D* 63.1 (2000). ISSN: 1089-4918. URL: http://dx.doi.org/10.1103/PhysRevD.63.014018 (on page 34).

[112] S. Biswas, E. Gabrielli, and B. Mele. "Single top and Higgs associated production as a probe of the $Ht\bar{t}$ coupling sign at the LHC". In: *Journal of High Energy Physics* 2013.1 (2013). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP01(2013)088 (on page 34).

[113] S. Biswas et al. "Direct constraints on the top-Higgs coupling from the 8 TeV LHC data". In: *Journal of High Energy Physics* 2013.7 (2013). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP07(2013)073 (on page 34).

[114] M. Farina et al. "Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson". In: *Journal of High Energy Physics* 2013.5 (2013). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP05(2013)022 (on page 34).

[115] P. Agrawal, S. Mitra, and A. Shivaji. "Effect of anomalous couplings on the associated production of a single top quark and a Higgs boson at the LHC". In: *Journal of High Energy Physics* 2013.12 (2013). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP12(2013)077 (on page 34).

[116] J. Ellis et al. "Disentangling Higgs-top couplings in associated production". In: *Journal of High Energy Physics* 2014.4 (2014). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP04(2014)004 (on page 34).

[117] C. Englert and E. Re. "Bounding the top Yukawa coupling with Higgs-associated single-top production". In: *Physical Review D* 89.7 (2014). ISSN: 1550-2368. URL: http://dx.doi.org/10.1103/PhysRevD.89.073020 (on page 34).

[118] M. Spira. "Higgs boson production and decay at hadron colliders". In: *Progress in Particle and Nuclear Physics* 95 (2017), pp. 98–159. ISSN: 0146-6410. URL: http://dx.doi.org/10.1016/j.ppnp.2017.04.001 (on page 35).

[119] C. Mariotti. "PoS (DIS2017) 015 Higgs Physics at the LHC". In: 2017 (on page 37).

[120] "The Nobel Prize in Physics 2013". en. In: *NobelPrize.org. Nobel Media AB* (2021) (on page 37).

[121] ATLAS Collaboration. "Measurements of Higgs boson production and couplings in the four-lepton channel in *pp* collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector". In: *Phys. Rev. D* 91 (1 2015), p. 012006. URL: https://link.aps.org/doi/10.1103/PhysRevD.91.012006 (on page 39).

[122] CMS Collaboration. "Measurement of the properties of a Higgs boson in the four-lepton final state". In: *Physical Review D* 89.9 (2014). ISSN: 1550-2368. URL: http://dx.doi.org/10.1103/PhysRevD.89.092007 (on page 39).

[123] ATLAS Collaboration. "Observation and measurement of Higgs boson decays to WW* with the ATLAS detector". In: *Physical Review D* 92.1 (2015). ISSN: 1550-2368. URL: http://dx.doi.org/10.1103/PhysRevD.92.012006 (on page 39).

[124] CMS Collaboration. "Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Physics Letters B* 791 (2019), pp. 96–129. ISSN: 0370-2693. URL: http://dx.doi.org/10.1016/j.physletb.2018.12.073 (on page 39).

[125] ATLAS Collaboration. "Cross-section measurements of the Higgs boson decaying into a pair of -leptons in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector". In: *Physical Review D* 99.7 (2019). ISSN: 2470-0029. URL: http://dx.doi.org/10.1103/PhysRevD.99.072001 (on pages 39, 244).

[126] CMS Collaboration. "Observation of the Higgs boson decay to a pair of leptons with the CMS detector". In: *Physics Letters B* 779 (Apr. 2018), pp. 283–316. ISSN: 0370-2693. URL: http://dx.doi.org/10.1016/j.physletb.2018.02.004 (on pages 39, 244).

[127] ATLAS Collaboration. "A search for the dimuon decay of the Standard Model Higgs boson with the ATLAS detector". In: *Physics Letters B* 812 (2021), p. 135980. ISSN: 0370-2693. URL: http://dx.doi.org/10.1016/j.physletb.2020.135980 (on page 39).

[128] CMS Collaboration. "Evidence for Higgs boson decay to a pair of muons". In: *Journal of High Energy Physics* 2021.1 (2021). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP01(2021)148 (on page 39).

[129] ATLAS Collaboration. "Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector". In: *Physics Letters B* 786 (Nov. 2018), pp. 59–86. ISSN: 0370-2693. URL: http://dx.doi.org/10.1016/j.physletb.2018.09.013 (on pages 39, 52, 216).

[130] ATLAS Collaboration. "Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector". In: *Physics Letters B* 784 (2018), pp. 173–191. ISSN: 0370-2693. URL: http://dx.doi.org/10.1016/j.physletb.2018.07.035 (on page 39).

[131] CMS Collaboration. "Observation of $t\bar{t}H$ Production". In: *Physical Review Letters* 120.23 (2018). ISSN: 1079-7114. URL: http://dx.doi.org/10.1103/PhysRevLett.120.231801 (on page 39).

[132] ATLAS Collaboration. "Evidence for the spin-0 nature of the Higgs boson using ATLAS data". In: *Physics Letters B* 726.1-3 (2013), pp. 120–144. ISSN: 0370-2693. URL: http://dx.doi.org/10.1016/j.physletb.2013.08.026 (on page 39).

[133] CMS Collaboration. "Study of the Mass and Spin-Parity of the Higgs Boson Candidate via Its Decays to Z Boson Pairs". In: *Physical Review Letters* 110.8 (Feb. 2013). ISSN: 1079-7114. URL: http://dx.doi.org/10.1103/PhysRevLett.110.081803 (on page 39).

[134] ATLAS Collaboration. "Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment". In: *The European Physical Journal C* 76.1 (2016). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-015-3769-y (on page 39).

[135] CMS Collaboration. "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV". In: *The European Physical Journal C* 75.5 (2015). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-015-3351-7 (on page 39).

[136] ATLAS and CMS Collaborations. "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". In: *Journal of High Energy Physics* 2016.8 (2016). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP08(2016)045 (on page 39).

[137] CMS Collaboration. "Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV". In: *Journal of High Energy Physics* 2017.11 (2017). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP11(2017)047 (on page 39).

[138] ATLAS Collaboration. *A combination of measurements of Higgs boson production and decay using up to* 139 fb^{-1} *of proton–proton collision data at* $\sqrt{s} = 13$ TeV *collected with the ATLAS experiment*. Tech. rep. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2725733 (on page 40).

[139] CMS Collaboration. *Combined Higgs boson production and decay measurements with up to 137 fb-1 of proton-proton collision data at* $\sqrt{s} = 13$ TeV. Tech. rep. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2706103 (on page 41).

[140] G. Esposito. *An introduction to quantum gravity*. 2011. arXiv: 1108.3269 [hep-th] (on page 42).

[141] G. Bertone, D. Hooper, and J. Silk. "Particle dark matter: evidence, candidates and constraints". In: *Physics Reports* 405.5-6 (2005), pp. 279–390. ISSN: 0370-1573. URL: http://dx.doi.org/10.1016/j.physrep.2004.08.031 (on page 42).

[142] M. Li et al. "Dark Energy". In: *Communications in Theoretical Physics* 56.3 (2011), pp. 525–604. ISSN: 0253-6102. URL: http://dx.doi.org/10.1088/0253-6102/56/3/24 (on page 42).

[143] L. Canetti, M. Drewes, and M. Shaposhnikov. "Matter and antimatter in the universe". In: *New Journal of Physics* 14.9 (2012), p. 095012. ISSN: 1367-2630. URL: http://dx.doi.org/10.1088/1367-2630/14/9/095012 (on page 42).

[144] G. Degrassi et al. "Higgs mass and vacuum stability in the Standard Model at NNLO". In: *Journal of High Energy Physics* 2012.8 (2012). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP08(2012)098 (on pages 43, 266, 267).

[145] C. Csáki and P. Tanedo. "Beyond the Standard Model". In: *2013 European School of High-Energy Physics*. 2015, pp. 169–268. arXiv: 1602.04228 [hep-ph] (on pages 43, 268).

[146] G. Bhattacharyya. "Hierarchy problem and BSM physics". In: *Pramana* 89.4 (2017), p. 53 (on pages 43, 268).

[147] Allanach, B.C. and Haber, H.E. "Supersymmetry, Part I (Theory)". In: (). URL: https://pdg.lbl.gov/2019/reviews/rpp2018-rev-susy-1-theory.pdf (on page 43).

[148] M. P. Hertzberg. "On inflation with non-minimal coupling". In: *Journal of High Energy Physics* 2010.11 (2010). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP11(2010)023 (on page 43).

[149] A. Djouadi. "The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model". In: *Phys. Rept.* 459 (2008), pp. 1–241. arXiv: hep-ph/0503173 (on page 43).

[150] G. Branco et al. "Theory and phenomenology of two-Higgs-doublet models". In: *Physics Reports* 516.1-2 (2012), pp. 1–102. ISSN: 0370-1573. URL: http://dx.doi.org/10.1016/j.physrep.2012.02.002 (on pages 43, 50).

[151] J. F. Gunion et al. *The Higgs Hunter's Guide*. Vol. 80. 2000 (on pages 43, 45).

[152] H. E. Haber. "Nonminimal Higgs sectors: The Decoupling limit and its phenomenological implications". In: *Joint U.S.-Polish Workshop on Physics from Planck Scale to Electro-Weak Scale (SUSY 94)*. Dec. 1994. arXiv: hep-ph/9501320 (on page 45).

[153] J. F. Gunion and H. E. Haber. "The CP conserving two Higgs doublet model: The Approach to the decoupling limit". In: *Phys. Rev. D* 67 (2003), p. 075019. arXiv: hep-ph/0207010 (on page 45).

[154] E. L. Berger et al. "Higgs boson decay into hadronic jets". In: *Phys. Rev. D* 66 (2002), p. 095001. arXiv: hep-ph/0205342 (on page 45).

[155] P. Draper, G. Lee, and C. E. M. Wagner. "Precise estimates of the Higgs mass in heavy supersymmetry". In: *Phys. Rev. D* 89.5 (2014), p. 055023. arXiv: 1312.5743 [hep-ph] (on page 45).

[156] L. H. C. S. W. Group et al. *LHC HXSWG interim recommendations to explore the coupling structure of a Higgs-like particle*. 2012. arXiv: 1209.0040 [hep-ph] (on page 46).

[157] J. de Blas et al. "Higgs Boson studies at future particle colliders". In: *Journal of High Energy Physics* 2020.1 (2020). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP01(2020)139 (on pages 46, 246).

[158] H. Pilkuhn. "The Interaction of Hadrons". In: *American Journal of Physics* 36.9 (1968), pp. 855–855. eprint: https://doi.org/10.1119/1.1975176. URL: https://doi.org/10.1119/1.1975176 (on page 46).

[159] D. Berdine, N. Kauer, and D. Rainwater. "Breakdown of the Narrow Width Approximation for New Physics". In: *Physical review letters* 99 (Oct. 2007), p. 111601 (on page 46).

[160] F. Wilczek. "Decays of Heavy Vector Mesons into Higgs Particles". In: *Phys. Rev. Lett.* 39 (21 1977), pp. 1304–1306. URL: https://link.aps.org/doi/10.1103/PhysRevLett.39.1304 (on page 47).

[161] T. Inami, T. Kubota, and Y. Okada. "Effective Gauge Theory and the Effect of Heavy Quarks in Higgs Boson Decays". In: *Z. Phys. C* 18 (1983), pp. 69–80 (on page 47).

[162] N. Greiner et al. "Full mass dependence in Higgs boson production in association with jets at the LHC and FCC". In: *JHEP* 01 (2016), 091. 32 p. arXiv: 1608.01195. URL: https://cds.cern.ch/record/2205473 (on page 47).

[163] M. Aaboud et al. "Search for new phenomena in events with same-charge leptons and b-jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Journal of High Energy Physics* 2018.12 (2018). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP12(2018)039 (on page 47).

[164] ATLAS Collaboration. "Search for four-top-quark production in the single-lepton and opposite-sign dilepton final states in *pp* collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Physical Review D* 99.5 (2019). ISSN: 2470-0029. URL: http://dx.doi.org/10.1103/PhysRevD.99.052009 (on page 47).

[165] A. M. Sirunyan et al. "Search for production of four top quarks in final states with same-sign or multiple leptons in proton–proton collisions at $\sqrt{s} = 13$ TeV". In: *The European Physical Journal C* 80.2 (2020). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-019-7593-7 (on page 47).

[166] CMS Collaboration. "Search for the production of four top quarks in the single-lepton and opposite-sign dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Journal of High Energy Physics* 2019.11 (2019). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP11(2019)082 (on page 47).

[167] CMS Collaboration. "Search for physics beyond the standard model in events with two leptons of same sign, missing transverse momentum, and jets in proton–proton collisions at $\sqrt{s} = 13$ TeV". In: *The European Physical Journal C* 77.9 (2017). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-017-5079-z (on page 47).

[168] K. Kudashkin et al. "Higgs bosons with large transverse momentum at the LHC". In: *Physics Letters B* 782 (2018), pp. 210–214. ISSN: 0370-2693. URL: https://www.sciencedirect.com/science/article/pii/S0370269318303733 (on pages 48, 49).

[169] C. Arnesen, I. Z. Rothstein, and J. Zupan. "Smoking Guns for On-Shell New Physics at the LHC". In: *Physical Review Letters* 103.15 (2009). ISSN: 1079-7114. URL: http://dx.doi.org/10.1103/PhysRevLett.103.151801 (on page 48).

[170] A. Banfi, A. Martin, and V. Sanz. "Probing top-partners in Higgs+jets". In: *Journal of High Energy Physics* 2014.8 (2014). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP08(2014)053 (on page 48).

[171] A. Azatov et al. "Taming the off-shell Higgs boson". In: *Journal of Experimental and Theoretical Physics* 120.3 (2015), pp. 354–368. ISSN: 1090-6509. URL: http://dx.doi.org/10.1134/S1063776115030140 (on page 48).

[172] C. Grojean et al. "Very boosted Higgs in gluon fusion". In: *Journal of High Energy Physics* 2014.5 (2014). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP05(2014)022 (on page 48).

[173] T. Neumann and M. Wiesemann. "Finite top-mass effects in gluon-induced Higgs production with a jet-veto at NNLO". In: *Journal of High Energy Physics* 2014.11 (2014). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP11(2014)150 (on page 48).

