UNIVERSITY OF CALIFORNIA
SANTA CRUZ

NOVEL TECHNIQUES IN THE SEARCH FOR HIGGS BOSONS PRODUCED VIA VECTOR BOSON FUSION IN ASSOCIATION WITH A HIGH-ENERGY PHOTON AND DECAYED TO BOTTOM QUARKS

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

DOCTOR OF PHILOSOPHY

in

PHYSICS

by

Hava Schwartz

March 2025

The Dissertation of Hava Schwartz
is approved:

_____
Jason Nielsen, Chair

_____
Mike Hance

_____
Stefania Gori

_____
Peter Biehl
Vice Provost and Dean of Graduate Studies

Contents

Abstract

Novel Techniques in the Search for Higgs Bosons Produced via Vector Boson Fusion in Association with a High-Energy Photon and Decayed to Bottom Quarks

by

Hava Schwartz

A search for the Standard Model Higgs boson produced in association with a high-energy photon is performed using 133 fb^{-1} of pp collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector at the Large Hadron Collider at CERN. The $H + \gamma$ final state is particularly promising to study because the photon requirement reduces the multi-jet background, and the $b\bar{b}$ final state is the dominant decay mode of the Higgs boson. Event selection requirements isolate vector boson fusion Higgs production, the dominant production mode in this channel. Several improvements enhance the search sensitivity compared to previous measurements, including better background modeling and characterization, use of a dense neural network classifier, and an updated signal extraction strategy adopting a binned-likelihood fit directly to the classifier discriminant. These advancements result in a Higgs boson signal strength measured as 0.2 ± 0.7 relative to the Standard Model prediction. This corresponds to an observed significance of 0.3 standard deviations, compared to 1.5 standard deviations expected signal significance.

To my parents, Jason and Shari Schwartz, who were moved together by a world ripping apart and filled their three children with enough curiosity and confidence to create their own.

Acknowledgements

I, the author of this dissertation, hold immense gratitude for the people and fortuitous circumstances that supported its completion. Thank you to Professor Jason Nielsen for countless hours of unconditional bolstering. You always found the perfect balance of joining me in the trenches for difficult calculations and coding errors and taking a backseat, providing advice and allowing me to spread my wings. Thank you to my committee members, Professors Mike Hance and Stefania Gori, not only for barreling through this dense dissertation, but for your generous teachings which were instrumental in the journey toward its fruition. Thank you to my grad cohort — especially Rene, Gabby, Heather, Adi, Kevin, Adrian, and Jacob — and to the SCIPP ATLAS group — particularly Carolyn, Nathan, Giordon, Mike, Jenni, Sam, Dongyi, Alex, Kirsten, Tony, Vitaliy, and Marco — for a journey in solitude would have been devoid of joy and meaning.

Further gratitude must be extended to my family. Thank you to my parents for gifts that word cannot do justice. Your support in every one of my scholastic endeavors is cherished, and I hope that this achievement is considered an extension of your own. Thank you to my siblings, Johanna and Ethan, for direct guidance in these pursuits, as well as the indirect guidance in maintaining motivation through your admiration.

My time at CERN would have been incredibly dull if it were not for the immense care and wisdom from Vinicius, James, Julia, Andrea, Eleanora, Caterina, Elisabeth, Sergi, Elia, Kerstin, Martin, Tobias, Kristin, Emily, Emily, Albert, Alan, and Patricia. Thank you to my CERN visitors — Emmy, Eric, Nashelli, Lauren, Oksana, Mirco, Anastasiia, Dave, Shahil, Savannah, Holly, and Tyler — for allowing me to share my passion.

Thank you all for sourcing the light that I reflect when I shine. Thank you to those

who would celebrate this accomplishment with me if they could: Baubie, Uncle Mikee, Uncle Sam, Uncle Carl, Bryant, Roman, Noah, Rita.

Chapter 1.

The Beginning

There was always a before, until there wasn't. During the time with no before, there was no inertial mass, and all particles traveled at the speed of light. Bound structures like atoms had no chance to exist. Around 10^{-12} seconds after its beginning, this hot soup of a universe expanded and cooled enough to undergo a phase transition. Interactions arose between fundamental particles and the Brout-Englert-Higgs field, a scalar field with a non-zero vacuum expectation, which resulted in the appearance of mass for most of these particles. This mechanism created the framework that has allowed for the complexities of our known universe today.

For millennia, humans have acted as appendages of this universe seeking to understand itself. In this time, models of the shared physical reality have evolved to become quite encyclopedic. Modern particle physicists have developed and tested a specific description of the world at its most fundamental scale known as the Standard Model. Almost every attempt to examine the validity of this model and pin down its free parameters have supported its efficacy and granted consistent results. Nevertheless, there are still experimental observations which the Standard Model fails to explain, motivat-

ing further study. There is an abundance of matter over anti-matter, even though the Standard Model predicts equal rates of matter and anti-matter creation and no particle to break this symmetry. Observations of distant galaxies demonstrate objects that are bound by more gravitational force than can be accounted for from the masses of visible matter. Dark matter, or matter that does not interact with light, has no place in the Standard Model. Interacting with and responsible for the mass of nearly all fundamental particles, the Higgs boson is one-of-a-kind in its appeal as a tool for probing physics beyond the Standard Model.

This dissertation presents a search of proton-proton collisions collected by the ATLAS detector for instances in which a Higgs boson was produced in a very specific way and decayed into its dominant decay mode, bottom quarks. The Higgs is to be produced in association with a high-energy photon, along with two extra jets with a large angle between them. Vector boson fusion is the primary mode of Higgs production in our selected phase space. Out of more than 10 million billion events collected during the data taking period targeted by this search, simulation estimates 80 times that the Higgs boson was created and destroyed in these particular ways and could have been reconstructed by ATLAS identifiably in the restricted phase space. The measurement presented here reports 18 signal events, with an uncertainty of $+55/-51$ events. The reported count is consistent with the Standard Model expectation, and if many such experiments were performed, a lower measurement is predicted to occur in over 12% of them. In order to even approach this absurd task, huge strides had already been taken in object and event reconstruction and identification, invoking advancements in theoretical modeling, machine learning methods, and computing resources. This work demonstrates contributions to the most precise search to date for this Higgs signature, employing techniques

in event selection and signal extraction that are novel for this channel.

Twelve chapters divide this manuscript, including the introduction and conclusion. Before reaching the data analysis procedures and results which merit the award of an advanced degree, the reader will encounter explanation regarding details behind the experiment. Chapter 2 dives into the theory of the Standard Model, focusing on elements of the model's construction, composition, and its implications for the scalar boson of the Brout-Englert-Higgs field. In order to further probe this model, high energy particle interactions are necessary, in which matter can be transformed and energy transferred through the universe's fabric. Sourcing the high energy interactions of this experiment, the Large Hadron Collider, described in Chapter 3, at CERN in Geneva, Switzerland, smashes protons together at relativistic speeds. The protons are accelerated until they have more energy than 6,500 times their resting mass energy in two dense beams circling in opposite directions. These beams are crossed at the interaction point, causing collisions every 25 nanoseconds inside an enormous cylindrical detector called ATLAS. Chapter 4 outlines the construction and function of the ATLAS detector, along with details surrounding the efforts invested in its operation and future upgrades. My own contributions to the experiment operation and to prototype testing for the detector's pending upgrade are also highlighted here. The identity of this work is very much tied to the name ATLAS, referring to both the experiment's apparatus and the members of its scientific collaboration.

Chapters 5 through 11 are devoted to an analysis of data collected by ATLAS, seeking the Higgs boson in a highly selective corner of the phase space. Motivating this search, Chapter 5 illustrates why this specific Higgs signature is interesting among all the possible paths to study its many production and decay modes. This chapter also

further elaborates on the novel techniques which separate this endeavor from previous efforts in this channel. Once motivated, the Monte Carlo simulation methods used to predict the composition and behavior of the collected dataset are described in Chapter 6. Selection criteria on the presence and kinematics of various final state elements are imposed on both simulated and collected data, cutting down the massive datasets into a restricted phase space that is most likely to contain our signal while eliminating as many non-signal events as possible. Chapter 7 summarizes the process of reconstruction and identification of physics objects from readings in detector components, followed by a synopsis of the requirements and restrictions placed on these objects to qualify for event selection. The selected events are evaluated by a finely-tuned neural network, with particularities found in Chapter 8, which outputs a quantifier of each event's likeness to signal. The implementation of this machine learning strategy for signal and background separation is a cornerstone for the improvements achieved in this analysis and marks a large part of this work. Methods performed to check and improve the success of the Monte Carlo simulation in representing the data are enumerated in Chapter 9. Uncertainties related to the limitations in dataset size, the theoretical modeling, and the experimental detector resolution and effects are cataloged in Chapter 10. Finally, the results of this search are presented in Chapter 11, including the extracted signal strength and background normalization. Post-fit plots of Monte Carlo and data data distributions for the fit observable (the neural net output score) and for relevant event kinematics (like the di-jet invariant mass of the two Higgs decay candidates) are included.

Chapter 2.

Theory of the Standard Model and the Higgs Boson

The quantum nature of our universe can be conveniently described by a single model, the Standard Model, with 17 fundamental particles. Section 2.1 gives the mathematical foundation for the quantum field theory description of particles. Then, Section 2.2 formally presents the Standard Model, including its symmetries and its constituents. Finally, in Section 2.3, special attention is given to the Higgs boson, the star of the show in this work.

§ 2.1. Standard Model Construction and Background

2.1.1. Quantum Field Theory

Small stuff behaves strangely. As early as 1801, the strangeness of the quantum world was illuminated by the double-slit experiment performed by Thomas Young, in which the duality between wave-like and particle-like behavior of light became evident [1].

This experiment and later double-path experiments demonstrated that, between measurements at point A and point B, the smallest energy quanta were not restricted to a single trajectory between A and B, but could take part in multiple paths simultaneously. In classical theory, an object's trajectory is well determined within some spatial symmetries by maximizing the action, or the total difference between kinetic and potential energies along the complete path. However, in quantum theory, the object is not so limited, and the probability a quantum object can be found at a point B after an initial measurement at point A must take into account all the possible trajectories that could lead the object between points A and B.

The path integral formulation of quantum mechanics sums all the possible ways in which a final state could have evolved from an initial state, as each possibility contributes to the probability of such a result occurring [2]. The classical action, a scalar quantity describing the balance of kinetic and potential energies for a trajectory, $q(t)$, in a physical system, can be seen more formally written in Equation 2.1.1.

$$S[q] = \int_{t_i}^{t_f} (\frac{1}{2}\dot{q}^2 - V(q))dt \tag{2.1.1}$$

While extremizing this quantity results in a particular trajectory known as the classical trajectory, the value of the action over a particular phase space path controls the weight with which that trajectory contributes to the final state probabilities. Namely, the probability amplitude, M in Equation 2.1.2, for a quantum object to be at q_f at time t_f given that it was observed at q_i at time t_i can be calculated by summing an action-associated weight over all possible paths that could bring the object between points q_i and q_f at

times t_i and t_f, known as the path integral, I in Equation 2.1.3.

$$M = \langle q_f | e^{-i\hat{H}(t_f - t_i)/\hbar} | q_i \rangle \tag{2.1.2}$$

$$I = \int [dq] e^{iS[q]/\hbar} \tag{2.1.3}$$

The probability amplitude in Equation 2.1.2 will later be referred to as the matrix element for a specified process, and further details on its calculation can be found in Appendix A.

This technique can be generalized even further by considering its application to fields, in which a physical quantity is assigned or attached to each point in space and time. In field theory, the operators $\hat{q}$ and $\hat{p}$ become $\hat{\phi}(\mathbf{x})$ and $\hat{\pi}(\mathbf{x})$, and the classical action for a D-dimensional space becomes Equation 2.1.4.

$$S = \int dt d^D x \left(\frac{1}{2}\dot{\phi}^2 - \frac{1}{2}(\nabla\phi)^2 - V(\phi)\right) \tag{2.1.4}$$

Quantum field theory (QFT) can then be defined through the path integral over fields $\phi(\mathbf{x}, t)$,

$$\int [d\phi] e^{iS[\phi(\mathbf{x},t)]},$$

with $\hbar = 1$ as is common convention. However, for this statement to be meaningful in the continuum limit where the distance between points in space reduces to 0, several properties [3] of the action must be satisfied.

- The action must be local, i.e. it can be written as $S = \int dt d^D x \, \mathcal{L}(\phi, \dot{\phi}, \nabla\phi)$, where the Lagrangian density $\mathcal{L}$ depends on the local value of a field and its derivatives.
- For this relativistic theory, the Lagrangian density must be Lorentz invariant.

- $\mathcal{L}$ must also be invariant under any other symmetries of the theory.
- The theory must also be renormalizable, such that all observables can be described in finite terms with the introduction of a few parameters.

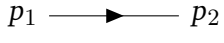
This last point is discussed a bit further in Section A.1. In this transition to the discussion of fields in the continuum limit, it is now fair to refer to these quantum objects as particles, excitations or ripples in the field. Skipping a bit of mathematical derivation, for a free non-interacting theory which satisfies the above requirements, the amplitude for a particle with mass m to propagate from x_1 to x_2 is given by Equation 2.1.5.

$$\langle\hat{\phi}(x_2,t_2)|\hat{\phi}(x_1,t_1)\rangle = (2\pi)^D\delta^D(p_1-p_2)\frac{e^{-i\sqrt{p_2^2+m^2}(t_1-t_2)}}{2\sqrt{p_2^2+m^2}} \tag{2.1.5}$$

This quantity is known as the propagator and is quite important in calculations of particle interactions [4]. However, it is also quite boring, as it implies that a freely moving particle without any interactions will not change momentum, let alone other quantum properties.

2.1.2. Feynman Diagrams

In order to move forward into interacting theories, it is compelling to introduce a visual method of organizing calculations of particle interactions called Feynman diagrams. For a non-interacting particle moving freely, the generic propagator written explicitly in Equation 2.1.5 would be depicted in a Feynman diagram as a straight line.



Particles can be categorized into two types according to the statistics obeyed when mixing indistinguishable elements. The probability density of the combined state must be constant under commutations of the individual states, which can happen in only two ways. For the combination of two indistinguishable particles ψ_a and ψ_b, there can be a symmetric combination, $\frac{1}{\sqrt{2}}(\psi_a(x_1)\psi_b(x_2) + \psi_b(x_1)\psi_a(x_2))$, or an antisymmetric combination, $\frac{1}{\sqrt{2}}(\psi_a(x_1)\psi_b(x_2) - \psi_b(x_1)\psi_a(x_2))$, with the symmetry referring to how the combined state behaves under particle commutation. Particles that combine antisymmetrically obey Fermi-Dirac statistics and are called fermions, and particles that combine symmetrically obey Bose-Einstein statistics and are known as bosons. The only experimentally observed fermionic field is the Dirac field, and it describes spin-$\frac{1}{2}$ particles, which are represented with solid lines in Feyman diagrams, . Three types of bosons appear in most basic QFT models: scalar bosons which have spin-0 and are represented with dashed lines , spin-1 vector bosons with no color interaction are represented with oscillating lines , and colored spin-1 vector bosons which appear as curled lines .

With these elements of construction defined, it is possible to look at interaction theories. A point of interaction is depicted in a Feynman diagram as a vertex with typically three (though sometimes four) intersecting lines, which can be either two fermions and a boson or three bosons as shown in Figure 2.1. Across any vertex, there is conservation of energy, momentum, angular momentum, spin, charge, and more [5].

The most simple and most elaborate particle interactions and decays can be represented through combinations of these vertices. A Feynman diagram is constructed with time running as a horizontal axis to the right in this dissertation, though time is sometimes represented as a vertical axis. Fermions often receive an arrow that points forward

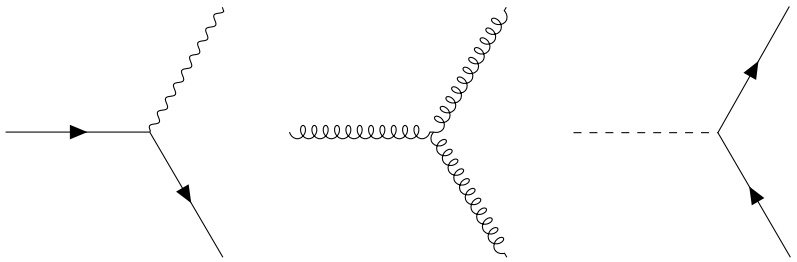


Figure 2.1.: Examples of allowed Feynman diagram vertices. With time evolving from left to right, vertex (a) shows a fermion radiating a vector boson (a concrete example of this would be bremsstrahlung radiation where a photon is radiated from an electron), vertex (b) shows an interaction of 3 colored vector bosons, and vertex (c) shows the decay of a scalar boson into fermions. This is certainly not an exhaustive list of allowed vertices.

in time for matter and backwards in time for anti-matter. Vertices are appended in such a way that the chronology of events for a given particle can be traced from left to right. For example, the Feynman diagram depicted in Figure 2.2 shows an initial state of a fermion and an anti-fermion, which annihilate and produce a vector boson, which then decays into a final state of a fermion and an anti-fermion. Several conservation laws constrain the allowed particles and kinematics in such an interaction, e.g. if the initial state fermions are of equal and opposite charge, then the vector boson must be neutral [6]. This depiction reflects the path integral formulation, in which all the possible ways in which such an interaction can happen contribute to its probability. For instance, if the final state kinematics are not specified, it is implied in this diagram that all possible fermion momenta which preserve conservation of momentum be summed (with weights related to the associated action). [7]

The Feynman diagram shown in Figure 2.2 is only Born-level, meaning it has the minimum number of vertices necessary to make such an interaction possible. However,

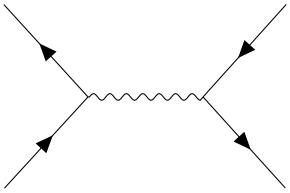


Figure 2.2.: Basic example of a tree-level Feynman diagram. With time oriented from left to right, a fermion and anti-fermion annihilate and produce a vector boson, which then decays into a final state fermion and anti-fermion.

the path integral formulation of QFT does not just imply that all momenta and trajectories should be considered, but also that all ways in which particles can interact without changing the initial and final observed states also contribute to the probability of the transformation. The intial and final states appear in Feynman diagrams as open legs that are not connecting two vertices, and these are the only measurable particles. Infinitely many loops can be added, which would not impact the observation of the initial or final states, but would represent additional ways in which the interaction can happen. With each additional loop in a Feynman diagram, a factor of the associated coupling strength squared comes along. If the coupling strength is small, this additional factor suppresses the contribution to the probability amplitude as the number of vertices increases [8]. Calculations reliant on the summation of these infinitely many paths can only be approximated, using perturbations in powers of coupling strength (though this perturbative method is not possible for high energy quantum chromodynamic, or QCD, interactions where α_s is large). In Feynman diagrams, this perturbation theory manifests as a correspondence between calculation orders and number of allowed loops or extra radiated legs in the accounted diagrams. The Born-level or tree-level diagram is also known as the leading-order diagram, and with the addition of each vertex beyond the

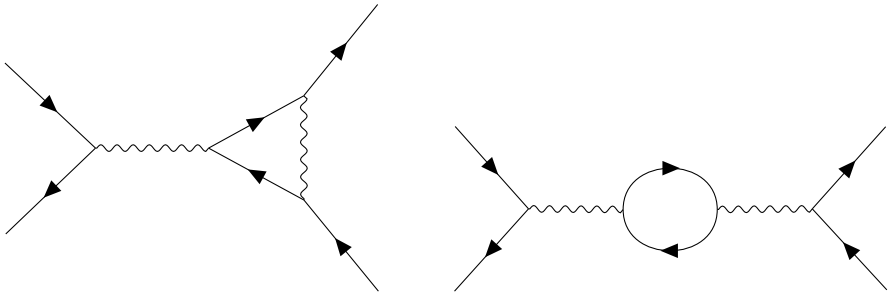


Figure 2.3.: A couple of examples of next-to-leading order (NLO) Feynman diagrams for the interaction represented at tree-level in Figure 2.2. Here, additional intermediate particles are introduced which are not part of the initial or final measurable states, and therefore represent a possibility in which the particle transformation could occur. With each additional vertex, the contribution to the overall probability amplitude are suppressed by powers of the coupling strength.

tree-level minimum, the diagrams are known as next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) and so on. A couple examples of NLO diagrams for the same transformation are shown in Figure 2.3. Further discussion of the calculation of process probabilities through the combination of Feynman diagrams is provided in Appendix A.

§ 2.2. Standard Model

To this point, the discussion of particle interactions has been left quite general and could be applied to a reality different than our own. Thanks to many theorists and experimentalists over the last century, it is possible to describe the physics of our known universe in terms of the interactions of specific particles. The Standard Model of particle physics is a quantum description of three of the known fundamental forces* in our universe: the

*A quantum description of gravity has not yet been devised successfully.

strong force which binds atomic nuclei, the weak force which mediates nuclear decays, and the electromagnetic force.

2.2.1. Fields and Gauge Symmetries

All particles are also described as excitations of fields, relating back to the wave-particle duality of nature demonstrated in the double-slit experiment. The values of these fields, however, are not measurable quantities, but instead can be inferred from the measurable aspects of these excitations, such as charge, energy, mass, etc. If we take electromagnetism as the most familiar example, the associated field is known as the 4-vector potential, A, which cannot be directly measured. Nonetheless, A can be inferred from the physical electric and magnetic fields. Though, it turns out, there are infinitely many vector potentials that could provide identical electric and magnetic fields, and from a given vector potential A, there is a set of operations that can be performed to arrive at another vector potential $\tilde{A}$ which produces the identical fields. This set of operations can be thought of as a symmetry group [9], since the measurable quantities of the electric and magnetic fields are invariant under these operations. In the case of electromagnetism, these such operations are infinite in number and can be matched one-to-one to the symmetries of a circle (e.g. rotations of the circle by any angle or reflections across any axis). This symmetry group is called $U(1)$, and it describes, or in some senses defines, electromagnetism. Physically, these symmetric operations can be thought of as shifts in the phase of the electric and magnetic waves. In the double-slit experiment, the observed interference pattern peaks in areas where the interfering waves are in phase with one another and dips where they are out of relative phase. An overall shift in the phases of both propagating fields would not change the position of the peaks and dips in the

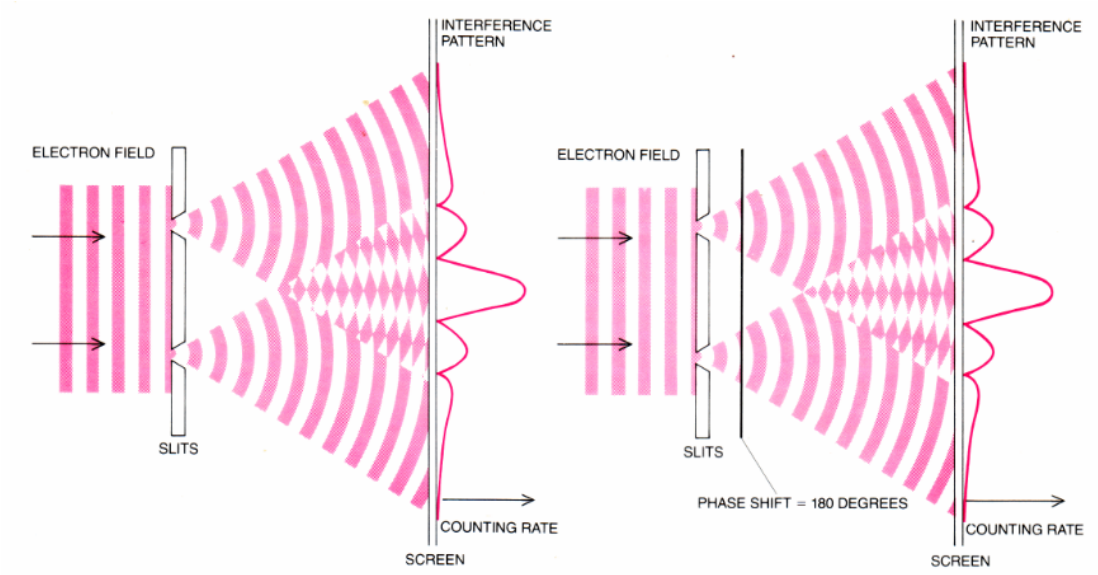


Figure 2.4.: Representation of the $U(1)$ global symmetry of electromagnetism. With rotations or reflections of the phase of both the propagating fields in this representation of the double-slit experiment, the resulting interference pattern is unchanged. [10]

resulting pattern, as depicted in Figure 2.4, representing a global $U(1)$ symmetry.

However, if the phase of only one of the two propagating waves is shifted, the resulting relative phase and interference pattern will indeed be modified. The same effect can be achieved by introducing a magnetic field between the two slits, with a result indistinguishable from the local phase transformation. Thus, the interaction of a field with matter through forces is actually a necessary component of maintaining such symmetries locally, independently at all points in space and time [10]. When the interactions between fields and matter are taken into account, a gauge theory is born. Gauge bosons are responsible for mediating these interactions of the gauge fields with matter. For electromagnetism, this gauge boson is the photon described in Section 2.2.2.

The symmetries of the weak and strong interactions are a bit more complex. Two

of the three gauge bosons of the weak force, W^+ and W^-, carry charge, implying some relationship with the electromagnetic force. In fact, these two forces can be unified into the electroweak force, which has the symmetry group $SU(2)_L \times U(1)_Y$. The group $SU(2)$, or the special unitary group of dimension 2, is the set of 2×2 matrices with determinant equal to 1, which can transform a complex dimension-2 vector without changing its magnitude [10]. This symmetry group has 3 generators, with a conventional form known as the Pauli matrices shown in Equation 2.2.1.

$$\sigma_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_z = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{2.2.1}$$

Together with the one generator for $U(1)$, this number of generators explains why there are collectively 4 gauge bosons for the electroweak force [10]. It is to be noted, however, that the four bosons that are known and loved ($W^\pm, Z, \gamma$) are not directly the generators of this symmetry group, but are instead representations in a different basis, the mass eigenbasis. The direct generators of the $SU(2)_L$ weak isospin group are represented as $W^{(1)}, W^{(2)}, W^{(3)}$, and the generator of the $U(1)_Y$ hypercharge group is B. More on how the mass eigenbasis relates to the symmetry generator basis will come in Section 2.3.

The color interactions of the strong force are a bit more complicated. For the strong force, the involved symmetry is $SU(3)$, which has 8 generators corresponding to the 8 types of gauge bosons, the gluons [11]. Together, these three fundamental forces can be unified into the symmetry group $SU(3) \times SU(2) \times U(1)$. However, this symmetry group misses a few key elements of our observed reality. Namely, it has no explanation of the distinction between the electromagnetic and weak forces, and it does not explain why the weak gauge bosons are massive unlike the photon and gluons. These discrepancies

motivate something known as electroweak symmetry breaking, and this is the origin of the Higgs boson, which is the only boson from Section 2.2.2 that is not directly motivated by the gauge symmetries here.

2.2.2. Particles of the Standard Model

There are 18 fundamental particles in the Standard Model, all of which have now been observed in experiments. These particles, which we define by their quantum numbers (i.e. spin, charge, mass, etc.), are shown in Figure 2.5.

Bosons

The bosons, characterized by having integer spin, act as force carriers and mediate the various interaction types. There is only one scalar (spin-0) boson, known as the Higgs boson, which takes part in electroweak interactions and is discussed further in Section 2.3. The remaining force carriers are vector (spin-1) bosons. The photon a.k.a. the quantum of light, the most well-known boson, mediates the electro-magnetic force. The massive vector bosons — W^+, W^-, and Z — are known as the weak bosons for their mediation of the weak force. Finally, the gluons, of which there are 8 unique eigenstates as described in Section 2.2.1, mediate the strong force and allow color interactions. Practically, each gluon carries a mixture of a color and an anti-color.

Fermions

The fermions, all with half-integer spin and with subdivisions of quarks and leptons, are the constituents of matter. Both subsets of fermions come in three generations, with different flavors and masses. Leptons, of which the most well-known is the electron,

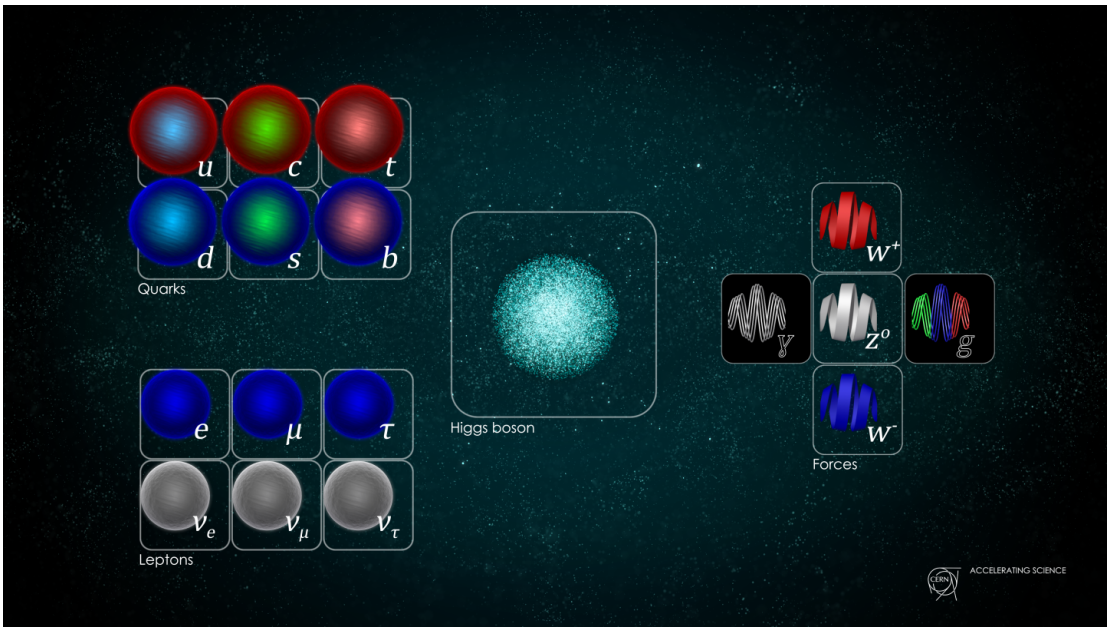


Figure 2.5.: Particles of the Standard Model, divided into classifications of color-interacting fermions (quarks), non-color-interacting fermions (leptons), and force carriers (bosons). [12]

participate in only electroweak interactions. The three generations of charged leptons are the electron (e), muon (μ), and tau (τ), along with their anti-matter counterparts, the positron ($\bar{e}$), anti-muon ($\bar{\mu}$), and anti-tau ($\bar{\tau}$). Each has unit charge, with the anti-matter components having positive charge. The neutral leptons, called neutrinos, come in association with the three generations of charged leptons, in an $SU(2)$ doublet. Mirroring the flavors of their charged partners, the neutral leptons are known as electron neutrinos (ν_e), muon neutrinos (ν_μ), and tau neutrinos (ν_τ). Though once thought to be massless, at least two neutrinos indeed have very small mass [13], orders of magnitude less than the electron, and this disparity from the other Standard Model particles is one of the model's current puzzles.

The quarks, while participating in electroweak interactions, also carry color and interact via the strong force. There are three generations of positively-charged quarks each with charge $+\frac{2}{3}$, the up (u), charm (c), and top (t), listed in order of increasing mass. Their partners, the three generations of quarks with charge $-\frac{1}{3}$, are the down (d), strange (s), and bottom (b). There are also anti-quarks for each of these six quarks, which are opposite in charge and operate in the anti-color space. The ‘‘charges" (color) of the strong force, which follows $SU(3)$ symmetry, come in three varieties, which get described as red, green, and blue, along with the anti-colors anti-red, anti-green, and anti-blue.

Standard Model Lagrangian

As mentioned in Section 2.1.1, the dynamics of a system can be described with just one equation, the Lagrangian density $\mathcal{L}$. Taking into account the three fundamental forces with quantum descriptions and the symmetry breaking by the Higgs, everything that we

know about particle physics can be summarized in Equation 2.2.2.

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}\not{D}\psi + \bar{\psi}_i y_{ij}\psi_j\phi + h.c. + |D_\mu\phi|^2 - V(\phi) \tag{2.2.2}$$

Of course, this is an extremely condensed version, for the same equation could just as easily take the full page to write. Each term of this Lagrangian corresponds to a set of interactions, or allowed vertices of the Standard Model. To make sense of $\mathcal{L}$, let's break it down term by term, and relate it back to the discussion of the previous two sections.

The first term, $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, describes the existence and interactions of all the gauge bosons. The second term, $i\bar{\psi}\not{D}\psi$, describes the interactions between the gauge bosons and matter (fermions). The fermionic quark and lepton fields are represented with ψ, while the covariant derivative $\not{D}$ contains the gauge bosons wrapped inside. The majority of our matter interactions are wrapped up in this second term, and it manifests in the Feynman diagram as a vertex with two fermions and a gauge boson. The third term, $\bar{\psi}_i y_{ij}\psi_j\phi$, brings in the interactions of the fermions with the Higgs boson, which is where fermions gain mass. The scalar Brout-Englert-Higgs field, discussed further in Section 2.3, is represented by ϕ, and the Yukawa couplings, y_{ij}, set the strength of the relationship between the fermions and the Higgs, and thereby the size of their mass. This third term is not self-adjoint, and so the Hermitian conjugate term ($h.c.$) must also be included to keep the $\mathcal{L}$ real. This fourth term, $h.c.$, physically describes the Higgs interactions with the anti-quarks and anti-leptons. The fifth term, $|D_\mu\phi|^2$, gives the interactions of the Higgs with the weak vector bosons, explaining the mystery of their mass. The final term, $-V(\phi)$, is the Higgs potential as described in Section 2.3, which also allows its self-interaction. [14]

§ 2.3. Higgs Boson

2.3.1. Electroweak Interaction without the Higgs Boson

Without the introduction of the Higgs Boson, the Standard Model is an incomplete theory that misses several natural phenomena witnessed in experiment and experience of electroweak interactions. The electroweak Lagrangian, prior to the introduction of the Higgs, can be split into two terms, one relating to the four gauge bosons and one relating to the fermions, as in Equation 2.3.1.

$$\mathcal{L}_{EW} = \mathcal{L}_{Gauge} + \mathcal{L}_{Fermions} \tag{2.3.1}$$

The major problems with this Lagrangian lie in the gauge term, which can be expanded in Equation 2.3.2.

$$\mathcal{L}_{Gauge} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} = -\frac{1}{h}(\partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g_W \epsilon^{abc} W_\mu^b W_\nu^c)^2 + \frac{1}{h}(g_Y \partial_\mu B_\nu - \partial_\nu B_\mu)^2 \tag{2.3.2}$$

Here, g_W and g_Y are the coupling strengths of the electroweak $SU(2)$ symmetry (isospin) group and electroweak $U(1)$ symmetry (hypercharge) group, respectively [15]. While dense, the striking example of an unexplained phenomenon here is the massive nature of the W and Z boson weak force carriers. A mass term appears in a Lagrangian as a term with a single field times its adjoint, with no derivatives or other fields involved. In this electroweak Lagrangian, no such terms exist for the gauge bosons, despite experimental evidence that some of these bosons have mass. [16]

Even before these weak vector bosons were observed, it was known that they could not be massless like the photon, as this would lead to much higher rates of nuclear beta

decay. For the stability of diverse atomic nuclei that can be observed in our natural universe, the weak force carriers must be massive, though this breaks the fundamental electroweak symmetry described in Section 2.2.1, leading to the conclusion that there must be another player in the game. Other than the W and Z bosons, it was also not understood how most other fundamental particles, the fermions, achieved their mass, though it was not a requirement that this occur through the identical mechanism. [15]

Separate from the mass dilemma, experimental observation of the angular distributions of beta decays reveals that these interactions produce only left-handed electrons and right-handed positrons. Similarly, studies of pion decays reveal only left-handed neutrinos and right-handed anti-neutrinos. Convincing data from low-energy experiments demonstrate that the $W^{\pm}$ bosons have a chirality preference in interactions, coupling only to left-handed fermions and right-handed anti-fermions, breaking parity symmetry [15, 16]. The $SU(3) \times SU(2) \times U(1)$ symmetry that is so elegantly motivated in electroweak unification does not fit this chiral-specific interaction, and this, too, requires another agent to explain. In beautiful simplicity, the Higgs mechanism completes all these puzzles simultaneously.

2.3.2. Higgs Mechanism and Spontaneous Symmetry Breaking

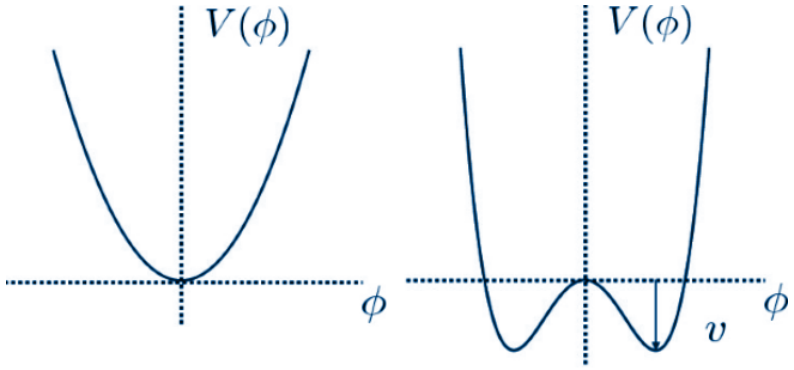
In 1964, Peter Higgs, and separately Francois Englert and Robert Brout, proposed a solution to the missing Standard Model elements with a simple question: what if these vector bosons gain mass through the interaction with another field [17, 18]? The simplest possible structure of a field that could explain the broken symmetry noted in Section 2.3.1 is a complex scalar field. The scalar nature of the field, meaning it has 0 spin, is supported by the isotropic angular distributions of particle interactions, impying the

field has no directional preference. The field must be complex in order to have a phase, which sets the gauge of these vector bosons and breaks their gauge invariance. Additionally, to explain the specific chirality observed in the weak interactions, the new field is introduced as a left-handed $SU(2)$ doublet. This complex scalar field became known as the Brout-Englert-Higgs field, $\phi = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi_1 + i\phi_2 \\ \phi_3 + i\phi_4\end{pmatrix}$.

The Higgs mechanism, by which the masses of the weak bosons arises from the introduction of the Higgs field, relies on the field having a non-vanishing vacuum potential. The only renormalizable potential that could result in such a non-vanishing lowest energy state is the Higgs potential in Equation 2.3.3.

$$V(\phi) = \mu^2(\phi^\dagger\phi) + \lambda(\phi^\dagger\phi)^2 \tag{2.3.3}$$

The first coefficient, μ^2, has some temperature dependence, and for extremely high temperatures $\mu^2 > 0$, implying a $V(\phi)$ like that shown in Figure 2.6(a) with a ground state at the origin. However, roughly a microsecond after the Big Bang, the universe cooled to below 160 GeV [19], and the value of μ^2 dropped below 0, leading to the potential landscape depicted in Figure 2.6(b). This potential is symmetric across phases of ϕ, with a ring of infinitely many equal lowest energy states away from origin, as can be seen in Figure 2.7 [4, 20]. At the time of the Higgs field ‘‘turn-on,” a single ground state emerged with non-zero expectation v, breaking this rotational symmetry and fixing the gauge of the interacting vector bosons, thereby breaking the $SU(2) \times U(1)$ symmetry from electroweak unification. One remaining degree of freedom remains from the four degrees of freedom of the complex scalar Higgs doublet, which would have resulted in three massless goldstone bosons, according to Goldstone’s conjecture [21]. However,



(a) $\mu^2 > 0$ (b) $\mu^2 < 0$

Figure 2.6.: Projected graphs of the Higgs potential in the early-universe case of extremely high temperature (a) and in the cooled universe case when symmetry breaking is forced (b).

due to the interaction with the gauge bosons, this is not the case. Three of the four degrees of freedom are absorbed into the interaction with the electroweak vector bosons, making three massive weak vector bosons, and the last remaining degree of freedom is left for the new scalar particle, the Higgs boson [22].

In order to look more closely at the Higgs mechanism and how it results in the masses of the weak vector bosons, we must include the kinetic term in the Higgs Lagrangian, Equation 2.3.4.

$$\mathcal{L}_{Higgs} = (D_\mu \phi)^\dagger (D_\mu \phi) - \mu^2 \phi^\dagger \phi - \lambda(\phi^\dagger \phi)^2 \tag{2.3.4}$$

The covariant derivative, $D_\mu \phi$ in Equation 2.3.4, maintains electroweak gauge invariance. Equation 2.3.5 writes the covariant derivative in terms of the generators of the electroweak $SU(2)_L \times U(1)_Y$ symmetry group noted in Section 2.2.1.

$$D_\mu \phi = \partial_\mu \phi - \frac{i}{2} g_W W_\mu^a \sigma^a \phi - 2 i g_Y B_\mu (Q - I_{W^{(3)}}) \phi \tag{2.3.5}$$

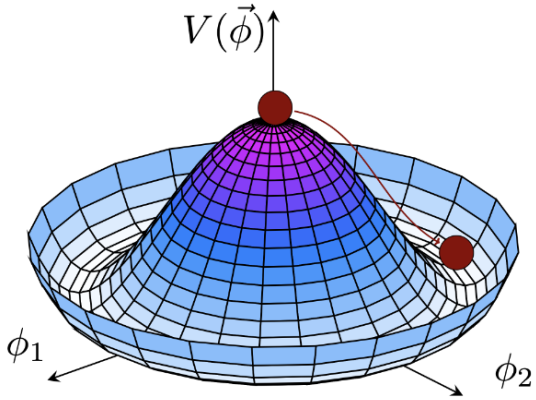


Figure 2.7.: Higgs potential in the complex plane in the case of non-vanishing vacuum potential, demonstrating the equipotential ring of infinitely many possible lowest energy states [23]. At the time of the universe cooling, a single Higgs ground-state arose, breaking this rotational symmetry and fixing the gauge of the electroweak vector bosons. The spontaneous symmetry breaking of the Higgs mechanism is illustrated through analogy with a ball rolling down from the top of a round hill, in which the direction the ball rolls is an equivalent representation of the broken symmetry.

If we consider the ground state of the Higgs with vacuum expectation value v, we can write this ground state without loss of generality as

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}.$$

Then,

$$|\phi_0|^2 = \frac{1}{2} v^2 = -\frac{\mu^2}{4\lambda}.$$

Plugging this expectation into the kinetic term of the the Higgs Lagrangian leads to Equation 2.3.6.

$$|D_\mu \phi|^2 = \frac{v^2}{8} (g_W^2 (W_\mu^{(1)})^2 + g_W^2 (W_\mu^{(2)})^2 + (g_W W_\mu^{(3)} - g_Y B_\mu)^2) \tag{2.3.6}$$

Here, we finally see the mass terms for the four bosons without any derivatives or other fields contaminating the terms.

We can further understand the relationship between the symmetry group generator basis and the mass eigenstate basis by putting this expectation on the kinetic Higgs Lagrangian term into matrix notation, where $|D_\mu \phi|^2 = \frac{1}{2} M_{ab}^2 A_\mu^a A^{\mu b}$ for M_{ab} given in Equation 2.3.7.

$$M_{ab} = \frac{v^2}{4} \begin{pmatrix} g_W^2 & 0 & 0 & 0 \\ 0 & g_W^2 & 0 & 0 \\ 0 & 0 & g_W^2 & -g_Y g_W \\ 0 & 0 & -g_Y g_W & g_Y^2 \end{pmatrix} \tag{2.3.7}$$

The matrix in Equation 2.3.7 has a degenerate eigenvalue of $M_W = \frac{g_W v}{2}$, as well as the non-degenerate eigenvalues $M_Z = \frac{1}{2} v \sqrt{g_W^2 + g_Y^2}$ and $M_\gamma = 0$. From here, we arrive at the relationship between the generator basis and the mass basis of the electroweak bosons. The $W^\pm$ bosons can be obtained from the first two weak isospin generators as $W^\pm = \frac{1}{\sqrt{2}}(W^{(1)} \pm W^{(2)})$, and the Z boson and photon are a mixture of the third weak isospin generator and the weak hypercharge group generator, such that $Z = \frac{g_W W^{(3)} - g_Y B}{\sqrt{g_W^2 + g_Y^2}}$ and $\gamma = \frac{g_W W^{(3)} + g_Y B}{\sqrt{g_W^2 + g_Y^2}}$.

While the Higgs mechanism was designed to explain the origin of the weak gauge boson masses and it does so quite elegantly, it also conveniently provides a mechanism for the masses of fermions. The Higgs-fermion term of the Standard Model Lagrangian is $\mathcal{L}_{Higgs-Fermion} = \bar{\psi}_i y_{ij} \psi_j \phi$. If the scalar Higgs field is expanded as a perturbation, h, around its ground state, v, such that $\phi = v + \frac{1}{\sqrt{2}} h$, then the Lagrangian term for the first

generation fermions becomes Equation 2.3.8.

$$\mathcal{L}_{Higgs-Fermion} = (v + \frac{1}{\sqrt{2}}h)(\bar{u_L} y_u u_R + \bar{d_L} y_d d_R + \bar{e_L} y_e e_R) \tag{2.3.8}$$

with neutrinos omitted. Here, for each fermion, we see two terms emerging: one mass term that is proportional to the Higgs vacuum expectation and the Higgs fermion coupling y_i (i.e. $m = y_i v$), and one fermion-Higgs interaction term with strength proportional to the fermion's mass through the coupling y_i. These couplings y_i are known as the Yukawa couplings. [24]

2.3.3. Open Questions Related to the Higgs

The Higgs boson was such an elegant solution to the remaining puzzles of the Standard Model that it motivated the investment of billions of dollars, thousands of minds, and decades of manpower to ensure its discovery. This investment paid off with the official Higgs boson discovery in 2012, nearly 50 years after the proposition of the theory. The machinery built toward this end is detailed in the next two sections.

Despite the incredible confirmation of the more complete Standard Model, there are still many important questions to ponder, motivating further investigation of the scalar boson. For example, is it really a single Higgs boson responsible for the masses of all generations of fermions [25]? While the larger coupling strength of the Higgs to the third generation quarks makes the interaction between the Higgs and b and t more abundant and thereby easier to study, there is still precision to be gained in measuring the couplings to these heaviest quarks. Furthermore, does the Higgs follow the same coupling strength trend to the second generation quarks and even lighter fermions? Does the Higgs' self-interaction match predicted behaviors? Could the Higgs boson couple

to new particles that have not yet been observed? Perhaps, in the process of probing this unique and important puzzle piece, clues to the physics that lies beyond our Standard Model can be illuminated.

In contribution to precisely probing the Higgs, this dissertation gives the details of a search for a rare Higgs production mode with a specific and interesting signature. This particular search through data is better motivated in Chapter 5, but first Chapters 3 and 4 detail the experiment source and apparatus that make the data possible.

Chapter 3.

The Large Hadron Collider at the European Center for Nuclear Research

Probing the fabric of the universe requires tearing it a bit, and for this, higher energies than those naturally being produced on our planet are necessary. To achieve particle interactions at higher energies, the largest machine ever built by mankind was erected. Sitting underground beneath Geneva, Switzerland, the Large Hadron Collider (LHC) is an octagonal ring 27 km in circumference. At the LHC, higher energy gives better resolution inside atomic nuclei, and grants higher probabilities for the creation of more massive particles [26], including ones we know but would like to understand better, like the Higgs boson, and potential particles that have not been met yet, like massive dark matter candidates. The method by which particles are accelerated at the LHC is described in Section 3.1. The collider's impressive performance in delivering frequent high energy interactions is discussed in Section 3.2, followed by the ongoing efforts to improve and advance its technologies in Section 3.3.

§ 3.1. Particle Acceleration at the LHC

3.1.1. Radiofrequency Cavities

To achieve these relativistic collision (center-of-mass) energies, charged particles are accelerated with strong electric fields, and in most accelerators like the LHC, this is done using radiofrequency (RF) cavities. These RF cavities generate a longitudinally oscillating electric voltage in a section of the accelerator, such that traversing charges experience the direction of voltage oscillation that gives positive acceleration [26]. As the voltage oscillates in the opposite direction (direction of deceleration), no ‘‘synchronous” charge (a charge perfectly synchronized with the RF frequency) is in the cavity, or in other words, the charges are elsewhere in the accelerator ring. In order to have resonant behavior, the RF frequency must be an integer multiple of the revolution frequency of charges around the ring, and that harmonic integer determines the number of ‘‘buckets” available for charges to occupy and stay synchronized with the accelerating voltage of the RF cavities. Once the charges are accelerated to the designed maximum energy, synchronous particles will experience no further force from the RF cavities, while particles moving faster or slower will experience deceleration or acceleration toward the path of the synchronous charge, keeping the charges tightly organized in ‘‘bunches” centered in RF buckets [26].

The LHC houses 16 RF cavities (8 per beam), each tuned to oscillate at 400 MHz and delivering up to 2 MV of potential difference. This oscillation frequency gives a harmonic number of ≈35,640 for protons traveling around the 27 km ring at roughly the speed of light. With 35,640 RF buckets, the LHC could in theory accelerate as many bunches of protons, but in reality, operators leave empty bucket gaps. Some gaps, known

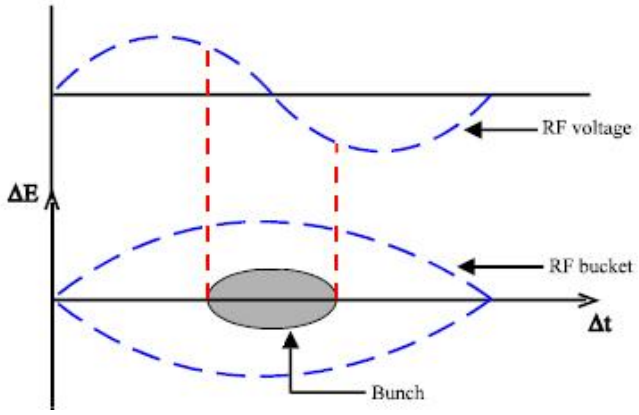


Figure 3.1.: Radiofrequency (RF) cavities deliver oscillating voltages that accelerate charged particles. The accelerating voltage oscillates with a frequency that is a harmonic of the revolution frequency of particles in the ring, and that harmonic number determines the number of segments of the ring circumference that can possess synchronized acceleration known as buckets. Traveling charges are accelerated longitudinally by these RF-induced electromagnetic fields, clumping in the centers of these RF buckets forming bunches. [26]

as abort gaps, are needed to allow time for kicker magnets to divert the beams when a dump is voluntarily performed [27]. Other gaps are necessary to ensure that the bunches collide in the intended locations at the centers of experiments. A typical LHC proton physics fill in Run 3 has 2808 proton bunches in each ring.

3.1.2. Magnets

While electric fields from RF cavities linearly accelerate the charged particles, another equally important tool, magnets accelerate charges tangentially to maintain the trajectories in the accelerator ring and focus the beams in the transverse directions. Bending the beams of charges in a flat circle, strong magnetic dipoles produce magnetic fields perpendicular to the velocity of the charges and exert a Lorentz force to curve their

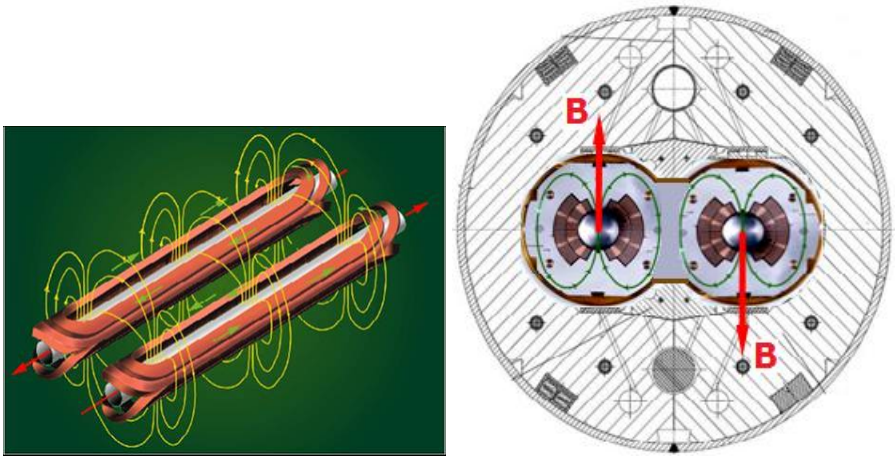


Figure 3.2.: LHC's 2-in-1 dipolar magnet design. The superconducting wires (pictured on the left) are arranged with currents circulating on either side of the beam pipes, through which the proton beams pass. These currents produce magnetic fields (pictured on the right) in opposite directions through the centers of the two beam pipes, one corresponding to the clockwise-circulating beam and one to the counter-clockwise beam. [26]

trajectories into an orbit around the ring. In order to produce the centripetal force necessary to maintain the orbit of the LHC ring, the dipole magnetic fields must be 8.3 T [26], which is roughly 100 thousand times the strength of Earth's magnetic field. The LHC dipoles have a specific design, shown in Figure 3.2, that make them a ''2-in-1" magnet, able to generate opposite fields in each ring, simultaneously controlling the clockwise and counter-clockwise circling beams. At an operating temperature of 1.9 K, the niobium-titanium cables producing these dipole magnets are super-conducting. The magnetic fields produced by these dipoles interact and exert a force that pushes the magnets to ''open." In order to counteract this destructive force, stainless steel collars surround the Nb-Ti cabling, and a magnetic iron yoke surrounds these non-magnetic collars. The LHC has 1232 such dipole magnets, each 14.3 m in length, positioned around the ring [26].

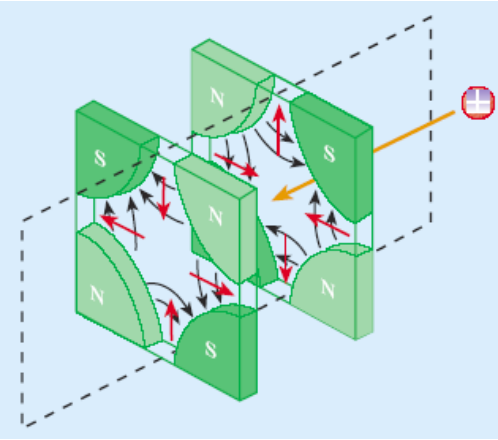


Figure 3.3.: Schematic of the focusing forces caused by successive quadrupole fields. For a positive charge coming from the right, the red arrows represent the direction of force the charge would experience depending on the region of the magnet traversed. The first magnet in the positive charge's path (right-most quadrupole) would focus the horizontal dispersion of the beam, as, for example, a particle traveling on the right of the beam center would receive a kick to the left. The second magnetic field (left-most quadrupole) would produce a focusing effect in the vertical direction. [26]

Acting on beams of charges like a lens acts on light, quadrupole magnets are used to focus the fast-moving beam. A given quadrupole magnet can either focus the horizontal width or the vertical height of the beam, but not both. To the opposite dimension, the quadrupole will give a de-focusing force. For this reason, quadrupoles are set up in successive pairs, one focusing on narrowing the beam dispersion in width (known as QF), and one defocusing the width while narrowing the spread in height (known as QD). Figure 3.3 shows the induced forces on the beam, demonstrating the one-dimensional focusing effects. The LHC uses a total of 858 magnetic quadrupoles. Two triplets of quadrupoles are used to focus the beams into a very narrow spot for insertion into an interaction point. These insertion magnets are also responsible for cleaning the beam to ensure no radiation damage to the most sensitive components of the LHC. In addition to quadrupoles, the LHC also uses sextuple, octopole, and even triplet magnets to further focus its beams [26].

3.1.3. Amplitude and Emittance

Two quantities are used to describe the beam size: transverse emittance, ε, and amplitude function, β.

Emittance, ε, describes the spread in position and momentum of particles along a particular beam axis, with the vertical and horizontal axes corresponding to transverse emittances. A smaller ε implies a tighter distribution of beam particle positions and momenta, leading to higher probabilities of interaction and higher collision luminosities. The transverse emittance of the beam is a constant intrinsic property that is reflective of the quality of bunch preparation process and cannot be altered by focusing inside the accelerator [26].

