October 2021

SEARCH FOR RESONANT PAIR PRODUCTION OF HIGGS BOSONS IN THE FOUR B-QUARK FINAL STATE WITH THE ATLAS DETECTOR

Dale Abbott
University of Massachusetts Amherst

SEARCH FOR RESONANT PAIR PRODUCTION OF HIGGS BOSONS IN THE FOUR B-QUARK FINAL STATE WITH THE ATLAS DETECTOR

A Dissertation Presented

by

DALE ABBOTT

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2021

Physics

SEARCH FOR RESONANT PAIR PRODUCTION OF HIGGS BOSONS IN THE FOUR B-QUARK FINAL STATE WITH THE ATLAS DETECTOR

A Dissertation Presented

by

DALE ABBOTT

Approved as to style and content by:

Stephane Willocq, Chair

Michael Ramsey-Musolf, Member

Verena Martinez Outschoorn, Member

Grant Wilson, Member

Narayanan Menon, Chair of the Faculty
Physics

DEDICATION

Dedicated to all those who made sacrifices to make my journey possible.

ACKNOWLEDGMENTS

From a perplexingly young age, I had ideas of existentialism and physics that would leave me confused and sleepless in my bed at night. Had it not been for the many people who entertained my naive yet introspective questions throughout my career in science, I would have faced a barrier of entry too high for me to overcome. In other words, I would not been able to do it without them.

First, I need to thank my mother and father, who realized my interest and facilitated it. From letting me dress up as Isaac Newton to give my book report in second grade or buying me a subscription to popular science in high school, they always found ways to sate my thirst for science or engineering.

While in high school, I found myself in Alan Kramer's physics class. I am forever grateful to Dr. Kramer who took time after school to explain what a career in physics would be like and ultimately inspired me to go on to study physics in college. During my upperclassman years at the University of Delaware, I met biophysics research Dr. Ed Lyman. Dr. Lyman broke the banality of the typical undergraduate experience of grinding out math equations on a chalk board, by allowing me into his research group. My love for research was found while simulating soft, interacting colloids on a two dimensional plane. Thank you to Dr. Lyman for igniting my passion, guiding me towards graduate school, and for maintaining an amazing working environment and team, which you allowed me to be part of.

I would like to thank Heath Hatch, for his encouragement, kindness, and guidance during my entry to the University of Massachusetts. Heath let me grow as a teacher and budding researcher, through the countless lectures and discussions.

There are so many people to thank who are part of the ATLAS experiment and CERN that it may be impossible for me to mention everyone but it will not keep me from trying nonetheless. First and foremost, I need to thank my research advisor Stephane Willocq for sharing his boundless knowledge of physics and experimental techniques with me. Stephane welcomed me into the UMass ATLAS group and gave me the opportunity to contribute to the greatest experimental endeavor by humankind to date. I believe Dr. Willocq brought out the best in me by letting me grow in my own way and guiding me when I needed it, and my life is forever changed because of it.

I would also like to thank some of the many other professors I worked with at UMass including Carlo Dallapiccola, Rafael Coelho Lopes de Sá, Verena Martinez Outschoorn, and Ben Brau. Dr. Dallapiccola always offered helpful advice and provided a comfortable, friendly environment for me to share questions about navigating graduate school and research. I would also like to thank my post-doc advisor Attilio Picazio. Attilio was a source of positivity and pragmatism for me in a chaotic world of research, presentations, and career building. I would also like to thank Peter Tornambe and Nora Pettersson who have helped me throughout my time at CERN.

There are countless employees, graduate students, and other friends I would like to thank as well. At UMass ATLAS Nathan Bernard, Zac Meadows, Guy Rosin, Margaret Lutz, Jackson Burzynski, Yuan-Tang Chou, Sam Krishnamurthy, Jay Sandesera, Makayla Vessella, and Cooper Wagner for all the helpful and fun discussions over the years. At CERN Huacheng Cai, Sean Gasiorowski, Nicole Hartman, Lucas Santiago Borgna, and Alex Emerman for all the collaborating and comradery during my time near and far. In the group of combined analysis editors not yet mentioned like Max Swiatlowski, Rafael Teixeira De Lima, Jana Schaarschmidt, Bill Balunas, Beojan Stanislaus, and Todd Huffman who poured countless hours into the analysis and were ideal exemplars of what a scientist should be. At UMass

Jake Shechter, Sara Feyzbakhsh, Leila Farhadi, and Mike Roberts for all the late night studying and long hours of grading together. The UMass administration Jane Knapp, Katie Bryant, Sara Cooper, Raquel Kirpan, and Narayanan Menon for being the unsung heroes of the graduate students. My CERN engineering friends Tuukka Lehtinen, Oto Kemularia, and Diego Zambelli for the amazing new perspectives they shared with me. At Amherst Dane Napoli and Galaxie Story for sharing their scientific knowledge of their unique fields of study with me. Lastly, Monica Coniglio for her relentless love and support during the many trying times over the course of my studies.

ABSTRACT

SEARCH FOR RESONANT PAIR PRODUCTION OF HIGGS BOSONS IN THE FOUR B-QUARK FINAL STATE WITH THE ATLAS DETECTOR

SEPTEMBER 2021

DALE ABBOTT, B.Sc, UNIVERSITY OF DELAWARE

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Stephane Willocq

A search for heavy resonances decaying to Higgs boson pairs in the $b\bar{b}b\bar{b}$ final state with 139 fb^{-1} of LHC proton-proton (pp) collision data collected at $\sqrt{s} = 13$ TeV obtained by the ATLAS detector during the 2015-2018 time period. Two benchmark signals are studied, both of which decay into a Higgs boson pair: a spin-2 Kaluza-Klein Graviton and a scalar resonance. This thesis studies high transverse momentum (boosted) resonances, which range from 900-5000 GeV. No statistically significant excesses are observed and the results are found to be compatible with the Standard Model. Upper limits are set on the production of Higgs boson pairs to $b\bar{b}b\bar{b}$ for each respective model.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

CHAPTER 1

INTRODUCTION

The focus of this thesis is on the search for rare Higgs pair events that decay into boosted bottom quark pairs, with the ATLAS experiment at CERN. First, I will provide some details about the theory that is the foundation of particle physics in Section 2. Then I will explain some more theory that motivates our experimental searches and offer a brief history of the experiment to date. I will end the section by giving a general outline of the approach and structure of the analysis.

With the mathematical and historical background laid out, I will go on to explain the hardware that makes up the ATLAS detector as well as an overview of the LHC at CERN in Section 3. Here I will give some details on hardware performance as well as the relevance of each component to physics. In Section 4 I will explain what data collected from the ATLAS detector is used in this analysis as well as how simulated physics is used. In Section 5 I will explain how physics objects used in the analysis are reconstructed from raw, detector level readouts. This will provide an overview of the software and algorithms specialized to do this task.

With the physics objects defined and our data and simulated samples collected and ready to use, we begin the highly specialized part of the analysis that looks for the rare Higgs pair events decaying into four bottom quarks. First, I outline the event selection in Section 6. What remains from the event selection is used to perform a background estimation, which is the most essential part of the analysis. The background estimation is partially data-driven and partially simulation-driven and is outlined in Section 7. Prior to the statistical analysis, sources of uncertainty in the analysis are quantified in Section 8. The background estimation provides a set of distributions of the mass of the Higgs boson pair, which is then used to perform a statistical analysis in Section 9. The section is finalized with our upper limits on the cross section as well as a limit comparison to other similar analyses.

CHAPTER 2

THEORY

The Standard Model (SM) is a robust theory that describes the elementary particles that make up the universe. Figure 2.1 shows the SM particles and their respective properties visualized. Particles consist of fermions with half-integer spin and bosons with whole-integer spin. Quarks and leptons are types of fermions. Protons, which are accelerated at the Large Hadron Collider ("LHC"), are made up of quarks. Leptons come in two varieties: charged and neutral. The charged leptons are electrons (e), muons (μ), and tau-leptons (τ), the latter two of which are in many ways heavier analogs to the electron. Each charged lepton is associated to an uncharged, weakly interacting neutrino, which has small mass.

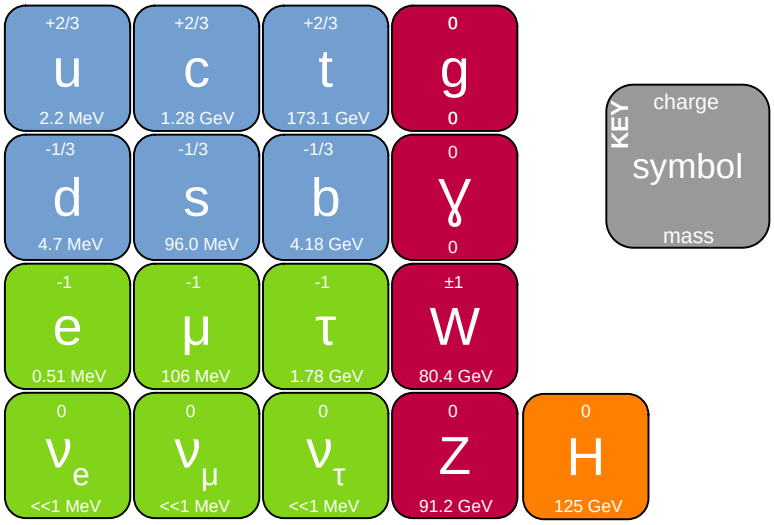


Figure 2.1. The particles of the Standard Model with their respective charge and mass are shown above and below the particle symbol. The quarks, leptons, vector bosons, and Higgs boson are shown in blue, lime green, magenta, and orange respectively.

The SM theory is founded upon Quantum Field theory (QFT), which describes elementary particles as fields and uses Lagrangian Densities ("$\mathcal{L}$") to describe the interactions between the elementary particles. The $W^{\pm}$, Z, gluon (g) and γ bosons are the force carriers or propagators for the weak ($W^{\pm}$ and Z), strong (g), and elec-

tromagnetic (γ) interactions. The five gauge bosons all have spin-1 and the gluon and photon are massless.

2.1 Gauge Theory and Massless Bosons

The use of the term gauge in physics has its origins in the early 1900s with the scientist Herman Weyl. The scientist supposedly coined the term "gauge" in reference to a measuring device used during the installation of railroads. The railroads were not dependent on which gauge was used (e.g. unit of measurement such as meters or feet) but only on that the measurements be consistent throughout the railroad construction. Hence the term, *gauge invariance* is coined and in physics, gauge invariance is meant to equilibrate different frames of reference despite the differences in formalisms or methods used to describe an object in each frame. That is, laws of physics should hold true regardless of where they are being measured or the formalisms employed to describe the physics. For example, a gluon is a superposition of color charges. The naming convention of red, green and blue is totally arbitrary, but the physics remains the same whether the gluon be in Geneva, Switzerland or in Amherst, Massachusetts.

An Informal Description of Gauge Invariance: The idea of gauge invariance is the cornerstone of particle physics and to really appreciate this, a philosophical description is helpful. One may naively expect the broad conceptual idea of limitations or rules to be, by definition, limiting. Actually, limitations are the grounds for any existence in this universe. For example, the board game Chess only exists by having limitations. We provide each piece a set of rules that it must obey, from which the competitive game is derived. This important game analogy can be applied to life and physics. Life has limitations that, like chess, bring it to existence. Likewise, for physics, all particles have a set of rules that they obey. The rules of physics are described by mathematical descriptions in QFT, which are often built from a quantized gauge symmetry (i.e. having gauge invariance). The limitations of each particle, make

it unique and distinguishable from the other particles, which is incredibly powerful in a realm of abstraction that is particle physics.

Although the formalism of gauge invariance was not canonized by Weyl until 1918, gauge fields and local gauge symmetries were found in physics well before that. Maxwell described electromagnetism with a U(1) symmetry,

$$\Psi(x) \rightarrow e^{i\theta(x)}\Psi(x), \tag{2.1}$$

which implies local gauge symmetry for a electromagnetic field $\Psi(x)$ about any arbitrary local rotation $\theta(x)$. To relate different reference frames, which may undergo this transformation, we use gauge fields. These gauge fields give rise to "forces" (e.g. electromagnetic, strong, weak, and gravitational), which in Maxwell's case was the electromagnetic force.

In 1954, Chen Ning Yang and Robert Mills expanded this idea with higher order transformation unitary "Lie groups". The groups are described by an $n \times n$ matrix that may transform a field like the previously stated example of the electromagnetic force, which is based on the unitary group U(1). There exists a special subgroup of these unitary groups, such that the order or sequence of operations is relevant (non-commutable). Such physics that corresponds to this is called *non-abelian*. The non-commutability gives rise to a new coupling constant that represents a gauge boson coupling to itself by a "charge". This coupling is zero for U(1) (commutable), which implies photons do not couple to other photons. The non-abelian physics is based on special unitary groups noted by "SU(n)", which represent an order dependent operation and where n is the number of degrees of freedom. Non-abelian gauge fields, unlike the photon, may couple to themselves.

Quarks make up six of the particles in the standard model seen in the blue part of Figure 2.1. Unlike charged leptons which only have electric charge, a quark may have three color-charges: red, blue, and green. The QFT that describes the physics

of quarks and their color gauge fields is called Quantum Chromodynamics (QCD). QCD derives from the SU(3) color symmetry group, where $n = 3$ for the three color flavors, and has a gauge field that corresponds to the strong force. The gluon is the propagator of the strong force and comes in eight varieties. The gluons are represented as a mix of two of the three color charges. The number of gauge bosons for any given unitary group is $n^2 - 1$, thus for QCD $n = 3$ and $3^2 - 1 = 8$.

QFT describes a system as a collection of fields at any point in space, where the lowest energy state of a given field is the vacuum state. An excitation to the field explains the existence of particles in the physical world. For the QCD and EM gauge fields, excitations are considered gluons and photons respectively. In general, mass terms arise in the Lagrangian density as $m^2\phi^2$. To preserve gauge invariance, the Lagrangian densities seen for photons and gluons are absent of such mass terms. Thus, gluons and photons as described in QED and QCD are massless.

2.2 Electroweak Theory and Symmetry Breaking

The weak gauge field or weak force is best described by the electroweak theory (EWT), which combines the electromagnetic force and weak force into one theory. The theory purports that in certain conditions, such as those seen in the early universe, the EM and weak forces are combined into one. This is not the case today, where the weak and EM force are separate. The EWT has a gauge symmetry of $SU(2) \times U(1)$.

The unique feature of EWT is that its symmetry is (spontaneously) broken by the Higgs boson field. This is called "spontaneous symmetry breaking" and is essential in many aspects of the SM, especially the Higgs, W, and Z bosons. To begin, the Lagrangian for transformations belonging to the U(1) symmetry group is,

$$\mathcal{L} = -F^{\mu\nu}F_{\mu\nu} + (D_\mu\phi)^*(D^\mu\phi) - \mu^2\phi^*\phi - \lambda(\phi^*\phi)^2, \tag{2.2}$$

where D_μ is the covariant derivative and $F^{\mu\nu}$ is the field strength tensor. This Lagrangian is invariant under the U(1) symmetry,

$$\phi(x) \to e^{i\alpha}\phi(x). \tag{2.3}$$

The potential part of the equations is

$$V(\phi) = \mu^2\phi^*\phi + \lambda(\phi^*\phi)^2, \tag{2.4}$$

which is minimized differently depending on the value of μ. For $\mu^2 > 0$, the potential is parabolic, with one minimum at $\phi = 0$. For $\mu^2 < 0$ however, the potential takes the shape of a sombrero ("Mexican hat"), both of which are shown in Figure 2.2. From the shape of the Mexican hat potential, it is apparent that the value that minimizes the potential is degenerate. This degeneracy breaks gauge symmetry as there exist multiple states that produce the minimum (vacuum) energy.

Solving for the vacuum energy by minimizing the potential,

$$v = \sqrt{\frac{-\mu^2}{2\lambda}}. \tag{2.5}$$

Then setting up a perturbative expansion about the vacuum state, so that

$$\phi(x) = (v + \eta(x))e^{i\xi(x)}, \tag{2.6}$$

where $\eta(x)$ and $\xi(x)$ are real fields and $\phi(x)$ is complex. Using the perturbative expansion in our original potential, the Lagrangian becomes

$$\mathcal{L} = -F^{\mu\nu}F_{\mu\nu} + \partial_\mu\rho\partial^\mu\rho + v^2\partial_\mu\xi\partial^\mu\xi - 4\lambda v^2\rho^2 + q^2v^2A_\mu^2 - 2qv^2\partial_\mu\xi A^\mu + ..., \tag{2.7}$$

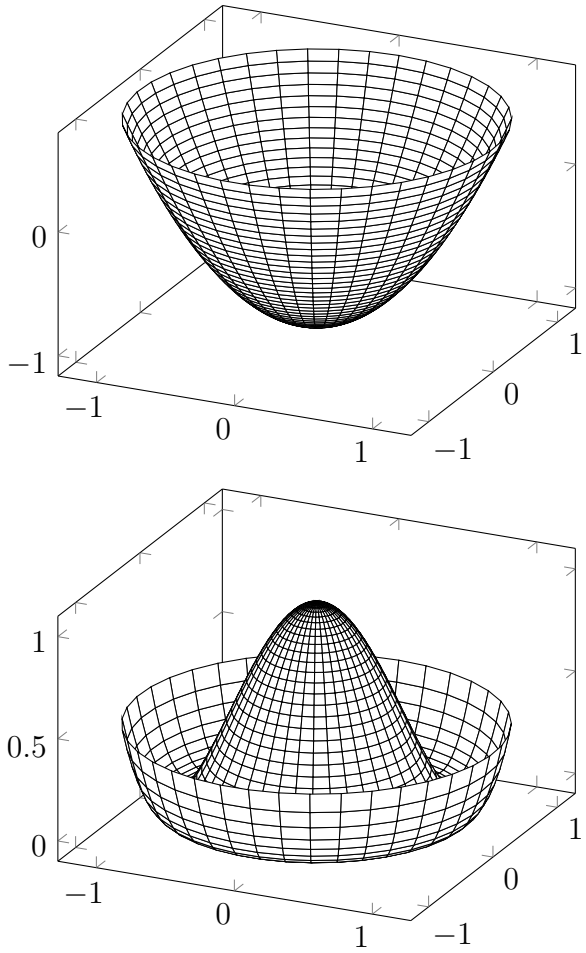


Figure 2.2. The two possible shapes for the potential in Eq. 2.4. The $\mu^2 > 0$ and $\mu^2 < 0$ scenarios are shown on the top and bottom respectively.

where A_μ is the spin-0 field, which corresponds to the photon or QED. The $q^2v^2A_\mu^2$ term in Equation 2.7 would give mass to the photon but recall that the gauge invariance for the U(1) photon, which may undergo a gauge transformation,

$$A_\mu \to A_\mu - \frac{1}{qv}\partial_\mu.\xi \tag{2.8}$$

The ξ component of the Higgs field has no mass term but instead corresponds the choice of gauge for a given system. The Higgs Mechanism is the introduction of broken gauge symmetry by a non-zero vacuum expectation value.[37] To extend our

understanding of the Higgs Mechanism to SU(2)×U(1), the covariant derivative and SU(2) Higgs doublet are defined as

$$D_\mu = \partial_\mu + \frac{i}{2} g \vec{\tau} \vec{W}_\mu + \frac{ig'}{2} B_\mu, \tag{2.9}$$

and

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{2.10}$$

where W_μ and B_μ represent massive and massless bosons, respectively. The complex scalar fields ϕ^+ and ϕ^0 may be represented as

$$\phi^+ = \frac{\phi_1 + i\phi_2}{\sqrt{2}} \tag{2.11}$$

and

$$\phi^0 = \frac{\phi_3 + i\phi_4}{\sqrt{2}}. \tag{2.12}$$

For SU(2)×U(1), the Lagrangian is

$$\mathcal{L} = -F^{\mu\nu} F_{\mu\nu} + (D_\mu \Phi)^* (D^\mu \Phi) - \mu^2 \Phi^* \Phi - \lambda (\Phi^* \Phi)^2. \tag{2.13}$$

The Lagrangian above is invariant under the gauge transformation

$$\Phi = \exp\left(\frac{ig}{2} \vec{\tau} \dot{\Delta} - \frac{ig'}{2} \Lambda \right) \Phi, \tag{2.14}$$

where Δ and Λ represent local gauge transformations in SU(2) and U(1) respectively and τ are the Pauli matrices. Minimizing the potential of the Lagrangian, three of the four components of the Higgs doublet are chosen to be zero. That is, $\phi_1 = 0$,

$\phi_2 = 0$, and $\phi_4 = 0$, which leaves a non-zero ϕ_3 which is minimized by the vacuum energy,

$$v = \sqrt{\frac{\mu^2}{\lambda}}. \tag{2.15}$$

The field Φ can be expanded about this vacuum as,

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ H + v \end{pmatrix}, \tag{2.16}$$

where H represents the Higgs field. When a local transformation is then applied to this equation, the three fields will arise that "eat" the ϕ_1, ϕ_2, and ϕ_4 from before. Similar to the U(1) case shown before, symmetry is broken by a non-zero expectation value for the vacuum state.[44]

In SU(2)×U(1) the field strength tensor becomes,

$$F^i_{\mu\nu} = \partial_\mu A^i_\mu - \partial_\mu A^i_\mu \rightarrow \partial_\mu A^i_\mu - \partial_\mu A^i_\mu + q\epsilon^{ijk} A^j_\mu A^i_\nu, \tag{2.17}$$

which is used for the Yang-Mills theory. Using these in the Lagrangian along with the Higgs doublet we find a combined Lagrangian that describes how the Higgs field interacts with the various other particles in the SM. The full Lagrangian density is broken into parts as,

$$\mathcal{L}_{EW} = \mathcal{L}_{Kinetic} + \mathcal{L}_{NC} + \mathcal{L}_{CC} + \mathcal{L}_H + \mathcal{L}_{HV} + \mathcal{L}_{WWV} + \mathcal{L}_{WWVV} + \mathcal{L}_{Yukawa}, \tag{2.18}$$

where $\mathcal{L}_{NC}$ is for the neutral current, $\mathcal{L}_{CC}$ is for the charged current, $\mathcal{L}_H$ is for the Higgs, $\mathcal{L}_{HV}$ is for the Higgs interacting with gauge vector bosons, $\mathcal{L}_{WWV}$ is for three-point gauge boson interactions, and the $\mathcal{L}_{WWVV}$ is for four-point gauge boson interactions.

The Lagrangian written in simple terms for the parts most relevant to this analysis is shown in Equation 2.19, with a full description of all the terms in the Lagrangian found in Reference [46]. The quadratic Higgs term $\frac{m_H^2}{2}H^2$ is a Higgs mass (m_H) term and the cubic Higgs $\frac{m_H^2}{2v}H^3$ is known as the Higgs self coupling. The $\frac{2m_W^2}{v}W_\mu^+W^{-\mu}H$ and $\frac{m_Z^2}{v}Z^\mu Z_\mu H$ terms couple the W and Z bosons to the Higgs boson by mass. The $\frac{g}{\sqrt{2}}\overline{u^i}(V_{CKM})_{ij}\gamma_\mu W^+ d^i$ term couples the W boson to quarks with the help of the Cabibbo-Kobayashi-Maskawa (CKM) matrix V_{CKM}, which quantifies the strength of weak interactions between the different quarks. The CKM matrix is a unitary matrix and is represented as a 3 by 3 matrix for the three generations of quarks in the SM. The $\frac{H}{v}\sum_f m_f\overline{f}f$ term couples the Higgs boson to fermions by the fermion mass. High order (h.o.) and hermitian conjugate terms (h.c.) are not shown here but may be found in Reference [46].

$$\begin{aligned}\mathcal{L} = \mathcal{L}_{Kinetic} &+ \frac{m_H^2}{2}H^2 + \frac{m_H^2}{2v}H^3 + \frac{2m_W^2}{v}W_\mu^+W^{-\mu}H + \frac{m_Z^2}{v}Z^\mu Z_\mu H \\ &+ \frac{g}{\sqrt{2}}\overline{u^i}(V_{CKM})_{ij}\gamma_\mu W^+ d^i - \frac{H}{v}\sum_f m_f\overline{f}f \\ &+ \texttt{h.o.and h.c.terms}\end{aligned} \tag{2.19}$$

With the Higgs field introduced to the potential that results in symmetry breaking in EWT, mass for the W and Z bosons become apparent. The photon does not break symmetry when gauge transformed, and thus remains massless like the gluon. The Higgs boson does not couple to the massless bosons at leading order, but does couple via virtual $t\bar{t}$ loops for the gluon and virtual W^+W^- pairs and virtual $t\bar{t}$ loops for the photon.

2.3 Higgs Boson Phenomenology

The Higgs boson is an excitation to the Higgs field. The first experimental observation of the Higgs boson was found at ATLAS and it's sister experiment CMS in 2012, with a mass $m_H = 125$ GeV.[7] Further measurements of the Higgs boson mass were published in 2018, which gives more confirmation to the mass measurement.[29] With QFT, QCD, and EWT quantifying the interactions in the SM, there are certain important takeaways when considering the Higgs boson. For example, the Higgs boson notably couples to fermions by their mass and to gauge bosons by the gauge coupling (g and g' for SU(2) and U(1) respectively) and the vacuum expectation value ("v"). With this understanding and *ceteris paribus* the Higgs boson should most commonly decay into top quarks, but the decay products are also affected by the Higgs boson mass. That is, conservation laws make it impossible for a lighter particle to decay into a heavier one. For example, a Higgs boson (125 GeV) decaying into top quarks (172 GeV) is impossible, but top quarks decaying into a Higgs boson is likely. The latter may be done through top loops decaying to a Higgs boson or a Higgs boson with Z boson or $t\bar{t}$ pair. Also, a rare decay channel of $t\bar{t}H$ is available to provide precise measurements of the top quark-Higgs boson Yukawa coupling.

The ratio of number of occurrences for a specific decay channel divided by the total occurrences of all possible decay channels is known as the "branching ratio" ("BR"). Figure 2.3 shows the BR of the various Higgs boson decays as a function of Higgs boson mass. If the Higgs boson has higher mass, it more easily decays into heavier particles like the top quark, or W or Z boson. The combined mass of W and Z boson pairs or top quark pairs are higher than that of the mass of the Higgs boson therefore lowering the likelihood of seeing these as a decay product of the Higgs. Table 2.1 provides exact values of BRs for the common Higgs boson decays.[39] [40]

It is apparent that Higgs bosons are most likely to decay to bottom quark pairs and therefore this decay mode presents itself as an opportune decay channel to study.

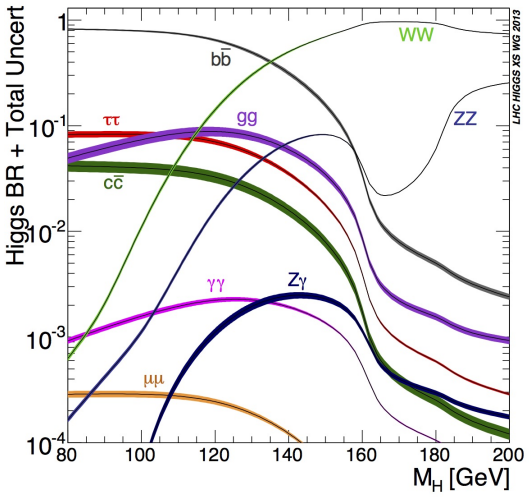


Figure 2.3. Higgs boson BR as a function of Higgs boson mass.

Decay Channel	Branching Ratio (%)
bb	$58.24^{+0.65}_{-0.65}$
WW	$21.4^{+0.99}_{-0.99}$
gg	$8.19^{+3.4}_{-3.4}$
$\tau\tau$	$6.27^{+1.17}_{-1.16}$
cc	$2.89^{+1.2}_{-1.2}$
ZZ	$2.62^{+0.99}_{-0.99}$
$\gamma\gamma$	$0.227^{+1.73}_{-1.72}$
$Z\gamma$	$0.153^{+5.7}_{-5.7}$
$\mu\mu$	$0.0212^{+1.23}_{-1.23}$

Table 2.1. Table of the most probably decays of the Higgs boson ordered by branching ratio (for Higgs boson mass of 125 GeV).

For Higgs pair events ("diHiggs" events), decay likelihood is multiplicative. The most common decay is $HH \to bbbb$, followed by $HH \to bbWW$. The full list of probabilities for each diHiggs decay channel is depicted in Figure 2.4.

The leading order Higgs boson production mode is gluon-gluon fusion ("ggf") and the subleading is vector boson fusion ("vbf"). The Feynman diagram for single Higgs boson ggf production is shown in Figure 2.5. The ggf production is the dominant production mode for diHiggs events as well.[38][39][40] The two leading order Feynman diagrams for diHiggs events in the ggf production mode are shown in Figure 2.6.

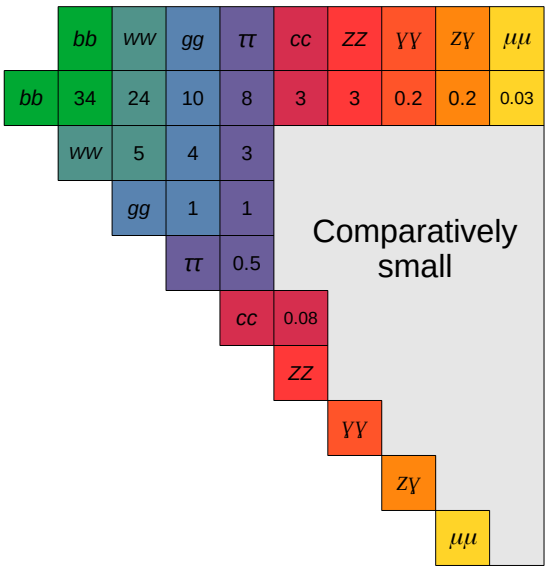


Figure 2.4. Probabilities for diHiggs events as percentages. The area marked "comparatively small" is for the channels that individually have a branching ratio less than 0.01%.

They are known as the "box" and "triangle" diagrams. The triangle diagram follows a triangle loop that emits an off-shell Higgs bosons ($X = H^*$), which then decays to two on-shell Higgs bosons. Through this process the Higgs self coupling may be tested. The box and triangle diagrams interfere destructively, which partially suppresses the ggf production mode.

2.4 BSM diHiggs Resonant Production

The term "resonance" is used in this analysis for a heavy particle that will appear as a bump or peak in the invariant diHiggs mass distribution – the final descriminant in this analysis. The resonance is expected to have a mass at least twice that of the SM Higgs boson, or $m_{resonance} > 2 \times m_{H_{SM}} = 250$ GeV. This helps define the lower bound of a resonant search, which starts at 251 GeV. Higher resonant masses will result in higher p_T constituents, which are also described as "boosted" or as having "boost". The resonant masses are not known prior to discovery, so a range of expected values

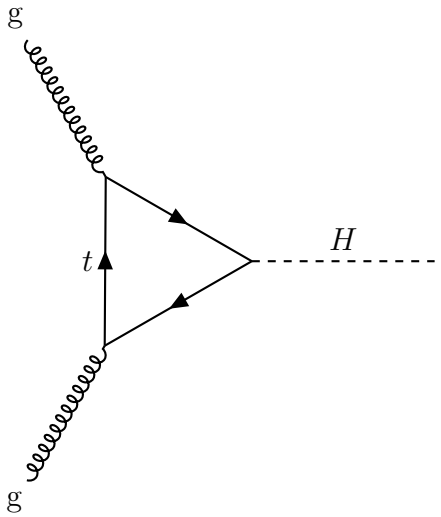


Figure 2.5. Feynman diagram for the leading order ggf production of Higgs Bosons.

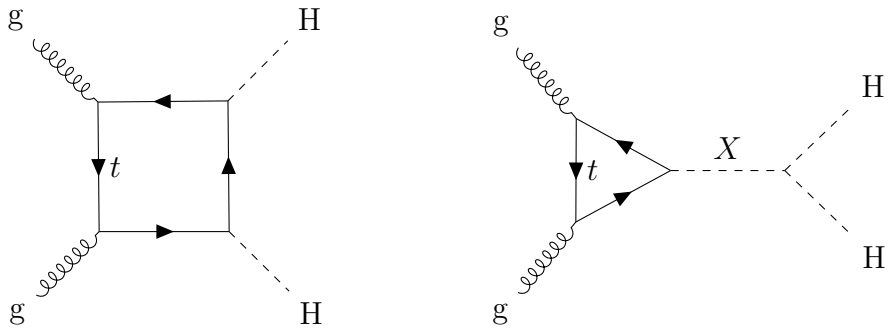


Figure 2.6. The two leading order Feynman diagrams for diHiggs ggf production. The box and triangle diagrams are on the left and right respectively.

are tested. The upper bound is theoretically limited by the energy of the colliding proton beams made by the LHC. For this analysis simulated samples are provided up to 5 TeV, because the number of events at the full `Run-2` luminosity expected to be seen at resonant masses higher than 5 TeV is small.

Three types of resonances are considered: spin-0, spin-1 and spin-2. Vector bosons (spin-1) are disallowed by the Landau-Yang theorem to decay into scalar bosons (spin-0) such as the Higgs boson, and thus are ruled out of three possibilities for diHiggs production. The scalar bosons are known as the "Spin-0" resonances in this analysis. The Spin-0 is model independent but well motivated by theories such as the two Higgs Doublet Model ("2HDM").[32] The 2HDM expands on the Higgs doublet seen in the SM SU(2) description by introducing two Higgs doublets. The Minimal Supersymmetric Standard Model ("MSSM") is a 2HDM, which theorizes several different particles including a "Heavy Scalar" that may decay into two Higgs Bosons. The Heavy Scalar may serve as a benchmark at times in this analysis.

Spin-2 bosons are exemplified in this analysis by the Kaluza-Klein Graviton ($G^*_{\mathtt{KK}}$), which is featured in the bulk Randall-Sundrum (RS) model.[45] The bulk RS model utilizes warped extra spacial dimensions ("WED") to explain why gravity is a force many orders of magnitude weaker than the Weak force (also known as the "hierarchy problem"). There is a free parameter for the bulk RS model k, which is meant to represent the curvature of the WED. Often, the free parameter is normalized by the Planck scale as $c \equiv k/\overline{M}_{Pl}$. When using $G^*_{\mathtt{KK}}$ samples in this analysis, the constant value is chosen for the free parameter as $c = 1.0$.

2.5 Previous Studies

The Higgs boson has a very short lifetime and therefore is never directly observed but rather reconstructed through it's decay products. This thesis is focused on the boosted $b\bar{b}b\bar{b}$ analysis, which is the high p_T part of the $HH \to b\bar{b}b\bar{b}$ analysis. Of the 45

possible diHiggs decay modes shown in Figure 2.4, six channels have been explored to date: $b\bar{b}b\bar{b}$, $b\bar{b}W^+W^-$, $b\bar{b}\gamma\gamma$, $b\bar{b}\tau^+\tau^-$, $W^+W^-\gamma\gamma$, and $W^+W^-W^+W^-$. Each channel offers some appeal for exploring.

Channels with $b\bar{b}$ have the highest branching ratio and therefore will yield the largest number of events to analyze. It is limited however by large amount of background produced via strong interactions, which is explained further in Section 7. Channels with W^+W^- depend on dilepton final states and therefore never fully utilize the high branching ratio of the channel. The channels with W^+W^- also have relatively poor Higgs boson mass resolution due to presence of neutrinos in their decay products. Channels with $\gamma\gamma$ tend to produce very "clean" results, in which there is very precise, high resolution reconstruction of Higgs bosons. If there is a significant number of events for $\gamma\gamma$ channels, resonances tend to be apparent amongst a smoothly falling background distribution. It suffers greatly, however, from the very low branching ratio and therefore low number of expected events for diHiggs searches. Channels with $\tau^+\tau^-$ have a decent number of events because of the branching ratio, but τ leptons are difficult to identify. The resolutions of the measurement of the (single) Higgs boson mass for the relevant decay channels are shown in Table 2.2.

$H \rightarrow$	m_H resolution
$b\bar{b}$	10 %
W^+W^-	20 %
$\gamma\gamma$	1-2 %
$\tau\tau$	15 %

Table 2.2. Higgs boson mass resolution for the common decay channels.

The conducted diHiggs analyses were combined with an partial data set of pp collision data from the 2015-2016 data taking period. The combined analysis sets the upper limit on the diHiggs cross-section for the Spin-0 resonance, as seen in Figure 2.7. The $b\bar{b}b\bar{b}$ analysis was the most sensitive channel for most of the resonant masses. For very low resonant masses the $b\bar{b}\gamma\gamma$ [18] and $b\bar{b}\tau^+\tau^-$ [19] analyses are more sensitive

than the $b\bar{b}b\bar{b}$ analysis [17]. Only the $b\bar{b}b\bar{b}$ and $b\bar{b}W^+W^-$ [19] analyses extend to a resonant mass of 3 TeV. As of December 2020, the $b\bar{b}\tau^+\tau^-$ analysis with data taken from the 2015-2018 period has also extended into 3 TeV as part of a new $b\bar{b}\tau^+\tau^-$ boosted analysis.

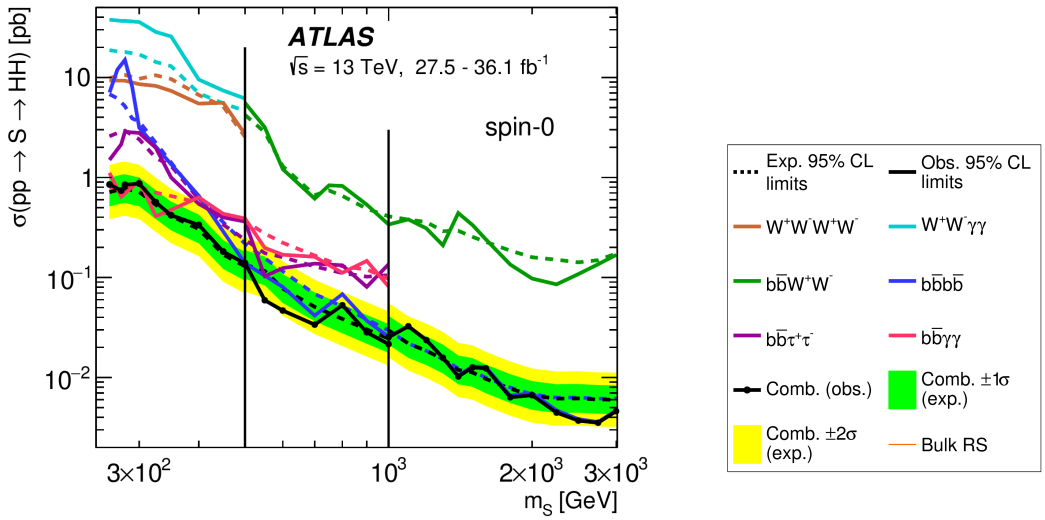


Figure 2.7. The 2015-16 result for the upper limit on the diHiggs cross section for the combined analyses.

2.6 Analysis Structure

The search for diHiggs events in the $b\bar{b}b\bar{b}$ decay channel is done between 251 and 5000 GeV, which is split up into two regimes: the lower resonance masses and higher resonance masses. The analysis that studies the former is known as the resolved analysis and the analysis that studies the latter is the boosted analysis, which is the focus of this thesis. Both $b\bar{b}b\bar{b}$ analyses look for two Higgs bosons, which decay into b-hadrons. The resolved analysis is sensitive between resonance masses of 251-1400 GeV. Events in the resolved analysis have low enough boost such that the Higgs bosons decay into two, well contained jets with a 0.4 radius. The resolved analysis

is also sensitive to the SM Higgs production, which is not included in the analysis presented in this thesis.

The boosted analysis is sensitive between 900-5000 GeV and is the focus of this thesis. Events in the boosted analysis have high enough transverse momentum such that if the same jets as the resolved analysis were used in the boosted analysis, it would result in overlap of the jets. Therefore, a larger radius jet is used to describe the Higgs boson and smaller, variable radius jets are used to describe the b-hadrons. This prevents the overlapping of jets and allows us to recover highly boosted diHiggs events. The boosted analysis also offers more multiple signal regions with less b-tags to compensate for b-tagging inefficiencies at high transverse momentum. Finally, the boosted analysis will combine with the resolved analysis to set an upper limit on the cross section of diHiggs resonances for the full range of 251-5000 GeV.

2.7 Analysis Approach

The general workflow of this analysis is outlined:

1. The ATLAS detector collects data from billions of particles emitted in pp collisions.

2. Advanced software reconstructs ATLAS readouts into physics data, to be used in analysis.

3. Rigorous analysis is performed to decipher actual $HH \rightarrow b\bar{b}b\bar{b}$ events from events that look very similar to it. We can remove many of these similar looking "background" events from the real resonant ("signal") events by exploiting tendencies experienced majorly by background and minimally by signal. One such way is to employ are a set of "cuts" that remove events that do not meet a certain kinematic requirement. Ultimately, it is very unlikely to remove all background so it is important to properly estimate how much background will

pass our selection process and will be observed in our final estimation of signal events. The region with possible signal events remains "blinded", or hidden from the analyzer to prevent bias, until the analysis structure is completely defined and proven.

4. Once the signal region selected data events are revealed (unblinded), the background estimation will be compared with the unblinded data. If there exists a resonance, it may be observed as an excess of data events when compared to background events, which must be of statistical significance to be acknowledged. If there is no excess and good agreement is seen between background and data, it validates the null-hypothesis.

5. In the absence of statistically significant resonances, a statistical analysis is performed to set an upper limit on the cross section of the resonances.

6. The $b\bar{b}b\bar{b}$ analysis upper limit is combined with other diHiggs analyses to create a combined upper limit on the cross section of Higgs boson pair production.

CHAPTER 3

LHC AND ATLAS

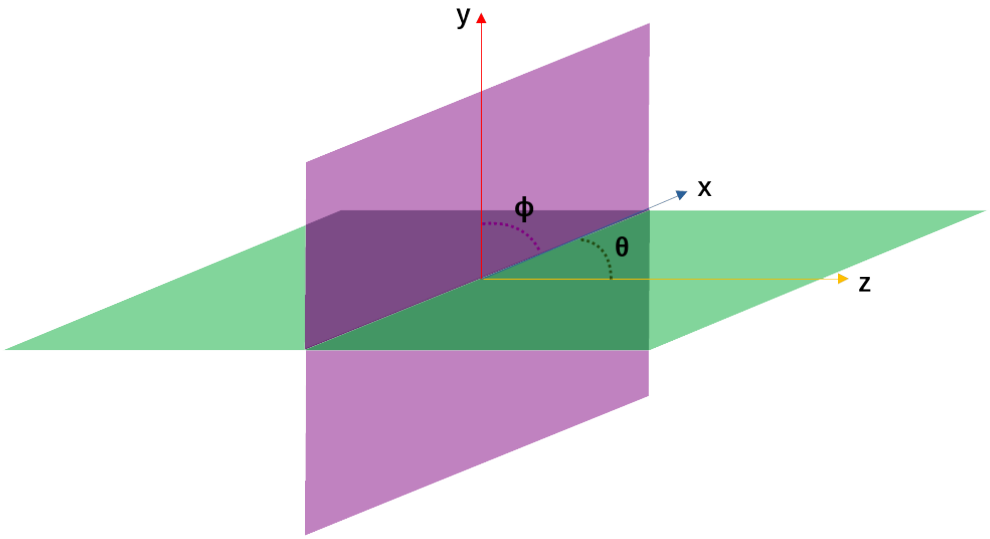


Figure 3.1. ATLAS coordinate system. The purple plane represents the transverse plane, x-axis points towards the middle of the LHC ring, z-axis points along the beam line, and the y-axis points upward.

The ATLAS detector is a multi-layered, multipurpose particle detector located on the LHC at CERN [11]. Shown in Figure 3.2, ATLAS contains cutting edge technology that makes up the several components of the detector, all of which serve a specific purpose in particle detection.

ATLAS uses a right-hand coordinate system, which is shown in Figure 3.1. For reference, the x-axis points towards the center of the collider ring, the z-axis points parallel to the beam axis, and the y-axis points upwards away from the center of Earth. The x-y plane is the transverse plane and is orthogonal to the beam axis. In polar coordinates, ϕ is the angle in the transverse plane and θ is measured in the x-z plane. Throughout this analysis, pseudo-rapidity is used instead of θ as it is Lorentz invariant and is defined in Equation 3.1.

$$\eta \equiv -\ln \tan\left(\frac{\theta}{2}\right) \tag{3.1}$$

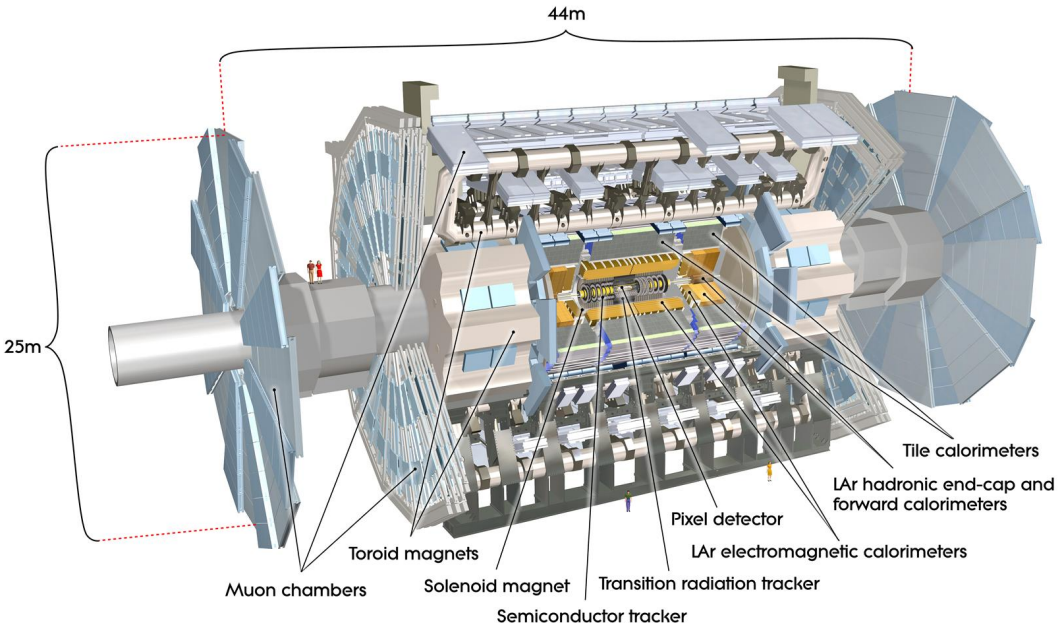


Figure 3.2. Image of ATLAS detector with a quarter cut away to see internal components.

3.1 Large Hadron Collider

The Large Hadron Collider (LHC) is a hadron collider at the European Organization for Nuclear research. The LHC spans a diameter of about 8.5km and is situated an average of 100 meters below the ground near the French border of Geneva, Switzerland. It features four major experiments, ATLAS, CMS, LHC-b, and ALICE, which are located at four separate interaction points (IP).

The LHC contains two parallel tubes called beam pipes, which house the protons during acceleration, before they reach the IP. Protons are supplied to the LHC beam pipe by an injection chain, which is outlined in Figure 3.3 [41]. Once protons are injected into the LHC, they are accelerated to achieve center-of-mass collisions of $\sqrt{s} = 13$ TeV (with considerations up to 14 TeV).

The luminosity L is described as,

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} F, \tag{3.2}$$

where N_b^2 is the number of particles per bunch, n_b is the number of bunches per beam, f_{rev} is the revolution frequency, γ_r is the Lorentz factor, ϵ_n is the normalized transverse beam emittance, β^* is the beta-function at the collision point, and F is the "geometric luminosity reduction factor". F accounts for a luminosity change caused by the crossing angle of the beams and is described as,

$$F = \left(1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*}\right)^2\right)^{-1/2}, \tag{3.3}$$

where θ_c is the angle of beam crossing, σ_z is the RMS bunch length, and σ^* is the transverse RMS beam size at the IP.