[174] C. Englert, M. McCullough, and M. Spannowsky. "Gluon-initiated associated production boosts Higgs physics". In: *Phys. Rev. D* 89 (1 2014), p. 013013. URL: https://link.aps.org/doi/10.1103/PhysRevD.89.013013 (on page 48).

[175] M. Buschmann et al. "Mass effects in the Higgs-gluon coupling: boosted vs. off-shell production". In: *Journal of High Energy Physics* 2015.2 (2015). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP02(2015)038 (on page 48).

[176] R. Contino et al. "On the validity of the effective field theory approach to SM precision tests". In: *Journal of High Energy Physics* 2016.7 (2016). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP07(2016)144 (on page 49).

[177] R. Penco. *An Introduction to Effective Field Theories*. 2020. arXiv: 2006.16285 [hep-th] (on page 49).

[178] J. R. Espinosa. "Higgs Effective Field Theory". In: *PoS* CORFU2015 (2016), p. 012 (on page 49).

[179] W. Buchmüller and D. Wyler. "Effective lagrangian analysis of new interactions and flavour conservation". In: *Nuclear Physics B* 268.3 (1986), pp. 621–653. ISSN: 0550-3213. URL: https://www.sciencedirect.com/science/article/pii/0550321386902622 (on page 49).

[180] K. Hagiwara et al. "Low energy effects of new interactions in the electroweak boson sector". In: *Phys. Rev. D* 48 (5 1993), pp. 2182–2203. URL: https://link.aps.org/doi/10.1103/PhysRevD.48.2182 (on page 49).

[181] B. Grzadkowski et al. "Dimension-six terms in the Standard Model Lagrangian". In: *Journal of High Energy Physics* 2010.10 (2010). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP10(2010)085 (on pages 49, 50, 241).

[182] R. Contino et al. "Effective Lagrangian for a light Higgs-like scalar". In: *Journal of High Energy Physics* 2013.7 (2013). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP07(2013)035 (on pages 50, 241).

[183] T. D. Lee. "A Theory of Spontaneous T Violation". In: *Phys. Rev. D* 8 (4 1973), pp. 1226–1239. URL: https://link.aps.org/doi/10.1103/PhysRevD.8.1226 (on page 50).

[184] T. D. Lee. “CP Nonconservation and Spontaneous Symmetry Breaking”. In: *Phys. Rept.* 9 (1974). Ed. by G. Feinberg, pp. 143–177 (on page 50).

[185] R. D. Peccei and H. R. Quinn. “CP Conservation in the Presence of Instantons”. In: *Phys. Rev. Lett.* 38 (1977), pp. 1440–1443 (on page 50).

[186] P. Fayet. “A Gauge Theory of Weak and Electromagnetic Interactions with Spontaneous Parity Breaking”. In: *Nucl. Phys. B* 78 (1974), pp. 14–28 (on page 50).

[187] K. Inoue et al. “Low-Energy Parameters and Particle Masses in a Supersymmetric Grand Unified Model”. In: *Prog. Theor. Phys.* 67 (1982), p. 1889 (on page 50).

[188] R. A. Flores and M. Sher. “Higgs Masses in the Standard, Multi-Higgs and Supersymmetric Models”. In: *Annals Phys.* 148 (1983), p. 95 (on page 50).

[189] J. F. Gunion and H. E. Haber. “Higgs bosons in supersymmetric models (I)”. In: *Nuclear Physics B* 272.1 (1986), pp. 1–76. ISSN: 0550-3213. URL: https://www.sciencedirect.com/science/article/pii/0550321386903408 (on page 50).

[190] F. J. Botella and J. P. Silva. “Jarlskog - like invariants for theories with scalars and fermions”. In: *Phys. Rev. D* 51 (1995), pp. 3870–3875. arXiv: hep-ph/9411288 (on page 50).

[191] M. Carena and H. E. Haber. “Higgs Boson Theory and Phenomenology”. In: *Prog. Part. Nucl. Phys.* 50 (2003), pp. 63–152. arXiv: hep-ph/0208209 (on page 50).

[192] H. Bélusca-Maïto et al. “Higgs EFT for 2HDM and beyond”. In: *The European Physical Journal C* 77.3 (2017). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-017-4745-5 (on page 50).

[193] S. Weinberg. “Effective gauge theories”. In: *Physics Letters B* 91.1 (1980), pp. 51–55. ISSN: 0370-2693. URL: https://www.sciencedirect.com/science/article/pii/0370269380906607 (on page 50).

[194] A. V. Manohar. “Effective field theories”. In: *Perturbative and Nonperturbative Aspects of Quantum Field Theory*. Ed. by H. Latal and W. Schweiger. Berlin, Heidelberg: Springer Berlin Heidelberg, 1997, pp. 311–362. ISBN: 978-3-540-49735-6 (on pages 50, 240).

[195] D. B. Kaplan. “Five lectures on effective field theory”. In: Oct. 2005. arXiv: nucl-th/0510023 (on pages 50, 240).

[196] M. Gorbahn, J. M. No, and V. Sanz. *Benchmarks for Higgs Effective Theory: Extended Higgs Sectors*. 2015. arXiv: 1502.07352 [hep-ph] (on page 50).

[197] J. Brehmer et al. “Pushing Higgs effective theory to its limits”. In: *Physical Review D* 93.7 (2016). ISSN: 2470-0029. URL: http://dx.doi.org/10.1103/PhysRevD.93.075014 (on page 50).

[198] S. Alte, M. König, and M. Neubert. “Effective Theory for a Heavy Scalar Coupled to the SM via Vector-Like Quarks”. In: *Eur. Phys. J. C* 79.4 (2019), p. 352. arXiv: 1902.04593 [hep-ph] (on page 50).

[199] M. Grazzini et al. "Modeling BSM effects on the Higgs transverse-momentum spectrum in an EFT approach". In: *Journal of High Energy Physics* 2017.3 (2017). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP03(2017)115 (on pages 51, 241–244).

[200] CMS Collaboration. "Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment". In: *The European Physical Journal C* 76.1 (2016). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-015-3769-y (on page 52).

[201] CMS Collaboration. "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV". In: *The European Physical Journal C* 75.5 (2015). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-015-3351-7 (on page 52).

[202] ATLAS and CMS Collaborations. "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. 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". In: *JHEP* 08 (2016), 045. 70 p. arXiv: 1606.02266. URL: https://cds.cern.ch/record/2158863 (on page 52).

[203] ATLAS Collaboration. "Measurements of the Higgs boson production cross section at 7, 8 and 13 TeV centre-of-mass energies and search for new physics at 13 TeV in the $H \to ZZ^* \to \ell^+\ell\ell'^+\ell'$ final state with the ATLAS detector". In: (Dec. 2015) (on page 52).

[204] LHCb Collaboration. "Search for $H^0 \to b\bar{b}$ or $c\bar{c}$ in association with a W or Z boson in the forward region of *pp* collisions". In: (2016). URL: https://cds.cern.ch/record/2209531 (on page 53).

[205] ATLAS Collaboration. "Search for the Standard Model Higgs boson produced by vector-boson fusion and decaying to bottom quarks in $\sqrt{s} = 8$ TeV *pp* collisions with the ATLAS detector". In: *Journal of High Energy Physics* 2016.11 (2016). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP11(2016)112 (on page 53).

[206] ATLAS Collaboration. "Search for Higgs bosons produced via vector-boson fusion and decaying into bottom quark pairs in $\sqrt{s} = 13$ TeV *pp* collisions with the ATLAS detector". In: *Physical Review D* 98.5 (2018). ISSN: 2470-0029. URL: http://dx.doi.org/10.1103/PhysRevD.98.052003 (on page 53).

[207] ATLAS Collaboration. *Measurements of Higgs Bosons Decaying to Bottom Quarks from Vector Boson Fusion Production with the ATLAS Experiment at* $\sqrt{s} = 13$ TeV. 2020. arXiv: 2011.08280 [hep-ex] (on pages 53, 244).

[208] E. Gabrielli et al. "Higgs boson production in association with a photon in vector boson fusion at the LHC". In: *Nuclear Physics B* 781.1-3 (2007), pp. 64–84. ISSN: 0550-3213. URL: http://dx.doi.org/10.1016/j.nuclphysb.2007.05.010 (on page 53).

[209] E. Gabrielli et al. “Asking for an extra photon in Higgs production at the LHC and beyond”. In: *Journal of High Energy Physics* 2016.7 (2016). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP07(2016)003 (on page 53).

[210] ATLAS Collaboration. “Search for Higgs boson production in association with a high-energy photon via vector-boson fusion with decay into bottom quark pairs at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector”. In: *JHEP* 03 (2021), p. 268. arXiv: 2010.13651 [hep-ex] (on page 53).

[211] ATLAS Collaboration. “Search for the Standard Model Higgs boson decaying into $b\bar{b}$ produced in association with top quarks decaying hadronically in pp collisions at $\sqrt{s} = 8\,\text{TeV}$ with the ATLAS detector”. In: *JHEP* 05 (2016), p. 160. arXiv: 1604.03812 [hep-ex] (on page 53).

[212] ATLAS Collaboration. “Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector”. In: *Eur. Phys. J. C* 75.7 (2015), p. 349. arXiv: 1503.05066 [hep-ex] (on page 53).

[213] ATLAS Collaboration. *Measurement of the Higgs boson decaying to b-quarks produced in association with a top-quark pair in pp collisions at* $\sqrt{s} = 13\,\text{TeV}$ *with the ATLAS detector*. Tech. rep. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2743685 (on page 53).

[214] CMS Collaboration. “Search for the associated production of the Higgs boson with a top-quark pair”. In: *Journal of High Energy Physics* 2014.9 (2014). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP09(2014)087 (on page 53).

[215] CMS Collaboration. “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”. In: *The European Physical Journal C* 75.6 (2015). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-015-3454-1 (on page 53).

[216] CMS Collaboration. *Measurement of* $\text{t}\bar{\text{t}}\text{H}$ *production in the* $\text{H} \to \text{b}\bar{\text{b}}$ *decay channel in* $41.5\,\text{fb}^{-1}$ *of proton-proton collision data at* $\sqrt{s} = 13\,\text{TeV}$. Tech. rep. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2675023 (on page 53).

[217] LHC-ILC Study Group. “Physics Interplay of the LHC and the ILC. Physics interplay of the LHC and the ILC”. In: *Phys. Rep.* 426.SLAC-PUB-10764. ANL-HEP-PR-2004-108. CERN-PH-TH-2004-214. DCPT-2004-134. DESY-04-206. DESY-2004-206. IFIC-2004-59. IISc-CHEP-2004-13. IPPP-2004-67. SLAC-PUB-10764. UB-ECM-PF-2004-31. UCD-2004-28. UCI-TR-2004-37. 2-6 (2004), 47–358. 312 p. URL: https://cds.cern.ch/record/800650 (on page 54).

[218] ATLAS Collaboration. *Search for boosted resonances decaying to two b-quarks and produced in association with a jet at* $\sqrt{s} = 13\,\text{TeV}$ *with the ATLAS detector*. Tech. rep. ATLAS-CONF-2018-052. Geneva: CERN, 2018. URL: https://cds.cern.ch/record/2649081 (on pages 54, 139, 163, 171).

[219] CMS Collaboration. "Inclusive Search for a Highly Boosted Higgs Boson Decaying to a Bottom Quark-Antiquark Pair". In: *Physical Review Letters* 120.7 (2018). ISSN: 1079-7114. URL: http://dx.doi.org/10.1103/PhysRevLett.120.071802 (on pages 54, 139).

[220] ATLAS Collaboration. *Study of Higgs-boson production with large transverse momentum using the $H \to b\bar{b}$ decay with the ATLAS detector*. Tech. rep. Geneva: CERN, 2021. URL: https://cds.cern.ch/record/2759284 (on pages 55, 139, 234, 240, 245).

[221] ATLAS Collaboration. Tech. rep. (on pages 55, 139, 234, 240, 245).

[222] CMS Collaboration. "Inclusive search for a highly boosted Higgs boson decaying to a bottom quark-antiquark pair at $\sqrt{s} = 13\,\text{TeV}$ with $137\,\text{fb}^{-1}$". In: (2020) (on pages 55, 139, 237, 238, 240, 245).

[223] T.-M. Yan and S. D. Drell. "The parton model and its applications". In: *International Journal of Modern Physics A* 29.30 (2014), p. 1430071. ISSN: 1793-656X. URL: http://dx.doi.org/10.1142/S0217751X14300713 (on page 55).

[224] F. Giuli. *Determination of proton parton distribution functions using ATLAS data*. 2019. arXiv: 1909.06702 [hep-ex] (on page 55).

[225] M. Rinaldi. "The proton structure via double parton scattering". In: *Journal of Physics: Conference Series* 1643 (2020), p. 012193. ISSN: 1742-6596. URL: http://dx.doi.org/10.1088/1742-6596/1643/1/012193 (on page 55).

[226] X. G. Wang and A. W. Thomas. "Refined analysis on the parton distribution functions of the proton". In: *Journal of Physics G: Nuclear and Particle Physics* 47.1 (2019), p. 015102. ISSN: 1361-6471. URL: http://dx.doi.org/10.1088/1361-6471/ab51b3 (on page 55).

[227] C. W. Bauer and F. J. Tackmann. "Gaining analytic control of parton showers". In: *Physical Review D* 76.11 (2007). ISSN: 1550-2368. URL: http://dx.doi.org/10.1103/PhysRevD.76.114017 (on page 55).

[228] Z. Nagy and D. E. Soper. "What is a parton shower?" In: *Physical Review D* 98.1 (2018). ISSN: 2470-0029. URL: http://dx.doi.org/10.1103/PhysRevD.98.014034 (on page 55).

[229] T. Sjostrand. "Monte Carlo Generators". In: *2006 European School of High-Energy Physics*. Nov. 2006. arXiv: hep-ph/0611247 (on page 55).

[230] P. Skands. *Introduction to QCD*. 2017. arXiv: 1207.2389 [hep-ph] (on page 56).

[231] M. Lavelle and D. McMullan. "Collinearity, convergence and cancelling infrared divergences". In: *Journal of High Energy Physics* 2006.03 (2006), pp. 026–026. ISSN: 1029-8479. URL: http://dx.doi.org/10.1088/1126-6708/2006/03/026 (on page 56).