On the other hand, the beam amplitude function, $\beta(s)$, is a function of the position, s, around the ring, as it reflects the setup and structure of accelerator's focusing magnets. The beam function β refers to the envelope of all particle trajectories in the beam, such that the maximum amplitude of oscillation of a single particle at position s of the ring is given by $x(s) = \sqrt{\varepsilon}\sqrt{\beta}(s)$. Of particular interest, the envelope of particle trajectories at the interaction point inside a particular detector is denoted by β^*. A narrower beam has a low β, while a high β denotes a wide and straight beam envelope [28].

3.1.4. Synchrotron Radiation

The benefit of circular colliders over linear colliders is that charges can witness the same accelerating cavities many times over, so shallower electric field gradients are needed as energy can be transferred in many tiny kicks. This would be absolutely ideal if it weren't for synchrotron radiation. The magnetic fields that keep the orbiting charges locked in the accelerator ring cause the relativistic charged particles to emit radiation, known as synchrotron radiation. As a result, the charges lose energy with each turn, and the RF cavities are forced to put that energy back into the beam. The energy lost in each second of revolution is given by Equation 3.1.1.

$$\left[\frac{e^4}{6\pi\epsilon_o m_o^2 c}\right]\gamma^2 B_\perp^2 \tag{3.1.1}$$

In Equation 3.1.1, $\gamma = \frac{E}{m_o c^2}$. Due to the mass dependence, energy loss due to synchrotron radiation is about 10^{13} times greater for electrons than for protons. For protons in the LHC, synchrotron radiation contributes an energy loss of about 10 keV per turn [26].

3.1.5. LHC's Acceleration Complex

In order to achieve the designed energy of beams in the LHC, which has protons traveling at 99.9999991% the speed of light, the proton acceleration does not occur all in one step. A network of accelerators is used in series, each delivering incremental impulses to the bunches of protons, increasing the velocity of travel progressively and reducing bunch size and emittance. The network of CERN machines that feed the LHC are displayed in Figure 3.4. The linear accelerator 3 (Linac3) and low energy ion ring (LEIR) are used to supply heavy ions to the LHC. The other machines are part of the proton acceleration sequence and are discussed below.

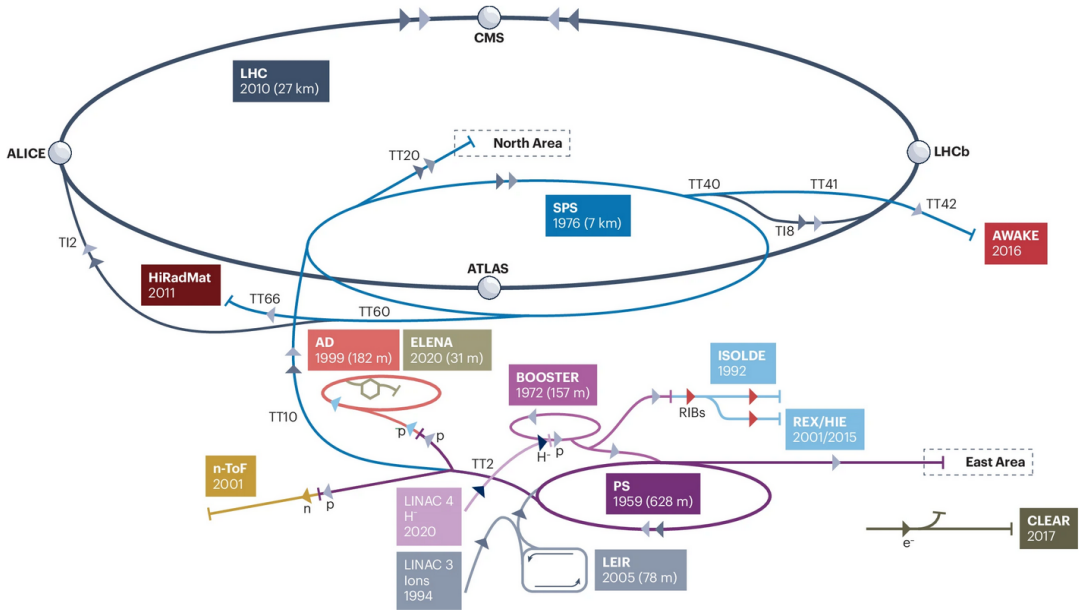


Figure 3.4.: Schematic diagram of CERN's accelerator machines . Protons are prepared in bunches through the Linac4, Booster, PS, and SPS, until they are finally fed into the LHC. [29]

Protons begin effectively at rest in the form of hydrogen gas in a bottle; a typical 750 mL water bottle could provide enough hydrogen atoms to fuel the LHC for over

200,000 years. A strong electric field is introduced to the hydrogen gas to split the diatomic molecules into H^+/H^- pairs. It is the H^- ions that are accelerated in CERN's Linear Accelerator 4 (Linac4) to an energy of 160 MeV [30]. Linac4, 86 m in length, became fully operational in 2020, replacing Linac2 in the LHC supply chain as a step in the direction of increasing beam luminosity, a goal which will continue over the course of this decade [31]. At the end of Linac4, a metal foil strips the electrons from the negative hydrogen ions, and protons are injected into the next accelerator, the Proton Synchrotron Booster (PSB). This charge exchange allows for effective beam extraction and for the accumulation of high current injection beams. The PSB accepted its first beams in 1972, and now accelerates protons to an energy of 2 GeV before injecting into the Proton Synchrotron (PS) [32]. The PS was CERN's first synchrotron, and has been doing its proud work since 1959, accelerating in its time not just protons but also alpha particles, oxygen, sulphur, argon, xenon and lead nuclei, electrons, positrons and antiprotons. In addition to continuing the LHC acceleration complex, the PS feeds the Antiproton Decelerator (AD), which fuels many low energy antimatter experiments [32]. Protons are brought to an energy of 26 GeV in the 628 m circumference of the PS, before entering the Super Proton Synchrotron (SPS). The SPS takes protons to their injection energy into the LHC of 450 GeV, but it doesn't just supply the LHC; it also feeds beams to other experiments like NA62, COMPASS, and the CERN North area test beam facility. A major accolade was earned by the SPS in 1983, when it operated as a proton-antiproton collider, and contributed to the Nobel-prize-winning discovery of W and Z bosons [32]. Once the beams arrive in the LHC (one going clockwise and one counter-clockwise), it takes about 20 minutes to ramp up the proton energy to 6.8 TeV. Table 3.1 summarizes the final energies reached by protons at each stage of the accelerator complex leading to the

Kinetic energy of a proton (K)	Speed (%c)	Accelerator
160 MeV	52	Linac 4
2 GeV	94.8	PS Booster
25 GeV	99.93	PS
450 GeV	99.9998	SPS
7 TeV (design energy)	99.9999991	LHC

Table 3.1.: Target proton energies and speeds in the CERN machines involved in the accelerator complex feeding protons to the LHC. [33]

LHC.

§ 3.2. Proton-proton Collisions at the LHC

3.2.1. LHC Turnaround

The time between when a beam dump has occurred and the LHC is back to delivering stable beams is known as turnaround and can take several hours. Protons are injected into each LHC ring from the SPS at an energy of 450 GeV in 39 batches of 72 proton bunches for nominal fills with a ‘‘full” 2808-bunch ring. This injection process is the most complicated phase of the LHC turnaround, taking nearly an hour as many pauses are taken to maintain cryogenic conditions and avoid overload [27].

Once fully injected, the ‘‘ramp” phase begins as protons are accelerated to collision energies, using the RF cavities mentioned in Section 3.1.1. In the 20 minutes of circulation it takes to accelerate the proton bunches to nearly 14 times their injection energy, the beams will pass through the RF cavities more than 10 million times [34]. After reaching 6.8 TeV (as the target energy of Run 3) or 6.5 TeV (operating energy of Run 2), a transitional ‘‘flattop” phase is declared as the beam quality is monitored

in preparation for the ‘‘squeeze” of the beams*. During the squeeze, the beam size β^* at the experiment interaction point, as described in Section 3.1.3, is minimized using two quadrupole magnets positioned before and after the ATLAS detector. Through this point in the LHC cycle, the beams are kept apart with local corrector magnets that provide separation bumps to avoid unstable collisions. During the ‘‘adjust” phase, these separation bumps are removed, and the instantaneous luminosity (see Section 3.2.2) is optimized for collisions [35].

Finally, stable beams are reached if all these steps go to plan. This implies that prior to even injecting, the four accelerators mentioned in Section 3.1.5 must have perfect execution of their duties, the cryogenic systems must keep the magnets cool enough for super-conduction, and the LHC beam pipes must maintain ultra-high vacuum pressures, ten times less than the atmosphere of the moon [35]. With the myriad ways reaching these high-energy collisions can go wrong, reaching the stable beam declaration is nothing short of a miracle.

3.2.2. Stable Beams

The Large Hadron Collider (LHC) at the European Center for Nuclear Research (CERN) circulates protons in two beams around a 27 km circumference ring at an energy of 6.8 TeV each to produce 13.6 TeV proton-to-proton collisions every $25\,\mathrm{ns}$. Note that 13.6 TeV is the LHC’s center-of-mass energy during Run 3 (2022-2026), but it was 13 TeV during Run 2 (2015-2018), and 7 and 8 TeV during Run 1 (2010-2013). While the LHC is the highest energy collider to ever exist, possibly more impressive is its

*My mother is quite disappointed that ‘‘Flattop Squeeze” has not yet been utilized as a name of a CERN rockband.

luminosity, or the number of collisions that can be produced at an interaction point per cm^2 per second. The instantaneous luminosity, $L_{inst} = \frac{n_b \langle \mu \rangle f_r}{\sigma_{inel}}$ is a measure of the rate of data production from a full ring with n_b colliding bunches revolving at a frequency f_r with average pile-up $\langle \mu \rangle$, normalized to the cross-section of pp inelastic scattering σ_{inel}. The instantaneous luminosity is measured very closely by LUCID, a Cherenkov detector positioned around the beamline 17 meters from ATLAS. It is important that the size of the collected datasets can be known as precisely as possible for giving context to yield measurements to calculate process cross-sections, which is very important to experiment. LUCID uses 16 photon multiplier tubes (PMTs) to detect the number of inelastic collisions in a given bunch crossing [36]. For a filled LHC ring, with 1.1×10^{11} protons per bunch, the collision luminosity is about $10^{34} \mathrm{cm}^{-2} \mathrm{s}^{-1}$, the highest for any proton collider ever [33]. The probable number of collisions can also be described per bunch crossing with a parameter known as pile-up, μ. The average pile-up values for each run are shown in Figure 3.5. The LHC was designed to produce an average pile-up of 20, but thanks to feats of engineering, the average pile-up during Run 3 is three times that number and climbing.

The measure of how much data has been collected over time is the integrated luminosity, effectively a count of how many collisions have been delivered to detectors during stable beam running time. The units of integrated luminosity are in inverse cross-sections such as fb^{-1}, such that the product of the integrated luminosity and a process cross-section gives the predicted number of times such a process has occurred in the data [26]. The integrated luminosities for each year of Runs 1, 2, and 3 (through 2024) are shown in Figure 3.6. The year 2024 truly marked record-breaking achievements, as nearly the same number of collisions produced in Run 2 as a whole were produced in a

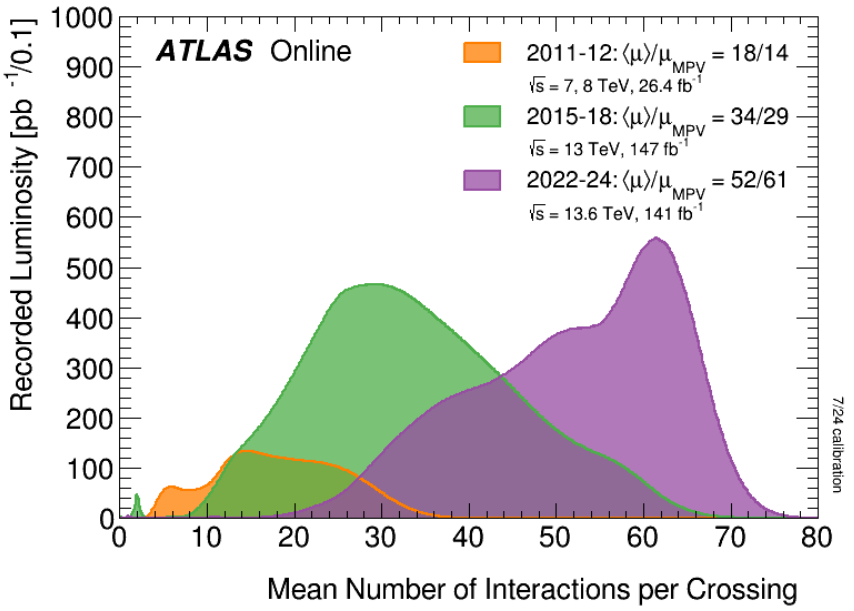


Figure 3.5.: Average pile-up for each LHC run. [37]

single year of Run 3.

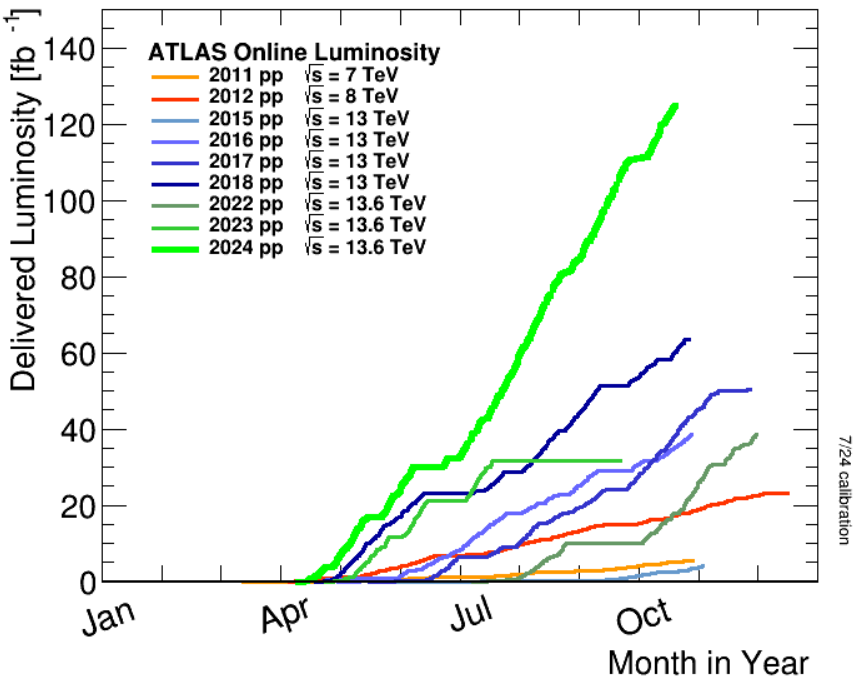


Figure 3.6.: Integrated luminosity of pp-collisions delivered by LHC for all run years from 2011 through 2024. [37]

§ 3.3. High-Luminosity LHC Upgrade

Beginning operation around 2030, a new phase of the LHC will bring a brand new era of data collection, upending the accelerator complex to produce the High-Luminosity LHC (HL-LHC). From the LHC to the HL-LHC, the instantaneous luminosity will increase by a factor of 5-10, approaching $10^{35}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, and the average pile-up will increase to up to 200 collisions per bunch crossing. At the new center-of-mass energy of $14\,\mathrm{TeV}$, the HL-LHC will produce 15 million Higgs bosons per year (compared to 3 million produced by the LHC in 2017). The HL-LHC is projected to produce $250\,\mathrm{fb}^{-1}$ of collision data per year, and by the end of its operation, ten times as much data will be collected than the entire LHC program to date [38].

A more intense and focused beam is necessary to achieve these advancements, and this requires new technology. Stronger niobium-tin quadrupole magnets will be added on either side of each experiment to squeeze the beams more tightly at the interaction points. New ‘‘crab cavities” will also be installed on either side of the experiments to introduce a tilt to the beams just before the interaction point, which will give the proton bunches some transverse momentum to enlarge the area of overlap between crossing bunches and increase the collision probability. These new cavities will be powered by newly-innovated 100 meter superconducting cables made of magnesium di-boride, which can conduct extremely intense currents at much warmer temperatures as high as 50 K. Additionally, collimators will be added and existing ones upgraded as reinforcements to absorb energy from stray particles. All in all, to transition the LHC for its high luminosity era, upgraded technology will be added to 1.2 km of the 27 km ring [38].

Chapter 4.

The ATLAS Detector

The ATLAS experiment has operated since 2008 as one of two all-purpose particle detectors at the LHC. One of many other forced acronyms in particle physics, ATLAS stands for ‘‘A Toroidal LHC ApparatuS,’’ a name which pays tribute to the iconic donut-shaped magnets which help to measure muon momenta. The ATLAS collaboration consists of over 3000 contributing authors who divide their time working on data taking and operations, data analysis, and detector upgrades. Details of the experiment’s design are discussed in Section 4.1, followed by some insights into its operation Section 4.2, and finally, a description of the plan to upgrade the detector to handle the High-Luminosity LHC is depicted in Section 4.3.

§ 4.1. Design and Components of the Current Detector

4.1.1. ATLAS Experiment Overview

The ATLAS detector acts as a catcher’s mitt for seizing what it can of the eruptive products from the proton collisions delivered by the LHC. The objective of the detector’s

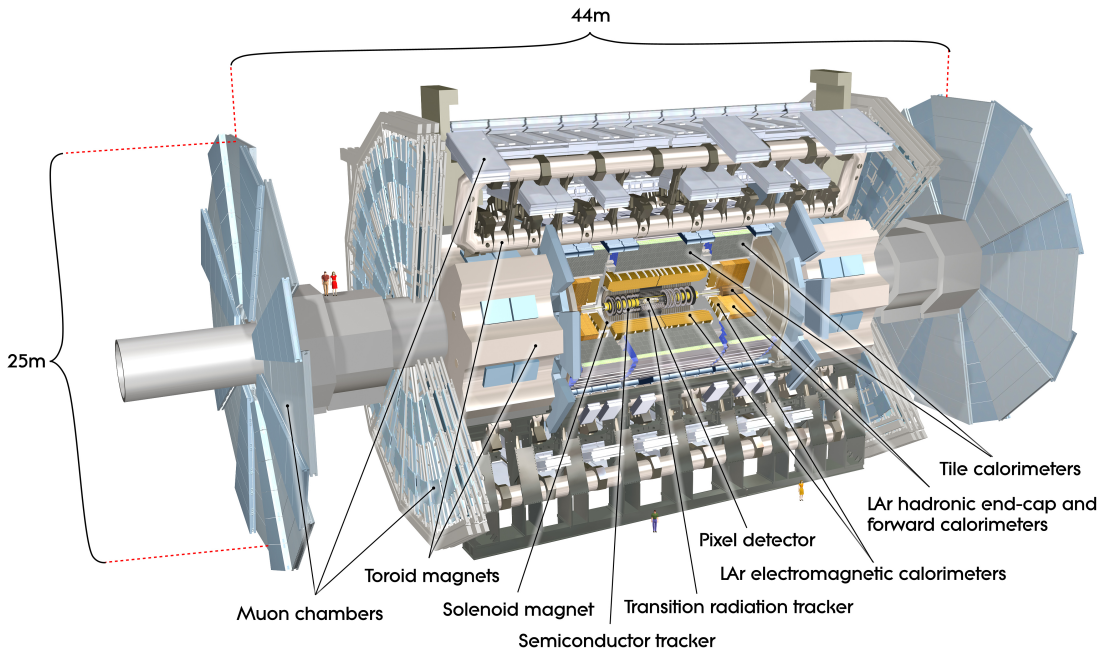


Figure 4.1.: A Toroidal LHC ApparatuS, ATLAS, sits 100 meters underground, where it collects proton-proton collisions from the LHC. It is composed of many subdetector elements, each designed to measure specific properties of particles traversing their volume. [39]

design is to identify the particles and their trajectories by measuring their energies and interactions. ATLAS as a whole is portrayed in Figure 4.1 as it sits 100 meters below the Earth's surface in Meyrin, Switzerland.

The experiment is constructed as a cylindrical onion, with layers of different detectors. Particles can largely be identified by the presence or absence of a signature or energy deposit in the layers traversed in the detector. Oriented from inward to outward radially relative to the collision site (which is roughly in the center of the detector), the layers of ATLAS are digested as follows. The Inner Detector, discussed in Section 4.1.2, provides tracking information for charged particles, but neutral objects are not

typically visible. Directly outside the inner detector, the electromagnetic calorimeters (Section 4.1.3) tend to absorb all the energy of light electromagnetic objects, like electrons and photons. Hadronic calorimeters, detailed in Section 4.1.3, make up the next layers, where the total energy of hadronic objects is deposited via strong and electromagnetic interactions. Lastly, with technologies expanded upon in Section 4.1.4, muon spectrometers measure the passage of heavy leptons, which typically bulldozes through other absorbing layers. Thanks to the information provided by these detectors, a method similar to 90's magazine quiz can be used to identify the particles in the experiment. Figure 4.2 summarizes this simplistic identification method, though it is not always as cut-and-dried as this flow chart presents.

A common detector coordinate system is defined for clarity in the communication of event kinematics. In this coordinate system, the z-axis runs parallel to the beam-line, as the x-axis runs parallel to the LHC plane with its positive direction oriented toward the center of the LHC ring. The y-axis points up. Event object kinematics are often discussed in terms of angular distributions and orientations. Rather than use θ, the angle with the beam-line directly, rapidity* gives a Lorentz invariant quantity, which is important as the particles being described travel at relativistic speeds. Pseudorapidity**, η, approximates rapidity in the low-mass regime and is most often used, since it is directly calculable from the Cartesian coordinates. η is 0 directly perpendicular to the beam-line, and is larger closer to the beam-line or in the ‘‘forward" direction. The azimuthal angle, ϕ, is measured around the beam relative to the x-axis. ATLAS is symmetric in ϕ [40]. Figure 4.3 gives a diagram of this coordinate system.

*Rapidity can be mathematically defined as $y = \frac{1}{2}\ln(\frac{E+p_z c}{E-p_z c})$.

**$\eta = -\ln\tan\frac{\theta}{2}$

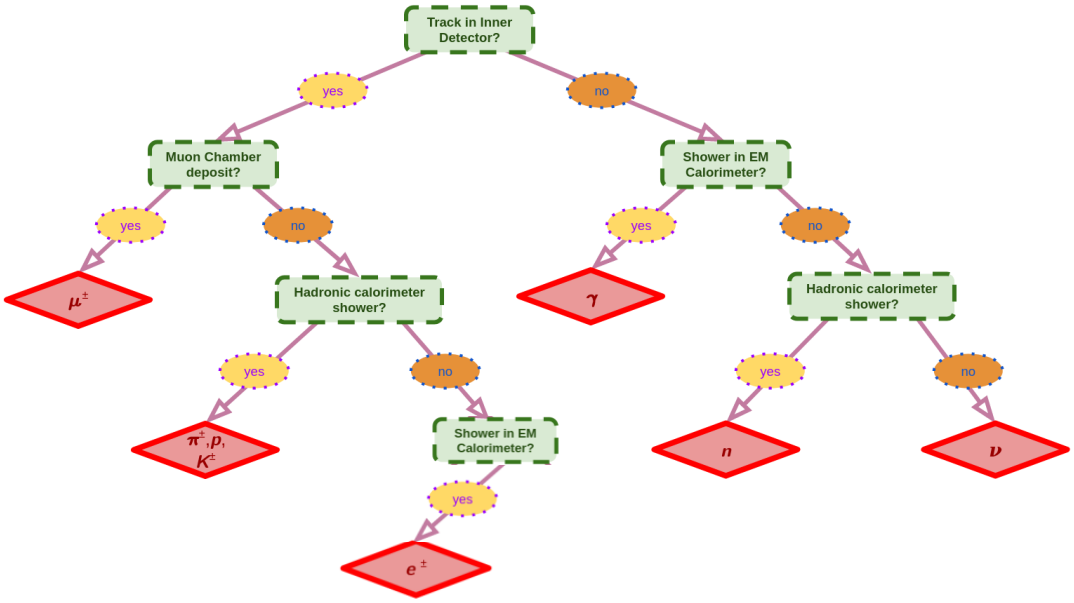


Figure 4.2.: Flow chart depiction of a simplification of ATLAS's methods of particle identification utilizing the experiment's layered design.

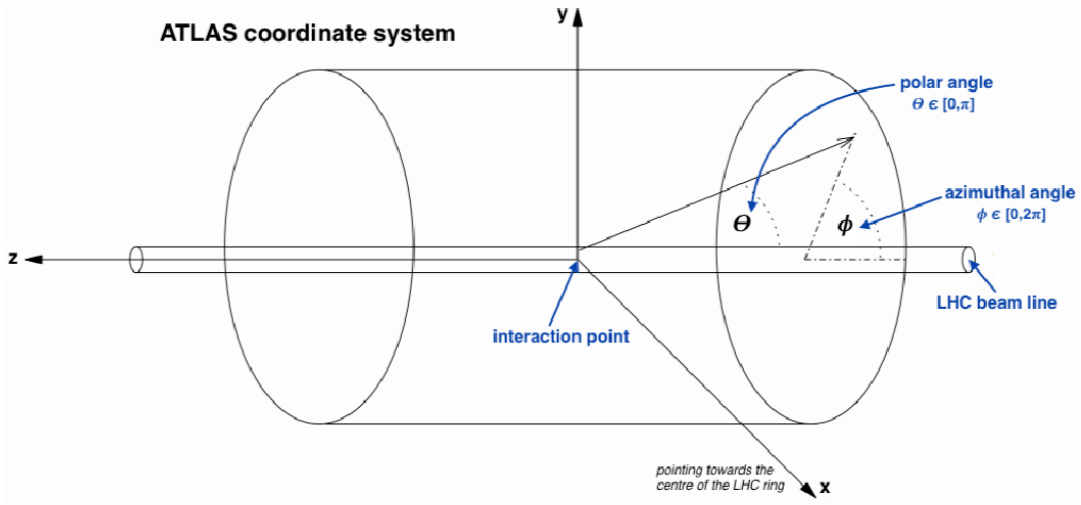


Figure 4.3.: Diagram depicting the coordinate system of ATLAS. The z-axis runs parallel to the beam-line, and the x-axis is oriented parallel to the Earth's surface toward the center of the LHC ring. Relevant pseudorapidity and azimuthal angles also demonstrated. [40]

4.1.2. Inner Detector

The inner detector (ID) consists of 3 subsystems: the pixel detector (PIX), the semiconductor tracker (SCT), and the transition radiation tracker (TRT). The ID sits at the smallest radius of ATLAS, closest to the beam line and primary interaction sites. Together, the components of the ID provide tracking of charged particles that pass through their included volume, which extends from a radius of 3.3 cm all the way to 115 cm from the beam-line and covers $|\eta|$ up to 2.5 [41, 42]. The layout of the ID is depicted in Figure 4.4, giving the position and coverage of each layer. Registered hits or charge deposits in segments of these layers are used to reconstruct the trajectories of charged particles. At least 7 hits must be registered among the layers of PIX and SCT to constitute a ‘‘track” in the pseudorapidity range $|\eta| \leq 2.5$.

Both SCT and PIX are silicon detectors and share a common technology. A bias

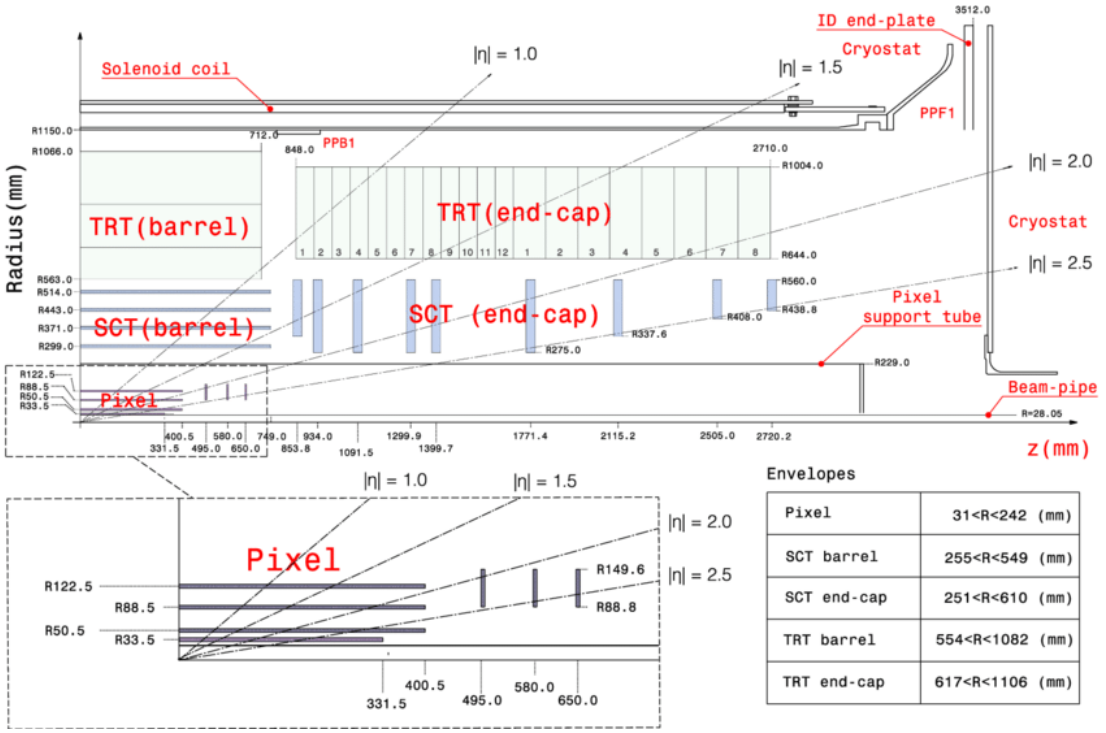


Figure 4.4.: Schematic of the layout of the layers of the inner detector, including coordinates in radius, z, and η. Schematic was created prior to the insertion of insertable b-layer, which is not shown but exists at a radius of 3.3 cm. [43]

voltage is applied to the thin semiconductor sensors until the bulk is depleted of charge carriers. Traversing charged particles excite electron-hole pairs which then accelerate toward the electrodes and generate a measurable signal to the readout electronics. Particularly in SCT, special attention is paid to the peak charge collection timing and charge collection spread at specific readout electrodes to better improve resolve the position of the traversing charges.

The innermost layer of ATLAS, PIX, may encompass the smallest volume but uses over half the data read-out channels of the entire experiment, due to its high density of readout elements which yield extremely precise position resolution. It consists of 4 layers in the barrel or central region of ATLAS and 3 disks covering the forward regions. The innermost layer, titled the insertable B-layer (IBL), was added during the technical shutdown before Run 2 to compensate for the increase in luminosity beyond the capabilities of the original 3-layer design. IBL has been taking data efficiently since 2015 [44]. PIX's layers are composed of planar silicon pixels, roughly 92 million in number, each with a size of 50 μm $\times$ 400μm^2 (50 μm $\times$ 250 μm^2 in IBL), i.e. about the size of a typical grain of sand. PIX uses n$^+$-in-n type sensors, meaning an n$^+$-implant readout is used on top of an n-type sensor bulk, which are doped by replacing silicon nuclei with phosphor to alter the conductivity [45]. Combining the precision of the PIX layers, this subdetector has a resolution of 10 μm [45].

Surrounding PIX, the semiconductor tracker (SCT) is comprised of 4 central barrel layers and 9 forward endcap disks. Like PIX, SCT is made up of silicon semiconductor, but rather than having 2D information in signal collection like pixels, SCT uses semiconductor strips with 1D information. Over 6 million ‘‘micro-strips" are positioned every 80 μm. To reduce noise and extend lifetime, both SCT and PIX are chilled to an

operating temperature below 0^oC with a shared cooling system pumped with octafluoropropane (C_3F_8). With the compounded precision of the multi-layer design, SCT has a resolution of 25 μm [41].

The outermost layer of the inner detector is the transition radiation tracker, TRT. TRT consists of 300,000 drift tubes (also known as ‘‘straws”), which are each 4 mm in diameter and are filled with a Xe-based or Ar-based gas mixture. A 31 μm gold-plated tungsten wire sits at the center, to which the thin Kapton walls are held at a relative potential of 1.5 kV. Traversing charged particles ionise the gas mixture, producing roughly 5-6 initial ion clusters per mm of path length. Accelerated toward the central anode by the electric field induced by the potential difference of the straw walls, the initial ion clusters undergo an avalanche effect with an amplification factor of 2.5×10^4, producing a measurable current [46]. The collection of these tightly packed drift tubes has an overall position resolution of 0.17 mm. In addition to registering charged particle tracking information, TRT provides useful particle identification information (not included in the simplified schematic of Figure 4.2). TRT's drift tubes are surrounded by stacks of polypropylene or polyethylene fibres acting as radiator foils, and as electrons cross these material interfaces, transition radiation is produced which can effectively be measured in the Xe-mixture filled drift tubes. Charged pions, on the other hand, have a much lower Lorentz factor than an electron at similar energies, and so transition radiation from a pion is far less common. Using measurements of transition radiation, TRT is not just a tracker but a powerful tool to discriminate between electrons and pions [47].

4.1.3. Calorimetry

Outside of the tracking detectors sit calorimeters, which perform a destructive measurement of the particle's energy. The calorimeters are designed to take advantage of a particular force or interaction, and are layered in such a way that the layers penetrated by a particle give great insight into the particle's identity. Just outside the inner detector, the electromagnetic (EM) calorimeter extends from a radius of 1.15 m to a radius of 2.25 m and covers $|\eta|$ up to 3.2 [48]. The EM calorimeter entirely uses Liquid Argon sampling technology, which is described further in Section 4.1.3. The Hadronic (HAD) calorimeter sits outside the EM calorimeter, looking for high-mass particles that manage to bulldoze their way through the EM calo layers without losing all their energy to bremsstrahlung. The HAD calorimeter also uses Liquid Argon calorimetry in the high-pseudorapidity endcap regions, and tile calorimetry in the central barrel, described in Section 4.1.3. Muons sail through both the EM and HAD calorimeters, so muon spectrometers are positioned at the largest radii of ATLAS with special technologies described in Section 4.1.4.

Liquid Argon Calorimeters

ATLAS's Liquid Argon (LAr) calorimeters are known for their accordion-like design, a shape which helps to ensure that there are no gaps where particles could escape undetected. The LAr calorimeter design can be observed in Figure 4.5. These calorimeters consist of liquid argon sandwiched between metal absorbing plates, made of either tungsten, copper, or lead. Fast-moving charged particles that hit the absorbing metals get converted through bremsstrahlung into ‘‘showers" of lower-energy particles, which ionize the liquid argon and can be counted by the proportional current generated on the

electrodes [49]. Lead absorbers are used in the EM calorimeter due to their short radiation length, X_o, so that space can be saved. Copper absorbers and tungsten absorbers in the far forward calorimeter (FCAL) have a short interaction length, λ_o, and are used for hadronic calorimetry to maximize nuclear interactions in the space available. To keep argon in its liquid form, these calorimeters are kept below -184^oC [48].

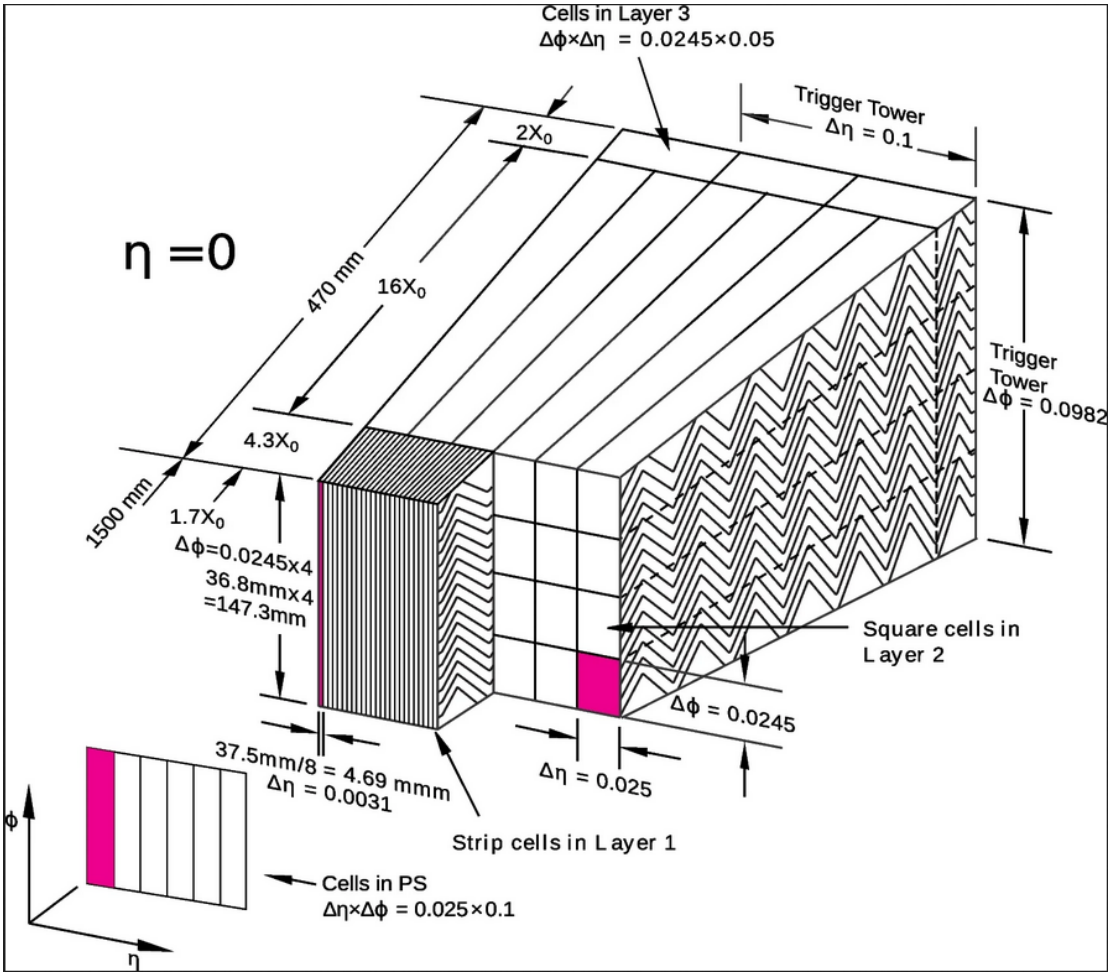


Figure 4.5.: Liquid Argon calorimeter design. The accordion-like layers are visible in the side cross-section. Careful considerations were taken on the depth required to provide a sufficient number of radiation lengths. Large energy deposits are registered by hardware ‘‘triggers” to aide in the filtration of data to be recorded by the experiment. [50]

Tile Calorimeters

The hadronic tile calorimeter surrounds LAr in the barrel region. A similar layering of absorbing and counting materials is employed, but this time the absorber is steel, a material with a relatively short nuclear interaction length, and the counter is a scintillating plastic. There are over 420,000 scintillating tiles used in ATLAS, which, combined with the steel, weigh over 3000 tons. Again, traversing particles hit the steel and produce showers of lower energy particles, which here induce the scintillating tiles to produce photons. The photons are then read out via nearly 10,000 photomultiplier tubes, which convert the light into a proportional current [49]. The steel and scintillator layered design of the tile calorimeter is displayed in Figure 4.6. Unfortunately, some of the showered particles are neutral and produce ‘‘invisible energy” in the form of neutral hadrons that will not be detected, and this has to be corrected for by a multiplicative factor on the visible energy recorded. Uncertainty on the amount of invisible energy does impact the detector’s energy resolution.

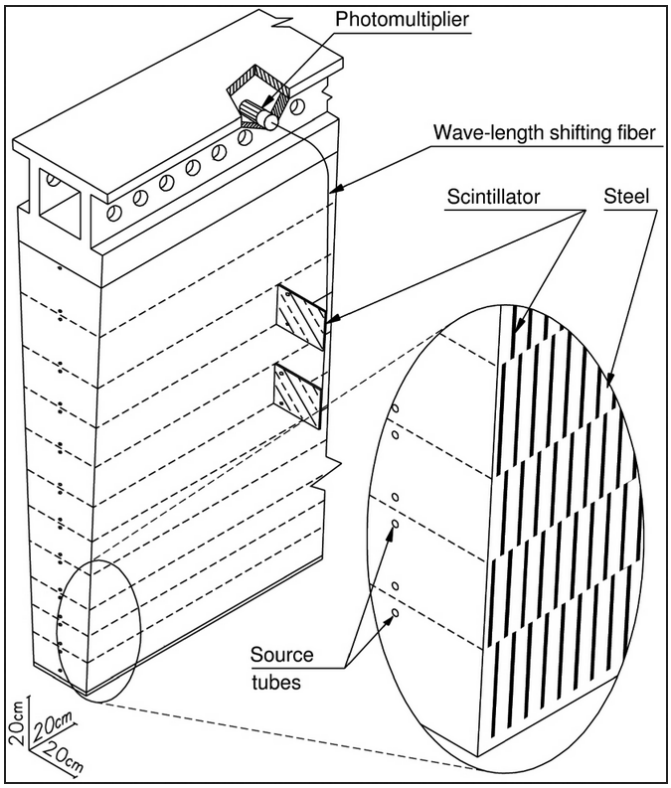


Figure 4.6.: Tile calorimeter design, which includes alternating steel and plastic scintillating crystal tiles. The steel tiles absorb some of the energy of traversing hadronic material, creating sprays of lower energy particles that are sampled and counted in the scintillating tiles. [51]

4.1.4. Muon Spectrometers

Muon spectrometers, sited at the farthest reaches of the ATLAS detector, use 4 unique technologies that can be divided into two purposes: speed and precision. The fast-response technologies include Resistive Plate Chambers (RPC) and Thin Gap Chambers (TGC), and these components are relied upon to make triggering decisions (discussed further in Section 4.2.1) within 2.5 μs. RPCs consist of parallel bakelite plastic plates held at a potential difference with an ionizing gas filling the gap in between, such that traversing muons generate a measurable current. TGCs exist at the endplates of ATLAS

and contain gold-plated tungsten wires at high voltage in a thin 1.4 mm gas-filled chamber sandwiched by cathode planes of a graphite-epoxy mixture. In TGCs, as well, the enclosed gas is ionized to generate a current for readout. The muon spectrometer precision technologies include Monitored Drift Tubes (MDT) and Cathode Strip Chambers (CSC). MDT uses a similar technology to that of TRT, gas-filled tubes with a central wire anode, though these tubes are metal. They get the name ‘‘monitored” due to the precise optical alignment devices (alignment precision of 30 microns) used to keep track of their precise positioning, which is necessary to achieve muon momentum resolution of 10% for a muon with a p_T of 1 TeV. The precise MDTs, with a spatial resolution of 80 μm are layered in a shared support structure with the fast RPCs. The CSCs are positioned at high-pseudorapidity and are even more precise with a spatial resolution of 60 μm. They achieve this through segmentation of the readout cathodes of multi-wired thin gap gas-filled proportional chambers. Together, these technologies, displayed in Figure 4.7, reconstruct the energy and trajectory of these heavy charged particles that otherwise pass straight through the other detector components [52].

Two ‘‘New” Small Wheels (NSW) were added to ATLAS in 2019 before the start of Run 3, replacing the previous small wheel CSCs in the endcaps in order to cope with the increase in luminosity and consequent event backgrounds while maintaining efficiency in performance. These not-so-small wheels are over 9 m in diameter, and their position and composition can be seen in Figure 4.8. NSW uses two layered technologies: small strip thin gap chambers (sTGC) and micromegas (MM), both of which are gas-filled chambers relying on ionization to produce a measurable signal [54]. sTGC, similar to TGC, is constructed from a gas-filled sandwich of cathode boards with a gold-plated tungsten anode in the center, and it gives fast signalling for the trigger system. MMs con-

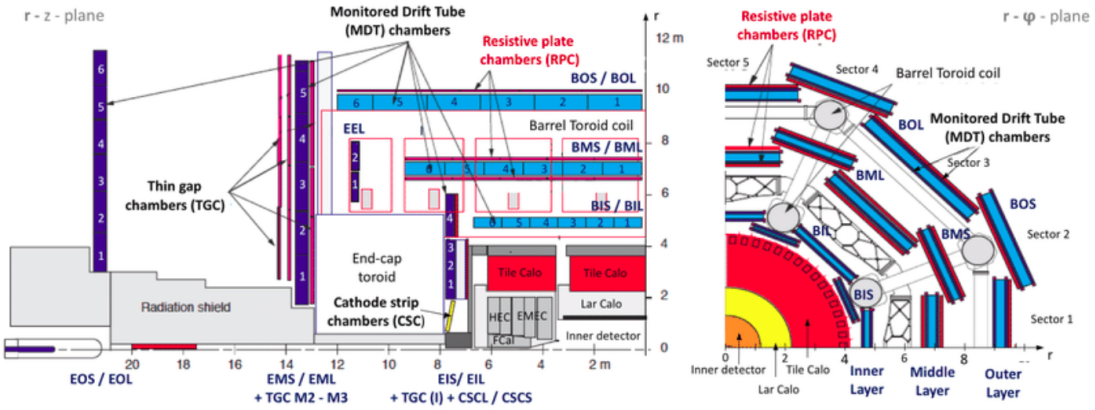


Figure 4.7.: Layout of muon spectrometer technologies in Runs 1 and 2. Left shows a cross-section displaying the end-cap technologies, and right shows a quadrant of the barrel. [53]

sist of planar electrodes with a gap divided by a thin mesh into a 5 mm drift region and a thin 120 μm amplification region to achieve better than 50 μm spatial precision [55]. I had the privilege of participating in ATLAS operations during NSW's commissioning and debut data taking, seeing first-hand the cooperation and flexible thinking necessary to achieve the designed performance of these grand particle detectors. Detector operations and my contributions are discussed further in Section 4.2, but due to my timeline, ATLAS's other systems were already optimized to behave harmoniously and operate at near-constant peak performance. The NSW upgrade is the first of the high-luminosity upgrades for ATLAS, which are discussed further in Section 4.3.

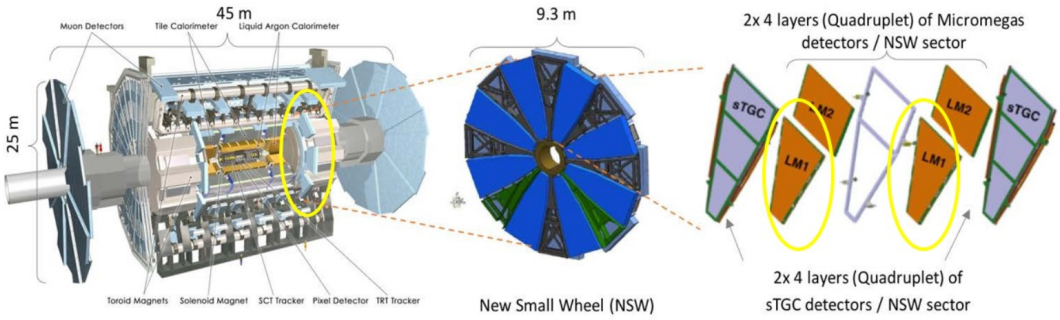


Figure 4.8.: ATLAS's New Small Wheel is not so small. This most recent upgrade to detector hardware sits in both end-caps. Each wheel holds 8 petal-shaped structures composed of sandwiched micromega and sTGC technologies. [55]

4.1.5. Magnet Systems

Momenta of charged particles can be resolved by their curvature in strong magnetic fields. As shown in Figure 4.9, ATLAS uses a central solenoid, a barrel toroid, and two end-cap toroids to bend moving charges in order to measure their charge and momentum. These electromagnets are cooled below -268^{o} C in order to maintain superconductivity in flat cables made from a composite of NbTi/Cu surrounded by aluminum [56, 57].

The central solenoid encloses the inner detector and provides a 2 T axial field using 9 km of the composite superconducting cables. Despite being only 4.5 cm thick, the solenoid supplies this large field to an encapsulated volume that is 5.8 m long and 2.5 m in diameter. It was constructed to be as thin as possible to minimize material to limit energy loss in particle measurements, and it was even possible to eliminate additional vacuum housing walls by mounting the magnet inside the vacuum vessel of the LAr calorimeter [56].

The toroidal magnets produce a magnetic field tangential to that of the solenoid, so the toroids and the solenoid are not magnetically coupled. Each of the three toroids

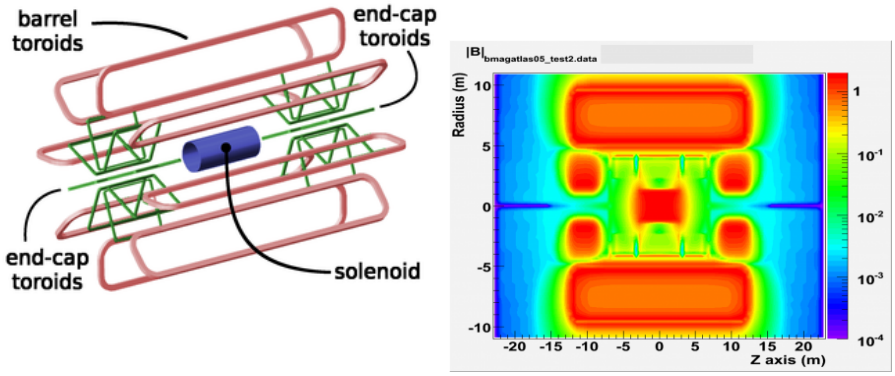


Figure 4.9.: Magnet systems of ATLAS. Left displays a schematic of the solenoid, barrel toroid, and end-cap toroid magnets as they are positioned and oriented in ATLAS [58]. Right shows a cross-section of the magnetic fields produced, demonstrating the concentrated fields in the inner detector from the solenoid and in the muon chambers from the toroids [59].

consist of 8 coils and can produce azimuthal fields up to 3.5 T to the muon spectrometers for measuring muon momenta. The barrel toroid, after which ATLAS is named, is over 25 m in length and is the largest toroidal magnet ever constructed. 56 km of cable go into the barrel toroid, along with 13 km in each end-cap toroid, collectively enough length to circle the entire LHC three times [56]. The end-cap toroids and the barrel toroid do experience Lorentz forces from one another which pull the end-cap toroids into the barrel, so robust support structures are in place to transfer the mechanical forces, as well as to supply cooling. For a muon passing through the barrel toroid, the bending power is 2-6 T m, and passing through an end-cap toroid, the bending power is 4-8 T m, allowing for precision in momentum resolution even for highly energetic muons [57].

§ 4.2. Detector Operation

4.2.1. Data Acquisition and Recording

LHC delivers 40 billion proton-proton beam bunch crossings or events to ATLAS each second, with each event stored in roughly 2 MB. An event of interest might look like the event display in Figure 4.10, which is identified as a Higgs boson decaying into two Z bosons. If CERN were to store every single collision on CDs, in three months, there would be enough CDs to stack to the moon [60]. Not only is it not feasible to store this much data, it is also not possible to write data to storage at the speed necessary to do so. Fortunately, not all of these collisions hold the interesting interactions that are the target of the ATLAS experiment. In order to filter the data into a manageable load, a two level trigger system rapidly discriminates events to record only events that have the potential to be interesting, based on detector energy deposit criteria that are optimized in coordination with various ATLAS physics teams.

The first level of filtering, known as the Level-1 trigger (L1), occurs in just 2.5 μs, shrinking the data output to just 100,000 events per second using specialized hardware electronics right on the detector components themselves. The L1 calorimeter (L1calo) is a hardware component that receives information from the liquid argon and tile calorimeters to identify events containing calorimeter physics objects, such as electrons, photons, jets, and missing energy. Similarly, L1 muon triggering hardware (L1muon) seeks muons and tau decays from the muon spectrometers. These systems send triggered event objects (TOBs) to the L1 topological trigger (L1Topo), which is able to perform kinematic cuts. The L1 central trigger processor (L1CTP) receives passing events from L1Topo and makes a final decision on whether an event should be accepted by L1 (L1A)

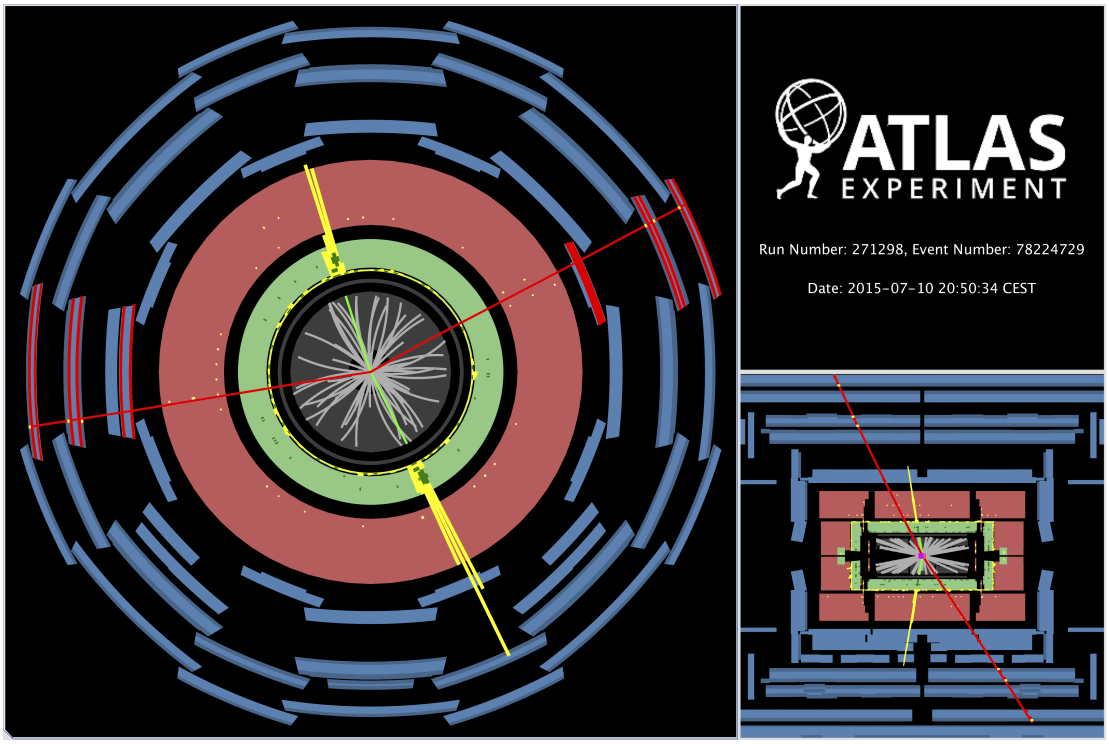


Figure 4.10.: An example of ATLAS's ATLANTIS event display, which depicts the ''hits" and energy deposits in each layer of the detector. This event, number 78224729 from run 271298 in 2015 at $\sqrt{s} = 13$ GeV, is suspected to have arisen from the production of a Higgs boson and its decay to two Z bosons. The first Z boson candidate has a mass of 94 GeV and p_T of 35 GeV, and is reconstructed from its decay to two oppositely charged muons shown in red. The other candidate has a mass of 86 GeV and p_T of 19 GeV and is reconstructed from to two oppositely charged electrons, shown as green tracks leading to energy deposits in the electromagnetic calorimeter. The gray curved tracks in the inner detector layers are from charges with $p_T \geq 1$ GeV; lower momentum tracks are omitted from this display. [61]

or rejected [62].

From there, a high level trigger (HLT), which lives on a grid system at CERN with 60,000 CPU cores, reduces the output all the way to 1000 events per second, and it does so in 200 μs. In this short amount of time, full event reconstruction is performed such that selection can be made on these reconstructed features using data from all detectors. The HLT functions on a system of algorithms using an early-rejection mechanism, such that selections with the largest amount of rejection are performed first to save CPU time. Recorded events retain an HLT flag that physics analyzers will use to select data subsets which contain the event objects of interest for their analysis final state signature [63].