Together with the cross section of a given event, σ_{event}, the number of events per second N_{event} is given in Equation 3.4.

$$N_{event} = L\sigma_{event} \tag{3.4}$$

We integrate the luminosity over a given time period to obtain integrated luminosity L_{int}. Described in Equation 3.5, L_{int} is one of the best metrics to see how much data is collected at the experiment. ATLAS is one of two high luminosity experiments at the LHC and recorded a total integrated luminosity of $139 fb^{-1}$ during the 2015-2018 `Run-2` data taking period of $\sqrt{s} = 13$ TeV collisions.

$$L_{int} = \int L dt \tag{3.5}$$

3.2 Inner Detector

The inner detector (ID) is the innermost component of the detector. Located around the beam pipe, it has four sub-components shown in Figure 3.4: The Insertable

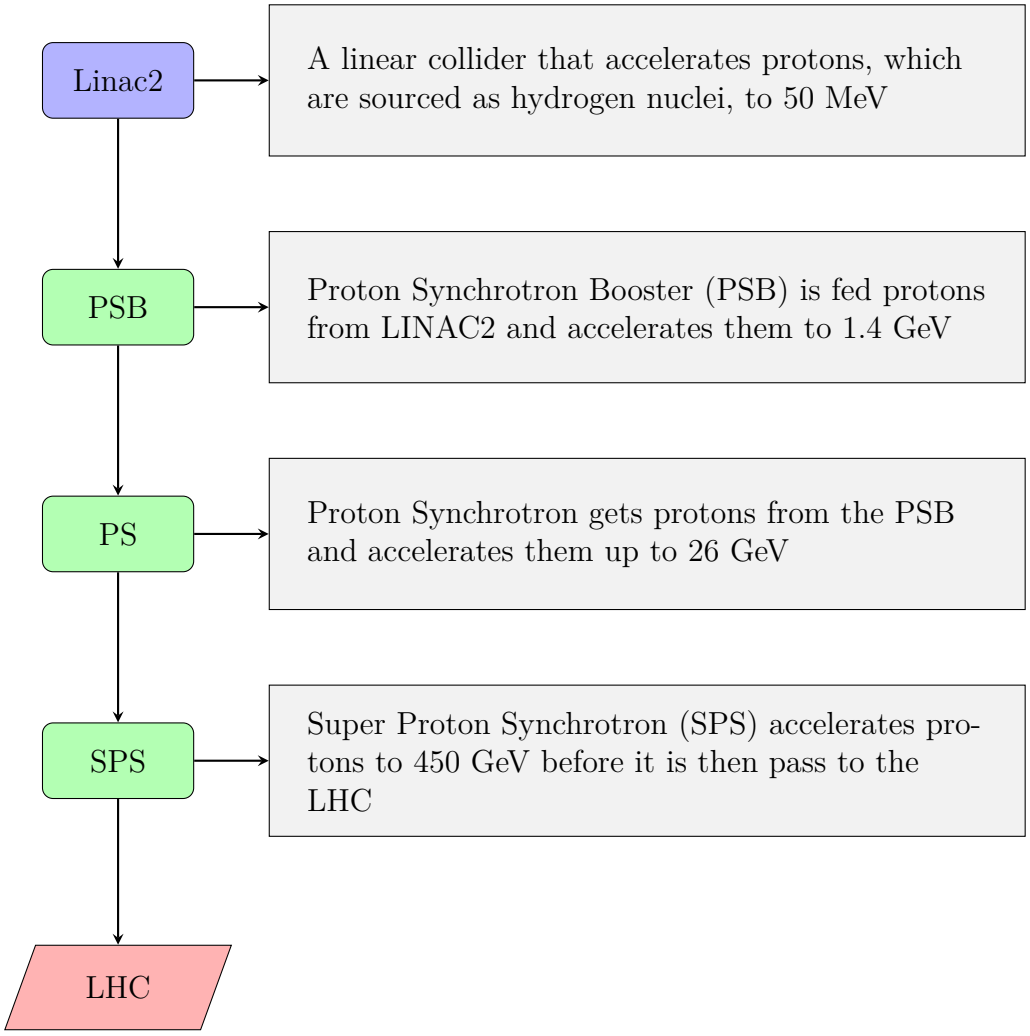


Figure 3.3. Visualization and description of injection chain for pp collision at the LHC for `Run-2` .

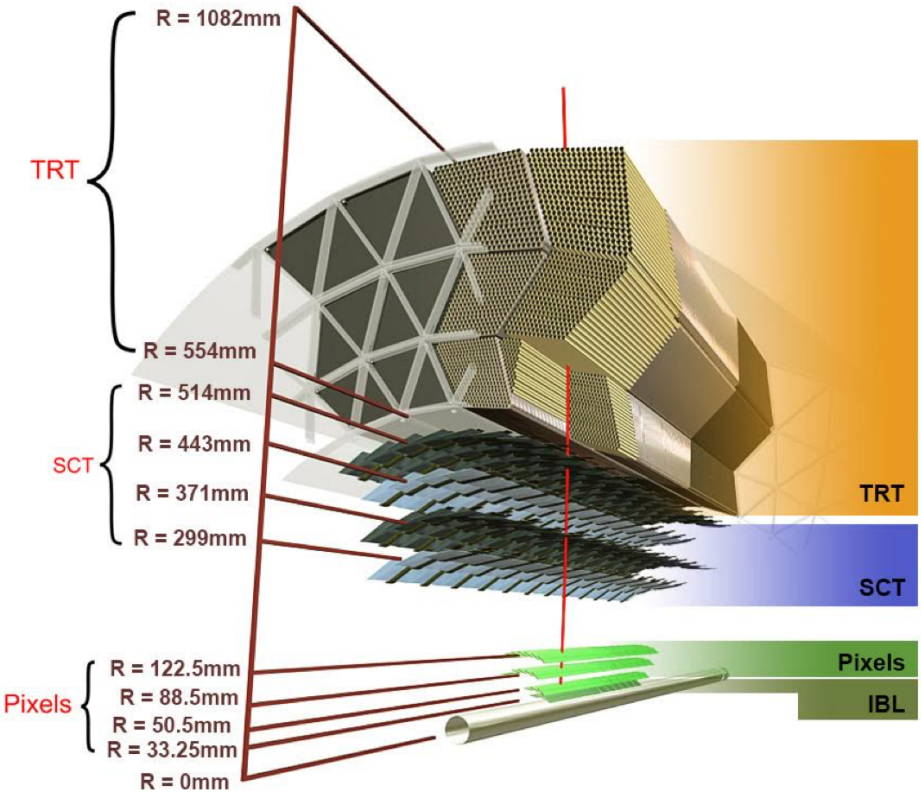


Figure 3.4. Cross section cut-away of the inner detector to show the barrel layers and the components of each respective layer.

B-Layer ("IBL", innermost) The Pixel detector ("Pixel detector", next-to-innermost), the Semiconductor Tracker ("SCT", next-to-outermost) and the Transition Radiation Tracker ("TRT", outermost). The ID covers a pseudorapidity of $|\eta| < 2.5$ and excels in tracking, which tracks the trajectory of charged particles. Specifically, the ID is designed to be highly efficient in b-hadron tagging (b-tagging) and tau tagging. The ID is surrounded by a solenoid, which immerses the ID in a 2 T magnetic field.[12]

The Pixel detector surrounds the beam pipe and has three barrel layers and three forward endcap disks. The locations of each layer is listed in Table 3.1. The name is derived from the pixels that make up the detector, which measure 50 by 400 μm^2. The Pixel is made up of modules, which contain silicon sensors bump-

layer	r [mm]	z [mm]
barrel-0	50.5	-400.5 to 400.5
barrel-1	88.5	-400.5 to 400.5
barrel-2	122.5	-400.5 to 400.5
disk-0	88.8 to 149.6	±495
disk-1	88.8 to 149.6	±580
disk-2	88.8 to 149.6	±650

Table 3.1. The position of Pixel components inside the detector. The disks are located at boths sides A and C of the detector, thus have two entries for z.

bonded to a front-end (FE) chip. The silicon senors house 47,268 pixel per module, which connect to one of 16 FE chips. Each pixel is $40400\mu m^2$ to $40600\mu m^2$ in $r\phi z$. The FE chip is then connected to the module-control chip (MCC), which enables communication with off-detector electronics. There are 1500 modules in the barrel and 700 modules in the endcap of the Pixel.[43]

The over 8 million pixels allow for the most precise tracking of any part of the detector. It is one of two essential components necessary for accurate vertexing, the other being the IBL. However preferment, the Pixel technology is very costly thus is prioritized to be located in the innermost part of the detector where the radiation will be greatest. At this part of the detector, radiation is considerable and will damage the integrity of the electronics over years of use.[11]

The IBL has been inserted in between the Pixel and beam pipe, where it excels at vertexing and b-tagging along with the Pixel. As it is so close to the beam pipe the IBL is designed to handle the radiation and luminosity present at the innermost part of the detector. The IBL uses pixel sensors bump-bonded to FE chips, which lie at a mean radius of $r = 33mm$. Each of the 32 FE chips are bonded to a respective 26880 pixel cells.[6]

The performance gain from the use of the IBL is demonstrated in Figure 3.5 [12]. The transverse impact parameter (d_0) resolution for tracks from a single particle

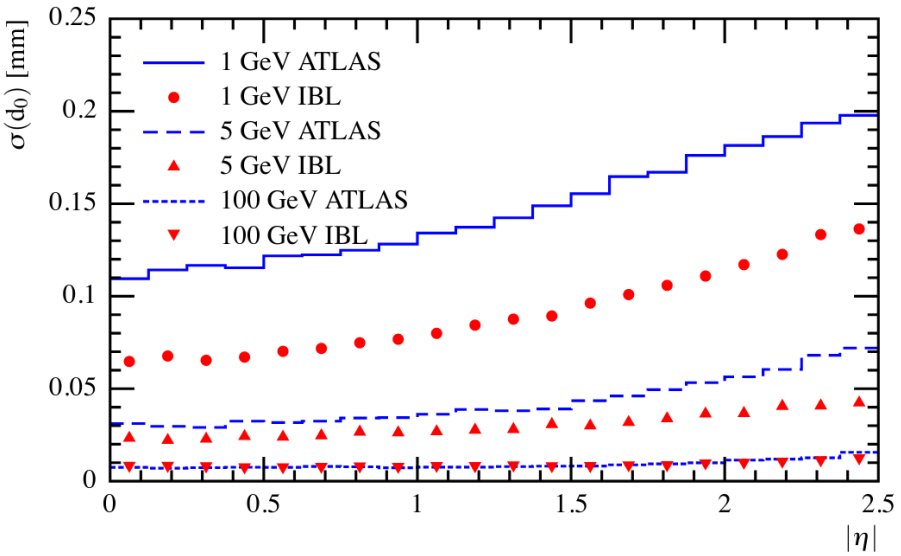


Figure 3.5. The transverse impact parameter resolution as a function of η for a 1, 5, and 100 GeV muon.

(a 1, 5, and 100 GeV muon) shows an improvement. A better impact resolution parameter results in improved b-tagging as demonstrated in Figure 3.6.

The SCT is the middle component of the ID and helps measure particle momentum. The SCT contains four concentric barrels and two endcaps that contain three disks respectively. The barrel has 2112 modules and the endcaps have 988 modules a piece. The modules are made up of four 12cm rectangular silicon strip sensors. In each module, two pairs are glued back to back to form a silicon strip double-layer.[16]

The outermost component of the inner detector is the TRT. It is a combination straw tracker and transition radiation detector. The TRT has a barrel containing 52,544 straws and endcaps that consist of 18 wheels and contain a total of 245,760 straws. The straws are drift tubes measuring 4 mm in diameter and up to 144cm in length. Each straw is filled with gas and has a fine wire down the axis of the cylindrical straw. Charged particles pass through the TRT and ionize the gas in the straws. A potential difference between the wire and the walls of the tube cause ions to drift to the center wire, which are then detected. Additionally, gaps between the cylindrical tubes

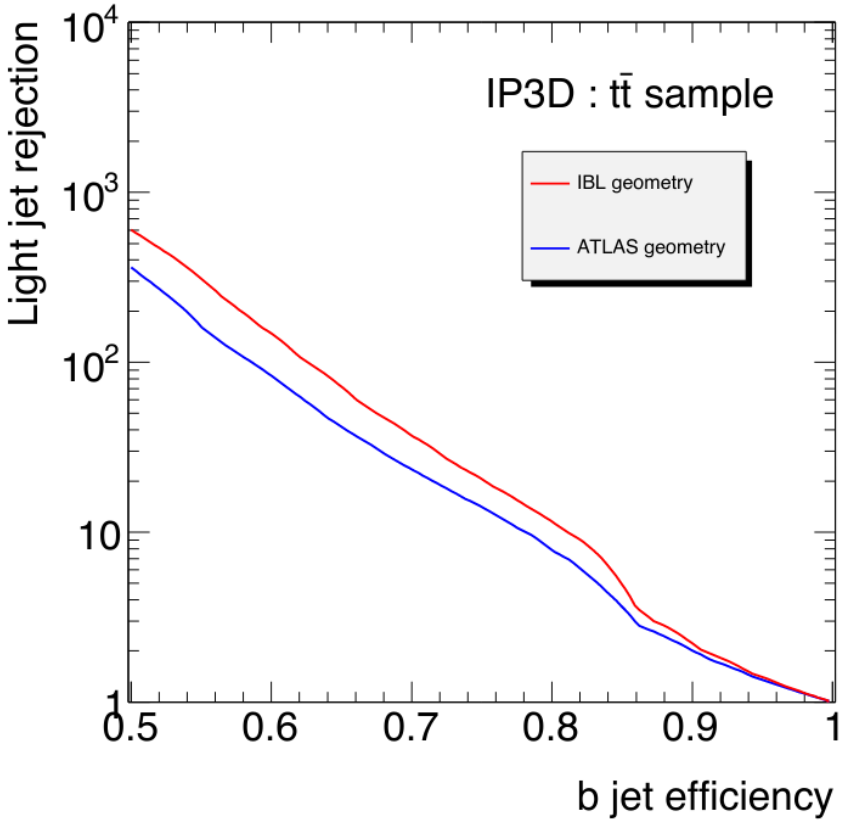


Figure 3.6. Efficiency and light-rejection rate of the b-tagging algorithm for the ATLAS detector with and without the IBL.

are filled with materials that have varying indexes of refraction, which facilitate the production of transition radiation. Transition radiation is the emission of photons from charged particles passing through media of different indexes of refraction.[1] The amount of transition radiation is larger for a lighter particle when compared to a heavier particle with similar transverse momentum. The charged particle passes through the TRT and emits transition radiation, which leaves a signal in nearby straws. Overall, the TRT provides a much coarser measurement when compared to the pixel components of the ID but offers a cost-efficient solution to handle the high luminosity of the LHC as well as a way to detect transition radiation.[15]

3.3 Calorimeter

The Electromagnetic (EM) calorimeter is composed of several lead-liquid argon (LAr) layers that absorb energy from passing charged particles via EM interactions and measure the energy lost. Heavy lead plates are separated by thin layers of LAr and electrodes, which are then structured as in an accordion pattern seen in Figure 3.7. Electromagnetically interacting particles pass through the dense calorimeter causing particle "showers", which occur when high energy particles cascade into many lower energy particles. Photons interact with the lead layers in the EM calorimeter, which result in production of electron-positron pairs and cause ionization in the LAr layer.[8]

The Hadron "Tile" calorimeter is outside of the EM calorimeter and operates by measuring energy losses from particles passing through many dense steel-scintillator layers. The steel and scintillating tiles are known as the "absorbing material" and "active material" respectively. The absorbing material precipitates a showering of particles and the active material measures the energy of those secondary particles. This is useful for gathering information on particles interacting via the strong force. Unlike the EM calorimeter, the Tile calorimeter is cylindrical and has radius of $2240 < r < 4230$ mm, which is shown in Figure 3.8. ([9])

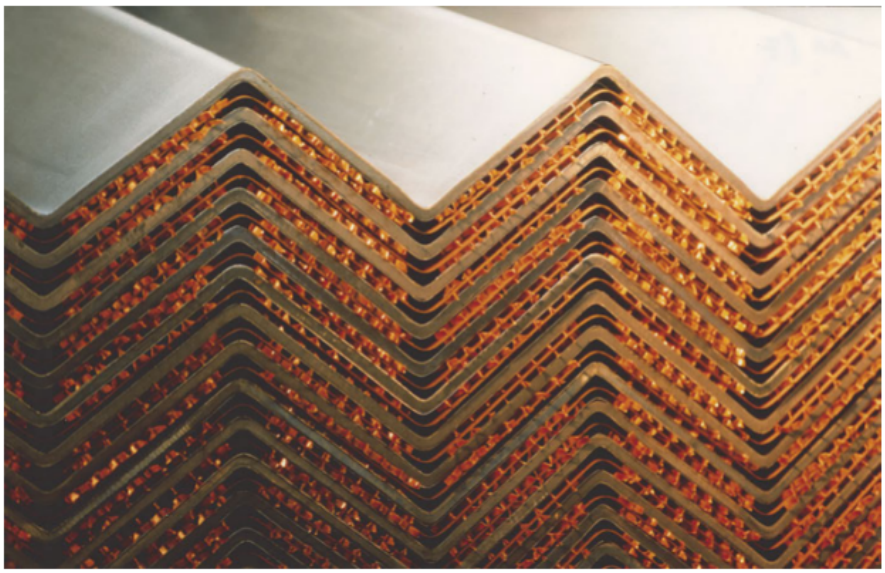


Figure 3.7. Lead plates of the EM calorimeter shown in the accordion pattern. Electrodes are in between the lead plates.

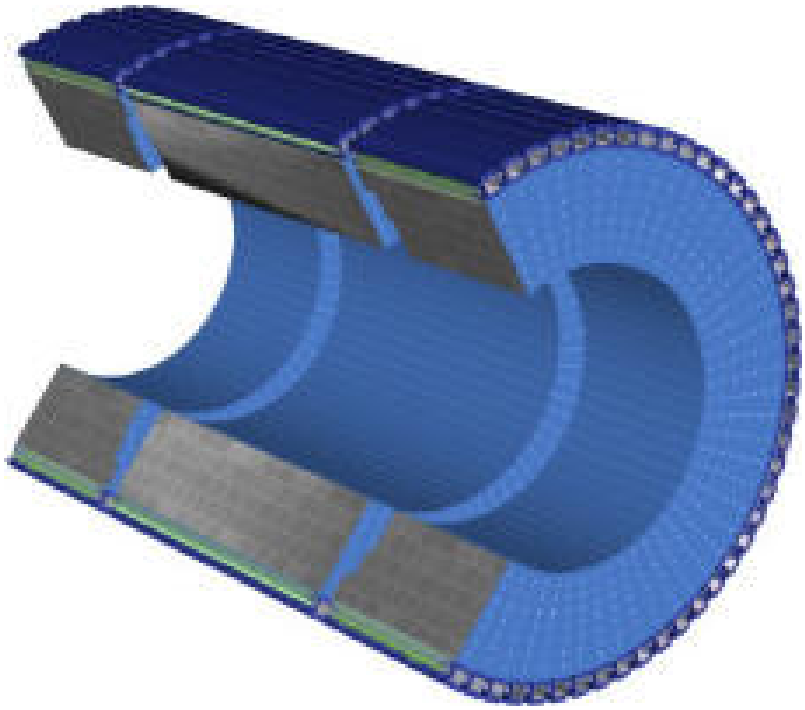


Figure 3.8. The barrel of the tile calorimeter. It has a long, central barrel and two extended barrels.

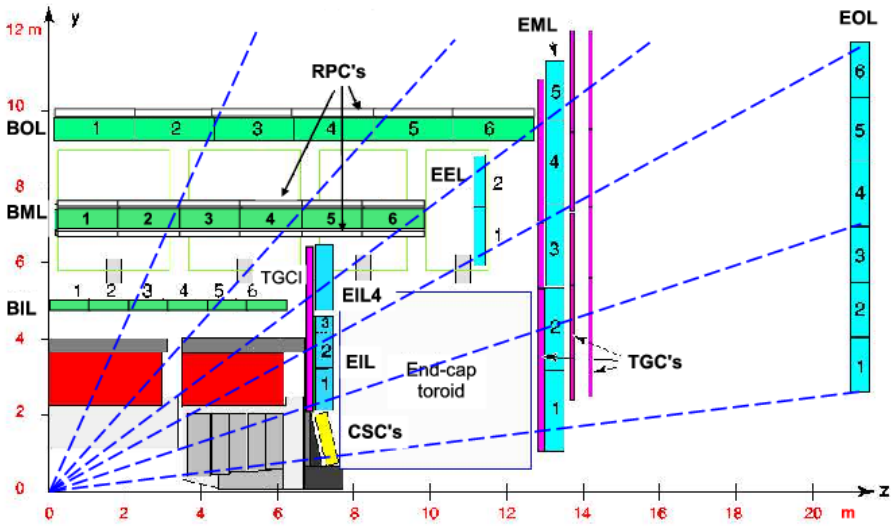


Figure 3.9. A cross section of one section of the Muon Spectrometer.

3.4 Muon Spectrometer

The outermost layer is the muon spectrometer (MS), which detects primarily muons that make it through the ID and calorimeters. It is made of three barrel layers and two endcaps each with three wheels. The MS contains tracking components Muon Drift Tubes (MDTs) and Cathode Strip Chambers (CSCs). The MDTs are drift tubes, which operate similarly to the TRT straws described in Section 3.2. The MDTs number about 350,000 for the entire MS and measure 3cm in diameter and of length varying from 85-650cm. Also similar to the straw tubes of the ID, the MDT's are a cost effective solution to cover large areas of the MS that require tracking. At the forward region, at the inner part of the endcap wheels, are the CSCs. The CSCs act to have standalone trigger and measurement capabilities and better performance with high rates of particles as compared to MDTs. The trigger components are the Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs). The layout of the MS is demonstrated in Figure 3.9.[10]

3.5 Magnet System

The detector has inter-layered magnets, which modify the trajectory of charged particles as they pass through the detector. This is important in reconstructing the momentum of a particle by observing the curve due to the magnetic field. There is a 2 tesla (T) solenoid outside the ID and a toroidal superconducting magnet embedded in the MS, which has a magnetic field integral that ranges from 2 to 8 Tm.

3.6 Trigger

ATLAS contains a two part trigger system, which parses the many collision events passing through the detector, only to keep and record the ones of interest. The Level-1 (L1) trigger is a hardware trigger, which uses information from the calorimeter and MS, to make a determination in about 2.5 s. The L1 trigger effectively reduces rate at which events need to be processed at the LHC from 40 MHz to 100 MHz. The High Level Trigger (HLT) is a software trigger, which is applied to events that pass the L1 trigger. The HLT further reduces the bunch crossing rate at the LHC to about 1KHz.[24]

CHAPTER 4

DATA AND MONTE CARLO SIMULATION

4.1 Data

This analysis uses the full ATLAS `Run-2` data set, which consists of $\sqrt{s} = 13$ TeV pp collision data taken during the operational 2015, 2016, 2017, and 2018 years. Figures 4.1 and 4.2 show the total integrated luminosity as a function of time and the bunch crossings per year, respectively. The amount of data since the start of data taking in 2011 has increased every year, except for the brief 2015 data set. The mean bunch crossing "μ" varies by year from an average of 13.4 to an average 37.8.

"Runs" are a period of time with continuous data taking, which are further broken into "luminosity blocks" (also known as "lumi-blocks" or LBs). A Good Runs List (GRL) is assembled by the review of offline and online data quality (DQ) of reconstructed data. Lumi-blocks that have data deemed "good-for-physics" are added to the GRL, which is then applied to data samples to be used in analysis. A total integrated luminosity of 139.1 fb^{-1} passes the full `Run-2` GRL, as compared to the 32.9 fb^{-1} that passed the 2015-2016 partial `Run-2` GRL, with the complete breakdown listed in Table 4.1. The DQ inefficiency for pp collisions at $\sqrt{s} = 13$ TeV is 4.4%, with the full breakdown per ATLAS component listed in Figure 4.3 [4]. The GRL xml files for 2015, 2016, 2017 and 2018 in descending order are:

- `data15_13TeV.periodAllYear_DetStatus-v89-pro21-02_Unknown_PHYS_StandardGRL_All_Good_25ns.xml`
- `data16_13TeV.periodAllYear_DetStatus-v89-pro21-01_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns.xml`
- `data17_13TeV.periodAllYear_DetStatus-v99-pro22-01_Unknown_PHYS_StandardGRL_All_Good_25ns_Triggerno17e33prim.xml`
- `data18_13TeV.periodAllYear_DetStatus-v102-pro22-04_Unknown_PHYS_StandardGRL_All_Good_25ns_Triggerno17e33prim.xml`

Year	Total Integrates Luminosity [fb^{-1}]
2015	3.2
2016	32.9
2017	44.3
2018	58.5

Table 4.1. Total integrated luminosities passing the GRL, collected by year.

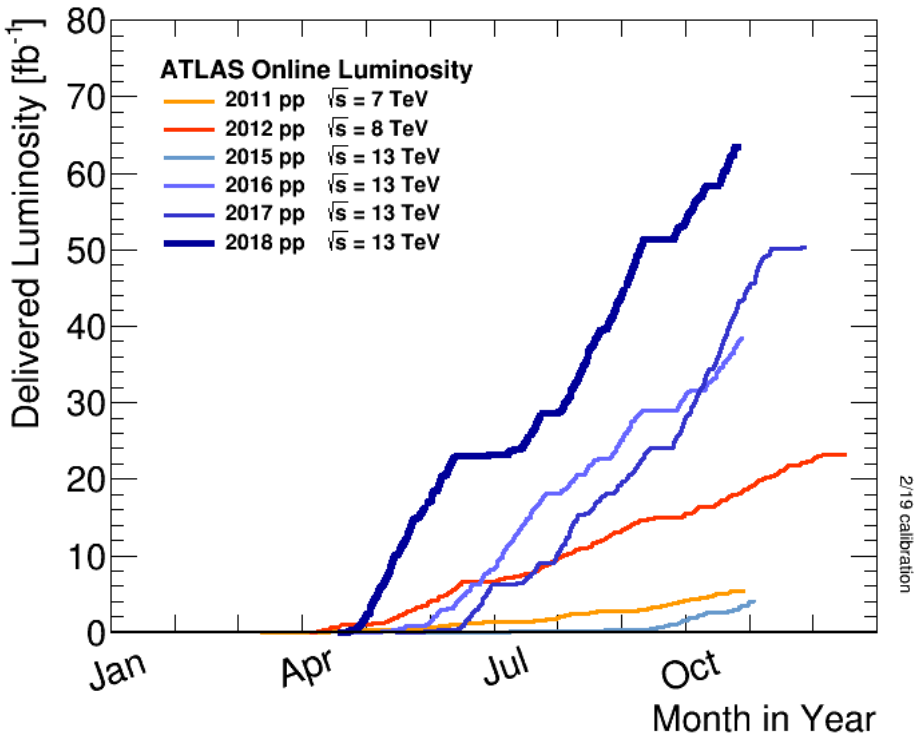


Figure 4.1. Luminosity per year as a function of time in the year for the entire data taking time period of ATLAS.

The boosted analysis uses a different trigger for each year of the `Run-2` data. Table 4.2 lists the combination of triggers and their plain-text description. All the triggers demand a selection of one "high p_T " large-R Jet. The 2017 and 2018 triggers have an additional mass requirement on one of the large-R Jets.

Low transverse momentum pp collisions are considered "pile-up" events and come in two varieties: out-of-time and in-time. The out-of-time time pile-up is from low-p_T particles passing through he detector from a preceding or following bunch crossing. The in-time pile-up is from pp collisions of the same bunch crossing. In-time and out-of-time pile-up is collectively called "pile-up" and may leave signatures in the detector such as tracks or calorimeter readings, which negatively affect reconstruction. The removal of pile-up is essential to the experiment especially as the average number interactions per bunch crossing increases, which will result in greater

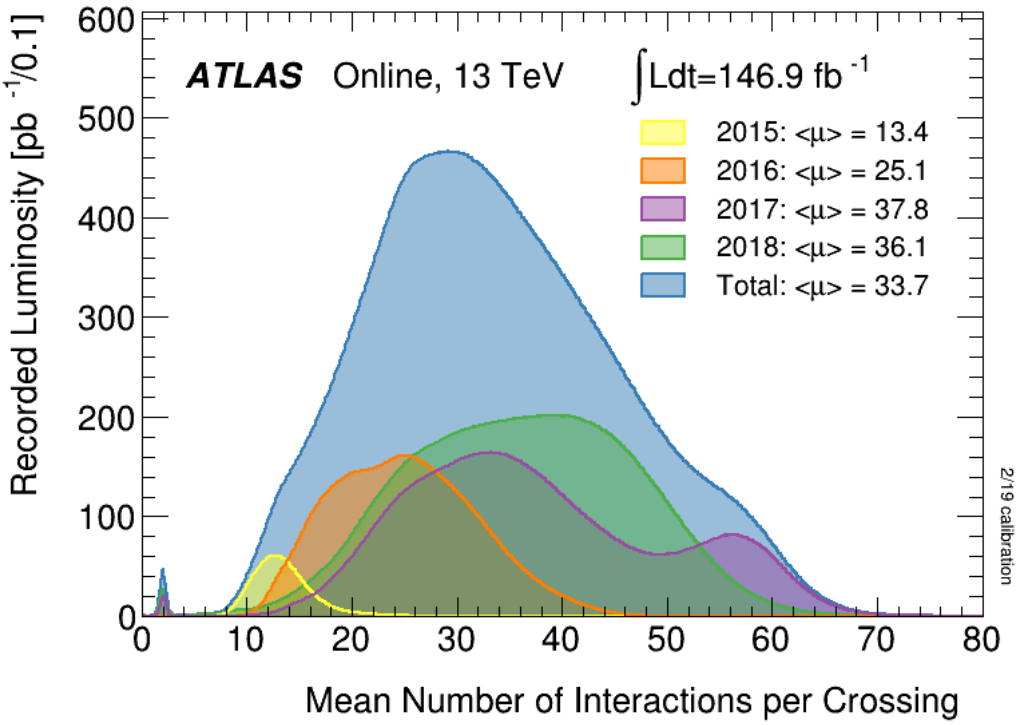


Figure 4.2. Mean number of interactions per crossing separated by year for `Run-2` .

pile-up. A selection and trigger requirement in this analysis will help remove much of the low-p_T events seen in pile-up.

The 2015 and 2016 triggers follow the recommended triggers in the previous partial `Run-2` analysis. The trigger are determined to be 100% efficient when applied in conjunction with the $p_T > 450$ GeV selection, which is further described in Section 6.1. The 2017 and 2018 triggers were chosen in attempt to maintain a 100% efficiency without drastically increasing the p_T threshold to the 2015 and 2016 triggers while simultaneously dealing with the higher pileup of the later years. This is achieved with the use of triggers that include the large-R jet mass requirement. As shown in Figure 4.4 and 4.5 , the trigger appears 100% after a p_T and m of over 450 GeV and 50 GeV respectively for a given 1 TeV signal sample. The metric of reverse cumulative trigger efficiency is explained in detail in Reference [5]. Furthermore, these triggers

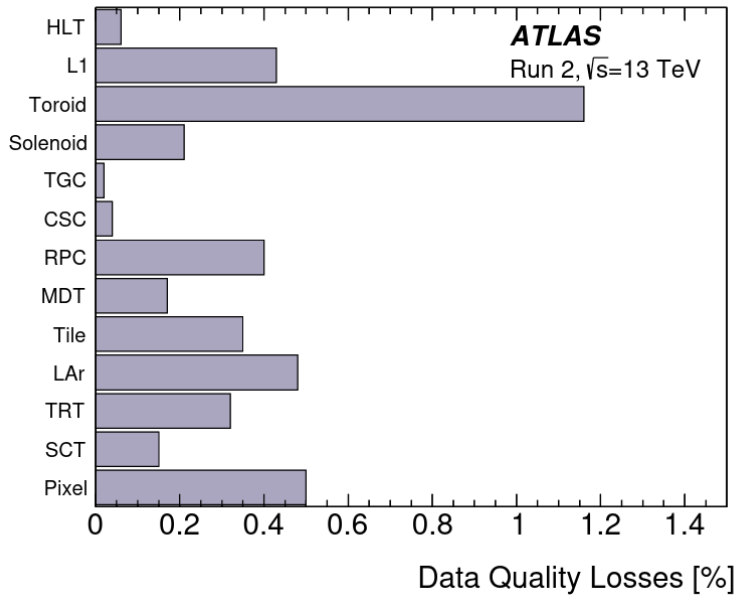


Figure 4.3. Data quality inefficiencies for each component of the ATLAS detector.

applied in conjunction with the large-R Jet p_T and mass described in 6.1 ensure 100% efficiency for the analysis.

4.2 MC Simulation

Monte Carlo (MC) simulation samples are created in three campaigns, which are labeled by mc16a, mc16d, and mc16e. The data-taking conditions for mc16a, mc16d and mc16e correspond to 2015-16, 2017 and 2018 respectively. Pileup reweighting is applied to match the MC samples to the interactions-per-bunch-crossing (μ) data distributions for each corresponding year. The value of μ for 2015-16 is found by averaging the μ distribution for each lumi-block of the 2015-16 data-taking time period. For 2017 and 2018, events in the MC samples are reweighted with the actual value of μ, as opposed to the average μ per lumi-block.

Although this analysis contains a primarily data-driven background estimation, MC samples are used in estimating $t\bar{t}$ as it is the source of subleading background in

Boosted Triggers	Description
2015	
HLT_j360_a10_lcw_sub_L1J100	$p_T{}^J > 360$GeV
2016	
HLT_j420_a10_lcw_L1J100	$p_T{}^J > 420$GeV
2017	
HLT_j420_a10t_lcw_jes_40smcINF_L1J100	$p_T{}^J > 420$GeV, $m^J > 40$GeV
2018	
HLT_j420_a10t_lcw_jes_35smcINF_L1J100	$p_T{}^J > 420$GeV, $m^J > 35$GeV

Table 4.2. Triggers for the Boosted analysis separated by year and their respective cuts.

our analysis. The other source of background is from QCD initiated jet-jet events that are not the b-hadron signal the analysis searches for. These "Dijet QCD" background events can be simulated as MC events and are useful in validating our background estimation method and serves as a quick reference for exploratory studies that need to be independent from our background estimation. Signal MC samples are used to test efficiencies, to understand the shape of signal distributions, and ultimately in limit setting, which is done for the two resonant signals of this analysis.

The spin-0 resonance is simulated as the "Narrow-width" Scalar sample, since it is assumed to have a width smaller than detector resolution. It is generated with a model independent fixed width of 1000 GeV. The Narrow-width Scalar MC samples in the boosted analysis are simulated at leading order in α_S with MadGraph, parton showering is done using Herwig 7, and the nominal PDF is NNPDF 2.3 LO. EvtGenis used to handle the decays of b-hadrons and c-hadrons after the showering and hadronization is modelled. There are fourteen Narrow-width Scalar samples used with masses of 900, 1100, 1200, 1300, 1400, 1500, 1600, 1800, 2000, 2500, 3000, 3500, 4000, and 5000 GeV. Similarly, the G^*_{KK} featured in the bulk RS model with c=1.0 samples are simulated at leading order in α_S with MadGraph and with the nominal PDF NNPDF 2.3 LO, but the parton showering done using Pythia 8 with EvtGen.

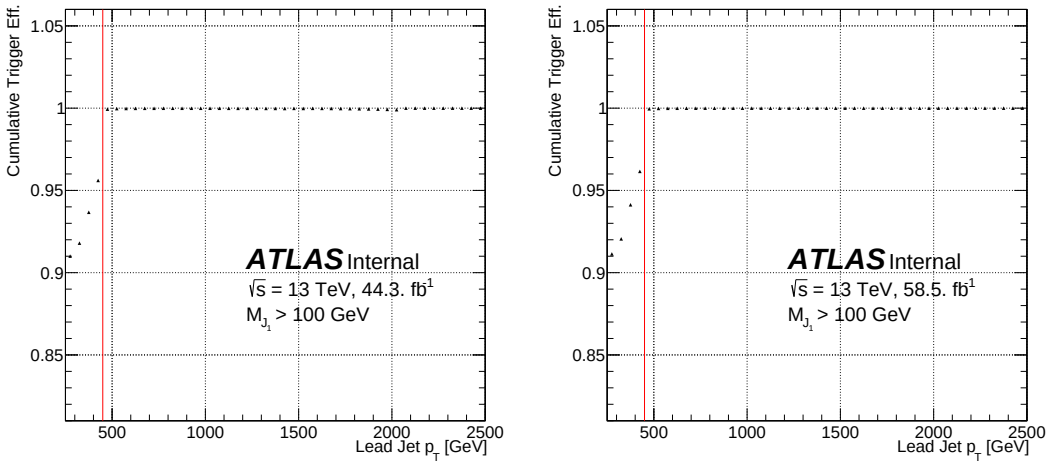


Figure 4.4. Reverse cumulative trigger efficiency as a function of transverse momentum for 2017 and 2018 data on the left and right respectively. The red line marks the 450 GeV cut in p_T applied to the boosted kinematic selection.

The G^*_{KK} samples have a width that depends on the resonance mass of the sample – varying from 3% at low mass to a wider 15% at high mass. There are fourteen G^*_{KK} samples used with masses of 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1800, 2000, 2250, 2500, 2750, and 3000 GeV. The cross sections for both signal sample sets are listed in Table 4.3.

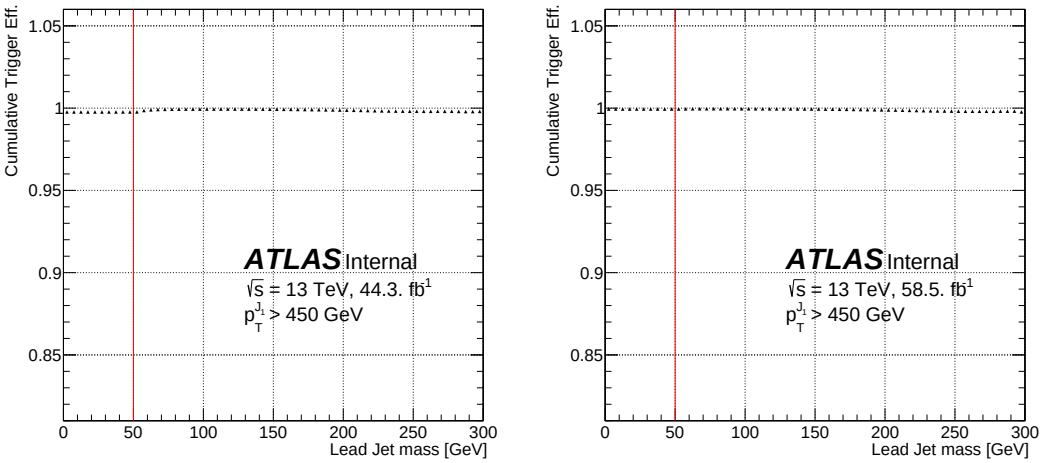


Figure 4.5. Reverse cumulative trigger efficiency as a function of mass for 2017 and 2018 data on the left and right respectively. The red line marks the 50 GeV cut in mass applied to the boosted kinematic selection.

Mass	Scalar (pb)	G^*_{KK} (pb)
900	1.6	0.020
1000	0.80	0.011
1100	0.41	6.9E-3
1200	0.22	4.3E-3
1300	0.12	2.7E-3
1400	0.070	1.8E-3
1500	0.041	1.2E-3
1600	0.025	7.9E-4
1800	9.6E-3	3.8E-4
2000	3.9E-3	1.9E-4
2250		8.4E-5
2500	5.3E-4	3.9E-5
2750		1.9E-5
3000	8.8E-5	9.4E-6
4000	3.4E-06	
5000	1.5E-07	

Table 4.3. Cross sections for simulated samples.

CHAPTER 5

EVENT RECONSTRUCTION

The complex and adept software used at ATLAS to turn detector read-outs into physics "objects" is described in this section. This process, known as "reconstruction", takes place in several steps before providing a final output to be analyzed. The prominent reconstruction objects used in this analysis are tracks, vertices, jets, and muons.

5.1 Tracks

The reconstructed trajectory of charged particles passing through the ID are called tracks. Potential track candidates are "seeded" off of measurements from the silicon detectors in the ID. Energy deposits in the silicon detectors above a threshold and close in proximity geometrically are grouped into "clusters". These clusters average about two pixels in $r-\phi$ and one to three pixels in η in the barrel. A track seed impact parameter is estimated by assuming a perfect helical trajectory in a uniform magnetic field in the ID. After clustering and seeding, a collection of track candidates are given a score, in which a higher score indicates the likelihood of a track to represent the true trajectory of the charged particle. Track candidates are scored once more after passing criteria to have $p_T > 500$ MeV, $|\eta| < 2.5$, at least 7 clusters, no more than two "holes" (an area in the detector that a track passes through, but has no respective cluster associated to it) in the SCT and Pixel, no track candidate sharing similar clusters, and no more than one hole in the Pixel. The transverse impact parameters is defined as distance of closest approach of the track trajectory to the primary vertex the transverse plane. The longitudinal impact parameter is the distance from the track to the z-axis (beam line) on the longitudinal plane. The transverse impact parameter of the track candidate must be smaller than 2mm and the difference between the longitudinal impact parameter and primary vertex to be less than 3mm.

Truth information is compared to data to validate track reconstruction. "True" reconstructed tracks have a truth-matching probability of 0.5 or greater, whereas

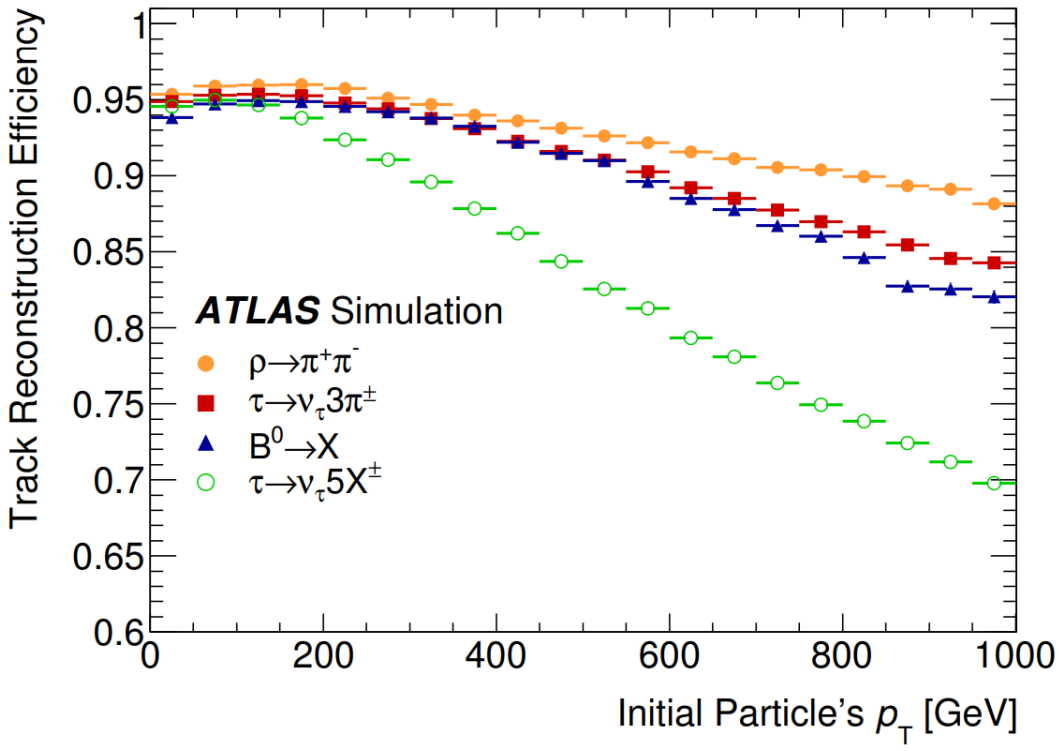


Figure 5.1. Reconstruction efficiency as a function of p_T for four different MC samples.

"fake" tracks have a truth-matching probability below 0.5. In a 2015 data sample test, only 1% of pp collisions produced fake tracks and for dijet MC test only 0.5%.

Track reconstruction performance is dependent on the p_T of the charged particle. Figure 5.1 shows how higher $p_T$tracks result in less precisely reconstructed tracks when compared to MC truth tracks. Notably, the track reconstruction efficiency for the b-hadron sample is 83-95% and is highest at low p_T . Higher p_T particle decays result in a more collimated collection of tracks and thus worsens clustering performance. This is demonstrated in figure 5.2, where the number of merged clusters for an MC sample increases as a function of p_T . Finally, the reconstruction efficiency is shown in Figure.[3]

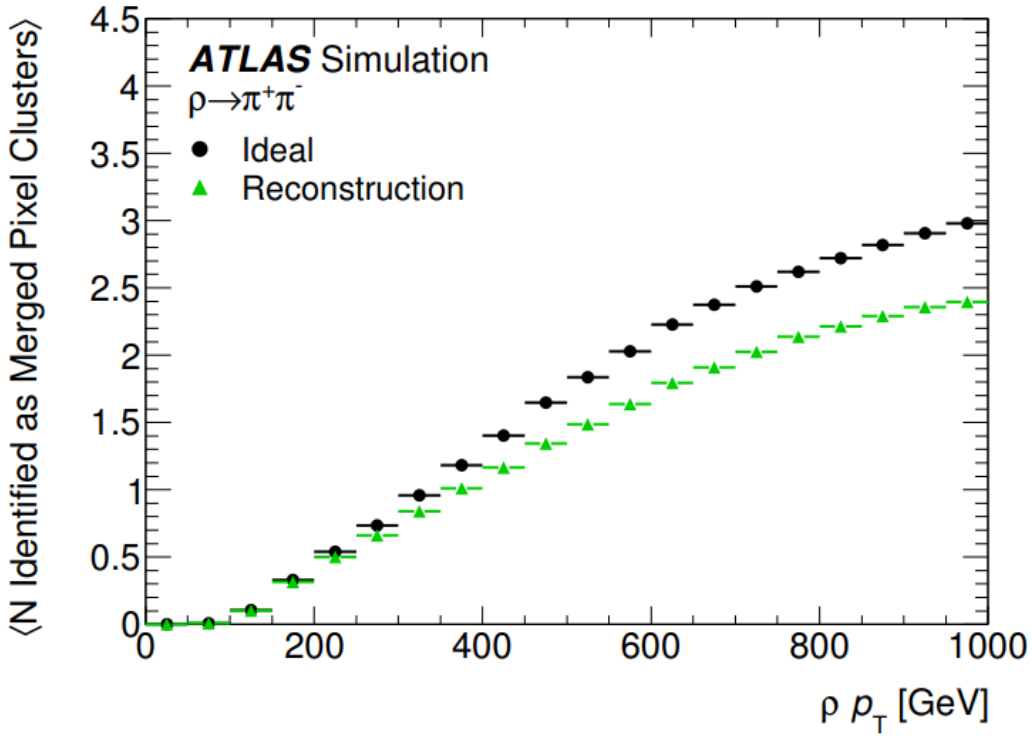


Figure 5.2. Ideal (truth) and reconstructed number of merged clusters for a $\rho \to \pi^+\pi^-$ MC sample.

5.2 Vertices

A vertex represents a point in space where an interaction occurs. Vertex reconstruction is done by collecting a set of tracks and performing an iterative fit to the set to determine where the tracks cross paths in space. First, a vertex seed position is selected as the center of the beam spot in the transverse plane and the point of closest approach of tracks with the center of the beam spot on the z-axis. Each track in the set is weighted based on likelihood of association to the vertex seed provided and tracks in the set that are not compatible with the seed are disregarded. The vertex seed is then recomputed and the process is repeated from the new seed and weighted tracks until a stable vertex is found.[25]

The "primary vertex" (PV) is used in this analysis and is essential in the suppression of pile-up and in b-tagging, which is described in Section 5.7. The sum of the squared p_T of the tracks associated to each vertex of an event is calculated and the vertex with the largest squared p_T sum is considered the PV. The PV is reconstructed with a 30 μm longitudinal resolution and 10-12 μm transverse resolution. Vertex reconstruction efficiency decreases as a function of interactions per bunch crossing as shown in Figure 5.3.[25]

5.3 Jets

A collimated spray of hadrons, which emerge from quarks or gluons, is known as a "jet". Topologically close calorimeter signals are clustered together to form a "topo-cluster". When jets are formed from topo-clusters, the direction of the jet is derived by pointing the cluster back to a PV in the ID.[27]

The "anti-k_t" jet clustering algorithm is used as an infrared and collinear safe jet finding algorithm [33]. The anti-k_t algorithm uses the transverse momenta, rapidity, and azimuth of topo-clusters to determine the distance between them. This distance, d_{ij}, is written as,

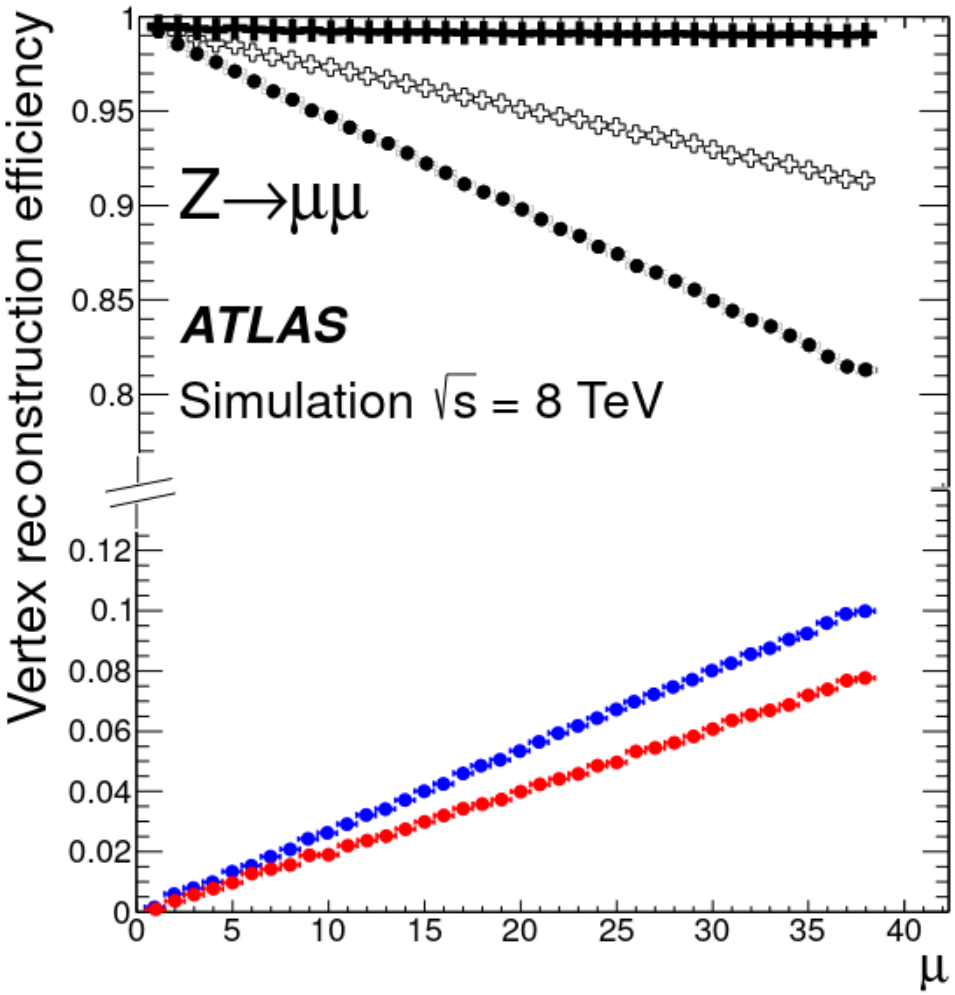


Figure 5.3. Vertex reconstruction efficiency for a $Z \to \mu\mu$ MC sample, which shows clean (•), low pile-up contamination (•), high pile-up contamination (•), clean+low pile-up contamination (empty +), and clean+low+high pileup contamination (solid +). The efficiency is a function of μ, the interactions per bunch crossing for a pp collision (average of 36.1 `Run-2`).

$$d_{ij} = min(1/k_{ti}^2, 1/k_{tj}^2)\frac{\Delta_{ij}}{R^2}, \tag{5.1}$$

where $k_{ti,j}$ is the transverse momentum of the i^{th} or j^{th} jet and Δ_{ij} is $(y_i^2 - y_j^2) + (\phi_i^2 - \phi_j^2)$ between the i^{th} or j^{th} jet. The value R is a radius factor, which represents the radius of a jet. R is tuned to a specific value depending on the optimal use of the jet, which varies typically from 0.2 for b-hadron jets to 1.0 for Higgs boson jets. The unique feature of equation 5.1 is the use of the inverse of transverse momentum, which prevents grouping of soft constituents with other soft constituents (i.e. it is "soft-resilient"). The result is "hard" jet constituents collecting other soft jet constituents and thus the formation of a well contained jet while being resilient to soft radiation.