[232] T. D. Lee and M. Nauenberg. "Degenerate Systems and Mass Singularities". In: *Phys. Rev.* 133 (1964). Ed. by G. Feinberg, B1549–B1562 (on page 56).

[233] F. Bloch and A. Nordsieck. "Note on the Radiation Field of the Electron". In: *Phys. Rev.* 52 (2 1937), pp. 54–59. URL: https://link.aps.org/doi/10.1103/PhysRev.52.54 (on page 56).

[234] T. Kinoshita. “Mass singularities of Feynman amplitudes”. In: *J. Math. Phys.* 3 (1962), pp. 650–677 (on page 56).

[235] S. Marzani, G. Soyez, and M. Spannowsky. “Looking Inside Jets”. In: *Lecture Notes in Physics* (2019). ISSN: 1616-6361. URL: http://dx.doi.org/10.1007/978-3-030-15709-8 (on page 56).

[236] G. P. Salam. “Towards jetography”. In: *The European Physical Journal C* 67.3-4 (2010), pp. 637–686. ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-010-1314-6 (on pages 56, 60, 116, 137).

[237] B. R. Webber. “Hadronization”. In: *Summer School on Hadronic Aspects of Collider Physics*. Nov. 1994. arXiv: hep-ph/9411384 (on page 56).

[238] *Particle Jet*. URL: https://cms.cern/tags/particle-jet (on page 56).

[239] N. Cabibbo. “Unitary Symmetry and Leptonic Decays”. In: *Phys. Rev. Lett.* 10 (12 1963), pp. 531–533. URL: https://link.aps.org/doi/10.1103/PhysRevLett.10.531 (on page 57).

[240] M. Kobayashi and T. Maskawa. “CP Violation in the Renormalizable Theory of Weak Interaction”. In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657 (on page 57).

[241] M. Artuso, E. Barberio, and S. Stone. “B meson decays”. In: *PMC Physics A* 3.1 (2009), p. 3. ISSN: 1754-0410. URL: https://doi.org/10.1186/1754-0410-3-3 (on page 57).

[242] “The CKM Quark-Mixing Matrix”. In: (2020). URL: https://pdg.lbl.gov/2020/reviews/rpp2020-rev-ckm-matrix.pdf (on page 57).

[243] “Production and Decay of b-flavored Hadrons”. In: *PTEP* 2020.8 (2020), p. 083C01 (on page 57).

[244] ATLAS Collaboration. “ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13\,\text{TeV}$”. In: *The European Physical Journal C* 79.11 (2019). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-019-7450-8 (on pages 57, 143, 144, 217).

[245] P. Zyla et al. “Top Quark”. In: (2019). URL: https://pdg.lbl.gov/2019/reviews/rpp2019-rev-top-quark.pdf (on page 57).

[246] ATLAS Collaboration. *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*. Tech. rep. ATL-PHYS-PUB-2016-012. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2160731 (on pages 57–59, 63, 105–107, 112).

[247] P Perret. “b Physics”. In: *AIP Conf. Proc.* 444 (1998), 395–415. 24 p. URL: https://cds.cern.ch/record/372680 (on page 57).

[248] DELPHI Collaboration. “Measurement of the charged particle multiplicity of weakly decaying B hadrons”. In: *Phys. Lett. B* 425 (1998), 399–412. 17 p. URL: https://cds.cern.ch/record/348599 (on page 57).

[249] L3 Collaboration. “Measurement of the average lifetime of b-hadrons in Z decays”. In: *Phys. Lett. B* 416 (1997), 220–232. 19 p. URL: https://cds.cern.ch/record/335781 (on page 57).

[250] A. Coccaro. *Track Reconstruction and b-Jet Identification for the ATLAS Trigger System*. Tech. rep. Geneva: CERN, 2011. arXiv: 1112.0180. URL: https://cds.cern.ch/record/1402986 (on page 58).

[251] A. Chisholm. *Introduction to Heavy Flavour Jet Tagging with ATLAS*. URL: https://indico.cern.ch/event/655628/contributions/2670400/attachments/1518249/2370617/ASC_FTagHbbWS_BTagging.pdf (on pages 58, 61).

[252] J. M. Butterworth et al. "Jet Substructure as a New Higgs-Search Channel at the Large Hadron Collider". In: *Physical Review Letters* 100.24 (2008). ISSN: 1079-7114. URL: http://dx.doi.org/10.1103/PhysRevLett.100.242001 (on pages 60, 137).

[253] M. Aaboud et al. "Search for resonances in diphoton events at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector". In: *Journal of High Energy Physics* 2016.9 (2016). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP09(2016)001 (on page 61).

[254] ATLAS Collaboration. *Search for scalar diphoton resonances with 15.4 fb^{-1} of data collected at $\sqrt{s}$=13 TeV in 2015 and 2016 with the ATLAS detector*. Tech. rep. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2206154 (on page 61).

[255] CMS Collaboration. *Search for resonant production of high mass photon pairs using 12.9 fb^{-1} of proton-proton collisions at $\sqrt{s}$ = 13 TeV and combined interpretation of searches at 8 and 13 TeV*. Tech. rep. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2205245 (on page 61).

[256] CMS Collaboration. "Search for Resonant Production of High-Mass Photon Pairs in Proton-Proton Collisions at $\sqrt{s} = 8$ and 13 TeV". In: *Physical Review Letters* 117.5 (2016). ISSN: 1079-7114. URL: http://dx.doi.org/10.1103/PhysRevLett.117.051802 (on page 61).

[257] L. Lista. "Practical Statistics for Particle Physicists". In: *2016 European School of High-Energy Physics*. Sept. 2016. arXiv: 1609.04150 [physics.data-an] (on pages 62, 213).

[258] L. Lista. *Statistical Methods for Data Analysis in Particle Physics*. Vol. 909. Springer, 2016. ISBN: 978-3-319-20175-7, 978-3-319-20176-4 (on pages 62, 213).

[259] *LuminosityPublicResultsRun2*. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2#Luminosity_summary_plots_for_201 (on pages 63, 73).

[260] Przysiezniak, H. "Gluon splitting into heavy quarks in e^+e^- annihilations". In: CERN-PPE-96-091G (1996), 5 p. URL: https://cds.cern.ch/record/320796 (on page 63).

[261] OPAL Collaboration. "Production rate of $b\bar{b}$ quark pairs from gluons and $b\bar{b}b\bar{b}$ events in hadronic Z decays". In: *Eur. Phys. J. C* 18 (2000), 447–460. 23 p. URL: https://cds.cern.ch/record/468615 (on page 63).

[262] DELPHI Collaboration. "Measurement of the rate of $b\bar{b}b\bar{b}$ events in hadronic Z decays and the extraction of the gluon splitting into $b\bar{b}$". In: *Phys. Lett. B* 462 (1999), 425. 18 p. URL: https://cds.cern.ch/record/393663 (on page 63).

[263] ALEPH Collaboration. "A measurement of the gluon splitting rate into $b\bar{b}$ pairs in hadronic Z decays". In: *Phys. Lett. B* 434 (1998), 437. 15 p. URL: https://cds.cern.ch/record/359635 (on page 63).

[264] DELPHI Collaboration. "Measurement of the Multiplicity of Gluons Splitting to Bottom Quark Pairs in Hadronic Z^0 Decays". In: *Phys. Lett. B* 405 (1997), 202–214. 18 p. URL: https://cds.cern.ch/record/325272 (on page 63).

[265] C. Patrignani et al. "Review of Particle Physics R". In: *Chin. Phys. CR* 40.10 (2016), p. 100001 (on page 64).

[266] ATLAS Collaboration. In: 3.08 (2008), S08003–S08003. URL: https://doi.org/10.1088/1748-0221/3/08/s08003 (on pages 66, 84).

[267] CMS Collaboration. In: 3.08 (2008), S08004–S08004. URL: https://doi.org/10.1088/1748-0221/3/08/s08004 (on page 66).

[268] ALICE Collaboration. In: 3.08 (2008), S08002–S08002. URL: https://doi.org/10.1088/1748-0221/3/08/s08002 (on page 66).

[269] LHCb Collaboration. In: 3.08 (2008), S08005–S08005. URL: https://doi.org/10.1088/1748-0221/3/08/s08005 (on page 66).

[270] C. Service graphique. "Overall view of the LHC. Vue d'ensemble du LHC". In: (2014). URL: https://cds.cern.ch/record/1708849 (on page 67).

[271] D. Boussard and T. P. R. Linnecar. *The LHC Superconducting RF System*. Tech. rep. Geneva: CERN, 1999. URL: https://cds.cern.ch/record/410377 (on page 67).

[272] Rossi, L. "Superconducting magnets for the LHC main lattice". In: *IEEE Trans. Appl. Supercond.* 14.LHC-Project-Report-718. CERN-LHC-Project-Report-718. 2 (2004), 153–158. 7 p. URL: https://cds.cern.ch/record/732330 (on pages 67, 68).

[273] M. Pezzetti. "Control of large helium cryogenic systems: a case study on CERN LHC". In: *EPJ Techniques and Instrumentation* 8.1 (2021), p. 6. ISSN: 2195-7045. URL: https://doi.org/10.1140/epjti/s40485-021-00063-w (on page 67).

[274] J. Vollaire et al. *Linac4 design report*. Ed. by M. Vretenar. Vol. 6/2020. CERN Yellow Reports: Monographs. Geneva: CERN, Sept. 2020. ISBN: 978-92-9083-579-0, 978-92-9083-580-6 (on page 68).

[275] K. H. Reich. "The CERN Proton Synchrotron Booster". In: *IEEE Transactions on Nuclear Science* 16.3 (1969), pp. 959–961 (on page 68).

[276] D. Cundy and S. Gilardoni. "The Proton Synchrotron (PS): At the Core of the CERN Accelerators". In: *Adv. Ser. Direct. High Energy Phys.* 27 (2017), 39–85. 47 p. URL: https://cds.cern.ch/record/2312566 (on page 68).

[277] M. Benedikt et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. URL: https://cds.cern.ch/record/823808 (on page 68).

[278] *Beam Lifetime: Taking a Closer Look at LHC*. URL: https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.beam_lifetime#:~:text=Due\%20to\%20this\%2C\%20finally\%20the,LHC\%20is\%20about\%2010\%20hours. (on page 68).

[279] E. Mobs. "The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019". In: (2019). URL: https://cds.cern.ch/record/2684277 (on page 69).

[280] L Rossi. *Manufacturing and Testing of Accelerator Superconducting Magnets*. 2015. arXiv: 1501.07164 [physics.acc-ph] (on page 68).

[281] AC Team. "Diagram of an LHC dipole magnet. Schéma d'un aimant dipôle du LHC". 1999. URL: https://cds.cern.ch/record/40524 (on page 70).

[282] W. Herr and B Muratori. "Concept of luminosity". In: (2006). URL: https://cds.cern.ch/record/941318 (on page 71).

[283] ATLAS Collaboration. *Luminosity determination in pp collisions at* $\sqrt{s} =$ 13 TeV *using the ATLAS detector at the LHC*. Tech. rep. ATLAS-CONF-2019-021. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2677054 (on pages 72, 217).

[284] ATLAS Collaboration. "ATLAS data quality operations and performance for 2015-2018 data-taking". In: *JINST* 15 (2019), P04003. 43 p. arXiv: 1911.04632. URL: https://cds.cern.ch/record/2700249 (on page 73).

[285] ATLAS Collaboration. "Luminosity determination in pp collisions at $\sqrt{s} =$ 8 TeV using the ATLAS detector at the LHC". In: *The European Physical Journal C* 76.12 (2016). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-016-4466-1 (on page 74).

[286] ATLAS Collaboration. "Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC". In: *Phys. Rev. Lett.* 117.18 (2016), p. 182002. arXiv: 1606.02625 [hep-ex] (on pages 74, 91).

[287] CMS Collaboration. "Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV". In: *JHEP* 07 (2018), p. 161. arXiv: 1802.02613 [hep-ex] (on pages 74, 91).

[288] TOTEM Collaboration. "First measurement of elastic, inelastic and total cross-section at $\sqrt{s} = 13$ TeV by TOTEM and overview of cross-section data at LHC energies". In: *Eur. Phys. J. C* 79.2 (2019), p. 103. arXiv: 1712.06153 [hep-ex] (on pages 74, 91).

[289] Boyd, J.T. *LHC Run-2 and Future Prospects*. Tech. rep. 2020. arXiv: 2001.04370. URL: https://cds.cern.ch/record/2707815 (on page 75).

[290] ATLAS Collaboration. *ATLAS Luminosity Public Results from Run 2*. Geneva: ATLAS, 2021. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2 (on page 74).

[291] CMS Collaboration. *CMS Luminosity Public Results from Run 2*. Geneva: CMS, 2021. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults (on page 74).

[292] P. Berta et al. "Pileup and underlying event mitigation with iterative constituent subtraction". In: *Journal of High Energy Physics* 2019.8 (2019), p. 175. ISSN: 1029-8479. URL: https://doi.org/10.1007/JHEP08(2019)175 (on pages 74, 75).

[293] ATLAS Collaboration. "Studies of the performance of the ATLAS detector using cosmic-ray muons". In: *European Physical Journal C* 71 (Mar. 2011), p. 1593 (on page 76).

[294] G. Perez. "Unitarization Models For Vector Boson Scattering at the LHC". In: (Jan. 2018) (on page 78).

[295] Wikipedia contributors. *Pseudorapidity — Wikipedia, The Free Encyclopedia*. 2021. URL: https://en.wikipedia.org/w/index.php?title=Pseudorapidity\&oldid=1018871377 (on page 78).

[296] ATLAS Collaboration. *ATLAS inner detector: Technical Design Report, 1*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331063 (on pages 79, 82, 83, 269).

[297] S Haywood et al. *ATLAS inner detector: Technical Design Report, 2*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331064 (on pages 79, 82, 83, 269).

[298] ATLAS Collaboration. *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1996. URL: https://cds.cern.ch/record/331061 (on pages 79, 85, 87).

[299] M. C. Aleksa et al. *ATLAS Liquid Argon Calorimeter Phase-I Upgrade: Technical Design Report*. Tech. rep. 2013. URL: https://cds.cern.ch/record/1602230 (on pages 79, 85).