4.2.2. ATLAS Operations Orchestra

The LHC cost roughly 3 billion euros for the construction of the machine alone, not to mention its annual energy consumption which is estimated to be nearly 1 TWh per year or a third of the energy produced in the entire canton of Geneva [64]. This investment from nearly 40 participating nations demands that the collider's outputs not be wasted. ATLAS takes its recording efficiency very seriously, and when the LHC is able to provide beams, ATLAS must be ready to operate and receive those collisions. In order to achieve close to 100% recording efficiency of the integrated luminosity delivered by the LHC, the ATLAS Control Room (ACR) is filled with shift-workers 24 hours a day and 7 days a week during operation season, which typically runs from April through November. ATLAS achieved 94.1% recording efficiency for data taking in the year 2024.

Collaboration members fill 8 positions in the ACR, commonly referred to as desks. Three of these desks correspond to specific subdetector sections: (1) inner detector sub-

systems, (2) calorimeters, and (3) muon spectrometers. Three more desks correspond to the general detector data taking: (4) run control ensures data taking is smoothly running, the (5) trigger shifter configures and monitors the triggers in use, and the (6) data quality shifter ensures that all detector subsystems are registering acceptable readings. Finally, there are two shift leaders: (7) a shift leader devoted to matters of safety and (8) a shift leader who communicates with ATLAS leadership and supervises activities at the other desks as the conductor of the orchestra. Together, the operators at these desks work in harmony for 8-hour blocks at a time, quickly adapting to problems that arise.

Urgent problems will make themselves apparent either to the detector control system (DCS) or to the data acquisition monitoring system (DAQ). Physical problems with a component of the detector appear in the DCS and may include currents, voltages, temperatures, or dew points outside of their targeted range. ATLAS's DCS monitoring system operates as a finite state machine (FSM), and any error or warning state propagates up the chain to ATLAS as a whole where it will not go unnoticed [65]. An example status of this DCS monitoring system is shown in Figure 4.11, as it would be seen in the control room. Subdetector experts optimize limit thresholds to distinguish between tolerable and dangerous detector behavior. Problems related to the data readout and its quality from these subdetectors appear in the DAQ interface and can range in severity from single events missing fragments of information to entire system readouts malfunctioning. It is the duty of the operators in the ACR to recognize a problem impacting the systems under their desk's and diagnose its severity, in order to call an appropriate expert and communicate the observed concern along with the status of operations in the ACR. Specialized system experts are on-call day and night to quickly rectify issues and return the detector to a tolerable state.

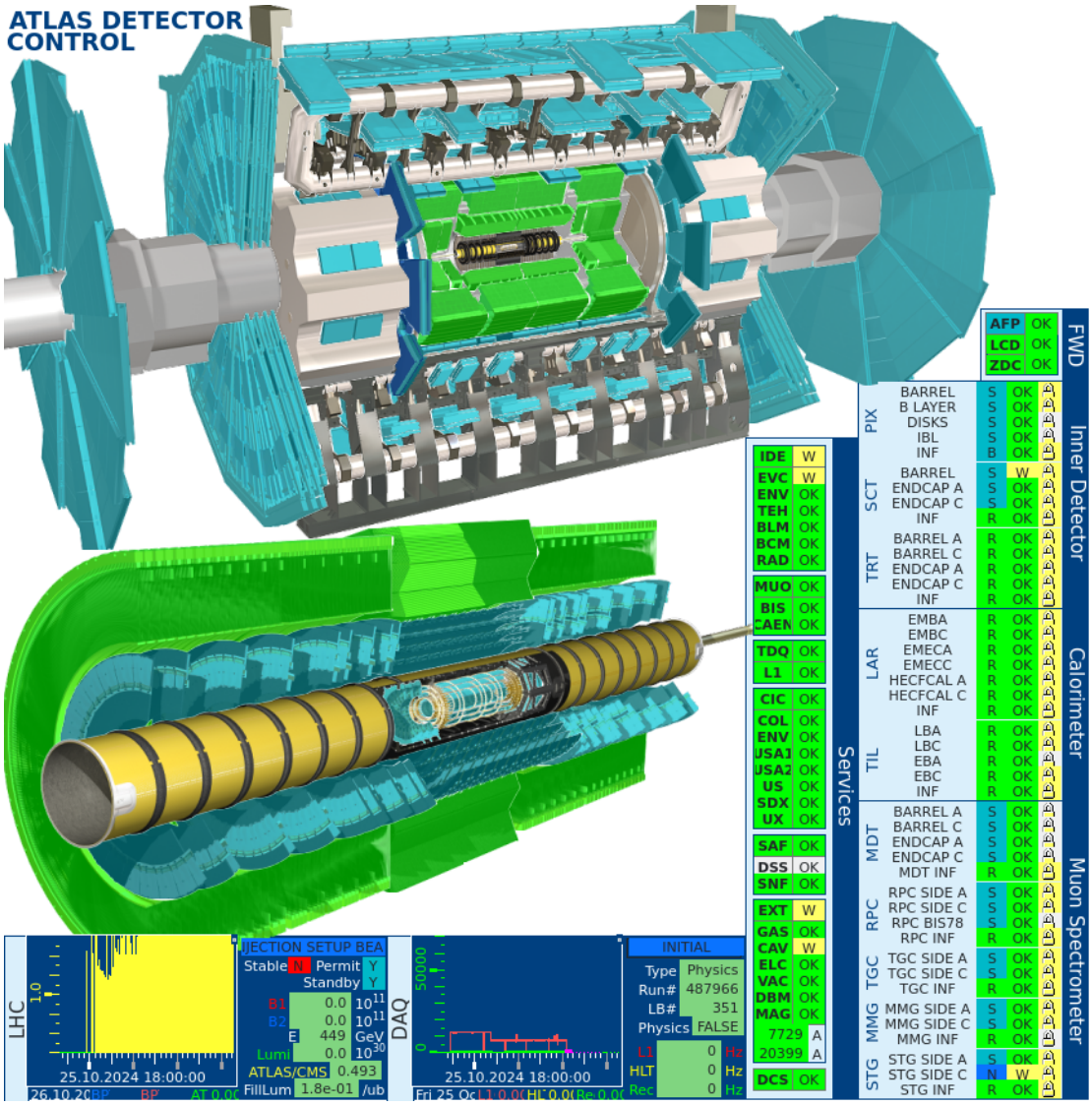


Figure 4.11.: Example status of ATLAS's finite state machine, which is observed in the control room 24 hours a day and 7 days a week, during LHC running periods, to monitor the safety of the detector hardware components. At the time of this FSM status, stable beams were not being delivered, so several subsystems are in standby (blue S). [65]

4.2.3. Inner Detector Pixel Operation

A closer look at detector operations requires a specific subsystem example. ATLAS's pixel detector operates nearly 2000 individual modules, each with high voltage sensor bias and low voltage power supplies, thermocouples, and multiple front-end readout chips. These modules wait in standby until stable beams are reached, at which point PIX transitions to a state ready for data-taking by ramping the high voltage sensor bias and powering the signal pre-amplifiers. With PIX ready, traversing charged particles induce a voltage in the sensor, and when this voltage passes a specified threshold, the readout front-ends store the time spent over this threshold (ToT) [66]. Careful calibrations are performed to achieve a uniform response across the detector, such that a minimum ionizing particle (MIP) registers the same ToT in all pixels. There are a number of parameters in the front-end electronics that can be tuned to minimize the spread in module response. To tune these parameters, a fake test charge with a predefined amplitude and timing is injected into the input of the signal pre-amplifiers, simulating charge deposition in the sensor, and the current threshold of each pixel is adjusted to homogenize the time spent over threshold in each channel. By design, the time over threshold reading is proportional to the charge deposited in that pixel sensor. These calibration scans are performed during detector downtime when no beam is present [67].

During stable beams, a number of problems can arise during regular operations of the pixel detector. For example, a spike in current will occasionally trigger an interlock, which will force a subsection of PIX to return to standby, turning off high voltage and pre-amplifiers in an effort to prevent any detector damage. Another frequent yet urgent issue occurs when ATLAS's trigger frequency gets close to the resonant frequency of the IBL wirebonds and initiates a fixed frequency trigger veto (FFTV). In the case of

an FFTV, all of ATLAS's data acquisition is halted until the cause of the trigger spike is handled. Occasionally, single event upsets will occur when a charge hits the readout electronics in such a way that it causes a bit flip, impacting the transmitted data. This can only be fixed by a reconfiguration or reset of the front-end, but luckily no external intervention is necessary because PIX performs an event counter reconfiguration every 5 seconds automatically during regular operation.

Sitting closest to the points of interaction, PIX receives the highest radiation dose of any of ATLAS's subsystems. PIX is expected to receive over 10^{15} MeV n_{eq}/cm^2 in non-ionizing radiation before it is replaced in the next long shutdown. Heavy particles and nuclei in flux cause damage to the sensor bulk by displacing silicon nuclei and leaving defects, which can modify the readout of MIP's. These defects can impact the electric field in the sensor and alter its depletion voltage. To stay ahead of the radiation damage, the high voltage sensor bias settings are checked twice per year and increased each year of data taking to maintain charge collection efficiency. This alone is not enough, however, and some radiation damage effects in the data are inevitable [68]. Intense effort has gone into the modeling of radiation damage effects such that Monte Carlo simulated data can reflect similar detector impacts.

4.2.4. My Contributions to Inner Detector Pixel Operation

I arrived at CERN in the summer of 2022, just in time to celebrate the start of Run 3 with an enthusiastic community. I quickly trained to take shifts in the control room at the inner detector desk, where I monitored the status of the PIX, SCT, and TRT subsystems, as well as the beam condition monitor (BCM) which closely watches the LHC beam conditions and luminosity nearest to interaction in ATLAS [69]. During smooth operation,

one almost feels bored, but soon a set of issues would arise and carry along intense pressure and stress, re-instilling an appreciation for the quiet periods. After many shifts day and night, the ATLAS jargon grew more familiar, and my face slowly became known. The pixel team took notice of my attentive nature, and due to their dearth in manpower, I was invited to join the squad as a first-level on-call expert, a position known as pixel run manager (PRM).

Suddenly, I was the person receiving the frantic calls from the ACR rather than delivering them, and the curtain was lifted. As PRM, for one week every six or so, I kept a CERN phone with me at all times, answering day and night, and I also logged the operations endeavors undergone that week. With the help of higher experts, I handled tripped current channels, failed stable beams transitions, glitched optical readouts, and many other issues in order to preserve the quality of ATLAS data taking. For many months, it felt as though the same problem would never come up twice, but I slowly gained more confidence and independence with the tools and became intimately familiar with the pixel system. Outside of stable beams, I performed threshold tunings and conducted investigatory scans to diagnose buggy modules. Eventually, I no longer felt as novice, and I welcomed new members and aided in training them in the art of keeping the ATLAS pixel detector safe.

§ 4.3. Detector Upgrade for High-Luminosity LHC

4.3.1. High-Luminosity Upgrade Requirements of ATLAS

The high-luminosity upgrade of the LHC will mark record-breaking pp-collision rates, which pose new challenges to the data collection and reconstruction capabilities of AT-

LAS. The luminosity will jump from 2×10^{34} protons per cm^2 per second to 7.5×10^{34}, bringing a mean number of interactions per bunch crossing or pileup to 200 from roughly 60 in the running conditions of Run 3, and increasing the L1 trigger rate from 100 kHz ten-fold to 1 MHz [70]. The current detector will be almost entirely replaced in order to deal with the additional pileup, as well as to withstand the immense radiation doses that the new running conditions will deliver. The endeavors undertaken to prepare ATLAS for this major change are detailed in Section 4.3.2, and particular attention is given to the upgrade of the inner tracking layers in Section 4.3.3, on which I have made contributions as discussed in Section 4.3.4.

4.3.2. Overview of Upgrade Projects

During the next long shutdown, the ATLAS detector will undergo a complete transformation. The inner detector will be entirely replaced by a new Inner Tracker (ITk), discussed further in the next section. New calorimeter electronics, forward detectors, muon chambers, and a new high granularity timing detector (HGTD) will be constructed and installed in place of the current systems [70]. For the calorimeters, the electronics on and off the detector will be upgraded to be able to tolerate the higher radiation doses, and these electronics will support a continuous readout rate of 40 MHz. Forward calorimeter detector components will be entirely replaced [71]. Entirely new muon spectrometers are being built for the barrel and inner end-cap regions to join the already-installed NSWs, and some chambers which have already been constructed show good performance and precision as designed [72]. The completely new HGTD is designed to have timing resolution on the order of 30-50 ps, which is essential for disentangling interactions of interest in the anticipated pileup conditions [73]. Figure 4.12 demon-

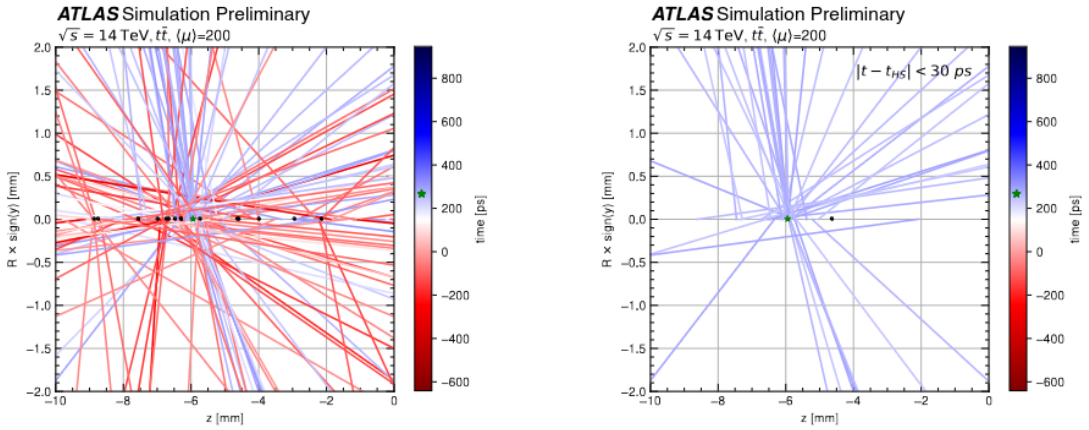


Figure 4.12.: Zoomed-in longitudinal view of a $t\bar{t}$ event with pile-up $\langle\mu\rangle = 200$. On the right, a timing cut is applied that removes tracks which diverge in recorded time from the hard-scatter tracks to the primary vertex by more than 30 ps. No timing cut is yet applied on the left. This simulation demonstrates the impact in pile-up rejection that the new High Granularity Timing Detector (HGTD) may achieve. [75]

strates the profound impact that 30 ps timing resolution can have on pile-up rejection, as the distance between pile-up interactions becomes on the same order of magnitude as the current resolution of the impact parameter of the primary vertex. The HGTD will accomplish this feat with low-gain avalanche diode (LGAD) sensors, which are constructed from very thin silicon substrates, such that MIPs induce fast signal pulses that are internally amplified [74]. Four layers of these silicon modules will be positioned in the high-pseudorapidity regions, $|2.4| < \eta < |4.0|$, not only to provide timing resolution, but also to provide a precise bunch-by-bunch luminosity measurement.

In addition to all these detector upgrades, the trigger and data acquisition system (TDAQ) will be completely revamped. The data throughput will increase from 0.2 TB/s to 4.6 TB/s, and the event size will increase to 4.6 MB/event. The hardware trigger, which will replace L1, will operate at 40 MHz, accepting events at a rate of 1 MHz [76].

Accelerated event reconstruction algorithms will already be available at this early filtering stage in the new Global Trigger, also passing information to the final decision-making Central Trigger Processor [76]. The software trigger (Event Filter), replacing HLT, will filter the final data collection down to a rate of 10 kHz or five times the output rate of ATLAS during Run 3. The Event Filter will make use of offline-like reconstruction algorithms, using FPGAs and GPUs to accelerate processing speeds. New algorithms are being designed using neural networks, such as a fast tracking algorithm which already shows similar performance to current Run 3 offline track reconstruction [77].

ATLAS will be ready to take almost 3 ab^{-1} of pp-collisions at $\sqrt{s}$ = 14 TeV over a decade of data-taking with the HL-LHC [70].

4.3.3. Inner Tracker Upgrade

ATLAS's inner detector will be decommissioned and replaced by a new Inner Tracker (ITk) consisting entirely of silicon detector pixel and strip modules. ITk will expand the tracking coverage of ATLAS to $\eta < 4$, and will provide similar tracking performance at $\langle\mu\rangle = 200$ to that which is achieved by the current ID at $\langle\mu\rangle = 40$. Production of ITk pixel and strip modules is underway, and ITk is expected to be assembled in 2027 for installation in the ATLAS cavern in 2028. The designed layout of ITk is shown in Figure 4.13. This new tracker will provide 13 hits to a particle traversing with $|\eta| < 2.6$ and 9 hits in the $2.7 < |\eta| < 4$ region [78].

The ITk pixels detector will consist of 5 billion individual channels using two sensor technologies: planar and 3D. The planar sensors will be between 100 μm and 300 μm thick, with the thinnest sensors closer to the beam-line since they are more radiation tolerant. Rather than the n^+-in-n type sensors used in the current ID, ITk's planar sensors

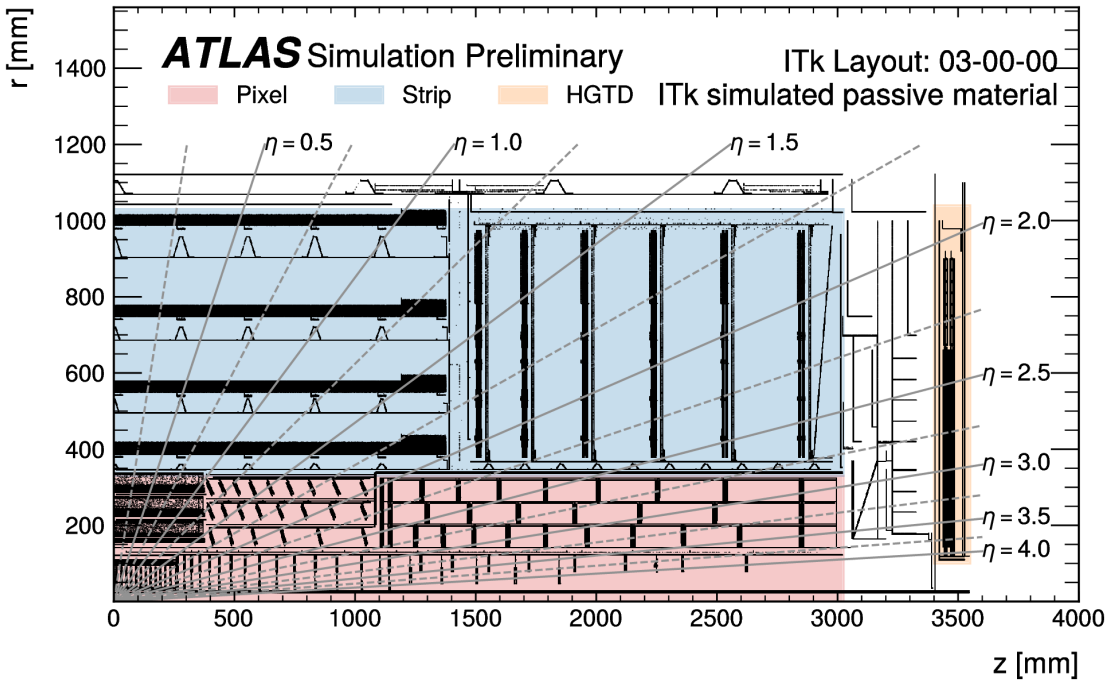


Figure 4.13.: Layout of Inner Tracker as designed. ITk Pixels are shown in red, and strips are shown in blue. These new inner layers of ATLAS will have full coverage up to a pseudorapidity of 4.0. [78]

will be n^+-in-p type, which produce faster signals, are more easily produced, and are more radiation hard [79]. In planar sensors, the readout electrodes sit outside the sensor, whereas in 3D sensors, the electrodes are inserted into the sensor bulk, which makes for smaller drift lengths and therefore improved radiation hardness. For this reason, 3D sensors will be placed in the innermost layers of the pixel tracker, and these layers will be designed to be replaced after 2000 fb^{-1} of data-taking. These 3D pixels will have geometry of 50×50 μm^2 and a power dissipation of 10 mW/cm^2, which is another critical performance element [80].

The ITk strips detector will consist of 4 barrel layers and 12 end-cap disks. Compared to SCT, the lengths of silicon strips in ITk are shorter to increase tracking granularity

and readout speeds, and strips in the inner barrel layers will be even shorter compared to those in the outer barrel layers and end-caps. ITk strips will extend further out than SCT, filling the space formerly occupied by TRT, for a total surface area of 165 m^2. The ITk strips detector will also have ten times as many channels as SCT, due to the shorter strip length and increased area [80].

4.3.4. My Contributions to the Inner Tracker Pixel Upgrade

Production of such a large-scale detector requires the multi-faceted participation of many institutions and individuals. The Santa Cruz Institute for Particle Physics (SCIPP) at UCSC has a role in many aspects of the ATLAS detector upgrade for HL-LHC. In particular, we perform electrical testing on fully-assembled quad modules, which consist of four readout chips bump-bonded to a single sensor and are designed for the outer layers of the inner system, barrel, and end-caps of ITk Pixel. During my time in Santa Cruz before moving to CERN, I served as a member of SCIPP's ITk pixel module testing team, as we were first building a functioning test-stand. In the cleanroom, I installed chiller lines, dry air circulation, radiation protection, and even shelving to prepare our test-stand for early stages of qualification. Controlled thermal cycling in the range of -55°C to 60°C is one such requirement for module testing, and the first full cycle achieved at SCIPP is plotted in Figure 4.14. Through trial and error and chiller geysers, I achieved the capabilities required for thermal cycling and performance testing of these quad modules, and I developed the environmental monitors to ensure the modules' safety during such tests.

During this time, I also completed my ATLAS qualification task, which allowed me to be a signing author of the collaboration. For this task, I designed and built a graphic

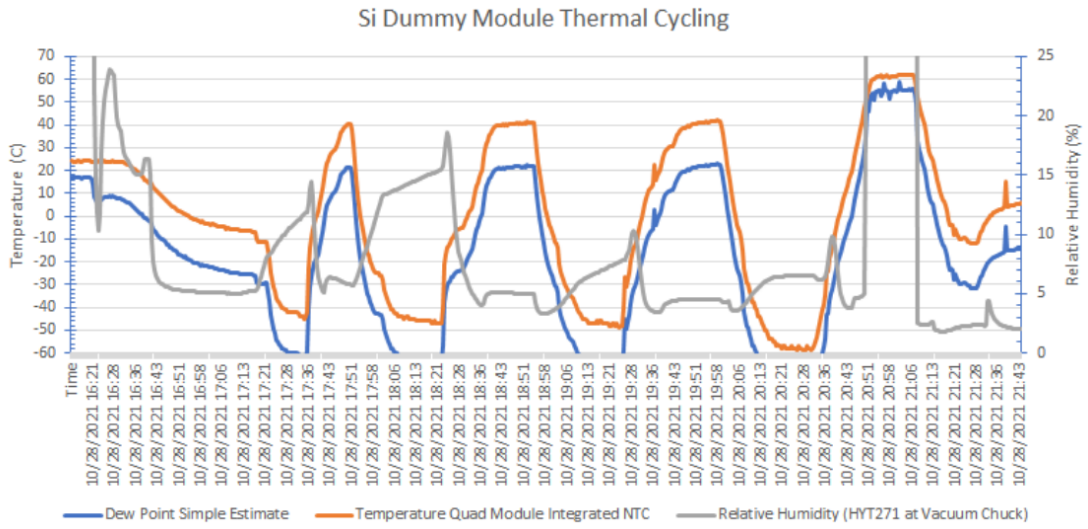


Figure 4.14.: Thermal cycling performed on a prototype ITk Pixel quad module at SCIPP. The temperature of the module (orange) demonstrates several cycles in the range -45°C to 40°C, followed by one large thermal cycle from -55°C to 60°C. The humidity (gray) is constantly monitored to ensure the dew point (blue) remains below the module temperature to avoid condensation.

user interface (GUI) for controlling lab hardware using a python backend dependent on a package called LabRemote [81]. This GUI is available broadly to other testing sites, and it has inherent support for chillers and power supplies as well as room for easy customization to fit a site's specific needs of automation [82].

After leaving Santa Cruz's cleanroom behind, I continued my work with ITk pixels by volunteering to test irradiated and non-irradiated pixel modules in synchronous operation at CERN's test beam site. There, we measured the performance of up to 4 modules at once, sandwiched between ''telescope" reference sensors, such that we could check how these modules maintain synchronized timing and position correlations. From working with one module at a time at SCIPP to several in sequence at the test beam to an entire detector in synchrony in ID operations, I have truly gained a rounded perspective on silicon detectors, and I am quite grateful for these opportunities.

Chapter 5.

Analysis Background and Motivations

In Section 2.3, the uniqueness and importance of the Higgs boson were made clear. While the particle was theorized to explain the massive nature of the weak vector bosons, the source of the masses of all fermions came essentially for free. To date, everything measured about the Higgs has fit perfectly into the simplest model, in which only one scalar complex doublet particle exists, giving mass to all particles. Observations of interactions between the Higgs boson and other particles support the model in which a particle's mass is directly proportional to that particle's coupling strength to the Higgs field. This is demonstrated by the straight line running through all the combined measurements of the Higgs couplings in Figure 5.1, which is why the particle's masses are attributed to the Higgs boson. Additionally, data used to calculate the rate, or cross-section, of Higgs production have been taken from proton-proton collisions at center of mass energies of 7, 8, 13, and 13.6 TeV. The standard model predicts that the cross-section of Higgs production increases linearly with center of mass energy, and the latest measurements are consistent with this prediction, as can be seen in Figure 5.2. Any departure from the Standard Model Higgs behavior must be small as it has not yet been

experimentally observed, and so precision measurements and rare searches are the evolution of the game.

Within the volume of ATLAS, the LHC has produced nearly 30 million Higgs bosons [83, 85], of which only about 0.3% are estimated to be accessible in the coverage of the detector and identifiable in the experimental reconstruction, so careful analysis techniques are extremely important to capitalize on those collected. The method by which the Higgs boson is further studied is by taking all the ways in which the Higgs can be produced and can decay, and dividing the large pie into small, extremely specific slices. Collectively, hundreds of ATLAS members study the Higgs boson, each specializing in a particular signature in most cases. Some recent endeavors of my ATLAS Higgs Physics Working Group colleagues are elaborated in Appendix B. In Section 5.1, the possible ways in which the Higgs boson can be produced and decay are detailed to set the landscape. Then, in Section 5.2, the specific signature of interest for this dissertation is motivated: vector boson fusion Higgs production in association with a high energy photon with the Higgs decaying to a bottom-quark pair.

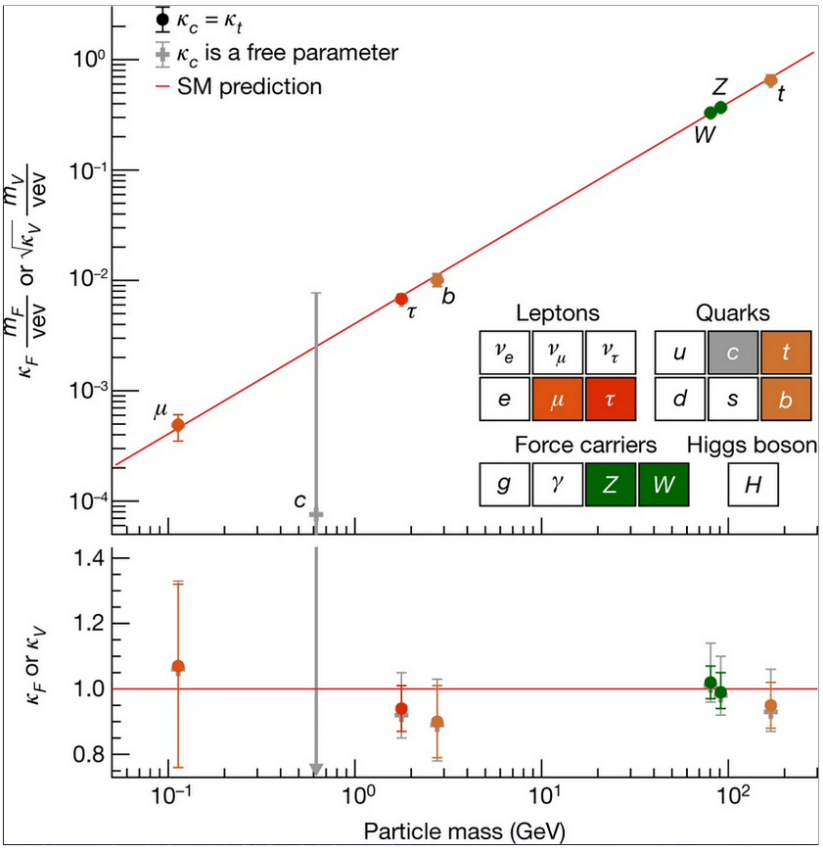


Figure 5.1.: Combined measurements of the interaction strength or coupling between the Higgs boson and the other standard model particles. Precision measurements have been performed for the third-generation fermions as well as for the massive vector bosons, and observations are gaining precision for some second-generation fermions like the muon and charm quark. If the single Higgs theory is supported as the source of the masses of all massive particles, then the interaction strength of a particle with the Higgs should be directly proportional to the particle's mass, forming a straight line on this plot. So far, observation supports this theory. [83]

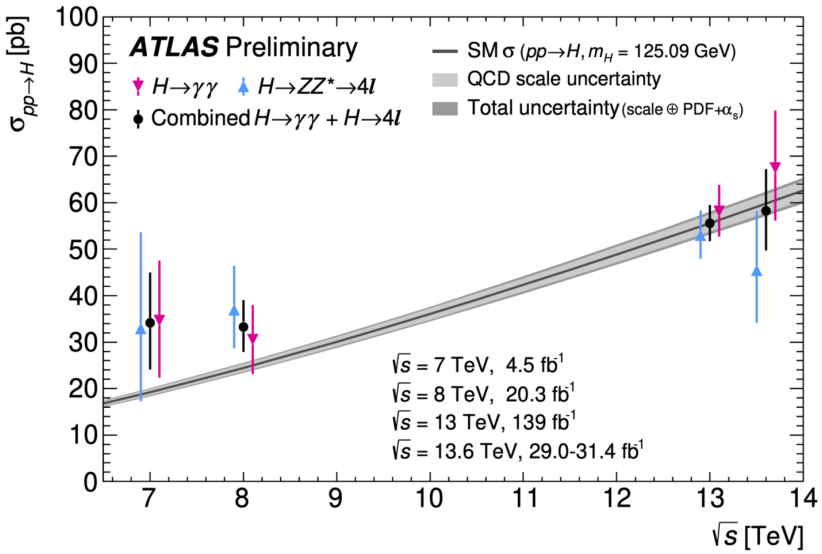


Figure 5.2.: Measured values of the Higgs cross-section from combined results in the most sensitive decay modes at center of mass energies 7, 8, 13, and 13.6 TeV. Standard Model predicts an increase in cross-section, σ, with center-of-mass energy, $\sqrt{s}$. [84]

§ 5.1. Higgs Production and Decay Modes

5.1.1. Production Modes

In proton-proton collisions, there are four main methods of producing a Higgs boson which account for over 99.95% of the Higgs bosons produced at the LHC. The Feynman diagrams for these four production modes are displayed in Figure 5.3. Gluon-gluon fusion (ggF) is the dominant production mode in the inclusive phase space, accounting for roughly 87% of Higgs produced in the LHC. In ggF Higgs production, the primary method of Higgs creation from the massless gluons is through the mediation of a virtual top quark loop, and at leading order, no additional final state products are created in this process. The next dominant production mode is vector boson fusion (VBF), in which two weak vector bosons, either W^+/W^- or ZZ radiating from quark initial states, fuse to create the massive neutral Higgs. VBF accounts for around 7% of Higgs produced, and in addition to the Higgs, it results in two final state quarks with large angular separation which are observed as ‘‘VBF jets” with high di-jet invariant mass. Next in prevalence, vector boson associated Higgs production (VH), in which quark fusion produces a weak boson which radiates a Higgs (a process known as Higgsstrahlung [86]), can be attributed to 4% of Higgs bosons. The final state signature of VH production is noted for having the decay products of a weak vector boson in addition to those of the Higgs. Completing the four dominant modes, Higgs production in association with a top or bottom quark pair (ttH/bbH) contributes another 1-2% of the Higgs bosons. Here. gluon splitting produces two top or bottom quark pairs, and one partner from each pair fuses to create the Higgs, leaving the other two top/bottom-quark partners to propagate to the final state signature [83]. The hierarchy of these production modes is

further displayed as their production cross-sections evolve with center of mass energy in Figure 5.4.

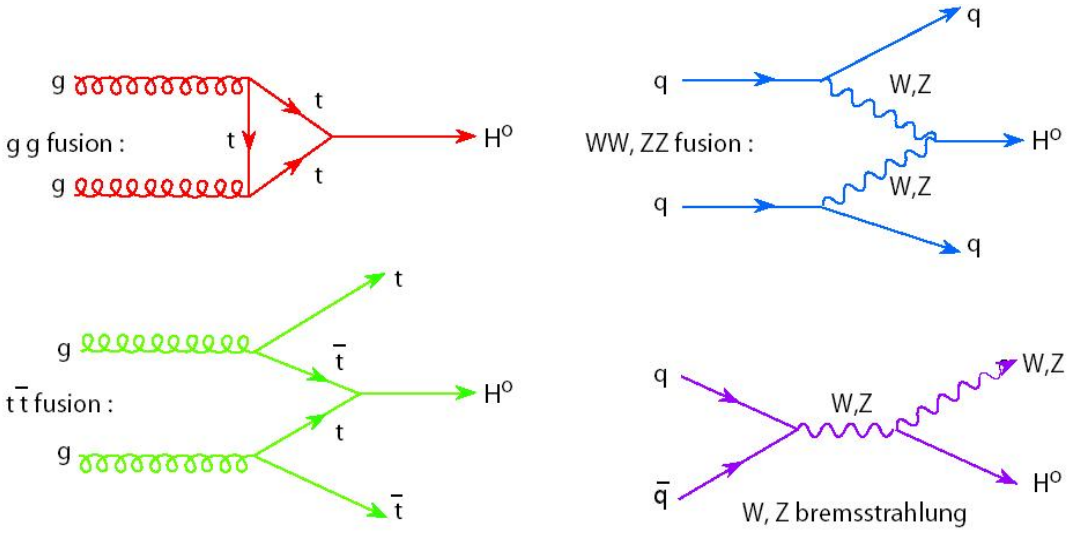


Figure 5.3.: Feynman diagrams of the 4 main Higgs production modes in proton-proton collisions. [87]

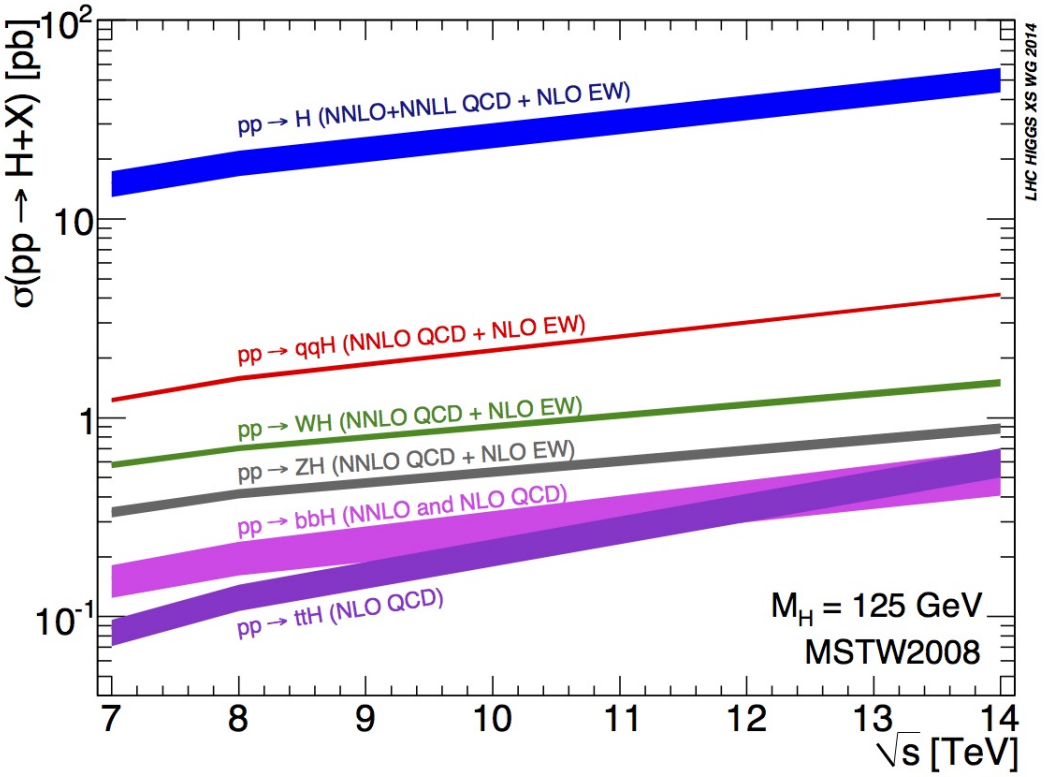


Figure 5.4.: Higgs production cross-sections as a function of center of mass energy. The dominant mode in blue represents gluon-gluon fusion. Second in red is VBF Higgs production, followed by VH production in green split between the associated vector bosons produced. Finally, Higgs production in association with a bottom/top quark pair are shown in pink/purple. [88]

5.1.2. Decay Modes

The ways in which the Higgs boson decays are equally important in determining what to look for in collected ATLAS events. Figure 5.5 shows the predictions of relative decay rates, or branching ratios, of the Higgs as a function of Higgs mass. For a mass of 125 GeV, the dominant decay mode is a bottom quark pair, since the bottom quark is the heaviest fundamental particle (largest coupling to the Higgs) that still is less massive than half the Higgs mass, such that two real bottom quarks can be produced without

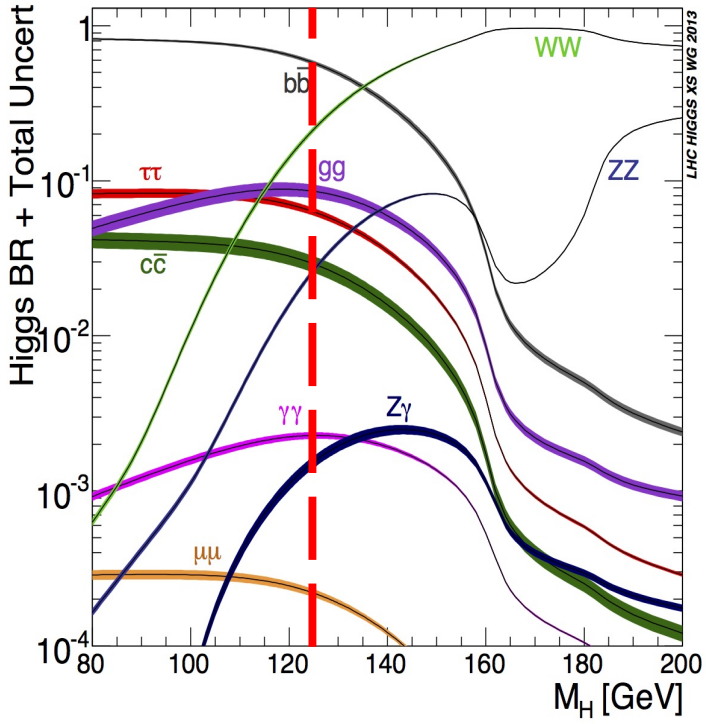


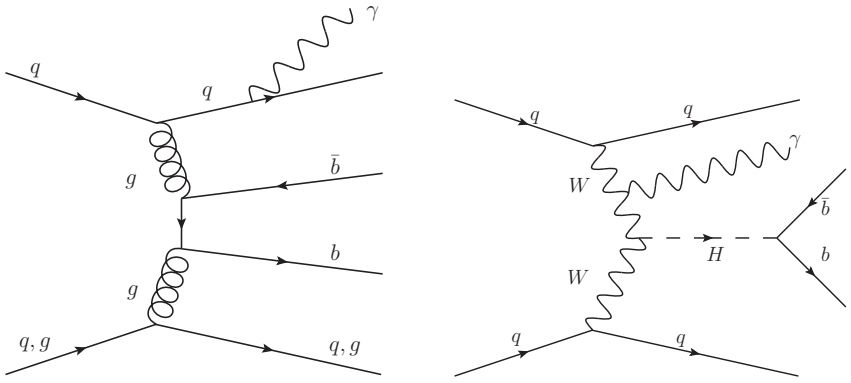
Figure 5.5.: Predicted branching fractions of Higgs boson as a function of Higgs mass. The known Higgs mass of 125 GeV is marked with red dashed line, fixing the hierarchy of Higgs decay modes, with the dominant decay mode being the bottom quark pair. [88]

violating energy conservation. Though the Higgs decays to bottom quarks almost 60% of the time, the decay modes most sensitive to Higgs measurement are actually $H \rightarrow \gamma\gamma$ (branching ratio of 0.2%) and $H \rightarrow ZZ \rightarrow 4l$ (branching ratio <2%) due to the precision of photon and lepton reconstruction and identification in the ATLAS detector. [88] It was in these less dominant decay modes that the Higgs boson reached its 5σ discovery significance.

§ 5.2. Motivation for the Search for VBF Higgs to Bottom-Quark Pair Production in Association with a Photon

Since its discovery in 2012 [89, 90], the scalar Higgs boson has been studied extensively through all the expected main production modes and through many of its decay channels. The analysis to which the rest of this dissertation will be devoted focuses on the production of a Higgs boson in association with a high energy photon, with the Higgs decay products being a pair of bottom quarks. The Higgs boson most frequently decays to a pair of b-quarks, so this is a powerful decay channel to study. However, this decay channel suffers from a large multijet non-resonant QCD background population, clouding the search. By requiring a final state with an associated high-energy photon, a destructive interference effect reduces this non-resonant background with Feynman diagrams visible in Figure 5.6a, making this an interesting final state signature to target [91].The photon also conveniently provides a clean signature on which to trigger.

With the requirement of an associated photon, VBF supplants ggF as the dominant production mode as demonstrated in Figure 5.7. This results from the increased number of electroweak vertices in VBF as well as from Furry's theorem [92], which implies that a single photon cannot arise from a closed fermion loop with an odd number of vertices. The requirement of a high invariant dijet mass further purifies the signal phase space toward VBF production, with vector boson associated (VH) Higgs production contributing less than 5% of the signal. An added benefit of the photon requirement is that a similar destructive interference effect to that which reduces the non-resonant QCD background



(a) Non-Resonant $b\bar{b}\gamma jj$ Background (b) $H\gamma jj$ Signal

Figure 5.6.: Feynman diagrams for non-resonant multijet background (a) and $H\gamma$ signal (b).

suppresses ZZ-fusion. This final state signature provides a unique opportunity to study WW-fusion Higgs production and the HWW interaction more closely.

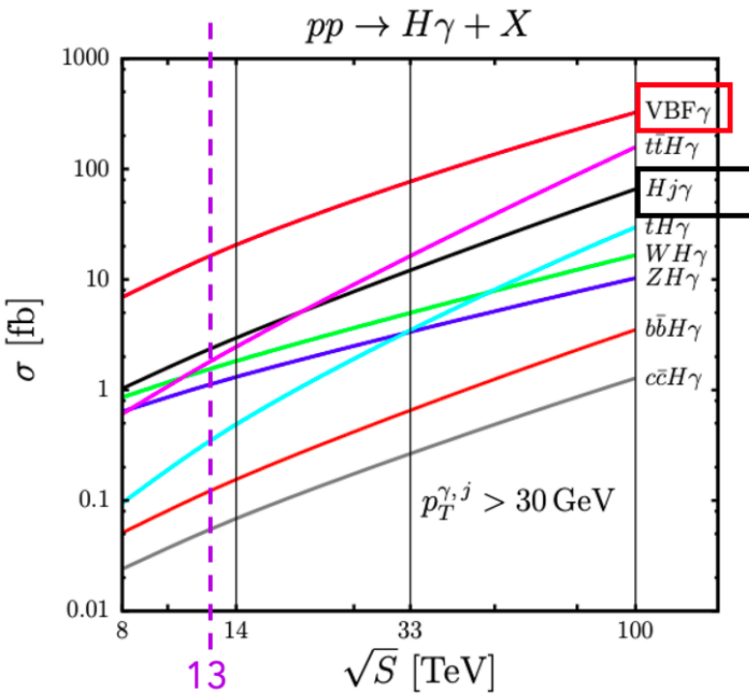


Figure 5.7.: $H + \gamma$ production mode cross-sections as a function of center of mass energy, with the Run-2 energy of 13 GeV highlighted with a pink dashed line. VBF is the dominant production mode when an associated photon is required. [93]

§ 5.3. Previous Work in this Channel

A previous full Run 2 analysis for this signature was performed and measured an observed (expected) signal strength of $\mu_H = 1.3 \pm 1.0$ (1.0 ± 1.0), with an expected significance of 1.0σ [94]. Compared to the previous effort, huge strides have been made in the multivariate analysis techniques employed as described in Section 8, replacing the previously-used boosted decision tree (BDT) classifier with a densely-connected neural network (DNN) able to handle the complex correlations of newly optimized kinematic input variables [95]. This novel effort also flips the script on the fitting approach that was adopted previously, extracting the signal measurement from the classifier output score

distribution directly and utilizing the di-b-jet invariant mass (m_{bb}) to define regions of purity, rather than fitting to the m_{bb} distribution like before. These advancements combined reduce the signal strength uncertainty allowing for improved signal sensitivity.

Chapter 6.

Monte Carlo Event Simulation

Predicting what could and should appear in the ATLAS data is a tricky endeavor. Monte Carlo (MC) event simulation uses computer-generated models to mimic the particle collisions and detector response to get a best estimate of what is expected in the experiment's data [96]. MC simulations play a role in nearly every phase of an analysis. They are used to optimize event selection by predicting the acceptance rate of the signal and rejection rate of various backgrounds. They are used to validate data and calibrate event selection efficiencies and even object reconstructions. MC events are often used to model backgrounds and fit the data, as well as to evaluate systematic uncertainties, from both theory and experimental sources. These simulations are even used in the design of the experimental detector itself, such as in setting standards for coverage and component specifications for optimal acceptance and efficiency. In later chapters, the importance of Monte Carlo event simulations will be self-evident as it continues to arise in all these roles for the VBF $H(\rightarrow b\bar{b})+\gamma$ analysis. Here in this chapter, an overview of the process of Monte Carlo simulation is given in Section 6.1, a bit of focus is granted to the methodology of MC event simulation within ATLAS in Section 6.2, and specific

details of the simulations used for the particular objectives of this analysis in Section 6.3.

§ 6.1. Overview of Monte Carlo Simulation

For high energy physics specifically, the the calculation of a process cross-section depends on the summation of all possible ways in which such a process can happen. For example, all of the possible initial partons that could contribute to the final state signature must be considered in the cross-section of a process $ab \rightarrow X$ in a proton-proton collider as in Equation 6.1.1.

$$\sigma = \sum_{a,b} \int_0^1 \int_0^1 dx_a dx_b PDF_a^{h_1}(x_a, \mu_F) PDF_b^{h_2}(x_b, \mu_F) \int d\sigma(ab \rightarrow X; \mu_F, \mu_R) \tag{6.1.1}$$

In Equation 6.1.1, the $PDF_a^{h_1}$ stands for the parton distribution function or the fraction of momentum of the original hadron h_1 carried by the quark or gluon parton a. The factorization and renormalization scales μ_F and μ_R, described further in Appendix A, are free parameters that must be chosen, and the uncertainty on that choice is importantly taken into consideration among many modeling systematics [97].

The contribution of each possible path in the total phase space to the cross-section of the general process is weighted by its action as described in Chapter 2. Additionally, as described in Chapter 2, cross sections can only be evaluated perturbatively in powers of field coupling as in Equation 6.1.2.

$$\sigma = \sigma_0 + \alpha_s \sigma_1(\mu_R) + \alpha_s^2 \sigma_2(\mu_R) + \ldots \tag{6.1.2}$$

In Equation 6.1.2, the tree-level estimate σ_0 is given a first order correction σ_1 from all

next-to-leading-order (NLO) diagrams, a second order correction σ_2 from next-to-next-to-leading-order diagrams, and so forth. The structure of the cross section calculation is independent of process and can therefore be written in the format of computer code allowing for MC simulation [98]. The calculation components can be defined in generality down to a few free parameters, such as particle mass and numbers [97].

Thus, two important ingredients are needed for an MC simulator, (1) a theory model and (2) a set of process-independent instruction building blocks [98]. With regards to (1), the model that is of most use in the case of this physics search is the Standard Model, but MC simulation is also used to simulate physics beyond the Standard Model using both model-specific and model-agnostic techniques. The matrix element calculations, which are discussed in further detail in Appendix A, are very good at handling scenarios involving hard large-angle emissions, but fail in cases of soft or low-angle emissions where infrared and collinear divergences are exhibited. Going from the hard final state signature elements to the sprays of hadrons observed in the detector components, there is a regime in which it can be more probable to have an added emission than not ($\sigma_{k+1} \not\ll \sigma_k$), and it is the responsibility of the parton shower program to perform an evolution of momentum transfer down to the scale of hadron confinement [99]. Effectively, for every gluon or quark parton, the parton shower algorithm is randomly selecting whether it will split off another quark or gluon from the probability landscape that comes from the conservation of event kinematics. The splitting probability is given by the DGLAP evolution equations and for a parton with energy between q^2 and $q^2 + dq^2$ to have an emission of energy Q_0^2, the probability is given in Equation 6.1.3.

$$d\mathcal{P}_i = \frac{\alpha_s dq^2}{2\pi q^2} \int_{Q_0^2/q^2}^{1-Q_0^2/q^2} dz \mathcal{P}_{ji}(z) \tag{6.1.3}$$

The functions $P_{ji}(z)$ describe the collinear splitting of parton i into parton j with momentum fraction z and are given by the Altareli-Parisi equations. For example, Equation 6.1.4 gives the splitting probabilities for a quark or gluon to create an additional collinear gluon.

$$\mathcal{P}_{q\to qg} = C_F \frac{1+z^2}{1-z}; \mathcal{P}_{g\to gg} = C_A \frac{(1-z(1-z))^2}{z(1-z)} \tag{6.1.4}$$

Equation 6.1.4 gives the master equations which every parton shower generator applies recursively until reaching a low-energy cutoff scale [99].

Factorization of cross section calculations is a necessary component to the methodology of separating the hard scatter and parton shower processes. As shown in Equation 6.1.5, the cross-section of the final state F with an additional parton (F+1) is a multiplicative factor times the cross section of the final state F with no added parton.

$$d\sigma_{F+1} = 4\pi\alpha_s(\mathcal{M}_{\mathcal{F}\to\mathcal{F}+\infty}\frac{ds_{a1}}{s_{a1}}\frac{ds_{b1}}{s_{b1}} + terms) \times d\sigma_F \tag{6.1.5}$$

Despite this factorability, there is a dilemma that arises when matching the matrix element and parton shower calculations beyond leading order. A double counting occurs between the matrix element calculation at NLO and the parton shower calculation of the leading order diagram, in which the case of an additional leg or emission on the LO diagram is considered in both calculations. To combat this double counting, there are three known methods to employ to correctly *match* the matrix element and parton shower calculations [100]:

- *Slicing* is a strategy in which the phase space is divided into two orthogonal regions. The hard scatter calculation covers one and the parton shower region covers the other.

- *Subtraction* utilizes the exact higher-order matrix element calculation and subtracts off the lower-order shower calculation to get a correction factor to the higher-order shower calculation

- Through imposing *Unitarity*, the third method gets a multiplicative correction factor to the shower splitting functions from the ratio of the matrix element to the shower approximation.

With a *matching* method selected and employed to remove the double counting between the hard scatter and shower calculations, *merging* of the tree-level matrix element calculation with the respective shower calculation is then performed such that the correct probability factor is implemented. The parton shower process must therefore be reversible such that it can be merged with the matrix element, which can be considered part of its branching history [99, 100].

The final phase of event simulation to form a reasonable approximation of experimental observables is hadronization. Near the hadronization scale (or the parton shower low-energy cutoff) where the quantum chromodynamic (QCD) calculations become non-perturbative, the final state partons form primary hadrons. To perform these calculations, the MC generators use phenomenological models like the string model of hadronization, which is based on the assumption that the potential between a quark and anti-quark in a meson increases linearly with separation distance. Also called the Lund fragmentation model, the idea is that mesons will stretch like a string until reaching enough tension to fragment, bringing an additional quark anti-quark pair from the vacuum and increasing the number of bound hadrons. Baryons can be similarly modeled by allowing string breaks from di-quark pairs. The fragmentation continues until no

further partitioning is possible [100].

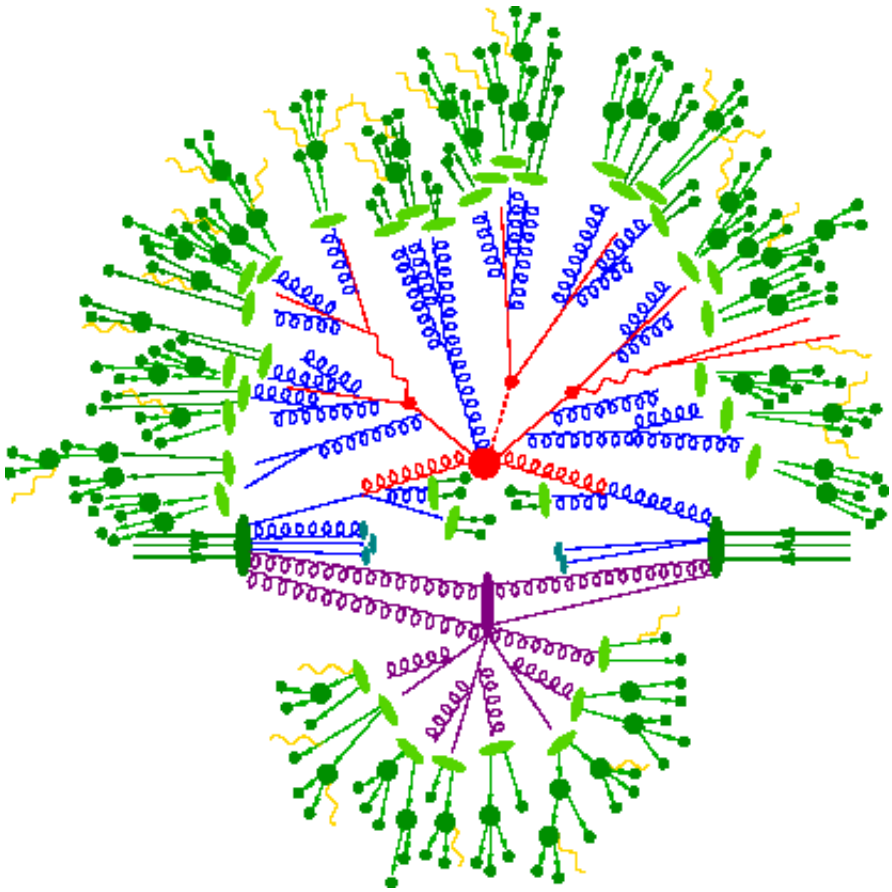


Figure 6.1.: Illustration of proton-proton collisions as simulated by Monte Carlo event generators. The hard scatter, which would be handled through matrix element calculations, can be seen in red at the center. Branching out from there in blue is the component handled by the parton shower. The hadronic transition is shown in light green followed by the hadronic decays in dark green. The purple represents a secondary hard-scattering vertex, and yellow represents photon radiation. [101]

To this point, everything has relied on the factorization of the cross-section calculation with additional interactions from the lower-order process [98]. However, in experiment, scattering from other hard proton-proton collisions and multi-parton interactions also contribute to process cross-sections. To properly simulate hadronic collisions, multi-

partonic interactions (MPI) and color connections/reconnections from beam remnants are also considered [100].