5.4 Small Radius Jets

Small radius jets are anti-k_t jets with a radius of $R = 0.4$. These jets are used in the resolved analysis and are "particle flow" or "PFlow" jets. PFlow jets are reconstructed by removing energy deposits in the calorimeter from charged hadrons and replaces it with tracking information, as opposed to a traditional method of using calorimeter or tracking information alone. The methods used to reconstruct PFlow jets have an advantage in using tracking information because the calorimeter is more likely to encounter pile-up. Thus, the use of PFlow jets have improved the energy and angular resolution of the small radius jets in the resolved analysis.

When the small radius jets in the resolved analysis are b-tagged, they represent the final, four bottom-quark state of the analysis as shown in Figure 5.4. Thus, the resolved analysis requires a signal event to have exactly four b-tagged small-R jets. A BDT pairing algorithm is applied to match two of the four b-tagged jets into two pairs, which are meant to represent the decay from each separate Higgs boson.

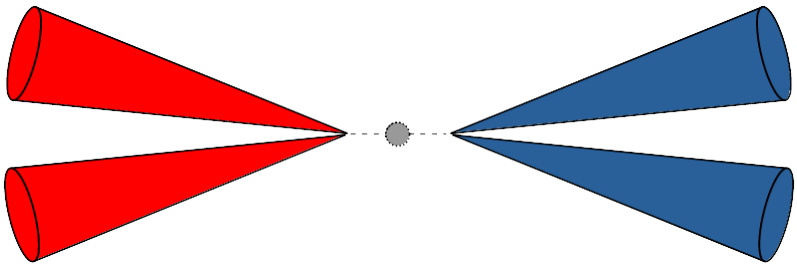


Figure 5.4. Depiction of the small radius jets in the Resolved analysis. Paired by color.

5.5 Large Radius Jets

Large radius jets, also know as "Large-R jets" or "fat jets", are anti-k_t jets with a radius $R = 1.0$. Large-R jets are the physics objects used to describe Higgs bosons in the boosted analysis and are made from locally calibrated topo-clusters.[27] A grooming procedure called "trimming" is applied to remove Large-R jets that are likely contaminated with pile-up. The trimming algorithm exploits the fact that jets with pile-up contamination typically observe more "subjets" (formed from topo-clusters inside the Large-R jet) with soft radiation. Using a k_t algorithm, the trimming procedure reclusters the Large-R jet into $R = 0.2$ subjets. Subjets not passing the criteria of

$$\frac{p_{Ti}}{p_T^{jet}} < f_{cut} \tag{5.2}$$

are removed, where p_{Ti} is the transverse momentum of the i^{th} subjet, p_T^{jet} is the transverse momentum of the untrimmed jet, and f_{cut} is the cutoff value set to 5%

(i.e. $f_{cut} = 0.05$). The subjets that remain after the trimming procedure form the final "trimmed jet".[13]

The mass of the Large-R jet is calculated by combining mass values from calorimeter and tracking information. First we start with the calorimeter mass m^{calo}, which is calculated as,

$$m^{calo} = \sqrt{(\sum_{i \epsilon jet} E_i)^2 - (\sum_{i \epsilon jet} \overrightarrow{p_i})^2}, \tag{5.3}$$

where each sum is over the i^{th} constituent in the jet, which is assumed to be massless. Then track information is used to calculate the track assisted mass m^{TA} as,

$$m^{TA} = \frac{p_T^{calo}}{p_T^{track}} m^{track}, \tag{5.4}$$

where p_T^{calo} is the transverse momentum of the jet from the calorimeter, p_T^{track} is the combined transverse momentum of the tracks associated to the jet, and the m^{track} is the combined invariant mass of the tracks associated to the jet. Tracks are associated to the jet via ghost association, which is explained further in Section 5.6. This is then combined to form a m^{comb} by,

$$m^{comb} = w * m^{calo} + (1 - w) * m^{TA}, \tag{5.5}$$

where w is a weight calculated for each Large-R jet from the resolution of the calibrated track and calorimeter mass terms. The inclusion of m_{TA} in m_{comb} improves the reconstruction of high p_T events, in which the calorimeter has poor angular granularity and the tracking excels in measuring the angular separation of nearby tracks. Ultimately, the use of m_{comb} improves the mass resolution as compared to m^{track} or m^{calo} alone.[20]

5.6 Track jets

Track jets are anti-k_t jets that are reconstructed from clusters of tracks in the ID. The tracks are meant to be found from hadrons passing through the ID, thus are used to describe b-hadrons that decay within the Large-R jet. Track jets are a physics object used primarily in the boosted analysis and have a variable radius (VR), which is defined by,

$$R = \frac{\rho}{p_T}, \tag{5.6}$$

where p_T is the transverse momentum of the jet and ρ is a free parameter. The algorithm that reconstructs VR track jets use the parameter ρ to control the influence of the transverse momentum on the effective radius of the jet. For the boosted analysis, the algorithm uses a ρ set to 30 GeV and is also constrained to a range of radii, $R_{min} = 0.02$ and $R_{max} = 0.4$.[28]

Other values of the ρ parameter, R_{min}, and R_{max} have been considered. Figure 5.5 plots the $h \to bb$ double subjet b-labelling efficiency for values of ρ from 10 to 50 GeV, where 30 GeV is highest efficiency for most values of p_T . Other values for R_{min} and R_{max} are also considered but 0.02 and 0.4 are chosen for maximal sensitivity while also having a minimal constraint on the radius.

Track jets are associated to Large-R jets by "ghost association" – the process of creating "ghosts", which are reconstructed track jets with transverse momentum set to an infinitesimally number (1 eV). The directionality of the ghosts is maintained and they are then passed into the collection of Large-R jet constituents. The negligible p_T of the ghosts do not effect the calorimeter measurements that the Large-R jet reconstruction depends on. Once the Large-R jet is then reclustered, it is apparent which tracks are clustered into the resulting Large-R jet and thus a track jet is associated to it accordingly.[14]

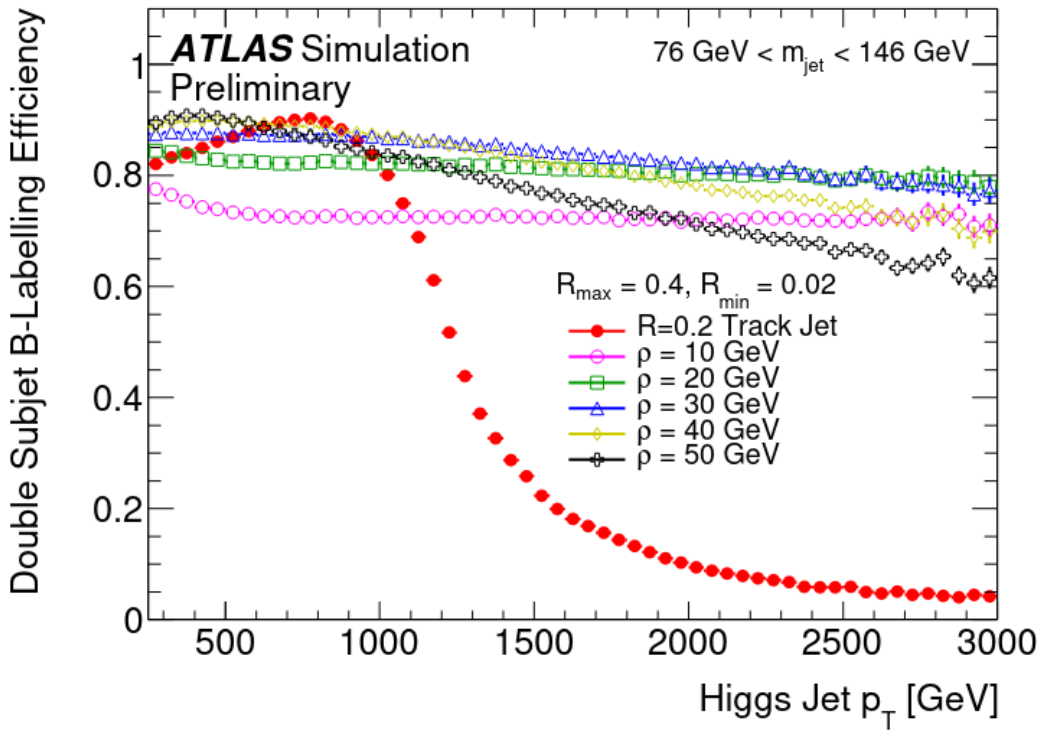


Figure 5.5. Plot of double subjet b-labelling efficiency for various values of ρ for VR track jets.

5.7 Flavor Tagging

Proper identification of b-quarks is a paramount aspect of this analysis as it is conducted as many as four times for each b-quark in the analysis. Identifying which flavor of quark a jet originates from is known as "flavor tagging". Jets deriving from b-quarks are most prominent in this analysis but jets deriving from c-quarks are also important and they both have corresponding jets called "b-jets" and "c-jets" respectively. Any other flavour of quark not including t-quarks that are described by a jet are know as a "light jet". The algorithmic flavor tagging for b-jets specifically is know as "b-tagging" and is directly used throughout this analysis.

B-tagging algorithms exploit the relatively long lifetime, high mass, and high decay multiplicity of b-hadrons to identify them in the ATLAS detector. The mean lifetime of a b-hadron is about 1.5 ps, with the lifetimes of all b-hadrons listed in Table 5.1. This lifetime is relatively "long" compared to particles such as the Higgs boson and closer to others such as c-hadrons. The decay length is determined by

$$L = \beta\gamma c\tau, \tag{5.7}$$

where $\beta = \frac{v}{c}$ and γ is the Lorentz factor. For b-hadrons with $p_T = 50$ GeV, the mean travel distance in the ID is $L_b \approx 5$ mm.

The identification of c-jets is an important part of b-tagging as a c-hadron may be the decay product of or may even falsely resemble a b-hadron. Table 5.2 shows the lifetimes of c-hadrons, which can be as long as 1 ps. The decay length for a $p_T =$ 50 GeV c-jet is $L_b \approx 1 - 3$ mm. So on average it has a shorter decay length, about 1/3 the mass, and lower decay multiplicity than that of the b-hadron, all of which is exploited to perform b-tagging.

The low level b-tagging algorithms have three components: the impact parameter algorithms, secondary vertexing, and decay chain finding (multi-vertex finding):

B-hadron	B_d	B^+	B_s	B_c^+	Λ_b	Ξ_b^0	Ξ_b^-	Ω_b^-
Composition	$\overline{b}d$	$\overline{b}u$	$\overline{b}s$	$\overline{b}c$	udb	usb	dsb	ssb
Mass (GeV)	5.27955	5.27925	5.3667	6.2745	5.6194	5.7878	5.7911	6.0710
Lifetime (ps)	1.519	1.638	1.512	0.500	1.451	1.477	1.599	1.54

Table 5.1. Mass, mean lifetime, and quark composition of b-hadrons.

C-hadron	D^+	D^0	D_s^+	Λ_c^+	Ξ_c^+	Ξ_c^0	Ω_c^0
Composition	cd	$c\overline{u}$	$c\overline{s}$	udc	usb	dsb	ssb
Mass (GeV)	5.27955	5.27925	5.3667	2.28646	2.5756	2.47088	2.6952
Lifetime (ps)	1.040	0.410	0.504	0.200	0.442	0.112	0.268

Table 5.2. Mass, mean lifetime, and quark composition of c-hadrons.

1. **Impact Parameter:** The "IP2D" and "IP3D" are the two impact parameter algorithms used in the first part of b-tagging. The algorithms consider the distance of closest approach in the r-ϕ plane also known respectively as the transverse and longitudinal impact parameters d_0 and z_0. The IP2D algorithm utilizes the impact parameters as d_0/σ_{d_0} while the IP3D algorithm additionally considers $z_0 \sin\theta/\sigma_{z_0 \sin\theta}$. These variables are known as the d_0 and z_0 "significances" and are simulated for b-jet, c-jet and light jets. These values are used to find the log-likelihood of each type of jet.[23]

2. **Secondary Vertex:** The secondary vertex algorithm ("SV1") reconstructs a vertex displaced from the primary vertex ("PV"). First, SV1 collects all possible two-track vertices in the jet. SV1 then iterates over all the tracks forming vertices and runs a fit to output the χ^2 it is associated to secondary vertex. Tracks with the lowest χ^2 are selected and a secondary vertex is reconstructed.

   The reconstructed secondary vertex ("SV") is distinguished by several discriminants, which include the number of tracks associated to the SV, the invariant mass of the SV, the fraction of energy of tracks associated to the SV over all the tracks in the jet, and the significances from the IP3D. Figure 5.6 demonstrates

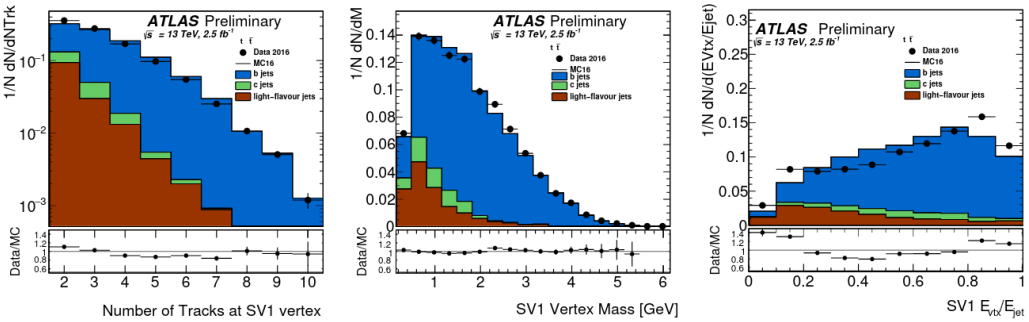


Figure 5.6. Plots of the number of tracks associated to a SV1 vertex, SV1 vertex mass, and the SV1 energy fraction are shown from left to right respectively. Data is plotted alongside $t\bar{t}$ MC in the form of b-jets, c-jets, and light jets.

the discriminating power the SV1 variables have towards SV1 in determining the type of jet tag. The first plot in Figure 5.6 compares data to $t\bar{t}$ MC for the number of tracks of b-jets, c-jets, light jets and data. The last two plots in Figure 5.6 show a similar comparison for SV1 vertex mass and energy ratio.

3. **Multi-vertex Algorithm:** The algorithm that reconstructs vertexes of heavy hadron decays inside a jet is JetFitter ("JF"). It is a modified Kalman filter that exploits the topological structure of weak b- and c-hadron decays inside a jet. It finds a common line that passes through the PV, SV, and tertiary vertex, which is then used as the flight path of the b-hadron.

   The JF has been reoptimized for `Run-2` to account for the effects of pile-up and a mass-dependent vertex selection has been introduced to increase tertiary vertex reconstruction efficiency. In a similar fashion to the SV1 algorithm, the relevant vertex information consists of the number of tracks, invariant mass of the vertex, the energy fraction of the vertex, and the significances from the IP3D algorithm.

The three aforementioned algorithms assist in the implementation of a high level tagger, "DL1r". DL1r is an artificial deep neural network ("NN"), which is trained

using Kera and Theanos. It uses IP2D, IP3D, SV1, and JF (b-tag and c-tag information) as well as jet p_T and η as input variables. DL1r is trained on b-jets, c-jets, and light jets selected from $t\bar{t}$ MC events. The output of DL1r are b-jet, c-jet and light jet probabilities, which can be combined to give a final DL1r score,

$$D_{DL1r} = \ln\left(\frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_{light}}\right), \tag{5.8}$$

where p_b, p_c, and p_{light} are the respective jet probabilities and f_c is the effective c-jet fraction in the training sample. The value for f_c can be tuned independently of NN training. For this analysis, it was optimized to 0.08 (i.e. 8% c-jet fraction in $t\bar{t}$ MC).

The comparison of b-tagging efficiency as a function of light and c-jet rejection is shown in Figure 5.7. High level b-tagging algorithms tune the b-tagging efficiency by cutting the DL1r score at certain values that result in said efficiency. Common cuts as trained on $t\bar{t}$ MC are at 70, 77, and 85% b-tagging efficiency working points ("WP"), of which the 77% WP is used for the analysis as it is considered optimal in most aspects of this analysis.

5.8 Muons

Muons are reconstructed from tracks in the MS, which are then matched to independently reconstructed tracks in the ID. Muons are required to have a $p_T > 4$ GeV, $|\eta| < 2.5$, and pass muon identification quality selection. Of the four muon identifications loose, medium, tight and high-p_T , medium is used for this analysis and is the default for all of ATLAS.[21]

A "muon-in-jet" correction is applied to the Large-R jets to account for energy missed from semileptonic b-hadron decays that result in a muon. If a muon is spatially matched to a track jet that is also b-tagged, the muon is used to correct the Large-R jet. The associated muon four-momentum is added to the four-vector of the

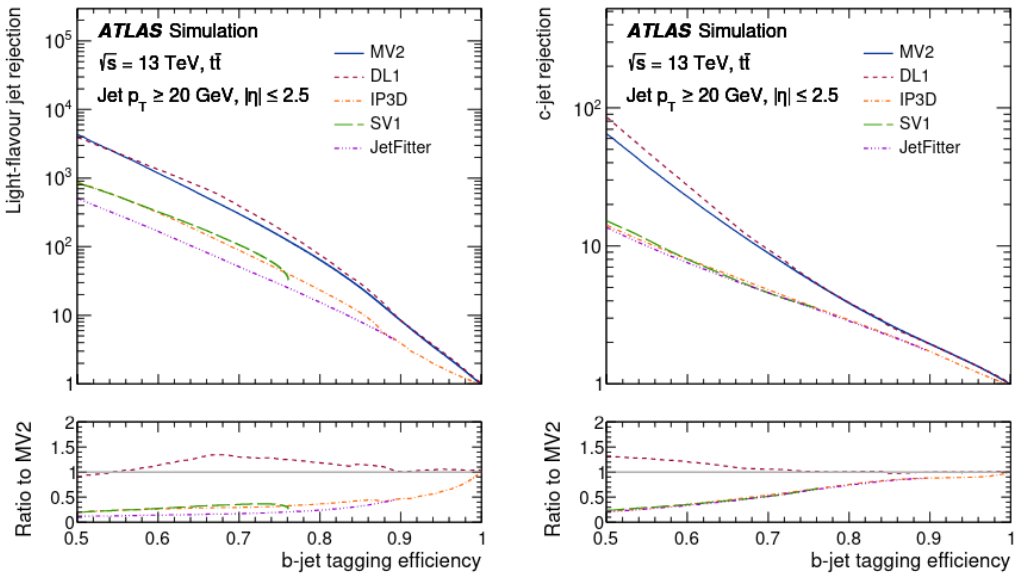


Figure 5.7. Plot of b-tagging efficiency as a function of light and c-jet rejection on the left and right respectively. The three low level b-tagging algorithms, DL1, and the deprecated high level tagger "MV2".

calorimeter information of the Large-R jet and the new m_{calo} is used to recalculate m_{comb} in equation 5.5. The muon-in-jet correction is applied to about 10% of events passing the basic event selection.

CHAPTER 6

EVENT SELECTION

The boosted analysis exists where the b-jets formed from Higgs decays become collimated. For this analysis, at least two Large-R jets are required on which a basic kinematic selection is applied to ensure this topology and to minimize the background. Subsequent selections are applied to remove pathological jets and veto events from the resolved analysis. A control region is defined to allow us to derive a background estimation that is extrapolated into our signal region and a validation region is defined to test the background estimation method.

6.1 Kinematic Selection

A series of kinematic cuts are summarized in Table 6.1. The first cut is on the lead jet large-R p_T , which is optimized to maximize trigger efficiency. The 2015 and 2016 data taking period is fully efficient for $p_T > 420$ GeV as optimized in the partial `Run-2` analysis here [31]. Special consideration is taken into account for the 2017 and 2018 years of data taking as the increased luminosity for this biennium results in greater pile-up. The trigger efficiency is calculated with a specific procedure of comparing the results with a trigger applied to another reference trigger. It is described below in detail:

- Apply a general preselection, which requires an event to have at least one R=1.0 jet prior to any trigger requirement.

- Define a reference sample, which is the 2017 or 2018 data sample with no boosted trigger selection applied and only has the resolved and boosted pre-selection cuts applied to the AOD data samples. The resolved preselection requires four Small-R jets, of which two or more are b-tagged. The boosted preselection requires two or more Large-R jets all with $m > 40$ GeV and one with $p_T > 400$ GeV.

- The boosted trigger has both a p_T and mass cut and therefore both need to be studied. To study the efficiency in leading Large-R jet p_T , cut on a high value of leading Large-R jet mass of 100 GeV. To study the efficiency in leading Large-R jet mass, cut on a high value of leading Large-R jet p_T of 450 GeV.

- Get the reverse cumulative distributions of the reference and nominal samples.

- Calculate the efficiency of the nominal sample to the reference sample – the (reverse) cumulative trigger efficiency.

This procedure is carried out on the 2017 and 2018 data for Large-R p_T and mass, as shown in Figure 4.4 and 4.5 respectively. The trigger efficiency plots demonstrate a cut after which the trigger efficiency approaches 100% for these data sets. To harmonize our method for the boosted analysis, the same cut of ${p_T}^{lead} > 450$ GeV and $m > 50$ GeV is applied to all years of `Run-2` .

Production of $G^*_{\mathtt{KK}}$ is primarily s-channel, which produces a distribution more central in η when compared to multijet background that includes t-channel events and therefore has a less central η distribution. The boosted analysis exploits this to reduce background by cutting on $\Delta\eta \equiv \eta_{J1} - \eta_{J2}$. The Scalar model does not exhibit as pronounced of an s-channel bias as the bulk RS model, thus a $\Delta\eta$ cut will not improve the sensitivity of the Scalar search as much. Improvements to both models are considered when optimizing the cut.

Figure 6.1 shows the values for $|\Delta\eta|$ that optimizes the significance for each mass point of each respective model. Significance is defined as $\frac{S}{\sqrt{B}}$ where S and B are signal and background yields respectively. The cut of $|\Delta\eta| < 1.3$ is chosen as it shows the most overall mutual benefit to both signal models. A m_{JJ} and model dependent cut was also considered but abandoned as the small gains seen were overshadowed by added complications of a more intricate cut.

Cut Name	Cut Details
Trigger	p_T^J (all) and m_J (2017,18)
large-R Jet p_T	$p_T^{lead} > 450\ GeV$ $p_T^{sub} > 250\ GeV$
large-R Jet mass	$m_J > 50\ GeV$
$\lvert\Delta\eta_{JJ}\rvert$	$\lvert\Delta\eta_{JJ}\rvert < 1.3$
Large-R jet $\lvert\eta\rvert$	$\lvert\eta\rvert < 2.0$
track jet $\lvert\eta\rvert$	$\lvert\eta\rvert < 2.5$
track jet p_T	$p_T\ > 5\ GeV$
b-tagging	≥ 1 b-tags

Table 6.1. Summary of the kinematic selection.

The final remaining kinematic selections on Large-R jets require $|\eta| < 2.0$, which requires the Large-R jets to be central in the detector, and require events to have $\geq$ 1 b-tag. Before checking any b-tagging criterion, track jets associated to Large-R jets are p_T ordered and only two track jets per Large-R jet are kept to be b-tagged. The requirement of at least one b-tag is simply to remove unused events as no event used in the boosted analysis beyond this selection have less than one b-tag. Each track jet is required to have a $|\eta| < 2.5$ and a $p_T\ > 5$ GeV, which helps suppress light jets.

6.2 Region Definition

6.2.1 Control, validation, and signal region

The control, validation, and signal regions, known as CR, VR, and SR, respectively, exist in the 2-D Higgs mass m_{H1} - m_{H2} plane. Figure 6.2 depicts the three regions overlaid on the plane. The three regions are defined by three variables,

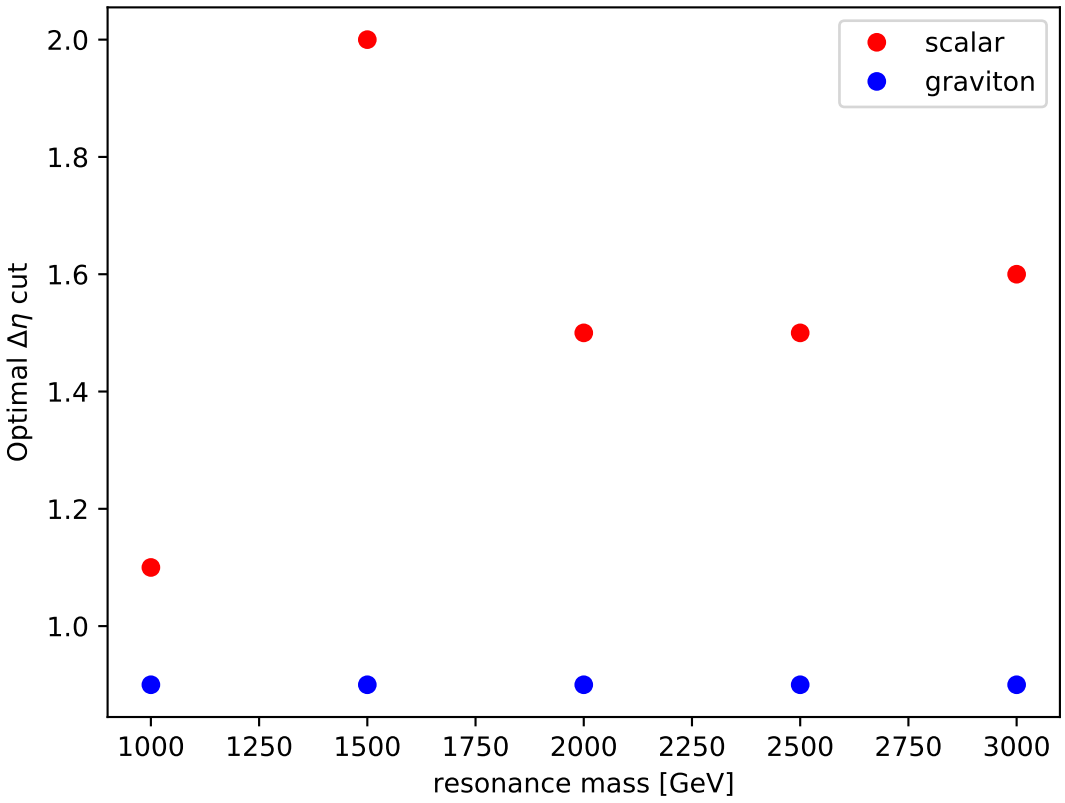


Figure 6.1. $\Delta\eta$ cut that maximizes the significance for each signal mass point.

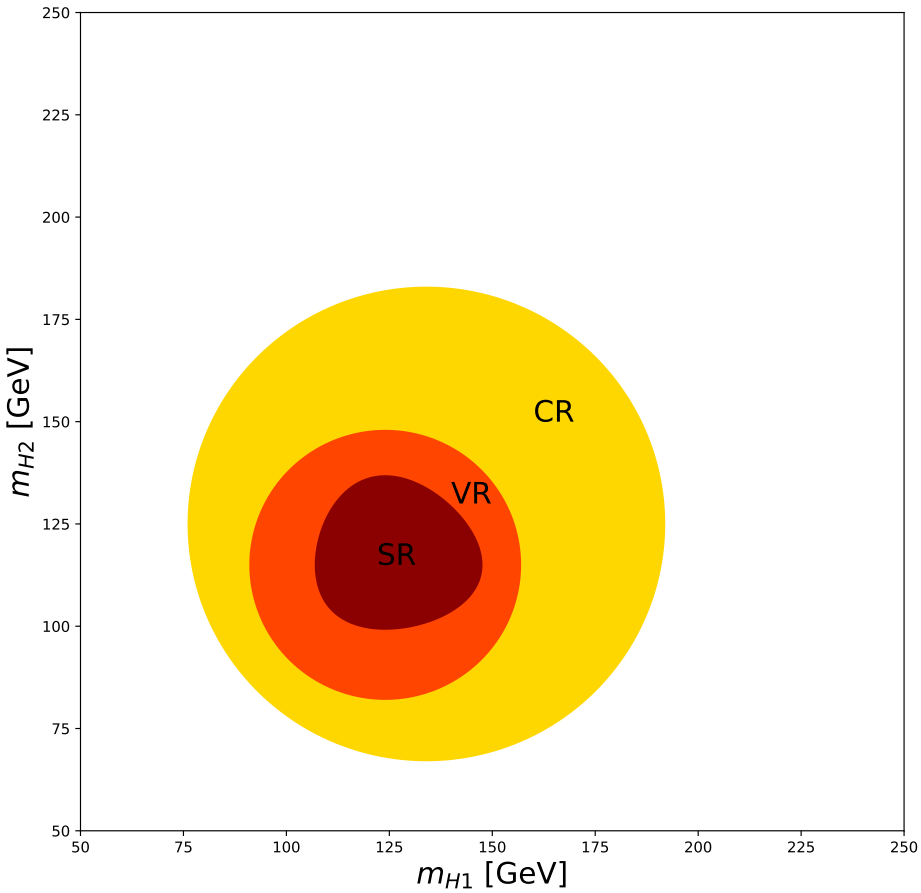


Figure 6.2. The SR, VR, and CR depicted on the m_{H1} - m_{H2} plane in red, orange and yellow respectively.

$$X_{HH} \equiv \sqrt{\left(\frac{m_{H1} - 124\text{ GeV}}{0.1 m_{H1}}\right)^2 + \left(\frac{m_{H2} - 115\text{ GeV}}{0.1 m_{H2}}\right)^2}, \tag{6.1a}$$

$$R_{HH}^{VR} \equiv \sqrt{(m_{H1} - 124\text{ GeV})^2 + (m_{H2} - 115\text{ GeV})^2}, \tag{6.1b}$$

$$R_{HH}^{CR} \equiv \sqrt{(m_{H1} - 134\text{ GeV})^2 + (m_{H2} - 125\text{ GeV})^2}, \tag{6.1c}$$

where $m_{H1(H2)}$ represents the mass of the leading (subleading) Higgs candidate.

Note on naming convention: The variable H_i is used to describe the standard model Higgs boson candidate in this analysis. The index i is 1 or 2, ordered by p_T . Occasionally for older plots when the convention is different, h_i is used but note that these represent the same thing. Before defining the regions on the m_{H1} - m_{H2}, I explicitly call label them by their jet, like J_i. Now that I have defined the regions I will use the H_i convention for the remaining parts of the thesis.

The SR is defined as any event with $X_{HH} < 1.6$, which is meant to keep the most signal while also being consistent with the Higgs boson hypothesis. The denominator is 10% of the mass, which is based on the resolution of the large-R jet masses. The leading Higgs mass is centered around the 124 GeV, which is 1 GeV less than the Higgs boson mass to account for energy losses during reconstruction. A source of energy loss occurs during semileptonic decays where it is only corrected in part with a muon-in-jet correction, but netrinos are not accounted for. The subleading Higgs mass is centered at 115 GeV, which accounts for a larger energy loss typical of subleading jets.

Throughout the development of this analysis, certain areas containing real, possible signal events remain hidden in order to prevent bias especially when performing operations such as background estimation. This process is know as “blinding” and is in kept place until all methods are tested, validated and approved for use on signal events. The SR of data with selection applied is “blinded” until the late stages of

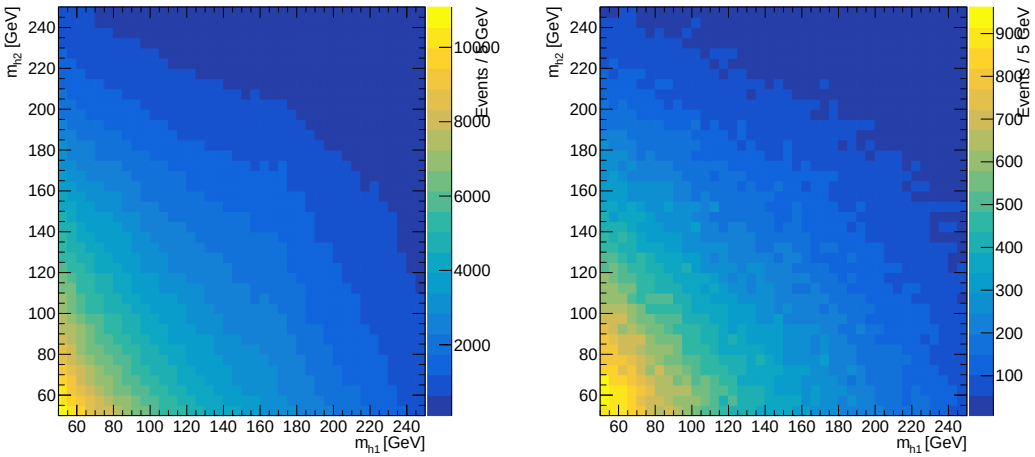


Figure 6.3. The $m_{H1} - m_{H2}$ plane for the 1b-1 and 2b-1 regions on the left and right respectively. The majority of background events are located on the low mass-mass region of the plane.

the analysis. The SR of the MC samples or data without a selection applied is not considered real or distinguishable signal and therefore need not be blinded.

The VR is defined by any event with $X_{HH} > 1.6$ and $R_{HH}^{VR} < 33GeV$. The VR is the closest region to the SR in the $m_{H1} - m_{H2}$ plane but excluded completely from the SR. This provides us a region to test analysis methods, such as background estimation, without unblinding.

The CR is defined by any event with $R_{HH}^{VR} > 33GeV$ and $R_{HH}^{CR} < 58GeV$. This region is the outermost ring/region of the $m_{H1} - m_{H2}$ plane used in the boosted analysis. Notably, the CR definition is shifted upward by 10 GeV and has an expanded radius, which is optimized to increase the number of events in this region while avoiding a background peak at low mass as seen in Figure 6.3. This figure shows the $m_{H1} - m_{H2}$ plane for a data set with a Large-R jet with 1-2 b-tags and the other Large-R jet with no b-tag, which is therefore orthogonal to a signal event and can shown unblinded.

A study was conducted to test the severity of signal contamination, or how many simulated signal events fall outside the signal region, in the VR and CR. Contamination in the VR was found to be within acceptable limits for all signal mass points and virtually zero in the CR. The study is further detailed in Appendix B.1.

6.2.2 Tagging Region Definition

A unique feature of the boosted analysis is the tagging regions, which allow signal events to have four, three or two b-tagged track jets. These regions, called 4b, 3b, and 2b-split, are distinguished from the resolved analysis, which only has one signal region with four b-tagged jets. These extra two regions are included in the boosted analysis to accommodate events that may fail our b-tagging but still contain signal. This situation is common at higher boost as b-tagging efficiency tends to reduce as a function of jet p_T . Figure 6.4 shows the p_T distributions for a Large-R jet with and without a b-tag requirement. The two plots are normalized by total events and the bottom plot represents the efficiency, which demonstrates the drops off of efficacy as a function of p_T . The jet, which contains a b-hadron, may mistakenly not be tagged and thus the event would be discarded if we were only to consider a four tag region as our signal region.

The 4b, 3b, and 2b-split regions have a specific topology, which are visualized in Figure 6.5. For each Large-R jet, only the two leading p_T track jets are considered when b-tagging. The three plots on the left are considered part of the boosted signal region but are also know as the "high-tag" regions. The exact logic definition for the high-tag region is defined in Table 6.2: The 4b region requires 2 b-tagged track jets in each Large-R jet; the 3b region requires 2 b-tagged track jets in one Large-R jet and 1 b-tagged track jet in the other Large-R jet; the 2b-split region requires 1 b-tagged track jet in each Large-R jet.

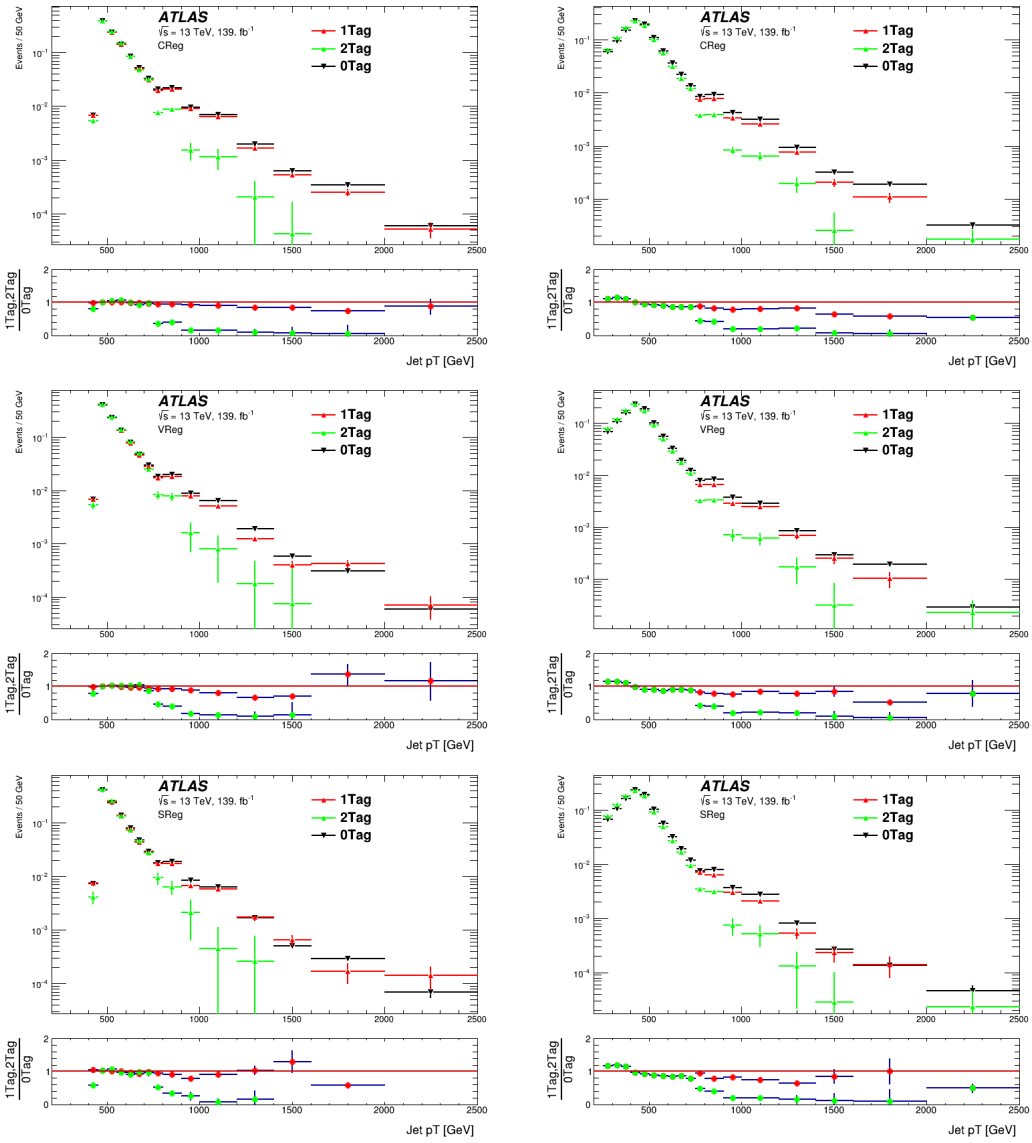


Figure 6.4. Distributions for Large-R jet p_T for with and without b-tagging. The CR, VR, and SR regions are shown on the top left, top right, and bottom respectively. The top distributions are normalized by event count and the bottom plot represents efficiency, which falls off as a function of p_T for all three regions.

Juxtaposed against the high-tag region is a "low-tag" region, which is shown on the right side of Figure 6.5. The low-tag region is used during our background estimation, which is detailed in Chapter 7. The three low-tag regions are 2b-2, 2b-1, and 1b-1, which are used model background for the 4b, 3b, and 2b-split regions respectively. The exact logic definition for the low-tag region is defined in Table 6.2: the 2b-2 region requires 2 b-tagged track jets in one Large-R jet and at least 2 track-jets in the other Large-R jet that are not b-tagged; the 2b-1 region requires 2 b-tagged track jets in one Large-R jet and at least 1 track-jet in the other Large-R jet that is not b-tagged; the 1b-1 region requires 1 b-tagged track jet in one Large-R jet and at least 1 track-jet in the other Large-R jet that is not b-tagged.

The low-tag region has a looser selection in track jets that are not b-tagged in order to increase our statistics to improve the background estimation. By allowing the 2b-1 region to have 1 Large-R jet with 1 *or more* track-jets with no b-tag, the statistics on the 2b-1 region is multiplied but an order of magnitude. This introduces an overlap of shared events between the 2b-2 and 2b-1 regions. To ameliorate this, "background sharing" is introduced to split up the overlapping 2b-2 events. The background sharing gives 80% of overlapping events to the 2b-1 region and the remaining 20% to the 2b-2 region. This "80/20" share is randomly selected from the shared events. The "80/20" share is the optimal value for the boosted analysis and maintains statistical independence between the two regions. The full study of background sharing is explained in detail in Appendix D.1.

6.3 NTrk

An additional cut on the number of charged-particle tracks ghost-associated to an untrimmed Large-R jet is heavily considered for the analysis. This value, known as "NTrk", represents the multiplicity of these tracks and is higher for gluon initiated jets. The analysis background, which is primarily multijet QCD, will result in more

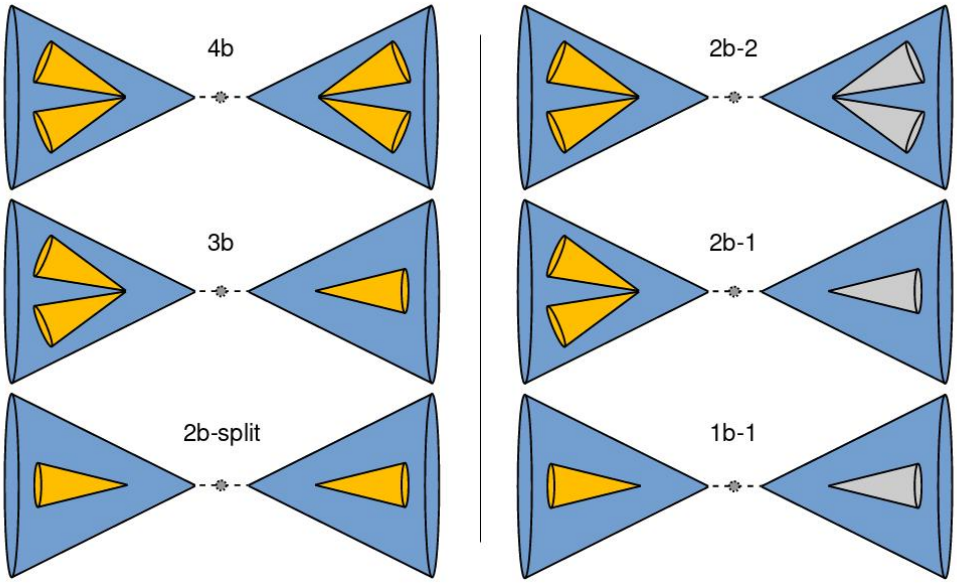


Figure 6.5. The six tagging regions for the boosted analysis. The three high-tag regions are on the left and the three low-tag regions are on the right. The blue cones represent a Large-R jet, the golden cone represents a b-tagged track jet, and the grey cone represents a track jet that is not b-tagged.

high-tag	4b	3b	2b-split
	2 b-tags	2 b-tags	1 b-tag
	2 b-tags	1 b-tag	1 b-tag
low-tag	2b-2	2b-1	1b-1
	2 b-tags	2 b-tags	1 b-tag
	$\geq$ 2 track-jets	$\geq$ 1 track-jets	1 track-jet

Table 6.2. The six tagging regions have a tagging requirement on each of the Large-R jets. To fall into a tagging region listed, one Large-R jet has to pass one of the two listed criterion while the other Large-R jet must pass the other remaining criteria.

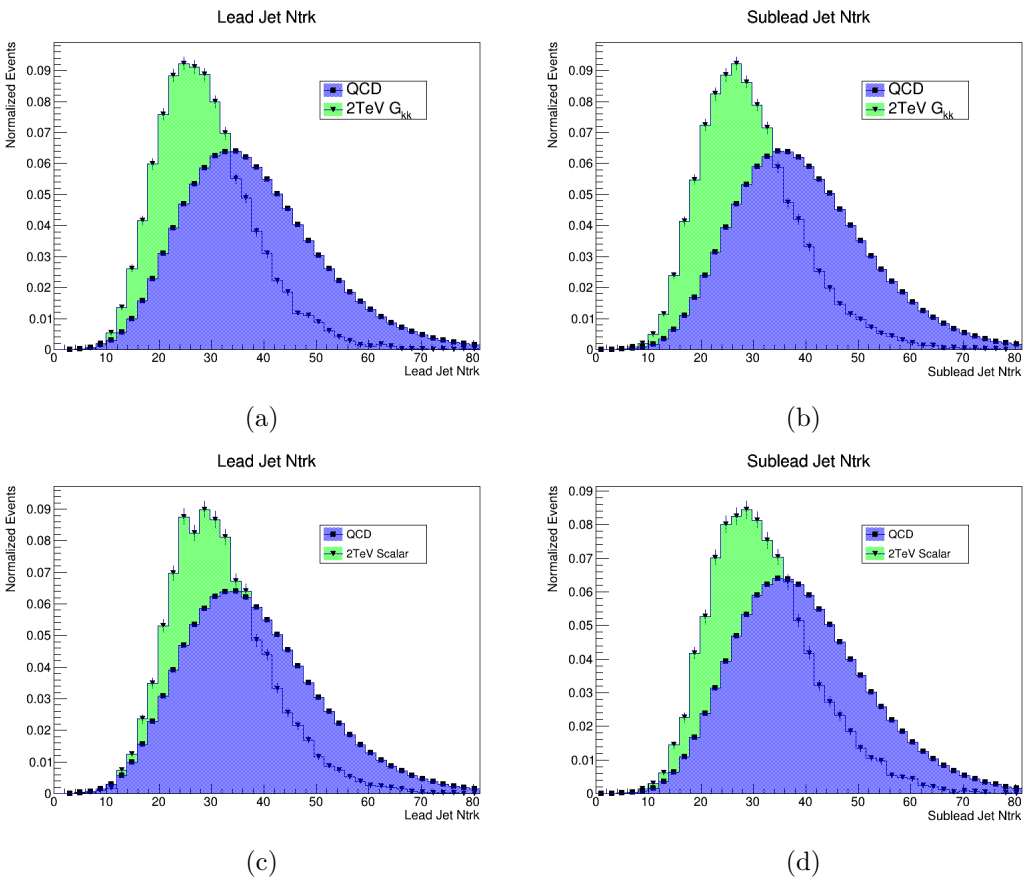


Figure 6.6. NTrk distribution for a 2 TeV G_{kk} (a, b) and Scalar (c, d) compared to QCD MC. Leading jet and subleading jet NTrk distributions are shown on the left and right, respectively.

gluon initiated jets when compared to our signal models that have quark initiated jets. Thus, the background has a higher NTrk value than our signal models as seen in Figure 6.6. NTrk provides a useful discriminant to remove background events without removing many signal events and therefore improves the sensitivity of the boosted analysis.