[300] ATLAS Collaboration. *ATLAS tile calorimeter: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1996. URL: https://cds.cern.ch/record/331062 (on pages 79, 85, 87).

[301] ATLAS Collaboration. *ATLAS muon spectrometer: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331068 (on pages 79, 89).

[302] ATLAS Collaboration. *ATLAS central solenoid: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331067 (on page 80).

[303] ATLAS Collaboration. "The ATLAS Inner Detector commissioning and calibration". In: *Eur. Phys. J. C* 70 (2010), pp. 787–821. arXiv: 1004.5293 [physics.ins-det] (on page 80).

[304] ATLAS Collaboration. "Experiment Briefing: Keeping the ATLAS Inner Detector in perfect alignment". 2020. URL: https://cds.cern.ch/record/2723878 (on page 81).

[305] M Capeans et al. *ATLAS Insertable B-Layer Technical Design Report*. Tech. rep. 2010. URL: https://cds.cern.ch/record/1291633 (on page 81).

[306] ATLAS Collaboration. *ATLAS Insertable B-Layer Technical Design Report Addendum*. Tech. rep. 2012. URL: https://cds.cern.ch/record/1451888 (on page 81).

[307] A. L. Rosa. *The ATLAS Insertable B-Layer: from construction to operation*. 2016. arXiv: 1610.01994 [physics.ins-det] (on page 81).

[308] F. Wang, B. Nachman, and M. Garcia-Sciveres. "Ultimate position resolution of pixel clusters with binary readout for particle tracking". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 899 (2018), pp. 10–15. ISSN: 0168-9002. URL: http://dx.doi.org/10.1016/j.nima.2018.04.053 (on page 81).

[309] Y. Takubo. *ATLAS IBL operational experience*. Tech. rep. ATL-INDET-PROC-2016-012. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2235541 (on pages 82, 270).

[310] ATLAS Collaboration. *ATLAS pixel detector: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1998. URL: https://cds.cern.ch/record/381263 (on page 82).

[311] ATLAS Collaboration. "Study of the material of the ATLAS inner detector for Run 2 of the LHC". In: *Journal of Instrumentation* 12.12 (2017), P12009–P12009. ISSN: 1748-0221. URL: http://dx.doi.org/10.1088/1748-0221/12/12/P12009 (on pages 82, 269).

[312] A Abdesselam. *ATLAS SCT Endcap Module Production*. Tech. rep. Geneva: CERN, 2006. URL: https://cds.cern.ch/record/973395 (on page 82).

[313] A. Abdelouahab et al. "Engineering for the ATLAS SemiConductor Tracker (SCT) end-cap". In: *Journal of Instrumentation* 3 (May 2008) (on page 82).

[314] X. Artru, G. B. Yodh, and G. Mennessier. "Practical theory of the multilayered transition radiation detector". In: *Phys. Rev. D* 12 (5 1975), pp. 1289–1306. URL: https://link.aps.org/doi/10.1103/PhysRevD.12.1289 (on page 83).

[315] U Egede. "The search for a Standard Model Higgs at the LHC and electron identification using transition radiation in the ATLAS tracker". 1998. URL: https://cds.cern.ch/record/344443 (on page 83).

[316] B. Dolgoshein. "Transition Radiation Detectors and Particle Identification". In: *Nucl. Instrum. Meth. A* 252 (1986), pp. 137–144 (on page 83).

[317] E. B. Klinkby. *Simulation of Transition Radiation and Electron Identification Ability of the ATLAS TRT*. Tech. rep. Geneva: CERN, 2011. URL: https://cds.cern.ch/record/1405697 (on page 83).

[318] ATLAS Collaboration. "Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties". In: *Journal of Instrumentation* 12.05 (2017), P05002–P05002. ISSN: 1748-0221. URL: http://dx.doi.org/10.1088/1748-0221/12/05/P05002 (on page 83).

[319] "Passage of Particles Through Matter". In: (2019) (on page 83).

[320] G. Stark. *The Search for Supersymmetry in Hadronic Final States Using Boosted Object Reconstruction*. URL: https://kratsg.github.io/thesis/ (on pages 85, 249).

[321] C. W. Fabjan and F Gianotti. "Calorimetry for Particle Physics". In: *Rev. Mod. Phys.* 75 (2003), 1243–1286. 96 p. URL: https://cds.cern.ch/record/692252 (on pages 85–87).

[322] F. Cavallari. "Performance of calorimeters at the LHC". In: *Journal of Physics: Conference Series* 293 (2011), p. 012001. URL: https://doi.org/10.1088/1742-6596/293/1/012001 (on pages 85, 87).

[323] A. P. Cheplakov et al. "Radiation hardness of GaAs preamplifiers for liquid argon calorimetry at LHC". In: (1996). URL: https://cds.cern.ch/record/324325 (on page 86).

[324] Y. Dai et al. "Electromagnetic Wave-Absorbing Properties of Steel Slag". In: *Journal of Materials Engineering and Performance* 28.1 (2019), pp. 535–542. ISSN: 1544-1024. URL: https://doi.org/10.1007/s11665-018-3831-7 (on page 86).

[325] T. G. Blackburn and M. Marklund. "Nonlinear Breit-Wheeler pair creation with bremsstrahlung γ rays". In: *Plasma Phys. Control. Fusion* 60.5 (2018), p. 054009. arXiv: 1802.06612 [physics.plasm-ph] (on page 86).

[326] L. M. de Andrade Filho et al. "Three-dimensional event visualization for the ATLAS calorimeter". In: *Computer Physics Communications* 183.2 (2012), pp. 245–250. ISSN: 0010-4655. URL: https://www.sciencedirect.com/science/article/pii/S0010465511003092 (on page 87).

[327] P. Adragna et al. "Testbeam studies of production modules of the ATLAS Tile Calorimeter". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 606.3 (2009), pp. 362–394. ISSN: 0168-9002. URL: https://www.sciencedirect.com/science/article/pii/S016890020900792X (on page 88).

[328] J. P. Badiou et al. *ATLAS barrel toroid: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331065 (on page 89).

[329] ATLAS Collaboration. *ATLAS end-cap toroids: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331066 (on page 89).

[330] E. Diehl. *Calibration and Performance of the ATLAS Muon Spectrometer*. 2011. arXiv: 1109.6933 [physics.ins-det] (on pages 89, 270).

[331] ATLAS Collaboration. "Resolution of the ATLAS muon spectrometer monitored drift tubes in LHC Run 2". In: *JINST* 14 (2019), P09011. 36 p. arXiv: 1906.12226. URL: https://cds.cern.ch/record/2680532 (on page 89).

[332] S. Aefsky. "Alignment of the Muon Spectrometer in ATLAS". In: *Physics Procedia* 37 (2012), pp. 51–56. ISSN: 1875-3892. URL: https://www.sciencedirect.com/science/article/pii/S1875389212016641 (on page 89).

[333] A. Aloisio et al. "The trigger chambers of the ATLAS muon spectrometer: production and tests". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 535.1 (2004), pp. 265–271. ISSN: 0168-9002. URL: https://www.sciencedirect.com/science/article/pii/S0168900204016511 (on page 90).

[334] ATLAS Outreach. "ATLAS Fact Sheet : To raise awareness of the ATLAS detector and collaboration on the LHC". 2010. URL: https://cds.cern.ch/record/1457044 (on page 91).

[335] M. zur Nedden. *The Run-2 ATLAS Trigger System: Design, Performance and Plan*. Tech. rep. ATL-DAQ-PROC-2016-039. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2238679 (on page 92).

[336] ATLAS Collaboration. "Operation of the ATLAS trigger system in Run 2". In: *Journal of Instrumentation* 15.10 (2020), P10004–P10004. ISSN: 1748-0221. URL: http://dx.doi.org/10.1088/1748-0221/15/10/P10004 (on page 92).

[337] ATLAS Collaboration. *ATLAS level-1 trigger: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1998. URL: https://cds.cern.ch/record/381429 (on page 92).

[338] ATLAS Collaboration. *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*. Tech. rep. 2013. URL: https://cds.cern.ch/record/1602235 (on page 92).

[339] S. Ask et al. "The ATLAS central level-1 trigger logic and TTC system". In: *JINST* 3 (2008), P08002 (on page 92).

[340] E. Simioni. "The Topological Processor for the future ATLAS Level-1 Trigger: from design to commissioning". In: (June 2014). arXiv: 1406.4316 [physics.ins-det] (on page 92).

[341] S. Artz et al. "The ATLAS Level-1 Muon Topological Trigger Information for Run 2 of the LHC". In: *JINST* 10.02 (2015), p. C02027 (on page 92).

[342] W Buttinger. *The ATLAS Level-1 Trigger System*. Tech. rep. ATL-DAQ-PROC-2012-024. Geneva: CERN, 2012. URL: https://cds.cern.ch/record/1456546 (on page 92).

[343] P. Jenni et al. *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 2003. URL: https://cds.cern.ch/record/616089 (on page 92).

[344] ATLAS Collaboration. "Performance of the ATLAS Trigger System in 2015". In: *Eur. Phys. J. C* 77.5 (2017), p. 317. arXiv: 1611.09661 [hep-ex] (on page 92).

[345] ATLAS Collaboration. *2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data*. Tech. rep. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2136007 (on page 93).

[346] ATLAS Collaboration. *Trigger Menu in 2016*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2242069 (on page 93).

[347] ATLAS Collaboration. *Performance of the ATLAS global transverse-momentum triggers at $\sqrt{s}$ = 8 TeV*. Tech. rep. Geneva: CERN, 2018. URL: https://cds.cern.ch/record/2311730 (on page 93).

[348] ATLAS Collaboration. *Trigger menu in 2018*. Tech. rep. ATL-DAQ-PUB-2019-001. Geneva: CERN, Oct. 2019. URL: https://cds.cern.ch/record/2693402 (on page 93).

[349] M. Cacciari and G. P. Salam. "Dispelling the N3 myth for the kt jet-finder". In: *Physics Letters B* 641.1 (2006), pp. 57–61. ISSN: 0370-2693. URL: https://www.sciencedirect.com/science/article/pii/S0370269306010094 (on pages 94, 142).

[350] M. Cacciari, G. P. Salam, and G. Soyez. "The anti-ktjet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (2008), pp. 063–063. ISSN: 1029-8479. URL: http://dx.doi.org/10.1088/1126-6708/2008/04/063 (on pages 94, 116, 121, 142).

[351] *Muon Trigger Public Results*. ATLAS (on page 95).

[352] F. Anulli et al. "The Level-1 Trigger Muon Barrel System of the ATLAS experiment at CERN". In: *JINST* 4 (2009), P04010 (on page 95).

[353] R. Brun and F. Rademakers. "ROOT — An object oriented data analysis framework". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389.1 (1997), pp. 81–86. ISSN: 0168-9002. URL: https://www.sciencedirect.com/science/article/pii/S016890029700048X (on page 96).

[354] ATLAS Collaboration. *ATLAS Computing: technical design report*. Technical design report. ATLAS. Geneva: CERN, 2005. URL: https://cds.cern.ch/record/837738 (on page 96).

[355] T Cornelissen et al. "The new ATLAS track reconstruction (NEWT)". In: *Journal of Physics: Conference Series* 119.3 (2008), p. 032014. URL: https://doi.org/10.1088/1742-6596/119/3/032014 (on pages 97, 108).

[356] ATLAS Collaboration. "Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2". In: *The European Physical Journal C* 77.10 (2017). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-017-5225-7 (on page 97).

[357] ATLAS Collaboration. "Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC". In: *The European Physical Journal C* 77.5 (2017). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-017-4887-5 (on pages 97, 99–101).

[358] ATLAS Collaboration. *Impact Parameter Resolution*. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2015-007/ (on page 98).

[359] ATLAS Collaboration. *Impact Parameter Resolutions in Di-Jet Events*. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2015-008/ (on page 98).

[360] T. Robertson and J. D. Cryer. "An Iterative Procedure for Estimating the Mode". In: *Journal of the American Statistical Association* 69.348 (1974), pp. 1012–1016. eprint: https://www.tandfonline.com/doi/pdf/10.1080/01621459.1974.10480246. URL: https://www.tandfonline.com/doi/abs/10.1080/01621459.1974.10480246 (on page 99).

[361] W. Waltenberger, R. Frühwirth, and P. Vanlaer. "Adaptive vertex fitting". In: *Journal of Physics G: Nuclear and Particle Physics* 34.12 (2007), N343–N356. URL: https://doi.org/10.1088/0954-3899/34/12/n01 (on page 99).

[362] ATLAS Collaboration. *Vertex Performance in 2018 data*. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2015-006/ (on page 101).

[363] P. Zyla et al. "Muon". In: (2020). URL: https://pdg.lbl.gov/2020/listings/rpp2020-list-muon.pdf (on page 101).

[364] S. Rettie. *Muon identification and performance in the ATLAS experiment*. Tech. rep. ATL-PHYS-PROC-2018-052. Geneva: CERN, June 2018. URL: https://cds.cern.ch/record/2626330 (on pages 102, 103).

[365] ATLAS Collaboration. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13\,\text{TeV}$". In: *The European Physical Journal C* 76.5 (May 2016). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-016-4120-y (on pages 103, 131, 145, 217, 274).

[366] A. Bell. *HEP Postgraduate Lecture Course*. URL: https://www.hep.ucl.ac.uk/postgrad/teaching/lhc/bTagging_2017.pdf (on page 105).

[367] ATLAS Collaboration. *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2273281 (on pages 105, 145).

[368] ATLAS Collaboration. "Performance of b-jet identification in the ATLAS experiment". In: *Journal of Instrumentation* 11.04 (2016), P04008–P04008. URL: https://doi.org/10.1088/1748-0221/11/04/p04008 (on page 106).

[369] ATLAS Collaboration. *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*. Tech. rep. ATLAS-CONF-2011-102. Geneva: CERN, 2011. URL: https://cds.cern.ch/record/1369219 (on page 107).

[370] ATLAS Collaboration. *Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13\,\text{TeV}$*. Tech. rep. ATL-PHYS-PUB-2015-018. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/2037683 (on page 108).

[371] ATLAS Collaboration. *Expected performance of the ATLAS experiment: detector, trigger and physics*. Geneva: CERN, 2009. URL: https://cds.cern.ch/record/1125884 (on page 108).