Altogether, Monte Carlo event simulation consists of matrix element calculations of a hard-scatter process at fixed order, matched and merged with parton shower branching calculations, followed by hadronization and fragmentation, with considerations and corrections from the underlying event. These steps are summarized in a schematic drawing in Figure 6.1, portraying proton-proton collisions as interpreted by MC generators.

§ 6.2. MC Event Simulation in ATLAS

There are over 30 MC/parton shower generators in use in ATLAS, but the main ones are MadGraph [98], Herwig [102, 103], Pythia [104], Sherpa [105], and POWHEG [106, 107, 108]. Each generator has a slightly different implementation of the collection of event simulation strategies discussed in Section 6.1. It is not out of the ordinary for an ATLAS MC production request to ask for hundreds of millions of simulated events for a particular process, but there are more elements to the process than the pure theoretical simulation strategies previously mentioned. The output of the matrix element calculation step is a file-type known as Les Houches Event File (LHE), which is often stored in case it is desired to be reshowered [109]. The output of the parton shower generation, called the EVNT file, is in the format HepMC, and it contains a graph structure with the particle and vertex information detailing the evolution of the simulated events [110].

Following the raw event simulation, it is necessary to simulate the detector response to such events, one sub-detector at a time, using tools based on Geant4 [111]. Given the truth information of particle energies and trajectories, the detector ‘‘hits” are approxi-

mated taking into consideration collection efficiencies and coverage specified down to the level of individual modules. The output of this simulation step is known as the HITS file. The full simulation is computationally expensive, taking lots of time and computational resources. Fast simulation software, like AtlFast3 [112], takes advantage of parametrized techniques and machine learning to quickly simulate a detector response to smeared truth-level quantities, and it can be used to generate large statistics in cases where more precise calibrations are not needed. After the simulation of detector interaction, the responses of ATLAS's readout electronics are estimated in a step called digitization. The digitization step also includes the effects of pile-up, or other proton-proton collisions occurring in the same event. Digitization takes the HITS files to simulated detector output files called Raw Data Objects (RDO) [110].

With the detector response approximated, the next step reconstruct event objects from the detector readings. This is a step that is not specific to simulated events, but occurs also for experimentally-collected data, and is discussed further in Chapter 7. Reconstruction brings the file-type to Analysis Object Data or AOD, a clunky structured data format which has a lot of excess information. Inside AOD's, all the analysis objects one could ever want are stored along with their associated truth information: jets, missing energy, electrons, photons, etc. To produce an analysis nTuple, most analyses perform an extra derivation step, in which a DAOD (Derived AOD) is produced with just the analysis objects of interest [113]. In the derivation of a DAOD from and AOD, entire events, analysis objects, or unneeded variables can be removed based on pre-set criteria. The Monte Carlo event simulation production pipeline and its intermediate file-types [114] are shown in the flow chart in Figure 6.2.

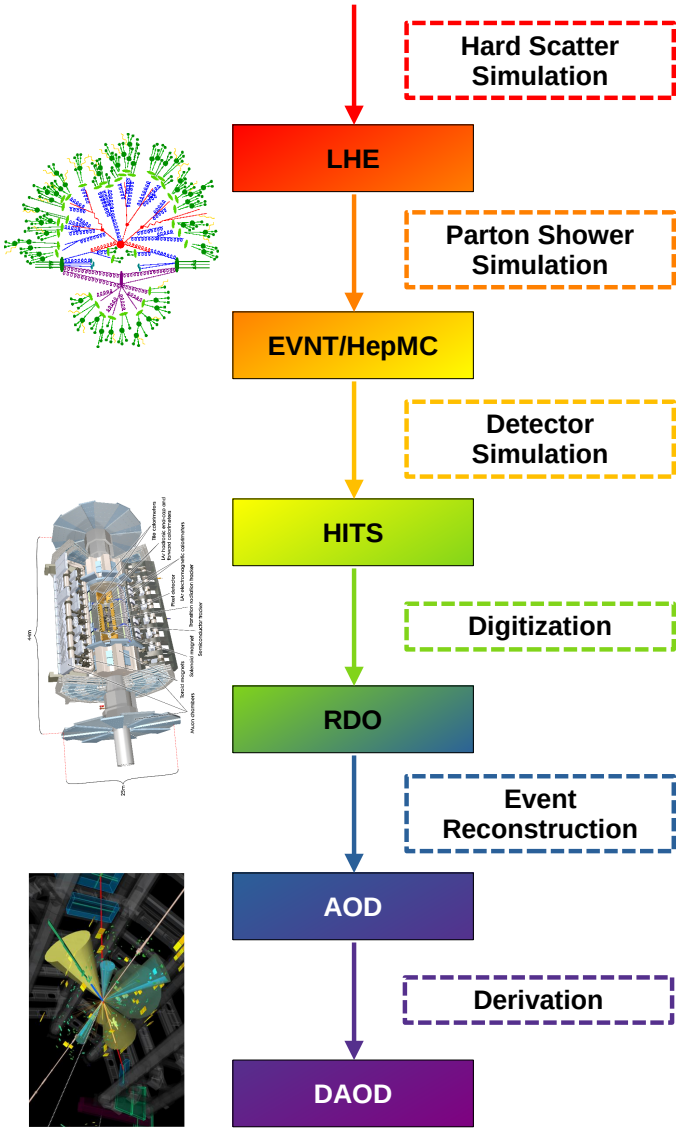


Figure 6.2.: Monte Carlo event simulation production pipeline and intermediate file-types.

§ 6.3. MC Simulation for VBF Higgs + Photon Analysis

Signal $H\gamma jj$ Monte Carlo (MC) events were generated at next-to-leading order (NLO) with MadGraph@NLO [115] and showered using Herwig7 [116, 117], with the PDF4LHC15 parton distribution function (PDF) set [118]. For this signal process, there was no other functional choice of generators. The process does not exist in POWHEG, and Pythia8 gives anomalous kinematic distributions for extra jets due to a bug in its modeling of ‘‘dipole recoil” [119]. This signal sample consists of Higgs bosons produced through VBF or VH, and contributions from Higgs processes that are more rare in the selected phase space are considered separately but are treated as background. To describe the $H\gamma$ component of other Higgs production processes, photon-filtered samples are derived from inclusive productions. The gluon-gluon fusion Higgs production (ggH) sample used Powheg [106, 120, 121] interfaced with Pythia8 [122]. Involvement of Higgs production in association with a $t\bar{t}$ pair (ttH) is described with MC simulated events also generated with Powheg interfaced with Pythia8.

High statistics fully-simulated Monte Carlo background samples were generated at Leading Order (LO) specifically for this analysis for the two dominant non-resonant backgrounds, $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$, using MadGraph@NLO for the matrix element calculation and Pythia8 for the showering to final state particles. Both of these samples use the PDF set PDF4LHC21_40_pdfas, and a filter requiring di-jet invariant mass above 500 GeV is applied to improve efficiency in the selected phase space discussed in Chapter 7. These non-resonant QCD backgrounds have about 37% efficiency under this high di-jet invariant mass filter, so without it, simulation of almost an extra two-thirds of the produced statistics would have been wasted. The final non-resonant background consid-

ered, $t\bar{t}$ production, is modeled by samples with all hadronic and partially or fully leptonic final states that were generated at NLO using Powheg and showered using Pythia8. These samples were renormalized to the next-to-next-to-leading order (NNLO) cross-section estimations [123]. Electroweak $Z\gamma jj$ was modeled specifically for this analysis at NLO using the identical parameters and PDF set to the signal sample. QCD $Z\gamma jj$ was generated at LO using MadGraph@NLO for the hard-scatter and Pythia8 for showering.

Certain simulation configurations are common to all samples. The decays of bottom and charm hadrons were performed by EvtGen [124]. Minimum bias events were simulated using the Pythia8 generator with the NNPDF2.3LO PDF set and the A3 tune [125]. Multiple events were overlaid on the hard-scatter interaction, according to the luminosity profile of the recorded data, to model pileup contributions. Generator choices, order of calculation, and PDF sets used for the considered MC samples are fully summarized in Table 6.1.

Sample	Generator	Calculation Order	PDF
$H(\rightarrow b\bar{b})\gamma jj$	aMC@NLO+Herwig 7	NLO	PDF4LHC15
$ggH + \gamma$	Powheg NNLOPS+Pythia 8	NNLO	PDF4LHC15
ttH (all-hadronic)	Powheg +Pythia 8	NNLO	NNPDF3.0
ttH (semi-leptonic)	Powheg +Pythia 8	NNLO	NNPDF3.0
$Z(\rightarrow b\bar{b})\gamma jj$ (EWK)	aMC@NLO+Herwig 7	NLO	PDF4LHC15
$Z(\rightarrow b\bar{b})\gamma jj$ (QCD)	MadGraph 5+Pythia 8	LO	PDF4LHC15
$b\bar{b}\gamma jj$ (non-resonant)	MadGraph 5+Pythia 8	LO	PDF4LHC21
$c\bar{c}\gamma jj$ (non-resonant)	MadGraph 5+Pythia 8	LO	PDF4LHC21
$t\bar{t}$ (non-allhadronic)	Powheg +Pythia 8	NLO	NNPDF3.0
$t\bar{t}$ (all-hadronic)	Powheg +Pythia 8	NLO	NNPDF3.0

Table 6.1.: Generators, order of calculation, and PDF sets used for the MC signal and background samples.

Chapter 7.

Object Reconstruction and Event Selection

For data either recorded in ATLAS or simulated via the methods described in Chapter 6, the signals in the detector readout must be reconstructed to form usable event objects. Then, events can be filtered by requisite criteria on the presence and kinematics of specific physics objects. The reconstruction and selection of electromagnetic objects and hadronic jets is described in Sections 7.1 and 7.2, respectively. Following object reconstruction, the first layer of data filtration through event triggers is specified in Section 7.3. The final prescription for offline event selection is given in Section 7.4.

§ 7.1. Electromagnetic Objects

The reconstruction of electromagnetic objects proceeds in phases. In the first phase, topo-clusters, or topologically connected energy deposits in the EM and hadronic calorimeters, are prepared, starting from proto-cluster seed cells which have high signal-to-noise ratio. The topo-clusters evolve by picking up neighboring cells that are above a signal-to-noise threshold. Once the topo-cluster is constructed, attempts are made to match it to a track in the inner detector. If a match is made, the track is re-fit with infor-

mation from the topo-cluster included, and this process helps to correct for electron energy losses due to bremsstrahlung radiation. An isolated topo-cluster with an associated matched track will be an electron candidate. Tracks that are very loosely matched to topo-clusters become candidates for conversion vertex reconstruction, which uses an algorithm to identify possible electron-positron pairs that may have arisen in the ID from a converted photon. Topo-clusters with an identified reconstructed conversion vertex or with no associated track become photon candidates [126].

These topo-clusters act as potential seeds for the formation of superclusters. This dynamic approach allows the recovery of low-energy photons produced in the Inner Detector. Topo-clusters that are near an initial seed could have emerged from bremsstrahlung or photon conversion, so they have the possibility of being merged into the supercluster. Tracks or conversion vertices are freshly matched to the identified superclusters, though trackless superclusters are also kept as unconverted photon candidates. Finally, the electron and photon event objects to be used in analysis are built and their energies calibrated [126].

7.1.1. Photons

A high energy photon is an important element to the final state signature of this analysis. Both converted and unconverted photons are used. Photon identification and isolation criteria are optimized to efficiently select prompt photons and avoid hadronic jets faking photons. Several discriminating variables are used to form cut-based criteria on the EM cluster shape for photon identification [126]. Some examples of such variables are:

- Hadronic leakage: ratio of the E_{T} in the hadronic calorimeter to the E_{T} in EM calorimeter

- Layer energy fraction: ratio of energy measured in each EM calorimeter layer to the total energy of the cluster
- Lateral shower width: $\sqrt{\frac{\sum_i E_i (i-i_{max})^2}{\sum_i E_i}}$, where i_{max} represents the index of the highest energy cell and i runs over indices of neighboring cells in the cluster

The ‘‘tight” identification criterion is the primary standard for photon identification, and it is built from discriminating variables, with requirements tuned in bins of $|\eta|$ and E_{T}. Photons in this analysis are required to pass tight identification in offline selection. ‘‘Medium” and ‘‘loose” identification criteria are used in the event selection triggers [126].

To further reduce the mistaken acceptance of jets faking photons, isolation requirements are imposed on the activity surrounding the photon candidate objects. This activity can be estimated from ID tracks near the photon supercluster or from nearby energy deposits in the calorimeters. In a cone (of radius $\Delta R = 0.4$ for our purposes) around the photon cluster direction, energy deposits in the calorimeters but outside the photon supercluster are summed to get a parameter $E_{\mathrm{T}}^{isol}_{cone}$. This analysis uses a fixed-cut tight isolation requirement for photon selection, which places an E_{T}-dependent criterion on the nearby energy, such that $E_{\mathrm{T}}^{isol}_{cone} < 0.022 \times E_{\mathrm{T}} + 2.45\,\mathrm{GeV}$ [126].

7.1.2. Electrons and Muons

Prompt electrons and muons are not of particular interest for the target signature, but they could be signs of overlap with the targets of other $H(b\bar{b})$ analyses. For example, VH Higgs production has the potential to look similar to our target but could have leptonic final states resulting from the vector boson. Any event containing isolated electrons or

muons with $p_{\mathrm{T}} > 25$ GeV are vetoed to avoid this potential overlap.

Electron candidates must satisfy $E_{\mathrm{T}} > 25$GeV, $|\eta| < 2.47$, and tight identification criteria [127]. The identification criteria is very similar to that for photons, but also includes a few discriminating variables related to the matched supercluster tracks. Muon analysis objects are reconstructed in a similar way to what was described for photons and electrons, but using energy deposits in the muon spectrometers, rather than the EM calorimeters [128]. Muon candidates have $p_{\mathrm{T}} > 25$ GeV and satisfy loose identification criteria [129]. For electrons, track-based isolation requirements depend on electron p_{T}, and for muons, the requirement for isolation is that tracks within a cone of $\Delta R = 0.2$ around the muon candidate track must have a total of less than 1.25 GeV.

§ 7.2. Jets

The target signature of this analysis contains at least four jets, two VBF jets and two b-tagged jets coming from the Higgs decay. Compared to the previous investigation into this channel, an update has been made to the jet reconstruction algorithm to include particle flow information, discussed further in Section 7.2.1. The methods of tagging b-jet candidates are relayed in Section 7.2.2. Further details to the considerations made for jets in this analysis are mentioned in Sections 7.2.3 and 7.2.4.

7.2.1. Jet Reconstruction

Similar to the construction of EM objects, jet reconstruction begins with an initial set of cluster candidates or seeds and engages in a course of merging neighboring clusters based on a criterion of closeness. The distance parameter of interest for neighboring

entities i and j is defined in Equation 7.2.1.

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}{R^2} \tag{7.2.1}$$

In Equation 7.2.1, k_t, y, ϕ are the transverse momentum, rapidity, and azimuthal angles of the merge candidates, and R is a radius parameter fixed for the algorithm (for our jets, $R = 0.4$) [130]. Clusters are merged in order of smallest d_{ij}, and the distance parameters are recalculated after merging. Merging continues until no further candidates within the radius parameter are available. ATLAS uses the anti-k_T algorithm in which $p = -1$ [131].

In the anti-k_T algorithm, small distance parameters will come more easily when at least one of the merge candidates is hard, and so soft clusters will be merged with hard ones long before two soft clusters are merged. From an initial hard seed, if there is not a second hard cluster within a distance of $2R$, the anti-k_T algorithm will result in a nice round jet. Regularity of boundaries are related to the sensitivity to non-perturbative effects like hadronization, so they are an important feature. If there are two hard entities within a radius of $2R$ and with similar transverse momentum, the surrounding soft deposits will be divided between them with a straight line down the middle. An example of the resulting clusters from the anti-k_T algorithm is illustrated in Figure 7.1 for a radius parameter $R = 1$ [131].

Traditionally, jets were reconstructed from topological clusters built solely from calorimeter information [132]. While this was true for the previous Run 2 analysis, this analysis updates the jet reconstruction algorithm to include particle flow information. In this alternative approach, the ParticleFlow [132] algorithm uses both track and calorimeter information, and its implementation has shown improvements in energy resolution

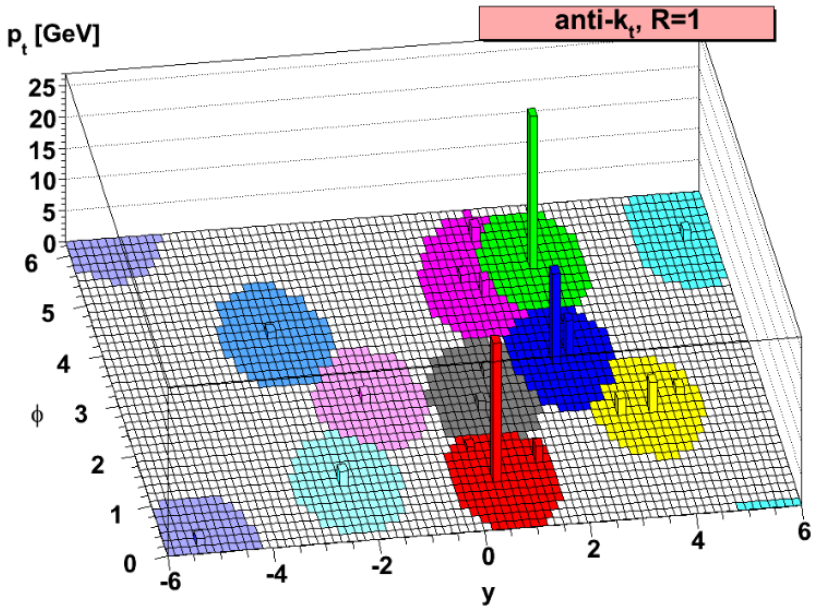


Figure 7.1.: Illustration of the finished product of cluster-merging for jet reconstruction using the anti-k_T algorithm. This parton-level event was simulated with Herwig together with many soft "ghosts" to match the jet boundaries. [131]

and pile-up stability. First, calorimeter topo-clusters are corrected by being matched to tracks and subtracting the expected energy that would be deposited from a particle with such a track. If the remaining energy of the system is consistent with the predicted level of fluctuations, the remnants are removed to form the corrected topo-clusters. Next, both the surviving positive-energy topo-clusters and the tracks matched to the hard vertex are used as inputs to the anti-k_T jet reconstruction. The track information allows the jet objects to be well matched to the primary hard scatter vertex, which greatly helps to reduce pile-up contamination. Finally, jet energy correction factors can be applied with higher specificity due to the added kinematic information from the inclusion of tracks, which allows for the improvements in jet energy resolution. Jet energies are calibrated sequentially, using correction factors related to the charged fraction of jet energy measured in tracks, then to the fraction measured in the third EM calorimeter layer, and lastly to the fraction of jet energy in the first Tile layer [132, 133].

Further assisting in pile-up rejection, jets are tightly tagged with the event primary vertex in the central regions using a likelihood-based discriminant called the Jet Vertex Tagger (JVT), and jets are loosely tagged in the forward regions using a similar discriminant dedicated to the forward region, the fJVT [134].

7.2.2. Flavor Tagging

This search for the Higgs boson looks in its dominant decay mode to bottom quarks. Hence, the ability to tag jets as b-flavor candidates is of utmost importance. b-tagging is used both in the trigger requirements and in the offline event selection requirements for this analysis. The triggers (discussed in Section 7.3) use MV2C10 and MV2C20 [135] algorithms, depending on the data taking period. The MV2 algorithms make use of

a boosted decision tree (BDT) which takes the outputs from lower level tagging algorithms, like IP2D/IP3D, SV1, and JetFitter, along with kinematic info as inputs [135]. The BDT produces a single-dimensional discriminant that allows for a cut-based b-tagging criterion. The selection of cut is calibrated using a $t\bar{t}$ sample to have predictable flavor efficiencies. The 77% working point is used by the triggers, which should correspond to about 77% b-flavor jet acceptance [135].

In offline selection (discussed in Section 7.4), this analysis replaced the MV2C10 b-tagging algorithm that was used in offline selection for the previous Run 2 analysis with the more advanced DL1r algorithm. This DL1r algorithm leans on a neural network (NN) to form its discriminant, using the same inputs as the MV2 algorithm [136]. The output discriminant of the DL1r algorithm is multidimensional, giving corresponding probabilities of the jet to be b-flavor, c-flavor, or light-flavor. The presence of multiple output nodes gives added flexibility in the final discriminant, which can be manipulated to be used not just for b-tagging but also for c-tagging. The final discriminant cut was optimized using both a $t\bar{t}$ sample and a $Z \rightarrow q\bar{q}$ sample. For the selection of b-tagged jets, a 77% efficiency fixed-cut working point is used, and the b-jet is required to fall in the central region with $|\eta| \leq 2.5$. The 77% working point has a rejection factor of 6 and of 134 on charm and light-jets, respectively [136, 137].

7.2.3. Jet Energy Corrections

Now that it is clear how b-jets are identified, another problem must be considered. The b-hadrons can decay through the weak interaction to charged leptons and neutrinos, and this energy may not be fully contained in the hadronic calorimeter. About 24% of the time, there is an electron or muon and neutrino inside the b-jet cone [138]. To account

for this possible source of missing energy, b-jets are corrected in this analysis using ‘‘muon-in-jet” (OneMu) plus ‘‘PtReco” methods.

In the muon-in-jet correction, if there is at least one jet within the radius of the b-jet cone, the muon most central in the cone will be added to the jet 4-vector [138]. The corrected jet 4-vector is called OneMu, and the choice of muon has been shown to result in negligible systematic uncertainties. Next, the PtReco [139] correction is applied, scaling the OneMu 4-vector by a factor that is dependent on p_{T} and η. The scale factors were derived from a $ZH \rightarrow llb\bar{b}$ sample, comparing the reconstructed jet energy to the target jet energy in bins of p_{T} and η. Since the scale factors are kinematically dependent, they should still be relevant for use in this analysis, and new scale factors were not derived [140].

7.2.4. Overlap Removal

Often, particles will leave energy deposits in multiple calorimeters; as mentioned in Chapter 4, particle identification is not always clear-cut. When this occurs, the same energy deposits might be used to reconstruct multiple event objects, and for example, the same physics object could get marked as both a lepton and a jet. To avoid this misreconstruction scenario, overlapping event objects are removed from consideration. Any jet within $\Delta R = 0.2$ of an electron is removed, and a jet within the same radius of a muon with fewer than three associated tracks is removed. Any photon identified within $\Delta R = 0.4$ of an electron or muon is removed. Any jet that is within a region of $\Delta R = 0.4$ around a photon with $E_{\mathrm{T}} > 30$ GeV is removed and not allowed to contribute to event selection.

§ 7.3. Data Samples and Triggering

As discussed in Section 4.2.1, there are too many proton-proton collision events to record every single one. Triggers are used as methods of filtering the events to maximize acceptance of hard scatter events by taking advantage of characteristic signatures, such as large deposits of transverse energy [141]. The particular trigger (or triggers) that an event passes to reach the ATLAS recorded dataset is tagged to that event. Then, particular analyses use these triggers to select data that might be interesting for their particular final state signature, so the triggers not only act as filtration for selecting events to record but also serve as a filter for combing through the recorded dataset to highlight those events that are most likely to contain the target process.

In this analysis, 133 fb^{-1} of ATLAS proton-proton collision data are considered, collected at a center of mass energy of $\sqrt{s} = 13$ TeV. This integrated luminosity was collected in Run 2, during the time period of 2015-2018. Although the complete Run 2 data set consists of 140 fb^{-1}, there were absences in the availability of certain VBF-related triggers and b-jet tagging triggers that caused reductions in the usable size of the data set. Table 7.1 summarizes the triggers used in this analysis.

Early on in Run 2 data taking, in 2015 and in the start of 2016 up to period F, the VBF $H + \gamma$ trigger available had no b-tagging requirement. This trigger is photon-seeded, requiring an isolated EM object with $E_{\text{T}} \geq 22$ GeV at the L1 stage [142]. This EM object is further refined at the HLT stage, requiring a photon passing the medium identification criterion with $E_{\text{T}} \geq 25$ GeV. The HLT trigger additionally requires at least four jets with $E_{\text{T}} \geq 35$ GeV in the pseudorapidity range $|\eta| \leq 4.9$. Of these four jets, at least one pair must have a dijet mass of at least 700 GeV.

In the rest of the data taking periods, from 2016 period G onwards through 2017 and 2018, b-tagging requirements at the HLT level for VBF $H+\gamma$ triggers became available. This analysis makes use of a trigger which requires at least one b-tagged jet at the 77% working point. In 2016, this tagging requirement is imposed using the MV2c20 b-tagging algorithm, and in 2017 and 2018 the MV2c10 algorithm is used [135]. Other than the b-tagging requirement, these triggers are quite similar to the 0-btag trigger used earlier in Run 2. Like the 0-btag trigger, the 1-btag triggers are also photon-seeded with an L1 trigger of a 24 GeV EM object, further refined in the HLT to be a photon passing medium identification criteria and with $E_{\mathrm{T}} \geq 27$ GeV. There is also forced to be a total of at least four jets of $E_{\mathrm{T}} \geq 35$ GeV (including the b-tagged jet), with at least one pair having $m_{jj} \geq 700$ GeV.

The last trigger of interest was added for this analysis and was not used in the previous full Run 2 analysis of this signature. In 2017 and 2018, a VBF $H + \gamma$ trigger became available which was meant to target events with lower photon energies. Unlike the other triggers, it is not photon-seeded at the L1 level, but is instead seeded by large dijet invariant mass with $m_{jj} \geq 300$ GeV, giving it the name L1Topo trigger. With this high invariant mass requirement, L1 EM threshold could be lowered to 18 GeV without trigger rates becoming too high. Like the other triggers, the L1Topo HLT requires at least four 35 GeV jets, including at least one b-tagged jet using MV2c10 at the 77% working point. The HLT also increases the dijet invariant mass requirement to 500 GeV and the photon energy minimum to 20 GeV.

Validations were performed to ensure that these triggers perform similarly for MC simulation as well as for data. A comparison of the MC and data trigger response as a function of offline photon p_{T} requirement is visible in Figure 7.2. With respect to

photon p_{T}, after a steep turn-on at 30 GeV, the efficiencies for MC and data match within statistical uncertainties, so no further calibration or correction is needed. With respect to the minimum di-jet invariant mass threshold, however, there remain some discrepancies in trigger efficiency between MC and data, even in the plateau regime, as can be seen in Figure 7.3. To account for this discrepancy, the MC is corrected in the analysis using a jet-level scale factor derived and recommended by the ATLAS *b*-jet trigger group [143, 144, 145].

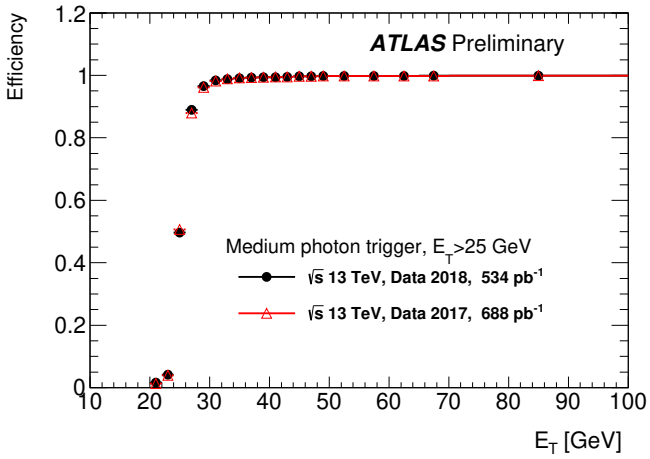
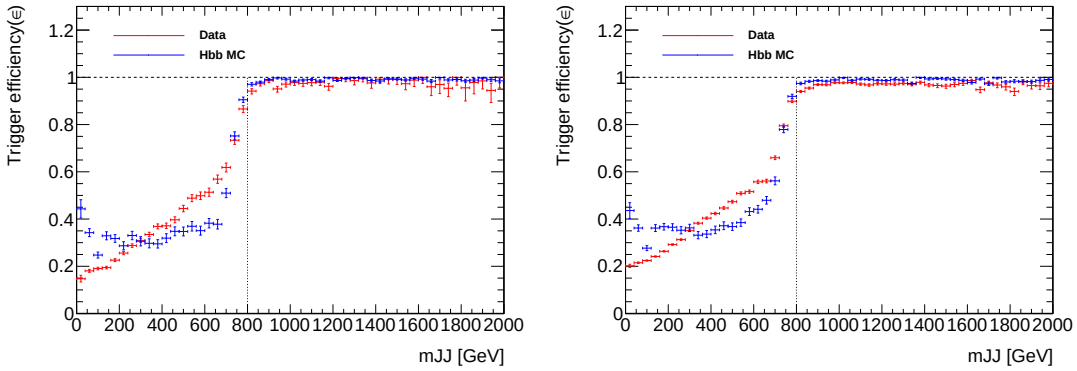


Figure 7.2.: Acceptance efficiency from photon-seeded triggers on data taken in 2017 and 2018 as a function of offline photon p_{T} selection. Photon selection requirements were set to capture only the efficiency plateau. Efficiency drops steeply for events with photons less than 30 GeV. [142]



(a) 0 *b*-jet (b) 1 *b*-jet

Figure 7.3.: Trigger efficiency for MC and data as a function of the offline selection threshold of minimum dijet invariant mass. Left comparison (a) is made for the 2015 and 2016 triggers which have no *b*-tagging requirement. Right plot (b) is for the photon-seeded trigger requiring at least one jet to pass the MV2c10 77% working point, and the comparison is nearly identical for the MV2c20 trigger used in 2016 as well as for the L1Topo trigger. The difference between MC and data after the steep turn-on around 800 GeV is corrected in the analysis by a jet-level scale factor. [146]

Period	Trigger Chain	Lumi	min m_{jj}	b-tags	min p_{T}^{γ}
Data15	HLT_g25_medium_L1EM22VHI _4j35_0eta490_invm700	2.6 fb^{-1}	700 GeV	0	25 GeV
Data16 (A-F)	HLT_g25_medium_L1EM22VHI _4j35_0eta490_invm700	15.1 fb^{-1}	700 GeV	0	25 GeV
Data16 (G-L)	HLT_g25_medium_L1EM22VHI _j35_0eta490_bmv2c2077_split _3j35_0eta490_invm700	13.3 fb^{-1}	700 GeV	1	27 GeV
Data 17	HLT_g25_medium_L1EM22VHI _j35_0eta490_bmv2c1077_split _3j35_0eta490_invm700	43.9 fb^{-1}	700 GeV	1	27 GeV
Data 17	HLT_g20_tight_icaloloose _j15_gsc35_bmv2c1077_split _3j35_0eta490_invm500	43.9 fb^{-1}	500 GeV	1	20 GeV
Data 18	HLT_g25_medium_L1EM22VHI _j35_0eta490_bmv2c1077_split _3j35_0eta490_invm700	58.1 fb^{-1}	700 GeV	1	27 GeV
Data 18	HLT_g20_tight_icaloloose _j35_bmv2c1077_split _3j35_0eta490_invm500	58.1 fb^{-1}	500 GeV	1	20 GeV

Table 7.1.: Summary of VBF $H(\rightarrow b\bar{b}) + \gamma$ triggers used in the analysis across different data taking periods. All triggers used require at least 4 jets with p_{T} of at least 35 GeV. The triggers in use also employ minimum di-jet invariant mass, MV2c10/MV2c20 b-tagging, and photon transverse energy requirements. The total integrated luminosity available in the Run 2 good run list suitable for these triggers is 133 fb^{-1}, and the integrated luminosity available for each trigger is shown in the table.

§ 7.4. Offline Event Selection

The target signature of this analysis consists of two b-tagged jets, two VBF-candidate jets, and a high energy photon. The selection requirements on these objects are summarized in Table 7.2, and aim at maximizing signal acceptance and background rejection.

Due to trigger requirements, each of the four offline selected jets are required to have $p_{\mathrm{T}} > 40$ GeV. The two b-tagged jets with highest transverse momentum are selected first as Higgs decay product candidates. The remaining jets are ordered by p_{T} and combined pair-wise, and the VBF-candidate jets are selected as the pair with highest invariant mass, above 800 GeV where the HLT efficiency reaches a plateau. The high-energy photon in this signature is required to have a minimum transverse energy of 30 GeV.

After applying the event selection criteria, candidate events are divided into two regions, which are both used simultaneously in classifier training and in the signal extraction fit. The signal region contains the majority of the signal sensitivity with candidate events that fall in the Higgs window of the invariant di b-jet mass, with m_{bb} between 100 and 150 GeV. The control region consists of events in the m_{bb} sidebands outside this mass window, with $m_{bb} \leq 100$ GeV and $150 \leq m_{bb} \leq 220$ GeV, and has the main purpose of fixing the background normalizations.

Trigger	L1	≥ 1 photon with $E_{\mathrm{T}} > 22\,\mathrm{GeV}$
	HLT	≥ 1 photon with $E_{\mathrm{T}} > 25\,\mathrm{GeV}$ ≥ 4 jets (or ≥ 3 jets and ≥ 1 b-jet) with $E_{\mathrm{T}} > 35\,\mathrm{GeV}$ and $\lvert\eta\rvert < 4.9$ $m_{jj} > 700\,\mathrm{GeV}$
Offline		≥ 1 photon with $E_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $\lvert\eta\rvert < 1.37$ or $1.52 < \lvert\eta\rvert < 2.37$ ≥ 2 b-jets with $p_{\mathrm{T}} > 40\,\mathrm{GeV}$ and $\lvert\eta\rvert < 2.5$ ≥ 2 additional jets with $p_{\mathrm{T}} > 40\,\mathrm{GeV}$ and $\lvert\eta\rvert < 4.5$ $m_{jj} > 800\,\mathrm{GeV}$ No electrons or muons with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ $m_{bb} < 220\,\mathrm{GeV}$
Region Selection	Signal Region	$100\,\mathrm{GeV} \leq m_{bb} \leq 150\,\mathrm{GeV}$
	Control Region	$m_{bb} \leq 100\,\mathrm{GeV}$ and $150\,\mathrm{GeV} \leq m_{bb} \leq 220\,\mathrm{GeV}$

Table 7.2.: Trigger and offline event selection criteria and region definitions for the $H(\rightarrow b\bar{b})jj + \gamma$ signature. L1 and HLT refer to the first-level trigger and the high-level trigger, respectively. The p_{T} and $|\eta|$ offline jet requirements are used to match trigger selections and b-tagging requirements. Criteria in m_{bb} define the control and signal fitting regions.

7.4.1. Direct-tagging vs Truth-tagging

Direct-tagging and truth-tagging are two methods of selecting MC events with signatures involving b-jets. Direct-tagging is the most straightforward selection strategy, to only accept events that have b-jet candidates with a tagging discriminant that passes the specified working point cuts. With truth-tagging, any event that passes the other offline selection requirements besides the b-tagging requirement is accepted, but a scale factor reduces its event weight such that its contribution is smaller than events passing the b-tagging requirement and decreases with lower b-flavor probability. A truth-tagged sample should be the superset of the direct-tagged sample, containing all the direct-tagged events and then some.

These tagging approaches show minimal difference in shape of kinematic distributions but noticeable effect on normalization due to the online b-tagging trigger requirements which forces a direct-tagging scheme for at least one of the b-jet candidates [147]. To maximize statistics for the NN training samples as well as for the MC fit templates, the dominant background MC samples for $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ use truth-tagging. Signal, data, and all other smaller backgrounds use direct-tagging. There is further discussion of the impacts of this and support for this choice provided in Chapter 9.

Chapter 8.

Multivariate Analysis Techniques

The multivariate analysis technique for event classification is a cornerstone of this analysis. A general discussion on machine learning classification algorithms is provided in Section 8.1. Following, a comparison is made in Section 8.2 between the technique used in the previous search in this channel, a boosted decision tree, and the neural network algorithm employed here. An alternative neural network strategy is considered in Section 8.3, in which an adversarial approach to maintaining independence between the output classification score and the di b-jet invariant mass, thereby preserving the output distribution shapes between the control and signal regions. Finally, details regarding the ultimate event classification model and training used in this analysis are in Section 8.4.

§ 8.1. Machine Learning Classification Algorithms

This analysis strategy is dependent on deep machine learning (ML) classification to aide in the separation of signal (Higgs b-decays) from background (non-resonant b-jet production). Such algorithms depend on several components: training datasets which contain a set of input variables and target labels, a model with tunable parameters and

optimization strategy, and an objective loss function. Input training datasets can come in various shapes, and there are different associated models that can handle data representations as images, variable-length sequences, or flattened columnar vectors. Deep learning algorithms are characterized by models which have many layers in order to progressively extract more information from the input features presented. In the case of supervised classification algorithms, which are touched upon here, the loss function should quantify the limitations of the trained model to correctly categorize new data points with their true target label. For our purposes, the loss function quantifies the misidentification of the process associated with an event based on a set of event kinematics. Gradient descent is the method by which the training is optimized to achieve the minimum evaluation according to the objective loss function [148]. This optimization is performed by altering the model parameters in the direction opposite to the gradient of the loss function with respect to those parameters, in other words in the direction of steepest downhill slope, with a step-size that is controlled by the learning rate model hyperparameter set by the user. Gradient descent continues iteratively, tuning the model parameters at each step, until a minimum is reached, which is characterized by little to no further change in the loss function evaluation. Of course, there is no guarantee that this minimum is a global minimum, which is why I would describe ML as more of an art than a science.

Boosted Decision Trees (BDTs) are an example of a ‘‘shallow” learning algorithm, as they combine information learned from input features in parallel, rather than progressively building upon that information in layers. A decision tree recursively splits data based on criteria of the given input features, where each split criterion is optimized to give the maximum gain in information [149]. Figure 4.2 could be considered a decision

tree, as it recursively splits input features related to the signatures of traversing particles in ATLAS. This information gain can be quantified by the entropy of the probability distribution of a particular target following a split, where a maximum entropy would correspond to a 50/50 split and minimum to a 0/100 split. Maximum information is gained when entropy is at a minimum. A boosted decision tree combines the information gained from multiple decision trees by summing them, each with an associated weight relative to its accuracy [150]. Though each decision tree might be weak on its own, boosting them together can create quite a powerful classifier. The limitation of BDTs, though, is that their structure does not allow for deep learning from layers in series, and this surface approach fails in scenarios with complicated correlations between the input features.

Neural network models consist of nodes or neurons and inter-layer neural connections, which have associated weights which hold the knowledge gained through the learning process [149]. Dense neural networks (DNNs), also called fully-connected, are the most flexible and general artificial neural networks and are characterized by all the nodes of a layer holding weighted connections to all of the nodes of the directly preceding and succeeding layers.

§ 8.2. Comparison of Boosted Decision Tree to Other MVA Algorithms

In the previous Run 2 analysis, a Boosted Decision Tree (BDT) was trained to separate signal from nonresonant QCD background. At the time, comparisons were not made with other MVA techniques. It was suspected that a neural network might outperform

a BDT in cases where input variables have complicated correlations. In this legacy analysis, studies were done to compare the performance of different machine learning techniques for signal and background separation. Only supervised learning strategies were investigated.

These preliminary studies were performed using Root Toolkit for Multivariate Data Analysis (TMVA) [151, 152]. To give the best chance of a fair comparison, identical input variables were provided. These input variables were selected to match those used in the final event selection of the previous analysis [94]. Input variables included m_{JJ}, $\cos\theta_C$, dRB_1J_1, photon centrality, pT_{JJ}, transverse momentum balance, $d\eta_{JJ}$, $dRB_2\gamma$, $dRB1\gamma$, and $d\phi_{BBJJ}$.

Without much optimization for either MVA type, signal and background separation performance was compared between TMVA's BDT and a Multilayer Perceptron (MLP), which is a basic neural network. Figures 8.1 and 8.2 show the outputs of these trainings. Other variables are also maintained between the two methods shown, such as training samples and statistics. Qualitatively, the improvement in signal and background separation is palpable in the output distribution plots, where the MLP demonstrates more of a peaking nature with a larger distance between the signal and background peaks than the BDT. Quantitatively, a useful metric of comparison is the integral under the Receiving Operator Characteristic (ROC) curve. Here, the MLP, with an area under the ROC curve of 0.821, outperforms the BDT, which has an area under the ROC curve of 0.806.

Other TMVA methods were also investigated, such as support vector machine, binary neural network, and MLP artifical neural network. Ultimately, these studies supported the idea that a complex neural network may achieve better signal purity and extraction, as has been shown in this legacy analysis.

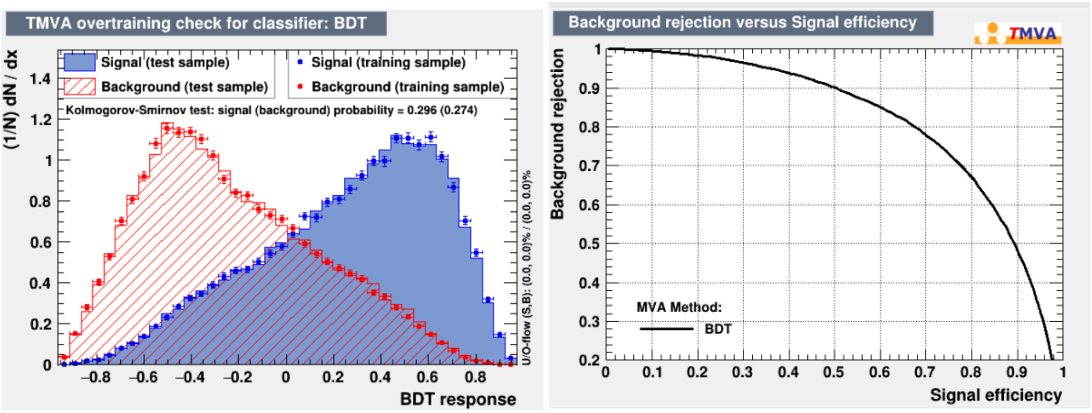


Figure 8.1.: Boosted Decision Tree performance using Root TMVA. BDT score output distributions for signal and background are shown on the left, and the signal acceptance in comparison to background rejection is shown in the ROC curve on the right. The area under the ROC curve is 0.806.

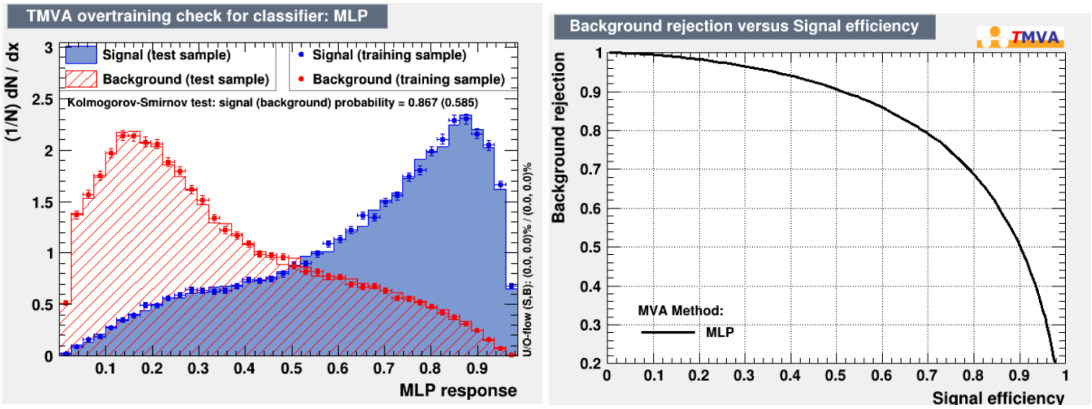


Figure 8.2.: MultiLayer Perceptron (a basic form of neural network) performance using Root TMVA. MLP score output distributions for signal and background are shown on the left, and the signal acceptance in comparison to background rejection is shown in the ROC curve on the right. The area under the ROC curve is 0.821.

§ 8.3. Alternative Approach Considered: Adversarial Neural Network

Once studies demonstrated that a neural network algorithm would have the best chance at handling the complex correlations between kinematic input variables, two strategies were considered to accomplish the task of separating signal-like from background-like events in data. One approach, which was ultimately not adopted, was to train an adversarial neural network (ANN) with a loss function which includes a penalty for sculpting differences in the NN output shape across bins of di-b-jet invariant mass. The motivation for this approach was to preserve a common output score distribution between the control and signal regions, which are defined by regions of m_{bb}, such that there would be no need for shape closure systematics between the regions. Additionally, with a common output shape between the fitting regions, correlation of systematics would be justified without further scrutiny, allowing the fit to retain maximum constraining power of uncertainty nuisance parameters. While the adversarial strategy had strong motivations, ultimately the discriminating power of the di-b-jet invariant mass proved too useful, and a neural network with the freedom to use the power of m_{bb} outperformed the ANN, even with the caveat that some uncertainties require decorrelation between the control and signal regions. In this section, the adversarial neural network approach is detailed, as it was pursued to completion before ultimately landing on the final event classification model, which is discussed in Section 8.4.

Since this strategy was proposed to conserve output score distributions across signal and control regions, m_{bb} itself could not be an input variable to the discriminant. Though m_{bb} was restricted from the training, other variables which contain strong cor-

relations with the invariant mass are included, such as $p_{T,bb}$, thus motivating this more complicated network. The ANN was trained and optimized for the best separation possible between $H\gamma jj$ signal and non-resonant QCD $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ backgrounds while being punished for ascertaining information about the invariant mass, m_{bb}. To accomplish this, a second output was defined. In addition to providing a the traditional output score giving the signal-likeness of an event, a second set of outputs gives the likelihood of an event to fall in a particular m_{bb} bin. The m_{bb} range in the selected phase space from $50 - 220$ GeV was split into 5 bins, though other binnings and numbers of bins were investigated. The ANN training was then performed through the maximization of a loss function with two terms, a positive term associated with the mean square difference between a training event's output score and its true label (0 for background, and 1 for signal) and an opposite sign term associated with the mean square difference of a second m_{bb} bin output and its true invariant mass bin (labeled 1 through 5).

A metric, which is discussed more thoroughly in Section 8.4.1, was defined to compare the performance of ANN models trained with different inputs and hyperparameters. This same metric was also used to compare the performance of the adversarial and direct approaches. The ANN model was iterated until there was confidence that not much further improvement in signal and background separation could be achieved with the given strategy and input datasets. The input variables that provided the best-achieved performance are listed in Table 8.1. However, despite best efforts the ANN model was outperformed by the strategy of training an NN exclusively for signal-background separation, and the final NN model chosen bested this ANN by an estimated 20% in expected signal significance. Poorer constraints on systematic uncertainties would not wash out such a large difference in discriminating power.

Variable Name	Description	Training Phase Space
p_T^{Bal}	Transverse momentum balance for selected final state objects	$0 \leq p_T^{Bal} \leq 1$
m_{jj}	Highest di-jet invariant mass with non-b-tagged jets	$m_{jj} \geq 800\,\mathrm{GeV}$
n_{Jets}	Number of resolved jets in the event	$n_{Jets} \geq 4$
$\min(dR(b,\gamma))$	Minimum angular separation between the photon and one of the b-tagged jets	$0.4 \leq \min(dR(b,\gamma))$
η_γ	Pseudorapidity of the photon	$-4.5 \leq \eta_\gamma \leq 4.5$
$d\eta_{jj}$	Difference in pseudorapidity between the two VBF candidate jets	$0 < d\eta_{jj} < 9$
η_{j5}	Pseudorapidity of the 3rd-highest p_T not b-tagged jet (extra 5th jet beyond the 4-jet required signature)	$-4.5 \leq \eta_{j5} \leq 4.5$, or $\eta_{j5} = -10$ if no extra jet exists
$cen(\gamma, jj)$	Photon centrality	$cen(\gamma, jj) \geq 0$
$d\eta_{bb}$	Difference in pseudorapidity between the two b-tagged jets	$-3.5 \leq d\eta_{bb} \leq 3.5$
$dR(b_1, \gamma)$	Angular separation between the highest momentum b-tagged jet and the photon	$0.4 \leq dR(b_1, \gamma) < 6$
p_T^{j2}	Transverse momentum of the second-highest momentum not b-tagged jet	$p_T^{j2} > 40$ GeV
$dR(b_1, j_1)$	Angular separation between the leading b-tagged jet and the leading not b-tagged jet	$0.4 \leq dR(b_1, j_1) < 6$

Table 8.1.: Ranked list of variables used as input features to the Adversarial Neural Network in the best-performing model studied. Phase space used in training included as it sometimes differs from that of the final event selection criteria.

§ 8.4. Final Event Classification Model: Densely-Connected Neural Network

A densely-connected neural network (DNN) replaces the Boosted Decision Tree (BDT) used in the previous full Run 2 analysis of this signature based on the studies described in Section 8.2, which showed improvement in signal and background separation using a DNN compared to a BDT with identical inputs. The final DNN uses all the discriminating power of the di-b-jet invariant mass and its correlated variables without concern for maintaining a common output score shape across control and signal regions. This strategy outperformed the adversarial technique considered in Section 8.3.

8.4.1. Training and Optimization

The Keras [153] high-level API of TensorFlow [154] was used to train the neural network to distinguish between signal $H\gamma jj$ and the dominant backgrounds, non-resonant QCD $bb\gamma jj$ and $cc\gamma jj$. The network was trained using supervised learning to give an output score between 0 and 1, evaluating each event on how signal-like (closer to 1) or background-like (closer to 0) it was. Background samples $bb\gamma jj$ and $cc\gamma jj$ share the label 0. Training terminates either when the maximum number of epochs is reached or when the evaluation of the loss function does not reduce significantly for 5 epochs in a row.

A 2-Fold method is used for training and evaluating the DNN, such that the full statistics of our Monte Carlo Samples can be utilized in the fit template. MC samples for signal, $bb\gamma jj$, and $cc\gamma jj$ are divided in half (by splitting across even and odd event numbers). Identical hyperparameters are used to train two separate models, each receiving

one half of the MC training samples. Then, to populate the NN output distribution template, MC events are evaluated on the opposite model, such that no event is evaluated with a model where the event was included in the training to avoid biasing the NN output distribution. MC samples from smaller backgrounds that are not used in training and data are also split in half and evaluated using both models.

Input variables and hyperparameters of the training (e.g. the number of hidden layers, number of training epochs, and number of neurons per layer) were optimized by rerunning the training with varied parameters and ranking the training performances using a number counting significance estimate, which assesses the achieved signal and background separation. The number counting significance estimate is obtained by dividing the NN output distribution of the test sample into 20 equal bins, calculating the sum of signal event weights divided by the square root of the sum of background event weights ($S/\sqrt{B}$) in each bin, and adding in quadrature and rescaling to the size of the complete sample set. Event weights (after kinematic reweighting) are used in the training such that the contribution to the loss function is proportional to an event's weight. Additionally, the overall weighting of different training samples with respect to one another became an import parameter to tune for training performance. Without intervention and taking the event weights at face value, the accuracy of signal labeling has relatively no impact in the loss function and therefore the optimization. Since the overall yields of the backgrounds are so much larger, the accuracy of background event labeling would strongly dominate the loss function if not for optimizing the relative sum of weights of the signal and background samples.

8.4.2. DNN Inputs and Hyperparameters

Using the signal VBF $H\gamma jj$ and background QCD $bbjj\gamma$ MC samples in the phase spaces described in Table 7.2, input variables for the DNN were optimized by checking the total number counting significance estimate after each variable was added one at a time. The variable with the highest significance in a 1-input NN was chosen first, and then each remaining variable was added individually to check which additional variable increased the significance the most as the second variable. This loop process then continued to choose the third, fourth, etc. variable that increased the significance the most. Once adding a new variable failed to increase the total significance, the optimization procedure was stopped. This provided a basis of variables from which to continue investigating. Further checks were performed manually to either confirm the maximum performance with variables obtained algorithmically or to alter the input variable combination for enhanced performance that could not be found algorithmically. It was found that this algorithmic method did not provide the optimal list of input variables, though it came close. Several more variables were added via manual searches, still using the significance estimate as a metric of comparison. Over 30 kinematic variables were considered as input.

The 14 kinematic variables chosen after optimizing, in order from most to least correlated with NN score, are shown in Table 8.2. In particular, new variables of note that were not included in the previous Run 2 analysis include m_{bb}, p_T^{bb}, $\Delta\eta_{b,b}$, η_{j5}, n_{Jets}, and p_T^{j1}. Low level variables like the transverse momentum of individual jets were not considered in the previous analysis, nor were variables that related to jets beyond the 4 required in the signature object selection. For events with only four jets, η_{j5} is undefined, so a random value from the corresponding kinematic distribution of 5+ jet events was

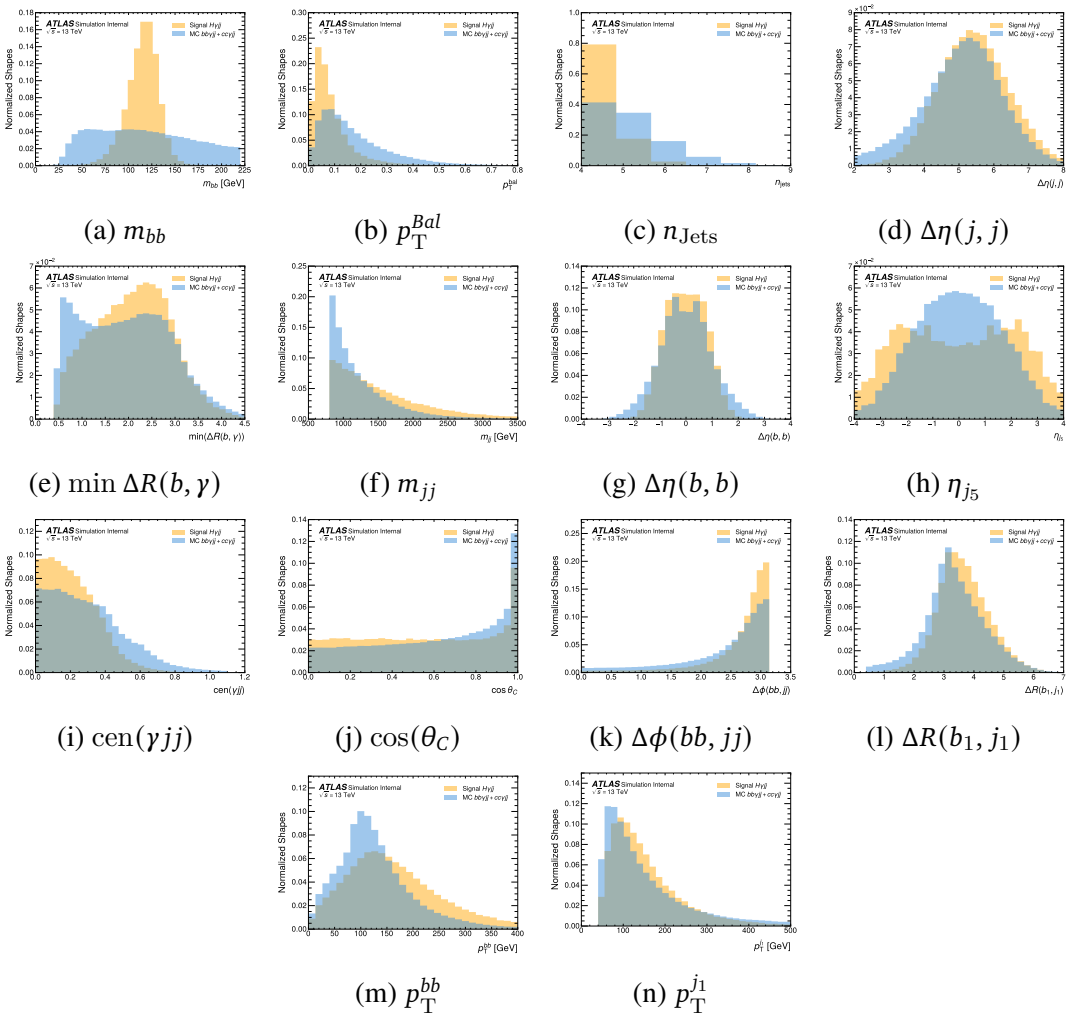
chosen. This maintained the overall η_{j5} distribution in order to minimize altering the DNN training, and was reproducible by using a seed. Variables highly correlated with m_{bb} were avoided to maintain orthogonality between the region discriminant and the fit variable, but in the new fit strategy of this analysis, the orthogonality is not necessary.