An NTrk cut is optimized for the boosted analysis by maximizing a sensitivity factor, which is defined by,

$$A = \frac{\epsilon}{\sqrt{B_{NTrk}}}, \tag{6.2}$$

where ϵ (cut efficiency) is the ratio of the signal yields with an NTrk cut divided by a signal yield without an NTrk cut ($\frac{S_{NTrk}}{S_{noNTrk}}$). This has the benefit of making this optimization procedure independent of signal cross section. B_{NTrk} is the background yield with an NTrk cut. The sensitivity factor is maximized for both leading and subleading jets, the procedure of which is outlined in detail in Appendix C.1.3.

Figure 6.7 shows the NTrk cut that maximizes the sensitivity factor as a function of signal mass. These values, however, differ between the $G^*_{\mathtt{KK}}$ and Scalar models, simply because the two samples use different parton showering (PS) generators during simulation. The $G^*_{\mathtt{KK}}$ model is simulated with Pythia 8 with EvtGen, which models NTrk well, but the Scalar model is simulated with Herwig 7, which does not simulate NTrk well. Upon further inspection of Figure 6.6, it is apparent the distribution of NTrk for the Scalar is shifted to a higher value of NTrk, which lessens the difference in the value of NTrk between signal and background and thus lessens the discrimination power of an NTrk cut for this model.

An optimization is performed for both models separately, which gives the maximal value of the sensitivity factor as a function of signal mass. A fit is performed to find a continuous functional cut to be applied to the boosted analysis. Several fits are tried and an optimization is run on both models independently. Ultimately, the cut that adds the greatest overall improvement to the analysis is shown in Figure 6.7.

Running the full analysis with NTrk showed an exceptional improvement to our sensitivity anywhere from 10-30% for the $G^*_{\mathtt{KK}}$ model. However, the PS differences are apparent for the Scalar which shows an improvement of 0-15%. After systematic uncertainties are applied for NTrk and to account for the PS differences in the scope of NTrk, most improvements from the NTrk cut to the Scalar model are negated and improvements from NTrk to the $G^*_{\mathtt{KK}}$ model are halved.

Ultimately, a decision is made to remove the NTrk cut from the final analysis despite the promising results demonstrated and having it fully integrated into

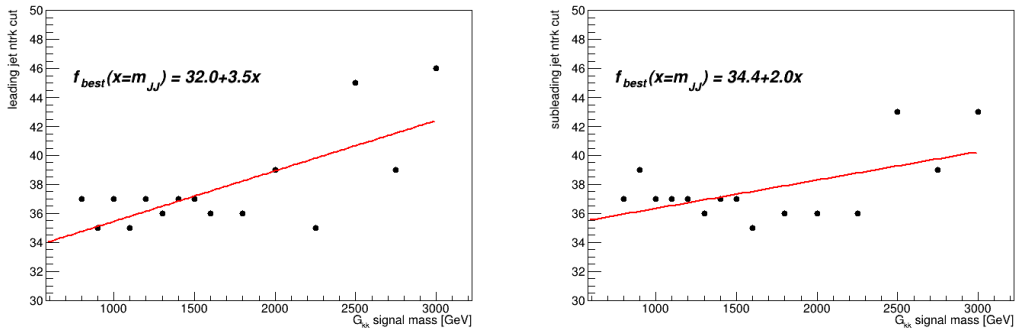


Figure 6.7. Optimized NTrk cut values for leading (left) and subleading (right) jets. All equations shown use m_{HH} in units of TeV.

the boosted analysis. Although PYTHIA 8 Scalar samples are available, HERWIG 7 was chosen by the diHiggs group some time prior to the introduction of NTrk to the boosted analysis. Thus, the NTrk cut is removed from the analysis in favor of harmonization with the other diHiggs groups, which is especially important when considering a combined diHiggs analysis. This is explained in further detail at the end of Appendix C.1.

6.4 Event Vetoes

6.4.1 Resolved Veto

The first event veto in the boosted analysis is meant to maintain orthogonality with the resolved analysis. Priority in overlapping events is given to the resolved analysis partially because of historical precedence but also because the VBF and non-resonant analysis is most similar to the resolved analysis. The veto procedure is to flag events in the boosted analysis that pass the signal region as defined in the resolved analysis by,

$$1.6 > \sqrt{\left(\frac{m_{H1} - 120\ GeV}{0.1 m_{H1}}\right)^2 + \left(\frac{m_{H2} - 110\ GeV}{0.1 m_{H2}}\right)^2}, \tag{6.3}$$

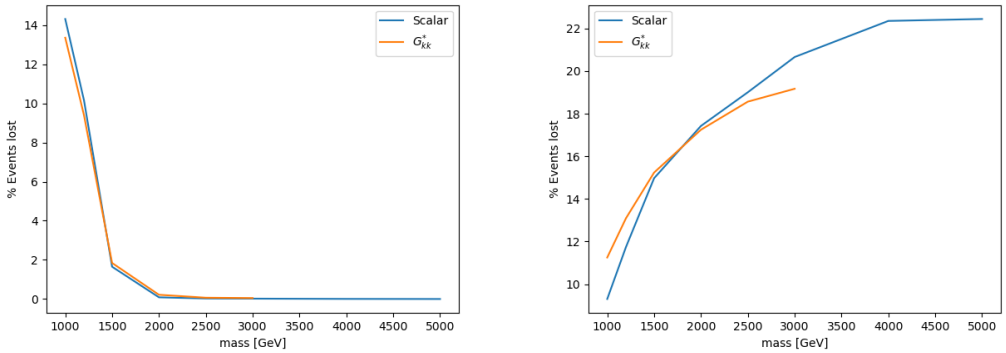


Figure 6.8. Rejection rate of the resolved veto and collinear track jet veto on the left and right respectively. The resolved veto rejection rate is found after the kinematic selection is applied and the collinear track jet veto rejection rate is found after the kinematic selection and resolved veto are applied.

where $m_{H1(H2)}$ is the mass of the resolved leading (subleading) Higgs candidate. Higgs candidates are reconstructed in the Resolved analysis by pairing two small-R jets together using a boosted decision tree (BDT) pairing. This pairing algorithm takes into account three variables: the spacial separation in R, the rapidity separation in η, and the angular separation in ϕ of the two jets. The 4-vector of the jet pair is combined to form a Higgs candidate.

The SR flag is combined with the requirement to have four or more b-tagged resolved jets, as signal events in the resolved analysis are expected to have exactly four b-tags. Similarly to the boosted analysis, the resolved analysis applies b-tagging at a 77% fixed cut working point using the DL1r flavour tagging algorithm.

Any event flagged for both the resolved SR and tagging criteria is removed from the boosted analysis. Figure 6.8 demonstrates how this cut mostly impacts low m_{HH}as the resolved analysis is most sensitive here and therefore most heavily overlaps with the boosted analysis. The impact of the resolved veto falls off to zero above 2 TeV, as we expect the resolved analysis to have no sensitivity above that point.

6.4.2 Collinear Track jet Veto

The mechanism of variable radius track jets, described in Section 5.6, is to adjust the radius of the jet as p_T increases. As opposed to fixed radius track jets, variable radius track jets allow us to improve our sensitivity especially for highly boosted jets, which become increasingly collimated. It is possible however for the jet axis of a high p_T jet to be contained within a lower p_T jet ("collinear"). The flavour tagging algorithm may incorrectly associate some of the tracks from the high p_T jet to that of the lower p_T one. This outcome is problematic to flavour tagging performance because flavour tagging relies heavily on track reconstruction to find the secondary vertex. Thus, this scenario is considered a pathological case and must be removed from use from any analysis using this track jet flavour tagging strategy, which includes the boosted analysis.

The "collinear track jet veto" is designed to remove the events with the pathological jet axis overlap. The criteria for this veto, described in Table 6.4, checks if the track jet axis falls within the radius of another track jet by,

$$\Delta R(j_i, j_j) < \min(Rad_{j_i}, Rad_{j_j}), \tag{6.4}$$

where $\Delta R(j_i, j_j)$ is the distance in R between the the two track jets and Rad_j represents the radius the track jet. The impact of the collinear track jet veto is shown in the right plot of Figure 6.8. The removal of events increases as a function of mass.

6.5 Cutflows and Efficiencies

The acceptance times efficiency values for the several selections are shown for several signal mass samples shown in Figure 6.9 for the bulk RS model and Scalar model on the left and right respectively. The value is found by dividing the event yield after each cut by the number of initial events. In general, the boosted analysis is most efficient between 1.5 and 2 TeV, where the acceptance times efficiency peaks

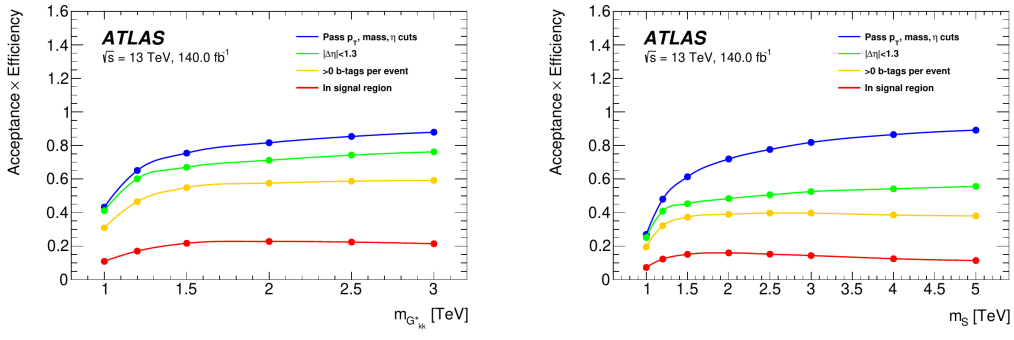


Figure 6.9. Acceptance times efficiency as a function of signal resonance mass. The bulk RS and Scalar models are on the left and right respectively.

at 22-23% and 15-16% for $G^*_{\mathtt{KK}}$ and Scalar respectively. The boosted analysis is least efficient at the lowest mass 900 GeV, where the acceptance times efficiency drops to 11% and 7% for the $G^*_{\mathtt{KK}}$ and Scalar respectively.

Similarly, the acceptance time efficiency for each tagging region is shown in Figure 6.10. The most sensitive region of the boosted analysis varies from 4b, to 3b, to 2b-split as a function of increasing signal mass. The higher mass resonances tend to have higher p_T , where b-tagging becomes less efficient and thus the 3b and 2b-split become more sensitive as mass increases.

The selection efficiency for each cut, or cutflow, is shown for several signal mass values in Tables 6.3, 6.4, and 6.5 for $G^*_{\mathtt{KK}}$ and Tables 6.6, 6.7, and 6.8 for Scalar samples. The event yield is expected to decrease as a function of signal mass as the signal production cross-section exponentially decreases as a function of signal mass.

	G^*_{KK} m = 1.5TeV	G^*_{KK} m = 2TeV	G^*_{KK} m = 3TeV
nInitialEventsAOD	68000.0	23333.33	45625.0
nInitialEvents	39.48	6.5	0.33
PassTrig	39.01	6.48	0.33
PassTwoFatJets	33.1	5.81	0.32
PassFatJetMass	32.57	5.73	0.31
PassFatJetPt	32.15	5.72	0.31
PassDijetEta	32.15	5.72	0.31
PassDetaHH	28.57	4.99	0.27
PassNTrk_lead	28.57	4.99	0.27
PassNTrk_sub	28.57	4.99	0.27
PassResVeto	27.98	4.98	0.27
PassCollinearVR	23.68	4.1	0.22
PassBJetSkim	23.32	4.01	0.21
PassSignal	9.17	1.58	0.08
PassSignal_min2bs	7.65	1.26	0.05
4b	4.24	0.59	0.02
3b	9.03	1.47	0.06
2b_Split	4.82	0.93	0.06
4b_bkgModel	0.5	0.08	0.0
3b_bkgModel	1.99	0.36	0.02
2bs_bkgModel	2.74	0.57	0.05
CReg_4b	0.44	0.05	0.0
CReg_3b	1.16	0.16	0.01
CReg_2b_Split	0.7	0.12	0.01
CReg_4b_bkgModel	0.07	0.01	0.0
CReg_3b_bkgModel	0.31	0.05	0.0
CReg_2bs_bkgModel	0.43	0.08	0.01
VReg_4b	1.07	0.15	0.0
VReg_3b	2.23	0.37	0.02
VReg_2b_Split	1.12	0.22	0.02
VReg_4b_bkgModel	0.11	0.02	0.0
VReg_3b_bkgModel	0.42	0.08	0.0
VReg_2bs_bkgModel	0.52	0.12	0.01
SReg_4b	2.23	0.31	0.01
SReg_3b	3.81	0.63	0.03
SReg_2b_Split	1.6	0.32	0.02
SReg_4b_bkgModel	0.16	0.03	0.0
SReg_3b_bkgModel	0.64	0.12	0.01
SReg_2bs_bkgModel	0.67	0.15	0.01

Table 6.3. Year 2015-2016 cutflow for three G^*_{KK} samples.

	$G^*_{\mathtt{KK}}$ m = 1.5TeV	$G^*_{\mathtt{KK}}$ m = 2TeV	$G^*_{\mathtt{KK}}$ m = 3TeV
nInitialEventsAOD	63400.0	57888.89	52948.72
nInitialEvents	48.52	7.99	0.41
PassTrig	47.38	7.93	0.41
PassTwoFatJets	40.58	7.12	0.39
PassFatJetMass	39.96	7.02	0.39
PassFatJetPt	39.55	7.01	0.39
PassDijetEta	39.55	7.01	0.39
PassDetaHH	35.18	6.12	0.33
PassNTrk_lead	35.18	6.12	0.33
PassNTrk_sub	35.18	6.12	0.33
PassResVeto	34.56	6.11	0.33
PassCollinearVR	29.3	5.05	0.27
PassBJetSkim	28.77	4.93	0.26
PassSignal	11.35	1.95	0.09
PassSignal_min2bs	9.41	1.55	0.07
4b	4.98	0.7	0.02
3b	11.25	1.78	0.07
2b_Split	5.99	1.16	0.08
4b_bkgModel	0.61	0.11	0.01
3b_bkgModel	2.44	0.44	0.02
2bs_bkgModel	3.49	0.74	0.06
CReg_4b	0.53	0.06	0.0
CReg_3b	1.46	0.2	0.01
CReg_2b_Split	0.88	0.14	0.01
CReg_4b_bkgModel	0.09	0.01	0.0
CReg_3b_bkgModel	0.37	0.06	0.0
CReg_2bs_bkgModel	0.53	0.1	0.01
VReg_4b	1.27	0.17	0.01
VReg_3b	2.78	0.43	0.02
VReg_2b_Split	1.38	0.28	0.02
VReg_4b_bkgModel	0.12	0.02	0.0
VReg_3b_bkgModel	0.5	0.1	0.01
VReg_2bs_bkgModel	0.72	0.15	0.01
SReg_4b	2.59	0.36	0.01
SReg_3b	4.75	0.77	0.03
SReg_2b_Split	2.07	0.41	0.03
SReg_4b_bkgModel	0.19	0.04	0.0
SReg_3b_bkgModel	0.79	0.15	0.01
SReg_2bs_bkgModel	0.87	0.2	0.02

Table 6.4. Year 2017 cutflow for three $G^*_{\mathtt{KK}}$ samples.

	$G^*_{\mathtt{KK}}$ m = 1.5TeV	$G^*_{\mathtt{KK}}$ m = 2TeV	$G^*_{\mathtt{KK}}$ m = 3TeV
nInitialEventsAOD	28000.0	30000.0	30000.0
nInitialEvents	63.9	10.53	0.54
PassTrig	62.48	10.46	0.54
PassTwoFatJets	53.32	9.41	0.52
PassFatJetMass	52.52	9.28	0.51
PassFatJetPt	52.0	9.26	0.51
PassDijetEta	52.0	9.26	0.51
PassDetaHH	46.24	8.1	0.44
PassNTrk_lead	46.24	8.1	0.44
PassNTrk_sub	46.24	8.1	0.44
PassResVeto	45.42	8.08	0.44
PassCollinearVR	38.54	6.72	0.36
PassBJetSkim	37.85	6.56	0.34
PassSignal	15.05	2.63	0.12
PassSignal_min2bs	12.39	2.05	0.08
4b	6.51	0.88	0.02
3b	14.22	2.37	0.1
2b_Split	8.18	1.54	0.1
4b_bkgModel	0.91	0.15	0.01
3b_bkgModel	3.26	0.6	0.03
2bs_bkgModel	4.77	1.03	0.09
CReg_4b	0.65	0.08	0.0
CReg_3b	1.79	0.25	0.01
CReg_2b_Split	1.2	0.21	0.01
CReg_4b_bkgModel	0.17	0.02	0.0
CReg_3b_bkgModel	0.49	0.08	0.0
CReg_2bs_bkgModel	0.76	0.14	0.01
VReg_4b	1.65	0.22	0.01
VReg_3b	3.45	0.57	0.03
VReg_2b_Split	1.86	0.36	0.03
VReg_4b_bkgModel	0.16	0.03	0.0
VReg_3b_bkgModel	0.71	0.13	0.01
VReg_2bs_bkgModel	0.96	0.22	0.02
SReg_4b	3.38	0.45	0.01
SReg_3b	6.11	1.06	0.04
SReg_2b_Split	2.9	0.54	0.03
SReg_4b_bkgModel	0.3	0.05	0.0
SReg_3b_bkgModel	1.05	0.22	0.01
SReg_2bs_bkgModel	1.21	0.27	0.03

Table 6.5. Year 2018 cutflow for three $G^*_{\mathtt{KK}}$ samples.

	scalar m = 1.5TeV	scalar m = 2TeV	scalar m = 3TeV
nInitialEventsAOD	2438.0	232.48	5.17
nInitialEvents	1311.91	127.49	2.88
PassTrig	1189.89	122.0	2.85
PassTwoFatJets	957.91	104.37	2.61
PassFatJetMass	946.24	102.96	2.58
PassFatJetPt	906.18	101.51	2.58
PassDijetEta	906.18	101.51	2.58
PassDetaHH	662.53	68.27	1.65
PassNTrk_lead	662.53	68.27	1.65
PassNTrk_sub	662.53	68.27	1.65
PassResVeto	650.24	68.18	1.65
PassCollinearVR	551.38	56.08	1.31
PassBJetSkim	542.85	54.87	1.25
PassSignal	221.1	22.48	0.45
PassSignal_min2bs	183.26	17.77	0.31
4b	102.35	8.13	0.1
3b	209.32	19.93	0.37
2b_Split	111.14	12.62	0.35
4b_bkgModel	11.3	1.23	0.03
3b_bkgModel	46.36	5.52	0.12
2bs_bkgModel	62.38	7.44	0.28
CReg_4b	11.17	0.71	0.01
CReg_3b	24.79	1.91	0.04
CReg_2b_Split	16.08	1.3	0.05
CReg_4b_bkgModel	1.45	0.16	0.0
CReg_3b_bkgModel	7.18	0.77	0.02
CReg_2bs_bkgModel	9.73	1.04	0.04
VReg_4b	24.43	1.81	0.02
VReg_3b	48.36	4.83	0.1
VReg_2b_Split	25.33	2.73	0.09
VReg_4b_bkgModel	2.39	0.26	0.01
VReg_3b_bkgModel	9.06	1.2	0.03
VReg_2bs_bkgModel	11.4	1.28	0.06
SReg_4b	53.61	4.28	0.04
SReg_3b	91.08	8.79	0.15
SReg_2b_Split	38.57	4.7	0.12
SReg_4b_bkgModel	3.33	0.43	0.01
SReg_3b_bkgModel	15.97	1.88	0.04
SReg_2bs_bkgModel	16.77	2.13	0.08

Table 6.6. Year 2015-2016 cutflow for three Scalar samples.

	scalar m = 1.5TeV	scalar m = 2TeV	scalar m = 3TeV
nInitialEventsAOD	2439.5	232.64	5.17
nInitialEvents	1585.42	154.82	3.52
PassTrig	1418.62	147.78	3.46
PassTwoFatJets	1155.87	127.19	3.18
PassFatJetMass	1140.2	125.38	3.14
PassFatJetPt	1100.17	123.95	3.13
PassDijetEta	1100.17	123.95	3.13
PassDetaHH	819.31	83.13	2.03
PassNTrk_lead	819.31	83.13	2.03
PassNTrk_sub	819.31	83.13	2.03
PassResVeto	806.55	83.06	2.03
PassCollinearVR	682.74	68.91	1.61
PassBJetSkim	672.18	67.36	1.54
PassSignal	277.65	27.41	0.56
PassSignal_min2bs	231.08	21.51	0.39
4b	125.35	9.92	0.11
3b	266.68	24.85	0.46
2b_Split	131.36	14.55	0.44
4b_bkgModel	13.73	1.62	0.03
3b_bkgModel	59.16	6.57	0.14
2bs_bkgModel	75.89	9.85	0.36
CReg_4b	12.48	0.82	0.01
CReg_3b	33.86	2.54	0.05
CReg_2b_Split	14.69	1.89	0.05
CReg_4b_bkgModel	1.83	0.2	0.0
CReg_3b_bkgModel	7.9	0.76	0.02
CReg_2bs_bkgModel	12.88	1.17	0.05
VReg_4b	30.15	2.14	0.03
VReg_3b	65.13	5.83	0.11
VReg_2b_Split	30.14	3.23	0.11
VReg_4b_bkgModel	2.64	0.39	0.01
VReg_3b_bkgModel	11.81	1.3	0.03
VReg_2bs_bkgModel	15.24	1.99	0.08
SReg_4b	66.17	5.48	0.05
SReg_3b	114.65	10.67	0.18
SReg_2b_Split	50.26	5.36	0.16
SReg_4b_bkgModel	4.43	0.5	0.01
SReg_3b_bkgModel	20.35	2.3	0.04
SReg_2bs_bkgModel	19.61	2.75	0.1

Table 6.7. Year 2017 cutflow for three Scalar samples.

	scalar m = 1.5TeV	scalar m = 2TeV	scalar m = 3TeV
nInitialEventsAOD	3659.2	348.95	7.5
nInitialEvents	2107.17	204.82	4.64
PassTrig	1876.18	194.51	4.56
PassTwoFatJets	1529.63	167.3	4.18
PassFatJetMass	1507.42	164.97	4.14
PassFatJetPt	1458.94	163.1	4.13
PassDijetEta	1458.94	163.1	4.13
PassDetaHH	1076.76	109.71	2.65
PassNTrk_lead	1076.76	109.71	2.65
PassNTrk_sub	1076.76	109.71	2.65
PassResVeto	1059.62	109.65	2.65
PassCollinearVR	905.57	90.44	2.11
PassBJetSkim	889.2	88.31	2.0
PassSignal	356.34	36.45	0.71
PassSignal_min2bs	290.81	28.64	0.48
4b	161.03	12.72	0.14
3b	347.29	31.76	0.55
2b_Split	180.61	20.07	0.57
4b_bkgModel	20.7	2.11	0.05
3b_bkgModel	75.55	8.74	0.19
2bs_bkgModel	104.02	12.92	0.49
CReg_4b	18.25	1.16	0.01
CReg_3b	45.36	3.52	0.07
CReg_2b_Split	26.44	2.66	0.07
CReg_4b_bkgModel	3.26	0.29	0.01
CReg_3b_bkgModel	10.86	1.27	0.02
CReg_2bs_bkgModel	14.81	1.83	0.07
VReg_4b	42.11	3.04	0.03
VReg_3b	86.68	7.35	0.14
VReg_2b_Split	39.73	4.5	0.14
VReg_4b_bkgModel	4.95	0.43	0.01
VReg_3b_bkgModel	16.17	2.05	0.04
VReg_2bs_bkgModel	20.44	2.32	0.1
SReg_4b	79.52	6.47	0.07
SReg_3b	146.98	14.41	0.22
SReg_2b_Split	64.31	7.76	0.19
SReg_4b_bkgModel	7.25	0.69	0.01
SReg_3b_bkgModel	26.62	2.75	0.06
SReg_2bs_bkgModel	28.33	3.76	0.14

Table 6.8. Year 2018 cutflow for three Scalar samples.

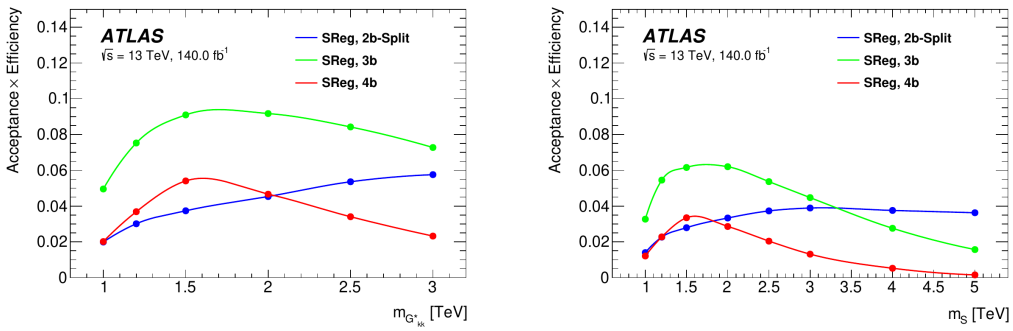


Figure 6.10. Acceptance times efficiency as a function of signal resonance mass for each b-tagging region. The b-tagging criteria are applied after the basic selection described in the beginning of this section and include a signal region cut.

CHAPTER 7

BACKGROUND ESTIMATION

The dominant source of background in the boosted analysis is from multijet QCD processes, such as gluon splitting, and the secondary background source is from production of pairs of top quarks "$t\bar{t}$ ". Other background sources include Z+jets and $ZZ \longrightarrow b\bar{b}b\bar{b}$, which account for $< 1\%$ of the total background and are therefore neglected in the background estimation. The exact composition of the background, shown in Table 7.1, varies depending on the tagging region and where the event lies in the $m_{H1} - m_{H2}$ plane. For the SR, multijet QCD makes up about 90-70% of background events, where 4b has the most multijet QCD and 2b-split has the least.

channel	**region**	**QCD multi-jet**	$t\bar{t}$	**total**	**% QCD**	**% $t\bar{t}$**
4 b-jets	CR	380	70	450	84	16
3 b-jets	CR	7000	2254	9300	76	24
2 b-jets (split)	CR	32000	25000	57000	56	44
4 b-jets	VR	160	15	170	91	9.0
3 b-jets	VR	2900	440	3300	87	13
2 b-jets (split)	VR	13000	5100	18000	72	28
4 b-jets	SR	75	7.0	82	92	9.0
3 b-jets	SR	1400	250	1600	85	15
2 b-jets (split)	SR	6400	2800	9300	70	31

Table 7.1. Estimated multijet QCD, $t\bar{t}$, and total background events and the fraction of QCD multi-jet and $t\bar{t}$ to the total background for each channel and region. This table is obtained from the background estimation, which uses the full data and simulated $t\bar{t}$ samples.

A data-driven method is used to model the shape and normalization of the m_{HH} distribution for the multijet QCD background instead of using simulated multijet QCD samples as it is difficult to model this background source through simulation. The final discriminant m_{HH} is used as it represents the Higgs boson pair and will be what is ultimately used in the statistical analysis. This shape is derived in the CR, tested in the VR, and then applied to the SR. The low-tag regions are used to derive the shape for the high-tag regions, both of which are defined in Section 6.2.2. The background in the 2b-split region is estimated using 1b-1 events, the background in

the 3b region is estimated using 2b-1 events, and the background in the 4b region is estimated using the 2b-2 events.

7.1 Derivation of Background Estimation Scale Factors

The background estimation is performed in the CR, where separate scale factors are derived for the two prominent sources of background. These scale factors are labeled as μ_{QCD} and $\alpha_{t\bar{t}}$ for multijet QCD and $t\bar{t}$, respectively. The scale factor μ_{QCD} is essentially an estimate of the ratio of high-tag to low-tag multijet QCD events whereas $\alpha_{t\bar{t}}$ is a normalization correction to account for differences between MC and data. Together, they are used to estimate the number of events in the high-tag region from the number of events in the low-tag region as:

$$HT_{bkg} = \mu_{QCD}(LT_{data} - LT_{t\bar{t}}) + \alpha_{t\bar{t}}\, HT_{t\bar{t}}\,, \tag{7.1}$$

where HT and LT represent the numbers of high-tag and low-tag events, respectively. The $t\bar{t}$ HT and LT events are from simulated $t\bar{t}$ samples.

With the values for μ_{QCD} and $\alpha_{t\bar{t}}$ derived in the CR, Equation 7.1 can then be applied in any of the CR, VR, or SR regions to estimate the background. This is only valid if the values for μ_{QCD} and $\alpha_{t\bar{t}}$ remain near constant across the $m_{H1}-m_{H2}$ plane. This proves to be mostly constant and any difference in μ_{QCD} along the $m_{H1} - m_{H2}$ plane are accounted for with an extrapolation uncertainty, which is outlined in Section 8.2.1.

A binned maximum likelihood fit is used to find the scale factors μ_{QCD} and $\alpha_{t\bar{t}}$ for all three tagging regions. Figure 7.1 shows the m_{H1} distributions for data, multijet QCD, and $t\bar{t}$ used to find μ_{QCD} and $\alpha_{t\bar{t}}$. The final scale factor values are found in Table 7.2. The $t\bar{t}$ sample lacks enough events in the 4b region to perform a proper fit and therefore $\alpha_{t\bar{t}}$ is fixed to 1.0 for this region. The two peak feature of the m_{H1}

multijet QCD distributions is an artifact of CR. This is caused by the exclusion of VR and SR events from the CR and thus giving it peaks at low and high mass. The higher mass of *top* quarks present in $t\bar{t}$ events results in a low-mass peak that is smaller than the one seen in the multijet QCD distribution. The difference in shape between the multijet QCD and $t\bar{t}$ m_{H1} distributions help differentiate the two backgrounds and make it an ideal variable to fit on.

This method gives an initial background estimation, which is shown in Figure 7.2. These plots show a slight downward trend when comparing data to the predicted background, which implies that the background estimation is under-predicting the background events as a function of increasing m_{HH} particularly for 3b and 2b-split. As this demonstrates a source of potential mismodelling in our background estimation, background reweighting is introduced to resolve this.

Scale Factor	4b	3b	2b-split
μ_{QCD}	0.0269 ± 0.0014	0.1191 ± 0.0023	0.536 ± 0.001
$\alpha_{t\bar{t}}$	1.0 ± 0.0	0.870 ± 0.046	0.900 ± 0.012
correlation	0.00	-0.74	-0.74

Table 7.2. Background scale factor values for the three tagging region. For the 4b region, $\alpha_{t\bar{t}}$ is fixed to 1.0.

7.2 Kinematic Reweighting

A proper background estimation relies heavily on comparing relative kinematic distributions with different b-tagging requirements. However, b-tagging will "sculpt" track jet kinematic distributions because b-tagging efficiency declines at high p_T. This is due to the decrease of the angular separation between tracks and the reconstruction efficiency of the secondary vertex. Thus, track jet variables such as track jet p_T and η are sculpted and therefore result in potential mismodelling of the background in the CR, VR, and SR. To accommodate these differences between b-tagging regions,

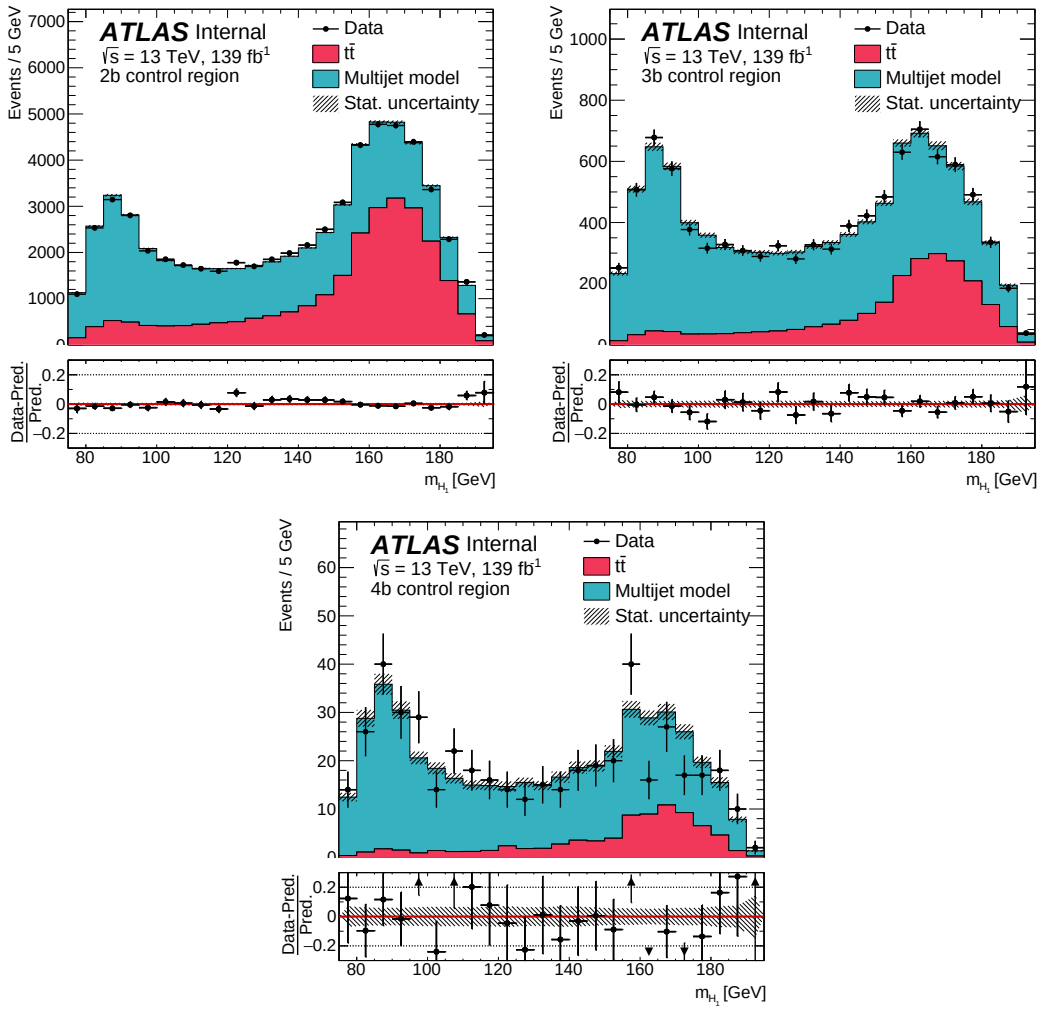


Figure 7.1. Background estimation fit result for the m_{H1} variable of multijet QCD and $t\bar{t}$ background compared to data in the CR. A two peak structure is apparent for the multijet QCD and not for $t\bar{t}$.

reweighting weights are derived by comparing different kinematic distributions of Higgs candidates with and without b-tags.

The regions with the most sculpting are the 3b and 2b-split – 4b does not have an adequate amount of events to imply any obvious trends in the background estimation so it is left as-is (with no reweighting applied to it). This means the reweighting that needs to be derived is: 2b-1 $\rightarrow$ 2b-1b and 1b-1 $\rightarrow$ 1b-1b. This implies that the reweighting need only be done by comparing the one b-tagged track jet in a Higgs

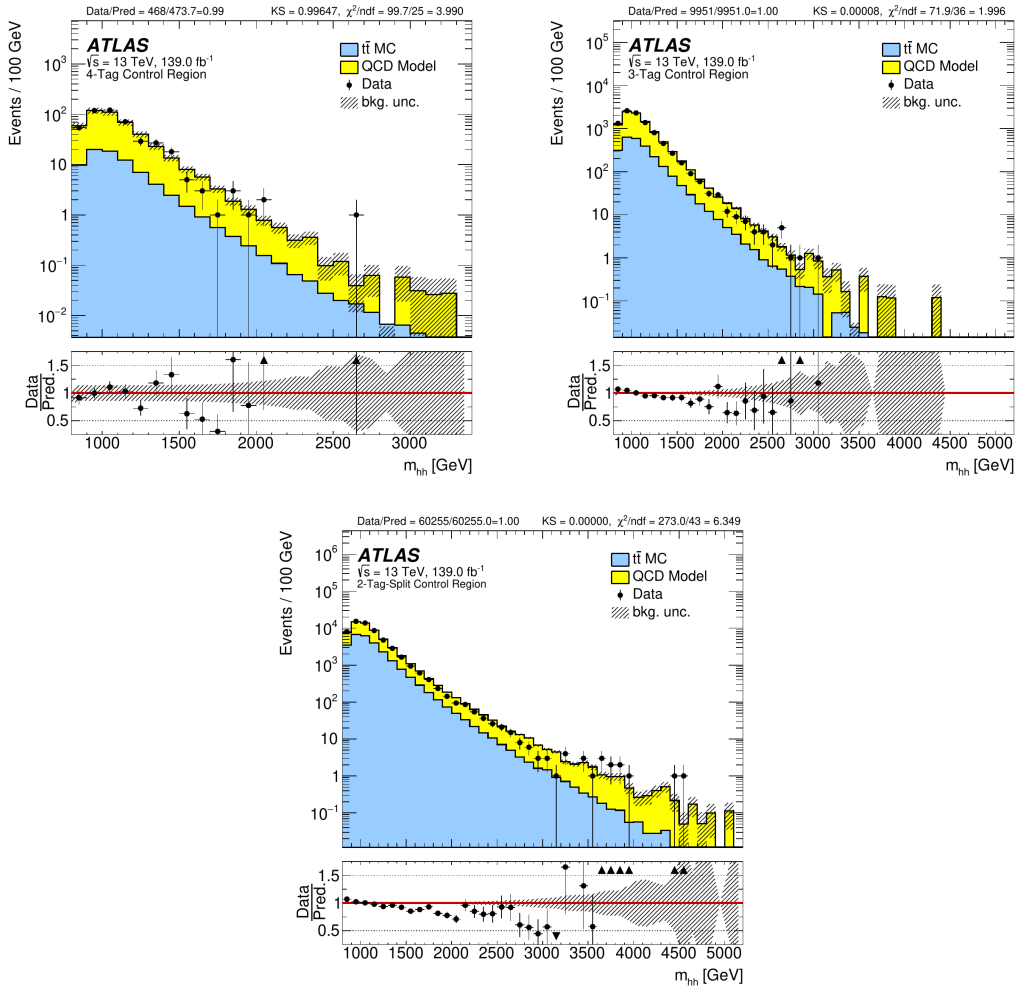


Figure 7.2. Plots of background estimation in the CR without reweighting applied. The 4b, 3b and 2b-split shown on the top left, top right, and bottom respectively.

candidate to a track jet without a b-tag ("untagged") in a Higgs candidate. If a Higgs candidate has two leading track jets in the untagged jet, which 2b-1 allows for (see *background sharing* in Section 6.2.2), one track jet is randomly chosen from the pair with a 50% chance of either of the two track jets being selected.

The reweighting procedure is iterative – after each iteration four ratios of kinematic distributions between the tagged and untagged variants are compared. The tagged distribution is scaled to the number of events in the untagged distribution

before the ratio of tagged to untagged is taken. The four variables that are most sculpted by b-tagging in this analysis and that are used for reweighting are:

1. p_T of the Higgs candidate,
2. p_T of the track jet in the Higgs candidate,
3. ΔR_{jj} between of the two track jets in a Higgs Candidate (only when applicable),
4. track jet η.

After each evaluation of the $\frac{tagged}{untagged}$ ratio for the aforementioned variables (ΔR_{jj} only when there are two track jets in a Higgs candidate), a fit is applied. Many fits have been considered, such as a linear fit, polynomial fit and cubic spline. A cubic spline, although most computationally resource intensive, performs the best of all methods tried and is therefore chosen for our reweighting scheme. With these splines, we derive a weight, which starts at one ($W_0 = 1$) and expressed as:

$$W_i = W_{i-1} \times [(\Pi_k f_{ik} - 1) \times lr_i + 1], \tag{7.2}$$

where f_{ik} is the spline for kinematic variable k of the i^{th} iteration and lr_i is a learning rate. The learning rate tempers the reweighting as to not over correct after every iteration and has been set to $lr_i = 1 - 0.5^i$. With this learning rate, the splines converge to near unity after ten iterations ($i = 10$). The weights are derived on data in the CR, which contains both multijet QCD and $t\bar{t}$ background, and are then applied to the low-tag data and $t\bar{t}$ MC samples.

Figures 7.3 and 7.4 show the m_{HH} distribution in the CR for the high-tag and low-tag regions before and after reweighting. A downward trend as a function of mass is apparent in the bottom half of the plots without reweighting for both 3 tag and 2 tag-split, which is corrected by reweighting. The impact of reweighting on the final background estimation is shown at the end of this section.

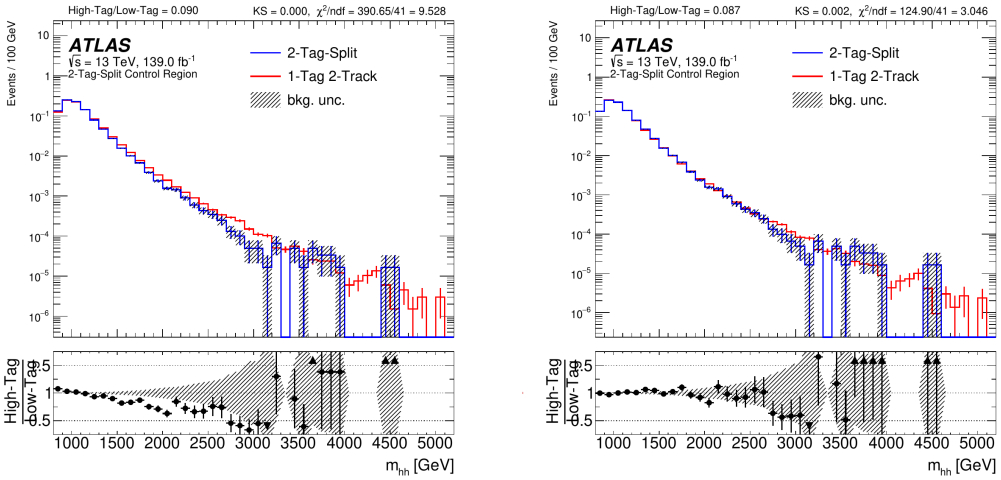


Figure 7.3. High-tag and low-tag m_{hh} distribution in the 2 tag-split region before and after reweighting on the left and right respectively. The shaded region is the background uncertainty, which is described in Section 8.2.

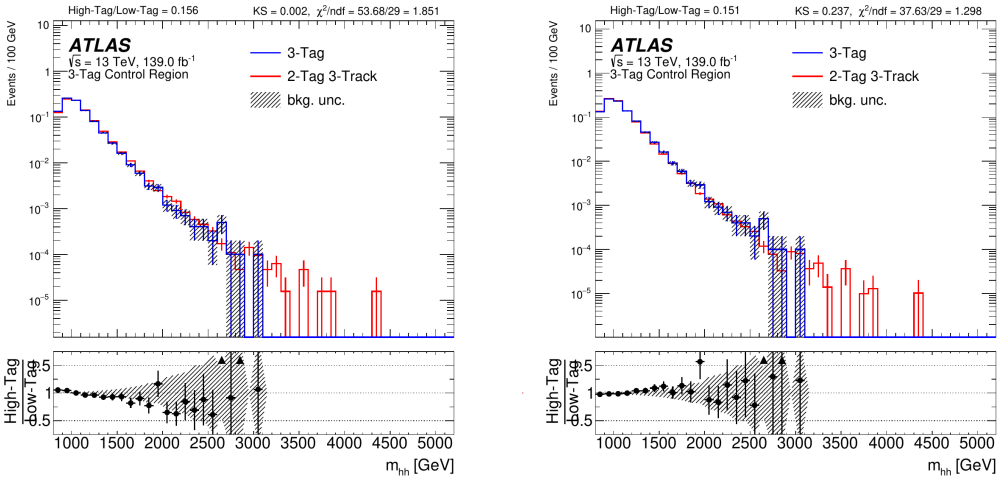


Figure 7.4. High-tag and low-tag m_{HH} distribution in the 3 tag region before and after reweighting on the left and right respectively.

7.3 Background smoothing

A background smoothing procedure is applied to the background estimation to reduce the impact of statistical fluctuations, which occur at high m_{HH}. The background distributions are fit to a three parameter function,

$$f(x) = \frac{e^{-p_0}}{x^2}(1-x)^{p_1 - p_2 \ln x}, \tag{7.3}$$

where $p_{0,1,2}$ are free parameters to be fit to the background distribution and x is the m_{HH} variable. This is the eighth of eight functions listed in Table 7.3, which are commonly used for fitting long-tail distributions. These functions are appropriately named "dijet" or "multijet" ("MJ") functions as they are often used in multi-jet analyses such as this one. The MJ8 function proves to be the most stable and accurate fitting function when applied to the background distributions of all the tagging regions.

Name	Function
MJ1	$f_1(x) = e^{-p_0}(1-x)^{p_1}x^{p_2}$
MJ2	$f_2(x) = e^{-p_0}(1-x)^{p_1}e^{p_2 x^2}$
MJ3	$f_3(x) = e^{-p_0}(1-x)^{p_1}x^{p_2 x}$
MJ4	$f_4(x) = e^{-p_0}(1-x)^{p_1}e^{p_2 \ln x}$
MJ5	$f_5(x) = e^{-p_0}(1-x)^{p_1}(1+x)^{p_2 x}$
MJ6	$f_6(x) = e^{-p_0}(1-x)^{p_1}(1+x)^{p_2 \ln x}$
MJ7	$f_7(x) = \frac{e^{-p_0}}{x}(1-x)^{p_1 - p_2 \ln x}$
MJ8	$f_8(x) = \frac{e^{-p_0}}{x^2}(1-x)^{p_1 - p_2 \ln x}$

Table 7.3. The MJ functions used to fit the background m_{HH} distribution.

The background distribution has nearly all events at low m_{HH} and therefore without careful consideration for this, the low-m_{HH} region can dominate the fit. The beginning of the distribution (below 1 TeV) has a positive slope upward followed by a point of inflection (1 TeV - 1.2 TeV) and a downward slope to the eventual long-tail at high-m_{HH} (> 2.5 TeV). The shape in the beginning of the distribution is partially

Tagging Region	QCD Range (GeV)	$t\bar{t}$ Range (GeV)
4b	1200-2800	(extrapolated)
3b	1200-2800	1200-2200
2b-split	1200-4300	1200-3900

Table 7.4. Background smoothing fit ranges for the background sources of the boosted analysis.

the result of the inefficiency of the boosted analysis at low mass and therefore the fit is set to begin at 1200 GeV to avoid having a bias.

The full set of "fit ranges", or the starting and ending point in m_{HH} of the fit, is listed in Table 7.4. The fit ranges are tuned for each channel to improve stability of the fitting procedure. The multijet QCD and $t\bar{t}$ background are fit separately and also have different fit ranges. The number of $t\bar{t}$ events in the 4b channel is too low to perform a fit, therefore the 3b region fit on $t\bar{t}$ is scaled to the number of events in the 4b region and used instead.

The resulting background smoothing is shown in Figures 7.5 and 7.6. The nominal line is used as the background estimation for each background type while the variations are used as a way to estimate the systematic uncertainty after smoothing, which is described in Section 8.2.3.

7.4 Background Estimation Result

The final background estimation plots for the boosted analysis are shown in Figures 7.7 - 7.8. The distributions shown have undergone the kinematic reweighting process and the distributions in the VR and SR have been smoothed. With the correcting of the reweighting and background smoothing, good agreement is seen between data and background in the VR and SR.

The background estimation plots shown in Figures 7.7 - 7.8 contain the background estimation uncertainties, which are described in Section 8.2. They do not, however, include the other sources of uncertainty explained in Section 8.