[372] ATLAS Collaboration. *Expected performance of the ATLAS b-tagging algorithms in Run-2*. Tech. rep. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/2037697 (on pages 108, 110, 111, 144).

[373] ATLAS Collaboration. *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*. Tech. rep. Geneva: CERN, 2018. URL: https://cds.cern.ch/record/2645405 (on page 109).

[374] A. Hocker et al. "TMVA - Toolkit for Multivariate Data Analysis". In: (Mar. 2007). arXiv: physics/0703039 (on page 110).

[375] ATLAS Collaboration. *Calibration of the ATLAS b-tagging algorithm in $t\bar{t}$ semi-leptonic events*. Tech. rep. Geneva: CERN, 2018. URL: https://cds.cern.ch/record/2638455 (on pages 113, 116, 144).

[376] ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *The European Physical Journal C* 77.7 (2017). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-017-5004-5 (on pages 114, 115, 120, 142).

[377] Y. Dokshitzer et al. "Better jet clustering algorithms". In: *Journal of High Energy Physics* 1997.08 (1997), pp. 001–001. ISSN: 1029-8479. URL: http://dx.doi.org/10.1088/1126-6708/1997/08/001 (on page 116).

[378] M. Wobisch and T. Wengler. "Hadronization corrections to jet cross-sections in deep inelastic scattering". In: *Workshop on Monte Carlo Generators for HERA Physics (Plenary Starting Meeting)*. Apr. 1998. arXiv: hep-ph/9907280 (on page 116).

[379] CMS Collaboration. *A Cambridge-Aachen (C-A) based Jet Algorithm for boosted top-jet tagging*. Tech. rep. Geneva: CERN, 2009. URL: https://cds.cern.ch/record/1194489 (on page 116).

[380] M. Cacciari, G. P. Salam, and G. Soyez. "FastJet user manual". In: *The European Physical Journal C* 72.3 (2012). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-012-1896-2 (on pages 116, 142).

[381] G. Soyez. *The SISCone and anti-kt jet algorithms*. 2008. arXiv: 0807.0021 [hep-ph] (on page 117).

[382] ATLAS Collaboration. *Variable Radius, Exclusive-k_T, and Center-of-Mass Subjet Reconstruction for Higgs($\to b\bar{b}$) Tagging in ATLAS*. Tech. rep. ATL-PHYS-PUB-2017-010. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2268678 (on pages 117, 118, 143).

[383] ATLAS Collaboration. "In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector". In: *The European Physical Journal C* 79.2 (2019). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-019-6632-8 (on pages 119, 120, 125, 126, 141, 216).

[384] A. Hrynevich. "ATLAS jet and missing energy reconstruction, calibration and performance in LHC Run-2". In: *JINST* 12.06 (2017). Ed. by L. Shekhtman, p. C06038 (on page 119).

[385] ATLAS Collaboration. "Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Physical Review D* 96.7 (2017). ISSN: 2470-0029. URL: http://dx.doi.org/10.1103/PhysRevD.96.072002 (on pages 119, 125).

[386] M Aharrouche et al. "Measurement of the response of the ATLAS liquid argon barrel calorimeter to electrons at the 2004 combined test-beam". In: *Nucl. Instrum. Methods Phys. Res., A* 614 (2010), 400–432. 33 p. URL: https://cds.cern.ch/record/1273537 (on page 119).

[387] M. Aharrouche et al. "Response uniformity of the ATLAS liquid argon electromagnetic calorimeter". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 582.2 (2007), pp. 429–455. ISSN: 0168-9002. URL: http://dx.doi.org/10.1016/j.nima.2007.08.157 (on page 119).

[388] M. Aharrouche et al. "Energy linearity and resolution of the ATLAS electromagnetic barrel calorimeter in an electron test-beam". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 568.2 (2006), pp. 601–623. ISSN: 0168-9002. URL: http://dx.doi.org/10.1016/j.nima.2006.07.053 (on page 119).

[389] ATLAS Collaboration. "Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data". In: *The European Physical Journal C* 72.3 (2012). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-012-1909-1 (on page 119).

[390] ATLAS Collaboration. "Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7\,\text{TeV}$". In: *The European Physical Journal C* 73.3 (2013). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-013-2304-2 (on page 120).

[391] D. Krohn, J. Thaler, and L.-T. Wang. "Jet trimming". In: *Journal of High Energy Physics* 2010.2 (2010). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP02(2010)084 (on pages 122, 141).

[392] ATLAS Collaboration. "Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s} = 7\,\text{TeV}$ using the ATLAS detector". In: *Journal of High Energy Physics* 2013.9 (2013). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP09(2013)076 (on pages 123, 126).

[393] ATLAS Collaboration. *Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector*. 2020. arXiv: 2007.02645 [hep-ex] (on page 123).

[394] R. Nayyar. "Measurement of the Jet Mass Scale and Resolution for Large Radius Jets at $\sqrt{s} = 8\,\text{TeV}$ using the ATLAS Detector". In: *PoS* ICHEP2016 (2017), p. 1130 (on page 123).

[395] ATLAS Collaboration. Tech. rep. (on pages 124, 125).

[396] ATLAS Collaboration. *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*. Tech. rep. ATLAS-CONF-2016-035. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2200211 (on pages 127, 129, 132, 201, 216).

[397] *xAODAnaHelpers Package*. URL: https://github.com/UCATLAS/xAODAnaHelpers (on pages 131, 149).

[398] J. Varela. *Validation of Higgs to $b\bar{b}$ tagging techniques via $Z \to b\bar{b}$*. URL: https://www.nevis.columbia.edu/reu/2019/talks/Varela_slides.pdf (on page 137).

[399] M. Cacciari and G. P. Salam. "Pileup subtraction using jet areas". In: *Physics Letters B* 659.1-2 (2008), pp. 119–126. ISSN: 0370-2693. URL: http://dx.doi.org/10.1016/j.physletb.2007.09.077 (on page 142).

[400] D. Krohn, J. Thaler, and L.-T. Wang. "Jets with variable R". In: *Journal of High Energy Physics* 2009.06 (2009), pp. 059–059. ISSN: 1029-8479. URL: http://dx.doi.org/10.1088/1126-6708/2009/06/059 (on page 143).

[401] ATLAS Flavour Tagging Group. *Tagger Calibration Recommendations for Release 21 TWiki*. URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/BTagCalibrationRecommendationsRelease21#Recommendations_for_variable_rad (on page 143).

[402] ATLAS Collaboration. "Identification of boosted Higgs bosons decaying into b-quark pairs with the ATLAS detector at 13 TeV". In: *Eur. Phys. J. C* 79.10 (2019), p. 836. arXiv: 1906.11005 [hep-ex] (on page 143).

[403] M. Battaglia et al. *Measurement of Inclusive Higgs Boson Production at High p_T^H in the $H \to b\bar{b}$ Decay Mode*. Tech. rep. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2703097 (on pages 145, 186, 205, 273, 308, 310).

[404] *RooBukinPdf Class Reference*. URL: https://root.cern/doc/v610/classRooBukinPdf.html (on page 151).

[405] ATLAS Collaboration. "Measurements of top-quark pair differential and double-differential cross-sections in the ℓ+jets channel with pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector". In: *Eur. Phys. J. C* 79 (2019), p. 1028. arXiv: 1908.07305 [hep-ex] (on pages 164, 165, 169, 221, 225, 229).

[406] Quality, Data. *ATLAS Data Quality Operations in 2012*. Tech. rep. ATL-DAPR-INT-2015-001. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/1982821 (on page 171).

[407] K. Hamilton, P. Nason, and G. Zanderighi. "MINLO: Multi-Scale Improved NLO". In: *JHEP* 10 (2012), p. 155. arXiv: 1206.3572 [hep-ph] (on pages 172, 174).

[408] J. M. Campbell et al. "NLO Higgs Boson Production Plus One and Two Jets Using the POWHEG BOX, MadGraph4 and MCFM". In: *JHEP* 07 (2012), p. 092. arXiv: 1202.5475 [hep-ph] (on pages 172, 174).

[409] K. Hamilton et al. "Merging H/W/Z + 0 and 1 jet at NLO with no merging scale: a path to parton shower + NNLO matching". In: *JHEP* 05 (2013), p. 082. arXiv: 1212.4504 [hep-ph] (on pages 172, 174, 237, 238).

[410] S. Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 06 (2010), p. 043. arXiv: 1002.2581 [hep-ph] (on pages 172–174).

[411] P. Nason. "A New method for combining NLO QCD with shower Monte Carlo algorithms". In: *JHEP* 11 (2004), p. 040. arXiv: hep-ph/0409146 (on pages 172–174).

[412] S. Frixione, P. Nason, and C. Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. arXiv: 0709.2092 [hep-ph] (on pages 172–174, 237).

[413] K. Hamilton, P. Nason, and G. Zanderighi. "Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator". In: *JHEP* 05 (2015), p. 140. arXiv: 1501.04637 [hep-ph] (on page 172).

[414] P. Nason and C. Oleari. "NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG". In: *JHEP* 02 (2010), p. 037. arXiv: 0911.5299 [hep-ph] (on pages 172, 174).

[415] H. B. Hartanto et al. “Higgs boson production in association with top quarks in the POWHEG BOX”. In: *Phys. Rev. D* 91.9 (2015), p. 094003. arXiv: 1501.04498 [hep-ph] (on pages 172, 174).

[416] G. Luisoni et al. “$HW^{\pm}/HZ + 0$ and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO”. In: *JHEP* 10 (2013), p. 083. arXiv: 1306.2542 [hep-ph] (on pages 172, 174, 237).

[417] G. Cullen et al. “Automated One-Loop Calculations with GoSam”. In: *Eur. Phys. J.* C72 (2012), p. 1889. arXiv: 1111.2034 [hep-ph] (on pages 172, 174).

[418] A. Denner et al. “Electroweak corrections to Higgs-strahlung off W/Z bosons at the Tevatron and the LHC with HAWK”. In: *JHEP* 03 (2012), p. 075. arXiv: 1112.5142 [hep-ph] (on pages 172, 174).

[419] A. Denner et al. “HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders”. In: *Comput. Phys. Commun.* 195 (2015), pp. 161–171. arXiv: 1412.5390 [hep-ph] (on pages 172, 174).

[420] K. Becker et al. “Precise predictions for boosted Higgs production”. In: (May 2020). arXiv: 2005.07762 [hep-ph] (on pages 172, 174, 287, 288).

[421] A. Djouadi, J. Kalinowski, and M. Spira. “HDECAY: A Program for Higgs boson decays in the Standard Model and its supersymmetric extension”. In: *Comput. Phys. Commun.* 108 (1998), p. 56. arXiv: hep-ph/9704448 (on page 172).

[422] M. Spira. “QCD effects in Higgs physics”. In: *Fortsch. Phys.* 46 (1998), pp. 203–284. arXiv: hep-ph/9705337 [hep-ph] (on page 172).

[423] A. Djouadi, M. M. Mühlleitner, and M. Spira. “Decays of supersymmetric particles: The Program SUSY-HIT (SUspect-SdecaY-Hdecay-InTerface)”. In: *Acta Phys. Polon. B* 38 (2007), pp. 635–644. arXiv: hep-ph/0609292 (on page 172).

[424] A. Bredenstein et al. “Radiative corrections to the semileptonic and hadronic Higgs-boson decays $H \to WW/ZZ \to 4$ fermions”. In: *JHEP* 02 (2007), p. 080. arXiv: hep-ph/0611234 (on page 172).

[425] A. Bredenstein et al. “Precise predictions for the Higgs-boson decay $H \to WW/ZZ \to 4$ leptons”. In: *Phys. Rev. D* 74 (2006), p. 013004. arXiv: hep-ph/0604011 (on page 172).

[426] A. Bredenstein et al. “Precision calculations for the Higgs decays $H \to ZZ/WW \to 4$ leptons”. In: *Nucl. Phys. Proc. Suppl.* 160 (2006), pp. 131–135. arXiv: hep-ph/0607060 [hep-ph] (on page 172).

[427] T. Gehrmann et al. “Jet cross sections and transverse momentum distributions with NNLOJET”. In: *PoS* RADCOR2017 (2018). Ed. by A. Hoang and C. Schneider, p. 074. arXiv: 1801.06415 [hep-ph] (on page 173).

[428] A. Gehrmann-De Ridder et al. “Precise QCD predictions for the production of a Z boson in association with a hadronic jet”. In: *Phys. Rev. Lett.* 117.2 (2016), p. 022001. arXiv: 1507.02850 [hep-ph] (on pages 173, 174).

[429] A. Gehrmann-De Ridder et al. "The NNLO QCD corrections to Z boson production at large transverse momentum". In: *JHEP* 07 (2016), p. 133. arXiv: 1605.04295 [hep-ph]. We thank Nigel Glover and Alexander Huss for providing $\sqrt{s} = 13$ TeV k-factors for the analysis fiducial volume. (On pages 173, 174).

[430] U. Blumenschein et al. "Pushing the precision frontier at the LHC with V+jets". In: Feb. 2018. arXiv: 1802.02100 [hep-ex] (on page 173).

[431] S. Frixione, P. Nason, and G. Ridolfi. "A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction". In: *JHEP* 09 (2007), p. 126. arXiv: 0707.3088 [hep-ph] (on pages 173, 174).

[432] E. Re. "Single-top Wt-channel production matched with parton showers using the POWHEG method". In: *Eur. Phys. J. C* 71 (2011), p. 1547. arXiv: 1009.2450 [hep-ph] (on pages 173, 174).

[433] R. Frederix, E. Re, and P. Torrielli. "Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO". In: *JHEP* 09 (2012), p. 130. arXiv: 1207.5391 [hep-ph] (on pages 173, 174).

[434] S. Alioli et al. "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions". In: *JHEP* 09 (2009), p. 111. arXiv: 0907.4076 [hep-ph] (on pages 173, 174).

[435] S. Frixione et al. "Single-top hadroproduction in association with a W boson". In: *JHEP* 07 (2008), p. 029. arXiv: 0805.3067 [hep-ph] (on page 173).

[436] T. Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), p. 159. arXiv: 1410.3012 [hep-ph] (on pages 173, 174).

[437] ATLAS Collaboration. "The ATLAS Simulation Infrastructure". In: *Eur. Phys. J. C* 70 (2010), p. 823. arXiv: 1005.4568 [physics.ins-det] (on page 173).