Signal and background distributions for the 14 selected NN input variables are shown in Figure 8.3. For some variables, like m_{bb} or p_{T}^{Bal}, the separation between signal and background is more evident and visible easily by eye. For others, the distinction might be more prevalent in a particular region of the phase space, as kinematics are often correlated. This is where the NN can shine, handling situations with complexly correlated variables.

The model hyperparameters that were selected through the optimization method described in Section 8.4.1 are listed in Table 8.3. Sparse Categorical CrossEntropy was used for the loss function, quantifying the distance between the true label and predicted label of training events, which is minimized in training. The ratio of the sum of weights for all training events for $bb\gamma jj$ and $cc\gamma jj$ was chosen to approximate the ratio of their predicted fiducial cross-sections in the phase space of our selection criteria, $4:1$. The ratio between the sum of weights of signal and of the two backgrounds was then optimized in the same way as the other hyperparameters, such that the final ratio is $12:4:1$ (Signal:$bb\gamma jj$:$cc\gamma jj$).

Input Variable	Description
n_{jets}	The total number of jets in the event with $p_T \geq 40\,\text{GeV}$
p_T^{bal}	Transverse momentum balance for selected final state objects
$\Delta\phi(bb, jj)$	Azimuthal angle between the plane of leading b-tag jets and the plane of the VBF jets
m_{bb}	Invariant mass of two leading b-jets
p_T^{bb}	Transverse momentum of the combination of b-jets
m_{jj}	Invariant mass of two VBF jets
$\cos\theta_C$	Cosine of the angle between the VBF jets and the two b-jets planes in the center of mass frame of the $bbjj$ system
$\min(\Delta R(b,\gamma))$	The minimum angular distance between one of the b-jets and the photon
$\Delta\eta(j,j)$	Pseudorapidity difference between two VBF jets
$\Delta R(b_1, j_1)$	Angular distance between the leading VBF jet and the leading b-tagged jet
$p_T^{j_1}$	Transverse momentum of the leading VBF jet
$\text{cen}(\gamma jj)$	Centrality of the photon with respect to the VBF jets
$\Delta\eta(b,b)$	Pseudorapidity difference between two b-jets
η_{j_5}	Pseudorapidity of a 5th jet (2 b-jets, 2 VBF jets, plus this extra 5th jet)

Table 8.2.: Kinematic variables used as input to neural network classifier. This list is ordered in decreasing correlation with between the input variable and NN score, *i.e.*, in decreasing order of importance in the NN training.



(a) m_{bb} (b) p_{T}^{Bal} (c) n_{Jets} (d) $\Delta\eta(j,j)$

(e) $\min \Delta R(b,\gamma)$ (f) m_{jj} (g) $\Delta\eta(b,b)$ (h) η_{j_5}

(i) $\mathrm{cen}(\gamma jj)$ (j) $\cos(\theta_C)$ (k) $\Delta\phi(bb,jj)$ (l) $\Delta R(b_1,j_1)$

(m) p_{T}^{bb} (n) $p_{\mathrm{T}}^{j_1}$

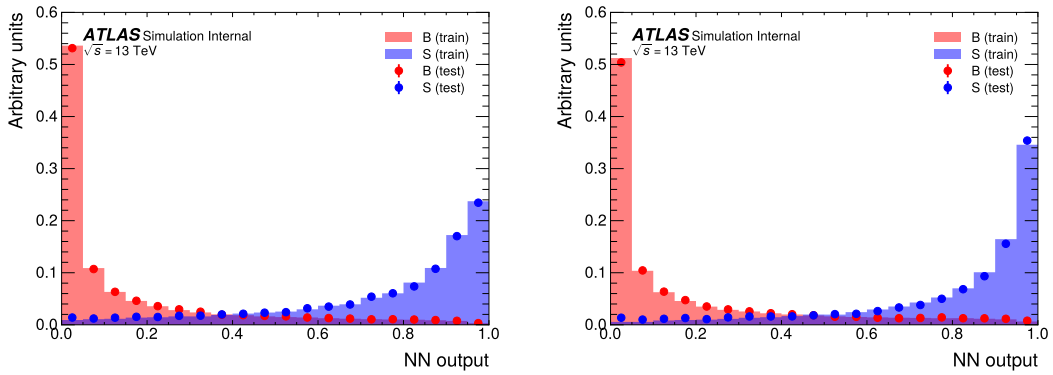
Figure 8.3.: Normalized kinematic distributions of the signal and backgrounds MVA training sets for the 14 selected input variables to the NN. Signal is in orange, and backgrounds in blue.

Hyperparameter	Value for Neural Net Training
Number of Hidden Layers	11
Number of Neurons per Layer in Bulk	256
Activation Function in Bulk	Rectified Linear Unit
Hidden Layer Activation Function	Softsign
Activation Function in Output Layer	SoftMax
Number of Neurons in Output Layer	1
Loss Function	sparse_categorical_crossentropy
Number of Training Epochs	Until Δ‘‘val_loss’’< 0.001 for 5 epochs
Batch Size	256
Learning Rate	$1e-3$
Total Weight Ratios between Samples $Signal : bb\gamma jj : cc\gamma jj$	$12 : 4 : 1$

Table 8.3.: Training hyperparameters for the deep neural network classifier.

8.4.3. DNN Performance

The outputs from the DNN training and testing is plotted in Figure 8.4. For two identical networks trained in the 2-Fold method, there is good separation between signal (S) and background (B), and also good agreement between the test (dots) and train (solid histogram) samples. Of the half of total statistics that is seen by each model, 20% of the training statistics are preserved to test for over-training and to measure training performance. The agreement between the score distributions of the training sets and test sets suggests that over-training is not an issue in the final trained model.



(a) Training Set 1

(b) Training Set 2

Figure 8.4.: DNN output scores for training and test samples show good separation between signal (S) and background (B). The two neural nets were each trained with half the MC statistics and with identical training parameters.

Chapter 9.

Signal and Background Modeling

The previous full Run 2 analysis of this Higgs signature was performed with a data-driven fit to the di b-jet invariant mass spectrum (m_{bb}), using a polynomial function for the falling non-resonant background. This strategy faced a large impact to signal sensitivity from uncertainty due to the potential presence of spurious signal. To avoid this spurious signal uncertainty and to limit statistical uncertainties as much as possible, a binned-likelihood fitting strategy is adopted here, using background and signal $H\gamma jj$ templates taken from Monte Carlo, where we profit from the higher statistics available compared to data. Flipping the script on observables, in this analysis, the fit is performed directly on the MVA classifier score distribution, using m_{bb} to define regions of purity as described in Chapter 7, as this method gives a larger signal significance even after including all systematic uncertainties.

With MC templates used to predict data in the signal extraction fitting strategy, care is taken to ensure good modeling of the signal and backgrounds. The compositions of the signal and backgrounds are estimated in Sections 9.1 and 9.2, respectively. Following, the dominant backgrounds, $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$, are in the spotlight in Sections 9.3 and

9.4, as the statistics and data agreement of these samples is maximized. Lastly, Section 9.5 takes a look at the correlations between the input kinematic variables and the DNN output score for MC and data.

§ 9.1. Signal Composition

The signal $H(\rightarrow b\bar{b})$ MC template is taken from an NLO event generation. While the emphasis of this analysis is on the VBF $H\gamma$ production mode, $VH\gamma$ contributes an estimated 8.8% of the signal sample, though it is expected to contribute less after final event selection phase space cuts. About 92.3% of the signal VBF events at generator-level are estimated to be from Higgs-production via WW-fusion, due to the suppression of ZZ-fusion by an interference effect mentioned in Section 5.2. While truth-tagged MC signal events are utilized in the MVA training to have ample training statistics, the signal fit template is extracted by evaluating the NN on only direct-tagged events.

§ 9.2. Background Composition

The templates for the $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ backgrounds use truth-tagging (see Section 7.4.1) to take advantage of as much of the statistics as possible, and tests were performed to ensure that the DNN output shapes are compatible with direct-tagged sample templates [147]. The remaining background and signal templates use direct-tagged samples (see Section 7.4.1), as is standard for analyses involving flavor-tagging.

Non-resonant backgrounds are estimated to contribute approximately 99% of the total background. Prior to any constraints from a fit to data, QCD multi-jet production

with b-quarks in association with a photon ($b\bar{b}\gamma jj$) dominates with a 70% contribution; multi-jet production with mis-tagged c-quarks in association with a photon ($c\bar{c}\gamma jj$) contributes an estimated 18%; non-resonant production of top-quarks ($t\bar{t}$) is estimated to contribute 11% of the total background cross-section. An NLO $t\bar{t}$-production sample is used after reweighting to the best-prediction NNLO cross-section. The non-resonant QCD production with light-quarks in association with a photon ($l\bar{l}\gamma jj$) is conservatively estimated to contribute $< 0.5\%$ of the total background cross-section, and an MC template is omitted.

Though resonant backgrounds contribute only 1% of the estimated total background, MC templates are still produced, since the kinematics of resonant backgrounds are more similar to signal than to the non-resonant backgrounds used in the NN training. Z-production in association with a photon and jets, $Z(\rightarrow b\bar{b})\gamma jj$, is considered with two sample productions: electroweak $Z\gamma jj$ production at NLO with no QCD couplings allowed, and QCD $Z\gamma jj$ production at LO. The contribution from $W\gamma$-production is estimated to be only 1% of that of $Z\gamma$, so no template is included for $W\gamma jj$. Other Higgs production modes have minimal contribution to the phase space, due to the required photon associated and the VBF-specific event selection cuts. Nevertheless, ggH and $t\bar{t}H$ MC templates are created and are treated as backgrounds to the predominantly-VBF Higgs signal.

§ 9.3. Truth-Tagging vs Direct-Tagging for $b\bar{b}/c\bar{c}\gamma jj$

The definitions of truth-tagging and direct-tagging are given in Section 7.4.1. To maximize statistics of simulated events in our low acceptance phase space, truth-tagging

samples are employed in the cases of signal and non-resonant background MVA training samples and in the fit templates for the two dominant backgrounds (though direct-tagging is used for the signal fit template). While the truth- and direct-tagged samples have differences in predicted yield as can be seen in Section 9.3.1, the samples do not show significant differences in the shapes of the fit observable as demonstrated in Section 9.3.2. With a floating norm factor to fix the non-resonant background normalization to data such that the starting yield is irrelevant, this evidence justifies the choice to use truth-tagging for these background fit templates to minimize MC statistical uncertainty.

9.3.1. Comparison of Yields of Truth- vs Direct-Tagged Samples

Table 9.1 and 9.2 show the differences in yield between direct-tagged and truth-tagged samples for the signal and dominant backgrounds, respectively. Typically, a truth-tagged for a signature with two tagged jets would contain all events passing the other event selection requirements, including events with 0 b-tagged jets (by direct-tagging standards). However, in the case of this analysis, at least one b-tagged jet is required by the dominant triggers in use, so a piece of the phase space is missing. Because the truth-tagging scale factors were not calibrated for this, the truth-tagged samples, while having more MC statistics, have a significantly lower yield (sum of event weights) than the direct-tagged counterparts. This is particularly evident for $c\bar{c}\gamma jj$.

For our fit templates, only truth-tagged $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ samples are used. For signal and all other backgrounds, direct-tagged MC events and scale factors are used to craft the fit templates. The observed difference in yield for our two dominant backgrounds is not of huge concern, though, since they share a floating normalization factor that fits their shared yield to data. Still, the ratios between the $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ differ by roughly

28% between the truth-tagged and direct-tagged cases, truth-tagged being the higher bb-to-cc ratio. To allow for the freedom in the ratio between these two samples in the fit, large normalization uncertainties on these backgrounds are decorrelated between the two samples.

Sample	mc16a	mc16d	mc16e	Total
Direct Tag Signal	17.0	23.5	30.9	71.3
Truth Tag Signal	16.3	22.5	29.7	68.5

Table 9.1.: Comparison of direct-tagging and truth-tagging methods show small differences in yield in signal for MC16a, d, and e. Signal yields match within 5%.

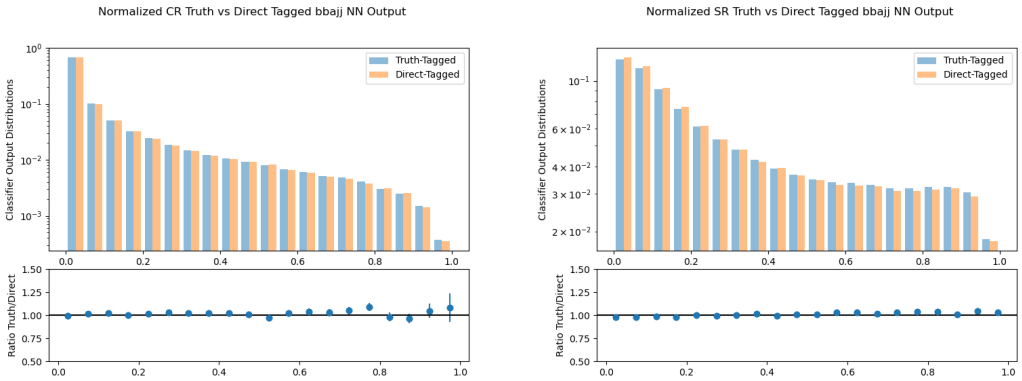
Sample	Direct-Tagged Yield	Truth-Tagged Yield
$b\bar{b}\gamma jj$	46245	43284
$c\bar{c}\gamma jj$	14669	10695

Table 9.2.: Comparison of direct-tagged and truth-tagged yields for the two dominant nonresonant QCD backgrounds. Differences are significant, but are not of concern due to the floating normalization factor fixing the yields of these two samples and the ratio freedom introduced from the theoretical uncertainties in the fit model.

9.3.2. NN Output Shape for Truth vs Direct-Tagged Backgrounds

To take full advantage of the large-scale Monte Carlo statistics generated for this analysis, the choice was made to use truth-tagged samples for the two dominant non-resonant backgrounds, $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$. It is known that there are yield differences between the truth-tagged and direct-tagged samples, but these are not of concern due to the floating

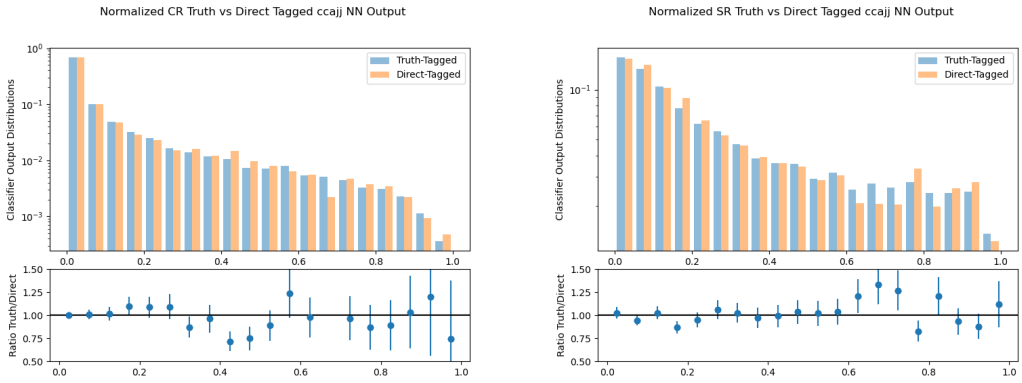
normalization factor for these samples in the fit. Here, we check that the NN output shapes are also compatible. Figures 9.1 and 9.2 show the neural net score shapes for the truth and direct-tagged $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ samples in the control and signal regions. The shapes are compatible within statistical uncertainties, and any differences are covered by existing systematics in the fit.



(a) Control Region

(b) Signal Region

Figure 9.1.: Comparison of Neural Net Score shape between truth- and direct-tagged $b\bar{b}\gamma jj$ MC16 samples in the control region (left) and the signal region (right).



(a) Control Region (b) Signal Region

Figure 9.2.: Comparison of Neural Net Score shape between truth- and direct-tagged $c\bar{c}\gamma jj$ MC16 samples in the control region (left) and the signal region (right).

§ 9.4. Kinematic Reweighting for Data and Monte Carlo Closure

The fitting strategy extracts the signal directly from the NN output score distribution, using a control and signal region defined by cuts on m_{bb}. This fit relies on MC template distributions, so it is important to ensure that the background MC agrees well with data. Discrepancies exist between data and MC for several kinematic variables, particularly noted for n_{jets}, $\min(\Delta R(b,\gamma))$, $\Delta\eta(j,j)$, and m_{bb}. This was also the case for the previous full Run 2 analysis, and is due to the difficulty of modeling non-resonant multijet backgrounds at leading-order.

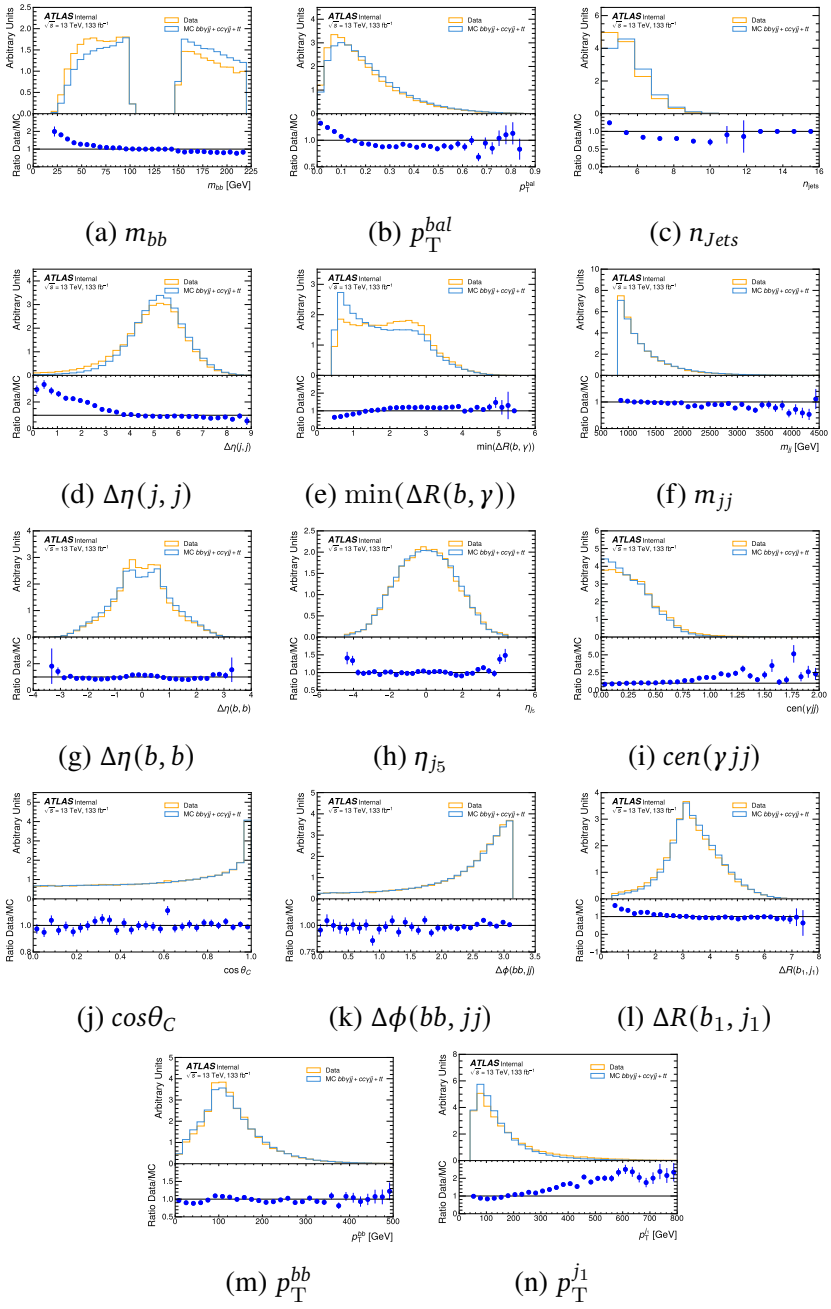
To improve data-to-MC agreement in the NN output distribution, reweighting scale factors are extracted from the kinematic distributions that are serving as the 14 MVA input variables (those listed in Table 8.2). Only the two dominant backgrounds, $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$, receive reweighting scale factors, and these backgrounds are reweighted to-

gether with identical kinematic scale factor functions. The next leading background, $t\bar{t}$, is generated at higher order, so it is not kinematically reweighted, though its contribution is considered. The 15 scale factors (photon centrality is reweighted twice due to lingering discrepancies remnant after the first round of scale factors) are extracted from the control region where there is negligible signal contamination, such that the signal will not be reweighted away. The scale factor functions are then extrapolated and applied to both the control and signal regions.

To perform reweighting, the ratio of the normalized difference between data and MC $t\bar{t}jj$ to the normalized sum of MC $b\bar{b}\gamma jj + c\bar{c}\gamma jj$ is considered for binned distributions of each kinematic input variable: $\frac{norm(\text{data}-t\bar{t})}{norm(b\bar{b}\gamma jj+c\bar{c}\gamma jj)}$. To this ratio in the m_{bb} sideband control region, a 6th-order polynomial is fit successively for each input variable, in the reverse order of their correlation with NN score. The function is then applied to the MC $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ samples, in both the control and signal regions, in order to scale these dominant backgrounds to better represent the data. These kinematic reweighting scale factors account for any mismodeling these non-resonant QCD backgrounds, as well as for any contaminations from backgrounds not otherwise considered, such as jets faking photons. Figure 9.3 shows the kinematic distributions in the control region before kinematic reweighting, comparing data with the sum of these three dominant backgrounds $b\bar{b}\gamma jj + c\bar{c}\gamma jj + t\bar{t}$. The same normalized comparisons after all the reweighting scale factors have been applied is shown in Figure 9.4 for the control region. The kinematic comparison is given for the signal region in Figure 9.5 after extrapolation and application of the reweighting scale factors.

NN training is redone after kinematic reweighting to obtain the final MVA model, such that the modified event weights can impact the loss function and training results

accordingly. Despite the large number of compounding scale factors, the overall impact on the NN output is washed out and relatively small. Additionally, the reweighted result is quite robust to procedural changes, as is reflected in the shape uncertainty due to this kinematic reweighting discussed in Section 10. Very good closure between MC and data shapes after kinematic reweighting can be seen in the control region and the background-rich areas of the signal region, *i.e.*, in the Higgs mass window $100 < m_{bb} < 150$ GeV but with a low NN score less than 0.6. Figure 9.6 shows the NN output distributions of the MC backgrounds before and after kinematic reweighting in comparison with data in the control and signal regions. In the background-rich bins where data and background MC events can be compared, the data and reweighted MC agree within statistical uncertainties, and no additional closure efforts or uncertainties are necessitated.



(a) m_{bb} (b) p_{T}^{bal} (c) n_{Jets}

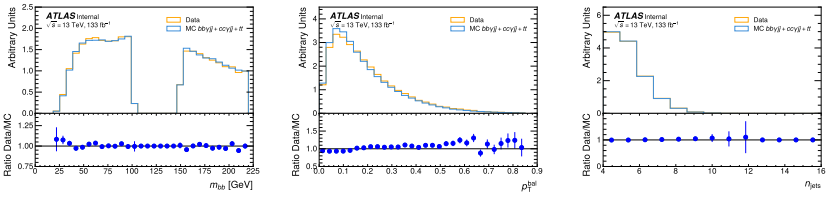
(d) $\Delta\eta(j,j)$ (e) $\min(\Delta R(b,\gamma))$ (f) m_{jj}

(g) $\Delta\eta(b,b)$ (h) η_{j_5} (i) $cen(\gamma jj)$

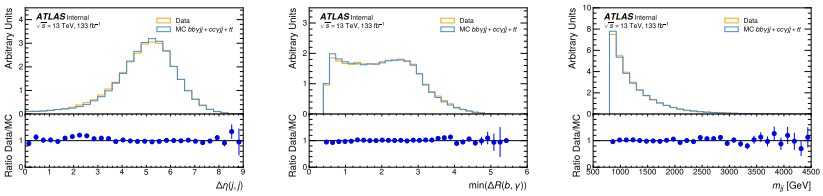
(j) $cos\theta_C$ (k) $\Delta\phi(bb,jj)$ (l) $\Delta R(b_1,j_1)$

(m) p_{T}^{bb} (n) $p_{\mathrm{T}}^{j_1}$

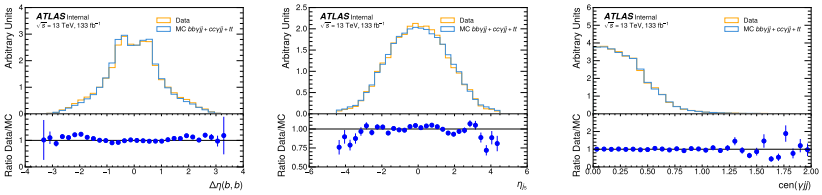
Figure 9.3.: Control region (m_{bb} sidebands) Monte Carlo $bb\gamma jj + cc\gamma jj$ versus data normalized comparison for all kinematic variable inputs to the MVA with $t\bar{t}$ MC subtracted from the data distribution. All input variables (except n_{Jets}) were fit with analytic functions for reweighting because MC and data have visible discrepancies in these distributions.



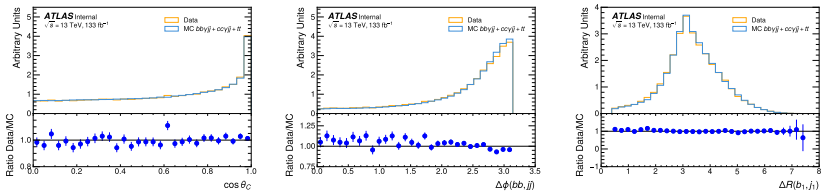
(a) m_{bb} (b) p_{T}^{bal} (c) n_{Jets}



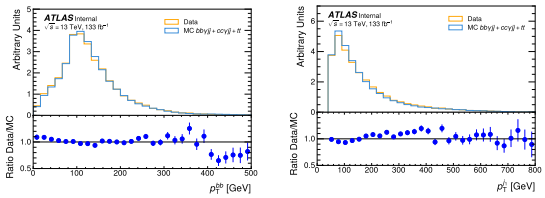
(d) $\Delta\eta(j,j)$ (e) $\min(\Delta R(b,\gamma))$ (f) m_{jj}



(g) $\Delta\eta(b,b)$ (h) η_{j_5} (i) $cen(\gamma jj)$

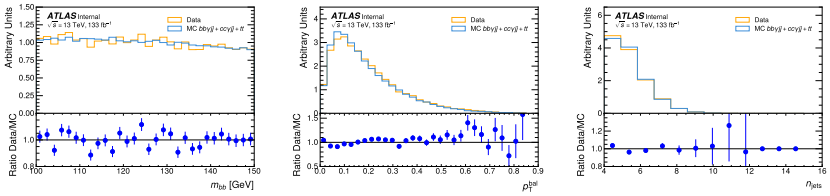


(j) $cos\theta_C$ (k) $\Delta\phi(bb,jj)$ (l) $\Delta R(b_1,j_1)$



(m) p_{T}^{bb} (n) $p_{\mathrm{T}}^{j_1}$

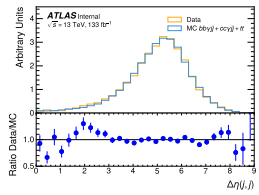
Figure 9.4.: Control Region Reweighted Monte Carlo $b\bar{b}\gamma jj + c\bar{c}\gamma jj$ and $t\bar{t}jj$ to Data normalized comparison kinematic variables after kinematic reweighting. The variables shown are those that are input into the MVA. All variables were used in kinematic reweighting.



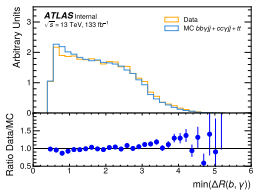
(a) m_{bb}

(b) p_{T}^{bal}

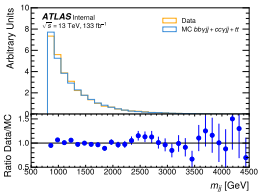
(c) n_{Jets}



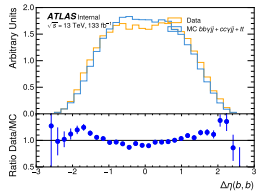
(d) $\Delta\eta(j,j)$



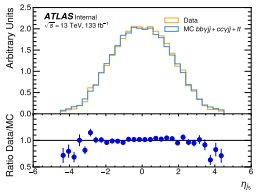
(e) $\min(\Delta R(b,\gamma))$



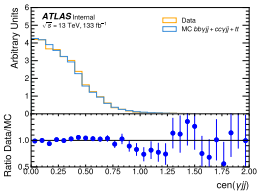
(f) m_{jj}



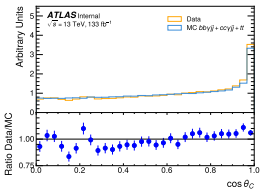
(g) $\Delta\eta(b,b)$



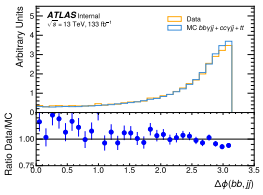
(h) η_{j_5}



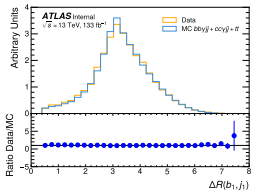
(i) $cen(\gamma jj)$



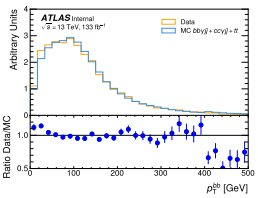
(j) $cos\theta_C$



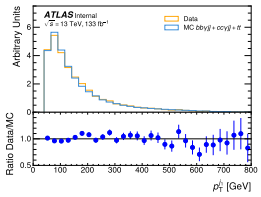
(k) $\Delta\phi(bb,jj)$



(l) $\Delta R(b_1,j_1)$

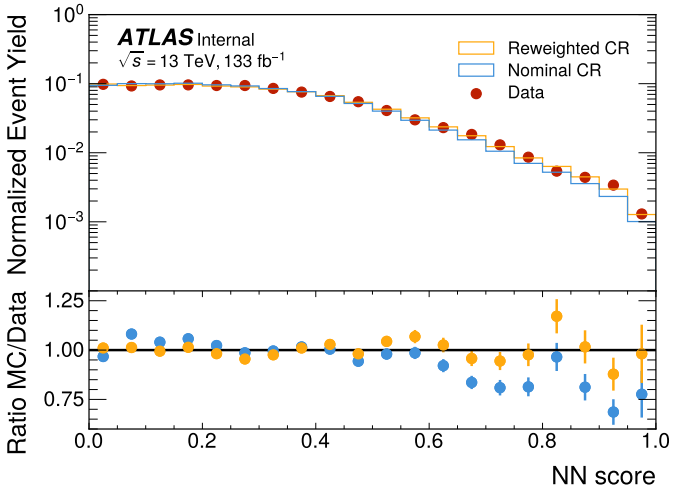


(m) p_{T}^{bb}

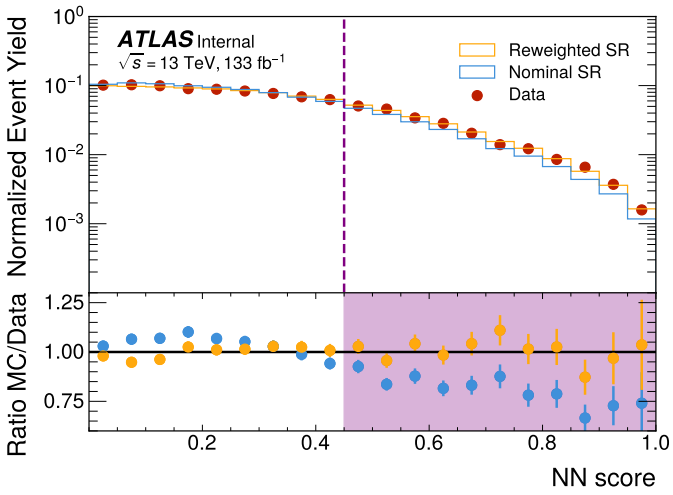


(n) $p_{\mathrm{T}}^{j_1}$

Figure 9.5.: Signal Region Reweighted Monte Carlo $b\bar{b}\gamma jj + c\bar{c}\gamma jj$ and $t\bar{t}jj$ to Data normalized comparison kinematic variables after kinematic reweighting. The variables shown are those that are input into the MVA. All variables were used in kinematic reweighting.



(a) CR (m_{bb} Sidebands)

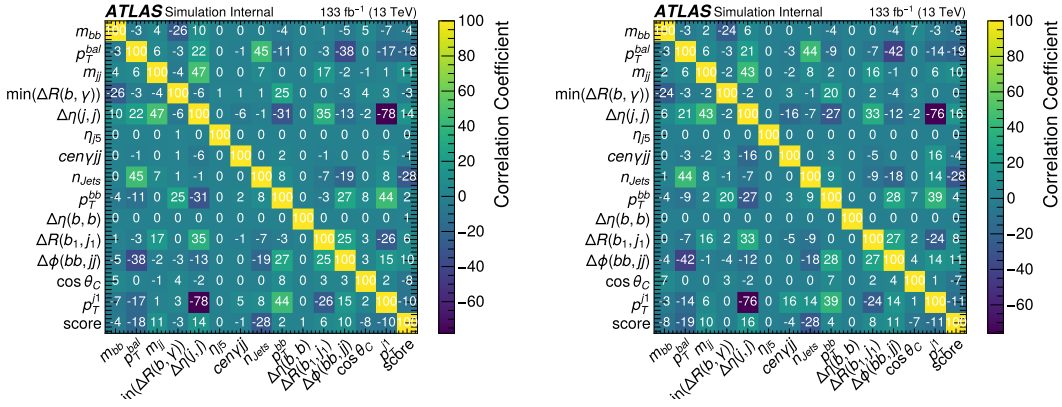


(b) SR (Higgs Mass Window)

Figure 9.6.: MC versus data comparison of the NN output distributions of all backgrounds with and without kinematic reweighting performed. Reweighting scale factors are extracted from the comparison with data in the control region m_{bb} sidebands (left), as well as in the low score bins of the signal region (right) where signal has negligible contamination. The purple dashed line marks the divide between the low score SR bins, which contribute to the validation of the reweighting scale factors, and the high score bins where signal contribution is present. MC and data are individually normalized, making this a direct comparison of neural net output shape.

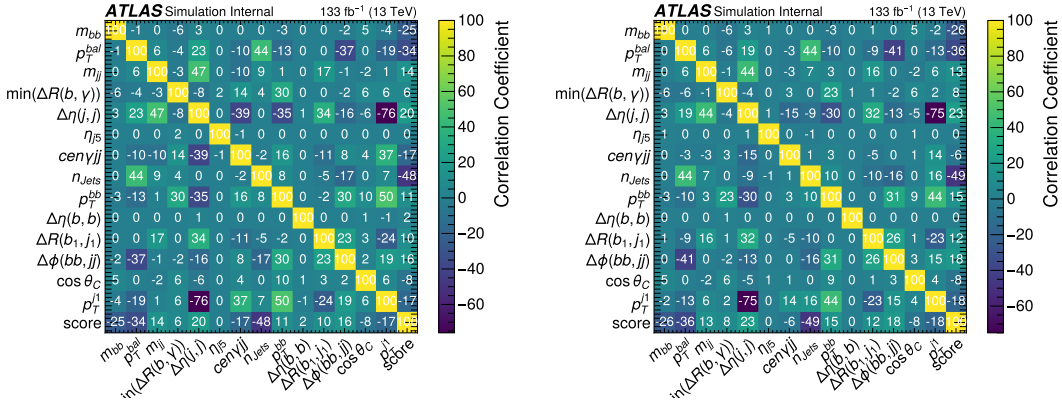
§ 9.5. Correlation between DNN Input Variables and Score for MC and Data

Within statistical uncertainties, there is excellent closure of the DNN output score shapes between data and MC after kinematic reweighting in both the control and signal regions, as was seen in Figure 9.6. Even still, there might be concern that there could be discrepancies persistent in the low and high m_{bb} sidebands of the control region, with opposite impacts which are canceling each other when combined. One test to reassure that the DNN is treating the kinematic information similarly between MC and data, and that the closure in output distributions is not a coincidence, is to look at the correlation between input variables and output score. Data and MC $b\bar{b}\gamma jj + c\bar{c}\gamma jj$+signal are compared in Figures 9.7, 9.8, 9.9, and 9.10, respectively for the inclusive phase space, signal region Higgs mass window, low control region sideband, and high control region sideband. The correlations between distributions are calculated using the linear correlation metric with event weights considered (after kinematic reweighting), though it is admitted that not all correlations can be described well with this linear metric. The most relevant row is the bottom one, which shows correlation with the output score. For all regions, data and MC show similar behavior across all input variables for linear correlation with output score, supporting the case that the NN is truly treating MC and data similarly, and the output score distribution closure is not coincidental.



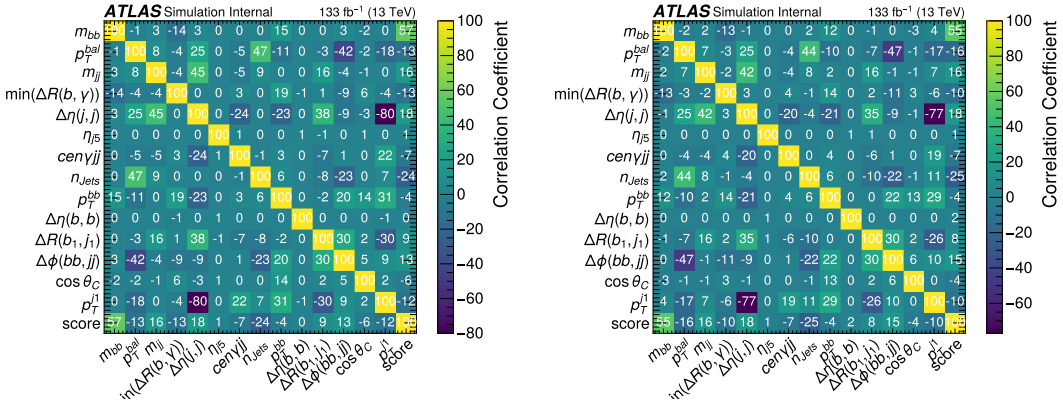
(a) MC (b) Data

Figure 9.7.: Correlation (r^2) matrices between kinematic input variables and NN score for MC and data inclusively for the entire range of m_{bb}, control and signal regions.



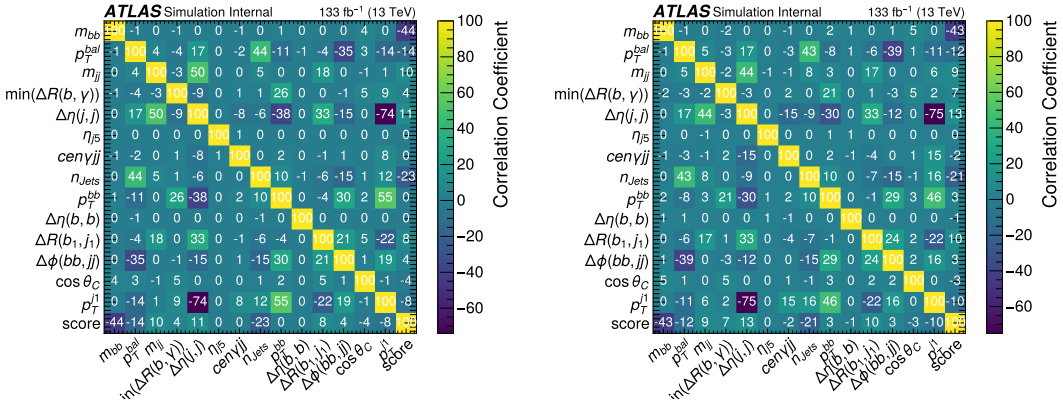
(a) MC (b) Data

Figure 9.8.: Correlation (r^2) matrices between kinematic input variables and NN score for MC and data within the signal region, $100 \leq m_{bb} \leq 150$ GeV.



(a) MC (b) Data

Figure 9.9.: Correlation (r^2) matrices between kinematic input variables and NN score for MC and data for the low sideband of the control region, $m_{bb} \leq 100$ GeV.



(a) MC (b) Data

Figure 9.10.: Correlation (r^2) matrices between kinematic input variables and NN score for MC and data for the high sideband of the control region, $150 \leq m_{bb} \leq 220$ GeV.

Chapter 10.

Systematic Uncertainties

Theoretical and experimental uncertainties in this analysis are characterized with a total of 78 nuisance parameters or degrees of freedom. These nuisance parameters (NPs) each have the ability to modify the normalizations and shapes of the signal and background MC templates. Unlike free-floating norm factors which receive flat priors, NPs have Gaussian restrictions, such that the target metric of the fit optimization, the profile likelihood, receives a penalty for large pulls away from 0 following a normal distribution. For each NP, histograms are input as systematic variations which represent a 1σ pull from the nominal template distributions, and in this way, these finite degrees of freedom give infinitely many linear combinations of morphings to be used to best describe the data. The systematic uncertainties *give* the freedom to the fit, so accurately estimating the ways in which various experimental and modeling uncertainties manifest in the analysis model and the impacts these have on the measurement is critical.

The impacts of these nuisance parameters on the signal strength uncertainty ($\sigma(\mu_H)$) are categorized and summarized in Table 10.1. The grouped impact of each uncertainty category is calculated by performing a fit with the category's corresponding nuisance

parameters fixed, and the difference in variance with that of the nominal fit with all nuisance parameters included is attributed to that group of systematics. Template smoothing is employed on all systematic variations to reduce the over-inflation of uncertainties due to correlated statistical fluctuations. Further details on the particular nuisance parameters in use are given in the following sections. Section 10.1 is devoted to the statistical uncertainties, while the theoretical modeling and experimental systematics are described in Sections 10.2 and 10.3, respectively.

§ 10.1. Statistical Uncertainties

Data statistics are limiting in this channel, due to the small cross section of $H\gamma$ production at $\sqrt{s} = 13\,\mathrm{TeV}$. Data statistical uncertainty contributes $\sigma(\mu_H) = 0.6$, dominating the overall uncertainty on signal strength. While data statistical uncertainty dominates, great attention was paid to reduce experimental and modeling uncertainties as much as possible. To reduce statistical uncertainties of the Monte Carlo background templates with respect to the previous publication, new and extended $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ samples were simulated with large statistics, reducing the statistical uncertainty contributions from MC to the signal strength uncertainty to a total of $\sigma(\mu_H) = 0.25$.

§ 10.2. Modeling and Theory Uncertainties

The uncertainty due to the absence of higher order calculation corrections is estimated through reweighted MC evaluated at varied renormalization and factorization scales, as expanded upon in Section 10.2.1 [155]. Uncertainties related to the choice in parton

distribution function are evaluated with reweighted MC evaluated on different PDF variations, as explained in Section 10.2.2. Sections 10.2.3 and 10.2.4 describe the methods for evaluating the uncertainties due to choice of parton shower generator and the uncertainties related to the kinematic reweighting procedure, respectively, which both supply significant shape freedoms to the fit. Combining these shape variations with the other modeling uncertainties on renormalization and factorization scale and PDF choice, the modeling uncertainties contribute a small fraction of the overall uncertainty on signal strength, with an impact of $\sigma(\mu_H) = 0.23$.

10.2.1. Theory systematic due to the choice of QCD scale

As alluded to in Sections 6.1 and A.1, elements of cross-section calculations are dependent on choices of factorization and renormalization scales [155]. Uncertainties on the signal cross section based on the choice of μ_R (Renormalization scale) and μ_F (Factorization scale) are evaluated by varying each parameter independently by scale factors of 0.5, 1.0, and 2.0. These variations were stored during event generation with MadGraph5 and were evaluated using 7-point variations [156]. Applying an envelope evaluation method to the variations, we eliminate the two off-diagonal variations ($\{\mu_R, \mu_F\}\times\{0.5, 2.0\}$, $\{2.0, 0.5\}$) and take the highest variation, in each bin, among ($\{\mu_R, \mu_F\}\times\{0.5, 0.5\}$, $\{1.0, 0.5\}$, $\{0.5, 1.0\}$, $\{1.0, 1.0\}$, $\{1.0, 2.0\}$, $\{2.0, 1.0\}$, $\{2.0, 2.0\}$) as our QCD scale uncertainty. Figure 10.1 demonstrates the impact of these variations on the signal NN output shape. Cross-section and acceptance uncertainties are included. The shape variation is also included in our implementation of this systematic, even though it is small.

The identical approach as with signal is used in the case of our $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ back-

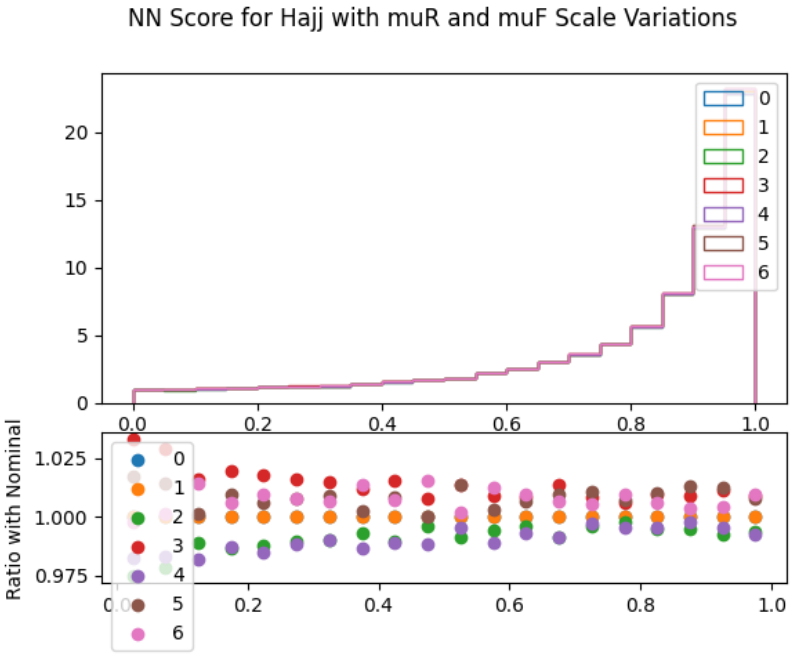


Figure 10.1.: μ_R and μ_F variations on DNN output from signal VBF Hajj MC sample. The 7 variations are from varying each parameter independently by scales of 0.5, 1.0, and 2.0. The upper and lower envelopes of these variations are taken as the up and down systematic variations.

ground samples, by taking the envelope of variations in factorization and renormalization scale parameters. Both overall cross-section and acceptance variations are considered, though the acceptance variation is quite small in comparison (0.5%). The effect from μ_R and μ_F variations on DNN output for the dominant backgrounds are shown in Figure 10.2.

To give freedom to the ratio of $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ backgrounds, scale uncertainties are decorrelated between the two samples, despite their identical simulation pipelines. In doing so, freedom is given to the ratio between $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ backgrounds, which otherwise are fixed together with a shared floating normalization factor. As seen in Section 9.3, there are significant normalization differences between the truth-tagged background samples and the direct-tagged, physically motivating this addition of ratio freedom. All samples are then uncorrelated, and have individual nuisance parameters, since different MC and parton shower generators are expected to have different responses to these scale uncertainties.

The μ_R and μ_F systematics have a a fairly negligible impact on the non-resonant QCD background shape, though shape variation is still considered in the fit. Though these shape effects are small, we did find that they were being utilized in our background-only fits to data to adjust the shapes of our dominant backgrounds. This created an instability due to a degeneracy between the large normalization impacts of the these scale uncertainties and the floating background normalization factor. To avoid this degeneracy, the choice was made to decorrelate the shape and norm impacts of these scale uncertainties on the two dominant backgrounds only, $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$. In doing so, the normalization-only nuisance parameters, labeled muRmuF_variation_Acc, adjust the ratio between the $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ samples, and the overall background normalization is

fixed by the floating background normalization factor as intended.

For $t\bar{t}$ production, the NLO sample is normalized to the best known cross-section at NNLO in QCD including resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms [157, 158] with the Top++2.0 program [123]. Therefore, we use the cross-section uncertainty on this higher-order estimate taken from the ATLAS Top Group [159]. From this recommendation, we implement an overall relative uncertainty due to scale variation of $+2.5\% - 3.6\%$ on $t\bar{t}$.

For the background of $Z\gamma jj$, the scale variations on event weight were not stored for the samples in use. Since the background is so small, though, we elected to take a conservative estimate on the magnitude of scale variations from another analysis. It is expected that the variations on our QCD sample would be comparable to the magnitude of variation quoted by the exotics search, which produced a $Z\gamma$ QCD sample with missing energy. Matching their finding, we implement an overall uncertainty of 16% for the entire $Z\gamma jj$ background, expecting this to be an overestimate for the $Z\gamma jj$ EWK sample.

10.2.2. Theory systematic due to the choice of PDF and α_s

Uncertainties due to the choice of PDF+α_s are evaluated by the event generator using PDF reweighting [160]. On-the-fly weights are stored during event generation with MadGraph5 for 43 error members of the PDF set `PDF4LHC21_40_pdfas` [161]. Member 0 is the nominal, while members 1-40 are PDF variations $\alpha_s(m_Z^2) = 0.118$ and members 41-42 are α_s variations for $\alpha_s(m_Z^2) = 0.117$ and $\alpha_s(m_Z^2) = 0.119$, respectively. The uncertainty is calculated as the standard deviation of these variations and is incorporated as an overall systematic in the fit, with no shape variation. The PDF and α_s combined uncertainty on the signal cross section is 2.2%.

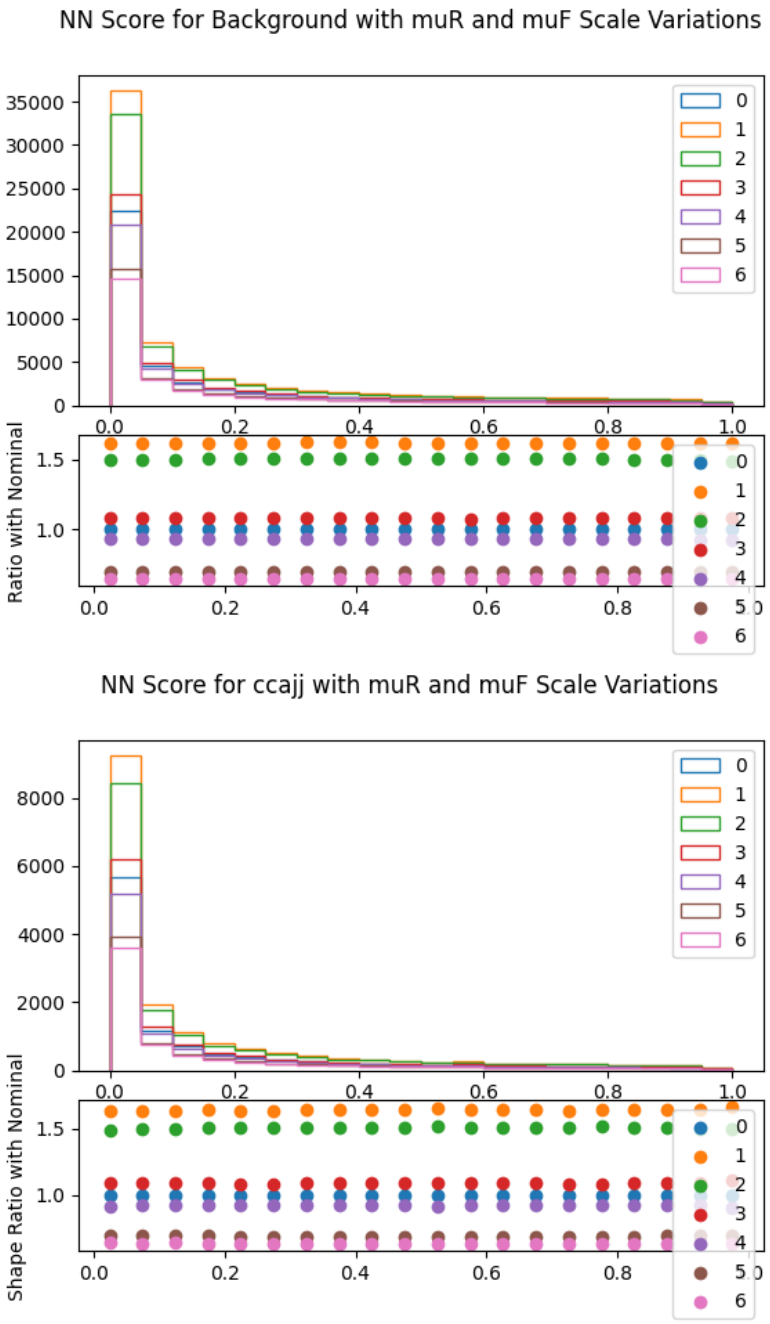


Figure 10.2.: μ_R and μ_F variations on DNN output from QCD $bb\gamma jj$ (left) and QCD $cc\gamma jj$ (right) MC samples. The 7 variations are from varying each parameter independently by scales of 0.5, 1.0, and 2.0.

For $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ backgrounds, PDF and α_s uncertainties are calculated independently. The PDF sets have 100 variations, and $\alpha_s(m_Z^2) = 0.117$ and $\alpha_s(m_Z^2) = 0.119$ are considered separately. Variation in generator level cross-section and acceptance are both considered. The uncertainty on $b\bar{b}\gamma jj$ background is 3.15% considering PDF and α_s uncertainties, and for $c\bar{c}\gamma jj$, the PDF and α_s variations impact the yield by 3.28%.

For $t\bar{t}$ production, the PDF and α_s uncertainties are taken from the results reported by the top group on the NNLO cross-section variation [159]. These variations contribute a relative 2.5% variation on the $t\bar{t}$ yield.

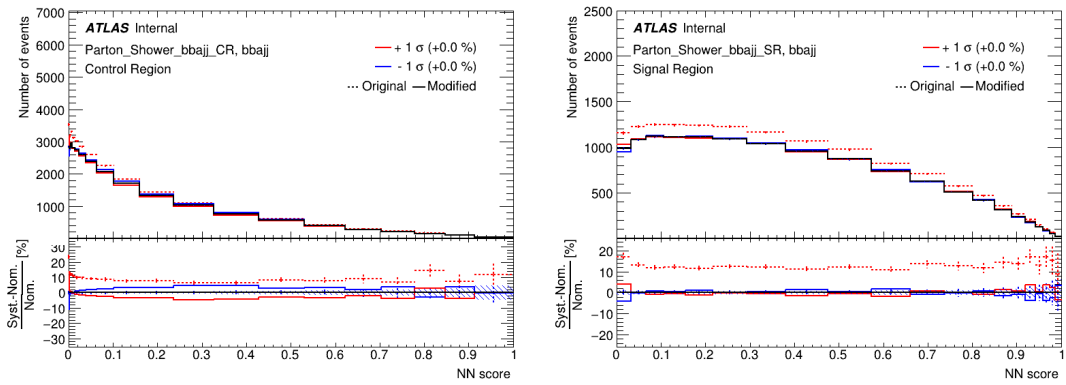
PDF and α_s variations are not considered for the $Z\gamma jj$, since variations on this small background would have very little impact on the signal measurement, and the μ_R and μ_F variations are already overestimated for the sample. For the other Higgs processes that serve as background to our VBF signal, an overall 20% normalization uncertainty is introduced to cover modeling uncertainties. These Higgs processes contribute less than 0.5% of the overall yield and tend to lie in the low NN score region, far from our signal peak. As a result, systematic variations on these processes have very little impact on the signal measurement, thus it was deemed unnecessary to study these variations more carefully. The introduction of a 20% normalization uncertainty on the smaller Higgs MC samples is believed to be a conservative approach.