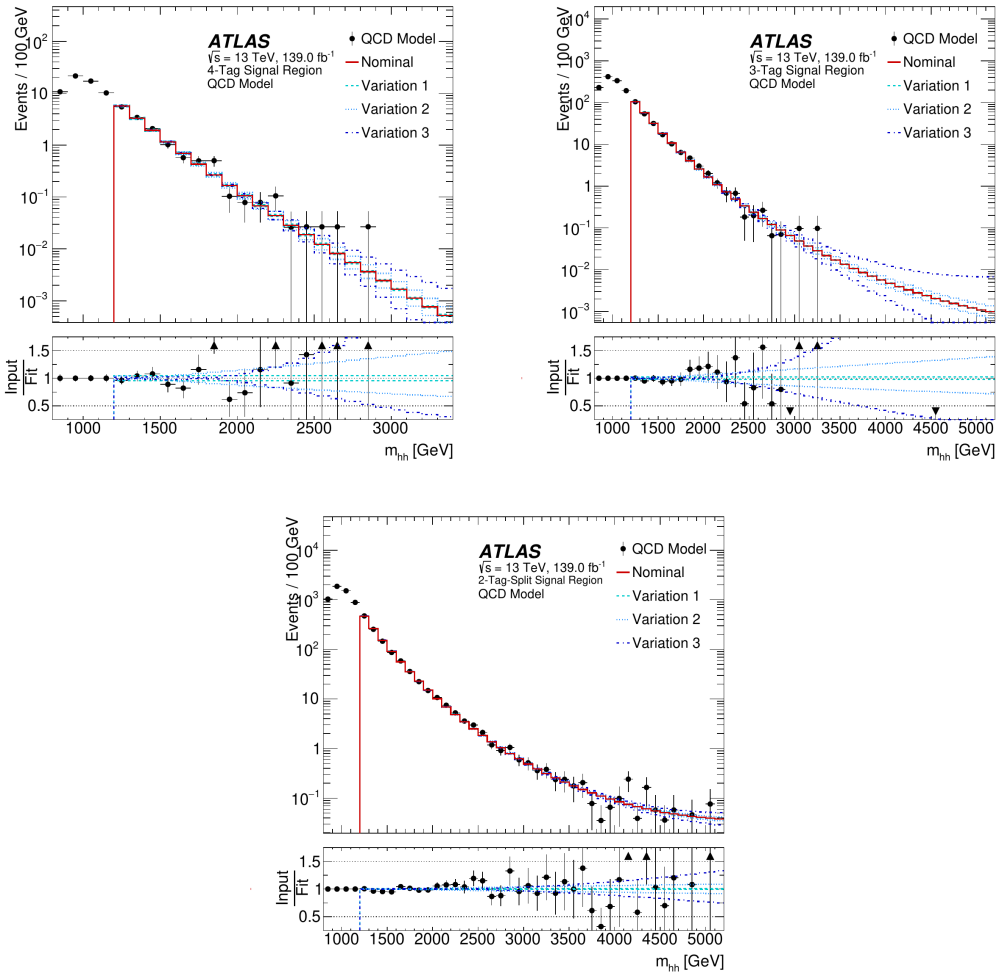


Figure 7.5. Smoothing of the multijet QCD background for the 4b, 3b, and 2b regions on the top left, top right, and bottom respectively.

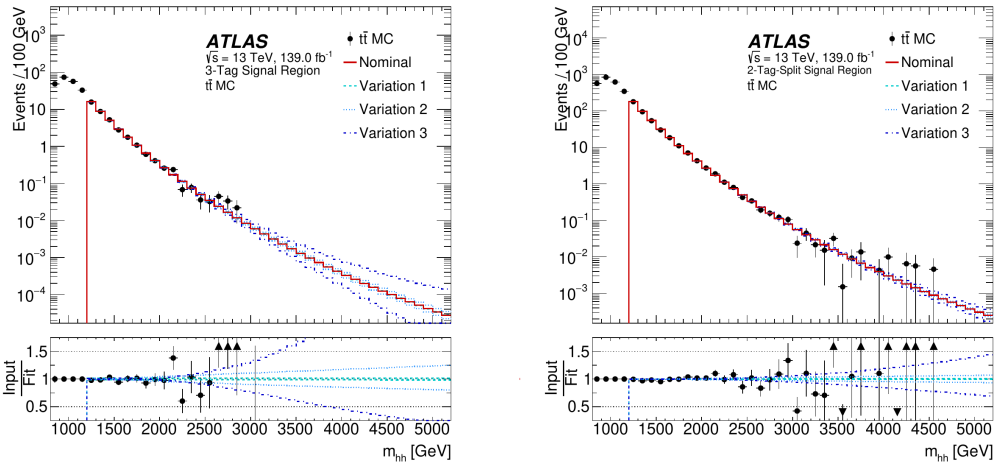


Figure 7.6. Smoothing of the $t\bar{t}$ background for the 3b, and 2b regions on the left and right respectively. The 4b region is not shown as it is derived from scaling the 3b region smoothing.

Figures 7.9 and 7.10 compare the background estimation derived with and without reweighting in the CR and VR respectively. Note that the boosted analysis was never unblinded without reweighting so a similar comparison figure for the SR is not shown. Without the reweighting, the background estimation tends to over predict the background as a function of m_{HH}, which is apparent in a downward trend in the bottom plots of the figures. This is due to the few single event bins at high m_{HH}, which prove difficult to fit properly and skew the estimation upwards. With the reweighting applied, the estimation better matches the bulk of the distribution, which lies between $m_{HH} = 1 - 3$ TeV but still accommodates the events above 3 TeV.

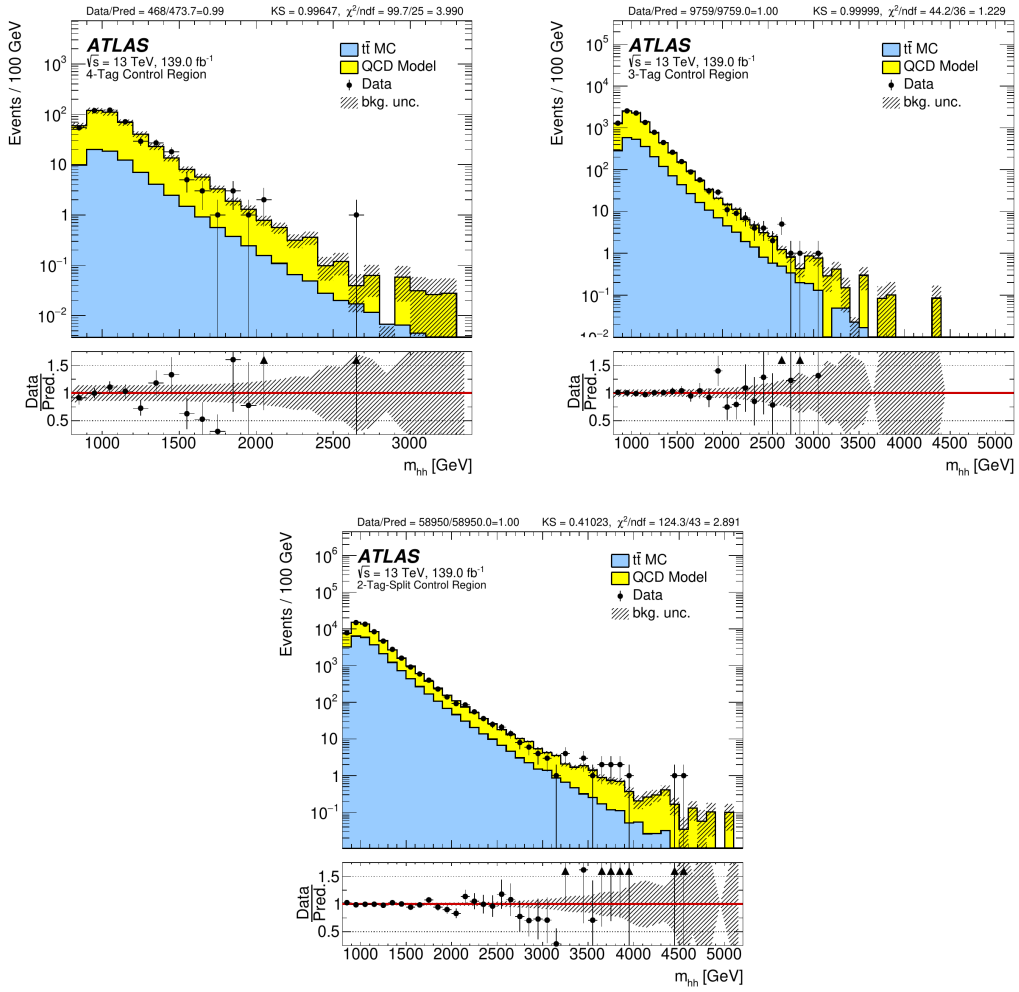


Figure 7.7. The m_{HH} distributions from the boosted analysis background estimation. Different tagging regions in the CR are shown: 4b, 3b, and 2b-split on the top left, top right, and bottom respectively.

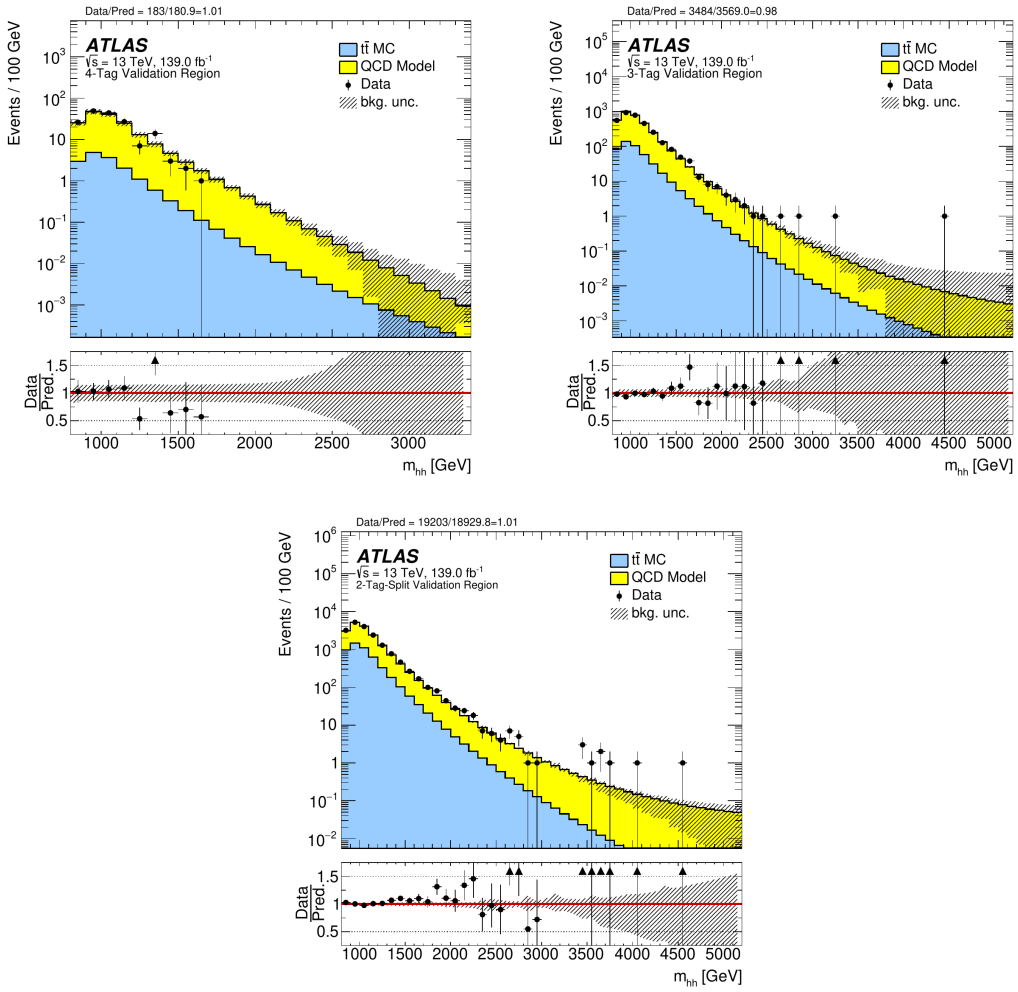


Figure 7.8. The m_{HH} distributions from the boosted analysis background estimation. Different tagging regions in the VR are shown: 4b, 3b, and 2b-split on the top left, top right, and bottom respectively.

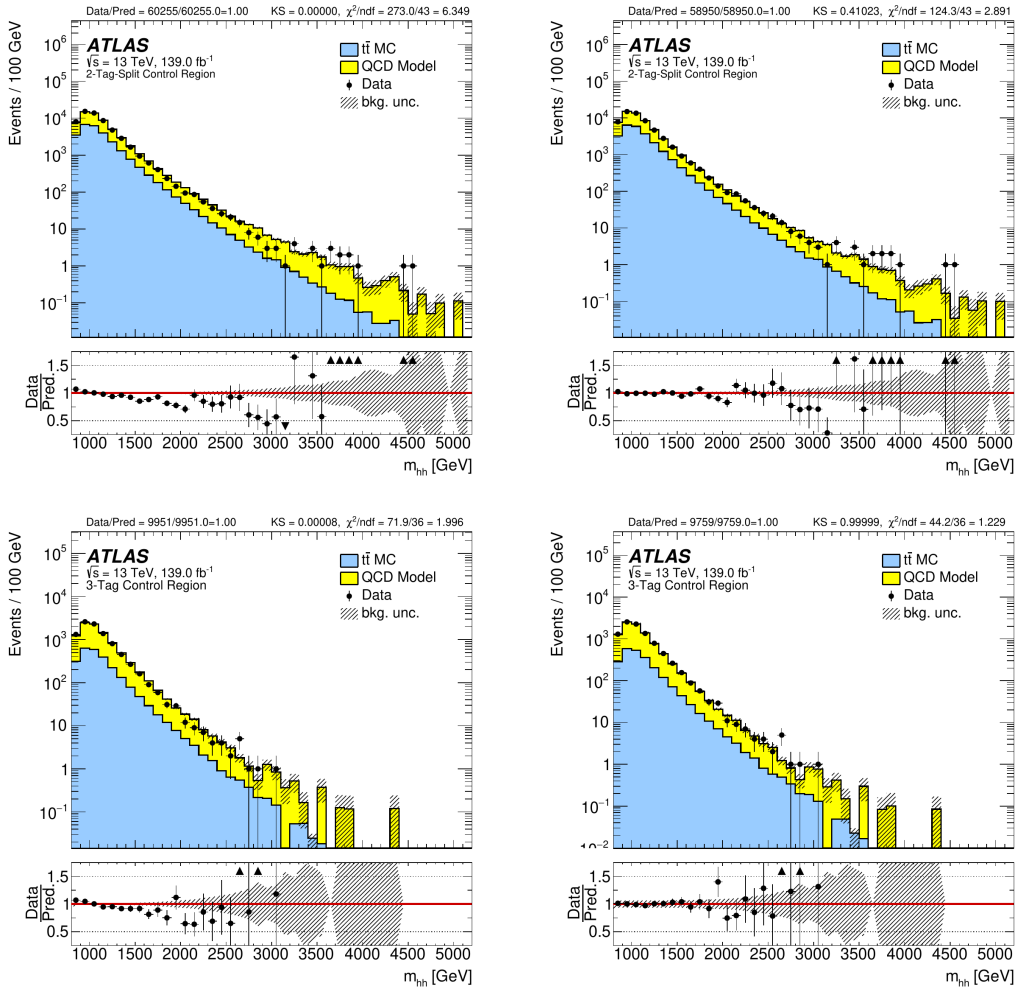


Figure 7.9. Plots of background estimation in the CR without and with reweighting applied on the left and right respectively. The 4b region plots are not reweighted and therefore not shown.

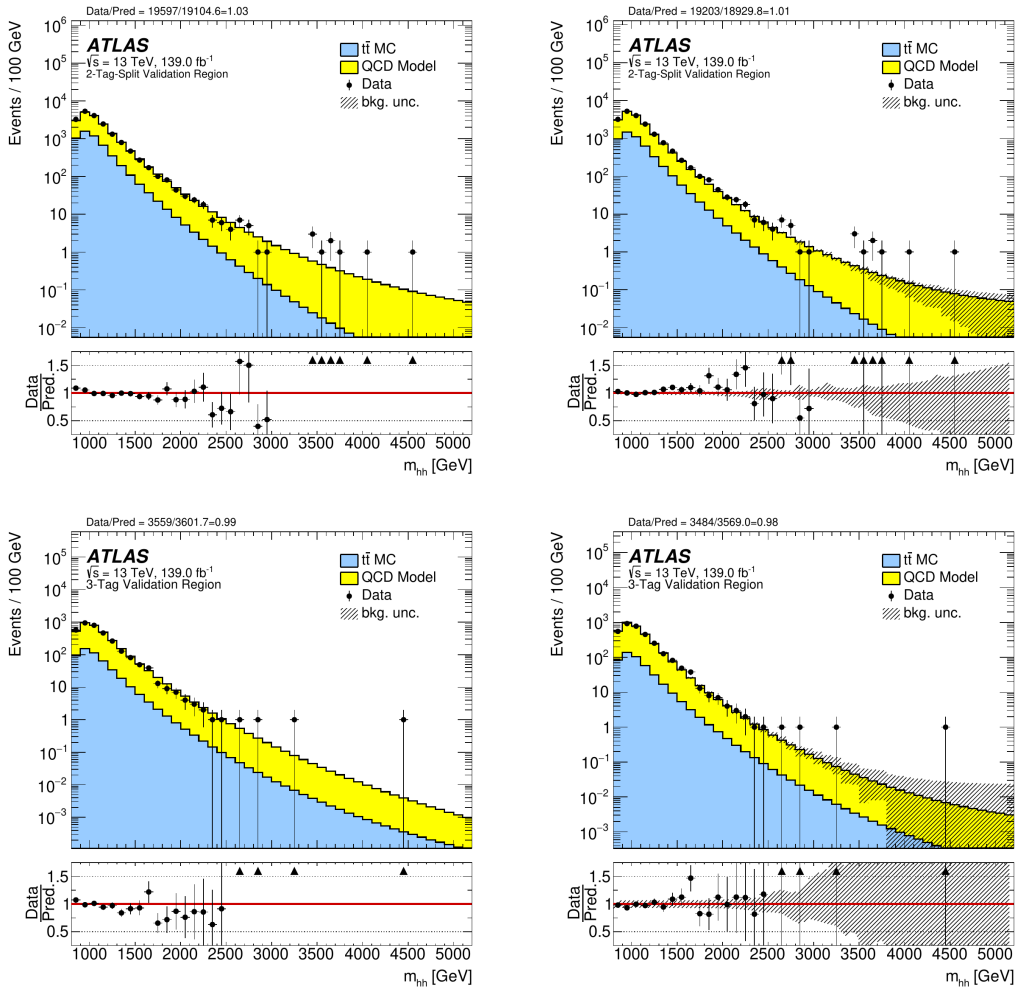


Figure 7.10. Plots of background estimation in the VR without and with reweighting applied on the left and right respectively. Smoothing is applied to this region and the 4b region plots are not reweighted and therefore not shown.

CHAPTER 8

SYSTEMATIC UNCERTAINTIES

The boosted analysis is statistically limited, especially at high resonance mass where the statistical uncertainties become large in the long tail of the backgrounds estimation. The biggest systematic uncertainties, however, come from b-tagging and jet mass resolution uncertainties. Other sources of uncertainty include MC simulation uncertainties for both $t\bar{t}$ and signal samples, background estimation uncertainties, theoretical uncertainties, and finally two non-closure uncertainties. The statistical uncertainty and the various other systematic uncertainties are described here.

8.1 Statistical Uncertainty

Statistical uncertainty is characterized by the expected fluctuations based on the number of repeated measurements. These fluctuations may theoretically be an event-by-event measurement, but in the case of this analysis and many of those at ATLAS the measurements of events are categorized by a discriminant (e.g. m_{HH}) and placed into bins of a histogram. The number of events in a given bin are described by a Poisson distribution, which approaches a Gaussian distribution for a substantial number of events, has standard deviation estimated by the $\sqrt{N_i}$, where N_i is the collected value of bin i. This is the statistical uncertainty associated to each bin, which becomes smaller relative to the number of events in the bin. That is, the expected value for a given bin has smaller expected fluctuations if we collect more events for that bin. Therefore, in the high m_{HH} region where bins have low, often singular values of N the statistical uncertainty is large relative to N. This ratio, $\sqrt{N}/N$, shrinks towards lower m_{HH} where N is substantial. Figure 8.1 shows two plots for 3b CR data: N and $\sqrt{N}/N$ as a function of m_{HH}. The statistical error dominates in the region of high m_{HH}.

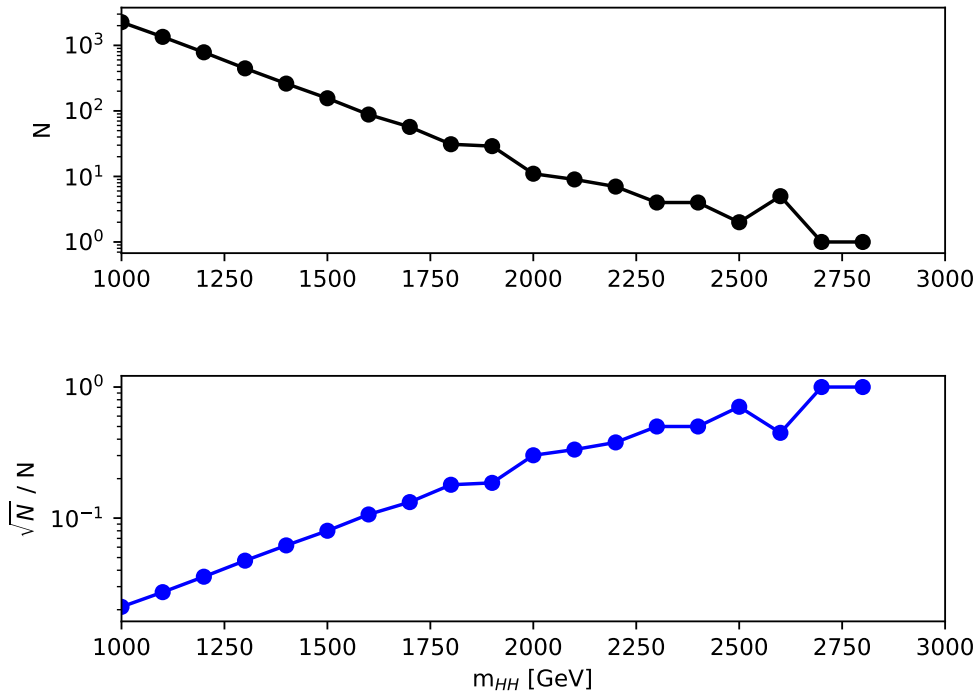


Figure 8.1. Distributions for the 3b CR data. Top and bottom respectively shows N and $\sqrt{N}/N$ as a function of m_{HH}.

8.2 Background Estimation Uncertainties

The uncertainties associated to the background estimation are split into eight parts: two normalization uncertainties and six shape uncertainties. The normalization uncertainties are the background estimation extrapolation uncertainty and the uncertainty due to possible variations in the definition of the CR. Four of the six shape uncertainties are associated to the background smoothing, which involve using alternative fits from varying the best fit value for μ_{QCD}, the best fit value for $\alpha_{t\bar{t}}$, the MJ function choice, and the fit range $\pm$100 GeV. The last two uncertainties are a "residual" (non-closure) uncertainty and a non-closure uncertainty, which are respectively for differences between CR and SR and differences between the data-driven background estimation and a background estimation that uses simulated multijet

QCD. Although simulated $t\bar{t}$ samples are used in the background estimation uncertainty, uncertainties pertaining directly to $t\bar{t}$ MC simulation are not described in this section but later in Section 8.3.

8.2.1 Extrapolation Uncertainty

The background estimation procedure is only valid in this analysis if the scale factor values μ_{QCD} and $\alpha_{t\bar{t}}$ remain constant across the $m_{H1} - m_{H2}$ plane. Any difference in the scale factors along the $m_{H1} - m_{H2}$ plane are accounted for by the background estimation extrapolation uncertainty defined here. Since $t\bar{t}$ is estimated from simulation, we may look at HT $t\bar{t}$ events in the SR without unblinding and therefore only an extrapolation uncertainty for μ_{QCD} is needed.

A two step Gaussian process technique ("GP") is used to estimate the data in the SR without unblinding the SR directly. The GP performs a regression (scikit-learn Gaussian Process Regression) in which several points (the VR and CR bins in this case) are used to interpolate data in an unknown area (SR) by estimating it as a multivariate Gaussian function. The GP is useful here because a Gaussian probability is supplied as a covariant matrix along with the interpolation.[34][42] The fit is done on the $m_{H1} - m_{H2}$ plane for $data - t\bar{t}$ and the resulting fit has a correlation length > 100 GeV, which is much larger than the SR and therefore allows a smooth, stable fit across the SR.

The GP is validated outside of the SR, where the values of the GP estimation closely match that of the actual $data - t\bar{t}$ values in that region. The resulting HT events in the SR are then available from the GP to use without unblinding the data. A double ratio comparing the HT and LT yields in the SR and CR is then calculated as,

$$R_{extrap} = \frac{N_{SR}^{HT}/N_{SR}^{LT}}{N_{CR}^{HT}/N_{CR}^{LT}}, \tag{8.1}$$

where N is the event yield for a given region. The uncertainty is then defined as $|R_{extrap} - 1|$, which is a flat normalization uncertainty of 2.0%, 5.7%, and 12% for the 2b-split, 3b, and 4b regions respectively.

8.2.2 Control Region Definition Uncertainty

The size and shape of the CR is well motivated, but any variation of the CR definition could impact the background estimation. Thus, a CR definition uncertainty is defined be varying the size and position of the CR in six ways. Essentially, every permutation of adding or subtracting 3 GeV to the two Higgs boson masses in the R_{HH}^{VR} and R_{HH}^{CR} definitions in Equation 6.1 accounts for four of the six variations. One of the last two variations make the CR larger by shrinking the VR by 3 GeV and expanding the CR by 3 GeV. For the last variation, the CR is made smaller by expanding the VR by 3 GeV and shrinking the CR by 3 GeV. The variations are depicted visually in Figure 8.2 where the first four variations are the "shifted" variations and the last two are the "big-little" variations. For every variation, a SR veto is done to ensure no overlap. For each variation the background estimation is repeated and the variation with the largest SR background yield difference is used to calculate the uncertainty. The largest yield difference between the nominal background estimation ("nom") and chosen variation ("var") is divided by the nominal yield to obtain the uncertainty, $\frac{|N_{SR}^{nom} - N_{SR}^{var}|}{N_{SR}^{nom}}$. The final normalization uncertainties found for 2b-split, 3b, and 4b respectively to be 0.75%, 1.5%, and 5.8%.

8.2.3 Background Smoothing Uncertainty

The background smoothing is added to the background estimation to reduce the impact of statistical fluctuations common at high values of the m_{HH} distributions. The smoothing introduces four sources of systematic uncertainty, all of which are accounted for in this Section. The uncertainties are shapes, listed as:

1. Altering the best-fit (nominal) value of μ_{QCD},

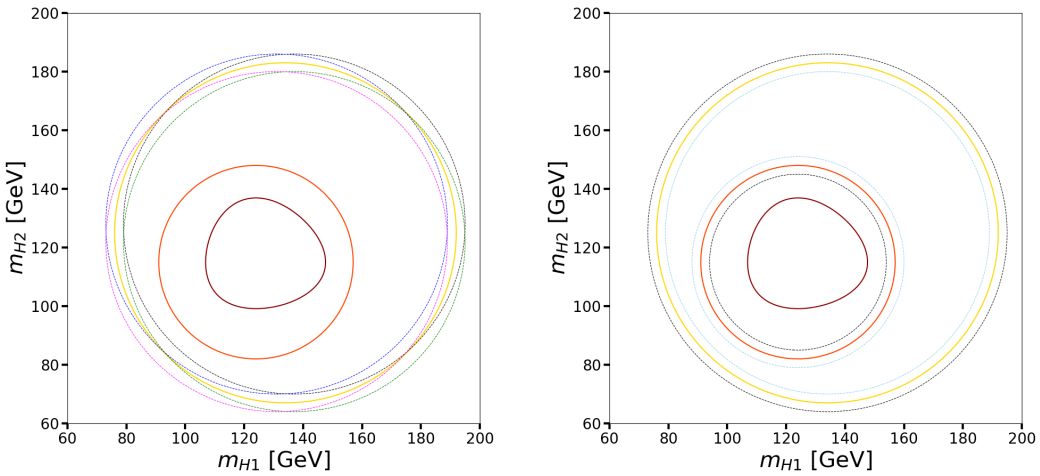


Figure 8.2. The CR, VR, and SR definitions on the $m_{H1} - m_{H2}$ plane with the CR variation overlaid as dashed contours. The left plot shows the "shifted" variations and the right plot shows the "big-little" variations.

2. Altering the best-fit (nominal) value of $\alpha_{t\bar{t}}$,

3. MJ fit function choice variation,

4. Fit range variation.

Variations on the best-fit are found by varying the μ_{QCD} and $\alpha_{t\bar{t}}$ scale factors along the eigenvector of the covariance matrix produced from the smoothing fit. There are three variations produced for μ_{QCD} and $\alpha_{t\bar{t}}$ respectively, which are subsequently propagated through to the background estimation to produce the alternative multijet QCD and $t\bar{t}$ smoothed background estimations. Each variation corresponds to a background estimation smoothed with one of the three varied scale factors and are shown in Figures 7.5 and 7.6. The resulting six total variations produce an envelope around the nominal background estimation, which is used in lieu of a systematic uncertainty in the area that is smoothed – above 1200 GeV.

Fit function choice and fit range uncertainties are added to account for the arbitrary nature of choosing one specific function and corresponding fit range over any

other. The fit function choice uncertainty is found by conducting the smoothing fit with every MJ function listed in Table 7.3. The result is several multijet QCD and $t\bar{t}$ distributions corresponding to the different MJ fit functions, shown in Figures 8.3 and 8.4. Note that if any of the MJ functions are missing from the figures, it is because the fit has failed to converge for that function and it is therefore excluded. The largest deviations from the nominal value are used to form an envelope, which is then used as the shape uncertainty.

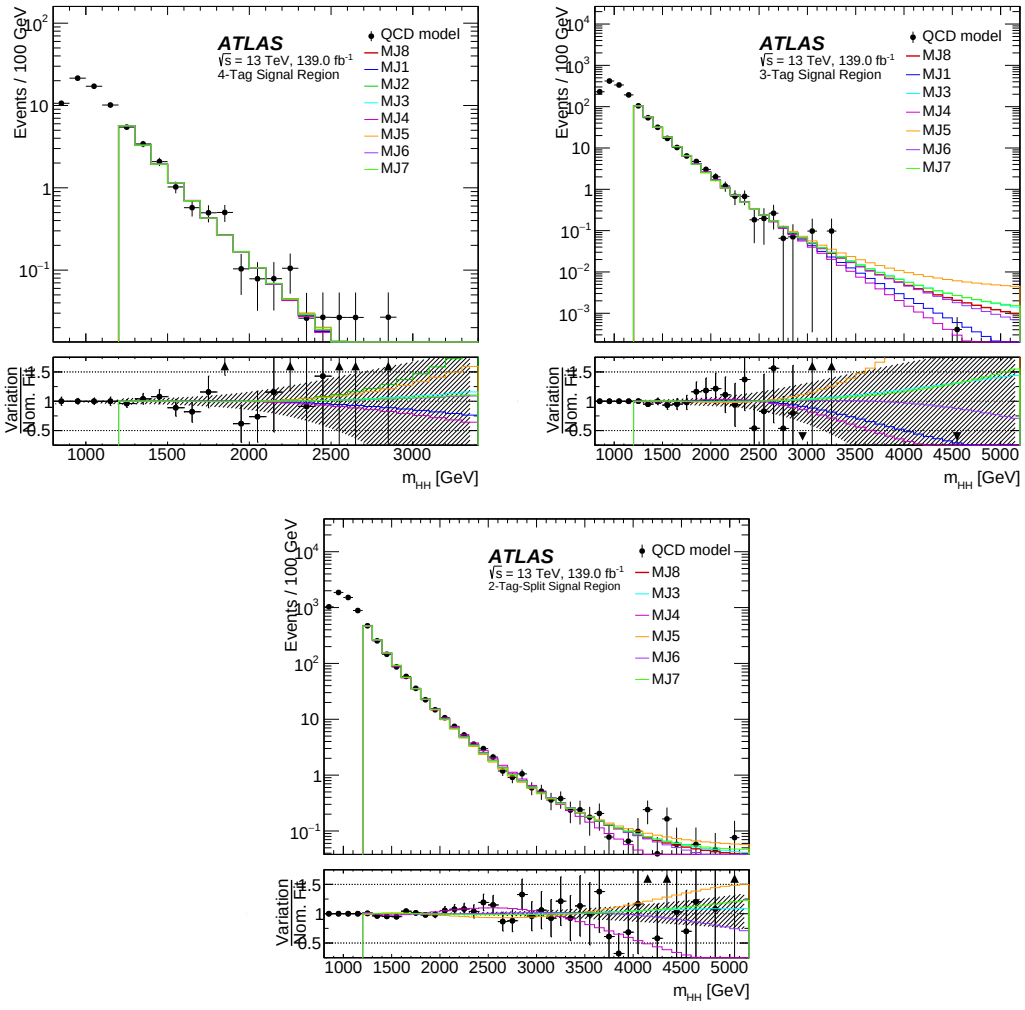


Figure 8.3. The several MJ fit functions applied to the mulitjet QCD background estimation in the SR. The 4b, 3b, and 2b-split shown on the top left, top right and bottom respectively.

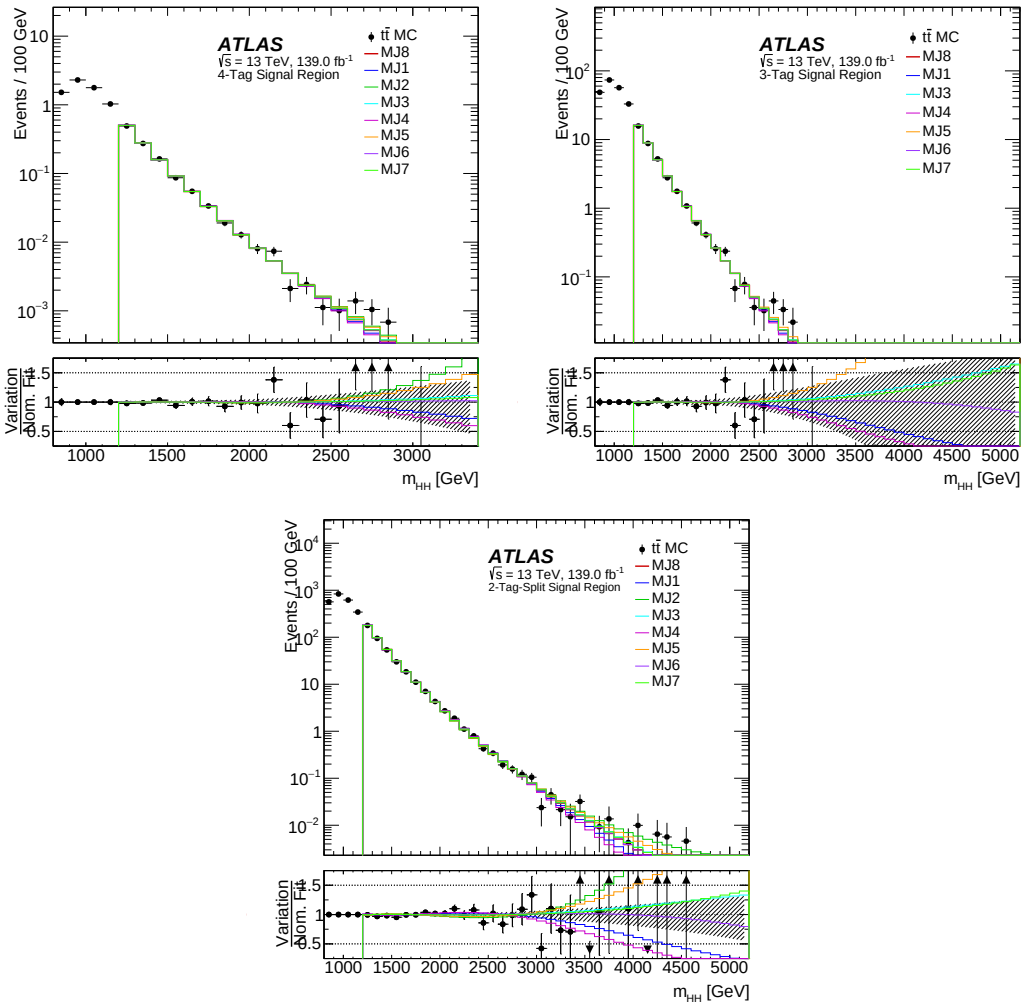


Figure 8.4. The several MJ fit functions applied to the $t\bar{t}$ background estimation in the SR. The 4b, 3b, and 2b-split shown on the top left, top right and bottom respectively.

The fit range uncertainty is found by repeating the smoothing fit with the nominal choice but with every permutation of upper and lower fit range bounds ±100 GeV. Figures 8.5 and 8.6 show the nominal smoothing fit along with each fit variation. Similarly to the fit function choice uncertainty, the largest deviations from the nominal value are used to form an envelope and is then used as the shape uncertainty for the smoothing fit range.

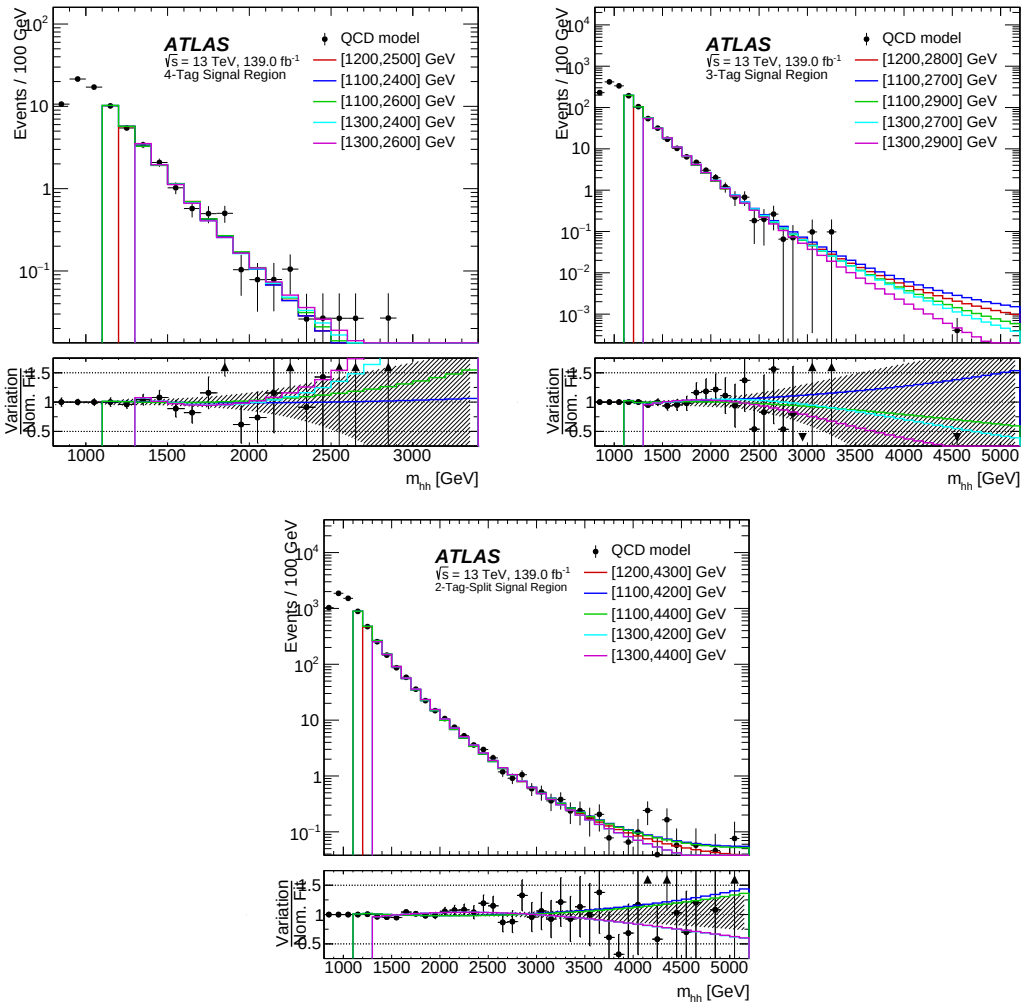


Figure 8.5. The multijet QCD background estimation in the SR with smoothing applied over several fit ranges, varied about the nominal fit range by 100 GeV. The 4b, 3b, and 2b-split shown on the top left, top right and bottom respectively.

8.2.4 Non-closure MC Uncertainty

A non-closure uncertainty is added to account for differences between the estimated multijet QCD background and the actual multijet QCD events. By using an identical multijet QCD background estimation method derived from simulated multijet QCD samples instead of data, HT SR events are available for comparison without unblinding. This is the only time simulated multijet QCD samples are directly used in the boosted analysis. Instead of using $data - t\bar{t}$ to run the background estimation,

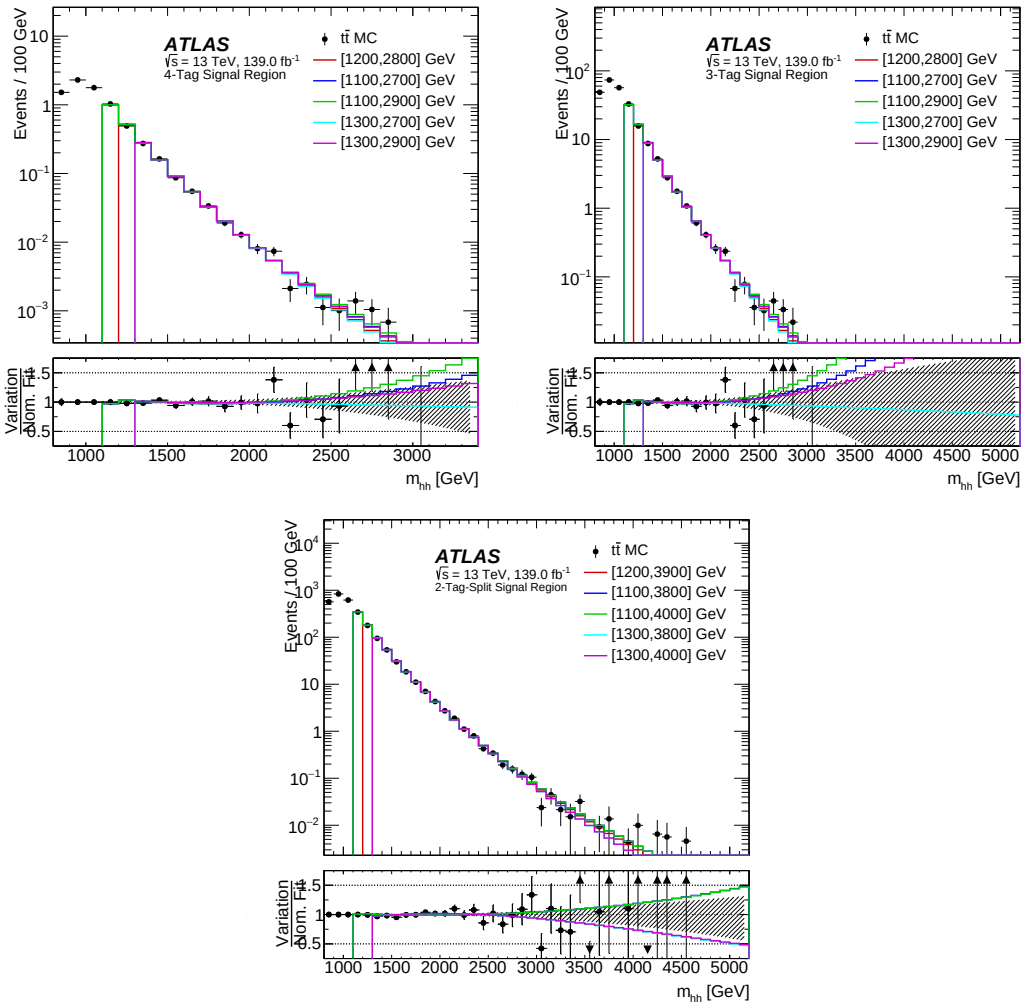


Figure 8.6. The $t\bar{t}$ background estimation in the SR with smoothing applied over several fit ranges, varied about the nominal fit range by 100 GeV. The 4b, 3b, and 2b-split shown on the top left, top right and bottom respectively.

simulated multijet QCD is propagated into the background estimation procedure. The resulting simulation-driven background estimation is then compared to the simulated multijet QCD events in the combined VR and SR (i.e. VR events are summed with SR events to make a "VR+SR" region) to increase the number of events. The ratio of the multijet QCD events from the simulation-driven background estimation to the actual simulated multijet QCD events in the VR+SR region as a function of m_{HH} is then used to derive a non-closure uncertainty. The resulting distribution

shown in Figure 8.7 is fit to a linear function to avoid statistical fluctuations and the resulting line is symmetrized about the nominal distribution to give an envelope to use as the uncertainty. The uncertainty is as large as 15% for most of 4b and the high m_{HH} values of 3b and 2b-split. It is as low as 5% for the other values of 3b and 2b-split.

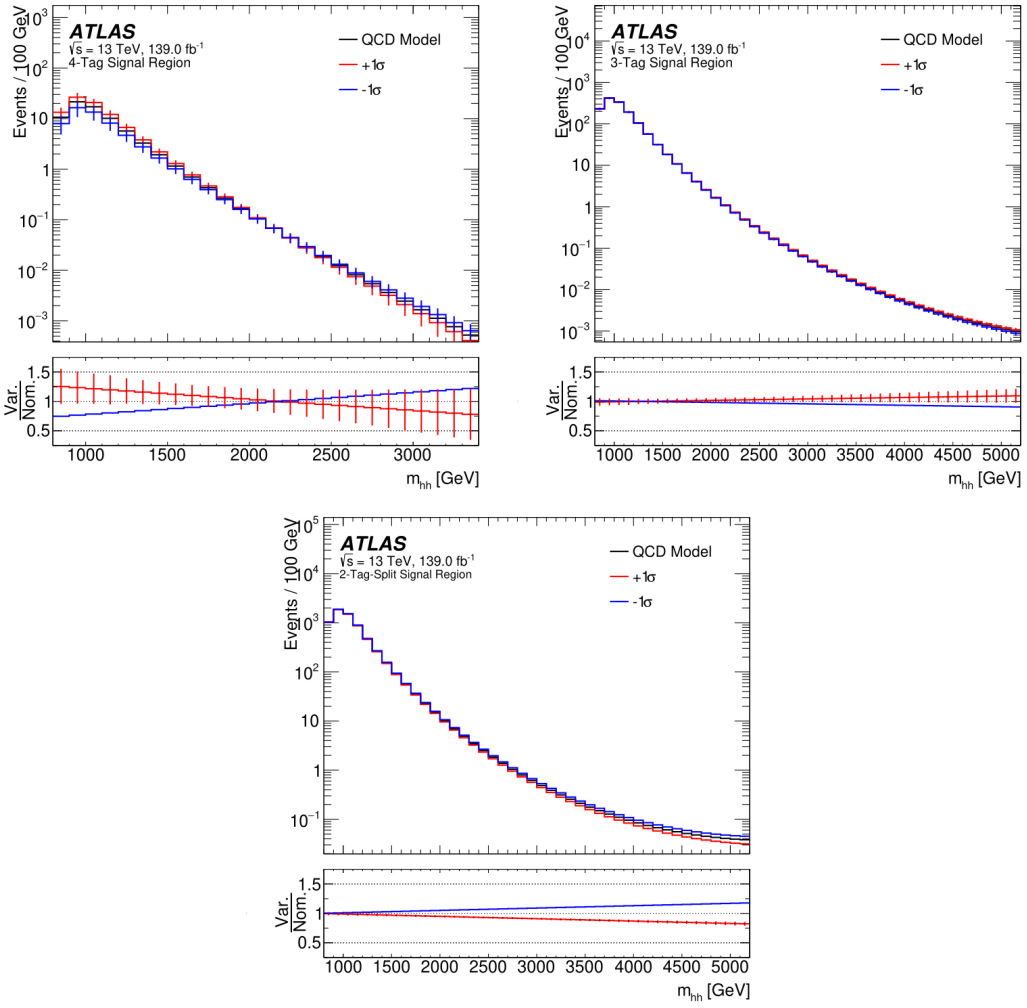


Figure 8.7. MC nonclosure uncertainties shown for 4b, 3b, and 2b-split on the left, middle and right respectively.

8.2.5 Residual Uncertainty

The last background estimation uncertainty is introduced to account for shape differences between the CR and SR regions. This is motivated by comparisons of $data - t\bar{t}$ distributions in the VR to data-driven background estimation distributions in the SR. Figure 8.8 shows this comparison for m_{HH} in the three tagging regions. With a bin-by-bin comparison some differences are seen, but singular bin values that deviate greatly compared to neighboring bins that have good agreement are considered statistical fluctuations and are thus ignored.