[438] S. Agostinelli et al. "GEANT4 – a simulation toolkit". In: *Nucl. Instrum. Meth. A* 506 (2003), p. 250 (on page 173).

[439] R. D. Ball et al. "Parton distributions with LHC data". In: *Nucl. Phys. B* 867 (2013), p. 244. arXiv: 1207.1303 [hep-ph] (on page 174).

[440] ATLAS Collaboration. *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie-Landshoff diffractive model*. Tech. rep. ATL-PHYS-PUB-2016-017. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2206965 (on pages 174, 285).

[441] D. J. Lange. "The EvtGen particle decay simulation package". In: *Nucl. Instrum. Meth. A* 462 (2001), p. 152 (on page 174).

[442] ATLAS Collaboration. *ATLAS simulation of boson plus jets processes in Run 2*. ATL-PHYS-PUB-2017-006. 2017. URL: https://cds.cern.ch/record/2261937 (on page 174).

[443] ATLAS Collaboration. *Studies on top-quark Monte Carlo modelling for Top2016*. ATL-PHYS-PUB-2016-020. 2016. URL: https://cds.cern.ch/record/2216168 (on page 174).

[444] ATLAS Collaboration. *Studies on top-quark Monte Carlo modelling with Sherpa and MadGraph5_aMC@NLO*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2261938 (on pages 174, 194).

[445] ATLAS Collaboration. *Improvements in $t\bar{t}$ modelling using NLO+PS Monte Carlo generators for Run 2*. ATL-PHYS-PUB-2018-009. 2018. URL: https://cds.cern.ch/record/2630327 (on page 174).

[446] ATLAS Collaboration. *Multijet simulation for* 13 TeV *ATLAS Analyses*. ATL-PHYS-PUB-2019-017. 2019. URL: https://cds.cern.ch/record/2672252 (on page 174).

[447] R. D. Ball et al. "Parton distributions for the LHC run II". In: *JHEP* 04 (2015), p. 040. arXiv: 1410.8849 [hep-ph] (on page 174).

[448] ATLAS Collaboration. "Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *JHEP* 09 (2014), p. 145. arXiv: 1406.3660 [hep-ex] (on page 174).

[449] T. Gleisberg et al. "Event generation with SHERPA 1.1". In: *JHEP* 02 (2009), p. 007. arXiv: 0811.4622 [hep-ph] (on page 174).

[450] F. Cascioli, P. Maierhöfer, and S. Pozzorini. "Scattering Amplitudes with Open Loops". In: *Phys. Rev. Lett.* 108 (2012), p. 111601. arXiv: 1111.5206 [hep-ph] (on page 174).

[451] T. Gleisberg and S. Höche. "Comix, a new matrix element generator". In: *JHEP* 12 (2008), p. 039. arXiv: 0808.3674 [hep-ph] (on page 174).

[452] S. Schumann and F. Krauss. "A parton shower algorithm based on Catani–Seymour dipole factorisation". In: *JHEP* 03 (2008), p. 038. arXiv: 0709.1027 [hep-ph] (on page 174).

[453] S. Höche et al. "QCD matrix elements + parton showers. The NLO case". In: *JHEP* 04 (2013), p. 027. arXiv: 1207.5030 [hep-ph] (on page 174).

[454] S. Catani et al. "Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO". In: *Phys. Rev. Lett.* 103 (2009), p. 082001. arXiv: 0903.2120 [hep-ph] (on page 174).

[455] S. Kallweit et al. "NLO electroweak automation and precise predictions for W+multijet production at the LHC". In: *JHEP* 04 (2015), p. 012. arXiv: 1412.5157 [hep-ph] (on page 174).

[456] B. Biedermann et al. "Automation of NLO QCD and EW corrections with Sherpa and Recola". In: *Eur. Phys. J. C* 77 (2017), p. 492. arXiv: 1704.05783 [hep-ph] (on page 174).

[457] ATLAS Collaboration. *ATLAS Pythia 8 tunes to* 7 TeV *data*. ATL-PHYS-PUB-2014-021. 2014. URL: https://cds.cern.ch/record/1966419 (on page 174).

[458] M. Czakon and A. Mitov. "Top++: A program for the calculation of the top-pair cross-section at hadron colliders". In: *Comput. Phys. Commun.* 185 (2014), p. 2930. arXiv: 1112.5675 [hep-ph] (on page 174).

[459] F. Buccioni et al. "OpenLoops 2". In: *Eur. Phys. J. C* 79.10 (2019), p. 866. arXiv: 1907.13071 [hep-ph] (on page 174).

[460] A. Denner, S. Dittmaier, and L. Hofer. "COLLIER: A fortran-based complex one-loop library in extended regularizations". In: *Comput. Phys. Commun.* 212 (2017), pp. 220–238. arXiv: 1604.06792 [hep-ph] (on page 174).

[461] I. Asimov and D. Shannon. *Franchise*. Creative Short Stories Series. Creative Education, 1989. ISBN: 9780886822323. URL: https://books.google.ca/books?id=WA_QPAAACAAJ (on page 178).

[462] G. Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *The European Physical Journal C* 71.2 (Feb. 2011). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-011-1554-0 (on pages 178, 213, 215).

[463] J. M. Lindert et al. "Precise predictions for V + jets dark matter backgrounds". In: *The European Physical Journal C* 77.12 (2017). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-017-5389-1 (on page 180).

[464] T. Sjöstrand. "Jet fragmentation of multiparton configurations in a string framework". In: *Nuclear Physics B* 248.2 (1984), pp. 469–502. ISSN: 0550-3213. URL: https://www.sciencedirect.com/science/article/pii/0550321384906072 (on page 180).

[465] ATLAS PmgTopProcesses Group. *PmgTopProcesses*. URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PmgTopProcesses#Systematic_uncertainties (on page 183).

[466] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *Annals Math. Statist.* 9.1 (1938), pp. 60–62 (on pages 186, 213).

[467] ATLAS Jet/Etmiss Group. *Large-R (R=1.0) jet energy scale uncertainties for full run 2 analyses (release 21)*. URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetUncertaintiesRel21Summer2019LargeR (on pages 190, 216).

[468] ATLAS Jet/Etmiss Group. *Talk pointed to on JMR recommendation TWiki to source of 20 percent relative JMR error*. URL: https://indico.cern.ch/event/441619/contributions/1099656/attachments/1148550/1647846/Jet_Substructure_Status_09_02_2015-Updated.pdf (on pages 190, 307).

[469] ATLAS Collaboration. *Measurement of the ATLAS Detector Jet Mass Response using Forward Folding with 80 fb^{-1} of $\sqrt{s}$ = 13 TeV pp data*. Tech. rep. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2724442 (on pages 190, 199, 203).

[470] ATLAS Jet/Etmiss Group. *Large-R ($R = 1.0$) jet energy scale uncertainties for full run 2 analyses (release 21) which points to the preliminary uncertainties for JMR given on this TWiki*. URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetUncertaintiesRel21Moriond2018LargeR (on pages 190, 216).

[471] *Definition of external constraints*. URL: https://root.cern.ch/doc/master/classRooAbsPdf.html#a536396deee60ae88762b306af45ec399 (on page 193).

[472] SLD Collaboration. "SLD DESIGN REPORT". In: (May 1984) (on page 204).

[473] P. C. Rowson, D. Su, and S. Willocq. "Highlights of the SLD physics program at the SLAC linear collider". In: *Ann. Rev. Nucl. Part. Sci.* 51 (2001), pp. 345–412. arXiv: hep-ph/0110168 (on page 204).

[474] SLD Collaboration. "Measurement of the charged multiplicities in b, c and light quark events from Z0 decays". In: *Phys. Lett. B* 386 (1996), pp. 475–485. arXiv: hep-ex/9608008 (on page 204).

[475] "New measurement of the charged particle multiplicity of weakly decay B hadrons with the DELPHI detector at LEP". In: (1997) (on page 204).

[476] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98.3 (2018), p. 030001 (on page 204).

[477] J. D. Bjorken. "Properties of hadron distributions in reactions containing very heavy quarks". In: *Phys. Rev. D* 17 (1 1978), pp. 171–173. URL: https://link.aps.org/doi/10.1103/PhysRevD.17.171 (on page 207).

[478] J. L. Rey Idler. "A study of the b-fragmentation function in $t\bar{t}$ production with the ATLAS experiment at the LHC". In: (2016). URL: https://cds.cern.ch/record/2209062 (on page 207).

[479] DELPHI Collaboration. "A study of the b-quark fragmentation function with the DELPHI detector at LEP I and an averaged distribution obtained at the Z Pole". In: *Eur. Phys. J. C* 71 (2011), p. 1557. arXiv: 1102.4748 [hep-ex] (on page 207).

[480] SLD Collaboration. "Measurement of the b quark fragmentation function in Z0 decays". In: *Phys. Rev. D* 65 (2002), p. 092006. arXiv: hep-ex/0202031 (on page 207).

[481] DELPHI Collaboration. "A study of the b-quark fragmentation function with the DELPHI detector at LEP I and an averaged distribution obtained at the Z Pole". In: *Eur. Phys. J. C* 71 (2011), p. 1557. arXiv: 1102.4748 [hep-ex] (on page 207).

[482] ATLAS Collaboration. *Measurements of b-jet moments sensitive to b-quark fragmentation in $t\bar{t}$ events at the LHC with the ATLAS detector*. ATLAS-CONF-2020-050. 2020. URL: http://cdsweb.cern.ch/record/2730444 (on page 207).

[483] L. Lyons. "Statistical Issues in Particle Physics". In: *Particle Physics Reference Library: Volume 2: Detectors for Particles and Radiation*. Ed. by C. W. Fabjan and H. Schopper. Cham: Springer International Publishing, 2020, pp. 645–692. ISBN: 978-3-030-35318-6. URL: https://doi.org/10.1007/978-3-030-35318-6_15 (on pages 211, 213).

[484] Cowan, G. "Statistics". In: (). URL: https://pdg.lbl.gov/2019/reviews/rpp2018-rev-statistics.pdf (on page 213).

[485] J. Ocariz. "Probability and Statistics for Particle Physicists". In: (2014), 253–280. 28 p. arXiv: 1405.3402. URL: https://cds.cern.ch/record/1701936 (on page 213).

[486] I. Narsky and F. Porter. *Statistical Analysis Techniques in Particle Physics: Fits, Density Estimation and Supervised Learning*. Nov. 2013, pp. 1–441. ISBN: 9783527410866 (on page 213).

[487] R. D. Cousins. "Lectures on Statistics in Theory: Prelude to Statistics in Practice". In: (July 2018). arXiv: 1807.05996 [physics.data-an] (on page 213).

[488] J. Fan, C. Zhang, and J. Zhang. "Generalized Likelihood Ratio Statistics And Wilks Phenomenon". In: *The Annals of Statistics* 29 (Aug. 2000) (on page 213).

[489] C. Gumpert et al. "Software for statistical data analysis used in Higgs searches". In: *J. Phys. Conf. Ser.* 490 (2014). Ed. by T. Christodoulakis, E. Vagenas, and D. Vlachos, p. 012229 (on page 213).

[490] G. Ranucci. "The Profile likelihood ratio and the look elsewhere effect in high energy physics". In: *Nucl. Instrum. Meth. A* 661 (2012), pp. 77–85. arXiv: 1201.4604 [physics.data-an] (on page 213).

[491] K. Al Khoury et al. *Measurement of $VH, H \to b\bar{b}$ with the ATLAS detector*. Tech. rep. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2692011 (on page 214).

[492] S. Hageboeck. *What the new RooFit can do for your analysis*. 2021. arXiv: 2012.02746 [physics.data-an] (on page 214).

[493] *RooBinSamplingPdf*. URL: https://root.cern.ch/doc/master/classRooBinSamplingPdf.html (on page 214).

[494] Andrea Sciandra. *Integration of Analytic PDFs in Binned Fits*. URL: https://indico.cern.ch/event/959448/contributions/4072603/attachments/2128398/3583863/AS_RooBinSamplingPdf_10_22_20.pdf (on page 214).

[495] W. Verkerke and D. Kirkby. *The RooFit toolkit for data modeling*. 2003. arXiv: physics/0306116 [physics.data-an] (on page 214).

[496] A. L. Read. "Presentation of search results: theCLstechnique". In: *Journal of Physics G: Nuclear and Particle Physics* 28.10 (Sept. 2002), pp. 2693–2704. URL: https://doi.org/10.1088/0954-3899/28/10/313 (on page 215).

[497] ATLAS $H \to b\bar{b}$ SubGroup. URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/Higgsbb (on page 216).

[498] Valerio Dao. *VH $(H \to b\bar{b})$ at ATLAS*. URL: https://indico.cern.ch/event/795581/contributions/3335194/attachments/1837111/3010715/CMS_FTAG_ValerioDao.pdf (on page 216).

[499] ATLAS Collaboration. "Jet mass and substructure of inclusive jets in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS experiment". In: *JHEP* 05 (2012), p. 128. arXiv: 1203.4606 [hep-ex] (on page 216).

[500] ATLAS Collaboration. *Measurement of b-tagging efficiency of c-jets in $t\bar{t}$ events using a likelihood approach with the ATLAS detector*. ATLAS-CONF-2018-001. 2018. URL: https://cds.cern.ch/record/2306649 (on page 217).

[501] ATLAS Collaboration. *Calibration of light-flavour b-jet mistagging rates using ATLAS proton–proton collision data at sqrts = 13 TeV*. ATLAS-CONF-2018-006. 2018. URL: https://cds.cern.ch/record/2314418 (on page 217).

[502] ATLAS Collaboration. *Simulation-based extrapolation of b-tagging calibrations towards high transverse momenta in the ATLAS experiment*. Tech. rep. ATL-PHYS-PUB-2021-003. Geneva: CERN, 2021. URL: https://cds.cern.ch/record/2753444 (on page 217).

[503] ATLAS Collaboration. "ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Eur. Phys. J. C* 79 (2019), p. 970. arXiv: 1907.05120 [hep-ex] (on page 217).

[504] G. Avoni et al. "The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS". In: *JINST* 13.07 (2018), P07017 (on page 217).