10.2.3. Theory systematics due to parton shower

Modeling uncertainties related to the choice in parton shower generator provide some of the largest shape variations in the fit. Parton shower variations are considered for our three dominant backgrounds (which collectively make up over 99% of our phase space), as well as for signal. For $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ backgrounds, the original sample production

utilized MadGraph and Pythia. Alternative samples have been produced at truth-level, using MadGraph and Herwig. These samples were produced with similar statistics to the original template samples, but due to the limitations in computing availability and time, they were only produced at truth-level, unlike the fully-simulated template samples. Unfortunately, truth-level and fully simulated MC have significantly different NN output shapes, so a method was designed to transfer the differences in kinematics between truth-level Pythia and Herwig -produced samples to the impacts on the fully simulated NN score distribution. In arbitrary order and one at a time (iteratively), each of the 14 kinematic input variables are compared at truth-level between the Pythia and Herwig -generated samples. The ratio of Herwig to Pythia is taken in 20 bin distributions, and applied as a transfer function on the event weights of the fully simulated Pythia samples. The kinematic comparisons at truth-level before and after the transfer functions are applied can be seen in Appendix D. Thus, the transfer functions adequately represent the differences in kinematics between the samples produced with different parton shower generators, and we extrapolate these differences to the fully-simulated NN output distribution. The NN output of the fully-simulated Pythia sample with these transferred event weights mimics the NN output of a fully-simulated Herwig sample and is shown in Figures 10.3 and 10.4. The variation is included in the fit as a shape-only systematic variation and symmetrized. A difference in behavior at low-score is noted compared the remaining distribution. To avoid constraint on the signal-like high-score bins in the signal region from this large departure in the background-like bins of the control region, the control and signal regions are decorrelated in the fit for the parton shower uncertainties of $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ and given unique nuisance parameters.

With the large nature of the shape variations of these background parton shower sys-

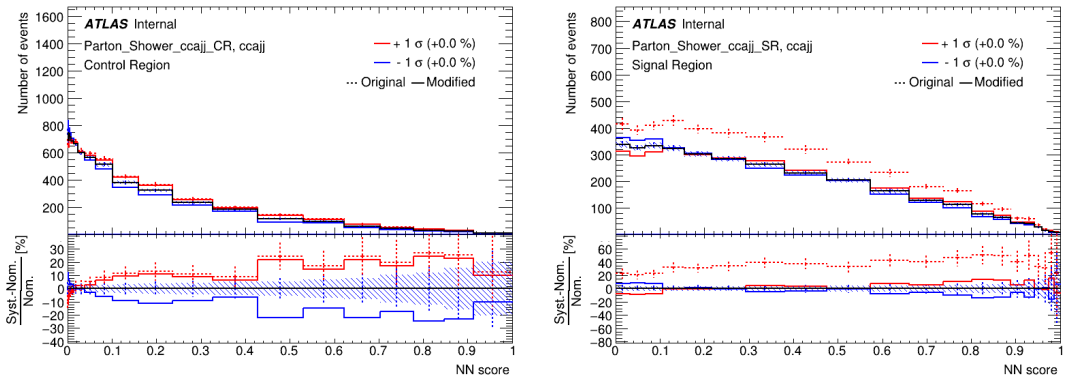


(a) CR (b) SR

Figure 10.3.: Parton shower systematic variations included in fit for $bb\gamma jj$ in the control and signal regions. Shape-only systematics are included, and impacts on normalization are not considered. Control and signal regions are given separate nuisance parameters, such that the behavior in the background-like bins of the control region does not constrain the pull for high score bins in the signal region.

tematics, it is comforting to know that this modeling uncertainty accurately represents a source of mismodeling in the MC templates. There are a few variables that, even after kinematic reweighting, show some residual discrepancies between MC and data in the signal region, as can be seen in Section 9.4 and Chapter 11. Some examples are $\Delta\eta(b,b)$, $\min \Delta R(b,\gamma)$, and p_{T}^{bal}. The propagation of the parton shower variations to those variables can be seen in Figure 10.5. Comparing these propagated variations to the lingering signal region MC-data discrepancies, the eigen-directions seem to match, suggesting that the variations appear to accurately grasp some aspect of mismodeling.

For $t\bar{t}$ production, the parton shower and matrix element variation samples were already officially produced. Both variations are incorporated, giving shape and normalization differences for $t\bar{t}$. Parton shower variation on $t\bar{t}$ varies the sample normalization by almost 10%. Matrix element variation is small in the control region at only 0.6%,



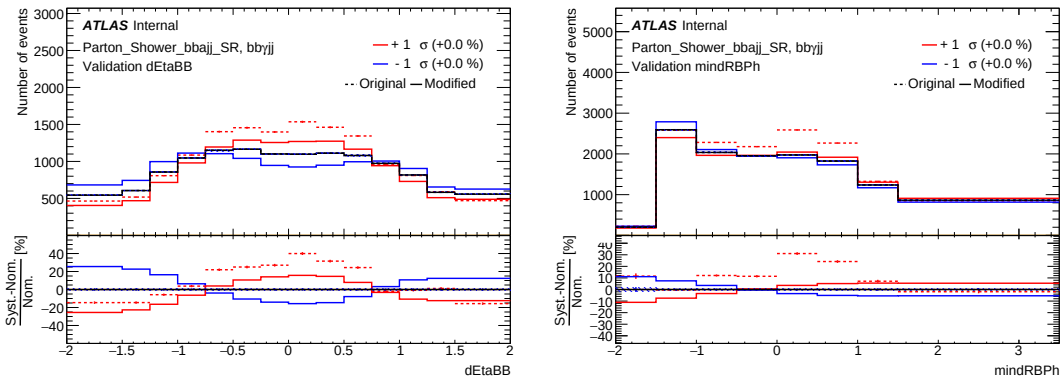
(a) CR (b) SR

Figure 10.4.: Parton shower systematic variations included in fit for $cc\gamma jj$ in the control and signal regions. Shape-only systematics are included, and impacts on normalization are not considered. Control and signal regions are given separate nuisance parameters, such that the behavior in the background-like bins of the control region does not constrain the pull for high score bins in the signal region.

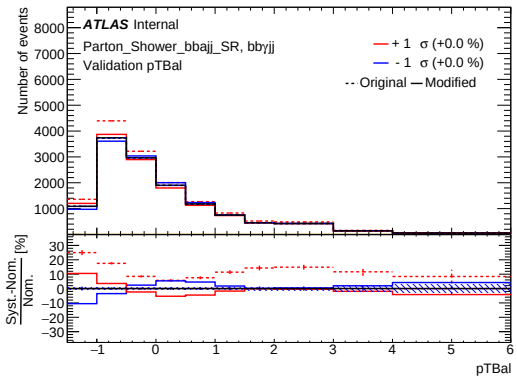
but is higher in the signal region at 5.3%. These $t\bar{t}$ production modeling variations have the largest impact on signal measurement uncertainty of the parton shower uncertainties considered.

The next-to-leading-order MC signal samples in this analysis use the Herwig 7 defaults of `H7-PS-MMHT2014LO` for the parton shower tune, and `H7-UE-MMHT` for the underlying event tune. Following the Physics Modeling Group recommendations, the theory uncertainties on the parton shower are studied within a single shower program by varying the ‘‘HardScaleFactor.” This changes the maximum allowed transverse momentum for shower emissions to $p_{\mathrm{T_{new}}}^{\mathrm{max}} = x \times p_{\mathrm{T_{new}}}^{\mathrm{max}}$, where x is the Hard Scale Factor. The Hard Scale Factor variations used in this analysis were $x = 0.5$ and $x = 2.0$.

A higher statistics formal MC production for signal with Hard Scale Factor variations

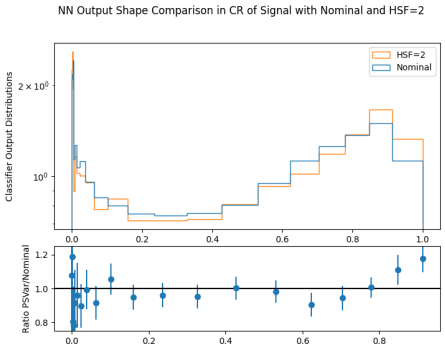


(a) $\Delta\eta(b,b)$ (b) $\min \Delta R(b,\gamma)$

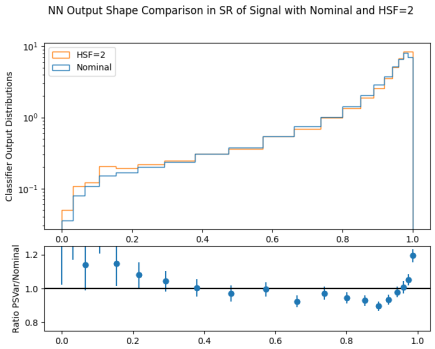


(c) p_{T}^{bal}

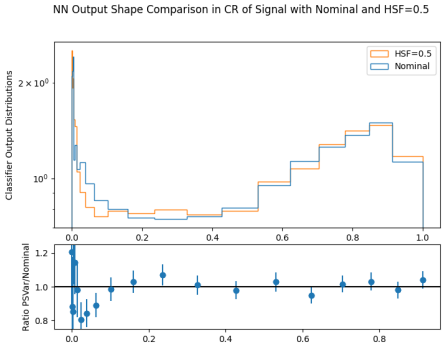
Figure 10.5.: Propagation of signal region parton shower systematic variations of $b\bar{b}\gamma jj$ to its kinematic distributions for $\Delta\eta(b,b)$, $\min \Delta R(b,\gamma)$, and p_{T}^{bal}.



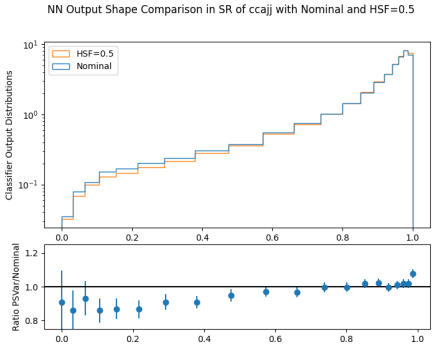
(a) CR, HSF=2



(b) SR, HSF=2



(c) CR, HSF=0.5



(d) SR, HSF=0.5

Figure 10.6.: Parton shower systematic variations included in fit for signal $H\gamma jj$ in the control and signal regions. Two parameter variations are considered using Herwig generator: a hard-scale factor of 2 (top) and 0.5 (bottom). Shape-only systematics are included, and impacts on normalization are not considered. Control and signal regions correlated with a single nuisance parameter.

was obtained at truth level and is incorporated into the fit systematics in a similar fashion as for the 2 dominant $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ backgrounds. Kinematically-dependent scale factors were derived from comparisons at truth-level with the nominal HardScaleFactor=1 Herwig signal sample. These scale factors were applied to the fully simulated nominal signal template to mimic fully-simulated NN output distributions with varied Hard Scale Factors. The main difference to the application of this strategy for the two dominant backgrounds is that there are two signal variations, rather than a single symmetrized variaion as for the backgrounds. Additionally, the spiked behavior that was seen at low NN score for the backgrounds and motivated the decorrelation of signal and control regions for those samples is not present for signal, so the control and signal regions remain correlated in the signal parton shower nuisance parameter, but again, a shape-only systematic is used. The differences in signal shape with varied Hard Scale Factors are depicted in Figure 10.6 for the control and signal regions.

10.2.4. Kinematic Reweighting Uncertainty

Kinematic reweighting on DNN input variables has effects on the DNN output shape. As shown in Figure 10.7, the nominal DNN output in the control region for $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ is affected by reweighting of the kinematic input variables. The kinematic reweighting is performed by deriving scale factors by fitting 6th-order polynomials to the ratio between the sum of Monte Carlo $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ and the difference between data and Monte Carlo $t\bar{t}$ in our control region, and then extrapolated to all regions included in our analysis. This is repeated for all 14 input variables to our neural net.

Due to the high order of polynomial used in fitting the ratios, it is not possible to extract an uncertainty by changing the choice of fit function. Comparing to no reweighting

at all would be an overestimate of the reweighting uncertainty. Systematic variations on our Kinematic Reweighting are found by repeating the process using the highest and lowest theoretical variations on Monte Carlo $t\bar{t}$ production predictions rather than the nominal event weights. More specifically, modeling normalization variations - taken from the NNLO factorization and renormalization scale variations, PDF and α_s variations, and mass scale variations - for $t\bar{t}$ are subtracted from data, and the kinematic reweighting scale factors are re-derived to the new ratio where $t\bar{t}$ is either over-estimated or under-estimated. The variations on $t\bar{t}$ used here are of the order 5% variation overall. Figure 10.7 shows the high and low variations in Kinematic Reweighting compared to the nominal for the control and signal regions.

This kinematic reweighting systematic is applied only to signal regions. This choice is physically motivated because the scale factors are extracted in the control region and extrapolated to the signal region, making it reasonable that the extrapolation uncertainty applies only to the signal region. Figure 10.7 depicts the variation in kinematic reweighting for both the control and signal regions, but only the signal region variations are used as systematic variations in the fit.

Due to the robustness of the kinematic reweighting to these changes in $t\bar{t}$ contribution, the variations in kinematic reweighting are not a large fraction of the bin yields. Further validation of this kinematic reweighting uncertainty evaluation procedure is available in Appendix C. Figure 10.8 shows the variations as input into the fit for the signal region. It is a bit clearer to see the shape variation in comparison with the nominal templates.

Incorporated at this level, kinematic reweighting, though ranking highly as a single nuisance parameter, has a relatively small impact on signal strength uncertainty compared to the grouped impact of other systematics. The kinematic reweighting accounts

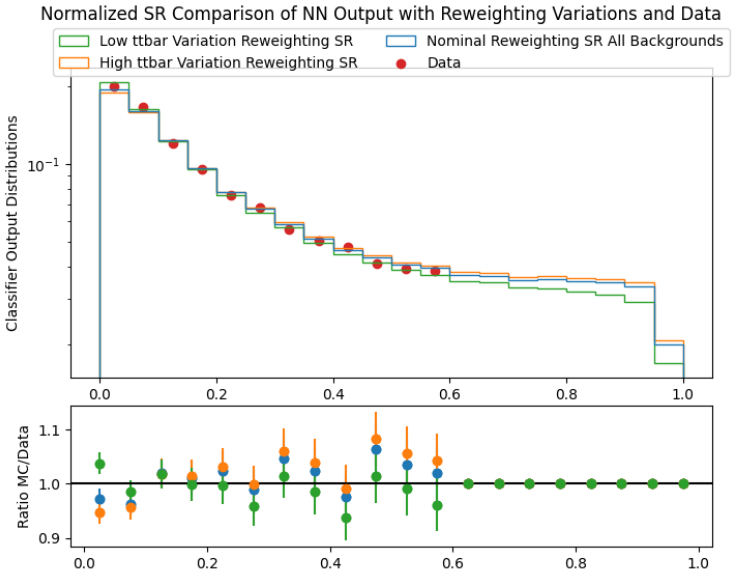


Figure 10.7.: The systematic variations capturing the kinematic reweighting extrapolation uncertainty are shown for the QCD $b\bar{b}\gamma jj$+$c\bar{c}\gamma jj$ plus all unreweighted MC backgrounds in normalized comparison with data where there is little signal contamination. The nominal reweighting (blue) is that which is used in the template for the fit. The two systematic variations are derived from repeating the kinematic reweighting strategy with an overestimated (orange) and under-estimated (green) $t\bar{t}jj$. It can be seen that the magnitude of variations sandwiches any lingering discrepancies between MC and data.

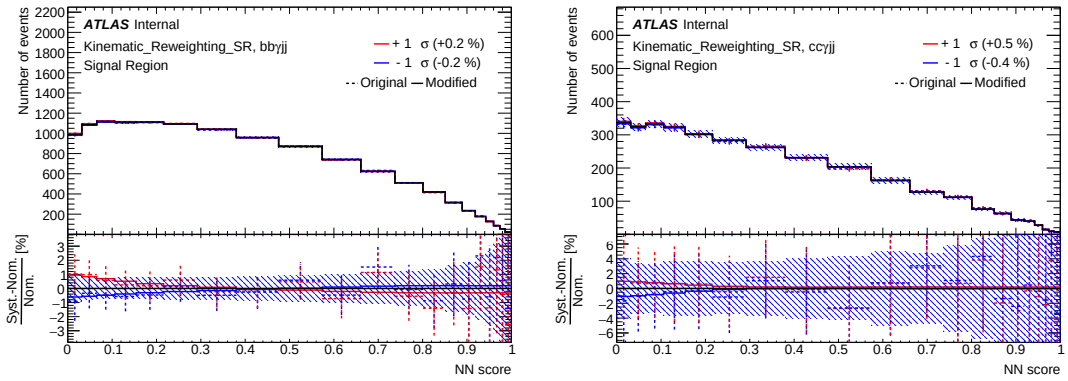


Figure 10.8.: Systematic variations for kinematic reweighting as used in the fit in the signal region only. Here we see the variations for $b\bar{b}\gamma jj$ (left) and $c\bar{c}\gamma jj$ (right) separately. Solid blue and red variations are after the smoothing of the raw dashed blue and red histograms to get rid of shared statistical fluctuations. Hashed blue region represents statistical uncertainties.

for roughly $\sigma(\mu_H) = 0.05$.

§ 10.3. Experimental Uncertainties

10.3.1. Luminosity Uncertainty

The measurement of the amount of data collected by the experiment, or the integrated luminosity, is extremely important in giving context to the event yields measured in analysis. The key to accessing the physical constants of the universe is in the cross-sections of various phenomena, not in raw counts. During Run 2, the luminosity measurements for ATLAS were primarily performed by LUCID, as mentioned in Section 3.2.2. LUCID's measurement of course comes with some uncertainties on response and calibration. That being said, careful calibrations performed in conjunction between ATLAS and LHC operations and clever processing techniques have allowed ATLAS to report

the most precise luminosity measurements to date at the LHC [162]. ATLAS's recorded Run 2 integrated luminosity and its associated uncertainty was updated in 2022 to be 140.1 ± 1.2 fb^{-1} [163]. This gives an uncertainty of just 0.83% on the cross-section measurement of a given process. All MC templates therefore receive a 0.83% normalization uncertainty including signal which directly translates to a 0.8% impact on signal strength uncertainty. This is a minor source of uncertainty for this analysis.

10.3.2. Photon and Electromagnetic Object Related Systematics

Compared to the precision of jet reconstruction and measurement, the uncertainty on the reconstruction of electromagnetic objects is quite small. With the only EM object of interest in this analysis being a high-energy photon, the signal measurement is not particularly sensitive to EM experimental uncertainties. Nevertheless, three NPs are dedicated to representing uncertainties related to the reconstruction of electromagnetic objects. Two pseudorapidity-dependent correction factors, calibrated using $Z \rightarrow ee$ simulation and data, for the energy scale and resolution of reconstructed EM objects are applied to the MC to better simulate the detector responses seen in data [126]. Two of the $e\gamma$ NPs are dedicated to systematic variations that represent how these uncertainties on the scale factors for electron and photon energy scale and resolution propagate to our final fit observable, the NN score. The last EM-related NP is dedicated to the uncertainty on photon identification efficiency, keeping in consideration that photon identification increases in efficiency and efficiency precision with E_T [164]. To account for the scenario of an electron being mistakenly identified as a photon, another η-dependent correction factor is applied, similarly calibrated using Z-decay data. The uncertainty on these correction factors are considered in a dedicated photon identification efficiency nuisance

parameter. Jets faking photons are not granted a dedicated systematic NP, and instead, this scenario is accounted for with the kinematic reweighting of the backgrounds as described in Section 9.4. Altogether, these parameters have a combined raw uncertainty of $\sigma(\mu_H)$=0.013 on the signal strength measurement, which is not a dominant experimental systematic uncertainty in comparison to the jet-related uncertainties described in Section 10.3.3.

10.3.3. Jet Related Systematics

The dominant experimental systematics impacting the measurement in this analysis are related to jets. To this end, over 54 systematic NPs are devoted to capturing a basis of the ways uncertainties in detector response and jet reconstruction can manifest in the signal extraction. The jet reconstruction described in Section 7.2.1 can err in a wide variety of ways, from differences in passive material thicknesses to those in simulation to fluctuations specific to an event like pile-up interference or a greater than average number of neutral hadrons in the jet shower. The possible variance of such departures from the perfect jet-measuring scenario result in uncertainties on the recorded jet energy scale, energy resolution, and tagged flavor [165]. The uncertainties on jet energy scale and resolution are detailed in Section 10.3.4, and details on flavor tagging systematics are provided in Section 10.3.5.

10.3.4. Jet Energy Scale and Resolution

The uncertainties on jet energy scale and resolution are implemented using ATLAS tools and are applied to all MC templates, backgrounds and signal alike [165]. There are 30

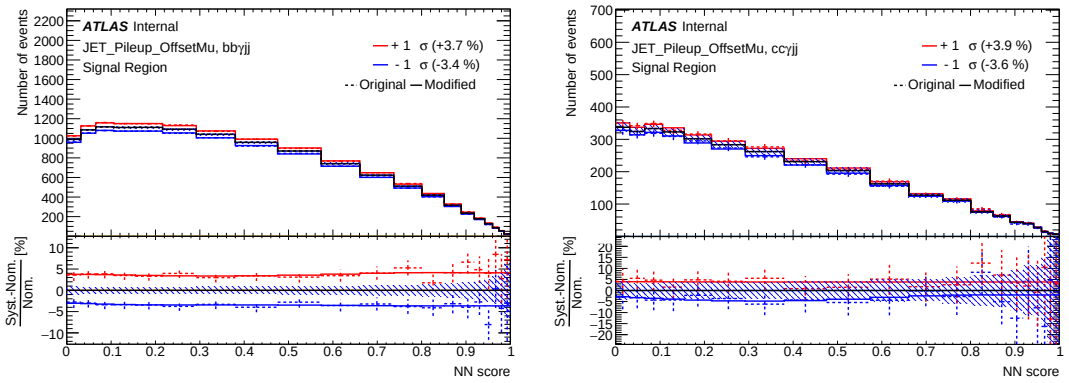
NPs (CategoryReduction scheme) devoted to jet energy scale (JES) uncertainties and the propagation of their effect into our fit observable, and these parameters are completely correlated across all input samples. The derivation of these systematics came in a multi-step approach. The forward jet scale is calibrated relative to the scale of central jets through η-intercalibration performed using dijet events. Z+jet and γ+jet samples are used to calibrate the jet scales for balance relative to the Z and to photons. Multi-jet samples help improve the high p_{T} regime by calibrating the balance of a jet relative to a recoiled system of jets. JES uncertainties arising from pile-up, flavor, and punch-through (a jet not fully contained by the calorimeters) are also all considered [166].

Jet energy resolution (JER) systematics receive 13 NPs in the ‘‘FullJER” scheme, which is more parameters than were implemented in the previous Run 2 analysis. JER uncertainties are derived by performing a Gaussian smearing of jets across various axes. Both MC and pseudodata are smeared by 1σ variations, and the fit JER NPs are constructed from their difference and symmetrized [167].

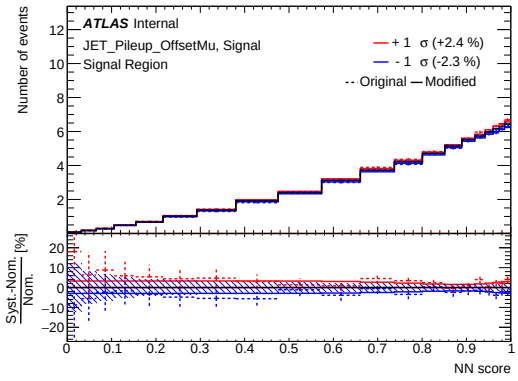
Of the experimental uncertainties, JES and JER systematics dominate with a contribution $\sigma(\mu_H) = 0.13$. One of the dominant JES uncertainties in its impact on the signal strength measurement comes from a pile-up related parameter, ‘‘JET_Pileup_OffsetMu,” and the systematics variations for this NP can be seen in Figure 10.9 for $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ backgrounds and signal in the signal region. One of the dominant JER uncertainties is shown in Figure 10.10 for $b\bar{b}\gamma jj$ and signal in the signal region. There are many such JES/JER NPs that share a similar magnitude of variation to those shown in Figures 10.9 and 10.10, so the total impact from these jet-related systematics is truly a cumulative effect and not a single responsible parameter.

Source of Absolute Uncertainty	$+\sigma(\mu_H)$	$-\sigma(\mu_H)$
Statistical		
Data Statistical	+0.58	-0.57
Gammas (MC Statistical)	+0.25	-0.22
Theory and Modeling		
Kinematic Reweighting	+0.001	-0.002
MC Modeling	+0.19	-0.15
Normalization Factors	+0.13	-0.08
Experimental		
Luminosity	+0.004	-0.006
Jet + Egamma + Photon	+0.15	-0.11
b-tagging	+0.03	-0.03
Total	$+0.69$	-0.64
Total MC Statistical	$+0.25$	-0.22
Total Theoretical	$+0.23$	-0.17
Total Experimental	$+0.16$	-0.12

Table 10.1.: The observed impacts of the statistical, theoretical, and experimental systematics on the signal strength uncertainty, as extracted from the final fit to observed data. Data statistical uncertainty is still the single largest contribution to the uncertainty.

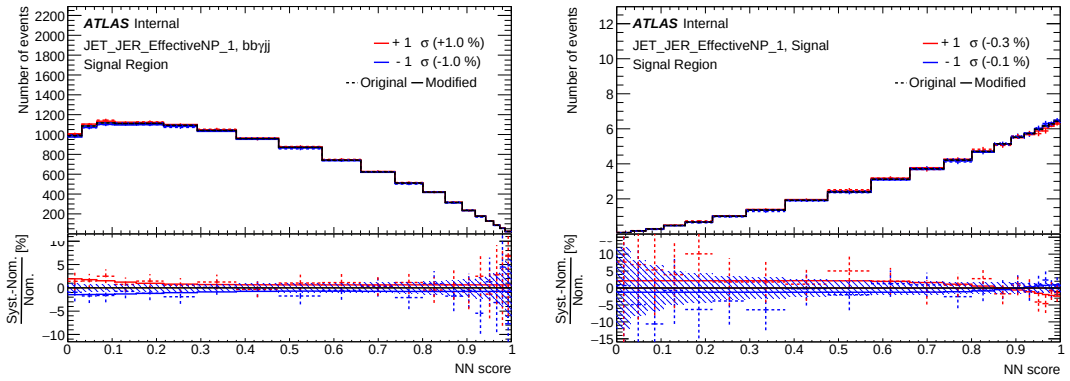


(a) $b\bar{b}\gamma jj$ (b) $c\bar{c}\gamma jj$



(c) Signal $H\gamma jj$

Figure 10.9.: Systematic variations for the JET_Pileup_OffsetMu nuisance parameter, one of the dominant experimental JES uncertainties. Variations are shown for the two dominant backgrounds, $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$, and for signal in the signal region. The black histogram represents the nominal MC template, and the solid red and blue histograms represent the up and down 1σ variations. The dashed red and blue lines visible in the ratio plot represent the raw variations before smoothing and symmetrization. The hashed blue area on the ratio plots represents statistical uncertainties only.



(a) $b\bar{b}\gamma jj$ (b) Signal $H\gamma jj$

Figure 10.10.: Systematic variations for the JET_JER_EffectiveNP_1 nuisance parameter, one of the dominant experimental JER uncertainties. Variations are shown for the dominant background $b\bar{b}\gamma jj$ and for signal in the signal region. The black histogram represents the nominal MC template, and the solid red and blue histograms represent the up and down 1σ variations. The dashed red and blue lines visible in the ratio plot represent the raw variations before smoothing and symmetrization. The hashed blue area on the ratio plots represents statistical uncertainties only.

10.3.5. Flavor Tagging

Uncertainties arising from the misidentification of the b-jet's flavor are an important consideration. There are 11 nuisance parameters devoted to this end, split across the contributions from each flavor: b-flavor with 2 NPs, c-flavor with 3, and light-flavor with 4, followed by two NPs derived from the extrapolation from charm jets. Scale factors were applied to the MC to calibrate the b-tagging efficiency relative to what is seen in data, and these systematics NPs capture the uncertainty on those scale factors from the different flavor contributions. For the dominant flavor components, B0 and B1, the fit uncertainties take into account matching between online and offline systematic variations, performed matching jet by jet in the events. The other flavor components include only offline scale factor variations [168].

Overall, the flavor tagging efficiency uncertainties are quite flat with NN score, as can be seen for two of the 11 components in Figure 10.11. Since the normalization of the dominant backgrounds is floating in the fit, the flavor-tagging systematics are not a dominant source of signal strength uncertainty. Collectively, these components contribute $\sigma(\mu_H) = 0.03$ to the signal measurement.

Unique to this analysis, b-tagging requirements are set at both the trigger and the offline selection levels, and different tagging algorithms are applied in each of the two phases. For this reason, care had to be taken to ensure that the jet matching and derived uncertainties seemed reasonable. In particular, since the dominant background MC templates used are truth-tagged, further checks were performed to ensure that the systematics are well behaved. As demonstrated in Figure 10.12, the flavor tagging systematics for the dominant flavor components behave practically identically between the truth-tagged and direct-tagged background samples, raising no red flags.

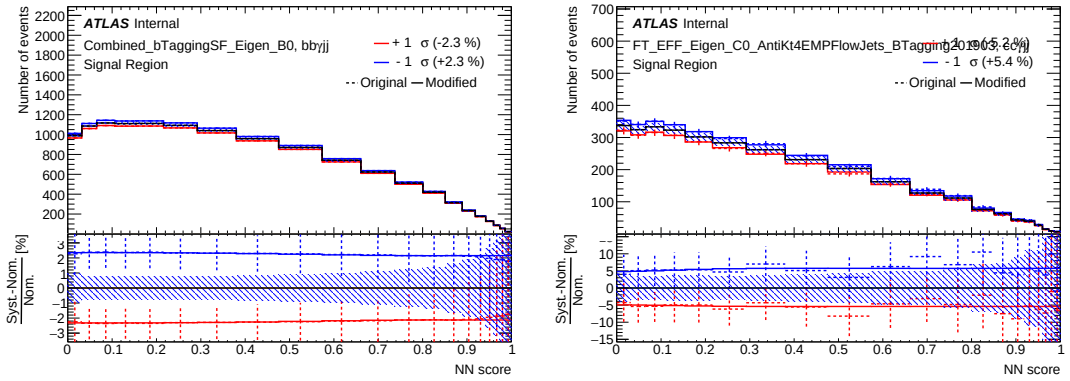


Figure 10.11.: b-tagging systematics in the signal region for two flavor components. On the left, the first b-flavor component is shown with its impacts on $b\bar{b}\gamma jj$, and on the right, the first c-flavor component is shown with its impacts on $c\bar{c}\gamma jj$.

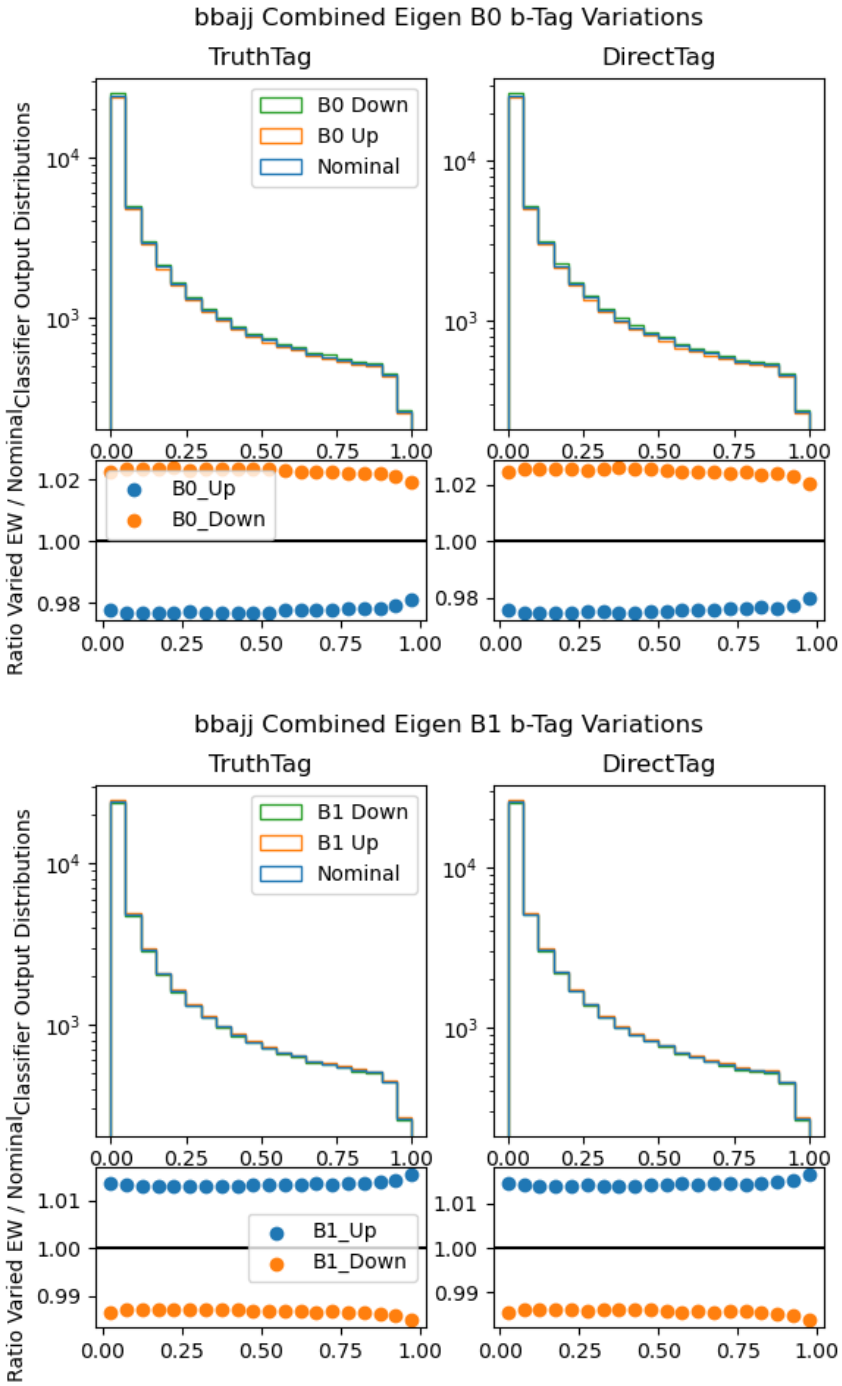


Figure 10.12.: Comparison of behavior of $b\bar{b}\gamma jj$ flavor tagging systematics for truth-tagged and direct-tagged background samples in the dominant flavor components, B0 (left) and B1 (right).

Chapter 11.

Results

To extract the signal strength, a binned likelihood fit is performed directly on the neural net classifier output distribution using TRExFitter [169, 170], simultaneously fitting control and signal regions which are separated by the invariant mass of the two leading b-tagged jets [171]. Floating normalization factors for signal and the combination of the two dominant non-resonant backgrounds are extracted by a simultaneous fit of the control (m_{bb} sidebands) and signal (Higgs mass window) regions, which are defined more explicitly in Chapter 9. The likelihood function gives a measure of the goodness of fit of MC to data with included penalization for strong pulls of any nuisance parameters away from their nominal values. These systematics nuisance parameters, described in Chapter 10, provide parametrizations for modifying the MC templates, and the likelihood function is optimized until their best-fit values are found.

Both control and signal regions consist of 20 NN score bins, the widths of which are set independently for each region using TRExFitter's transformation type ''D'' algorithm, which maximizes signal sensitivity while maintaining total statistical uncertainties below a certain threshold. The result of this autobinning algorithm produces

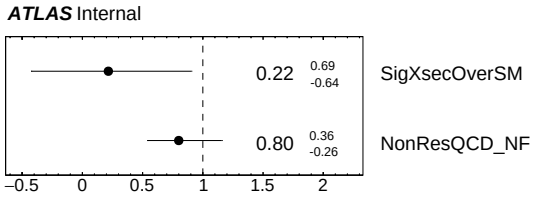


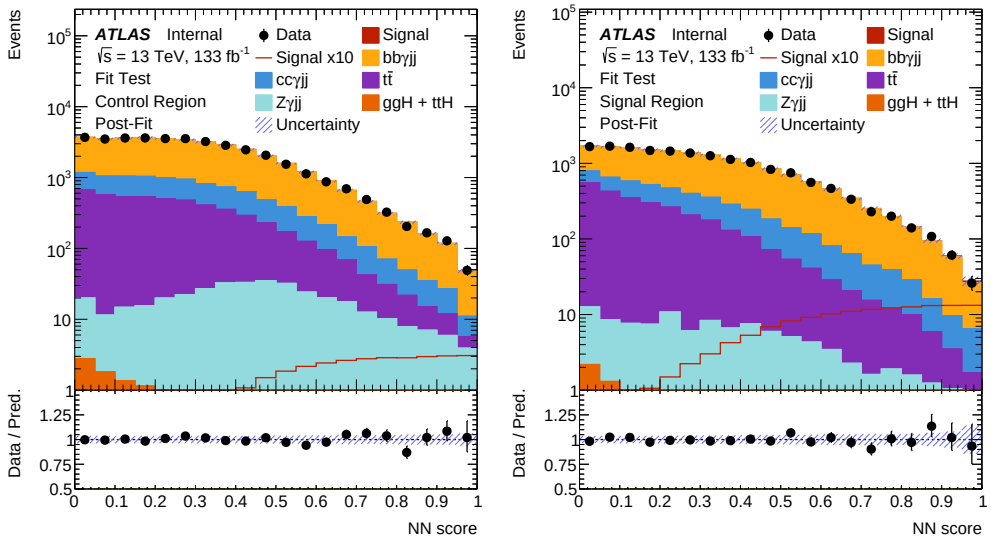
Figure 11.1.: Extracted norm factors for signal and for the two dominant $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$ backgrounds from the final fit to data.

very fine binning in the high score range of the signal region where there is the most signal purity, as well as fine binning in the low score range of the control region, where background statistics are quite high. The highest score bins of the signal region hold an overwhelming majority of the signal sensitivity and purity in our fit, with more than 60% of estimated signal lying in the signal region with an NN score above 0.6. Observed NN score distributions* for the control and signal regions are shown in Figure 11.2. Expected and observed yields are provided numerically for each sample in Table 11.1.

The Higgs boson signal strength and its uncertainty is defined relative to its standard model prediction. The observed signal strength is reflected in Figure 11.1, along with the final fit normalization for the dominant backgrounds, which is about 80% of the leading-order MC estimation. An inclusive signal strength of $\mu_H = 0.22 \pm 0.65$ was observed, compared to an expected value of 1.0 ± 0.7. This corresponds with an observed signal significance of 0.33 standard deviations, compared with an expected significance of 1.48 standard deviations.

Checks were performed to assess any possible bias in this signal measurement, and negligible bias was observed. Details are provided in Appendix E. The measured signal

*Note that the NN score has been renormalized in the post-fit distributions to give regular bin widths. For this reason, the binning does not reflect the raw score binning resultant from the auto-binning algorithm, which is more fine in high score signal region bins and in low score control region bins.



(a) CR (b) SR

Figure 11.2.: Post-fit plots to the observed dataset, showing control region (left) and signal region (right). Background and signal MC templates are shown with all post-fit pulls to nuisance parameters and normalization factors incorporated. Data bin yields are represented with black markers. The red line histogram represents the post-fit signal bin yields multiplied by a scale factor of 10, while the stacked red signal histogram is normalized the fitted signal strength of $\mu = 0.22$. The ratio plots (bottom) display the bin-by-bin quotient between the data and the post-fit total signal and background yields. Post-fit uncertainties are depicted with a blue hashed band in both the yield and ratio plots.

Yields	CR		SR Bins 0-16		SR Bin 17		SR Bin 18		SR Bin 19	
	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre
$b\bar{b}\gamma jj$	27400	30600	11100	12400	78	86	49	55	20	23
$c\bar{c}\gamma jj$	5100	7400	2200	3200	10	15	6	9	5	7
$t\bar{t}$	4800	4600	2700	2600	5	4	2	2	1	1
$Z\gamma jj$	400	400	100	100	1	1	1	1	0.4	0.4
$ggH+t\bar{t}H$	10	10	8	8	<0.1	<0.1	<0.1	<0.1	<0.1	<0.1
Total Bkg	37700 ± 1400		16100 ± 600		94 ± 6		59 ± 5		27 ± 4	
Signal	3	14	9	41	1	6	1	6	1	6
Data	37701		16070		108		61		26	

Table 11.1.: Post-fit and pre-fit MC yields and data yields shown in bulk in the background dominated control region and low score signal region bins and broken down in the last three signal region bins, where signal purity is the highest. The MC backgound total and its uncertainty are given post-fit.

strength is not in tension with either the Standard Model or the previous search for this signature in this dataset. A description of the compatibility with the previous result is available in Appendix G.

§ 11.1. Systematic Nuisance Parameter Pulls and Rankings

The 78 nuisance parameters that give freedom to the fit can be seen with their final pulls and constraints in Figure 11.3. The pulls are mostly left close to their nominal values, and no NP has a pull larger than a standard deviation. This suggests that, in general, the nominal templates represent the data well, and there is no egregious mismodeling.

The significance of each nuisance parameter in terms of its impact to the signal strength uncertainty is ranked in Figure 11.4 for the top 25 NPs. Data statistics dominate, followed by the normalization of the largest backgrounds. Then, some of the modeling uncertainties, which give significant shape variations to the NN output distribution, rank highly. One possible reason the $t\bar{t}$ modeling systematics rank so highly is that they are larger for the signal region than the control region, though the control region offers the most constraining power. A few of the jet-related systematics NPs from Section 10.3.3 reach highly on the list.

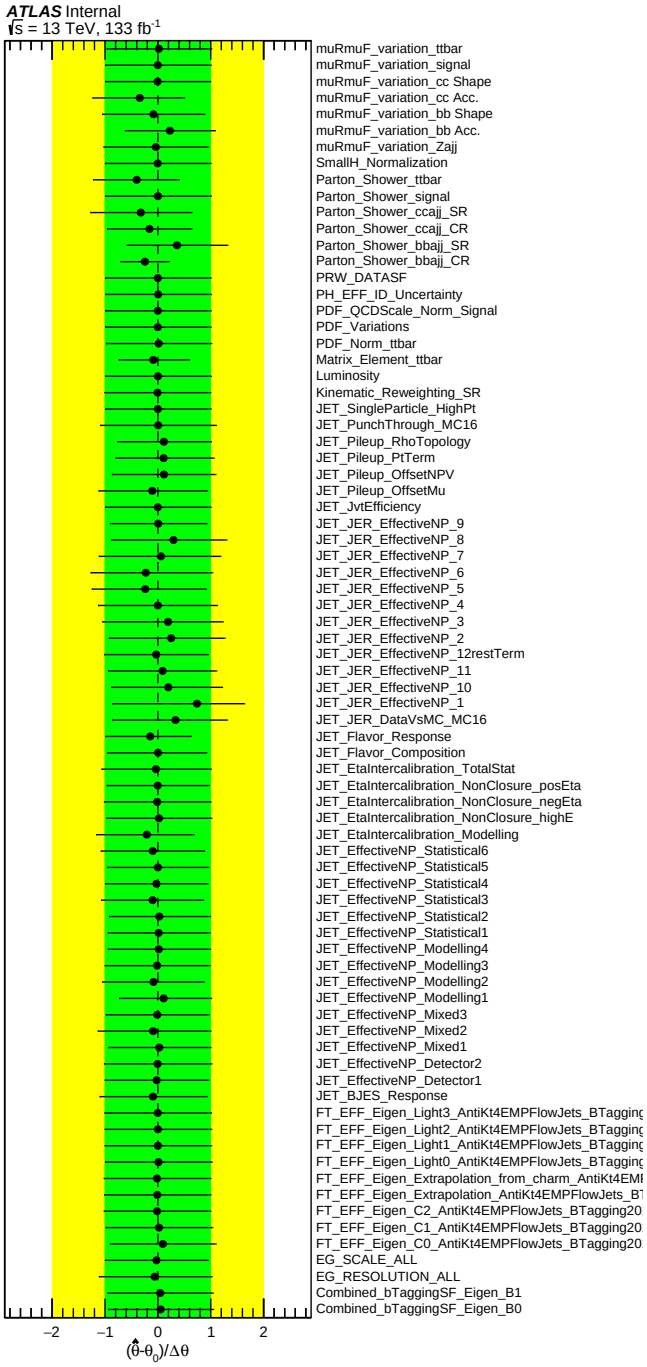


Figure 11.3.: Post-fit to data observed pulls and uncertainties of systematics nuisance parameters.

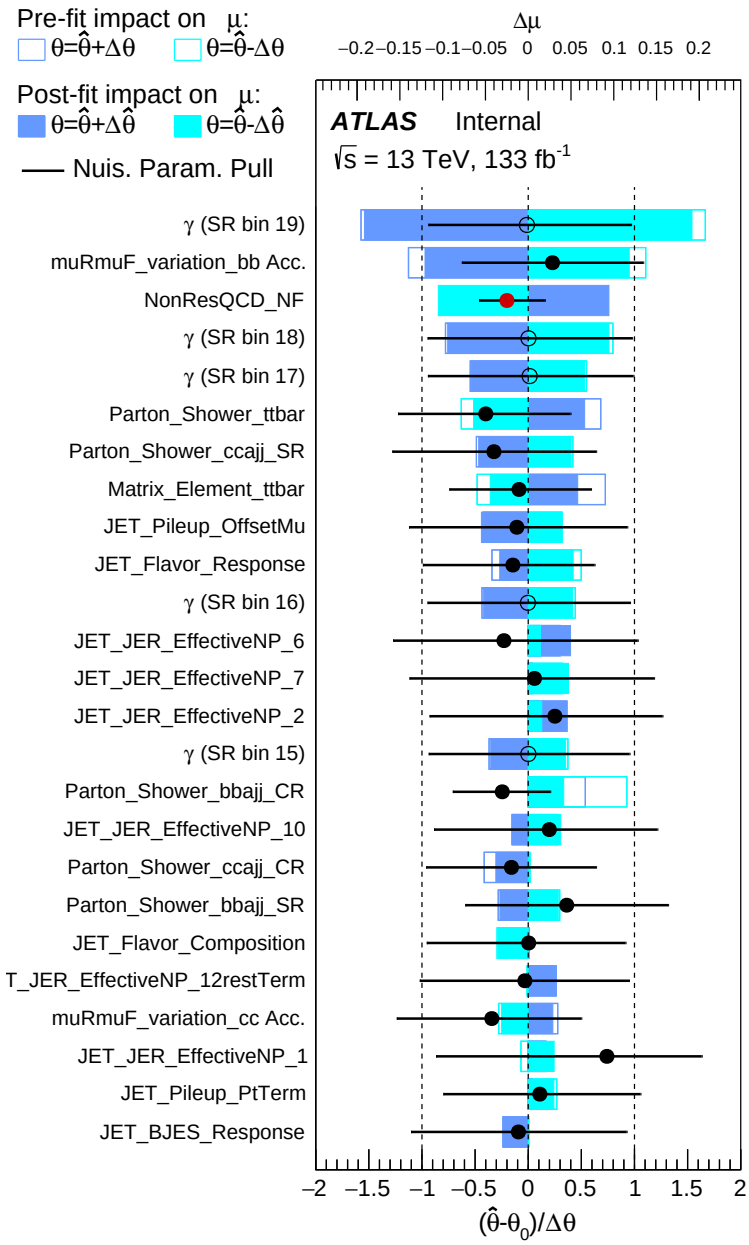
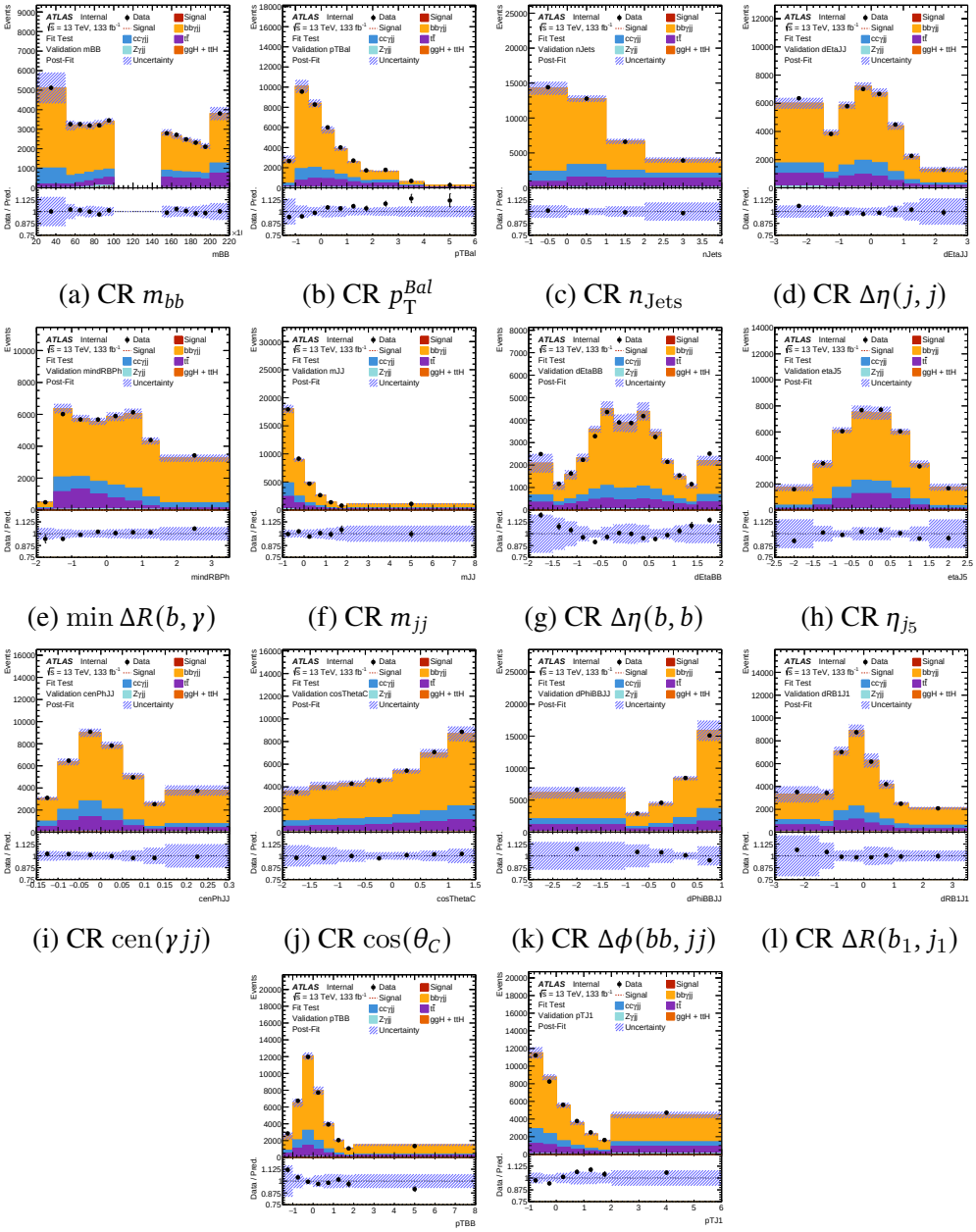


Figure 11.4.: Ranking of the top 25 systematic nuisance parameters in order of impact on signal strength uncertainty. Pre-/post-fit pulls and uncertainty impacts are shown for each NP.

§ 11.2. Post-Fit MC-Data Closure of Kinematic Distributions

Important to demonstrating sufficient MC modeling closure with data within uncertainties, post-fit distributions have been created for all the input kinematic variables. Figures 11.5 and 11.6 show the kinematic distributions, in the control region and signal region respectively, of MC and data after propagating all post-fit pulls and norm factors. Post-fit uncertainties overlay the ratio plots as blue hashed bands. Essentially all post-fit Data/MC ratios lie within 1 standard deviation of the post-fit uncertainties, suggesting that any differences are well enclosed within the included systematics. It is very much non-trivial that these plots would depict such good agreement, considering the fit is performed using only the NN output score as an observable. The excellent closure strongly supports the signal extraction model. Further checks were performed on specific kinematic input variables to ensure this good modeling persists in the signal sensitive bins at high NN score in the signal region. Post-fit kinematic comparisons between MC and data in NN score slices in the signal region can be observed in Appendix F.



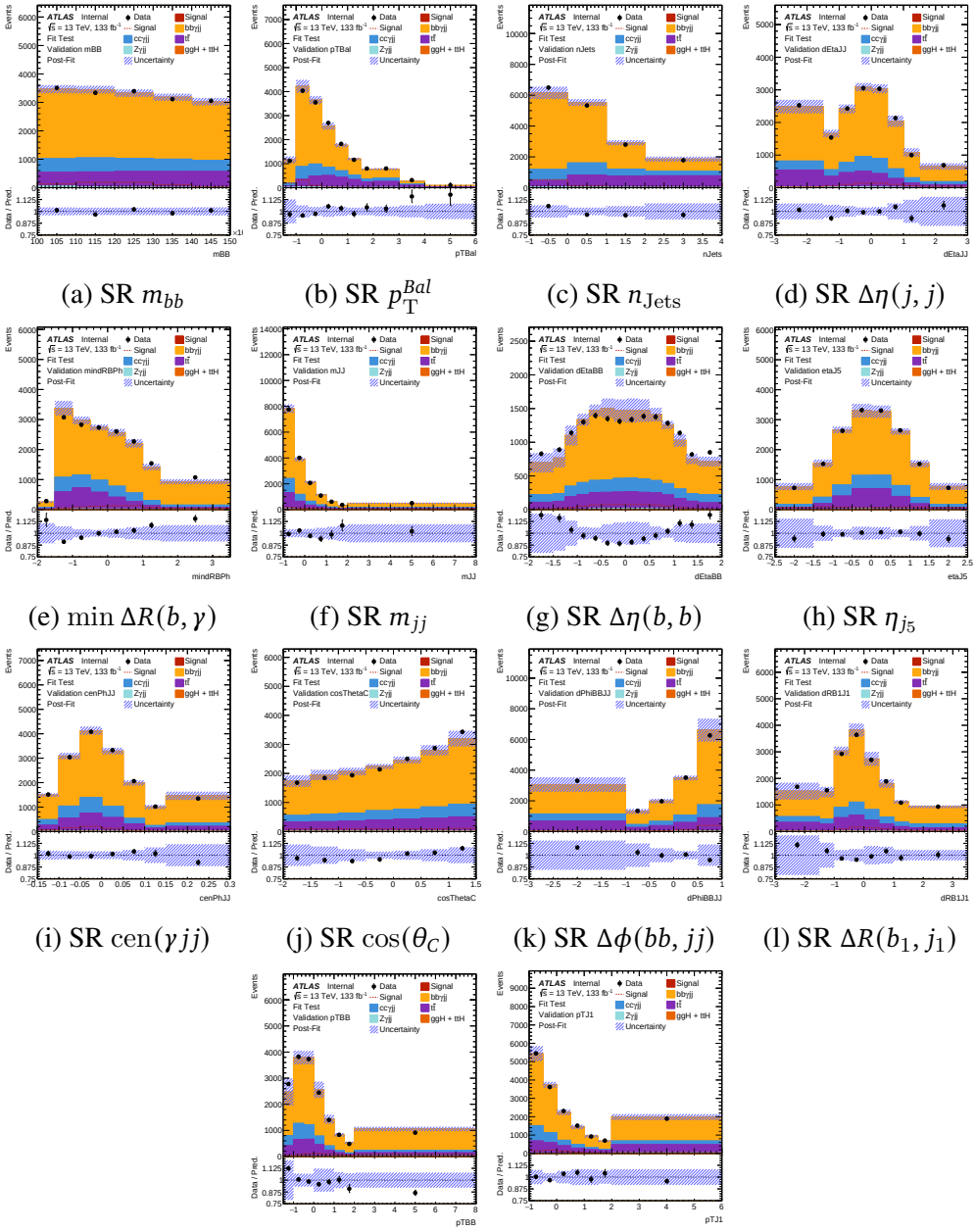
(a) CR m_{bb} (b) CR p_{T}^{Bal} (c) CR n_{Jets} (d) CR $\Delta\eta(j,j)$

(e) min $\Delta R(b,\gamma)$ (f) CR m_{jj} (g) CR $\Delta\eta(b,b)$ (h) CR η_{j_5}

(i) CR $\mathrm{cen}(\gamma jj)$ (j) CR $\cos(\theta_C)$ (k) CR $\Delta\phi(bb,jj)$ (l) CR $\Delta R(b_1,j_1)$

(m) CR p_{T}^{bb} (n) CR $p_{\mathrm{T}}^{j_1}$

Figure 11.5.: Control region post-fit kinematic distributions with the propagation of all norm factors and systematics NP pulls resultant from the convergence of the NN score fit of MC to data. The dashed red signal line is the final signal yield inflated by a factor of 10. The variable x-axes shown are after the Gaussian regularization transformation performed before input into the NN (except for m_{bb}, which is shown with the untransformed m_{bb} axis). MC and data agreement is enclosed within post-fit uncertainties.