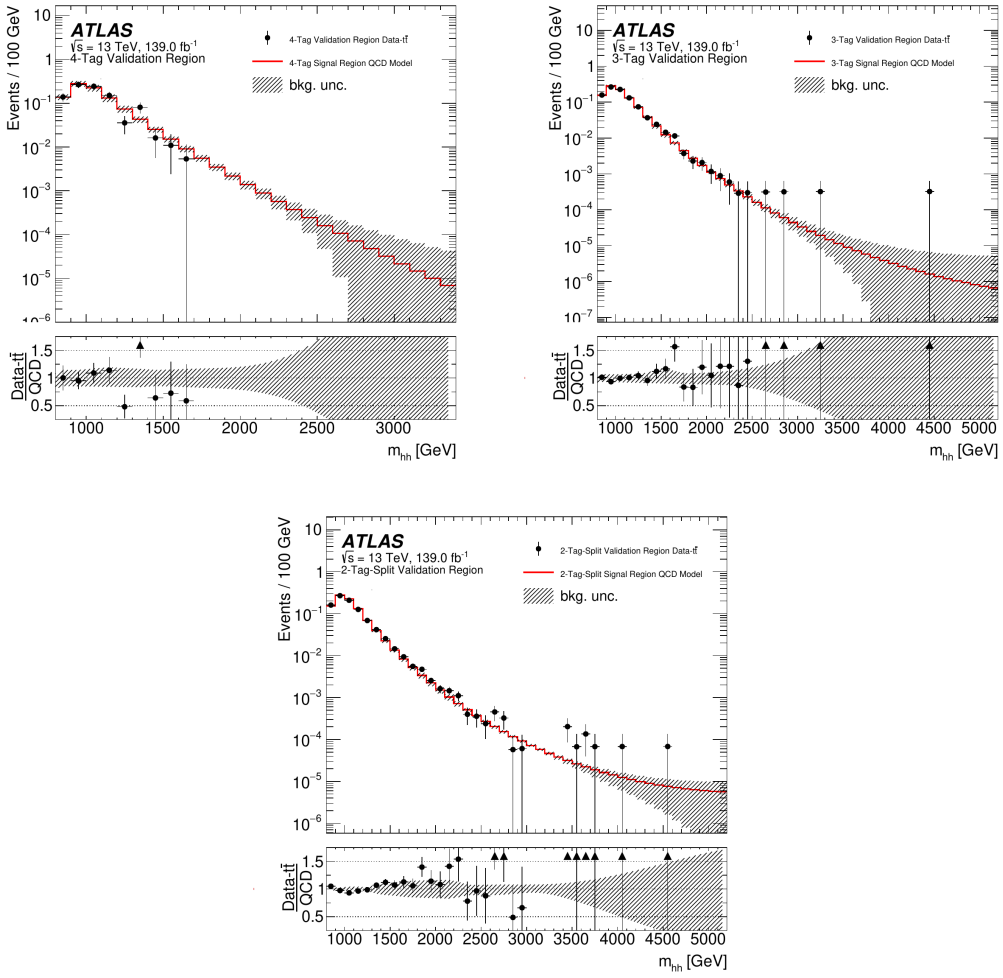


Figure 8.8. Comparison of the simulation driven multijet QCD estimation in the SR to the $data - t\bar{t}$ in the VR. The 4b, 3b and 2b-split shown top left, top right, and bottom respectively.

Only the 2b-split region demonstrates consistent, considerable differences and therefore is the only tagging region in which this uncertainty is applied. To derive the shape uncertainty, first any bin considered a statistical fluctuation is filtered from the distribution, then the ratio plot is smoothed. The shape is then symmetrized about nominal and used as the uncertainty.

8.3 MC Simulation Uncertainties

Experimental uncertainties only directly affect simulation samples, as the rest of the analysis uses data. The multijet QCD estimation does however use simulated $t\bar{t}$ subtracted from data and therefore uses simulation disseminates into the multijet QCD estimation and is impacted by MC simulation uncertainties. To handle the indirect nature of this, a given uncertainty type is passed through the analysis chain to produce a respective background or signal variation. The variations that overlap in both multijet QCD and $t\bar{t}$ are correlated and may produce either a shape or normalization uncertainty.

There are several MC simulation uncertainties included in this analysis. Uncertainty on the measurement of luminosity at LHC, uncertainties on jet measurements, uncertainty on b-tagging efficiencies, and other uncertainties pertaining to MC simulation are all accounted for.

8.3.1 Luminosity Uncertainty

The uncertainty on the total integrated luminosity for `Run-2` is a flat 1.7% normalization. This was obtained from the LUCID-2 detector, which is a detector along the LHC that measures the luminosity. This normalization is applied to simulated samples such as the signal samples and $t\bar{t}$ sample, which is propagated to the background.

8.3.2 Jet uncertainties

Four uncertainties are assigned to the Large-R jet: jet energy scale (JES), jet energy resolution (JER), jet mass scale (JMS), and jet mass resolution (JMR). All the jet uncertainties are modelled using the *in-situ* method described in Reference [22]. Most the jet uncertainties are relatively small, varying from 0-4% as a function of p_T , except the JMR, which is substantial for this analysis. These uncertainties follow the "in situ" method described in References [30] and [22]. Jet calibrations are meant to correct for detector response differences in different parts of the detector. Ratios of the jet energy or jet mass are compared between well defined physics objects and physics objects in the forward regions of the detector. The uncertainties of the calibration scaling and resolution are the JES, JER, JMS, and JMR, which is outlined in detail in Reference [30].

8.3.3 Flavor Tagging Uncertainties

The b-tagging uncertainties for this analysis are measured from track jets in $t\bar{t}$ events. The uncertainties associated to b-tagging are applied to b-tagging efficiency calibrations. These calibrations are optimized to best correct differences between simulation and data, as b-tagging applied to a simulated $t\bar{t}$ event may produce a different set of events passing a certain b-tagging working point than it would if applied to data instead. The prominent uncertainties for b-tagging are MC generator modelling uncertainties and experimental uncertainties. The first is found by comparing the nominal sample to a $t\bar{t}$ sample with a variation in the PS and also a $t\bar{t}$ enriched data sample. The difference between the resulting calibration scale factors produced are measured and used as an uncertainty.

The detector and reconstruction uncertainties are applied due to potential detector mismodelling during simulation. The impact of mismodelling is probed by repeating the simulation with reduced tracking efficiency, increased fake rate, and

biased tracking curvature (q/p). The p_T distributions for the alternative simulations are compared to the nominal one, and the ratio is used as the uncertainty. The impact of b-tagging uncertainties is moderate for most ranges of p_T in this analysis, anywhere from 0.2-10% and increasing with p_T .

8.3.4 $t\bar{t}$ MC Simulation Uncertainty

The several uncertainties on $t\bar{t}$ MC simulation ("$t\bar{t}$ MC uncertainties") are all evaluated by producing variations of the $t\bar{t}$ simulated samples with slightly changed components of the simulation. The three most significant sources of $t\bar{t}$ MC uncertainty are found by varying the matrix element calculation, parton showering generator, and the additional hard radiation parameter. Other uncertainties to the PDF and final state radiation renormalization scale were studied but found to be negligible for all parts of the boosted analysis and therefore not included.

The first uncertainty is attributed to the arbitrariness in choice of parton showering generator ("$t\bar{t}$ PS uncertainty"). The nominal $t\bar{t}$ sample uses Pythia8, which is compared to a $t\bar{t}$ sample simulated with Herwig7. The plots of Figure 8.9 show the nominal m_{HH} distribution compared the alternative m_{HH} distribution for the 4b, 3b, and 2b-split regions. The ratio is used as the $t\bar{t}$ PS uncertainty.

The matrix element ("ME") calculation uncertainty is found by varying the ME generator from Powheg to aMC@NLO. Only all-hadronic samples for the aMC@NLO are available, thus these samples are scaled by $N_{all}/N_{all-hadronic}$ to properly compare it to the nominal sample. The plots of Figure 8.10 shows the comparison of the alternative sample to the nominal for the 4b, 3b, and 2b-split regions. The ratio is used as the $t\bar{t}$ ME uncertainty.

The last uncertainty used is to account for the amount of hard radiation in $t\bar{t}$ decays, which is tuned during the simulation. The Powheg damping parameter h_{damp} is adjusted by a factor of two to change the hard radiation. Only all-hadronic samples

are available with varied h_{damp}, thus these samples are scaled by $N_{all}/N_{all-hadronic}$ to properly compare it to the nominal sample. The plots of Figure 8.11 shows the comparison of the m_{HH} distribution for the nominal $t\bar{t}$ sample to the m_{HH} distribution for the $t\bar{t}$ sample with varied h_{damp} factors (up and down). The ratio is shown for the 4b, 3b, and 2b-split regions and is used as the $t\bar{t}$ h_{damp} uncertainty.

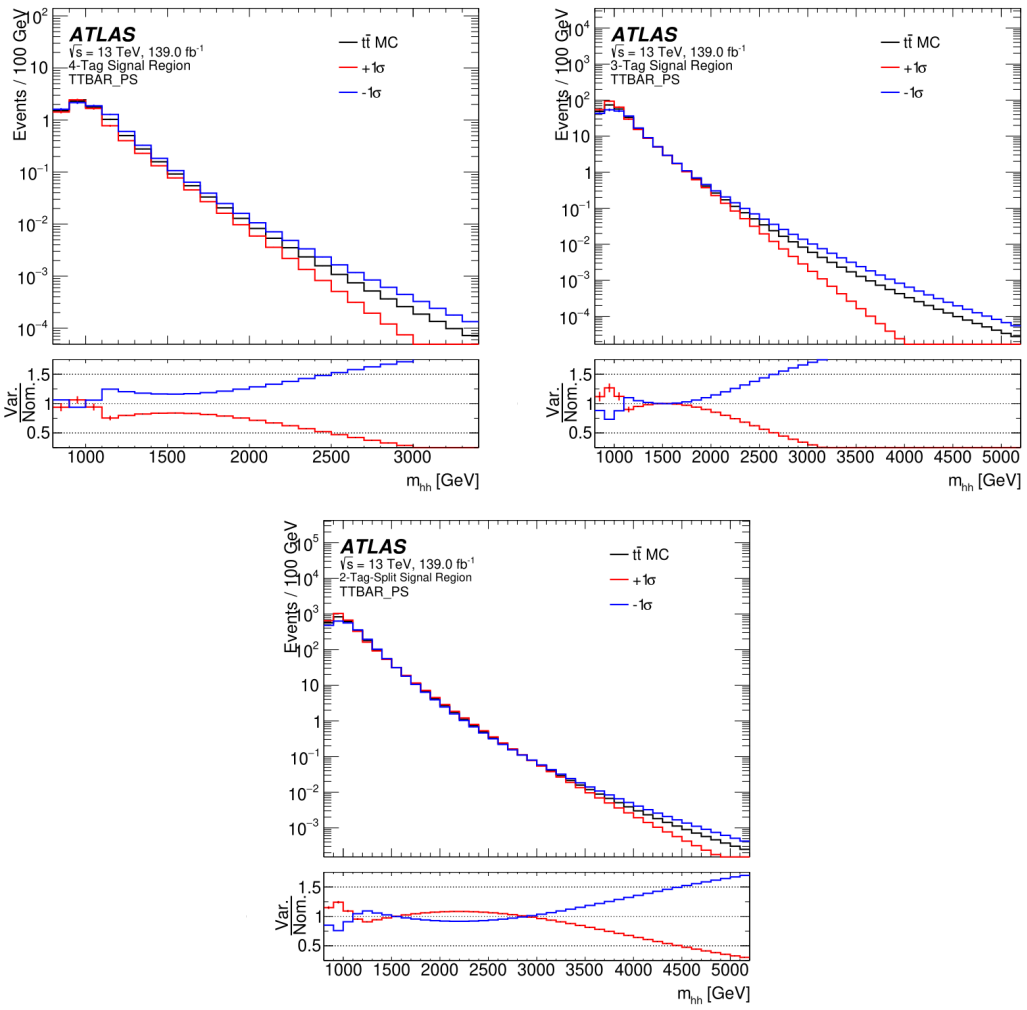


Figure 8.9. Ratio plots of the PS variation m_{HH} distributions to the nominal $t\bar{t}$. The 4b, 3b, and 2b-split regions shown respectively on the top left, top right, and bottom.

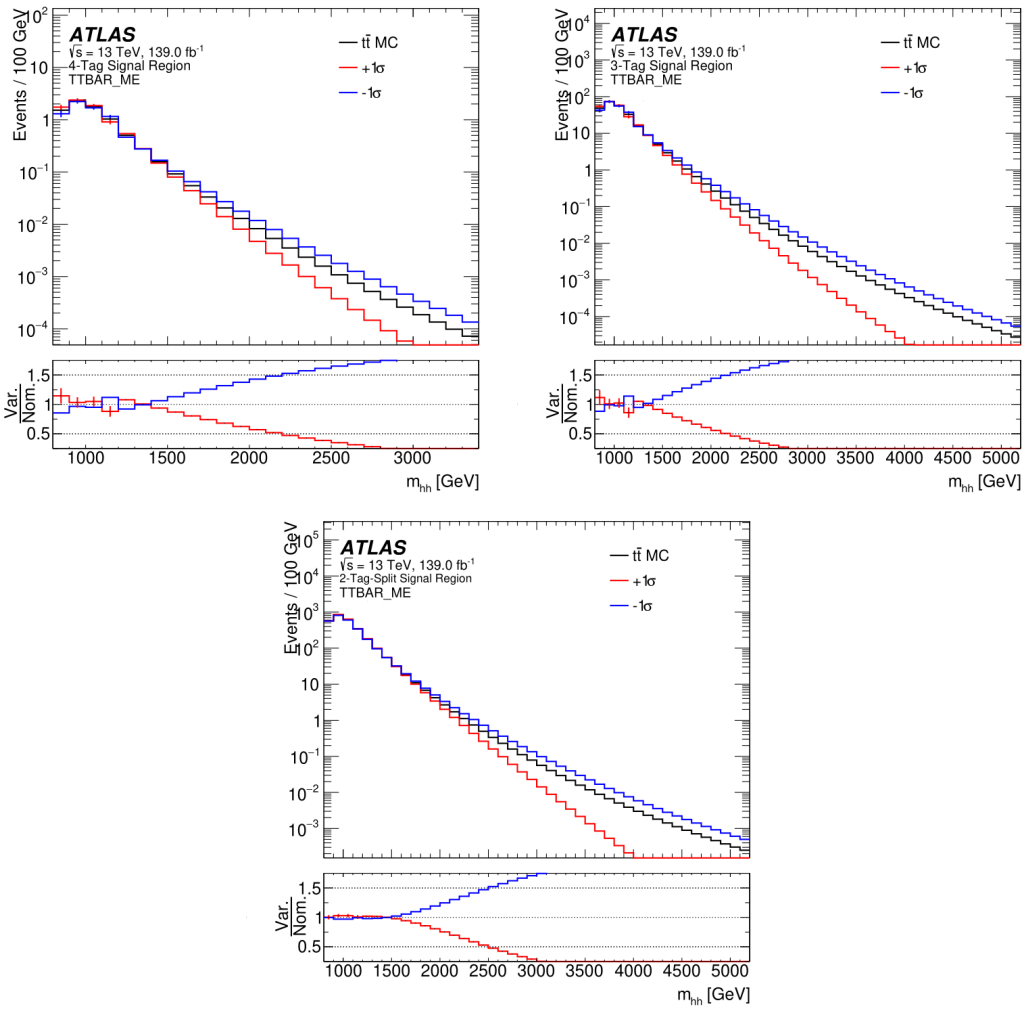


Figure 8.10. Ratio plots of the ME variation m_{HH} distributions to the nominal $t\bar{t}$. The 4b, 3b, and 2b-split regions shown respectively on the top left, top right, and bottom.

8.4 Signal Theory Uncertainties

The primary source of uncertainty in our signal samples comes about due to the choice of PS generator. A signal uncertainty on the ME calculation and a PDF uncertainty were considered but found to be negligible and therefore were excluded from the boosted analysis. The PS generator uncertainty is handled similarly to the $t\bar{t}$ uncertainty, by comparing signal samples made with the nominal to an alternative sample. The nominal Bulk RS model with c=1.0 sample uses PYTHIA8 and the Scalar

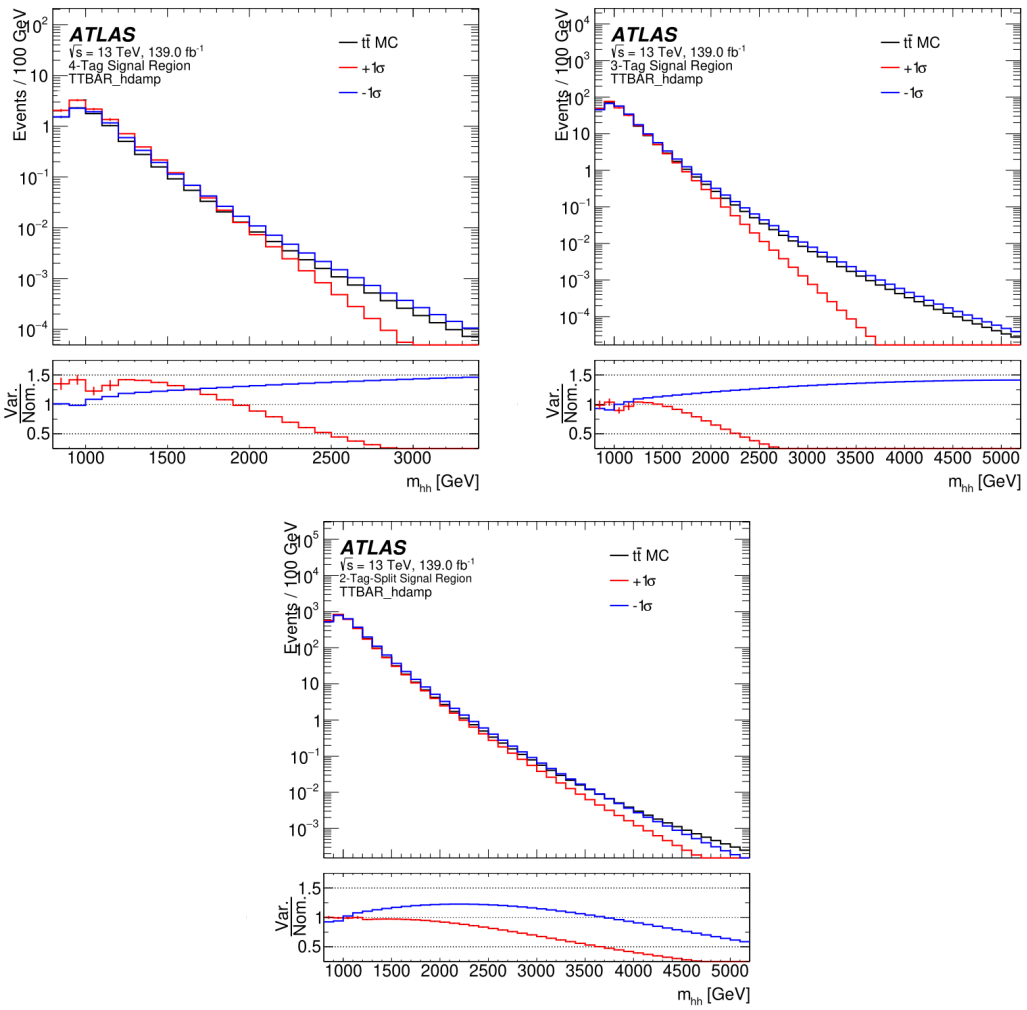


Figure 8.11. Ratio plots of h_{damp} variation m_{HH} distributions to the nominal $t\bar{t}$. The h_{damp} has two variations for the up (green) and down (red). The 4b, 3b, and 2b-split regions shown respectively on the top left, top right, and bottom.

model sample uses HERWIG7, thus a comparison need only be done between PYTHIA8 and HERWIG7. Three alternative samples for the Scalar model made with PYTHIA8 are available at 1, 2, and 4 TeV. Any shape difference in m_{HH} is not apparent enough to be considered, so only a normalization uncertainty is applied. The ratio of the event yields between PYTHIA8 and HERWIG7 for the three tagging regions are studied, with an estimated 10% uncertainty to the overall number of signal events is given to all tagging regions.

8.5 Summary of Uncertainties

A study of the systematic uncertainties is done to show the impact each one has on the final result. Table 8.1 is made by only allowing one source of systematic uncertainty to be applied at a time and assessing the relative impact on the limit (more on the limits in Section 9). This is roughly a sensitivity evaluation for each uncertainty source, and is done for six mass points for the Scalar model. The "Combined" uncertainty is the other sources added in quadrature. The most ubiquitous source of systematic uncertainty is the JMR uncertainty, followed by the b-tagging calibration uncertainty. The b-tagging uncertainty scales with number of b-tags so at 3000 GeV where 4b is disregarded, the uncertainty becomes small. For the 4000 and 5000 GeV mass points, the uncertainty reemerges as a dominant one when b-tagging efficiency begins to degrade.

Uncertainty Source	$m_X = 1000GeV$	1600	2000	3000	4000	5000	Mean
Luminosity	0.039%	0.12	0.35	0.53	0.17	0.18	0.23
JMR	4.3	7.4	16	9.5	7.5	6.7	8.6
JES	0.013	0.096	0.29	0.53	0.005	0.25	0.20
JMS	0.074	1.4	1.3	5.3	1.6	0.59	1.7
b-tagging	0.69	1.8	3.4	0.78	7.4	6.5	3.4
$t\bar{t}$ Modelling	7.0	0.66	0.86	0.79	0.17	0.44	1.7
Signal Theory	1.8	1.5	1.4	1.6	1.1	1.2	1.4
Background Estimation	20	1.2	0.97	1.3	1.7	0.94	4.4
Combined	22	8	17	11	11	9	11

Table 8.1. The percentage impact of each systematic uncertainty on the final limit for six Narrow-width Scalar signal samples.

CHAPTER 9

RESULTS

With the procedure of the analysis outlined, the boosted analysis demonstrates proper validation and proof of concept to officially unblind the data SR. Data and background are compared in the SR to look for significant differences between them. Furthermore, upper limits on the cross-section of $HH \to b\bar{b}b\bar{b}$ on the two signal models are evaluated separately and limit plots are generated to demonstrate both.

9.1 Statistical Analysis

A full statistical analysis is conducted to determine any existence of signal this analysis. Furthermore, the upper limit on the cross section both signal models is set. The goal of the statistical analysis is provide a model to describe the data according to different assumptions such as the hypothesis of signal in addition to background (signal+background hypothesis) or background only hypothesis. The statistical analysis framework provides a median, expected significance of a signal model and demonstrates the total expected fluctuations about the median. The results of this statistical analysis is displayed as a "limit plot", where the expected limit is shown along with one and two standard deviation ("σ") "bands". This is done for separately for each signal model, where the significance is found for the mass points of a signal model and combined into one limit plot. The bands are characterized by the uncertainties, which include statistical and systematic sources. The significance of the data is then analyzed in a similar way to give an observed limit. The observed limit is compared to the expected limit to look for statistical excesses or to validate the null hypothesis. If the observed limit differs from the expected limit by $\geq 5\sigma$, it may be an indication of a discovery.

A test statistic $\hat{q}_\mu$ is introduced to study the compatibility of data and the signal model as defined in Reference [36]. The parameter μ represents the signal strength, where $\mu = 0$ is the background-only hypothesis. The test statistic uses a one-sided negative log-likelihood and presents three possible scenarios, which are,

$$\hat{q}_\mu = \begin{cases} -2\ln\frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(0,\hat{\theta}(0))} \text{ for } \hat{\mu} < 0 \\ -2\ln\frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(\hat{\mu},\hat{\theta}(\mu))} \text{ for } 0 < \hat{\mu} < \mu \\ 0 \text{ for } \mu < \hat{\mu} \end{cases} . \tag{9.1}$$

The function L is the likelihood function that is the product of the signal plus background Poisson probabilities of all the bins in the m_{HH} histogram. The function θ characterizes the probability density function shape for signal and background and contains a combination of nuisance parameters ("NPs") of the background and signal. In general, NPs are other parameters associated to the statistical analysis such as sources of uncertainties and their values are notably not know a priori but only after fitting to the data. The parameters in the denominator $\hat{\mu}$ and $\hat{\theta}$ are the values of μ and θ that maximize L. The parameter in the numerator $\hat{\hat{\theta}}$ is the value of θ that maximizes the likelihood given a value of μ.

This test statistic definition is chosen with the expected behaviour of our hypotheses. That is, the presence of signal in the boosted analysis is expected to be seen as additional events to the background only distribution. Upward fluctuations of the signal ($\mu < \hat{\mu}$) are made sure to not serve as evidence against the signal by setting $\hat{q}_\mu = 0$. For downward fluctuations in data ($\hat{\mu} < 0$), μ in the denominator likelihood function is set to zero to make sure it does not serve as evidence against the background.

The test statistic is then used to calculate the 95% Confidence Level statistic (CL_S) [36]. This is defined as,

$$CL_S = \frac{P_{s+b}(\hat{q}_\mu \leq \hat{q}_{\mu obs})}{P_b(\hat{q}_\mu \leq \hat{q}_{\mu obs})}, \tag{9.2}$$

where $P_{s+b}(\hat{q}_\mu \leq \hat{q}_{\mu obs})$ is the probability of the signal+background model to produce better agreement to data than observed and $P_b(\hat{q_\mu} \leq \hat{q_{\mu_{obs}}})$ is the probability of the

background only model to produce better agreement to data than observed. The variable μ is varied until $CL_S = 0.95$ is found. This is done for each mass point in their respective signal model and an upper limit on the cross section is derived.

9.2 Unblinded Signal Region

The statistical analysis is applied to the background estimation, which ultimately pulls the "pre-fit" background distribution derived with the background estimation procedure closer to that of the data. The resulting "post-fit" distributions allow for a better comparison of data to background but will still allow for extra events not from the background to be apparent. There has been no excess of events seen in either the pre-fit or post-fit plots, which indicates no observation of signal events are seen with any confidence for this analysis.

Figure 9.1 shows the pre-fit plots for background, data, and signal. Simulated Spin-0 resonant samples are scaled to arbitrary values and plotted to demonstrate their shape and position compared to the background estimation. The background has good agreement with the data for most points and any data point that deviates from the background appears to be the result of statistical fluctuations.

Figure 9.2 shows the post-fit plots for background, data, and signal. Simulated Scalar samples are scaled to arbitrary values and plotted to demonstrate their shape and position compared to the background estimation. Similar to the pre-fit plots, the background has good agreement with the data for most points and any data point that deviates from the background appears to be the result of statistical fluctuations.

9.3 Testing the Background-only Hypothesis

The probability that a set of data is compatible with a given model, is quantified by its p-value. The p-value "p_0" tests the compatibility of data with the background-only hypothesis. Another test statistic "q_0" is defined as,

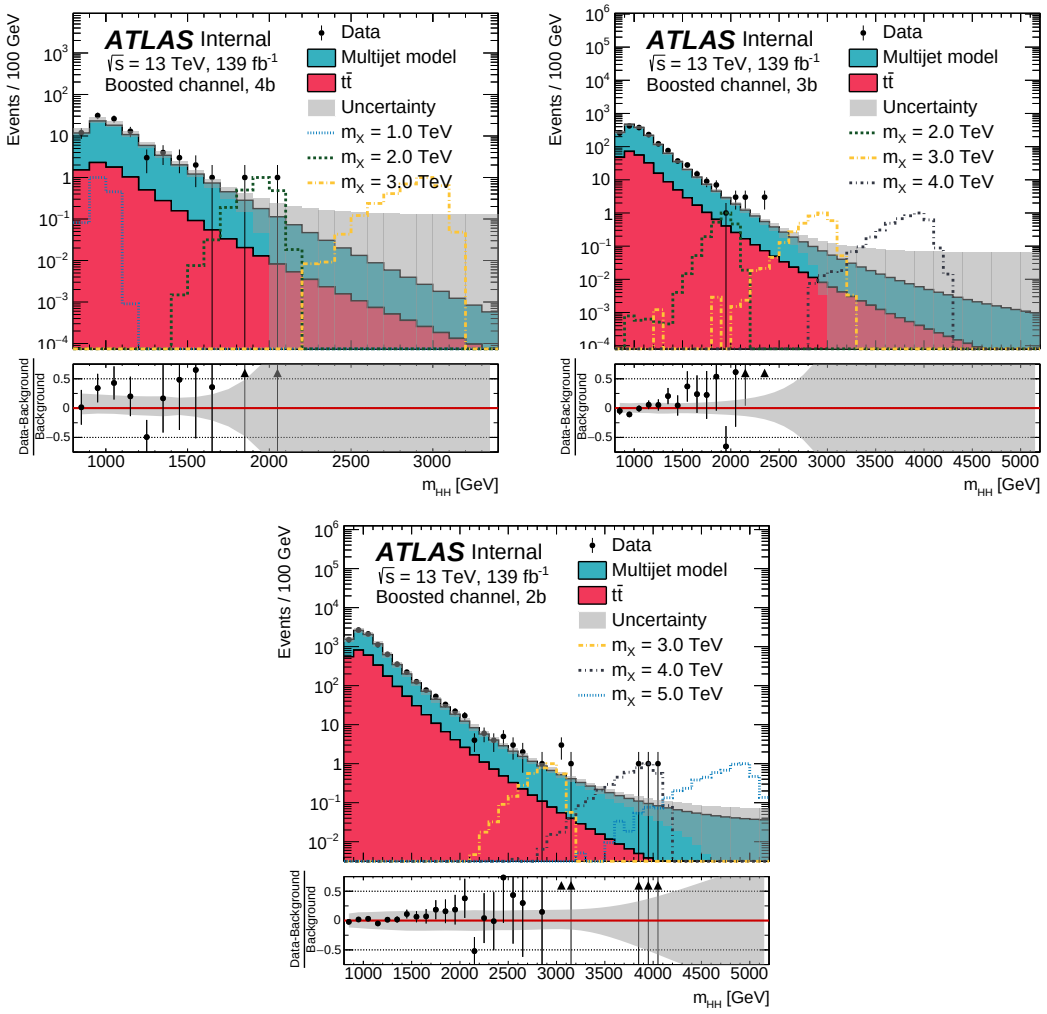


Figure 9.1. Pre-fit plots for the 4b, 3b, and 2b-split regions on the top left, top right, and bottom respectively. Three simulated Scalar distributions are shown, overlaid for reference.

$$q_0 = \begin{cases} -2\ln\frac{L(0,\hat{\hat{\theta}}(0))}{L(\hat{\mu},\hat{\theta}(0))} & \text{for } \hat{\mu} \geq 0 \\ 0 \quad \text{for } \hat{\mu} < 0 \end{cases}, \tag{9.3}$$

which corresponds to the background only hypothesis, $\mu = 0$. For this test statistic, a value of $\hat{\mu} > 0$ may constitute a discovery, albeit only for adequate standard deviations. A value of $\hat{\mu} < 0$ is, although not compatible with the background only hypothesis, likely the result of a systematic error.

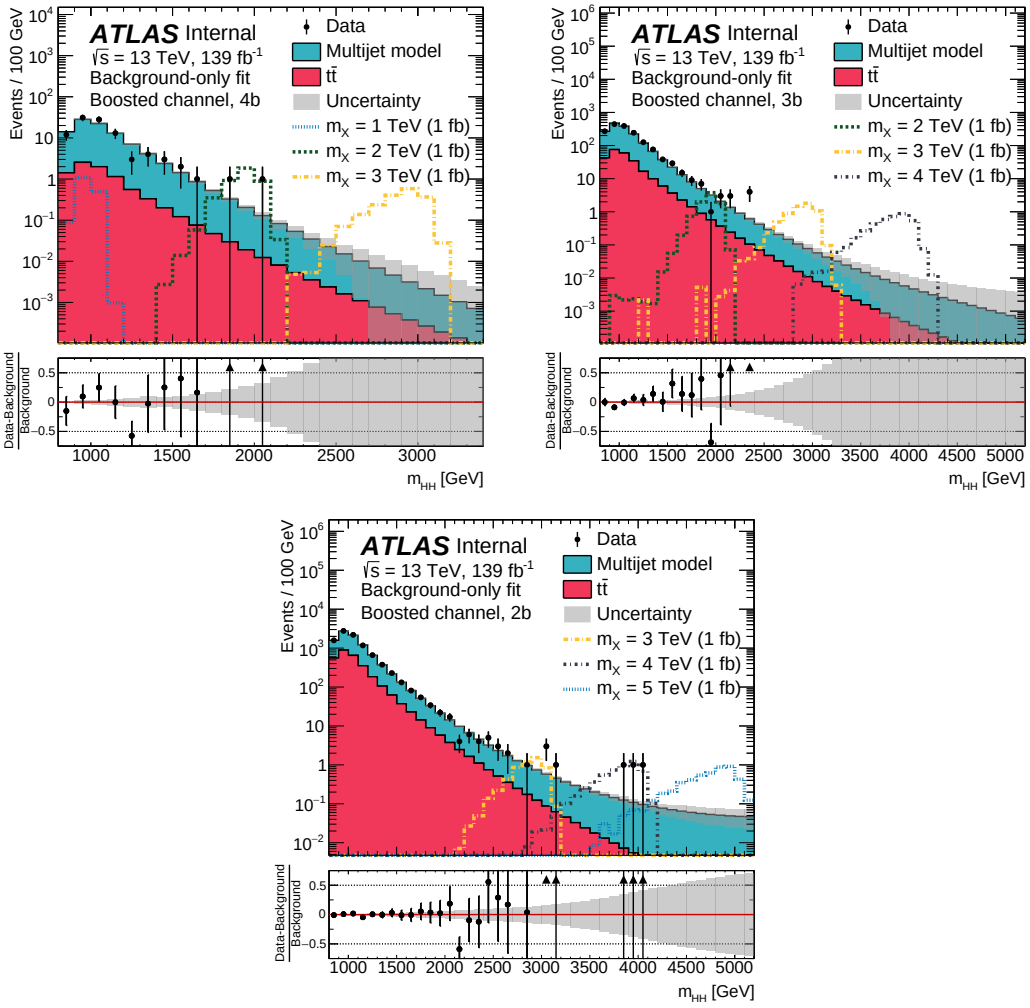


Figure 9.2. Post-fit plots for the 4b, 3b, and 2b-split regions on the top left, top right, and bottom respectively. Three simulated Scalar distributions are shown, overlaid for reference.

The background estimation is extrapolated into the VR, as opposed to the SR, and used as the basis of the background-only hypothesis since the VR is expected to have very low signal contamination. From this sample, q_0 for the background-only hypothesis is derived and compared to real data. This is done for both signal models with the largest excess appearing above 4 TeV for the Scalar model of just above 1σ. There is also a slight excess starting around 1.4 TeV for the both models, but it is

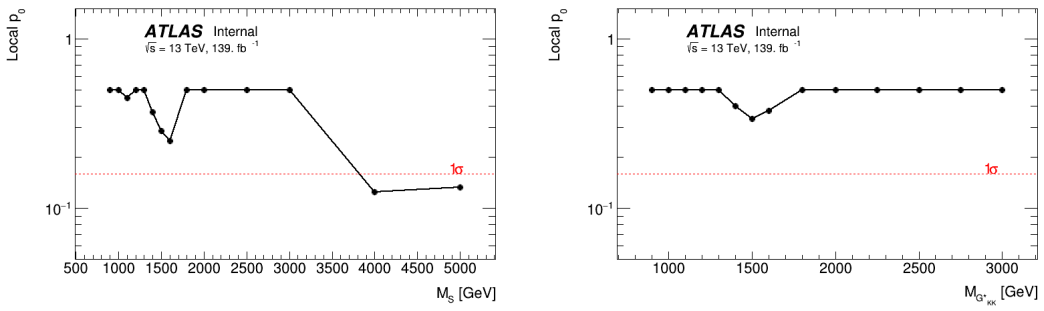


Figure 9.3. Local p_0 plots for the Scalar model and bulk RS model with c=1 shown on the left and right respectively.

well below 1σ and is not enough to suggest a signal. The rest of the points show very good compatibility with the background-only hypothesis as seen in Figure 9.3.

9.4 Expected and Observed Limits

Expected limits are drawn from pseudo experiments, which are generated with the background only hypothesis.[36] Figures 9.4 to 9.5 show the resulting limit plots for the Scalar model and bulk RS model with c=1.0 respectively. The figures show the combined (limit for the combined 4b, 3b, and 2b-split regions) expected limit, one and two sigma bands about the combined expected limit, and the expected limits produced with the individual tagging regions (dashed lines). The 4b region is most sensitive from 0.9 to 2 TeV signal mass and above 2 TeV the 3b becomes most sensitive. At very high resonant mass, the 2b-split channel begins to dominate in sensitivity. Note that only the 3b region is contributing to the entire mass range of the limit plots. The sensitivity of the 4b and 2b-split regions above 3 TeV and below 2 TeV respectively is essentially zero. Thus, they are neglected in the mass ranges where they do not contribute as to avoid any problems during the limit setting procedure, especially in 4b where the number of background events approach zero at high mass.

Figure 9.4 shows the limits obtained with only the statistical uncertainty considered and Figure 9.5 shows the limits obtained with all the sources of uncertainty listed

in Section 8 included. The statistical-uncertainty-only upper limit is better than that of the one with full systematic uncertainties included, as demonstrated by Figure 9.6. The systematic uncertainties worsen the limits by 10-30% for most masses and towards the lowest masses the limits degrade by as much as 50%. Also, the limits with full systematic uncertainties have wider bands than the limits with only statistical uncertainties.

9.5 Pull and Impact plots

Pull and impact plots are made to assess the effect NPs have during the fitting procedure. A pull is meant to quantify any deviations from the nominal value of a given NP, normalized by it's uncertainty and are demonstrated with "pull plots" for each individual mass point. An impact measures the effect the varying a given NP by $\pm 1\sigma$ has on $\hat{\mu}$, written as μ_{var}. This is then compared to $\hat{\mu}$ as $\hat{\mu} - \mu_{var}$ and plotted as an "impact plot".

Figures 9.7 and 9.8 show the pull and impact plots for both signal models for three different mass points. Only the top ten ranked NPs are shown for each mass point. The impacts are shown as the red and blue bars whereas the pulls are the black dots and lines. For every mass point the JMR (JET_MassRes_Hbb_comb) is a top six ranking source of uncertainty. The non-closure (DijetShape) is large in the 4b region for low mass points, which is likely due to turn on effects of the boosted analysis efficiency in that mass range. The flavor tagging extrapolation uncertainty is written as "FT_EFF_extrapolation", which is not to be confused with the background estimation uncertainty written like "ExtrapUncert_bXb".

9.6 Combined Results

Post unblinding results from both the boosted and resolved analyses are combined. The resolved analysis is entirely orthogonal to the boosted analysis, thanks to the

resolved veto, and features one SR with 4 b-tagged pflow jets. The background estimation is done using machine learning, which trains CR data events with 2 b-tags to target data events with 4 b-tags. The resolved analysis is sensitive to resonant masses between 250 and 1400 GeV, which means it shares six mass points with the boosted analysis: 900, 1000, 1100, 1200, 1300, and 1400 GeV.

Figure 9.9 shows the resolved limits combined with the boosted limits. These figures fully account for the systematic uncertainties in both analyses, which are mostly uncorrelated between the analyses except for the signal modelling luminosity uncertainties. A 2σ bump in the observed limit is seen around the 1100 GeV mass point and a $1.5 - 2.0\sigma$ bump is seen for the 1200, 1300, and 1400 GeV mass points. This is vetted by both analyses and the reason is attributed to applying the resolved background estimation at a relatively high mass (1100 GeV or greater). A high mass extension of the resolved analysis to 1400 GeV is relatively new to help recover some events in the overlap region and allow a smooth transition into the boosted only limit. The desired effect is achieved, but it pushes the resolved background estimation to the edge of it's sensitivity, where there proves to be a slight under-prediction in background events and therefore a slight loss of sensitivity. Nonetheless, the excesses are still within an acceptable threshold to not be considered problematic and therefore the combined limits remain valid. The values from the combined limits are shown in Tables 9.1 and 9.2. The bulk RS model with c=1.0 is shown on the respective limit and is excluded for a $G^*_{\mathtt{KK}}$ mass range of 299 GeV to 1450 GeV.

9.7 Previous Result: Comparison

A comparison is made of the current boosted analysis to the partial `Run-2` boosted analysis from 2018 done with 36.1 fb^{-1} of data from the ATLAS detector. Although the systematic uncertainties defined may be somewhat similar between both iterations of the analysis, the statistical uncertainty only limits provide the best point of

mass	observed	expected	-2 σ	-1 σ	-1 σ	-2 σ
251	1592.64	1226.73	572.08451	786.605	2324.89	6234.26
260	3606.59	2369.64	1194.83	1621.71	3785.55	6297.23
280	3007.66	1706.11	832.694	1134.71	2573.07	4429.77
300	1386.3	1040.16	462.745	644.49	1814.38	3595.64
400	58.5206	102.826	49.31231	67.41372	181.996	418.975
500	31.1503	34.93576	17.23582	23.49	58.886	116.4852
600	9.71486	16.83130	9.03377	12.1279	25.898	49.78110
700	7.92956	9.03748	4.42449	6.03788	15.44134	32.91360
800	4.20776	5.73665	2.83989	3.87598	9.55585	18.52957
900	4.24107	3.95143	2.12083	2.84722	5.94492	10.31780
1000	2.39099	2.87652	1.54390	2.07269	4.30491	7.28766
1100	5.76745	2.18924	1.17502	1.57747	3.29673	5.59981
1200	3.32786	1.80702	0.96987	1.30206	2.67113	4.21109
1300	1.99533	1.52448	0.81823	1.09847	2.23396	3.32406
1400	2.68177	1.30766	0.70186	0.94225	1.94966	2.95545
1500	2.55517	1.14572	0.61494	0.82555	1.74177	2.71242
1600	1.83856	1.08390	0.58176	0.78101	1.70590	2.82807
1800	1.17845	0.84575	0.45394	0.60941	1.38013	2.43143
2000	1.00842	0.68541	0.36788	0.49388	1.14780	2.12967
2500	0.69866	0.50185	0.26936	0.36161	0.84816	1.61330
3000	0.41157	0.41639	0.22349	0.30003	0.68510	1.26885
4000	0.92170	0.43815	0.23517	0.31571	0.72140	1.37039
5000	0.42825	0.49268	0.26444	0.35500	0.82107	1.58822

Table 9.1. Table of limits for the spin-0 cross section.

comparison. Note that the current, full `Run-2` analysis provides over three times the number of events, which will by default improve the limit without any other changes. Therefore to facilitate a fair comparison of the sensitivities of both analyses, the current analysis has been scaled to 36.1 fb^{-1} , which is the total integrated luminosity of the partial `Run-2` analysis.

Figures 9.10 and 9.11 show a comparison between the partial `Run-2` limits and the down-scaled version of the current limits for both signal models. The current analysis demonstrates an improved limit in this comparison. The biggest change that results in an improvement is attributed to the integration of VR track jets. The previous iteration of the boosted analysis used fixed radius track jets, which are less efficient

mass	observed	expected	-2 σ	-1 σ	-1 σ	-2 σ
251	220.528	451.231	236.007	319.748	703.863	1155.0
260	255.479	552.453	296.516	398.073	819.538	1363.33
280	507.339	777.71	417.417	560.383	1139.48	1819.8
300	539.73275	704.966	344.779	469.419	1157.03	1891.12
400	42.573	81.806	40.69330	55.3215	137.545	278.0787
500	23.7056	30.3855	15.35256	20.83095	49.4737	89.03102
600	8.46964	15.01471	7.61368	10.32869	24.1678	42.95937
700	7.41843	8.64468	4.33954	5.89981	14.04319	25.31185
800	4.16766	5.55428	2.82901	3.84424	8.73169	14.58303
900	4.24107	3.95143	2.12083	2.84722	5.94492	10.31780
1000	2.39099	2.87652	1.5439	2.07269	4.30491	7.28766
1100	5.76745	2.18924	1.17502	1.57747	3.29673	5.59981
1200	3.32786	1.80702	0.96987	1.30206	2.67113	4.21109
1300	1.99533	1.52448	0.81823	1.09847	2.23396	3.32406
1400	2.68177	1.30766	0.70186	0.94225	1.94966	2.95545
1500	2.55517	1.14572	0.61494	0.82555	1.74177	2.71242
1600	1.69384	0.95281	0.51140	0.68655	1.50882	2.53345
1800	1.12791	0.73426	0.39410	0.52908	1.19701	2.10239
2000	0.99973	0.59902	0.32151	0.43163	1.00303	1.85060
2250	0.84706	0.48326	0.25938	0.34822	0.81458	1.52756
2500	0.56258	0.41454	0.22249	0.29870	0.69549	1.30200
2750	0.29911	0.37237	0.19986	0.26831	0.62022	1.15300
3000	0.33052	0.33963	0.18229	0.24473	0.55676	1.01451

Table 9.2. Table of limits for the G^*_{KK} cross section.

especially at high boost. The next biggest sensitivity improvement to the analysis is from the latest flavor-tagging. The DL1r flavor-tagging algorithm was introduced to the analysis in late 2018 and shows better b-tagging efficiency for higher p_T jets, when compared to it's predecessor MV2c10 flavor tagging algorithm. Another contribution that affects the limit is the reoptimized cut on $\Delta\eta$. Note that the new analysis introduces two additional signal mass points for the Spin-0: 4000 and 5000 GeV.

9.8 Full Run-2 diHiggs Analyses

As of Spring in 2021, there are two diHiggs analyses with published results for the full Run-2 data set. The $b\bar{b}\gamma\gamma$ channel and $b\bar{b}\tau\tau$ boosted channel have publications that can respectively be found in References [2] and [35]. Neither publication appears to have any significant indication of signal and have set upper limits on the cross section of the Scalar model, much like the $b\bar{b}b\bar{b}$ channel. The limits are shown together in Figure 9.12. The $b\bar{b}\tau\tau$ has only published results for the boosted channel, thus the limit starts at 1 TeV. The $b\bar{b}\gamma\gamma$ analysis is only sensitive up to 1 TeV so its full range is from 251-1000 GeV. Note that the $b\bar{b}\gamma\gamma$ analysis also considers several other mass points of the Spin-0 resonance. For much of the mass spectrum the $b\bar{b}b\bar{b}$ is the most sensitive diHiggs channel, except for the mass points below 400 GeV where $b\bar{b}\gamma\gamma$ is dominates. The $b\bar{b}\tau\tau$ limit is sensitive from 1000-3000 GeV and has discontinuities at 1600 and 2500 GeV due to different resonant mass dependent requirements imposed on the analysis.

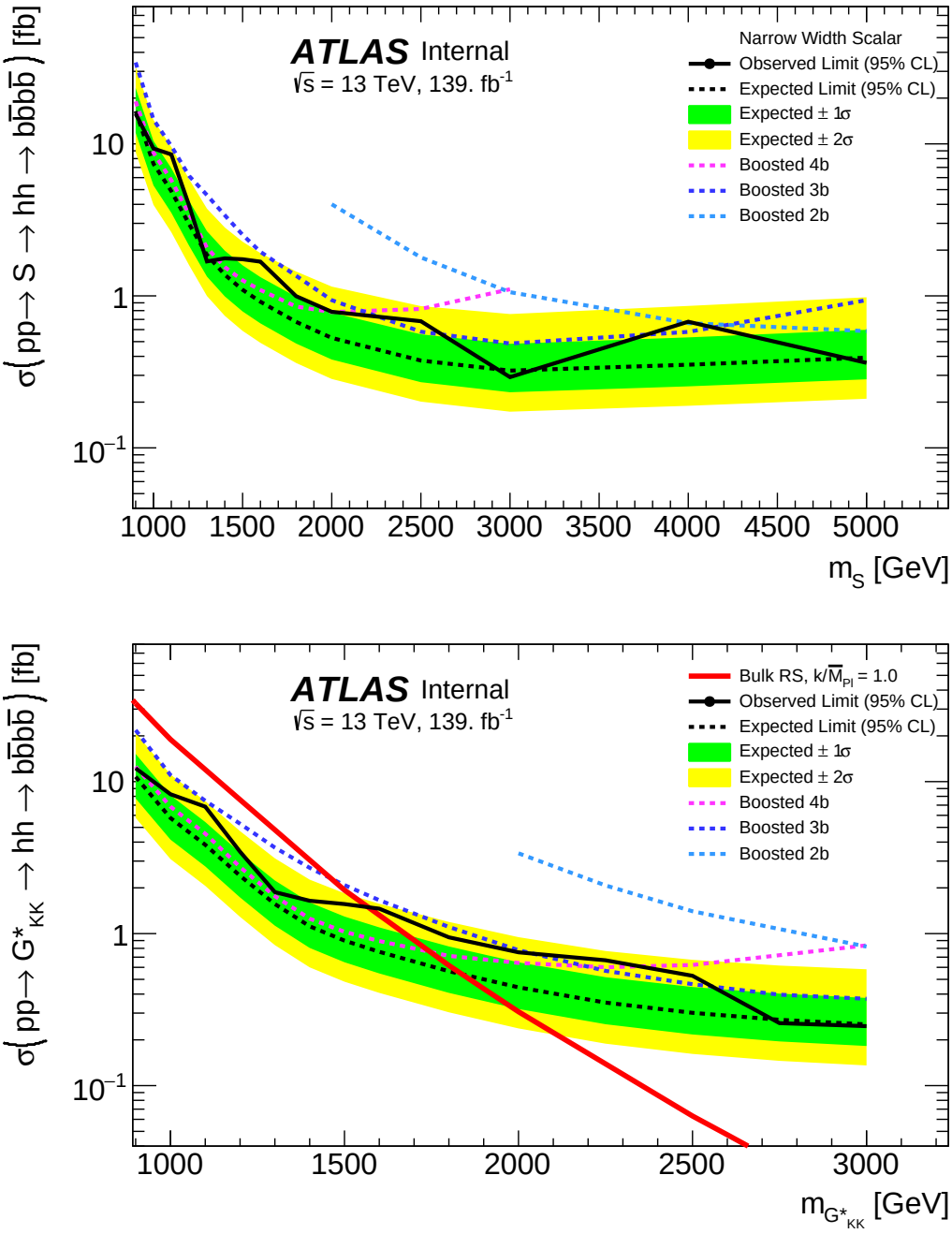


Figure 9.4. Limit plots with statistical uncertainties only. The Scalar model and the Bulk RS model with $k/\bar{M}_{Pl} = 1.0$ are shown on the top and bottom respectively. The expected limits for the 4b, 3b, and 2b-split channels are shown in pink, blue, and azure, respectively.