[505] R. Barlow and C. Beeston. "Fitting using finite Monte Carlo samples". In: *Computer Physics Communications* 77.2 (1993), pp. 219–228. ISSN: 0010-4655. URL: https://www.sciencedirect.com/science/article/pii/001046559390005W (on page 218).

[506] K. Becker et al. *Recommended predictions for the boosted-Higgs cross section.* Tech. rep. LHCHXSWG-2019-002. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2669113 (on pages 237, 238).

[507] ATLAS Collaboration. *Combined measurement of the total and differential cross sections in the $H \to \gamma\gamma$ and the $H \to ZZ^* \to 4\ell$ decay channels at $\sqrt{s}$ =13 TeV with the ATLAS detector.* Tech. rep. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2682844 (on pages 239, 240, 245).

[508] I. Z. Rothstein. "TASI lectures on effective field theories". In: Aug. 2003. arXiv: hep-ph/0308266 (on page 240).

[509] S. Weinberg. "Effective Field Theory, Past and Future". In: *PoS* CD09 (2009), p. 001. arXiv: 0908.1964 [hep-th] (on page 240).

[510] W. Buchmuller and D. Wyler. "Effective Lagrangian Analysis of New Interactions and Flavor Conservation". In: *Nucl. Phys. B* 268 (1986), pp. 621–653 (on page 241).

[511] C. J. C. Burges and H. J. Schnitzer. "Virtual Effects of Excited Quarks as Probes of a Possible New Hadronic Mass Scale". In: *Nucl. Phys. B* 228 (1983), pp. 464–500 (on page 241).

[512] C. N. Leung, S. T. Love, and S. Rao. "Low-Energy Manifestations of a New Interaction Scale: Operator Analysis". In: *Z. Phys. C* 31 (1986), p. 433 (on page 241).

[513] A. Pomarol and F. Riva. "Towards the Ultimate SM Fit to Close in on Higgs Physics". In: *JHEP* 01 (2014), p. 151. arXiv: 1308.2803 [hep-ph] (on page 241).

[514] A. Falkowski and F. Riva. "Model-independent precision constraints on dimension-6 operators". In: *JHEP* 02 (2015), p. 039. arXiv: 1411.0669 [hep-ph] (on page 241).

[515] J. Ellis, V. Sanz, and T. You. "The Effective Standard Model after LHC Run I". In: *JHEP* 03 (2015), p. 157. arXiv: 1410.7703 [hep-ph] (on page 241).

[516] B. Dumont, S. Fichet, and G. von Gersdorff. "A Bayesian view of the Higgs sector with higher dimensional operators". In: *JHEP* 07 (2013), p. 065. arXiv: 1304.3369 [hep-ph] (on page 241).

[517] A. Falkowski. "Effective field theory approach to LHC Higgs data". In: *Pramana* 87.3 (2016), p. 39. arXiv: 1505.00046 [hep-ph] (on page 241).

[518] A. Butter et al. "The Gauge-Higgs Legacy of the LHC Run I". In: *JHEP* 07 (2016), p. 152. arXiv: 1604.03105 [hep-ph] (on page 241).

[519] ATLAS Collaboration. *Constraints on an effective Lagrangian from the combined $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels using 36.1 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data collected with the ATLAS detector*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2293084 (on page 241).

[520] S. Willenbrock and C. Zhang. "Effective Field Theory Beyond the Standard Model". In: *Ann. Rev. Nucl. Part. Sci.* 64 (2014), pp. 83–100. arXiv: 1401.0470 [hep-ph] (on page 241).

[521] E. Masso. "An Effective Guide to Beyond the Standard Model Physics". In: *JHEP* 10 (2014), p. 128. arXiv: 1406.6376 [hep-ph] (on page 241).

[522] A. Pomarol. "Higgs Physics". In: *2014 European School of High-Energy Physics*. Dec. 2014. arXiv: 1412.4410 [hep-ph] (on page 241).

[523] B. Henning, X. Lu, and H. Murayama. "How to use the Standard Model effective field theory". In: *JHEP* 01 (2016), p. 023. arXiv: 1412.1837 [hep-ph] (on page 241).

[524] A. David and G. Passarino. "Through precision straits to next standard model heights". In: *Rev. Phys.* 1 (2016), pp. 13–28. arXiv: 1510.00414 [hep-ph] (on page 241).

[525] D. B. Franzosi and C. Zhang. *Probing top-quark chromomagnetic dipole moment at next-to-leading order in QCD*. 2015. arXiv: 1503.08841 [hep-ph] (on page 242).

[526] H. Mantler and M. Wiesemann. "Top- and bottom-mass effects in hadronic Higgs production at small transverse momenta through LO+NLL". In: *Eur. Phys. J. C* 73.6 (2013), p. 2467. arXiv: 1210.8263 [hep-ph] (on page 243).

[527] M. Grazzini and H. Sargsyan. "Heavy-quark mass effects in Higgs boson production at the LHC". In: *JHEP* 09 (2013), p. 129. arXiv: 1306.4581 [hep-ph] (on page 243).

[528] M. Grazzini, A. Ilnicka, and M. Spira. "Higgs boson production at large transverse momentum within the SMEFT: analytical results". In: *Eur. Phys. J.* C78.10 (2018), p. 808. arXiv: 1806.08832 [hep-ph] (on page 244).

[529] ATLAS Collaboration. "Higgs boson production cross-section measurements and their EFT interpretation in the 4ℓ decay channel at $\sqrt{s}$ =13 TeV with the ATLAS detector". In: *Eur. Phys. J. C* 80.10 (2020), p. 957. arXiv: 2004.03447 [hep-ex] (on page 244).

[530] CMS Collaboration. "Measurements of production cross sections of the Higgs boson in the four-lepton final state in proton–proton collisions at $\sqrt{s} =$ 13 TeV". In: *Eur. Phys. J. C* 81.6 (2021), p. 488. arXiv: 2103.04956 [hep-ex] (on page 244).

[531] ATLAS Collaboration. "Measurements of WH and ZH production in the $H \to b\bar{b}$ decay channel in pp collisions at 13 TeV with the ATLAS detector". In: *Eur. Phys. J. C* 81.2 (2021), p. 178. arXiv: 2007.02873 [hep-ex] (on page 244).

[532] CMS Collaboration. "Observation of Higgs boson decay to bottom quarks". In: *Phys. Rev. Lett.* 121.12 (2018), p. 121801. arXiv: 1808.08242 [hep-ex] (on page 244).

[533] A. G. et al. *High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1*. CERN Yellow Reports: Monographs. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2284929 (on page 246).

[534] I Béjar Alonso et al. *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*. Ed. by I Béjar Alonso. CERN Yellow Reports: Monographs. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2749422 (on page 246).

[535] HL-LHC Industry. *LHC/HL-LHC Plan*. URL: https://project-hl-lhc-industry.web.cern.ch/sites/project-hl-lhc-industry.web.cern.ch/files/inline-images/HL-LHC-plan-2021-1.pdf (on page 246).

[536] ATLAS Collaboration. *Projections for measurements of Higgs boson signal strengths and coupling parameters with the ATLAS detector at a HL-LHC*. Tech. rep. Geneva: CERN, 2014. URL: https://cds.cern.ch/record/1956710 (on page 246).

[537] ATLAS Collaboration. *ATLAS Phase-II Upgrade Scoping Document*. Tech. rep. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/2055248 (on page 246).

[538] ATLAS Collaboration. *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2285585 (on page 246).

[539] ATLAS Collaboration. *Technical Design Report: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade*. Tech. rep. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2719855 (on page 246).

[540] ATLAS Collaboration. *Technical Design Report for the ATLAS Inner Tracker Strip Detector*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2257755 (on page 246).

[541] ATLAS Collaboration. *ATLAS Liquid Argon Calorimeter Phase-II Upgrade: Technical Design Report*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2285582 (on page 246).

[542] ATLAS Collaboration. *Technical Design Report for the Phase-II Upgrade of the ATLAS Tile Calorimeter*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2285583 (on page 246).

[543] ATLAS Collaboration. *Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2285580 (on page 246).

[544] ATLAS Collaboration. *Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2285584 (on page 246).

[545] CMS Collaboration. *The Phase-2 Upgrade of the CMS Beam Radiation Instrumentation and Luminosity Detectors*. Tech. rep. Geneva: CERN, 2021. URL: https://cds.cern.ch/record/2759074 (on page 246).

[546] CMS Collaboration. *The Phase-2 Upgrade of the CMS Data Acquisition and High Level Trigger*. Tech. rep. Geneva: CERN, 2021. URL: https://cds.cern.ch/record/2759072 (on page 246).

[547] CMS Collaboration. *The Phase-2 Upgrade of the CMS Level-1 Trigger*. Tech. rep. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2714892 (on page 246).

[548] CMS Collaboration. *A MIP Timing Detector for the CMS Phase-2 Upgrade*. Tech. rep. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2667167 (on page 246).

[549] CMS Collaboration. *The Phase-2 Upgrade of the CMS Endcap Calorimeter*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2293646 (on page 246).

[550] CMS Collaboration. *The Phase-2 Upgrade of the CMS DAQ Interim Technical Design Report*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2283193 (on page 246).

[551] CMS Collaboration. *The Phase-2 Upgrade of the CMS L1 Trigger Interim Technical Design Report*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2283192 (on page 246).

[552] CMS Collaboration. *The Phase-2 Upgrade of the CMS Muon Detectors*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2283189 (on page 246).

[553] CMS Collaboration. *The Phase-2 Upgrade of the CMS Barrel Calorimeters*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2283187 (on page 246).

[554] CMS Collaboration. *The Phase-2 Upgrade of the CMS Tracker*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2272264 (on page 246).

[555] CMS Collaboration. *Technical proposal for a MIP timing detector in the CMS experiment Phase 2 upgrade*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2296612 (on page 246).

[556] CMS Collaboration. *The Phase-2 Upgrade of the CMS Beam Radiation, Instrumentation, and Luminosity Detectors: Conceptual Design*. Tech. rep. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2706512 (on page 246).

[557] L. Calligaris. *Status of the Phase-2 Tracker Upgrade of the CMS experiment at the HL-LHC*. Tech. rep. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2765457 (on page 246).

[558] S.-C. Hsu. "ATLAS ITK Upgrade Project". In: (2018). URL: https://cds.cern.ch/record/2302625 (on page 246).

[559] A. Blue et al. "Magnetic triggering — time-resolved characterisation of silicon strip modules in the presence of switching DC-DC converters". In: *Journal of Instrumentation* 16.06 (2021), P06012. URL: https://doi.org/10.1088/1748-0221/16/06/p06012 (on page 246).

[560] S. S. et al. "Measuring the border of the active area on silicon strip sensors". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 985 (2021), p. 164665. ISSN: 0168-9002. URL: http://www.sciencedirect.com/science/article/pii/S0168900220310627 (on page 246).

[561] L. Poley et al. “The ABC130 barrel module prototyping programme for the ATLAS strip tracker”. In: *Journal of Instrumentation* 15.09 (2020), P09004–P09004. ISSN: 1748-0221. URL: http://dx.doi.org/10.1088/1748-0221/15/09/P09004 (on page 246).

[562] L. P. et al. “Mapping the in-plane electric field inside irradiated diodes”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 980 (2020), p. 164509. ISSN: 0168-9002. URL: http://www.sciencedirect.com/science/article/pii/S0168900220309062 (on page 246).

[563] M. M. et al. “Electrical characterization of surface properties of the ATLAS17LS sensors after neutron, proton and gamma irradiation”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* (2020), p. 164456. ISSN: 0168-9002. URL: http://www.sciencedirect.com/science/article/pii/S0168900220308536 (on page 246).

[564] C. H. et al. “Strip sensor performance in prototype modules built for ATLAS ITk”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 978 (2020), p. 164402. ISSN: 0168-9002. URL: http://www.sciencedirect.com/science/article/pii/S0168900220307993 (on page 246).

[565] J. F.-T. et al. “Humidity sensitivity of large area silicon sensors: Study and implications”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 978 (2020), p. 164406. ISSN: 0168-9002. URL: http://www.sciencedirect.com/science/article/pii/S0168900220308032 (on page 246).

[566] C. H. et al. *Study of n-on-p sensors breakdown in presence of dielectrics placed on top surface*. Tech. rep. ATL-ITK-PROC-2018-011. Geneva: CERN, 2018. URL: https://cds.cern.ch/record/2305615 (on page 246).

[567] J. Ellis. “The Future of High-Energy Collider Physics”. In: *38th International Symposium on Physics in Collision*. Oct. 2018. arXiv: 1810.11263 [hep-ph] (on page 247).

[568] M. Cepeda et al. “Report from Working Group 2: Higgs Physics at the HL-LHC and HE-LHC”. In: *CERN Yellow Rep. Monogr.* 7 (2019). Ed. by A. Dainese et al., pp. 221–584. arXiv: 1902.00134 [hep-ph] (on page 247).

[569] W. C. Röntgen. “On a New Kind of Rays”. In: () (on page 249).

[570] A. Einstein. “Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt”. In: *Annalen der Physik* 322.6 (1905), pp. 132–148. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.19053220607. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19053220607 (on page 249).

[571] G. N. LEWIS. “The Conservation of Photons”. In: *Nature* 118.2981 (1926), pp. 874–875. ISSN: 1476-4687. URL: https://doi.org/10.1038/118874a0 (on page 249).

[572] H. J. Kimble, M. Dagenais, and L. Mandel. “Photon Antibunching in Resonance Fluorescence”. In: *Phys. Rev. Lett.* 39 (11 1977), pp. 691–695. URL: https://link.aps.org/doi/10.1103/PhysRevLett.39.691 (on page 249).

[573] J. T. M. F.R.S. "XL. Cathode Rays". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 44.269 (1897), pp. 293–316. eprint: https://doi.org/10.1080/14786449708621070. URL: https://doi.org/10.1080/14786449708621070 (on page 249).

[574] Professor Sir E. Rutherford F.R.S. "LIV. Collision of α particles with light atoms. IV. An anomalous effect in nitrogen". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 37.222 (1919), pp. 581–587. eprint: https://doi.org/10.1080/14786440608635919. URL: https://doi.org/10.1080/14786440608635919 (on page 249).

[575] J. Chadwick. "Possible Existence of a Neutron". In: *Nature* 129.3252 (1932), pp. 312–312. ISSN: 1476-4687. URL: https://doi.org/10.1038/129312a0 (on pages 249, 252).