(a) SR m_{bb} (b) SR p_{T}^{Bal} (c) SR n_{Jets} (d) SR $\Delta\eta(j,j)$

(e) min $\Delta R(b,\gamma)$ (f) SR m_{jj} (g) SR $\Delta\eta(b,b)$ (h) SR η_{j_5}

(i) SR $\mathrm{cen}(\gamma jj)$ (j) SR $\cos(\theta_C)$ (k) SR $\Delta\phi(bb,jj)$ (l) SR $\Delta R(b_1,j_1)$

(m) SR p_{T}^{bb} (n) SR $p_{\mathrm{T}}^{j_1}$

Figure 11.6.: Signal region post-fit kinematic distributions with the propagation of all norm factors and systematics NP pulls resultant from the convergence of the NN score fit of MC to data. The dashed red signal line is the final signal yield inflated by a factor of 10. The variable x-axes shown are after the Gaussian regularization transformation performed before input into the NN (except for m_{bb}, which is shown with the untransformed m_{bb} axis). MC and data agreement is enclosed within post-fit uncertainties.

Chapter 12.

Summary

This dissertation presents a search for predominantly vector-boson fusion Higgs boson production in association with a high-energy photon in the $b\bar{b}$ Higgs decay channel. Proton-proton collisions at $\sqrt{s}$ = 13 TeV produced by the LHC at CERN are investigated. The volume of data probed is an integrated luminosity of $133\,\text{fb}^{-1}$ collected by the ATLAS detector at CERN. It is particularly powerful to study the $H(\rightarrow b\bar{b})\gamma$ final state because the photon requirement greatly reduces the multijet background. Advancements in this analysis over previous efforts in this channel include updated and extended Monte Carlo background datasets, finely-optimized multi-variate analysis techniques for signal-background separation, and a novel approach to fitting data for signal extraction. With these developments, an expected signal significance of 1.5σ was achieved, marking an improvement of nearly 50% beyond the expected significance of the previous full Run-2 analysis. An observed Higgs boson signal strength of 0.22 ± 0.65 was measured, with an observed significance of 0.3σ.

My most noteworthy contributions to this analysis include: the construction, optimization, and training of the DNN classifier and alternative approaches (Chapter 8);

background modeling checks and kinematic reweighting (Chapter 9); evaluation of systematic uncertainties and implementation of fit nuisance parameters (Chapter 10); and the performance of the final fit and all fit checks (Chapter 11).

Appendix

A. Standard Model Calculations

§ A.1. Probabilities and Matrix Element Calculations

Matrix Element Calculation

The probability of a final state resulting from an initial state is given by the square of a matrix element, where the matrix corresponds to the transformation that would transition such an initial to such a final. The full matrix element contains infinitely many Feynman diagrams. Because of the dependency of the particle coupling strengths on the interaction energy, a scale is selected for the calculations. For the purposes of the enclosed analysis, the renormalization and factorization scales, μ_R and μ_F, for the evaluation of the coupling strengths are set to the Higgs mass, 125 GeV.

A.1.1. Combining Feynman Diagrams

It has been mentioned that every possible way in which an observed initial and final state measurement can occur contributes to its probability. It would be nice and convenient if every possible Feynman diagram contributed a positive small piece to the overall transformation probability, but the universe is not so simple. Instead, some diagrams interfere and contribute to the matrix element with opposite sign. An example of this

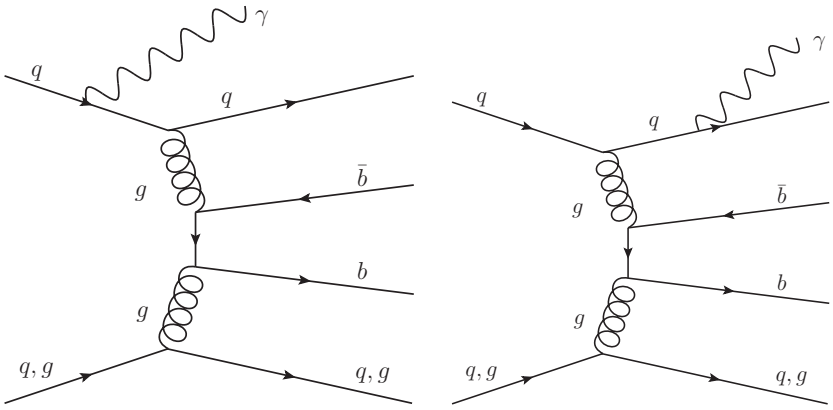


Figure A.1.: Feynman diagrams for non-resonant $b\bar{b}\gamma jj$ QCD background. Destructive interference occurs between these two diagrams, and they are combined with opposite sign.

is in the interference of a photon radiating from an initial state leg and a final state leg in the dominant background of this analysis, non-resonant QCD $b\bar{b}\gamma jj$. These diagrams are depicted in Figure A.1. When combining the probabilities of these two diagrams, they are summed with opposite sign, resulting in destructive interference and lessening the cross-section of this dominant background.

A.1.2. Cross-Section Calculation

The cross-section of a proton collision that gives the interaction of two initial partons i and j and results in a hadron from a final-state parton k with some other species X not specified can be calculated as

$$\frac{d\sigma}{dp^3} = \sum_{i,j,k} \int dx_i f_i(x_i) \int dx_j f_j(x_j) \int dz_k F_k(z_k) |\mathcal{M}(ij \to kX)|^2,$$

where $f_i(x_i)$ gives the probability of finding parton i inside the proton with the momentum fraction x_i provided by the parton distribution function (PDF). The matrix element,

$\mathcal{M}(ij \rightarrow kX)$, is a parton-level calculation that evaluates the likelihood of the initial state partons i and j interacting to produce a final state parton k and some other species. The probability of that final state parton k to hadronize into a particular hadron with momentum fraction z_k is given by the parton fragmentation function, $F_k(z_k)$.

This probability calculation diverges in two cases: (1) contributions of virtual particles in the loops of Feynman diagrams can have larger and larger momentum resulting in an ultraviolet (UV) divergence, and (2) contributions from very low momentum (soft) or low relative angle (collinear) radiative legs can be greater and greater in number resulting in an infrared (IR) divergence. These divergences can rectified such that finite probability amplitudes can be calculated through the processes of regularization and renormalization. In the case of UV divergences, the calculation can be regularized by introducing a cutoff or renormalization scale, μ_R, where contributions from particles with small momentum compared to μ_R behave as they typically would and contributions from high-momentum particles on the scale of μ_R or higher are suppressed by energy-dependent modifications to the interaction couplings. The soft IR divergences cancel out when including perturbations of increasingly higher order, but to handle the collinear IR divergences, a second scale must be introduced, a collinear cutoff known as the factorization scale, μ_F. With these renormalization and factorization scales introduced, the cross-section becomes a finite and calculable quantity, but it is important to note that the couplings and parton distribution and fragmentation functions become dependent on the choice of these scales.

B. Recent Higgs Results from ATLAS

When participating in a large collaborative science initiative like the ATLAS experiment, it is incredibly apparent that the work of one individual exists within the context of and built upon the countless labors and triumphs of fellow scientists. In particular, many other Higgs searches and analyses came to fruition on a similar timeline to this thesis work. Amongst these endeavors were other Higgs searches targeting different production and decay channels, as well as analyses seeking to constrain Higgs properties such as its resonant mass width or its charge conjugation parity (CP) symmetry. I was lucky enough to be invited to the LHC Days 2024 conference to represent ATLAS and speak on some of these amazing results. In this chapter, the work in this thesis is contextualized by the other recent published Higgs boson studies by members of the experiment collaboration.

§ B.1. Recent Higgs Search Results

Search for the associated production of charm quarks and a Higgs boson decaying into a photon pair with the ATLAS detector The first search for inclusive Higgs production in association with a charm quark was published in July 2024, looking in the di-photon decay channel using the full Run 2 dataset. The standard model prediction

for the cross-section of $H + c$ production is $2.9\,\mathrm{pb}$, and a cross-section of $5.2 \pm 3.0\,\mathrm{pb}$ was measured for an observed (expected) significance of $1.7\sigma(1.0\sigma)$ consistent with the SM. This search benefited from an advanced background estimation approach, known as Gaussian Process Regression (GPR). GPR is a non-parametric approach that is not restricted to a particular function-form of the background being estimated in a data driven approach. Instead, the background form is fit with its parameters simultaneously from a probabilistic distribution of possible functions. This was applied for the total non-resonant di-photon background simultaneously in the lower and upper invariant mass sidebands.

Study of High-Transverse-Momentum Higgs Boson Production in Association with a Vector Boson in the $qqbb$ Final State In April 2024, the first study of Higgs production in association with a vector boson (VH) into a fully hadronic final state was published. Simultaneously, this study also marked the first look at high Higgs transverse momentum for the VH production channel. Unfortunately, the differential measurement as a function of p_T^H had little sensitivity from Run 2 statistics and will benefit from further data collection. The inclusive cross-section measured compared the standard model prediction gave a signal strength of $\mu = 1.4(+1.0/-0.9)$ for an observed (expected) significance of $1.7\sigma(1.2\sigma)$. The analysis employed two different taggers to identify the Higgs boson and vector boson decay candidates. Careful attention was also paid to the background modeling, where a transfer factor method was employed to capture the differences in background shape between the signal and control regions as a function of jet mass and p_T.

Measurement of the associated production of a top-antitop-quark pair and a Higgs boson decaying to a $b\bar{b}$ pair The most precise $t\bar{t}H$ measurement in a single decay channel thus far was brought to us July 2024, looking both at an inclusive cross-section measurement and a differential measurement as a function of Higgs transverse momentum. The search benefited from improved b-tagging and looser signal acceptance requirements to triple the signal acceptance of $t\bar{t}H$ events over previous efforts on this Run 2 dataset. It also employed advanced multiclass NN classifiers to define orthogonal signal-enriched signal regions and background-dominated control regions. The modeling of $t\bar{t}+jets$ was improved for this effort, such that the dominant systematic for this approach is not from background modeling but from uncertainty in the modeling of final state radiation in the $t\bar{t}H$ signal. With these improvements, the measured $t\bar{t}H$ cross-section was $\sigma_{ttH} = 411 \pm 54(stat.)^{+85}_{-75}(syst.)$ fb, corresponding to an observed (expected) significance of $4.6\sigma(5.4\sigma)$. The SM prediction for this cross-section is $507\,\text{fb}$.

Evidence for the VH, $H \to \tau\tau$ process with the ATLAS Run 2 dataset In December 2023, ATLAS achieved evidence of the Higgs boson produced in association with a vector boson in the tau lepton decay channel. The measured signal strength with breakdown of uncertainty was $\mu_H^{\tau\tau} = 1.28^{0.3}_{-0.28}(stat.)^{+0.25}_{-0.21}(syst.)$ from the combination of ZH and WH production channels, giving an observed significance of 4.2σ compared to 3.6σ expected. The ZH and WH signals were measured independently and were found to have compatible extracted signal strengths with a p-value of 56%. Beyond the increased dataset with about 7 times the statistics since the last measurement in Run 1, major improvements were achieved with a Neural Network classifier discriminator along with improvements in $\tau_{had-vis}$ reconstruction. The dominant systematic uncertainties still

impacting this channel come from the hadronic τ-decay and the simulated background sample size.

Integrated and differential cross-section measurements for VBF Higgs production in the $H \to WW^* \to e\nu\mu\nu$ final state In November 2023, a study measured the integrated and differential fiducial cross-sections of VBF Higgs production in the $WW^* \to e\nu\mu\nu$ final state using ATLAS Run 2 data. The differential cross-sections as functions of kinematic variables were used to constrain effective field theory coefficients, results of which are detailed in Section B.3. Comparing with an SM predicted cross-section of $2.07\,\mathrm{fb}$, the measured integrated cross-section $\sigma = 1.68 \pm 0.33(stat.) \pm 0.23(syst.)$ underestimates the SM expectation by roughly one standard deviation. This measurement was dominated by data statistical uncertainties, and limitations of fixed-order NLO simulations, which overestimated the data, were also noted.

Measurements of differential cross sections of Higgs boson production through gluon fusion in the $H \to WW^* \to e\nu\mu\nu$ final state Published in September 2023, measurements of differential cross sections as functions of many kinematic variables were performed for Higgs bosons produced via gluon fusion in the $H \to WW^* \to e\nu\mu\nu$ final state using Run 2 ATLAS data. In order to compare directly with theory, collected event objects are reconstructed to particle level using likelihood unfolding with Tikhonov regularisation. Differential measurements of m_T, y_{ll}, p_T^{ll}, p_T^H, y_{j_0}, $p_T^{l_0}$, $\Delta\phi_{ll}$, and $\cos\theta$ are all studied and found to be compatible with SM expectations. The systematic uncertainties of this analysis strategy are dominated by jet and muon reconstruction uncertainties, modeling uncertainties of the WW and $V\gamma$ backgrounds, and by the data-driven background estimates of mis-identified objects. The integrated fiducial

cross-section is measured to be $\sigma = 56.0^{+10.0}_{-9.5}$ fb.

Evidence for the Higgs boson decay to a Z boson and a Photon with ATLAS and CMS In a combined result between CMS and ATLAS in March 2024, the first evidence ($> 3\sigma$ result) of the Higgs decaying to a Z boson and a photon was recognized. This decay doesn't occur at tree-level of course, but the leading diagrams consist of a top-loop or a W boson loop to lead to the Z and photon final states. This result found a significance of 3.4σ above the background-only $\mu = 0$ hypothesis. The measured signal was also a bit above the SM prediction, with a signal strength of $\mu = 2.2 \pm 0.7$ times the the SM cross-section estimate. This combined the ATLAS signal strength measurement of $\mu = 2.0^{+1.0}_{-0.8}$ with the CMS signal strength measurement of $\mu = 2.4^{+1.0}_{-0.9}$. Together, the measured branching fraction to this rare decay channel is $(3.4 \pm 1.1) \times 10^{-3}$. The systematic uncertainties still dominant in this measurement are the uncertainty in the $Z\gamma$ branching fraction impacting the cross-section measurement, as well as uncertainties in the background modeling.

Measurement of the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4l$ cross-sections in a first look at Run 3 data The first Run 3 look at the $H \to \gamma\gamma$ and $H \to 4l$ cross-sections at the higher center of mass energy of $13.6\,\mathrm{TeV}$ was published in February 2024, looking at $31.4\,\mathrm{fb}^{-1}$ and $29.0\,\mathrm{fb}^{-1}$ respectively. In the diphoton decay channel, an unbinned analytic function fit is performed to the invariant mass, $m_{\gamma\gamma}$, spectrum. In the diphoton signal phase space, a fiducial cross-section of $\sigma_{fid} = 76^{+14}_{-13}$ fb was measured, which can be extrapolated to a total cross-section of $\sigma(pp \to H) = 67^{+12}_{-11}$ pb. In the quadruplet lepton channel, a binned fit to the invariant mass, m_{4l}, spectrum yielded a fiducial cross-section of $\sigma_{fid} = 2.8 \pm 0.74\,\mathrm{fb}$, which extrapolates to a total cross-section of $\sigma(pp \to$

$H) = 46 \pm 12\,\mathrm{pb}$. This fit used MC templates for the signal and nonresonant background and dedicated control regions to constrain the reducible backgrounds. Combined, these measurements mark a Higgs cross-section off $\sigma(pp \to H) = 58.2 \pm 8.7\,\mathrm{pb}$, which is consistent with the SM prediction of $\sigma(pp \to H)_{SM} = 59.9 \pm 2.6\,\mathrm{pb}$.

§ B.2. Recent Constraints on the Higgs Mass and Mass Width

The Higgs represents a very narrow particle resonance, and total width of the Higgs boson (Γ_H) is theoretically predicted to be just $4.1\,\mathrm{MeV}$ in the SM. This is less than the experimental resolution of the ATLAS detector in most scenarios, so probing this number takes quite some tricks. One such trick to probe the Higgs width is to check for evidence of Higgs decays that would not be allowed from an on-shell Higgs, hopefully finding close to none and thereby constraining the off-shell Higgs presence.

Evidence of off-shell Higgs boson production from ZZ leptonic decay channels and constraints on its total width Brought to us in December 2023, ATLAS published experimental evidence ($> 3\sigma$ measurement) of off-shell Higgs boson production, looking in the decay channels $ZZ^* \to 4l$ and $ZZ^* \to 2l2\nu$, where CMS already demonstrated similar precision. In the $ZZ^* \to 4l$ channel where mass can be more easily reconstructed without large losses, the quadruplet invariant mass was required to be above $220\,\mathrm{GeV}$ in the signal region, avoiding any on-shell Higgs domain, and $180 \leq m_{4l} \leq 220\,\mathrm{GeV}$ was reserved for control regions. The $ZZ^* \to 2l2\nu$ channel, characterized by high missing transverse energy and exactly two oppositely-charged leptons, benefits from a higher

branching fraction but also faces larger background contributions. Combining these two channels, a signal strength of $\mu_{off-shell} = 1.1^{+0.7}_{-0.6}$ was measured, which can then be combined with the on-shell $H \rightarrow ZZ^* \rightarrow 4l$ analysis to give a constraint on the Higgs total mass width, Γ_H. The ratio of the off-shell to on-shell signal strengths is equivalent to the ratio of Higgs total width to its SM prediction, $\frac{\mu_{off-shell}}{\mu_{on-shell}} = \frac{\Gamma_H}{\Gamma_H^{SM}}$. Using this assumption, the total width measured by this analysis was $\Gamma_H = 4.5^{+3.3}_{-2.5}$ MeV, giving the 95% confidence level upper bound of $\Gamma_H \leq 10.5(10.9)$ MeV. Besides statistical uncertainties, the dominant systematics impacting this analysis were parton shower uncertainties, NLO electroweak uncertainties, and jet-related systematics.

Constraint on the total width of the Higgs boson from Higgs boson and four-top-quark measurements in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ Brought to us in July 2024, this study performed the first Higgs width constraint using both on-shell and off-shell Higgs production. The achieved 95% confidence upper limit was observed (expected) to be $\Gamma_H \leq 450(75)$ MeV. This upper limit lowers to $\Gamma_H \leq 160(55)$ MeV, with the assumption that only SM particles enter in all involved loop processes.

Measurement of the Higgs boson mass with $H \rightarrow \gamma\gamma$ decays in ATLAS Run-2 data and in combination with Run-1 measurement Marking the most precise measurement of the Higgs mass at the time of its publication in December 2023, an analysis exploited the high resolution of the invariant mass of photon pairs reconstructed from the decays of Higgs bosons to achieve a mass measurement with a precision of 0.11%. Using only the Run-2 dataset, the Higgs mass measured by ATLAS was reported to be $m_H = 125.17 \pm 0.11(stat.) \pm 0.09(syst.)$. When combined with the Run 1 result through a simultaneous fit to both datasets, including 16 different signal categories, the central value

measured shifts by $50\,\mathrm{MeV}$ for a measurement of $m_H = 125.22 \pm 0.11(stat.) \pm 0.09(syst.)$ GeV. In addition to the larger dataset, the improvements in this measurement are made possible by an improvement in the photon energy scale uncertainties by nearly a factor of four. This was largely achieved by deriving E_T-dependent scale factors, where previously scale factors were flat across transverse energy and the E_T-dependence of photon energy calibrations was accepted as an uncertainty.

Combined measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4l$ decay channels using Run-1 and Run-2 data Succeeding the previous Higgs mass measurement by just a month and coming out in January 2024, a combination from the diphoton and $ZZ^* \to 4l$ decay channels using Run-1 and Run-2 ATLAS datasets quickly became the most precise Higgs mass measurement to date with a precision of 0.09%. This measurement yielded a Higgs mass of $m_H = 125.11 \pm 0.09(stat.) \pm 0.06(syst.)$ GeV. This analysis also benefited from the improvements in photon energy calibration, as well as 50% reduction in muon momentum uncertainty.

§ B.3. Recent Constraints on the CP Symmetry and other Quantum Properties of the Higgs Boson

The Standard Model predicts the spin-parity of the Higgs boson to be 0^+ with even C-parity. Sources of violation of combined charge conjugation and parity (CP) invariance have the potential to aid in the understanding of the matter-antimatter asymmetry or baryon asymmetry of the universe, which is quite an important open question. The only known source of CP violation in the SM to date is through quark mixing due to

a complex phase in the Cabibbo-Kobayashi-Maskawa (CKM) matrix, but this does not begin to come close to the magnitude of CP violation necessary to explain our universe's asymmetry.

Constraints on EFT coefficients provided by differential cross-section measurements of VBF Higgs production in the $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ final state In addition to its measurement of the integrated fiducial cross section, this November 2023 publication utilizes differential cross-section measurements as a function of $\Delta\phi_{jj}$ to place the strongest constraints to date on the Warsaw coefficients c_{HWB}, $c_{H\tilde{B}}$, and c_{Hq3}. Furthermore, careful attention as paid to the correlations between differential fiducial cross-sections, such that these results can be used in future global fits.

Study of the CP property of the Higgs boson to electroweak boson coupling in the VBF $H \rightarrow \gamma\gamma$ channel In this study published in December 2023, constraints are provided on the effective field theory coefficients $\tilde{d}$ in the HISZ basis and c_{HW} in the Warsaw basis, both describing potential CP-odd couplings between the Higgs and electroweak gauge bosons. This represents the first time ever that a 95% confidence interval is set for $\tilde{d}$. Though quickly outdone in a study described later in this subsection, this publication also marked the strongest constraint to date on the c_{HW} coefficient, twice as restrictive as the previous ATLAS or CMS $H \rightarrow 4l$ analyses. The EFT coefficient constraints were achieved through an analytic function fit to $m_{\gamma\gamma}$ in bins of optimal observables. No evidence of CP violation was detected, and the measurement's sensitivity was dominated by data statistical uncertainties.

Test of CP-Invariance of the Higgs boson in vector-boson fusion production and in its decay into four leptons Published in May 2024, a test of the CP-invariance of the Higgs boson was performed using ATLAS Run 2 data, looking at VBF Higgs production with its decay into four leptons ($H \rightarrow ZZ^* \rightarrow 4l(l = e, \mu)$). This study set constraints on possible BSM CP-odd Higgs couplings using the shapes of "optimal observables", which are CP-odd by construction. These optimal observables are constructed by normalizing the interference term between the SM and BSM interactions for three coefficients in the Warsaw basis: $c_{H\tilde{B}}$, $c_{H\tilde{W}B}$, and $c_{H\tilde{W}}$, as well as for the corresponding coefficients in the Higgs basis: $\tilde{c}_{ZZ}$, $\tilde{c}_{Z\gamma}$, and $\tilde{c}_{\gamma\gamma}$. Because these observables are CP-odd by construction, any observed asymmetry in their shape would be direct evidence for CP-invariance violation. This analysis is more sensitive than previous efforts due to its use of shape-only distributions, avoiding some of the uncertainty involved in measurements of cross-sections and yields. Tighter constraints on the CP-odd Wilson coefficients relevant to the HVV vertex were achieved compared to previous ATLAS $H \rightarrow ZZ^* \rightarrow 4l$ and VBF $H \rightarrow \gamma\gamma$ measurements, and all observations remained so far consistent with the Standard Model.

Probing the CP nature of the top-Higgs Yukawa coupling in $t\bar{t}H$ and tH events with $H \rightarrow b\bar{b}$ decays Uniquely utilizing both Higgs production in association with a top-quark pair and with a single top quark as signals, an effort was made to constrain the mixing angle, α, between CP-odd and CP-even components in the top-Higgs Yukawa coupling. The standard model prediction for this mixing angle is $\alpha = 0$, and the analysis of Run 2 data measured a best-fit value of $\alpha = 11°(+52°/-73°)$. This constraint was achieved by fitting CP-observables, as well as by measuring cross-section yields in

many segmented phase space regions. Lots of attention was paid to classification into these regions, and several MVA techniques were employed to handle event selection, as well as to assign jet reconstructions due to the high jet multiplicity of this signature. A limit was also made on the top-Higgs Yukawa coupling strength parameter, $\kappa_t = 0.84(+0.30/-0.46)$ compared with the SM expectation of $\kappa_t = 1$. While these constraints do not represent the most sensitive to date, they are orthogonal to and can be combined with the more sensitive constraints from the $H \rightarrow \gamma\gamma$ channel. Modeling uncertainties on the $t\bar{t}+ \geq 1b$ background dominate the uncertainties, and further sensitivity will be available with better background modeling and more statistics.

Differential cross-section measurements of Higgs boson production in the $H \rightarrow \tau^+\tau^-$ decay channel While simultaneously making the most precise VBF Higgs-production cross-section measurement to date with a measured VBF signal strength of $\mu_{VBF} = 0.93^{+0.17}_{-0.15}$, an effort published in July 2024 made differential cross-section measurements as a function of Higgs p_T to make the strongest constraint yet on the CP-odd Wilson coefficient, $c_{H\tilde{W}}$.

Determination of the relative sign of the Higgs boson couplings to W and Z bosons using WH production via Vector Boson Fusion Excitingly appearing in February 2024, a clever analysis approach to a search for WH production via vector boson fusion allowed for the exclusion of opposite sign couplings of the W and Z bosons to the Higgs boson at more the 5 standard deviations. The relative sign of these couplings determine how matrix elements from various Feynman diagrams combine in the calculation of the cross-section of a given process. In the case of VBF WH production, opposite sign W and Z boson couplings to the Higgs would lead to constructive interference, thereby en-

hancing the VBF WH cross-section. Two analyses were performed of the data events, which contain two b-jet Higgs decay candidates, one lepton W decay candidate, and two jets with high difference in rapidity. One analysis constrains the couplings, while the other makes a measurement and upper limit of the process cross-section. In the signal region of the analysis constraining the couplings, 70 data events were observed, where the SM same-sign coupling scenario would predict 80.6 ± 8.6 events and the opposite-sign coupling scenario with constructive interference of diagrams would yield a predicted 361 ± 46 events. The low yield in this signal region largely constrains the relative sign of these two couplings, and when combined with the constraint of the previous Higgs combination best fit, the opposite-sign scenario is completely excluded at more than 5σ. With this publication, the relative sign of the W and Z couplings to the Higgs is determined to be positive.

C. Validation of Kinematic Reweighting Uncertainty Approach

Relative to the size of the kinematic reweighting scale factors and impact on the NN distributions, the uncertainty associated is quite small. However, as evidenced by the post-fit kinematic plots shown in Chapter 11, there is no need for a closure-related systematic, so the only uncertainty is in the extrapolation of the scale factors derived in the control region and applied in the signal region. Since the reweighting strategy is quite robust (large CR statistics, High order polynomial reweighting functions, performed in reverse order of input variable importance), little is left undetermined. Additionally, impacts of reweighting of various variables wash out or cancel in the bins of the NN output, and the same happens with the reweighting's uncertainty. To validate this uncertainty strategy, several checks were performed. A check in which the parton shower variation samples underwent the same kinematic reweighting procedure, and a comparison of the output result is given in Section C.1.

§ C.1. Kinematic Reweighting of Parton Shower Variation Check

To assess the robustness of the kinematic reweighting strategy, a reweighting of the parton shower variations of the backgrounds is performed. Theoretically, the reweighting result should be more or less independent of the MC starting point, and the output should match between the reweighted Pythia and Herwig backgrounds. In practice, the ‘‘Herwig” sample already contains kinematically-correlated scale factors arising from transferring the truth-level distribution to the expectation of the fully-simulated, complexifying the result a tad. The output reweighted Pythia and ‘‘Herwig” distributions are shown in Figure C.1. The reweighted results of both sets of samples are well within statistical, kinematic reweighting, and parton shower uncertainties, further justifying that additional kinematic reweighting systematics are not necessary.

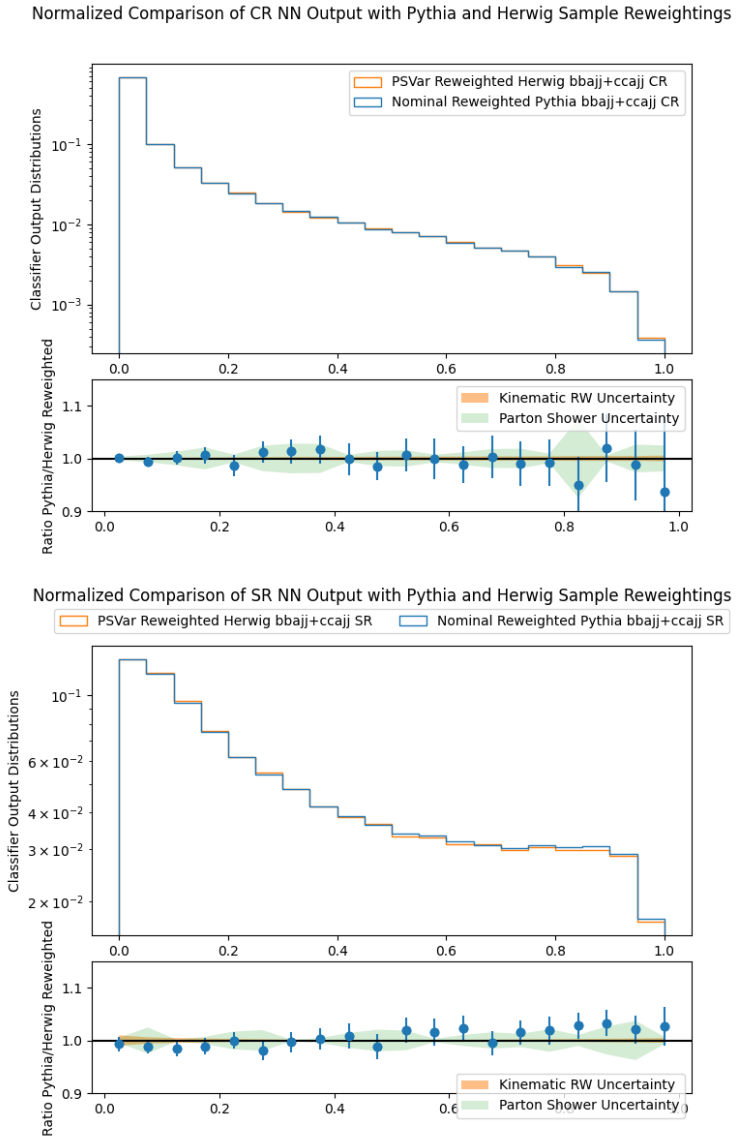
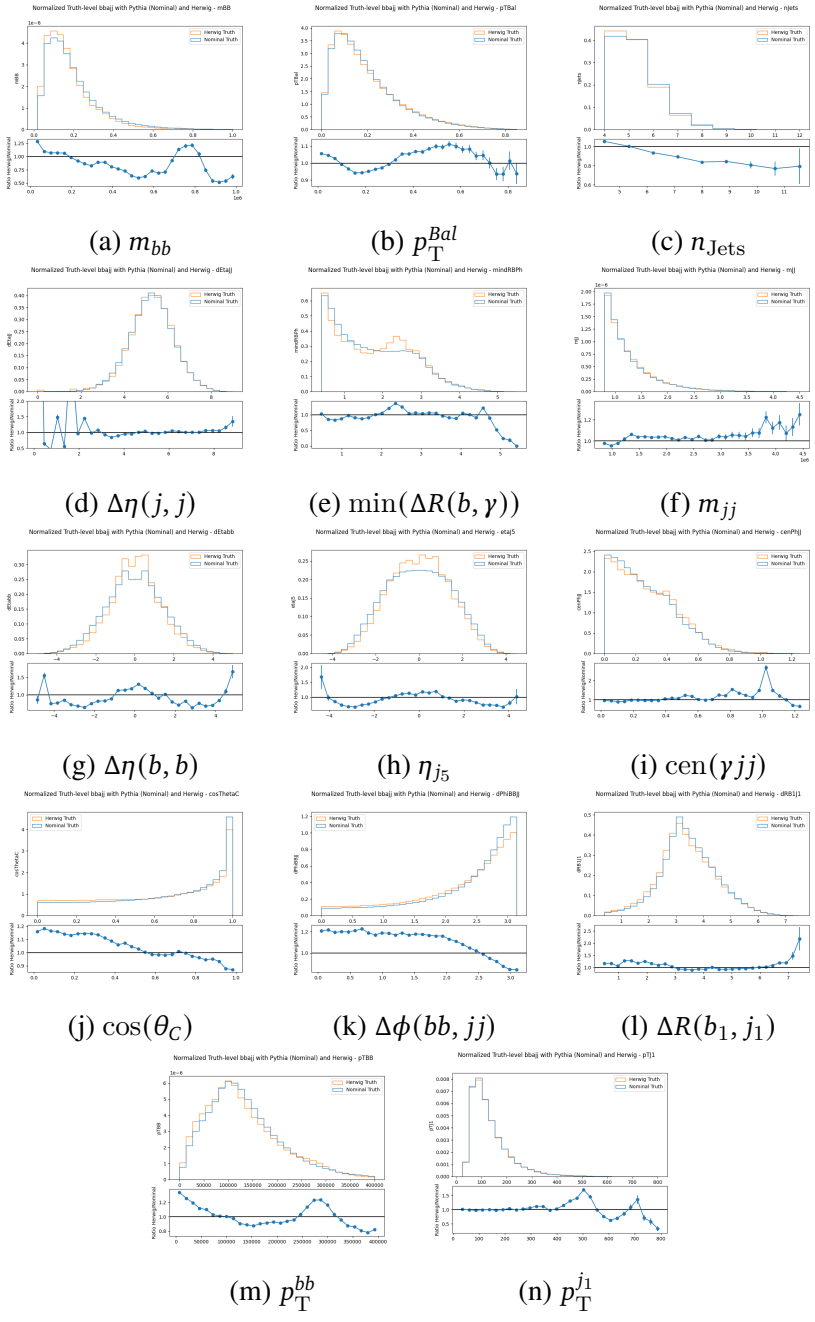


Figure C.1.: Comparison of the NN output distributions after kinematic reweighting of the nominal Pythia samples and of the Herwig parton shower variations. Kinematic reweighting and parton shower uncertainties (without smoothing) are overlaid on the ratio plot.

D. Parton Shower Uncertainty Derivation

The difference in kinematic distributions between the Pythia and Herwig -showered truth-level MC samples are shown in Figures D.1 and D.3 for $b\bar{b}\gamma jj$ and $c\bar{c}\gamma jj$, respectively. The transfer functions are also applied to the truth-level MC to show decent closure after the reweighting, as in Figures D.2 and D.4, despite ignoring the correlations between input variables.



(a) m_{bb} (b) p_{T}^{Bal} (c) n_{Jets}

(d) $\Delta\eta(j,j)$ (e) $\min(\Delta R(b,\gamma))$ (f) m_{jj}

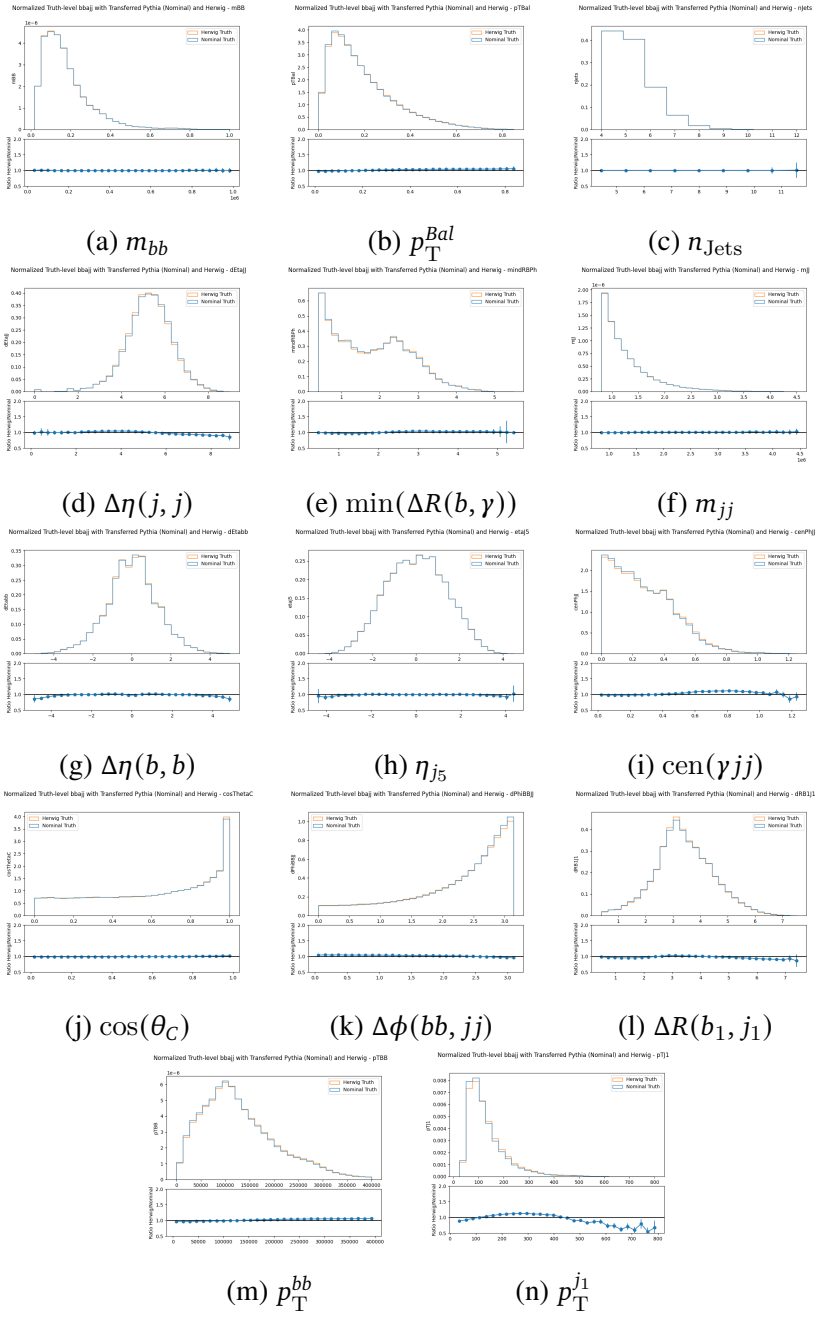
(g) $\Delta\eta(b,b)$ (h) η_{j_5} (i) $\mathrm{cen}(\gamma jj)$

(j) $\cos(\theta_C)$ (k) $\Delta\phi(bb,jj)$ (l) $\Delta R(b_1,j_1)$

(m) p_{T}^{bb} (n) $p_{\mathrm{T}}^{j_1}$

Figure D.1.: Truth-Level Monte Carlo comparison of kinematic MVA input variables for $bb\gamma jj$ showered with Pythia (as in nominal MC templates) and Herwig7 (Parton Shower Variation). Ratios are extracted as transfer functions to apply to the fully-simulated Pythia-showered NN output distribution to mimic a fully-simulated Herwig-showered NN output, used for evaluation of parton shower uncertainty for dominant $bb\gamma jj$ background.

D. Parton Shower Uncertainty Derivation



(a) m_{bb} (b) p_{T}^{Bal} (c) n_{Jets}

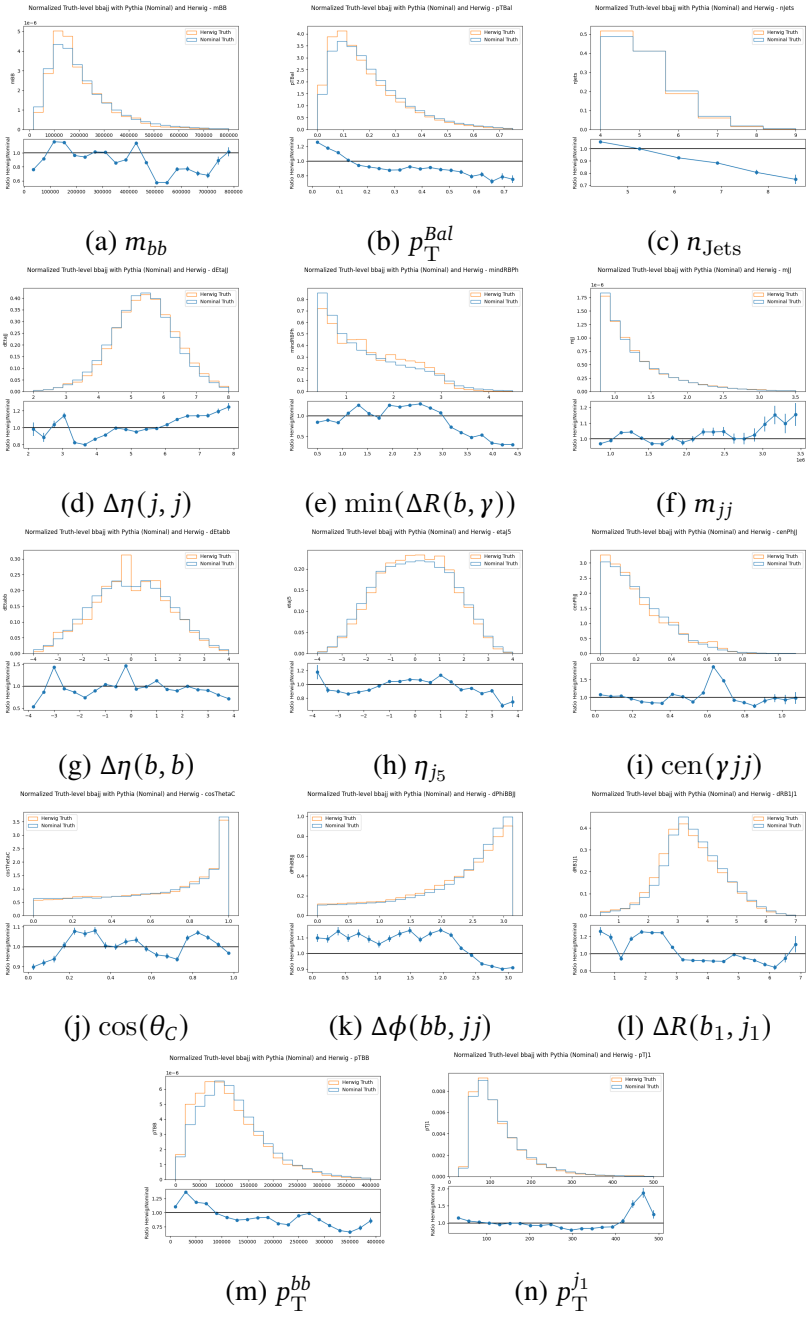
(d) $\Delta\eta(j,j)$ (e) $\min(\Delta R(b,\gamma))$ (f) m_{jj}

(g) $\Delta\eta(b,b)$ (h) η_{j_5} (i) $\mathrm{cen}(\gamma jj)$

(j) $\cos(\theta_C)$ (k) $\Delta\phi(bb,jj)$ (l) $\Delta R(b_1,j_1)$

(m) p_{T}^{bb} (n) $p_{\mathrm{T}}^{j_1}$

Figure D.2.: Comparison of kinematic MVA input variables after transfer factors are applied for truth-level $bb\gamma jj$ showered with Pythia and Herwig7. Ratios show good closure after all 14 transfer functions are applied, despite variable correlations being ignored in the iterative reweighting process. Only the variable pTJ1 shows appreciable differences remaining after the series of reweightings, and this is not a concern because it is not a very powerful variable in the NN.



(a) m_{bb} (b) p_{T}^{Bal} (c) n_{Jets}

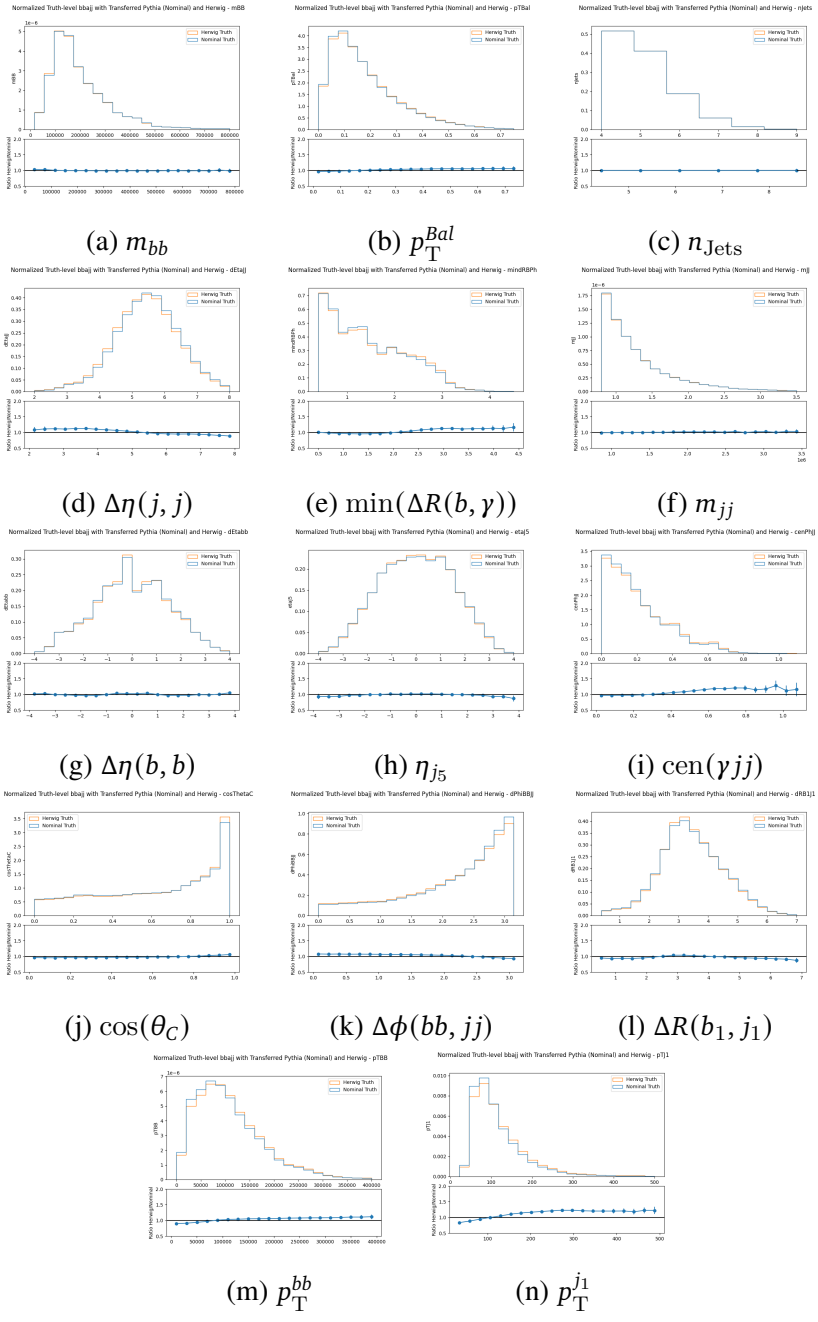
(d) $\Delta\eta(j,j)$ (e) $\min(\Delta R(b,\gamma))$ (f) m_{jj}

(g) $\Delta\eta(b,b)$ (h) η_{j_5} (i) $\mathrm{cen}(\gamma jj)$

(j) $\cos(\theta_C)$ (k) $\Delta\phi(bb,jj)$ (l) $\Delta R(b_1,j_1)$

(m) p_{T}^{bb} (n) $p_{\mathrm{T}}^{j_1}$

Figure D.3.: Truth-Level Monte Carlo comparison of kinematic MVA input variables for $cc\gamma jj$ showered with Pythia (as in nominal MC templates) and Herwig7 (Parton Shower Variation). Ratios are extracted as transfer functions to apply to the fully-simulated Pythia-showered NN output distribution to mimic a fully-simulated Herwig-showered NN output, used for evaluation of parton shower uncertainty for dominant $cc\gamma jj$ background.



(a) m_{bb} (b) p_{T}^{Bal} (c) n_{Jets}

(d) $\Delta\eta(j,j)$ (e) $\min(\Delta R(b,\gamma))$ (f) m_{jj}

(g) $\Delta\eta(b,b)$ (h) η_{j_5} (i) $\mathrm{cen}(\gamma jj)$

(j) $\cos(\theta_C)$ (k) $\Delta\phi(bb,jj)$ (l) $\Delta R(b_1,j_1)$

(m) p_{T}^{bb} (n) $p_{\mathrm{T}}^{j_1}$

Figure D.4.: Comparison of kinematic MVA input variables after transfer factors are applied for truth-level $cc\gamma jj$ showered with Pythia and Herwig7. Ratios show good closure after all 14 transfer functions are applied, despite variable correlations being ignored in the iterative reweighting process. Most ratios that depart from 1 are in bins with very low statistics, so not of concern.

E. Fit Bias Checks

§ E.1. Linearity of Extracted Signal Strength with Injected Signal Strength

In order to ensure the fitting strategy has not left us biased to the particular signal strength measurement of $\mu = 1$, we performed Asimov fits in which the signal strength was artificially increased and decreased. The ability of the fit to output a matching signal strength measurement to that injected demonstrates that our modeling is unbiased in this respect. Figures E.1 and E.2 depict the Asimov fit results when a signal strength of $\mu = 0.5$ and $\mu = 2$ are injected, respectively. Indeed, the fit is able to land right on the mark, extracting the same signal strength as was injected, demonstrating that our fit model is not biased in this respect.

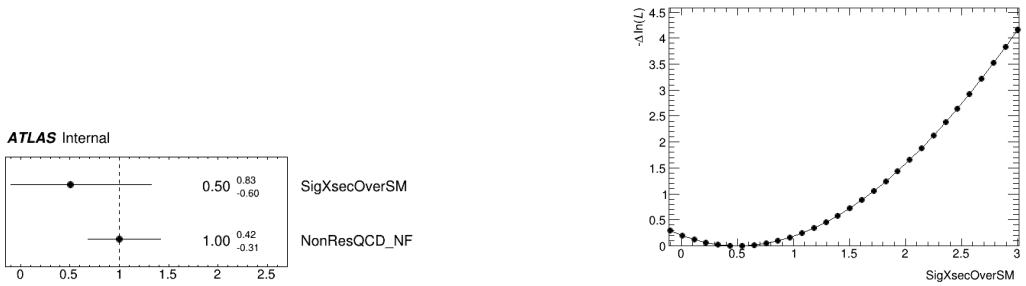


Figure E.1.: Post-fit results for the signal strength parameter of interest when a signal strength of $\mu = 0.5$ is injected into the Asimov fit.

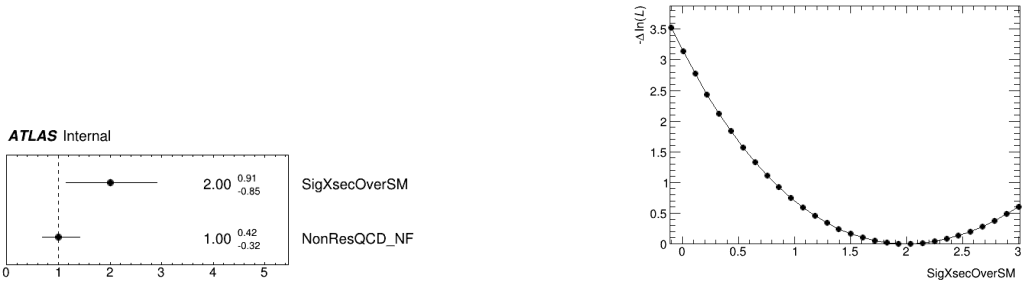


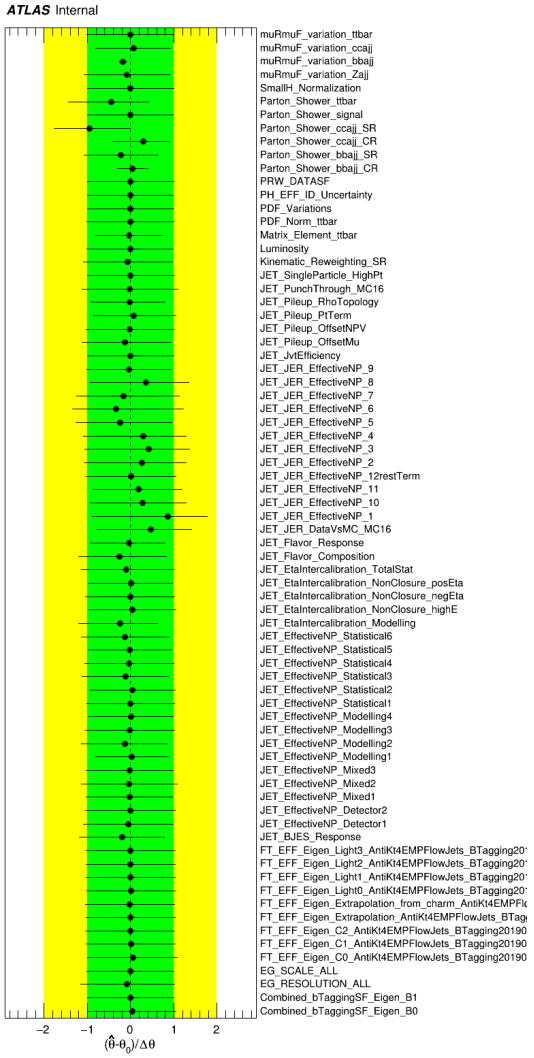
Figure E.2.: Post-fit results for the signal strength parameter of interest when a signal strength of $\mu = 2$ is injected into the Asimov fit.

§ E.2. Observation of Background NP Pulls with Exposure to Signal Region Bins

We test the stability of the background nuisance parameter pulls as the fit is allowed to use signal region bins to make its constraint. First, the low NN bins of the signal region were included, up to a score of 0.6, where there is very little signal contamination. Finally, with the permission from our editorial board and conveners, we exposed the fit to the entire phase-space, but with any signal nuisance parameters removed and with the signal strength fixed. In this way, this study only observes the stability of the background modeling and does not bias our search before unblinding.

The constraint and pulls of the nuisance parameters at these three different stages can be observed in Figure E.3. Overall, it is observed that the fit is quite stable as it is opened up to the signal region. No striking differences in pulls were noted from the control region only fit, except a regression towards the central value in some cases like the Jet Energy Resolution systematics. Some differences in the final fit, where the entire phase space is included, are due to the removal of the signal nuisance parameters, which was not done in the blinded fit regions.