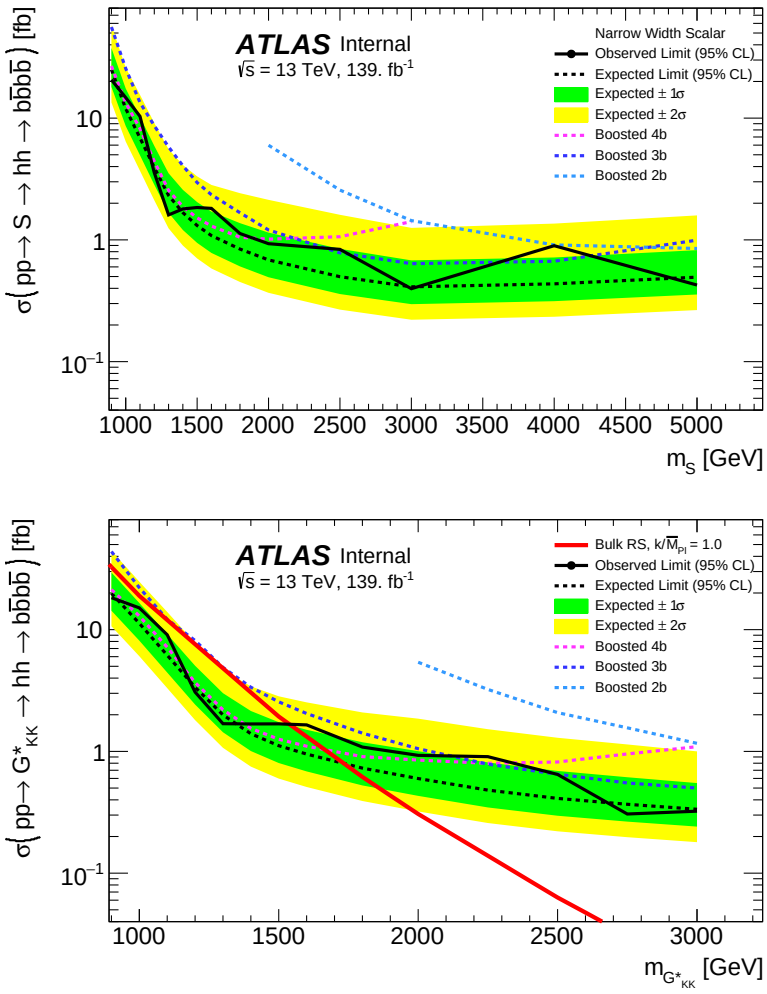


Figure 9.5. Limit plots with fully accounted for systematic and statistical uncertainties. The Scalar model and the Bulk RS model with $k/\bar{M}_{Pl} = 1.0$ are shown on the top and bottom respectively. The expected limits for the 4b, 3b, and 2b-split channels are shown in pink, blue, and azure, respectively.

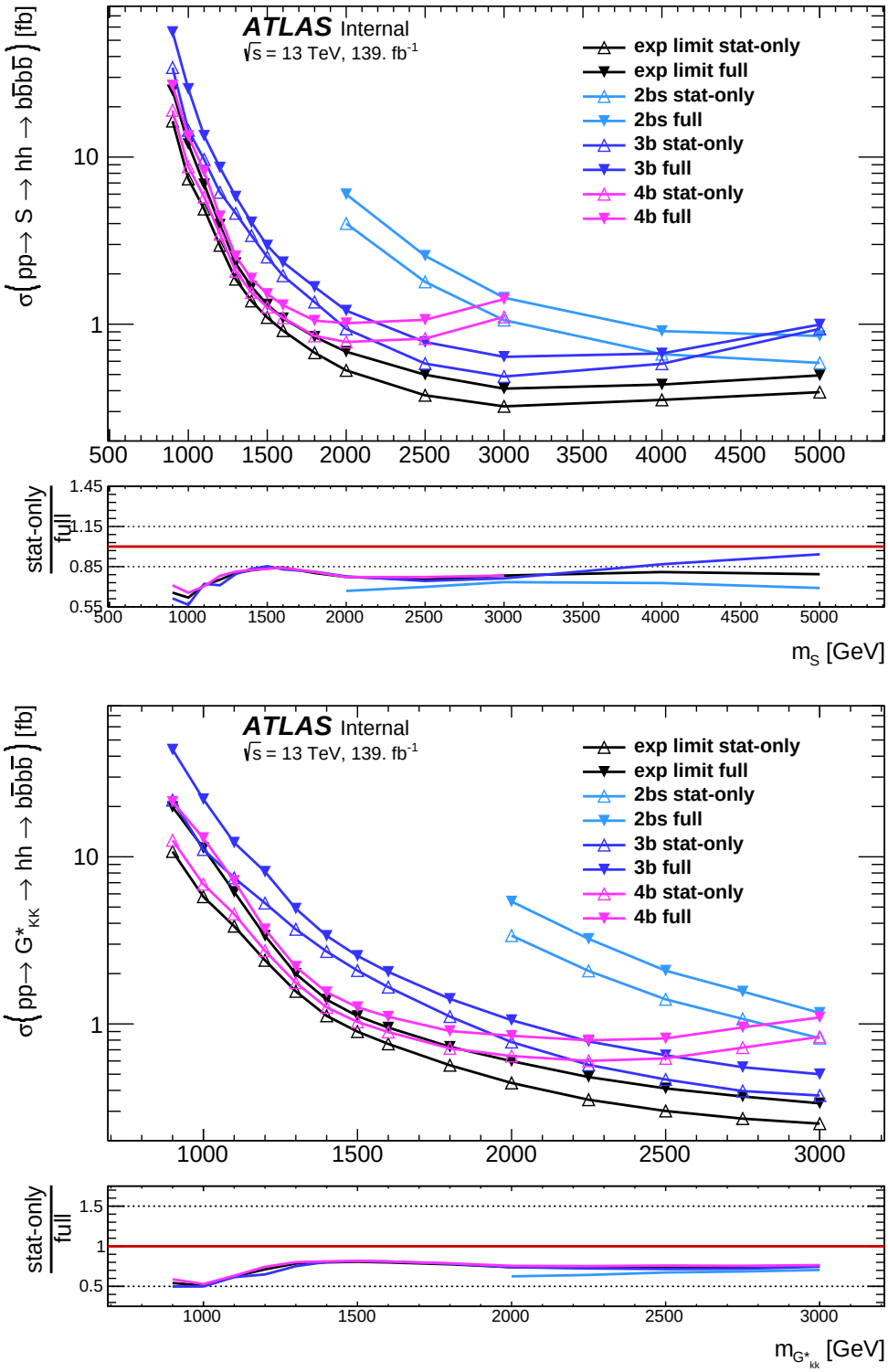


Figure 9.6. Comparison of the nominal lines for the combined, 4b, 3b, and 2b-split limts with ("full") and without ("stat-only") systematic uncertainties included.

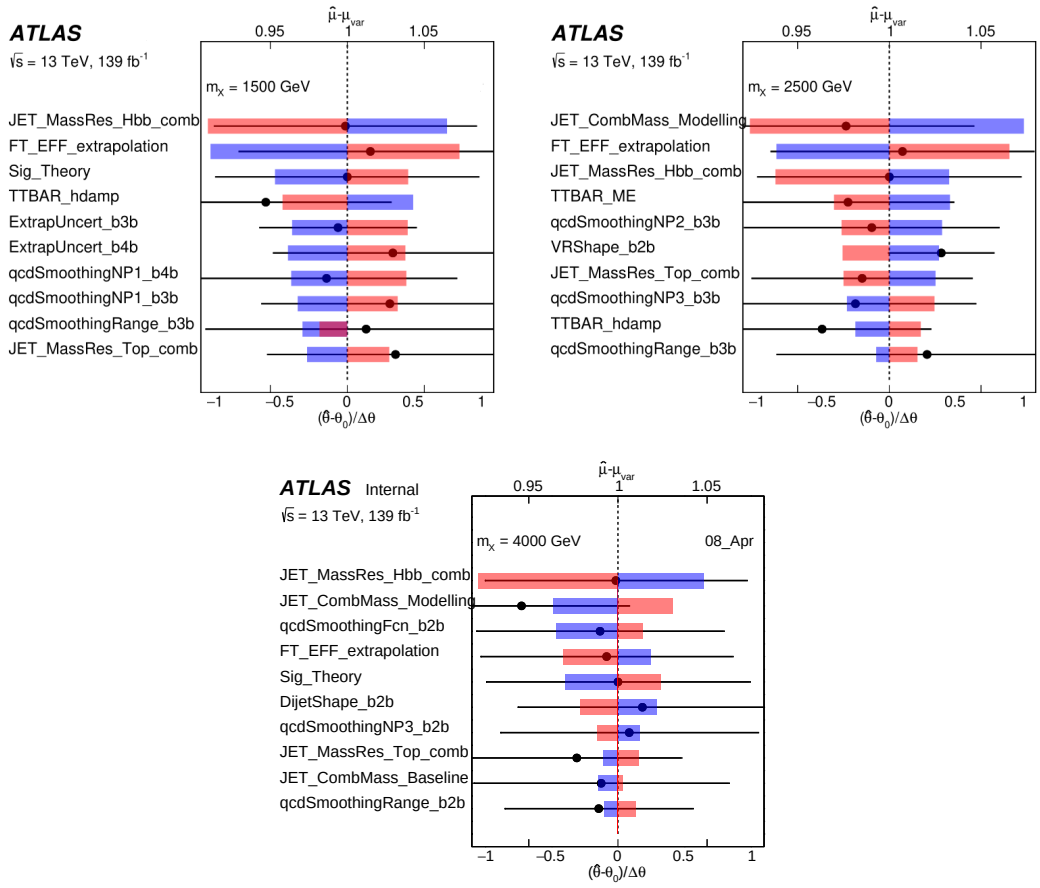


Figure 9.7. Pull and impact plots for the Spin-0 for three mass points. Only the top ten ranking NPs are shown.

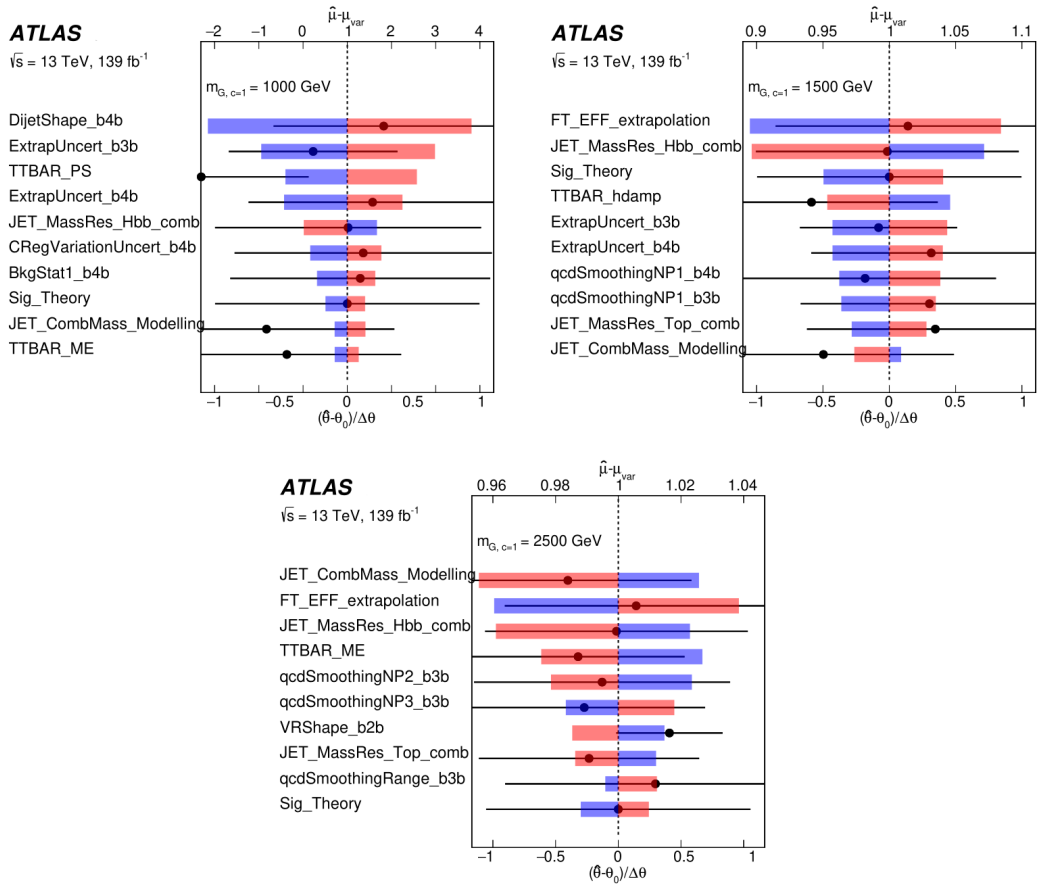


Figure 9.8. Pull and impact plots for the G^*_{KK} for three mass points. Only the top ten ranking NPs are shown.

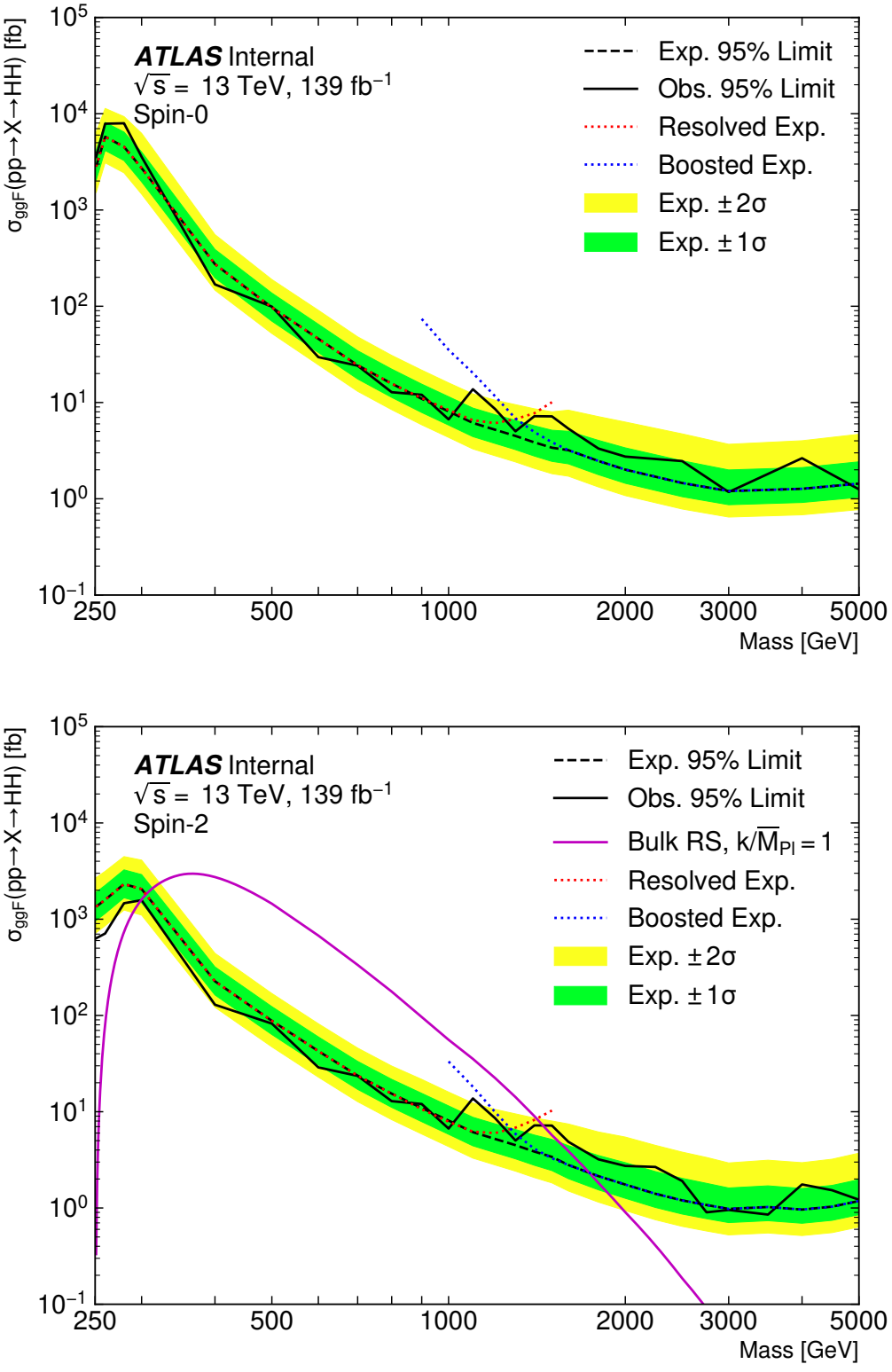


Figure 9.9. Resolved and boosted combined limits for the Scalar (top, "spin-0") and bulk RS model with c=1.0 (bottom, "spin-2") models.

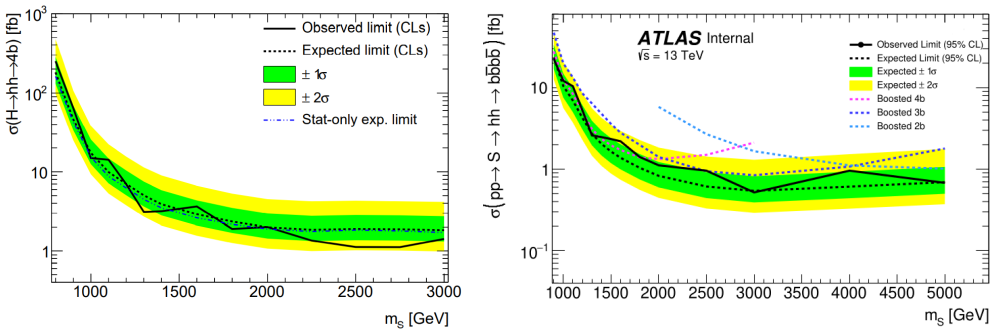


Figure 9.10. The statistical uncertainty only limits for the Scalar model. The left is the partial `Run-2` result and the right is the current result scaled the number of events of partial `Run-2` . Note that the current result also has two additional mass points 4000 and 5000 GeV.

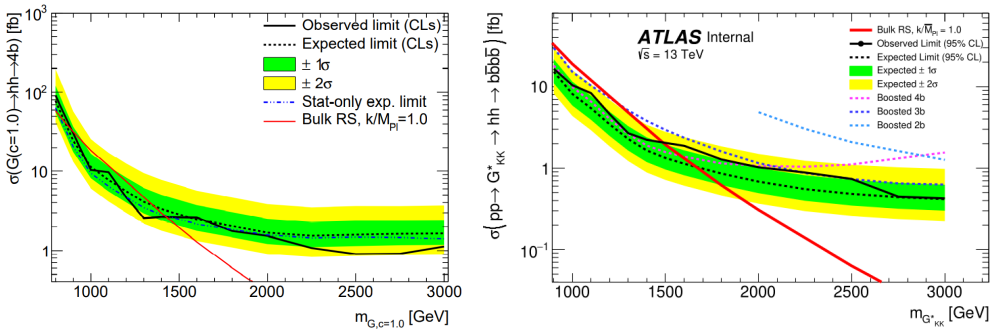


Figure 9.11. The statistical uncertainty only limits for the bulk RS model with c=1.0. The left is the partial `Run-2` result and the right is the current result scaled the total integrated luminosity of partial `Run-2` .

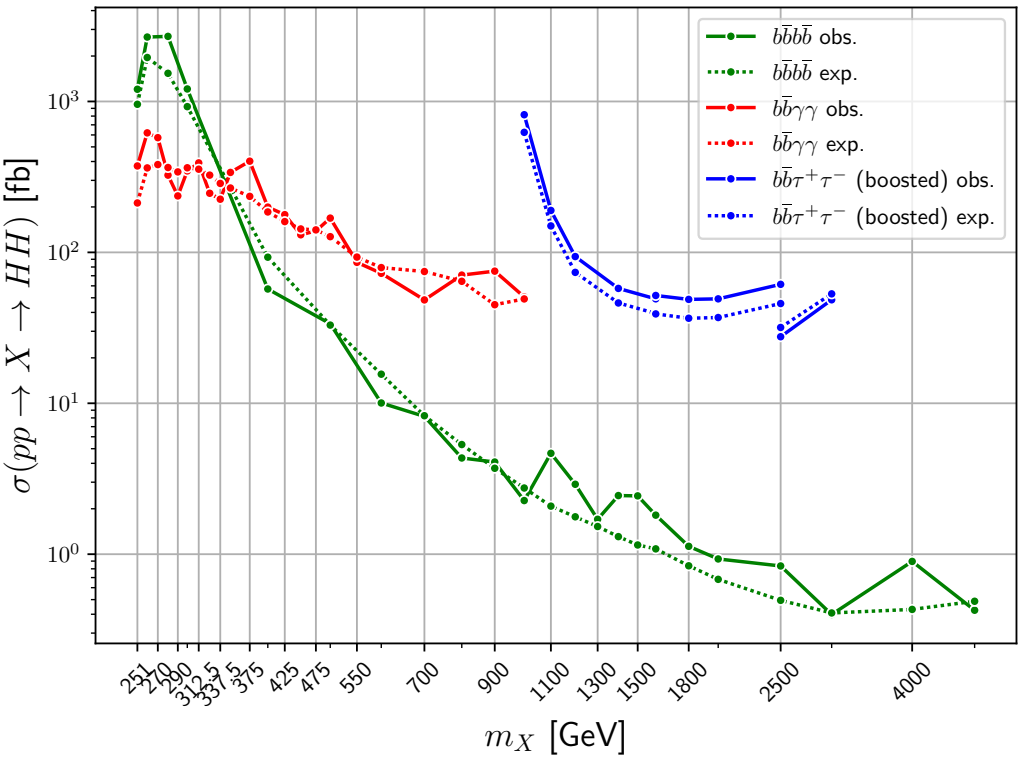


Figure 9.12. Observed ("obs.") and expected ("exp.") limits for three diHiggs analyses: $b\bar{b}b\bar{b}$, $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau\tau$ (boosted) shown in green red and blue respectively. All three analyses use the full `Run-2` dataset.

CHAPTER 10

CONCLUSION

The $HH \to b\bar{b}b\bar{b}$ boosted analysis for the full Run-2 data set has been an overall success. Expanding on partial Run-2 results at 36.1 fb^{-1} , the analysis has improved sensitivity greatly due to the addition of variable radius track jets, improved b-tagging, and larger data set of 139 fb^{-1} . We continue to set limits along with the resolved analysis on the $pp \to HH \to b\bar{b}b\bar{b}$ cross section, excluding resonances for a wide range of masses.

The future for physics analyses throughout CERN is bright. The official Run-3 data taking time period is expected to commence in late 2021, which will last about three years and will double the amount of data collected. By the eventual Run-4 and Run-5 is scheduled to commence by some time in late 2020 or early 2030, bringing the LHC and ATLAS up to it's full potential with the High Luminosity LHC ("HL-LHC"). During the time of the HL-LHC the number of events collected should increase by an order of magnitude. For the boosted $b\bar{b}b\bar{b}$ analysis, the increase in data from future runs will improve the sensitivity of the analysis in the areas particularly limited by low number of events. This will dampen the affect of statistical uncertainty and improve our background estimation, which will ultimately help to exclude the observation of high mass resonances that we study in this analysis.

Aside from the increase in data, the boosted analysis can utilize some emerging techniques in future searches. The first should be an NTrk cut, which is outlined in Appendix C.1. Harmonizing the PS generators between signal samples will allow us to utilize this discriminant to increase the sensitivity of the analysis. Another strategy to adopt is background reweighting done with machine learning. The spline reweighting technique outlined in Appendix E proves a sufficient method for this analysis but in future analyses it should be compared to a method with machine learning to check for sensitivity gains and effects on the uncertainties. Lastly, b-tagging is a difficult and complicated process in which improvements have been made to date in b-tagging

efficiency and fake rejection. Progress will continue to be made, especially with the introduction of increased pile-up expected to come with the HL-LHC.

Standard model non-resonant $HH \rightarrow b\bar{b}b\bar{b}$ searches probe the important Higgs self-coupling term. The boosted analysis, however, is currently on the fringe of the sensitivity to the SM non-resonant search and at this time is negligible. With improvements the LHC upgrades, ATLAS improvements, and possibly several more iterations of this analysis, the sensitivity of the boosted channel may hopefully be able to contribute SM non-resonant search in the future.

Contributing to this analysis and being part of such an outstanding and brilliant team has been a unique privilege that few people will ever experience. The hard work I have poured into this project was worth the effort. Whether in failure or in success, the chance to do pure research, exploration, and expression of creativity is the truest path to meaning one can follow and I am grateful for the opportunity. Ishmael writes,

> *"There is a wisdom that is woe; but there is a woe that is madness. And there is a Catskill eagle in some souls that can alike dive down into the blackest gorges, and soar out of them again and become invisible in the sunny spaces. And even if he for ever flies within the gorge, that gorge is in the mountains; so that even in his lowest swoop the mountain eagle is still higher than other birds upon the plain, even though they soar."*

APPENDIX A

SEMIMERGED STUDY

A.1 Studies on a Semi-Merged Region for the Analysis

A.1.1 Introduction

As part of the full run 2 search of the channel, studies on an intermediate regime between the previously defined resolved and boosted analyses have been conducted. The studies of this new regime, which is named semi-merged, is motivated by previous evidence of a decline in sensitivity for the resolved and boosted analyses between a dihiggs mass of 800GeV and 1400GeV. The ultimate goal of this study is to see if semi-merged should be included in the full run 2 analysis by demonstrating that it adds a worthwhile contribution to our combined analysis.

A.1.2 Samples

Samples used in the study were Monte Carlo signal samples. Bulk RS KK Graviton (Madgraph, Pythia8) with $c = 1.0$ and standard model (Herwig7) monte carlo samples were used for the resonant and non-resonant studies respectively. Of the 20 masspoints generated, 600, 800, 1000, 1200, and 1400GeV (DSIDs 301491, 301493, 301495, 301497, and 301499 respectively) were chosen for the resonant study as a semi-merged analysis displayed virtually no sensitivity gain beyond those mass points. The non-resonant study features a full m_{hh} range from 0 to 3000GeV.

A.1.3 A Three Analysis Study

Studies of both resonant and non-resonant signals have been conducted. The resolved, boosted, and semi-merged analyses have been implemented up until but not including the signal region definition. Thus the selection process is an essential kinematic one, which will allow for a proper comparison of all three analyses' contributions without conducting semi-merged analysis in it's entirety.

For these analyses, variable radius trackjets were used with a 70% b-tagging working point. The boosted anaylsis was allowed one to two b-tags per both large-radius Jets (LR Jet), resulting in the standard boosted 4-tag, 3-tag and 2-tag split tagging scenarios. The resolved analysis requires four b-tagged small-radius jets (SR jet). The semi-merged analysis uses one LR Jet with two b-tags and two b-tagged SR jets.

Higgs candidates are formed from a pair of SM jets (resolved candidate) or a LR Jet (boosted candidate) passing the selection. The semi-merged analysis is constructed out of one Higgs boson candidate from the boosted analysis and another from the resolved analysis. Much of the semi-merged selection (table 1) is borrowed from the other two analyses for their respective Higgs candidates. An addition requirement of $\Delta R_{Jj} > 1.0$ is added to ensure separation between Higgs candidates. This cut was optimized on a 1TeV Bulk RS KK Graviton signal sample to minimize mismatching small-radius jets. To guarantee orthogonality events passing the resolved and boosted selection are vetoed from the semi-merged analysis. The boosted analysis also contains a resolved veto.

A yield study was performed on both resonant (A.1) and non-resonant (A.2) samples. For the resonant, the three selections were applied to five signal mass points

Resolved	Boosted	Semi-merged	
		Boosted Candidate	Resolved Candidate
$\|\Delta R_{jj}\| < 1.0$	$\|\Delta\eta_{JJ}\| < 1.7$	$\|\Delta R_{Jj}\| > 1.0$	$\|\Delta R_{jj}\| < 1.0$
$p_T^j > 40GeV$	$p_T^{lead,sub} > 450GeV, 250GeV$	$p_T^J > 250GeV$	$p_T^j > 40GeV$
$m^j > 50GeV$	$m^J > 50GeV$	$m^J > 50GeV$	$m^j > 50GeV$
$\|\eta_j\| < 2.5$	$\|\eta_J\| < 2.0$	$\|\eta_J\| < 2.0$	$\|\eta_j\| < 2.5$
1 b-tag per jet	1 or more b-tag per Jet	2 b-tags per Jet	2 b-tagged jets

Table A.1. The selections applied to the three analyses. The semi-merged has two separate selections for the boosted and resolved candidates.

600, 800, 1000, 1200, and 1400GeV. A.1 shows the yields as normalized to the amount of events originally in each respective mass point sample. By normalizing, we are able to observe a comparison between the mass points despite the slight differences in sample sizes. The semi-merged analysis shows a contribution, albeit small in comparison to the resolved plus boosted analyses, which peaks at 800GeV. For the non-resonant, A.2 shows the m_{hh} spectrum for three selections. The bulk of the yield is found in the resolved selection at low mass, with a contribution from the semi-merged between 600 and 1000GeV.

A.1.4 Conclusion

A preliminary yet fair comparison was performed on a semi-merged regime, which is made up of one boosted Higgs candidate and one resolved Higgs candidate. A selection process was used to ensure orthogonality and to select only potential Higgs candidates. Resonant and non-resonant signal yield studies show no significant contribution provided by a semi-merged regime as compared to the other analyses. The results do not warrant a semi-merged analysis to be included in the full run 2 paper for the channel at this time.

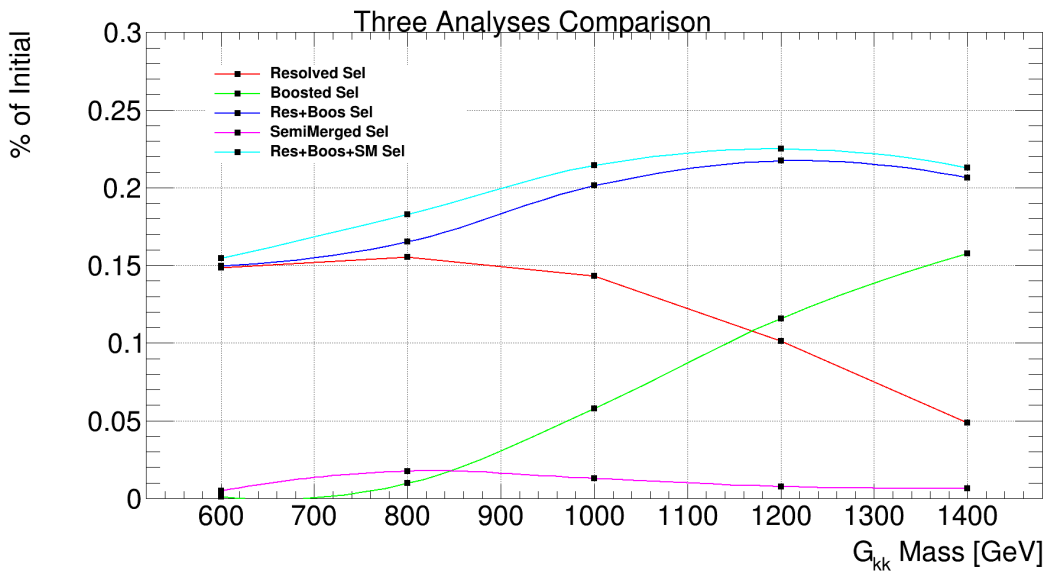


Figure A.1. Each selection is applied on same unselected 600, 800, 1000, 1200, and 1400GeV Monte Carlo G_{kk} signal samples. Boosted selection has a resolved veto and the semi-merged selection has a resolved and boosted veto.

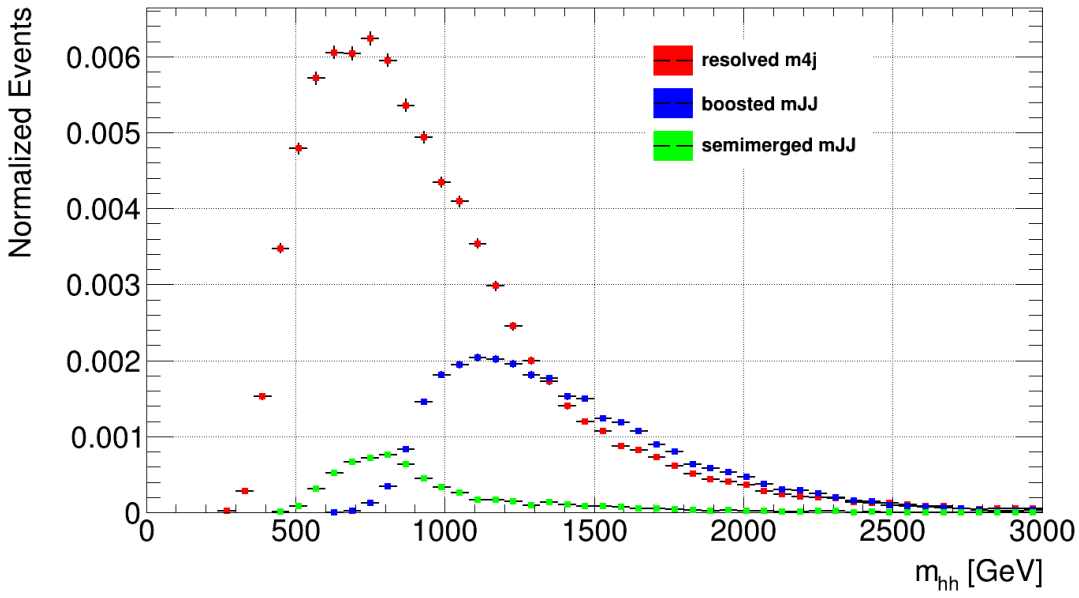


Figure A.2. The m_{hh} distribution divided by the total number of events. The unitless distribution compares the yield of events each analysis can provide for a standard model search. Boosted selection has a resolved veto and the semi-merged selection has a resolved and boosted veto.

APPENDIX B

SIGNAL CONTAMINATION

B.1 Signal Contamination

A test was conducted to check for the contamination of signal events in the VReg. The metric for signal contamination is the double ratio $\frac{S_{VReg}/B_{VReg}}{S_{SReg}/B_{SReg}}$, where S_{VReg} (S_{SReg}) represents the yield of events from a signal sample in the VReg (SReg) and B_{VReg} (B_{SReg}) represents the yield of estimated background events in the VReg (SReg). Fig. B.1 shows the double ratio for the full range of G_{kk} and Scalar signal samples. The contamination is most pronounced for the highest and lowest signal samples in the available range, but never exceeds an unacceptable level.

Similarly and for completeness, the study was repeated to test the signal contamination in the CReg with the double ratio, $\frac{S_{CReg}/B_{CReg}}{S_{SReg}/B_{SReg}}$. Fig. B.2 shows the double ratio for the full range of G_{kk} and Scalar signal samples. The contamination is low and is within an acceptable level for the entire range of signal samples.

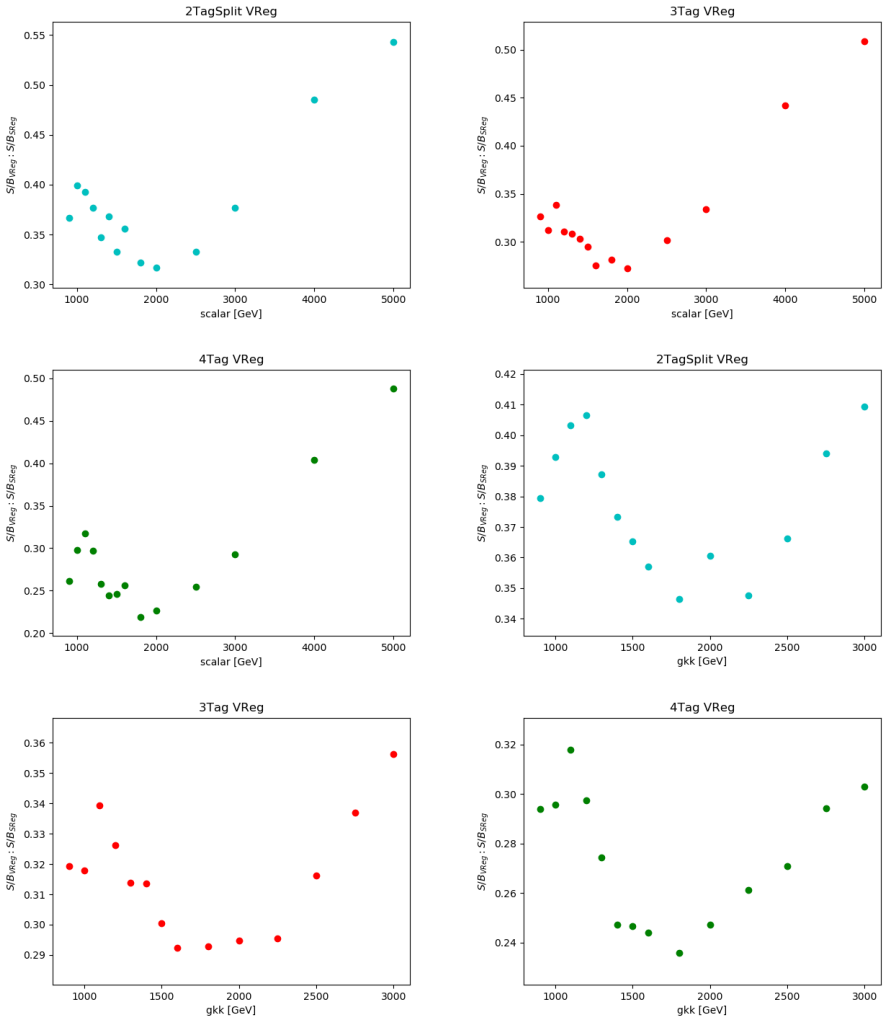


Figure B.1. Plots of the VReg signal contamination double ratio for each mass point of of the Scalar and G_{kk} signal samples. Plots for 2Tag, 3Tag, and 2TagSplit are in blue, red, and green respectively.

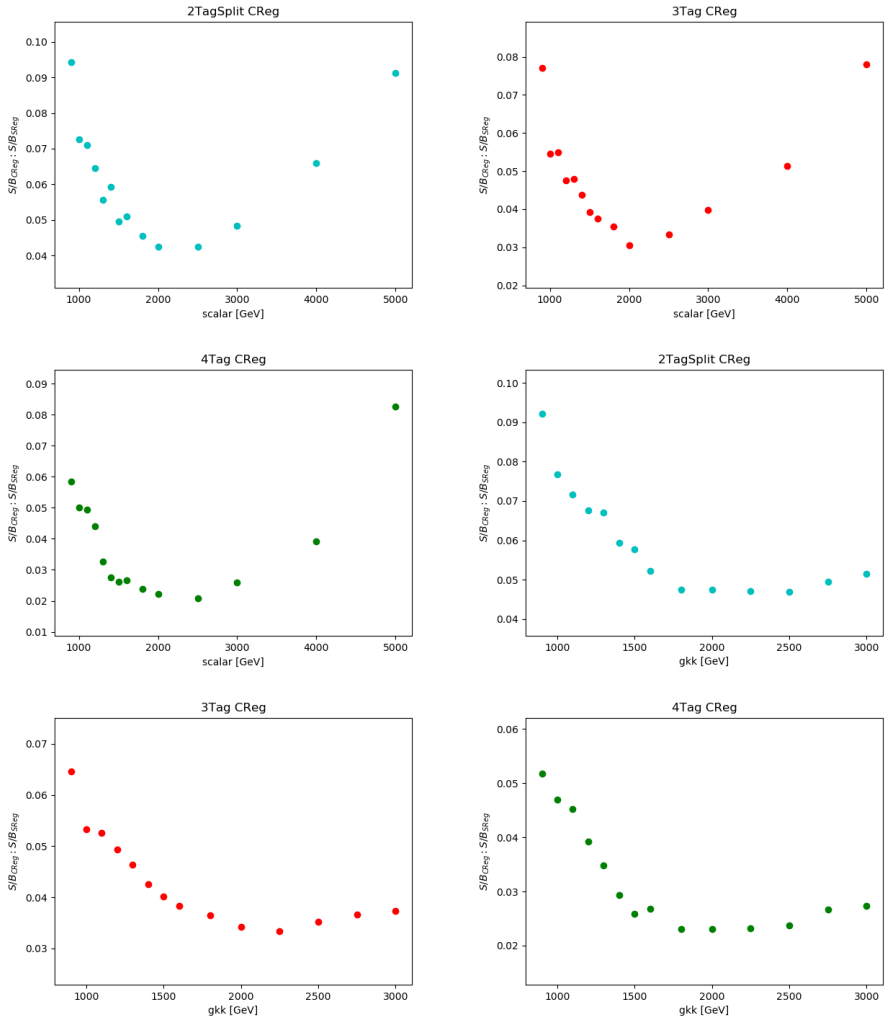


Figure B.2. Plots of the CReg signal contamination double ratio for each mass point of of the Scalar and G_{kk} signal samples. Plots for 2Tag, 3Tag, and 2TagSplit are in blue, red, and green respectively.

APPENDIX C

NTRK

C.1 NTrk

C.1.1 Introduction

The number of charged-particle tracks ghost-associated to the primary vertex of an untrimmed large radius jet is used as a background discriminant. This track multiplicity value ("NTrk") is higher for gluon initiated jets, which compose most of our primary background source of the form of dijet QCD. All tracks associated with the large-R jets are required to have $p_T > 0.5$GeV and $|\eta|<2.5$.

C.1.2 Motivation

The discrimination power of using NTrk is demonstrated Fig.C.1 and C.2. NTrk distributions for resonant signal MC and dijet MC are compared. In these distributions a small selection is required to remove "soft" jets, which would normally be removed through a standard selection process. The selection, of which the m_{JJ} requirement supplies the greatest soft jet discrimination, includes p_T^{Lead} >500GeV, $p_T^{Subleading}$ >250GeV, $|\eta|$ <2.0, and m_{JJ} >1500GeV. The selection is applied to trimmed jets, although the NTrk value is nevertheless derived from it's untrimmed parent.

C.1.3 Cut Optimization

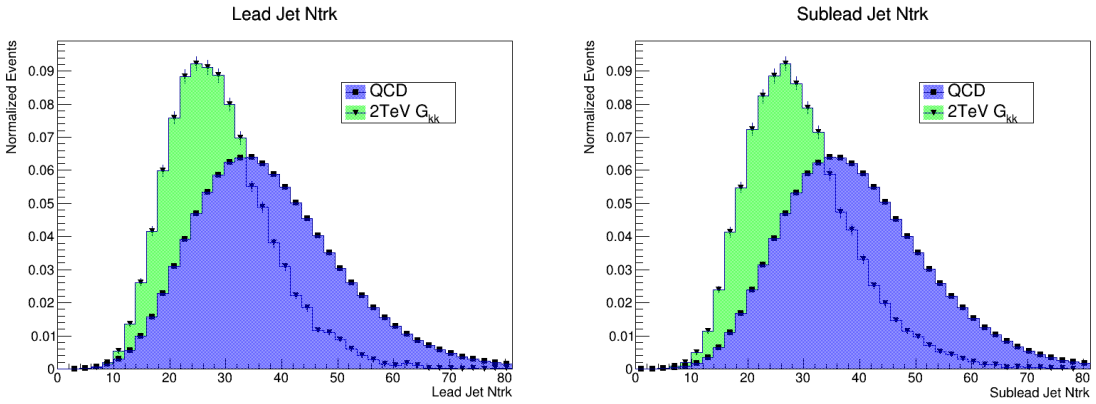
The difference between the mean NTrk value between background and signal MC is demonstrated in Fig.C.1 and C.2 is shown for 2 TeV signal mass poles. The difference is nearly 10 for G_{kk} and 8 for Scalar samples, which varies slightly depending on resonant mass point. The difference in NTrk distribution and NTrk mean is accounted for by mismodelling of the NTrk variable depending on the parton showering generator for each respective sample. This is explained in section C.1.5. A study was performed to find an optimal NTrk cut that will cater to all resonance masses of the two motivating signal models. No consideration was made for the standard model analysis.

C.1.3.1 Optimization Method

Optimization proceeds from the aforementioned MC samples with the small selection applied. Both G_{kk} and Scalar resonant signal samples are considered for the mass points 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1800, 2000, 2250, 2500, 2750, 3000, 3500, and 4000GeV and 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1800, 2000, 2500, and 3000GeV respectively. A sensitivity factor, A, is chosen as the variable to optimize on as to consider both background and signal in a potential sensitivity gain.

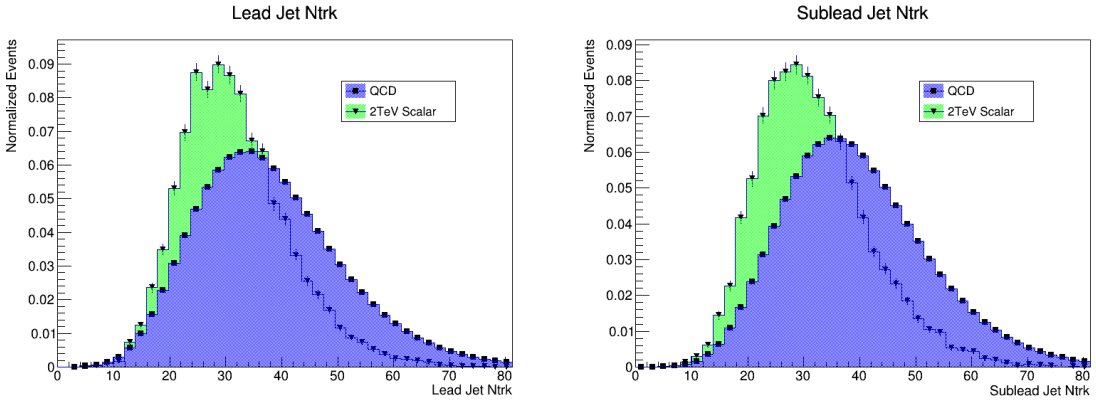
$$A = \frac{\epsilon}{\sqrt{B_{NTrk}}} \tag{C.1}$$

where ϵ (cut efficiency) is the ratio of the signal yields with an NTrk cut divided by a signal yield without an NTrk cut ($\frac{S_{nTrk}}{S_{nonTrk}}$). B_{NTrk} is the background yield with an NTrk cut. To avoid bias from low jet mass particles that dominate the background,



(a) (b)

Figure C.1. NTrk distribution for a 2 TeV G_{kk} compared to QCD MC. Lead jet and sublead jet NTrk distributions shown left and right respectively.



(a) Leading (b) Subleading

Figure C.2. NTrk distribution for a 2 TeV Scalar compared to QCD MC. Lead jet and sublead jet NTrk distributions shown left and right respectively.