[576] S. H. Neddermeyer and C. D. Anderson. "Note on the Nature of Cosmic-Ray Particles". In: *Phys. Rev.* 51 (10 1937), pp. 884–886. URL: https://link.aps.org/doi/10.1103/PhysRev.51.884 (on page 249).

[577] F. REINES and C. COWANjun. "The Neutrino". In: *Nature* 178.4531 (1956), pp. 446–449. ISSN: 1476-4687. URL: https://doi.org/10.1038/178446a0 (on page 249).

[578] G. Danby et al. "Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos". In: *Phys. Rev. Lett.* 9 (1 1962), pp. 36–44. URL: https://link.aps.org/doi/10.1103/PhysRevLett.9.36 (on page 249).

[579] M. Breidenbach et al. "Observed Behavior of Highly Inelastic Electron-Proton Scattering". In: *Phys. Rev. Lett.* 23 (16 1969), pp. 935–939. URL: https://link.aps.org/doi/10.1103/PhysRevLett.23.935 (on page 249).

[580] E. D. Bloom et al. "High-Energy Inelastic $e-p$ Scattering at 6° and 10°". In: *Phys. Rev. Lett.* 23 (16 1969), pp. 930–934. URL: https://link.aps.org/doi/10.1103/PhysRevLett.23.930 (on page 249).

[581] G. D. Rochester and C. C. Butler. "Evidence for the Existence of New Unstable Elementary Particles". In: *Nature* 160 (1947), pp. 855–857 (on page 249).

[582] J. J. Aubert et al. "Experimental Observation of a Heavy Particle J". In: *Phys. Rev. Lett.* 33 (23 1974), pp. 1404–1406. URL: https://link.aps.org/doi/10.1103/PhysRevLett.33.1404 (on page 249).

[583] J. E. Augustin et al. "Discovery of a Narrow Resonance in e^+e^- Annihilation". In: *Phys. Rev. Lett.* 33 (23 1974), pp. 1406–1408. URL: https://link.aps.org/doi/10.1103/PhysRevLett.33.1406 (on page 249).

[584] M. L. Perl et al. "Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation". In: *Phys. Rev. Lett.* 35 (22 1975), pp. 1489–1492. URL: https://link.aps.org/doi/10.1103/PhysRevLett.35.1489 (on page 249).

[585] D. P. Barber et al. "Discovery of Three-Jet Events and a Test of Quantum Chromodynamics at PETRA". In: *Phys. Rev. Lett.* 43 (12 1979), pp. 830–833. URL: https://link.aps.org/doi/10.1103/PhysRevLett.43.830 (on page 249).

[586] K. Kodama et al. “Observation of tau neutrino interactions”. In: *Physics Letters B* 504.3 (2001), pp. 218–224. ISSN: 0370-2693. URL: https://www.sciencedirect.com/science/article/pii/S0370269301003070 (on page 249).

[587] R. P. Feynman. “The principle of least action in quantum mechanics”. 1942. URL: https://cds.cern.ch/record/101498 (on page 250).

[588] W. Heisenberg. “Über den Bau der Atomkerne. I”. In: *Zeitschrift für Physik* 77.1 (1932), pp. 1–11. ISSN: 0044-3328. URL: https://doi.org/10.1007/BF01342433 (on page 252).

[589] E. Wigner. “On the Consequences of the Symmetry of the Nuclear Hamiltonian on the Spectroscopy of Nuclei”. In: *Phys. Rev.* 51 (2 1937), pp. 106–119. URL: https://link.aps.org/doi/10.1103/PhysRev.51.106 (on page 253).

[590] J. L. Rosner. *The Eightfold Way*. Tech. rep. hep-ph/0109241. EFI-2001-41-[CHICAGO]. 2001. URL: https://cds.cern.ch/record/519831 (on page 254).

[591] M Gell-Mann. “The Eightfold Way: A Theory Of Strong Interaction Symmetry”. In: (Mar. 1961). URL: https://www.osti.gov/biblio/4008239 (on page 254).

[592] M. Teklishyn. “Measurement of the c (1S) production cross-section via the decay c to proton-antiproton final state”. In: (Sept. 2014) (on page 254).

[593] M. Gell-Mann. “A schematic model of baryons and mesons”. In: *Physics Letters* 8.3 (1964), pp. 214 –215. ISSN: 0031-9163. URL: http://www.sciencedirect.com/science/article/pii/S0031916364920013 (on page 255).

[594] G Zweig. *An SU_3 model for strong interaction symmetry and its breaking; Version 1*. Tech. rep. CERN-TH-401. Geneva: CERN, 1964. URL: https://cds.cern.ch/record/352337 (on page 255).

[595] Stefan Kluth. “Tests of quantum chromo dynamics at e+e- colliders”. In: *Reports on Progress in Physics* 69.6 (2006), 1771–1846. URL: {https://doi.org/10.1088/0034-4885/69/6/r04} (on page 256).

[596] E. Fermi. “Versuch einer Theorie der β-Strahlen. I”. In: *Zeitschrift für Physik* 88.3 (1934), pp. 161–177. ISSN: 0044-3328. URL: https://doi.org/10.1007/BF01351864 (on page 257).

[597] Wikipedia contributors. *Fermi’s interaction — Wikipedia, The Free Encyclopedia*. 2020. URL: https://en.wikipedia.org/w/index.php?title=Fermi\%27s_interaction\&oldid=995802369 (on page 257).

[598] F. L. Wilson. “Fermi’s Theory of Beta Decay”. In: *American Journal of Physics* 36.12 (1968), pp. 1150–1160. URL: https://ui.adsabs.harvard.edu/abs/1968AmJPh..36.1150W (on page 257).

[599] C. S. Wu et al. “Experimental Test of Parity Conservation in Beta Decay”. In: *Phys. Rev.* 105 (4 1957), pp. 1413–1415. URL: https://link.aps.org/doi/10.1103/PhysRev.105.1413 (on page 257).

[600] C. N. Yang and R. L. Mills. “Conservation of Isotopic Spin and Isotopic Gauge Invariance”. In: *Phys. Rev.* 96 (1 1954), pp. 191–195. URL: https://link.aps.org/doi/10.1103/PhysRev.96.191 (on page 259).

[601] R. P. Feynman. "Relativistic Cut-Off for Quantum Electrodynamics". In: *Phys. Rev.* 74 (10 1948), pp. 1430–1438. URL: https://link.aps.org/doi/10.1103/PhysRev.74.1430 (on page 261).

[602] S. Tomonaga. "On a relativistically invariant formulation of the quantum theory of wave fields". In: *Prog. Theor. Phys.R* 1 (1946), pp. 27–42 (on page 261).

[603] M. F. Zoller. "Vacuum stability in the SM and the three-loop -function for the Higgs self-interaction". In: *What We Would Like LHC to Give Us* (2014). URL: http://dx.doi.org/10.1142/9789814603904_0034 (on page 266).

[604] M. Lüscher and P. Weisz. "Scaling laws and triviality bounds in the lattice 4 theory: (I). One-component model in the symmetric phase". In: *Nuclear Physics B* 290 (1987), pp. 25–60. ISSN: 0550-3213. URL: https://www.sciencedirect.com/science/article/pii/0550321387901775 (on page 266).

[605] M. Lüscher and P. Weisz. "Scaling laws and triviality bounds in the lattice 4 theory: (II). One-component model in the phase with spontaneous symmetry breaking". In: *Nuclear Physics B* 295.1 (1988), pp. 65–92. ISSN: 0550-3213. URL: https://www.sciencedirect.com/science/article/pii/0550321388902283 (on page 266).

[606] H. D. Politzer and S. Wolfram. "Bounds on particle masses in the Weinberg-Salam model". In: *Physics Letters B* 82.2 (1979), pp. 242–246. ISSN: 0370-2693. URL: https://www.sciencedirect.com/science/article/pii/0370269379907469 (on page 266).

[607] P. Q. Hung. "Vacuum Instability and New Constraints on Fermion Masses". In: *Phys. Rev. Lett.* 42 (14 1979), pp. 873–876. URL: https://link.aps.org/doi/10.1103/PhysRevLett.42.873 (on page 266).

[608] V. Branchina, E. Messina, and M. Sher. "Lifetime of the electroweak vacuum and sensitivity to Planck scale physics". In: *Phys. Rev. D* 91 (1 2015), p. 013003. URL: https://link.aps.org/doi/10.1103/PhysRevD.91.013003 (on page 266).

[609] F. Bezrukov et al. "Higgs boson mass and new physics". In: *Journal of High Energy Physics* 2012.10 (2012). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP10(2012)140 (on page 266).

[610] S. Alekhin, A. Djouadi, and S. Moch. "The top quark and Higgs boson masses and the stability of the electroweak vacuum". In: *Physics Letters B* 716.1 (2012), pp. 214–219. ISSN: 0370-2693. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.024 (on page 266).

[611] D. Buttazzo et al. "Investigating the near-criticality of the Higgs boson". In: *Journal of High Energy Physics* 2013.12 (2013). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP12(2013)089 (on page 266).

[612] J. Elias-Miró et al. "Higgs mass implications on the stability of the electroweak vacuum". In: *Physics Letters B* 709.3 (2012), pp. 222–228. ISSN: 0370-2693. URL: http://dx.doi.org/10.1016/j.physletb.2012.02.013 (on page 267).

[613] A. Hook et al. “Probable or improbable universe? Correlating electroweak vacuum instability with the scale of inflation”. In: *Journal of High Energy Physics* 2015.1 (2015). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP01(2015)061 (on page 267).

[614] J. Kearney, H. Yoo, and K. M. Zurek. “Is a Higgs vacuum instability fatal for high-scale inflation?” In: *Physical Review D* 91.12 (2015). ISSN: 1550-2368. URL: http://dx.doi.org/10.1103/PhysRevD.91.123537 (on page 267).

[615] G Aad et al. “ATLAS pixel detector electronics and sensors”. In: *Journal of Instrumentation* 3.07 (2008), P07007–P07007. URL: https://doi.org/10.1088/1748-0221/3/07/p07007 (on page 270).

[616] ATLAS IBL Collaboration. “Production and integration of the ATLAS Insertable B-Layer”. In: *Journal of Instrumentation* 13 (Mar. 2018) (on page 270).

[617] A Abdesselam and T Akimoto. *The Barrel Modules of the ATLAS SemiConductor Tracker*. Tech. rep. ATL-INDET-PUB-2006-005. ATL-COM-INDET-2006-009. CERN-ATL-COM-INDET-2006-009. Geneva: CERN, 2006. URL: https://cds.cern.ch/record/974073 (on page 270).

[618] J. Pequenao and P. Schaffner. “How ATLAS detects particles: diagram of particle paths in the detector”. 2013. URL: https://cds.cern.ch/record/1505342 (on page 272).

[619] ATLAS Collaboration. “Performance of the ATLAS muon triggers in Run 2”. In: *Journal of Instrumentation* 15.09 (2020), P09015–P09015. ISSN: 1748-0221. URL: http://dx.doi.org/10.1088/1748-0221/15/09/p09015 (on page 273).

[620] S. Agostinelli et al. “Geant4—a simulation toolkit”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. URL: https://www.sciencedirect.com/science/article/pii/S0168900203013688 (on page 285).

[621] T. Sjöstrand, S. Mrenna, and P. Skands. “A brief introduction to PYTHIA 8.1”. In: *Computer Physics Communications* 178.11 (2008), pp. 852–867. ISSN: 0010-4655. URL: http://dx.doi.org/10.1016/j.cpc.2008.01.036 (on page 285).

[622] J. M. Campbell et al. “NLO Higgs boson production plus one and two jets using the POWHEG BOX, MadGraph4 and MCFM”. In: (2012). eprint: arXiv:1202.5475 (on pages 286, 287).

[623] K. Hamilton, P. Nason, and G. Zanderighi. *Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator*. 2015. arXiv: 1501.04637 [hep-ph] (on page 286).

[624] T. Sjöstrand et al. “An introduction to PYTHIA 8.2”. In: *Computer Physics Communications* 191 (2015), pp. 159–177. ISSN: 0010-4655. URL: http://dx.doi.org/10.1016/j.cpc.2015.01.024 (on pages 286–289).

[625] J. Pumplin et al. “New Generation of Parton Distributions with Uncertainties from Global QCD Analysis”. In: *Journal of High Energy Physics* 2002.07 (2002), pp. 012–012. ISSN: 1029-8479. URL: http://dx.doi.org/10.1088/1126-6708/2002/07/012 (on pages 286, 287).

[626] D. J. Lange. “The EvtGen particle decay simulation package”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 462.1 (2001), pp. 152–155. ISSN: 0168-9002. URL: https://www.sciencedirect.com/science/article/pii/S0168900201000894 (on pages 286–289).

[627] P. Nason and C. Oleari. “NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG”. In: *Journal of High Energy Physics* 2010.2 (2010). ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP02(2010)037 (on pages 286, 289).

[628] H. B. Hartanto et al. “Higgs boson production in association with top quarks in the POWHEG BOX”. In: *Phys. Rev.* D91.9 (2015), p. 094003. arXiv: 1501.04498 [hep-ph] (on page 287).

[629] S. Carrazza, S. Forte, and J. Rojo. “Parton Distributions and Event Generators”. In: *Proceedings, 43rd International Symposium on Multiparticle Dynamics (ISMD 13)*. 2013, pp. 89–96. arXiv: 1311.5887 [hep-ph] (on pages 288, 289).

[630] E. Bothmann et al. “Event generation with Sherpa 2.2”. In: *SciPost Physics* 7.3 (2019). ISSN: 2542-4653. URL: http://dx.doi.org/10.21468/SciPostPhys.7.3.034 (on page 288).

[631] R. D. Ball et al. “Parton distributions for the LHC Run II”. In: *JHEP* 04 (2015), p. 040. arXiv: 1410.8849 [hep-ph] (on pages 288, 289).

[632] E. Re. “Single-top Wt-channel production matched with parton showers using the POWHEG method”. In: *The European Physical Journal C* 71.2 (2011). ISSN: 1434-6052. URL: http://dx.doi.org/10.1140/epjc/s10052-011-1547-z (on page 289).

[633] *XML Analytic Workspace Builder*. URL: https://gitlab.cern.ch/atlas-phys-exotics-dijetisr/xmlAnaWSBuilder (on page 307).