(a) CR + Signal Region below Score of 0.6



(b) Full Control and Signal Regions

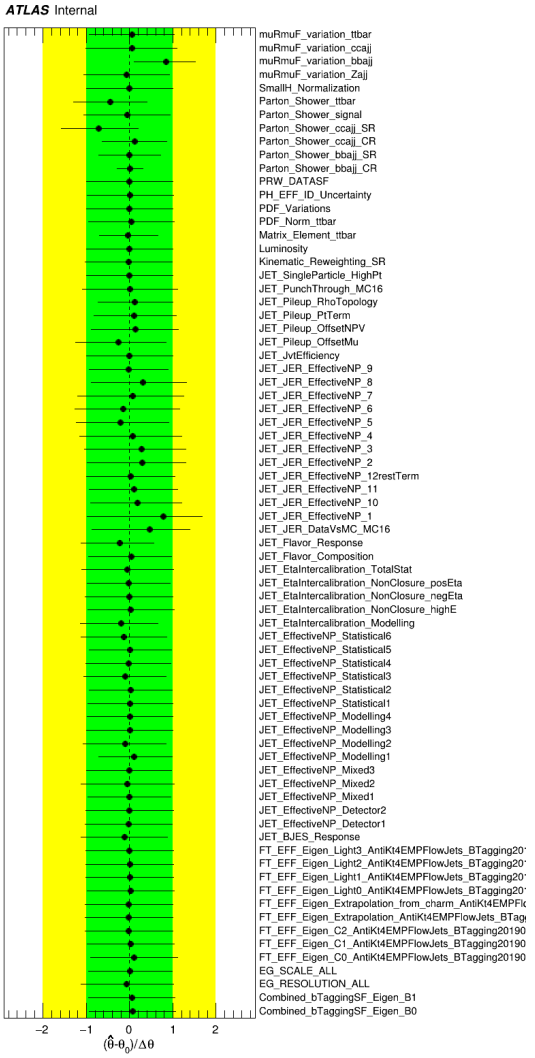


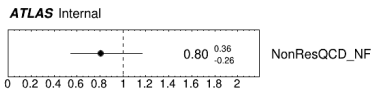
Figure E.3.: Pulls and constraints of nuisance parameters in background-only fits as the fit is opened to more of the phase space in the signal region. It is generally remarked that the fit is quite stable, and this test did not bring about major surprises or strong deviations in NP pulls.

§ E.3. Performing S+B Fit to Custom Asimov Built from Post-Fit Pulls in Background-Only Fit to Data

We perform a study to assess the bias in our fit model, or the tendency for the measured signal strength to be pulled higher or lower than the injected strength. Due to the lack of data statistics, it was not possible simulate signal region background data by transferring the control region data, as very few data events land in the high score bins in the control region. Instead, we designed a bias check that uses the post-fit results of the $\mu = 1$ conditional background-only fit. The post-fit nuisance parameter pulls from the conditional fit to data are injected to form a custom Asimov pseudo-dataset for this bias check with an injected signal strength, μ, of 1. The signal-only systematics are re-introduced, and the signal normalization is allowed to float once again. To this custom pseudo-dataset, a signal+background fit is performed, matching the fit architecture that is to be used for the final fit to data. If the fit model has no bias whatsoever, then a post-fit signal strength measurement of 1 is expected.

The extracted normalization factors from the conditional background-only fit and from the signal+background fit to the custom Asimov dataset are shown in Figure E.4. The extracted signal strength of 1.02 demonstrates that there is very minimal to negligible bias in our fit model. This bias is less than 3% of the size of our uncertainty on signal strength. Figure E.5 depicts that the signal extraction is smooth in the fit model as well, with no sharp jumps or discontinuities. Lastly, it is also comforting to see that the pulls of nuisance parameters are fairly consistent between the background-only data fit and signal+background Asimov fit, as shown in Figure E.6. With these details, it is evident that our fit model does not suffer from significant bias.

(a) Background-Only Conditional Data-Fit



(b) S+B Custom Asimov Fit

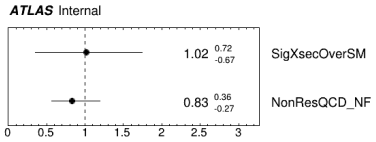


Figure E.4.: Post-fit normalization factors for $\mu = 1$ conditional background-only fit to complete set of control and signal region data (left), and for the signal+background bias check fit to a custom pseudo-dataset (right), built from the post-fit distributions of the conditional fit. An extracted signal strength of 1.0 marks no bias in the fit model.

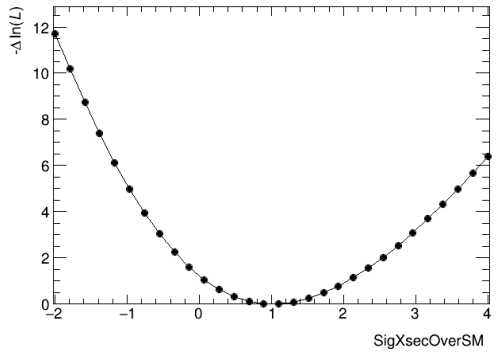


Figure E.5.: Likelihood scan for best-fit value of signal strength in fit bias test. Signal strength of 1 was injected, so an extracted strength of 1 represents no bias in the fit model.

(a) Background-Only Conditional Data-Fit

(b) S+B Custom Asimov Fit

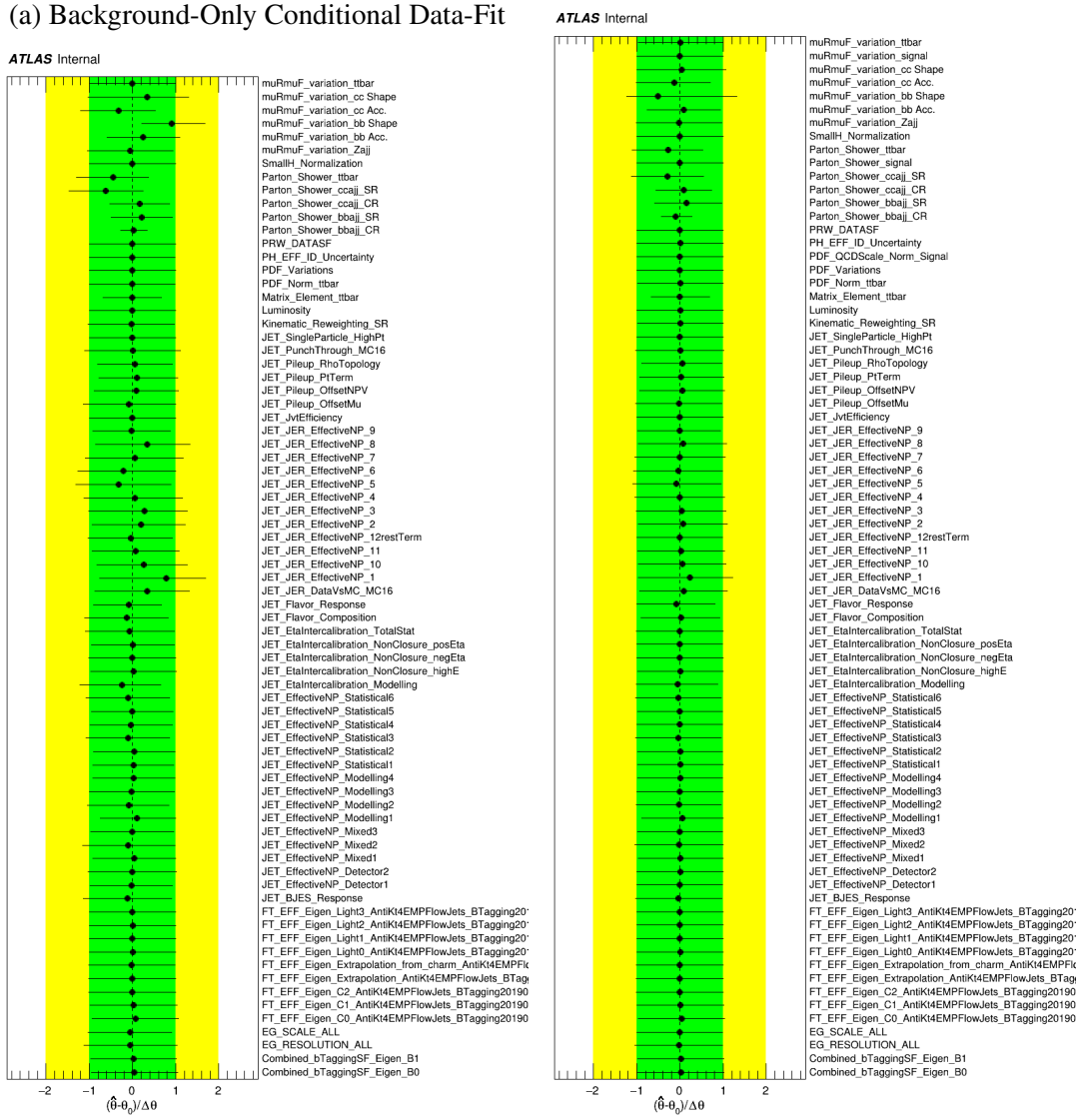
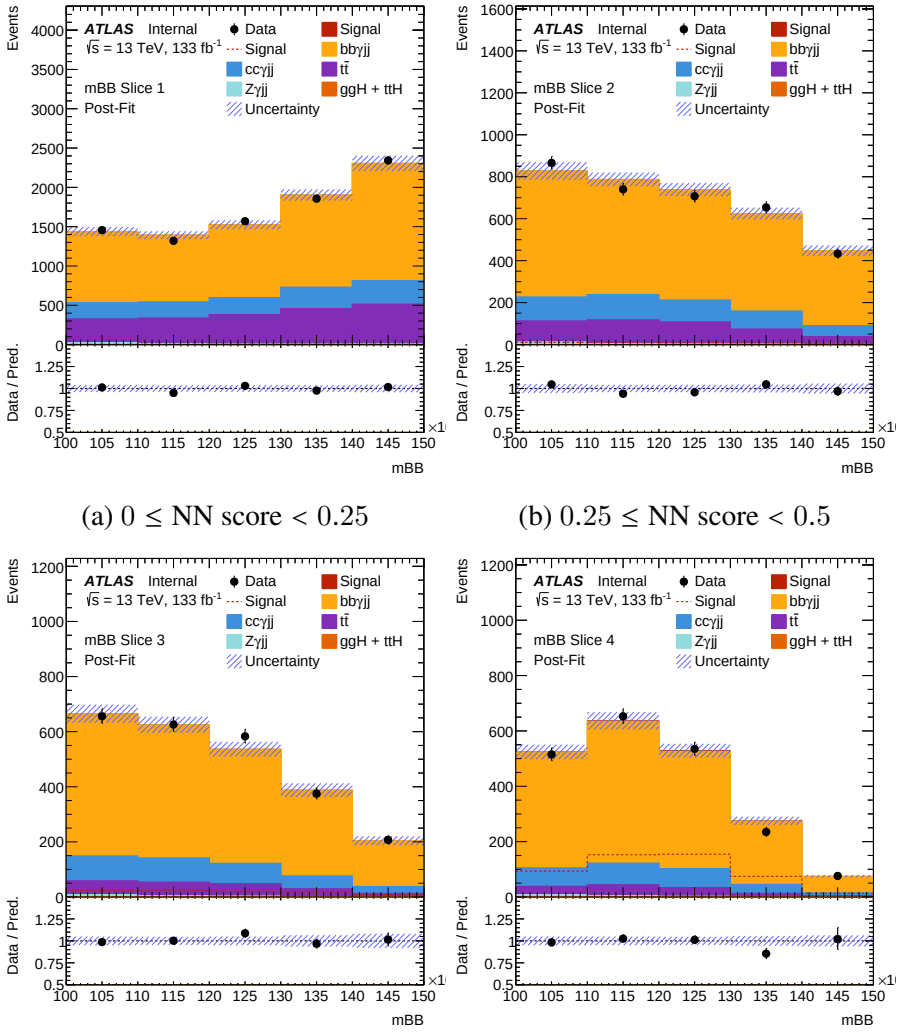


Figure E.6.: Post-fit pulls of all nuisance parameters in a background-only conditional fit to data (left) and in the signal+background bias test fit to a custom Asimov pseudo-dataset (right). It is comforting to see mostly consistent pulls between the two fits, suggesting that the re-introduction of signal systematics does not disturb the fit stability.

F. Post-fit Signal Region Kinematic Distributions in NN Score Slices

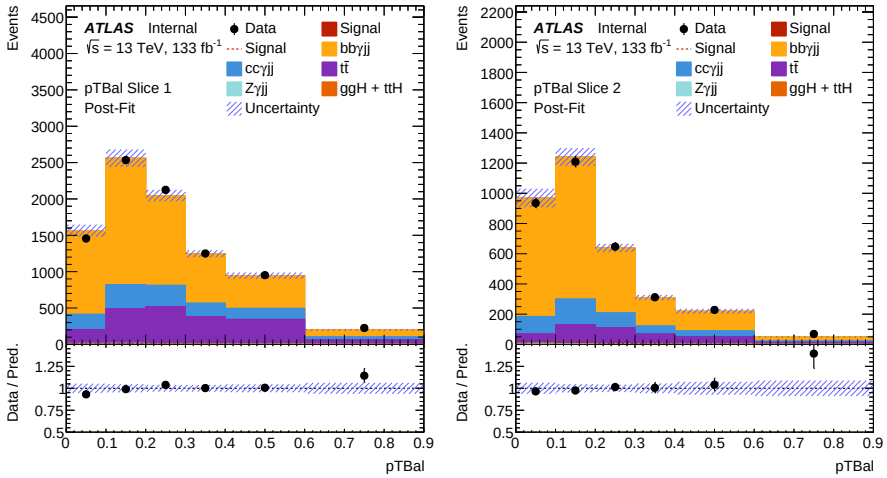
The post-fit agreement between MC and data for the input kinematic distributions shown in Section 11.2 is reassuring. However, to further confirm the modeling in the high score signal-sensitive bins of the signal region, post-fit kinematic comparisons are made in slices of NN score for some of the most important input variables. Figures F.1, F.2, and F.3 show the post-fit agreement in the signal region between MC and data in four slices of NN score for m_{bb}, p_{T}^{Bal}, and n_{Jets}, respectively. MC and data agree within post-fit uncertainties, and no divergence in agreement is observed in correlation to NN score. This is non-trivial due to the use of kinematic reweighting scale factors that were extracted in the control region and extrapolated to this signal region. This check demonstrates that the modeling, even in the most signal sensitive bins, can be trusted.



(a) $0 \leq$ NN score < 0.25 (b) $0.25 \leq$ NN score < 0.5

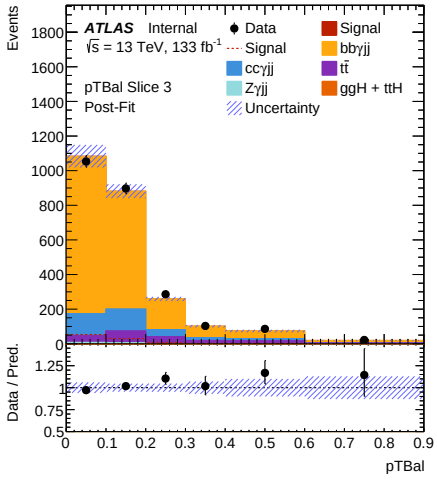
(c) $0.5 \leq$ NN score < 0.75 (d) $0.75 \leq$ NN score ≤ 1

Figure F.1.: Post-fit signal region m_{bb} in slices of NN score. Post-fit uncertainties are depicted by blue-hashed lines. MC and data agree within post-fit uncertainties, and no divergence in agreement is observed in correlation to NN score. The red dashed histograms represent pre-fit signal distributions scaled by a factor of 10.

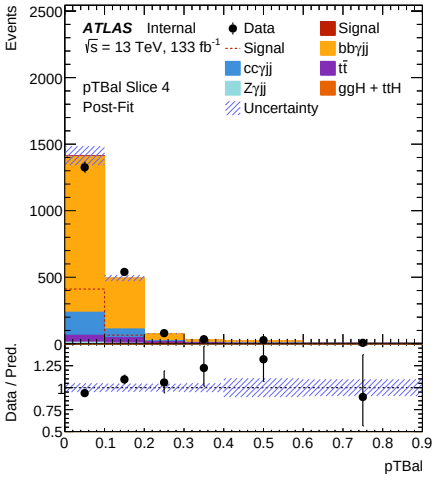


(a) $0 \leq$ NN score < 0.25

(b) $0.25 \leq$ NN score < 0.5

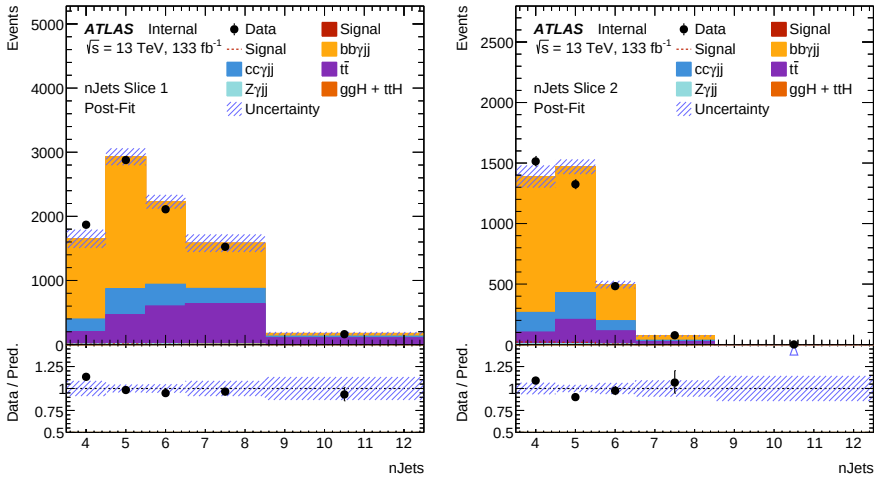


(c) $0.5 \leq$ NN score < 0.75



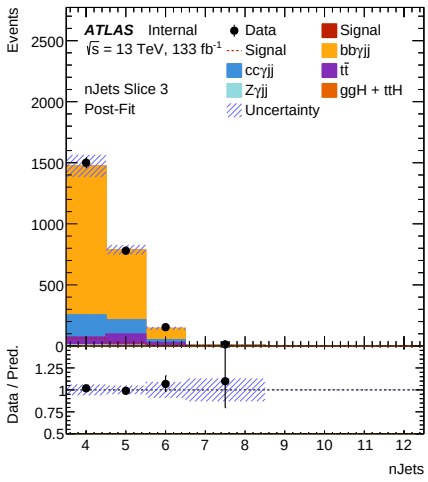
(d) $0.75 \leq$ NN score ≤ 1

Figure F.2.: Post-fit signal region p_{T}^{Bal} in slices of NN score. Post-fit uncertainties are depicted by blue-hashed lines. MC and data agree within post-fit uncertainties, and no divergence in agreement is observed in correlation to NN score. The red dashed histograms represent pre-fit signal distributions scaled by a factor of 10.

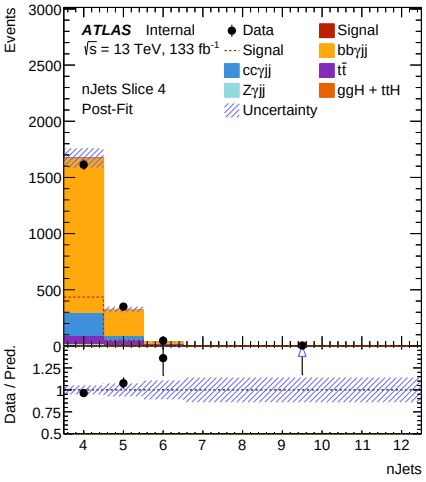


(a) $0 \leq$ NN score < 0.25

(b) $0.25 \leq$ NN score < 0.5



(c) $0.5 \leq$ NN score < 0.75



(d) $0.75 \leq$ NN score ≤ 1

Figure F.3.: Post-fit signal region n_{Jets} in slices of NN score. Post-fit uncertainties are depicted by blue-hashed lines. MC and data agree within post-fit uncertainties, and no divergence in agreement is observed in correlation to NN score. The red dashed histograms represent pre-fit signal distributions scaled by a factor of 10.

G. Compatibility with Previous Result

A previous analysis of the full Run 2 dataset for this VBF $H + \gamma$ signature achieved a signal strength of $\mu_H = 1.3 \pm 1.0$. Here, we check the compatibility of our signal measurement with this previous result.

A fit with the signal strength fixed to the best-fit value of the previous analysis (μ=1.3) achieved a NLL value which was 1.18 above the NLL from our final fit. This would correspond to a p-value of 27.7% if treated as a chi-squared value with 1 degree of freedom and no correlation taken into account. We also attempt to assess the amount of correlation that it would take to make the difference in observed results significant (p-value less than 5%). Comparing the two measurements with an assumption of 0 correlation, a chi-squared of 0.8 is found which corresponds to a p-value of 37% for one degree of freedom. It would take a correlation of at least 84% to achieve a p-value of 5% or lower. The overlap of data events landing in the signal sensitive regions of the two analyses is detailed in Table G.1. Of 317 data events in the Higgs mass window and in the high BDT region of the previous analysis, only 7 have a high output score for the NN in this analysis, making for an overlap of around 2%. With this low overlap of data in the signal sensitive regions and the extreme differences in signal extraction strategy, there is not sufficient correlation to mark incompatibility between the current and previous measurements.

	High BDT	Medium BDT	Low BDT	None	Total
SR (score>0.8)	7	57	2038	5461	7563
None	310	1453	7259		
Total	317	1510	9297		

Table G.1.: Overlap in data events between the signal sensitive regions of this analysis with the previous. The previous analysis used regions of purity defined by BDT score, and the signal sensitive bins are those in a Higgs mass window range of 100 GeV $\leq m_{bb} \leq 140$ GeV. For the current work, the signal sensitive region is in the Higgs mass window of $100\,\text{GeV} \leq m_{bb} \leq 150\,\text{GeV}$ in a NN score range of greater than 0.8.

Bibliography

[1] T. Young. ‘‘The Bakerian Lecture. Experiments and calculations relative to physical optics’’. en. In: *Philos. Trans. R. Soc. Lond.* 94.0 (Dec. 1804), pp. 1–16.

[2] John Cardy. *Introduction to Quantum Field Theory*. https://www-thphys.physics.ox.ac.uk/people/JohnCardy/qft/qftcomplete.pdf.

[3] Giampiero Esposito, Alexander Yu Kamenshchik, and Giuseppe Pollifrone. ‘‘Effective Action in Quantum Field Theory’’. In: *Euclidean Quantum Gravity on Manifolds with Boundary*. Dordrecht: Springer Netherlands, 1997, pp. 51–79.

[4] Michael Edward Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Reading, USA: Addison-Wesley (1995) 842 p. Westview Press, 1995.

[5] Gerardus ’t Hooft and Martinus J G Veltman. *Diagrammar*. CERN Yellow Reports: Monographs. Geneva: CERN, 1973. DOI: 10.5170/CERN-1973-009. URL: https://cds.cern.ch/record/186259.

[6] Tina Potter. *Feynman Diagrams*. https://www.hep.phy.cam.ac.uk/~chpotter/particleandnuclearphysics/Lecture_05_FeynmanDiagrams.pdf.

[7] David A Edwards. ‘‘Mathematical foundations of quantum field theory: Fermions, gauge fields, and supersymmetry part I: Lattice field theories’’. en. In: *Int. J. Theor. Phys.* 20.7 (July 1981), pp. 503–517.

[8] Robert V Harlander and Jean-Philippe Martinez. ‘‘The development of computational methods for Feynman diagrams’’. en. In: *Eur. Phys. J. H* 49.1 (Dec. 2024), pp. 1–45.

[9] Fionn Fitzmaurice. *Symmetry in Particle Physics*. https://www.maths.tcd.ie/~fionn/misc/symmetry/symmetry.pdf.

[10] Gerard ’t Hooft. ‘‘Gauge Theories of the Forces Between Elementary Particles’’. In: *Sci. Am.* 242N6 (1980). Ed. by Richard A. Carrigan and W. P. Trower, pp. 90–116.

[11] Michael M. Rijssenbeek. *Unitary Groups and SU(N)*. https://sbhep.physics.sunysb.edu/~rijssenbeek/PHY557/SU(2)-SU(N)_YoungTableaux.pdf.

[12] *The Standard Model*. https://home.cern/science/physics/standard-model.

[13] Y Fukuda et al. ''Evidence for oscillation of atmospheric neutrinos''. In: *Phys. Rev. Lett.* 81.8 (Aug. 1998), pp. 1562–1567.

[14] Julia Woithe, Gerfried J Wiener, and Frederik F Van der Veken. ''Let's have a coffee with the Standard Model of particle physics!'' In: *Phys. Educ.* 52.3 (May 2017), p. 034001.

[15] C. S. Wu et al. ''Experimental Test of Parity Conservation in Beta Decay''. In: *Phys. Rev.* 105 (4 Feb. 1957), pp. 1413–1415. DOI: 10.1103/PhysRev.105.1413. URL: https://link.aps.org/doi/10.1103/PhysRev.105.1413.

[16] Antonio Pich. ''The Standard Model of Electroweak Interactions; rev. version''. In: (2008). DOI: 10.5170/CERN-2006-003.1. URL: https://cds.cern.ch/record/819632.

[17] P.W. Higgs. ''Broken symmetries, massless particles and gauge fields''. In: *Physics Letters* 12.2 (1964), pp. 132–133. ISSN: 0031-9163. DOI: https://doi.org/10.1016/0031-9163(64)91136-9. URL: https://www.sciencedirect.com/science/article/pii/0031916364911369.

[18] F Englert and R Brout. ''Broken symmetry and the mass of gauge vector mesons''. en. In: *Phys. Rev. Lett.* 13.9 (Aug. 1964), pp. 321–323.

[19] Michela D'Onofrio and Kari Rummukainen. ''Standard model cross-over on the lattice''. In: *Physical Review D* 93.2 (Jan. 2016). ISSN: 2470-0029. DOI: 10.1103/physrevd.93.025003. URL: http://dx.doi.org/10.1103/PhysRevD.93.025003.

[20] Steven Weinberg. *The Quantum Theory of Fields, Volume 1: Foundations*. Cambridge University Press, 2005. ISBN: 0521670535.

[21] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. ''Broken Symmetries''. In: *Phys. Rev.* 127 (3 Aug. 1962), pp. 965–970. DOI: 10.1103/PhysRev.127.965. URL: https://link.aps.org/doi/10.1103/PhysRev.127.965.

[22] Suzanne Dam. *Spontaneous Symmetry Breaking in the Higgs Mechanism*. https://philsci-archive.pitt.edu/9295/1/Spontaneous_symmetry_breaking_in_the_Higgs_mechanism.pdf.

[23] Mathilda Turtola. *Electroweak Symmetry Breaking and Higgs Pair Production in Effective Field Theories*. https://uu.diva-portal.org/smash/get/diva2:1882276/FULLTEXT01.pdf. 2024.

[24] Arcadi Santamaria. ‘‘Masses, mixings, Yukawa couplings and their symmetries’’. In: *Phys. Lett. B* 305 (1993), pp. 90–97. DOI: 10.1016/0370-2693(93)91110-9. URL: https://cds.cern.ch/record/246574.

[25] Howard E. Haber. *Higgs Boson Physics – The View Ahead*. 2023. arXiv: 2210.00449 [hep-ph]. URL: https://arxiv.org/abs/2210.00449.

[26] CERN. *Taking a Closer Look at LHC*. https://www.lhc-closer.es/taking_a_closer_look_at_lhc.

[27] R Bailey and Paul Collier. *Standard Filling Schemes for Various LHC Operation Modes*. Tech. rep. Geneva: CERN, 2003. URL: https://cds.cern.ch/record/691782.

[28] Bernhard Holzer. *Transverse Beam Dynamics*. https://indico.cern.ch/event/817381/contributions/3412311/attachments/1835890/3178447/Transverse_Beam_Dynamics_2p.pdf. 2019.

[29] Esma Mobs. ‘‘The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018’’. In: (2018). General Photo. URL: https://cds.cern.ch/record/2636343.

[30] Dan Faircloth and Scott Lawrie. ‘‘An overview of negative hydrogen ion sources for accelerators’’. In: *New Journal of Physics* 20.2 (Feb. 2018), p. 025007. DOI: 10.1088/1367-2630/aaa39e. URL: https://dx.doi.org/10.1088/1367-2630/aaa39e.

[31] Karlheinz Schindl. ‘‘The injector chain for the LHC’’. In: (1999). URL: https://cds.cern.ch/record/397574.

[32] CERN. *Accelerators*. https://home.cern/science/accelerators.

[33] Ana Lopes and Melissa Loyse Perrey. ‘‘FAQ-LHC The guide’’. 2022. URL: https://cds.cern.ch/record/2809109.

[34] Matteo Solfaroli Camillocci. *LHC Nominal Cycle*. https://indico.cern.ch/event/434129/contributions/1917195/attachments/1205096/1765722/Nominal_cycle.pdf.

[35] Mike Lamont. *The LHC's first long run*. https://cerncourier.com/a/the-lhcs-first-long-run-2/. Aug. 2013.

[36] ATLAS Collaboration. ‘‘Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC’’. In: *The European Physical Journal C* 83.10 (Oct. 2023). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-023-11747-w. URL: http://dx.doi.org/10.1140/epjc/s10052-023-11747-w.

[37] ATLAS. *Public Luminosity Results*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun3. 2025.

[38] CERN. *High-Luminosity LHC FAQ*. https://home.web.cern.ch/resources/faqs/high-luminosity-lhc.

[39] Joao Pequenao. ''Computer generated image of the whole ATLAS detector''. 2008. URL: http://cds.cern.ch/record/1095924.

[40] Sebastian Murk. ''Trigger Algorithms and Monte Carlo Event Generation for Dijet Searches in ATLAS and Beyond''. In: 2017. URL: https://api.semanticscholar.org/CorpusID:54570830.

[41] ATLAS. *The Inner Detector*. https://atlas-public.web.cern.ch/Discover/Detector/Inner-Detector.

[42] 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.

[43] *Glossary | ATLAS Open Data*. https://opendata.atlas.cern/docs/8TeVDoc/glossary.

[44] Yosuke Takubo. *ATLAS IBL operational experience*. Tech. rep. Geneva: CERN, 2017. DOI: 10.22323/1.287.0004. URL: https://cds.cern.ch/record/2235541.

[45] Norbert Wermes and G Hallewel. *ATLAS pixel detector: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1998. URL: https://cds.cern.ch/record/381263.

[46] Bartosz Mindur. *ATLAS Transition Radiation Tracker (TRT): Straw tubes for tracking and particle identification at the Large Hadron Collider*. Tech. rep. Geneva: CERN, 2017. DOI: 10.1016/j.nima.2016.04.026. URL: https://cds.cern.ch/record/2139567.

[47] T Akesson et al. ''Electron identification with a prototype of the Transition Radiation Tracker for the ATLAS experiment: ATLAS TRT collaboration''. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 412.2 (1998), pp. 200–215. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(98)00457-4. URL: https://www.sciencedirect.com/science/article/pii/S0168900298004574.

[48] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical design report. ATLAS. Geneva: CERN, 1996. DOI: 10.17181/CERN.FWRW.FOOQ. URL: https://cds.cern.ch/record/331061.

[49] ATLAS. *Calorimeter*. https://atlas-public.web.cern.ch/Discover/Detector/Calorimeter.

[50] ATLAS Collaboration. ''Deep generative models for fast photon shower simulation in ATLAS''. In: *Comput. Softw. Big Sci.* 8 (2022), p. 7. DOI: 10.1007/s41781-023-00106-9. arXiv: 2210.06204 [hep-ex].

[51] ATLAS Collaboration. ''Study of energy response and resolution of the ATLAS Tile Calorimeter to hadrons of energies from 16 to 30 GeV''. en. In: *Eur. Phys. J. C Part. Fields* 81.6 (June 2021).

[52] ATLAS. *Muon Spectrometer*. https://atlas-public.web.cern.ch/Discover/Detector/Muon-Spectrometer.

[53] Fabian Kuger. ''Signal Formation Processes in Micromegas Detectors and Quality Control for large size Detector Construction for the ATLAS New Small Wheel''. In: (Aug. 2017). DOI: 10.48550/arXiv.1708.01624.

[54] T. Kawamoto et al. *New Small Wheel Technical Design Report*. Tech. rep. ATLAS New Small Wheel Technical Design Report. 2013. URL: https://cds.cern.ch/record/1552862.

[55] Nicolas Viaux Maira. *The ATLAS New Small Wheel new Muon Stations Ready for LHC Run3*. Tech. rep. Geneva: CERN, 2023. DOI: 10.1016/j.nima.2022.167574. URL: https://cds.cern.ch/record/2812136.

[56] ATLAS. *Magnet System*. https://atlas.cern/Discover/Detector/Magnet-System.

[57] H. H. J. ten Kate. ''Superconducting magnet system for the ATLAS detector at CERN''. In: *IEEE Trans. Appl. Supercond.* 9.2, pt.1 (1999), pp. 841–846. DOI: 10.1109/77.783428. URL: https://cds.cern.ch/record/409763.

[58] IRFU. *ATLAS-Muons*. https://irfu.cea.fr/en/Phocea/Vie_des_labos/Ast/ast_sstechnique.php?id_ast=387.

[59] Golabz. *Measurement Of Magnetic Field With The Giant Atlas Solenoid*. https://www.golabz.eu/ils/measurement-of-magnetic-field-with-the-giant-atlas-solenoid.

[60] Abha Eli Phoboo. *Dealing With Data*. https://atlas.cern/updates/news/dealing-data. 2014.

[61] ATLAS Collaboration. ''ATLAS Event Displays: Higgs Boson candidates decaying to two Z bosons (ZZ) and four leptons''. General Photo. 2015. URL: https://cds.cern.ch/record/2854878.

[62] Emily Ann Smith. *The phase-1 upgrade of the ATLAS level-1 calorimeter trigger*. Tech. rep. Geneva: CERN, 2022. DOI: 10.22323/1.398.0754. URL: https://cds.cern.ch/record/2789104.

[63] Aleksandra Poreba. ''Operational experience with the new ATLAS HLT framework for LHC Run 3''. In: *EPJ Web Conf.* 295 (2024), p. 02021.

[64] CERN. *Facts and figures about the LHC*. https://home.cern/resources/faqs/facts-and-figures-about-lhc.

[65] S. Schlenker et al. *The ATLAS Detector Control System*. https://accelconf.web.cern.ch/icalepcs2011/papers/mobaust02.pdf.

[66] Markus Keil. ''Online calibration and performance of the ATLAS Pixel Detector''. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 650.1 (2011). International Workshop on Semiconductor Pixel Detectors for Particles and Imaging 2010, pp. 9–13. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2010.11.150. URL: https://www.sciencedirect.com/science/article/pii/S0168900210026975.

[67] B Di Girolamo. *The ATLAS Pixel Detector*. Tech. rep. Geneva: CERN, 2011. URL: https://cds.cern.ch/record/1385028.

[68] ATLAS. ''Modelling radiation damage to pixel sensors in the ATLAS detector''. In: *Journal of Instrumentation* 14.06 (2019). ISSN: 1748-0221. DOI: 10.1088/1748-0221/14/06/p06012. URL: http://dx.doi.org/10.1088/1748-0221/14/06/P06012.

[69] A Gorisek. ''The ATLAS Beam Condition Monitor Commissioning''. In: (2008). DOI: 10.5170/CERN-2008-008.264. URL: https://cds.cern.ch/record/1158638.

[70] Oleg Solovyanov. *ATLAS upgrades for High Luminosity LHC*. Tech. rep. Geneva: CERN, 2024. URL: https://cds.cern.ch/record/2904688.

[71] ATLAS Collaboration. *ATLAS Tile Calorimeter Phase-II Upgrade: Technical Design Report*. ATLAS-TDR-028; CERN-LHCC-2017-019. 2017. URL: https://cds.cern.ch/record/2285583.

[72] ATLAS Collaboration. *ATLAS Muon Spectrometer Phase-II Upgrade: Technical Design Report*. ATLAS-TDR-026; CERN-LHCC-2017-017. 2017. URL: https://cds.cern.ch/record/2285580.

[73] ATLAS Collaboration. *A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade: Technical Design Report*. ATLAS-TDR-031; CERN-LHCC-2020-007. 2020. URL: https://cds.cern.ch/record/2719855.

[74] Gabriele Giacomini. ‘‘LGAD-Based Silicon Sensors for 4D Detectors’’. In: *Sensors* 23.4 (2023). ISSN: 1424-8220. DOI: 10.3390/s23042132. URL: https://www.mdpi.com/1424-8220/23/4/2132.

[75] *Investigating the impact of 4D Tracking in ATLAS Beyond Run 4*. Tech. rep. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2023-023. Geneva: CERN, 2023. URL: http://cds.cern.ch/record/2870326.

[76] ATLAS Collaboration. *ATLAS TDAQ Phase-II Upgrade: Technical Design Report*. ATLAS-TDR-029; CERN-LHCC-2017-020. 2017. URL: https://cds.cern.ch/record/2285584.

[77] Francesca Pastore. *ATLAS TDAQ upgrades for Phase-2*. Tech. rep. Geneva: CERN, 2024. URL: https://cds.cern.ch/record/2907275.

[78] Sebastien Roy-Garand. *ATLAS Inner TracKer Upgrade. 11th Large Hadron Collider Physics Conference 2023*. Tech. rep. Geneva: CERN, 2023. DOI: 10.22323/1.450.0002. URL: https://cds.cern.ch/record/2869708.

[79] Y. Unno et al. ‘‘Development of n-on-p silicon sensors for very high radiation environments’’. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 636.1, Supplement (2011). 7th International ””Hiroshima”” Symposium on the Development and Application of Semiconductor Tracking Detectors, S24–S30. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2010.04.080. URL: https://www.sciencedirect.com/science/article/pii/S0168900210009344.

[80] *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*. Tech. rep. Geneva: CERN, 2017. DOI: 10.17181/CERN.FOZZ.ZP3Q. URL: https://cds.cern.ch/record/2285585.

[81] Elisabetta Pianori et al. *LabRemote*. https://gitlab.cern.ch/berkeleylab/labRemote/-/tree/main. 2018.

[82] Hava Rhian Schwartz. *Graphical User Interface for Accessing LabRemote Equipment (GALE) Documentation: Details regarding the installation, customization, and use of this app for remotely accessing hardware lab equipment*. Tech. rep. This document is in relation to a Qualification Task. The goal of the task was to develop a user-friendly interface for accessing hardware lab equipment. Geneva: CERN, 2022. URL: https://cds.cern.ch/record/2813206.

[83] ATLAS Collaboration. ‘‘A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery’’. en. In: *Nature* 607.7917 (July 2022), pp. 52–59.

[84] ATLAS. *ATLAS measures the Higgs boson at 13.6 TeV*. https://atlas.cern/Updates/Briefing/Run3-Higgs. May 2023.

[85] *ATLAS and CMS release results of most comprehensive studies yet of Higgs boson's properties*. https://home.cern/news/news/physics/atlas-and-cms-release-results-most-comprehensive-studies-yet-higgs-bosons.

[86] Robert V. Harlander, Stefan Liebler, and Tom Zirke. ''Higgs Strahlung at the Large Hadron Collider in the 2-Higgs-doublet model''. In: *Journal of High Energy Physics* 2014.2 (Feb. 2014). ISSN: 1029-8479. DOI: 10.1007/jhep02(2014)023. URL: http://dx.doi.org/10.1007/JHEP02(2014)023.

[87] Dennis Silverman. *Higgs Production and Decay Channels*. https://sites.uci.edu/energyobserver/2012/11/26/higgs-production-and-decay-channels/. 2012.

[88] *The Higgs Boson | ATLAS Open Data*. https://opendata.atlas.cern/docs/visualization/the-higgs-boson.

[89] ATLAS Collaboration. ''Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC''. In: *Phys. Lett. B* 716 (2012), p. 1. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[90] ''Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC''. In: *Physics Letters B* 716.1 (Sept. 2012), pp. 30–61. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.021. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.021.

[91] Emidio Gabrielli et al. ''Asking for an extra photon in Higgs production at the LHC and beyond''. In: (2016). DOI: 10.1007/JHEP07(2016)003. eprint: arXiv:1601.03635.

[92] W. H. Furry. ''A symmetry theorem in the positron theory''. In: *Phys. Rev.* 51.2 (Jan. 1937), pp. 125–129.

[93] Emidio Gabrielli et al. ''Asking for an extra photon in Higgs production at the LHC and beyond''. In: *Journal of High Energy Physics* 2016.7 (July 2016). ISSN: 1029-8479. DOI: 10.1007/jhep07(2016)003. URL: http://dx.doi.org/10.1007/JHEP07(2016)003.

[94] 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 $\sqrt{s} = 13$ TeV with the ATLAS detector''. In: *JHEP* 03 (2021), p. 268. DOI: 10.1007/JHEP03(2021)268. arXiv: 2010.13651 [hep-ex].

[95] B. Denby. ''Neural networks and cellular automata in experimental high energy physics''. In: *Computer Physics Communications* 49.3 (1988), pp. 429–448. ISSN: 0010-4655. DOI: https://doi.org/10.1016/0010-4655(88)90004-5. URL: https://www.sciencedirect.com/science/article/pii/0010465588900045.

[96] *Introduction to Monte Carlo Simulations | ATLAS Open Data*. https://opendata.atlas.cern/docs/documentation/monte_carlo/introduction_MC.

[97] Matthew Gignac. *Monte Carlo generators, and their usage/modeling in ATLAS*. https://indico.cern.ch/event/801520/contributions/3530704/attachments/1912970/3162045/2019_09_23_SUSYIntroMC_MGignac.pdf. 2019.

[98] J. Alwall et al. ''The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations''. In: *Journal of High Energy Physics* 2014.7 (July 2014). ISSN: 1029-8479. DOI: 10.1007/jhep07(2014)079. URL: http://dx.doi.org/10.1007/JHEP07(2014)079.

[99] Stefan Höche. *Introduction to parton-shower event generators*. 2015. arXiv: 1411.4085 [hep-ph]. URL: https://arxiv.org/abs/1411.4085.

[100] Peter Z. Skands. ''QCD for Collider Physics''. In: (2012). Comments: Lectures from the 2010 European School of High Energy Physics (ESHEP 2010), Raseborg, Finland. DOI: 10.5170/CERN-2012-001.51. arXiv: 1104.2863. URL: https://cds.cern.ch/record/1345043.

[101] T. Gleisberg et al. ''Event generation with SHERPA 1.1''. In: *Journal of High Energy Physics* 2009.02 (Feb. 2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. URL: https://dx.doi.org/10.1088/1126-6708/2009/02/007.

[102] Manuel Bähr et al. ''Herwig++ physics and manual''. In: *The European Physical Journal C* 58.4 (Nov. 2008), pp. 639–707. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-008-0798-9. URL: http://dx.doi.org/10.1140/epjc/s10052-008-0798-9.

[103] Johannes Bellm et al. ''Herwig 7.0/Herwig++ 3.0 release note''. In: *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].

[104] Christian Bierlich et al. *A comprehensive guide to the physics and usage of PYTHIA 8.3*. 2022. arXiv: 2203.11601 [hep-ph]. URL: https://arxiv.org/abs/2203.11601.

[105] Aneta Siemiginowska et al. ''Sherpa: An open-source Python fitting package''. In: *Astrophys. J. Suppl. Ser.* 274.2 (Oct. 2024), p. 43.

[106] Paolo Nason. ''A New method for combining NLO QCD with shower Monte Carlo algorithms''. In: *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146.

[107] Stefano Frixione, Paolo Nason, and Carlo Oleari. ''Matching NLO QCD computations with parton shower simulations: the POWHEG method''. In: *Journal of High Energy Physics* 2007.11 (Nov. 2007), pp. 070–070. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2007/11/070. URL: http://dx.doi.org/10.1088/1126-6708/2007/11/070.

[108] Simone Alioli et al. ''A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX''. In: *Journal of High Energy Physics* 2010.6 (June 2010). ISSN: 1029-8479. DOI: 10.1007/jhep06(2010)043. URL: http://dx.doi.org/10.1007/JHEP06(2010)043.

[109] Leif Lönnblad. *Next Steps in Les Houches Event format (LHEF)*. https://indico.cern.ch/event/277733/contributions/1623403/attachments/506390/699108/CERN13a.pdf. 2013.

[110] J. Chapman et al. ''Challenges of the ATLAS Monte Carlo production during Run 1 and beyond''. In: *Journal of Physics: Conference Series* 513.3 (June 2014), p. 032029. DOI: 10.1088/1742-6596/513/3/032029. URL: https://dx.doi.org/10.1088/1742-6596/513/3/032029.

[111] S. Agostinelli et al. ''Geant4—a simulation toolkit''. en. In: *Nucl. Instrum. Methods Phys. Res. A* 506.3 (July 2003), pp. 250–303.

[112] ''AtlFast3: The Next Generation of Fast Simulation in ATLAS''. In: *Computing and Software for Big Science* 6.1 (Mar. 2022). ISSN: 2510-2044. DOI: 10.1007/s41781-021-00079-7. URL: http://dx.doi.org/10.1007/s41781-021-00079-7.

[113] A. Rimoldi et al. *The Simulation of the ATLAS Experiment: Present Status and Outlook*. Tech. rep. Geneva: CERN, 2005. DOI: 10.5170/CERN-2005-002.227. URL: https://cds.cern.ch/record/803313.

[114] James Catmore. *Introduction to ATLAS data preparation and analysis*. https://indico.cern.ch/event/472469/contributions/1982677/attachments/1220934/1785823/intro_slides.pdf. 2016.

[115] J. Alwall et al. ''The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations''. In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[116] M. Bähr et al. ‘‘Herwig++ physics and manual’’. In: *Eur. Phys. J. C* 58 (2008), p. 639. DOI: 10.1140/epjc/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph].

[117] Johannes Bellm et al. ‘‘Herwig 7.0/Herwig++ 3.0 release note’’. In: *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].

[118] Jon Butterworth et al. ‘‘PDF4LHC recommendations for LHC Run II’’. In: *J. Phys. G* 43 (2016), p. 023001. DOI: 10.1088/0954-3899/43/2/023001. arXiv: 1510.03865 [hep-ph].

[119] Barbara Jäger et al. ‘‘Parton-shower effects in Higgs production via vector-boson fusion’’. In: *The European Physical Journal C* 80.8 (Aug. 2020). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-020-8326-7. URL: http://dx.doi.org/10.1140/epjc/s10052-020-8326-7.

[120] Stefano Frixione, Paolo Nason, and Carlo Oleari. ‘‘Matching NLO QCD computations with parton shower simulations: the POWHEG method’’. In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[121] Simone Alioli et al. ‘‘A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX’’. In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[122] Torbjörn Sjöstrand et al. ‘‘An introduction to PYTHIA 8.2’’. In: *Comput. Phys. Commun.* 191 (2015), p. 159. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[123] Michal Czakon and Alexander Mitov. ‘‘Top++: A program for the calculation of the top-pair cross-section at hadron colliders’’. In: *Comput. Phys. Commun.* 185 (2014), p. 2930. DOI: 10.1016/j.cpc.2014.06.021. arXiv: 1112.5675 [hep-ph].

[124] D. J. Lange. ‘‘The EvtGen particle decay simulation package’’. In: *Nucl. Instrum. Meth. A* 462 (2001), p. 152. DOI: 10.1016/S0168-9002(01)00089-4.

[125] ATLAS Collaboration. *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model*. ATL-PHYS-PUB-2016-017. 2016. URL: https://cds.cern.ch/record/2206965.

[126] ''Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data''. In: *Journal of Instrumentation* 14.12 (Dec. 2019), P12006–P12006. ISSN: 1748-0221. DOI: 10.1088/1748-0221/14/12/p12006. URL: http://dx.doi.org/10.1088/1748-0221/14/12/P12006.

[127] ATLAS Collaboration. ''Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data''. In: *JINST* 14 (2019), P12006. DOI: 10.1088/1748-0221/14/12/P12006. arXiv: 1908.00005 [hep-ex].

[128] ''Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ s =13 TeV''. In: *The European Physical Journal C* 76.5 (May 2016). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4120-y. URL: http://dx.doi.org/10.1140/epjc/s10052-016-4120-y.

[129] ATLAS Collaboration. ''Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ = 13 TeV''. In: *Eur. Phys. J. C* 76 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[130] Ryan Atkin. ''Review of jet reconstruction algorithms''. In: *Journal of Physics: Conference Series* 645.1 (Sept. 2015), p. 012008. DOI: 10.1088/1742-6596/645/1/012008. URL: https://dx.doi.org/10.1088/1742-6596/645/1/012008.

[131] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. ''The anti-ktjet clustering algorithm''. In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), pp. 063–063. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2008/04/063. URL: http://dx.doi.org/10.1088/1126-6708/2008/04/063.

[132] ''Jet reconstruction and performance using particle flow with the ATLAS Detector. Jet reconstruction and performance using particle flow with the ATLAS Detector''. In: *Eur. Phys. J. C* 77.7 (2017), p. 466. DOI: 10.1140/epjc/s10052-017-5031-2. arXiv: 1703.10485. URL: https://cds.cern.ch/record/2257597.

[133] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. ''FastJet user manual''. In: *Eur. Phys. J. C* 72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].

[134] ATLAS Collaboration. *Tagging and suppression of pileup jets with the ATLAS detector*. ATLAS-CONF-2014-018. 2014. URL: https://cds.cern.ch/record/1700870.

[135] ''ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV''. In: *The European Physical Journal C* 79.11 (Nov. 2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7450-8. URL: http://dx.doi.org/10.1140/epjc/s10052-019-7450-8.

[136] ''ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset''. In: *The European Physical Journal C* 83.7 (July 2023). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-023-11699-1. URL: http://dx.doi.org/10.1140/epjc/s10052-023-11699-1.

[137] ATLAS Collaboration. ''ATLAS flavour-tagging algorithms for the LHC Run 2 *pp* collision dataset''. In: *Eur. Phys. J. C* 83 (2023), p. 681. DOI: 10.1140/epjc/s10052-023-11699-1. arXiv: 2211.16345 [physics.data-an].

[138] Tatsuya Masubuchi et al. *Object selections for analyses involving a Higgs boson produced in association with a vector boson in which $h \to b\bar{b}$ and V decays leptonically with Run-2 data collected with the ATLAS detector at $\sqrt{s} = 13$ TeV*. Tech. rep. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2151844.

[139] ''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. DOI: 10.1016/j.physletb.2018.09.013. URL: http://dx.doi.org/10.1016/j.physletb.2018.09.013.

[140] Zhijun Liang et al. *Search for Higgs Boson Production via Weak Boson Fusion in Association with a High-Energy Photon*. Tech. rep. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2234295.

[141] Pekka Sinervo. *Hard Scattering in Hadron-Hadron Collisions: Physics and Anatomy*. https://hep.physics.utoronto.ca/PekkaSinervo/phy2407s/PHY2407Introduction1.pdf.

[142] Tetiana Hryn'ova and Fernando Monticelli. *Electron and photon trigger efficiencies measured on early 2018 data for LHCC*. Tech. rep. Geneva: CERN, 2018. URL: https://cds.cern.ch/record/2318523.

[143] ATLAS Collaboration. ''ATLAS *b*-jet identification performance and efficiency measurement with $t\bar{t}$ events in *pp* collisions at $\sqrt{s} = 13$ TeV''. In: *Eur. Phys. J. C* 79 (2019), p. 970. DOI: 10.1140/epjc/s10052-019-7450-8. arXiv: 1907.05120 [hep-ex].

[144] 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.

[145] ATLAS Collaboration. *Calibration of light-flavour b-jet mistagging rates using ATLAS proton–proton collision data at* $\sqrt{s}$ = 13 *TeV*. ATLAS-CONF-2018-006. 2018. URL: https://cds.cern.ch/record/2314418.

[146] Han Cui et al. *Search for Higgs Boson Production via Weak Boson Fusion in Association with a High-Energy Photon at* $\sqrt{s}$ = *13 TeV with the ATLAS detector*. Tech. rep. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2673667.

[147] ATLAS Collaboration. *Flavor Tagging Efficiency Parametrisations with Graph Neural Networks*. ATL-PHYS-PUB-2022-041. 2022. URL: https://cds.cern.ch/record/2825433.

[148] Sebastian Ruder. *An overview of gradient descent optimization algorithms*. 2017. arXiv: 1609.04747 [cs.LG]. URL: https://arxiv.org/abs/1609.04747.

[149] Georgia Karagiorgi et al. *Machine Learning in the Search for New Fundamental Physics*. 2021. arXiv: 2112.03769 [hep-ph]. URL: https://arxiv.org/abs/2112.03769.

[150] Sunil Ray. *Introduction to Boosting Algorithms in Machine Learning — analyticsvidhya.com*. https://www.analyticsvidhya.com/blog/2015/11/quick-introduction-boosting-algorithms-machine-learning/. 2024.

[151] Andreas Hocker et al. ''TMVA - Toolkit for Multivariate Data Analysis''. In: (Mar. 2007). arXiv: physics/0703039.

[152] Jan Therhaag. ''TMVA Toolkit for multivariate data analysis in ROOT''. In: *PoS* ICHEP2010 (2010). Ed. by Bernard Pire et al., p. 510. DOI: 10.22323/1.120.0510.

[153] François Chollet et al. *Keras*. https://keras.io. 2015.

[154] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015. URL: https://www.tensorflow.org/.

[155] Emanuele Bagnaschi et al. ''An extensive survey of the estimation of uncertainties from missing higher orders in perturbative calculations''. In: *Journal of High Energy Physics* 2015.2 (Feb. 2015). ISSN: 1029-8479. DOI: 10.1007/jhep02(2015)133. URL: http://dx.doi.org/10.1007/JHEP02(2015)133.

[156] ATLAS Physics Modeling Group. *Scale Uncertainties*. URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PmgSystematicUncertaintyRecipes#Scale_uncertainties.

[157] Michal Czakon and Alexander Mitov. ''NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels''. In: *JHEP* 12 (2012), p. 054. DOI: 10.1007/JHEP12(2012)054. arXiv: 1207.0236 [hep-ph].

[158] Michal Czakon and Alexander Mitov. ''NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction''. In: *JHEP* 01 (2013), p. 080. DOI: 10.1007/JHEP01(2013)080. arXiv: 1210.6832 [hep-ph].

[159] A. D. Martin et al. ''Uncertainties on α_S in global PDF analyses and implications for predicted hadronic cross sections''. In: *Eur. Phys. J. C* 64 (2009), pp. 653–680. DOI: 10.1140/epjc/s10052-009-1164-2. arXiv: 0905.3531 [hep-ph].

[160] Andy Buckley et al. ''LHAPDF6: parton density access in the LHC precision era''. In: *The European Physical Journal C* 75.3 (2015), p. 132. DOI: 10.1140/epjc/s10052-015-3318-8. URL: https://doi.org/10.1140/epjc/s10052-015-3318-8.

[161] *LHAPDF sets*. URL: https://lhapdfsets.web.cern.ch/current/PDF4LHC21_40_pdfas/PDF4LHC21_40_pdfas.info.

[162] ATLAS Collaboration. *ATLAS delivers most precise luminosity measurement at the LHC*. https://home.cern/news/news/physics/atlas-delivers-most-precise-luminosity-measurement-lhc.

[163] ''Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC''. In: *The European Physical Journal C* 83.10 (Oct. 2023). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-023-11747-w. URL: http://dx.doi.org/10.1140/epjc/s10052-023-11747-w.

[164] ''Electron and photon efficiencies in LHC Run 2 with the ATLAS experiment''. In: *Journal of High Energy Physics* 2024.5 (May 2024). ISSN: 1029-8479. DOI: 10.1007/jhep05(2024)162. URL: http://dx.doi.org/10.1007/JHEP05(2024)162.

[165] *Small-R (R=0.4) jet energy scale and jet energy resolution uncertainties for full run 2 analyses (release 21)*. https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetUncertaintiesRel21Summer2018SmallR.

[166] ''Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector''. In: *The European Physical Journal C* 81.8 (2023). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-021-09402-3. URL: http://dx.doi.org/10.1140/epjc/s10052-021-09402-3.

[167] Maximilian Swiatlowski. *FullJER Scheme*. https://indico.cern.ch/event/1051376/contributions/4420484/attachments/2270608/3856133/jer.pdf. 2021.

[168] ATLAS. *Tagger Calibration Recommendations for Release 21*. https://twiki.cern.ch/twiki/bin/view/AtlasProtected/BTagCalib2017#Recommendation_October_2018.

[169] Kyle Cranmer et al. ''HistFactory: A tool for creating statistical models for use with RooFit and RooStats''. In: (June 2012).

[170] Wouter Verkerke and David Kirkby. *The RooFit toolkit for data modeling*. 2003. arXiv: physics/0306116 [physics.data-an].

[171] Wolfgang A. Rolke, Angel M. López, and Jan Conrad. ''Limits and confidence intervals in the presence of nuisance parameters''. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 551.2 (2005), pp. 493–503. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2005.05.068. URL: https://www.sciencedirect.com/science/article/pii/S016890020501291X.