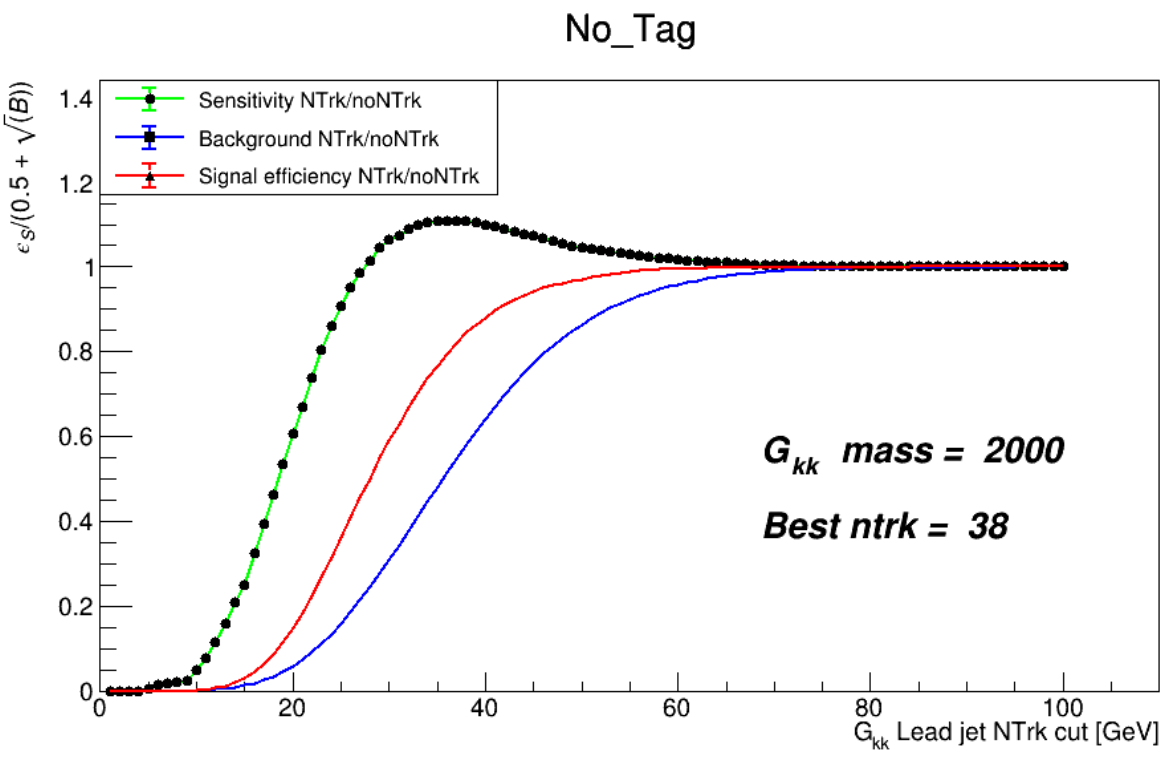
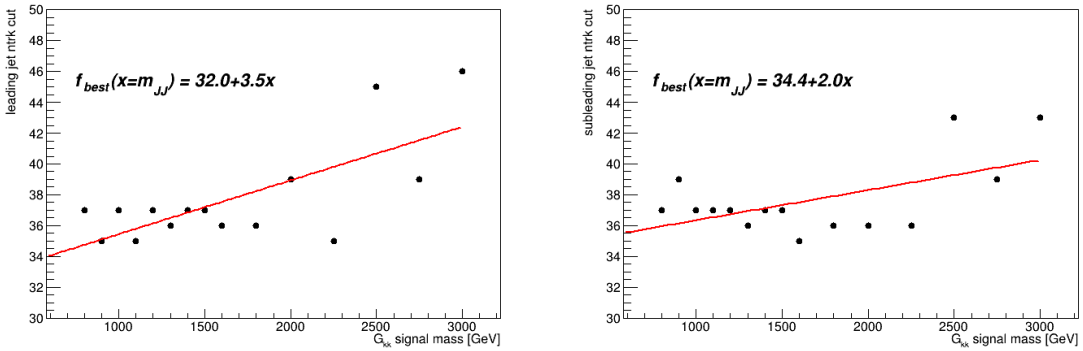


Figure C.3. Plot shows sensitivity factor A (green), background efficiency (blue) and signal efficiency growth for stepping NTrk cuts from 0 to 100. They converge to a unity as the NTrk grows larger, which approaches no significant cut. Sensitivity consistency peaks between NTrk cuts of 28 to 40 for all samples tested. The three plots converge to unity as the NTrk cut grows larger, which approaches no relevant cut. A small, negligible value is added the the denominator to prevent divergence at very small NTrk cut values.

only yields within a mass window of $0.90M_{sig}$ to $1.10M_{sig}$, where M_{sig} is $0.95\times$ the signal pole mass.

The sensitivity value A is then maximized by stepping through several cuts, which remove anything with NTrk larger than a specified cut value. This cut value is iterated from 0 to 100, from which the maximal value of A is obtained for each mass points. The sensitivity gain, signal efficiency, and background efficiency as a function of NTrk cut is shown for a 2000 GeV G_{kk} sample in Fig. C.3. This type of plot has been repeated and studied for each mass point of G_{kk} and Scalar sample for both leading and subleading jets, in which a similar trends are exhibited.

In order to harmonize differences between the two signal models, an identical cut is to be selected and applied to the analysis. To further increase sensitivity gains, both leading and subleading jet NTrk are cut on. Furthermore, leading and subleading jet NTrk cuts are optimized as a pair that maximizes total sensitivity gain for a given signal mass point. For instance, leading and subleading jet NTrk cut values 37 and 37 maximize A for the G_{kk} 1.5 TeV mass point. These optimal



(a) Leading (b) Subleading

Figure C.4. Optimized NTrk cut values for leading (left) and subleading (right) jets. This is optimized on G_{kk} samples, which use PYTHIA8 as a parton showering generator. All equations shown use m_{JJ} in units of TeV.

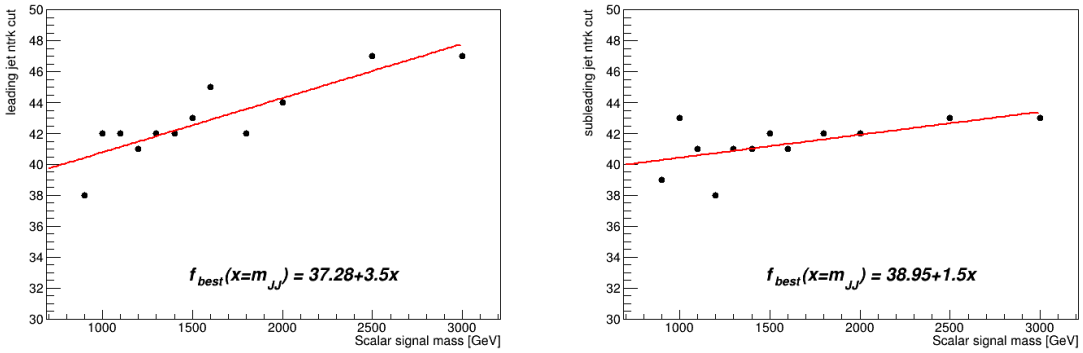
Model	Lead Jet Cut	Sublead Jet Cut
Pythia (linear)	$32 + 3.5x$	$34.4 + 2x$
Herwig	$37.28 + 3.5x$	$38.95 + 1.5x$
Pythia (cubic)	$35.42 + 2.7x - 2.9x^2 + 1.03x^3$	$44.83 - 11x + 3.7x^2 - 0.058x^3$

Table C.1. Top three functional form of optimized cuts. x is m_{JJ} in units of TeV.

values are then individually fit for each jet to a polynomial that is a function of diHiggs mass. Fig. C.4 and Fig. C.5 shows a linear polynomial fit applied to the optimized G_{kk} and Scalar samples respectively. As our G_{kk} and Scalar samples use hadronization and parton showering done with PYTHIA8 with EVTGENand HERWIG7 with EVTGENrespectively, functional fits on these signal samples will be referred to by their according generator name. Further consideration was made to other fits, which include higher order polynomials, constant cuts values, and a mix of the several fits, before finalizing functional form of the cut.

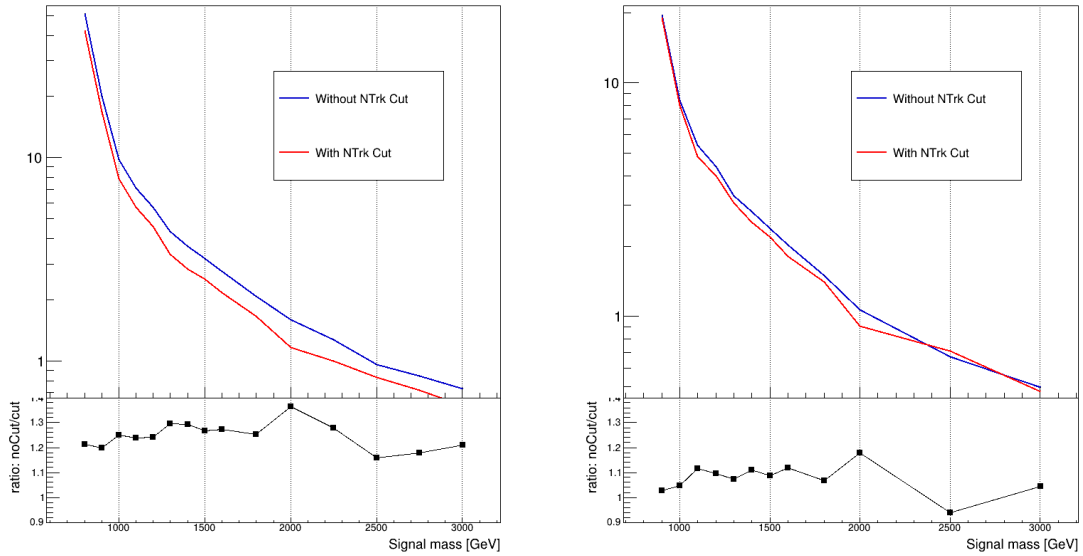
C.1.3.2 Optimization Results

Table C.1 shows the three top performing functional cuts in consideration. The final function choice is picked by observing the expected limit with and without the NTrk cut for the two signal models. Fig.C.6, C.7, and C.8 shows various expected limit (nominal line only) for the different functional cuts compared to the expected limit without the cut. The PYTHIA8 Model (linear) shows the best limit with greatest improvements demonstrated for both signal models, thus it is used as the final NTrk cut.



(a) Leading (b) Subleading

Figure C.5. Optimized NTrk cut values for leading (left) and subleading (right) jets. This is optimized Scalar samples, which use HERWIG7 as a parton showering generator. All equations shown use m_{JJ} in units of TeV.



(a) (b)

Figure C.6. Nominal limit plot comparison for analysis with and without NTrk PYTHIA8 (linear) functional cut. G^*_{KK} and Scalar limits are shown left and right respectively.

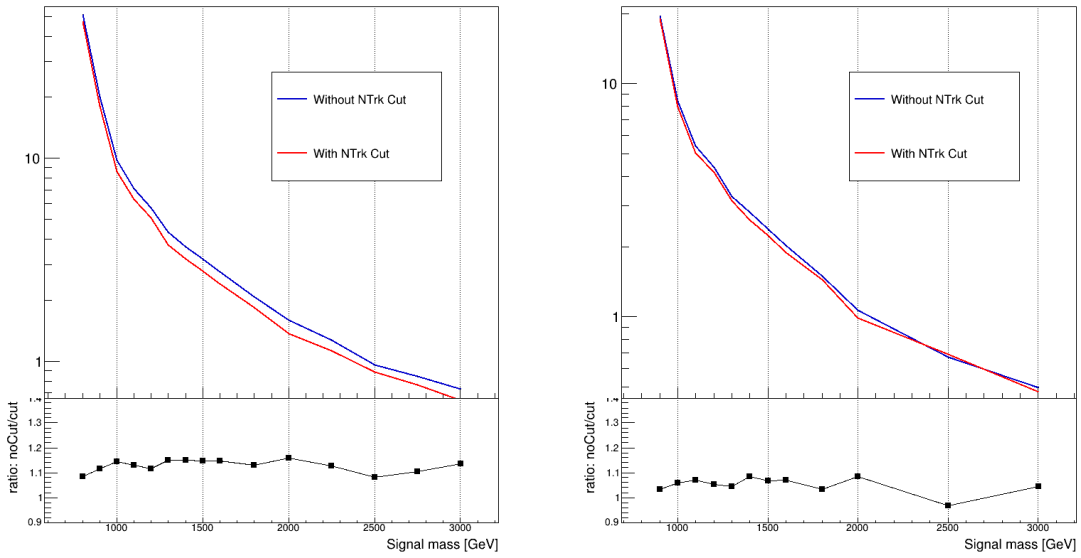
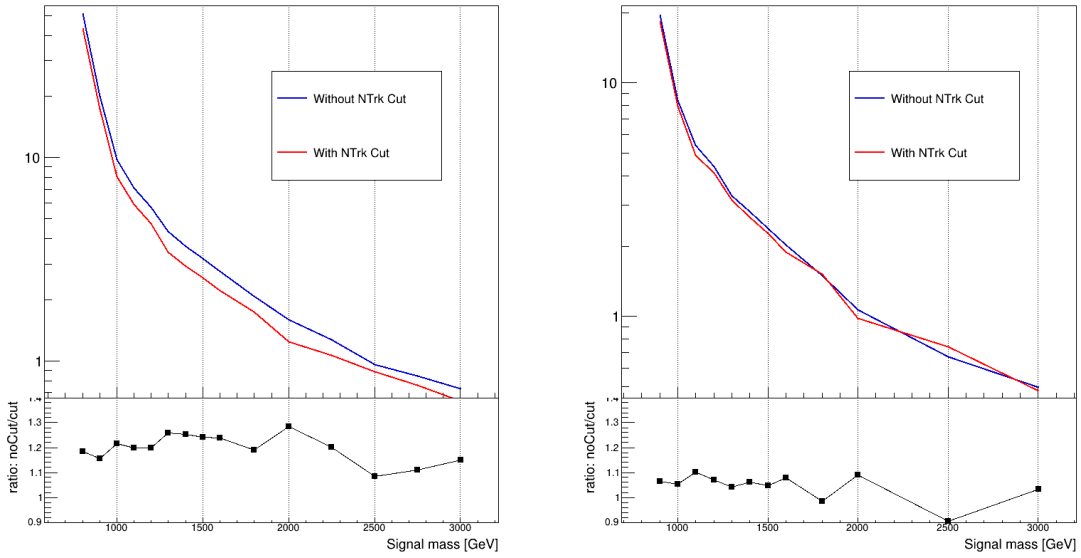
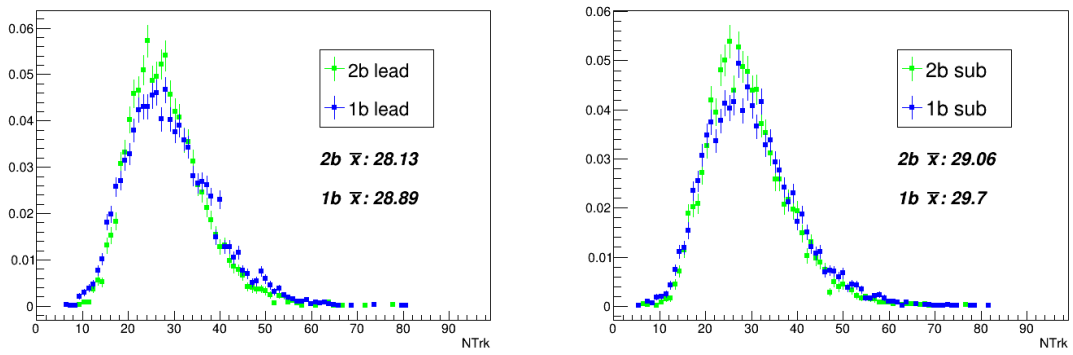


Figure C.7. Nominal limit plot comparison for analysis with and without NTrk HERWIG7 functional cut. G_{kk} and Scalar limits are shown left and right respectively.



(a) (b)

Figure C.8. Nominal limit plot comparison for analysis with and without NTrk PYTHIA8 (cubic) functional cut. G^*_{KK} and Scalar limits are shown left and right respectively.



(a) Leading (b) Subleading

Figure C.9. NTrk distributions for leading and subleading jets on the left and right respectively. Distribution with a 2-tag and 1-tag requirement are overlaid with average NTrk denoted by $\bar{x}$.

C.1.4 Differences in b-tagging

A study of the effect of b-tagging with NTrk selection has been conducted. B-tagging provides a powerful discriminant in our analysis. The study intends to explore the impact of the NTrk selection on the 4tag, 3tag, and 2tag-split regions to find if special consideration should be made to optimize for each region individually.

NTrk distributions for leading and subleading jets were examined. Fig. C.9 shows the NTrk distributions for a 2 TeV G_{kk} sample with one or two tags. There are only slights differences between the NTrk distributions for different tags, with a mean NTrk value difference less than one. This difference is slight enough to not demand a separate optimization for each b-tag region. This study further validates the use of only one harmonized optimization choice for the analysis regardless of b-tag region or signal model.

C.1.5 Parton Showering Generator Differences

G_{kk} and Scalar signal samples are modelled by different parton showering generators (PS). G_{kk} and Scalar samples use Pythia8 and Herwig7 respectively. Studies into generator differences are generally well motivated, but is further prompted by the differences seen in the NTrk distributions of the signal samples between Fig. C.1 and Fig. C.2.

Alternative Scalar samples are simulated using Pythia8 for three mass points relevant to the boosted regime: 1, 2, and 4 TeV. Fig. C.11 shows a comparison of NTrk distribution for the two differently showered Scalar samples. The Herwig7 sample shows an NTrk distribution with higher values than that of Pythia8, which therefore provides less NTrk difference from background and thus is a less powerful discriminant for Herwig7 samples.

A cutflow comparison between scalar samples with HERWIG7 and PYTHIA8 reveals the major differences between the samples. The cuts most profoundly affected by PS differences are b-tagging and NTrk cuts. Fig C.10 show the near final effect of PS differences on b-tagging and NTrk. We account for an approximate PS difference in b-tagging of $< 10\%$ for all tagged regions and Scalar samples. The PS difference for NTrk appears to be anywhere from $10 - 20\%$ depending on the tagged region and Scalar sample, which is considerable in our analysis and cannot be ignored. Two methods are considered in handling PS differences due to NTrk.

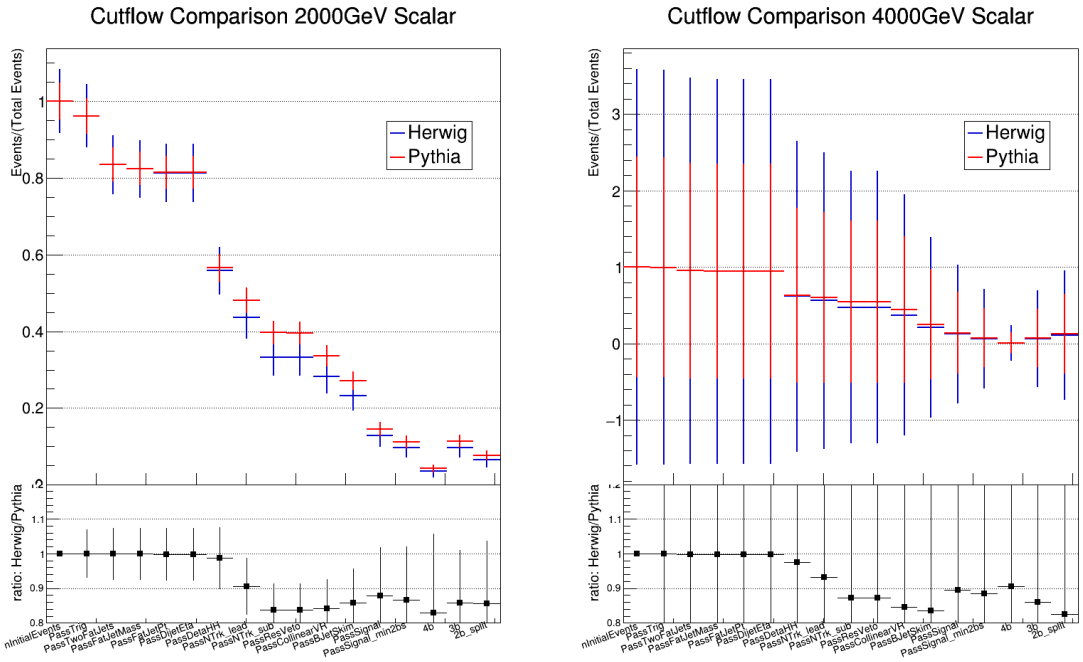
The first considered method in handling PS differences in NTrk is treating it as a source of uncertainty. The Scalar samples most relevant in investigating this are the 2 and 4 TeV Scalar samples – the 1 TeV Scalar sample is too close to the edge of our sensitivity for the boosted analysis so it will not be considered for this uncertainty evaluation. The evaluation is also considered preliminary as two mass points are not enough to assess any shape uncertainty, thus only a normalization factor is used. If we observe the differences in efficiency between PS in Fig C.10, a fair normalization factor of 15% can be assigned to NTrk. This will be added in quadrature with other uncertainties.

The second considered method is based off the better NTrk modeling by PYTHIA8. The PYTHIA8 Scalar samples can be used as the proper simulation of NTrk and scaling of HERWIG7 efficiencies to that of PYTHIA8 may be performed. This scaling would only be done for the relevent NTrk cuts, which would significantly improve the limit and efficiencies for the HERWIG7 samples. After lengthy consideration, this method, although potentially most benefitial to our analysis, is discounted. Scaling efficiencies may be too drastic of an untertaking in the late stageds of this analysis and harmonizing this method with other diHiggs group analyses would prove difficult. This method may be considered in the future or for analyses with similar situations, which can more easily harmonize without conflict.

C.1.6 NTrk Uncertainty

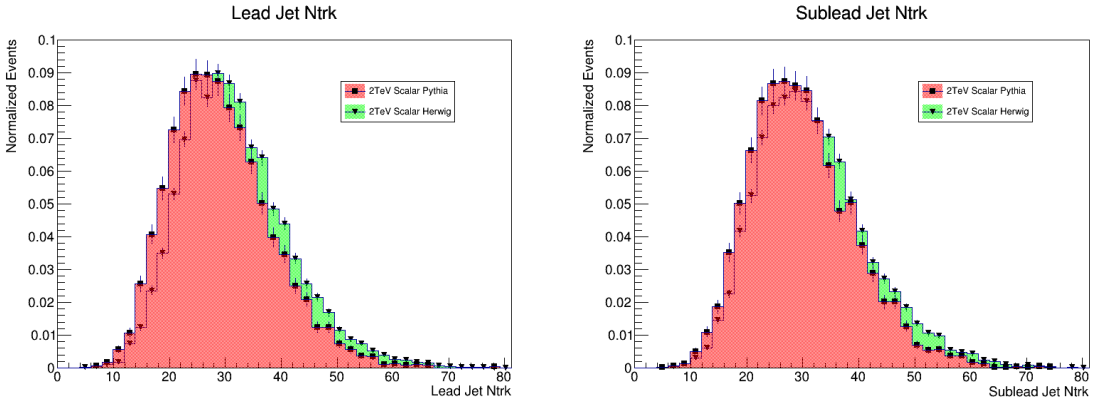
The small cross section of the Higgs boson makes directly evaluating data-to-MC scale factors for NTrk unfeasible. Instead the larger cross section of the W/Z boson is exploited to get the data-to-MC scale factor. The method used is described in depth in the "Systematic uncertainties for Higgs tagging" section of [26]. The topology of the Higgs candidate in the referenced analysis is exactly that of this analysis, thus the described NTrk uncertainty can be safely propagated into this analysis. The result is a conservative 15% normalization uncertainty attributed to data-to-MC differences in NTrk. The expected limit including this 15% uncertainty is calculated and compared to the stat-only limit without NTrk to help quantify sensitivity gains. Fig. C.12 shows a comparison of the nominal lines from these limits. The G_{kk} limit shows great improvement whereas the Scalar limit shows little to no improvement, which is accounted for by HERWIG7 mismodelling NTrk described previously.

If an additional uncertainty is assigned to the PS differences as previously described, we may see a profound effect on the limit. An additional 15% normalization uncertainty is attributed to the PS difference based on the 2 and 4 TeV Scalar sam-



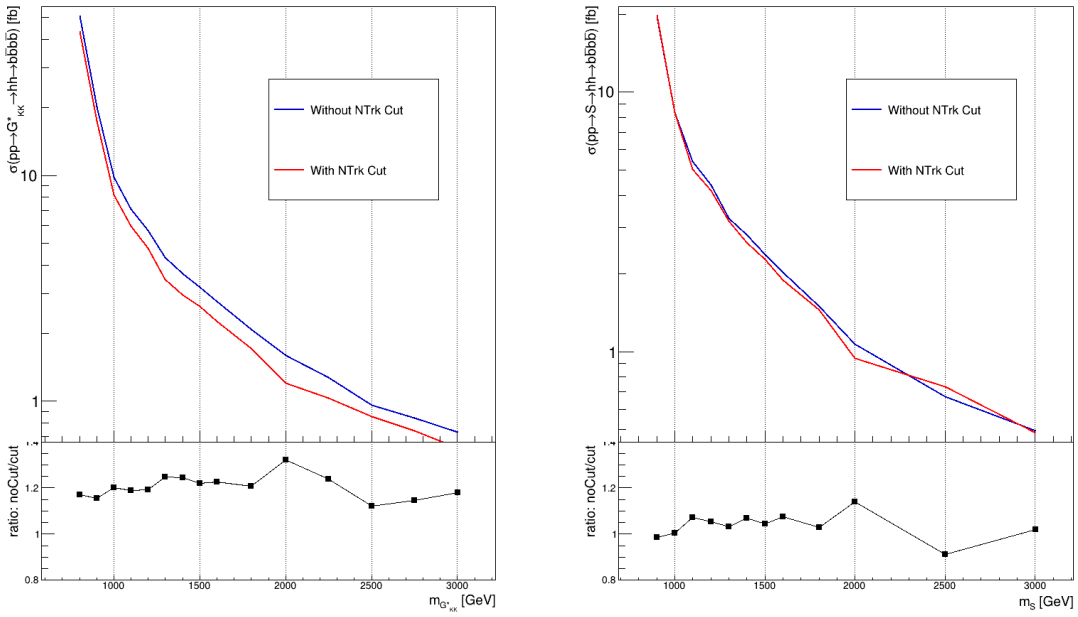
(a) (b)

Figure C.10. HERWIG7 and Pythia cutflow comparison divided by total events in the sample. A major deviation begins around the NTrk selection and is seen in the b-tagging as well. The 2 and 4 TeV Scalar samples are shown left and right respectively. Note that the cutflow is sequential up until the tagging, which includes cuts up until colinear trackjet veto.



(a) (b)

Figure C.11. NTrk distribution comparison for leading and subleading jets between a 2 TeV Scalar signal sample using HERWIG7 and another using PYTHIA8.



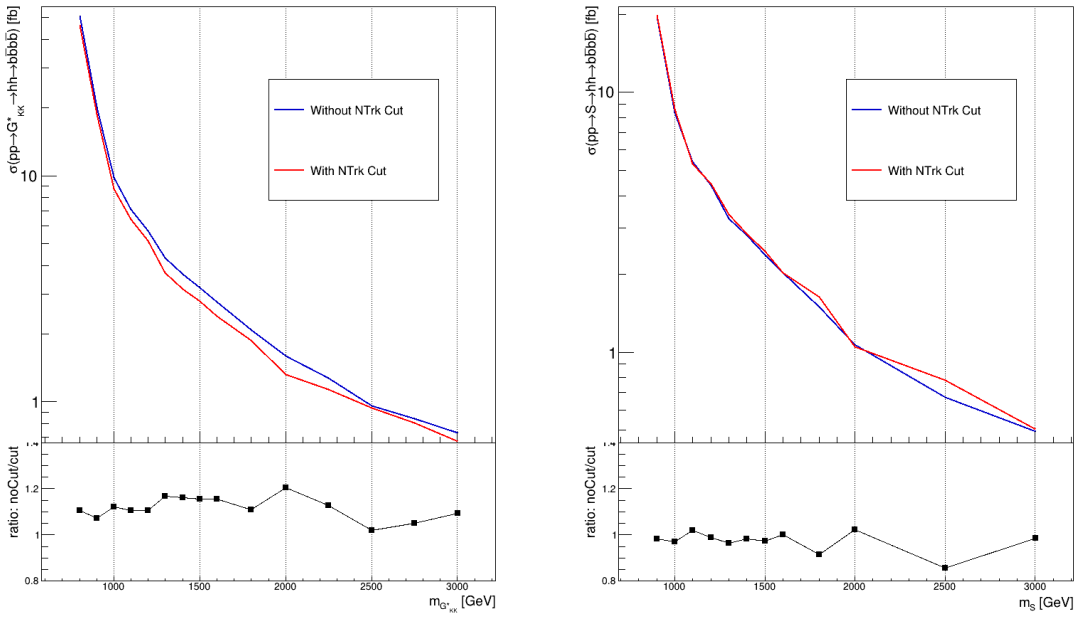
(a) G_{KK} (b) Scalar

Figure C.12. Limit plots with a 15% data-to-MC normalization uncertainty applied. G_{KK} (left) and Scalar (right) limits are shown. The nominal line is plotted alongside the nominal line from the stat-only baseline analysis without NTrk. The G_{KK} shows significant improvement and the Scalar limit shows modest improvement.

ple PS comparisons. Fig. C.13 shows the nominal line comparison with both NTrk and PS difference uncertainties included. The the resulting G_{kk} limit shows a small improvement and the Scalar limit shows no improvement.

C.1.7 Conclusion

Using NTrk as a disciminant in the Boosted analysis shows great potential. A stat-only study of our baseline analysis with NTrk shows vast improvements to the G_{kk} model and modest improvements to the Scalar model. This difference in G_{kk} and Scalar models leads to the conclusion that the Parton showering generator used in the simulated Scalar samples does not model the NTrk variable properly. This mismodelling has proven irreconsilable, both from standpoint of the boosted analysis and large scale harmonization. The only viable option presented is to add a large normalization uncertainty to account for PS differences, which would negate most sensitivity gains NTrk would provide. After much deliberation, because of these PS differences difficulties as well as the added complication of including this cut, it will not be included in the final analysis. Further consideration should be given to NTrk



(a) G_{KK} (b) Scalar

Figure C.13. Limit plots with a 15% data-to-MC and a 15% PS normalization uncertainty applied. G_{KK} (left) and Scalar (right) limits are shown. The nominal line is plotted alongside the one from the stat-only baseline analysis without NTrk. The G_{KK} shows modest improvement and the Scalar limit shows no improvement.

in future analyses, especially if said analyses follow a similar topology as the Boosted analysis and can maintain a PYTHIA8 only analysis as its MC PS.

APPENDIX D

BACKGROUND SHARING

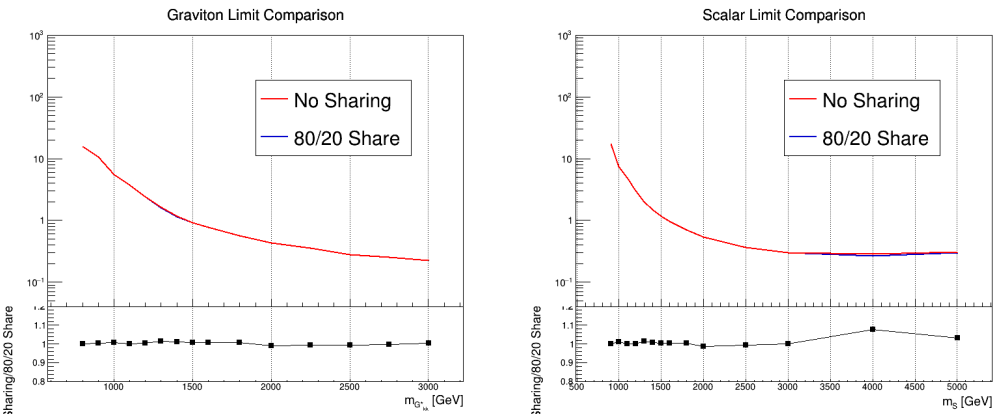


Figure D.1. G_{kk} and Scalar stat-only limit comparison on the left and right respectively. The comparison is done on the analysis with and without sharing. The impact on the analysis is minimal.

D.1 Background Sharing

The 3b region uses the 2b-1 "low-tag" region to model background, which is meant to have similar topology to the 3b region while being orthogonal to signal events. Events with the exact topology of having one large-R Jet with two b-taggedtrack jets and the other large-R Jet with a single track jet with no b-tagging, are scarce. The number of events is insufficient to apply fitting and obtain a background estimation. Thus, we allow for the scenario of an event with one large-R Jet with two b-taggedtrack jets and the other large-R Jet with one or more track jets without b-taggingto pass as a 2b-1.

This revised definition of the low-tag region effectively increases the yield by an order of magnitude. The revised definition also causes overlap with the 4b low-tag region (2b-2), which uses one large-R Jet with two b-taggedtrack jets and one large-R Jet with two track jets with no b-tagging. In order to maintain statistical independence between the 3b and 4b regions, the 2b-2 region is shared between the two regions. Events are selected randomly to obtain a 80/20 share between 3b and 4b respectively. A 50/50 share was also considered.

The random selection is done by performing modular division on the second digit of the event number. The second digit is used to vary the selection from a similar modular division that is done on the first digit during the reweighting process. A comparison between background estimation with no sharing and with sharing is shown in Figure D.1. The nominal line for the stat-only limit is compared to demonstrate any changes that may arise from background sharing. It is clear that the 80/20 share has only a small impact on the limit, thus is chosen as the optimal share selection for our analysis.

The background sharing results in less events in which the background is modelled, especially for the 4b region. The smoothing process needs to be optimized accordingly

to the new dataset. The newly optimized background smoothing is responsible for the slight increase in sensitivity at 4 TeV in Figure D.1.

APPENDIX E

BACKGROUND REWEIGHTING

The reweighting is applied to jet p_T , ΔR_{jj}, jet η to correct for any mismodelling that might take place during background estimation. The below plots compare the HT and LT distributions, which by definition have different number of b-tags. The effect b-tagging has on these distributions is apparent in the left plots, which are not reweighted. The reweighting is applied to the right plots, which helps remove the sculpting tagging has on the distributions.

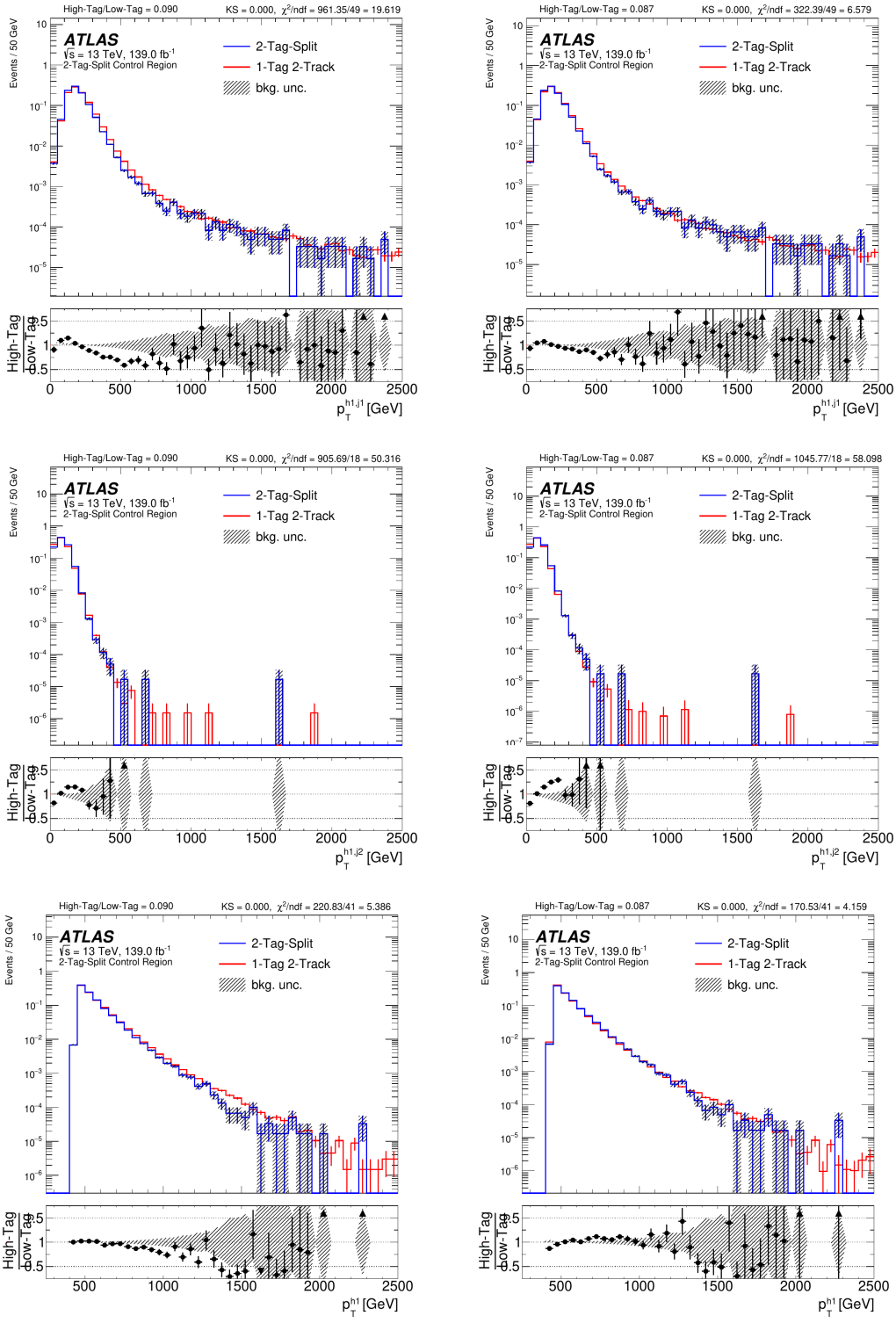


Figure E.1. High-tag and low-tag leading Higgs candidate p_T distributions in the 2 tag-split region before and after reweighting on the left and right respectively.

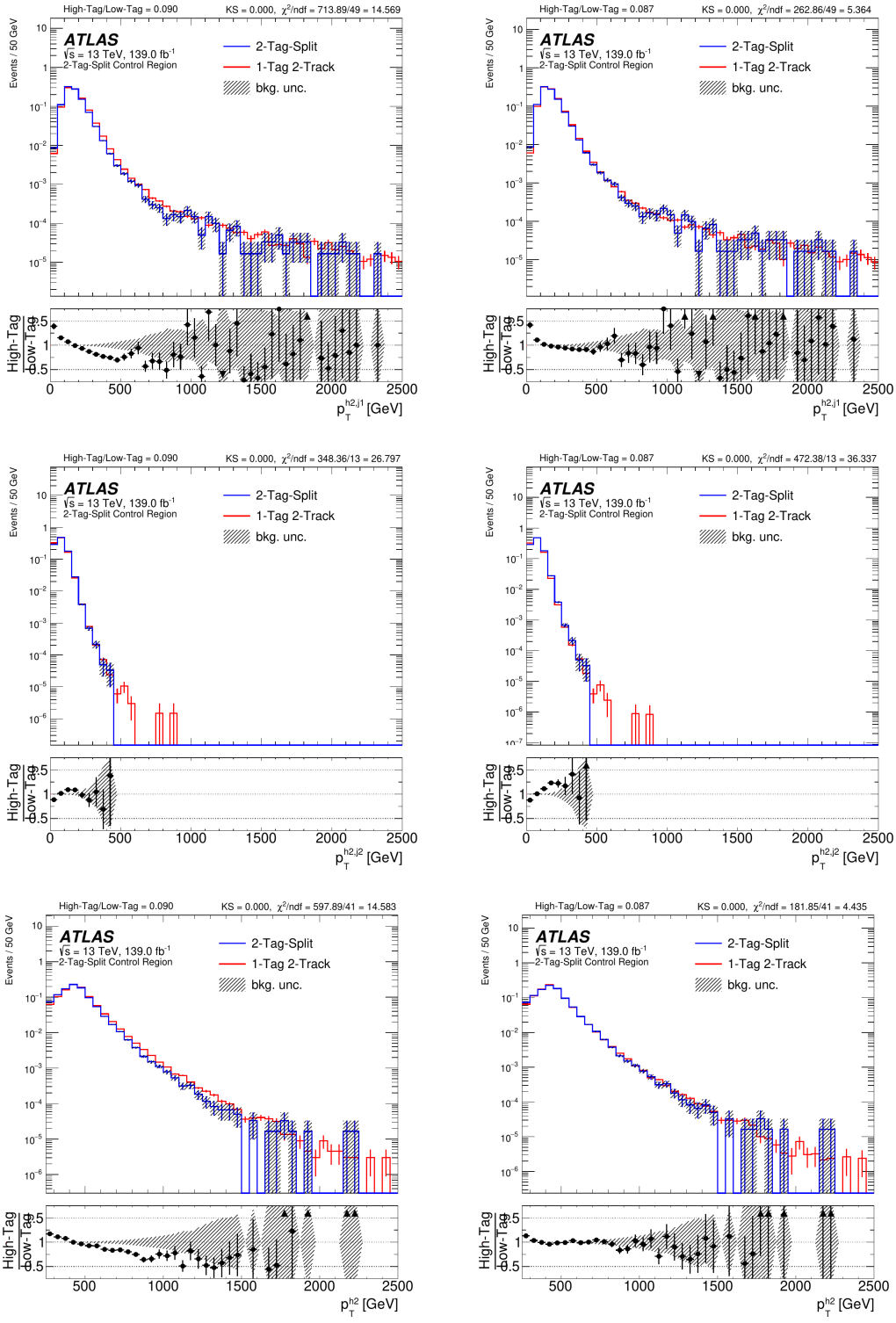


Figure E.2. High-tag and low-tag subleading Higgs candidate p_T distributions in the 2 tag-split region before and after reweighting on the left and right respectively.

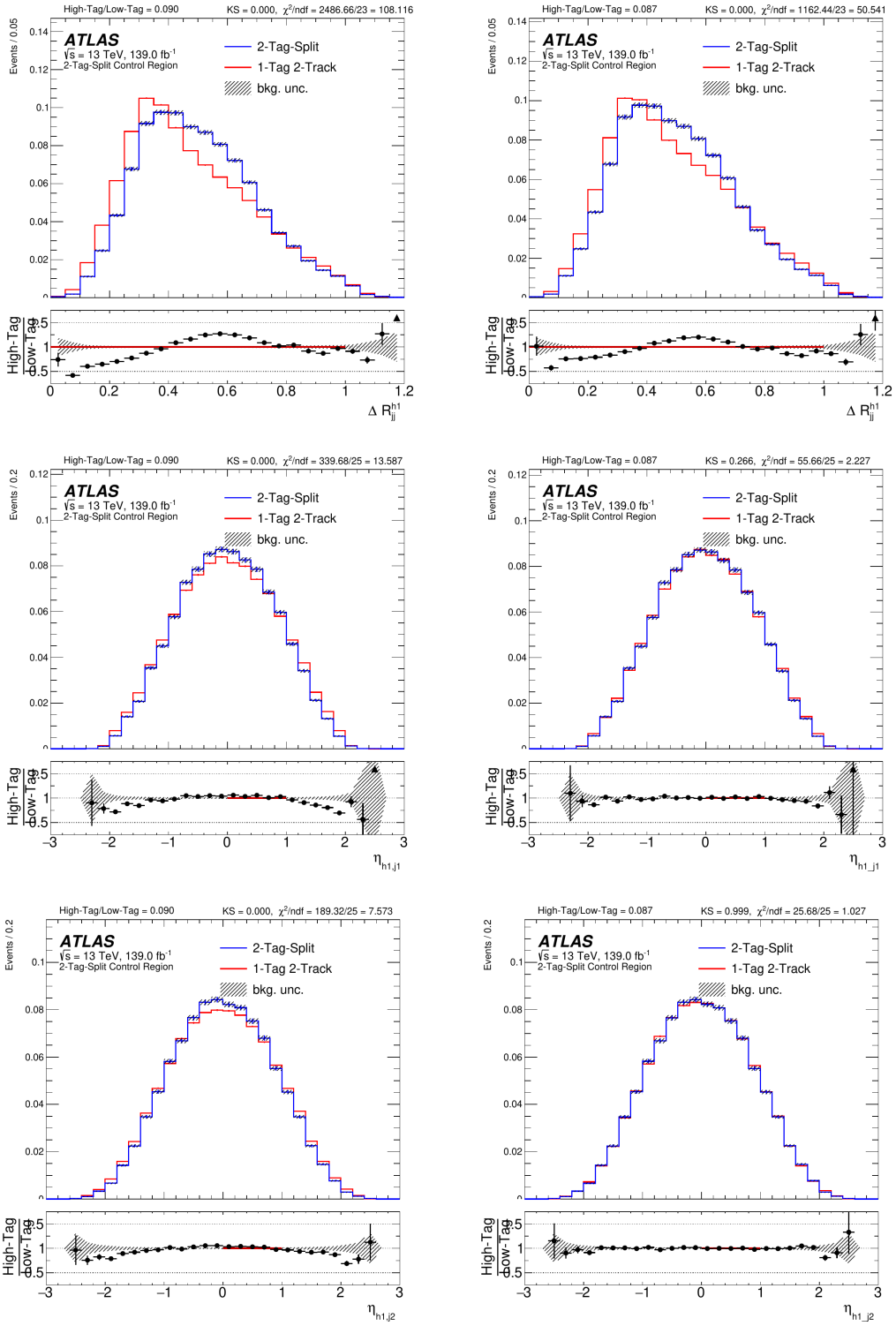


Figure E.3. High-tag and low-tag leading Higgs candidate angular distributions in the 2 tag-split region before and after reweighting on the left and right respectively.

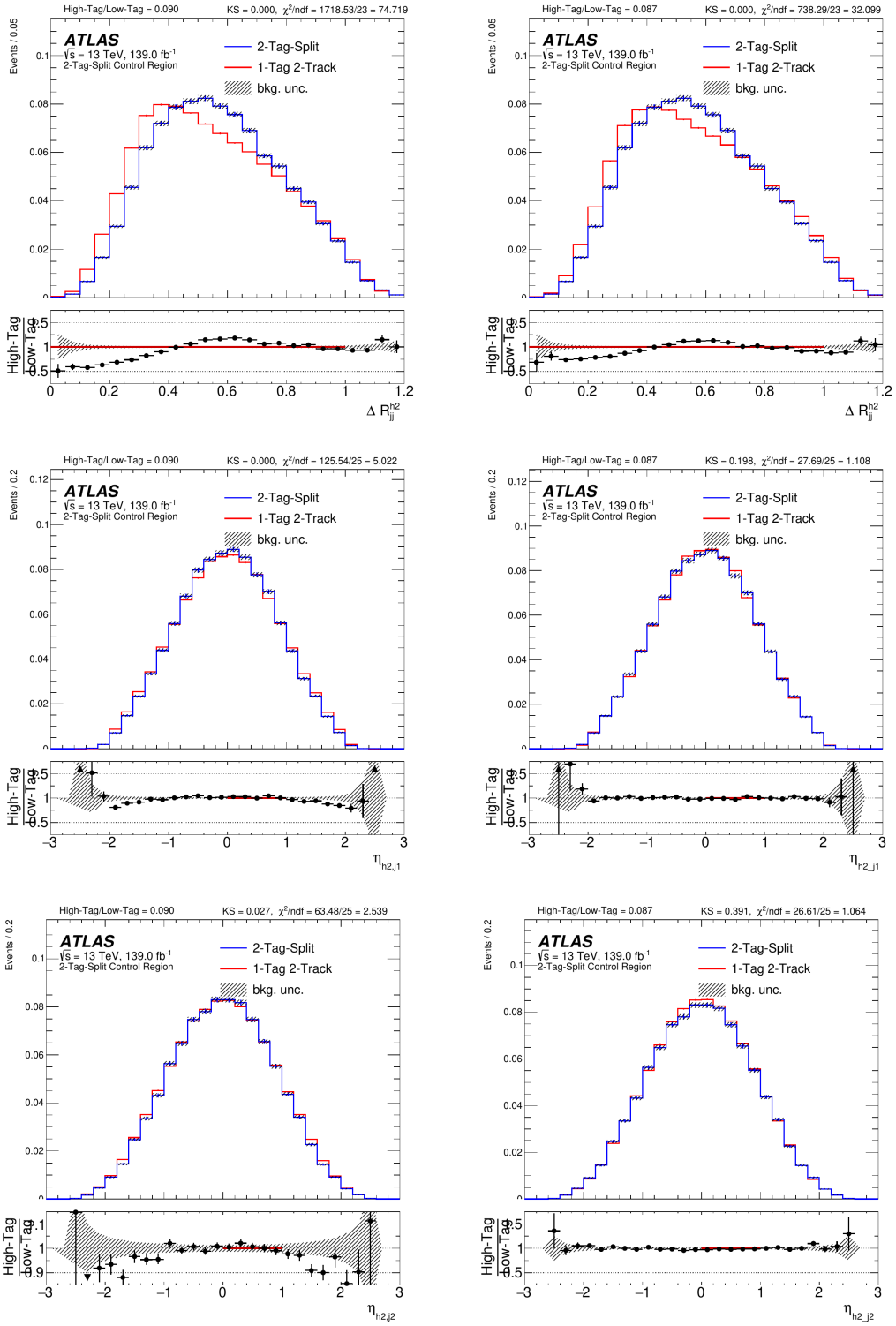


Figure E.4. High-tag and low-tag subleading Higgs candidate angular distributions in the 2 tag-split region before and after reweighting on the left and right respectively.

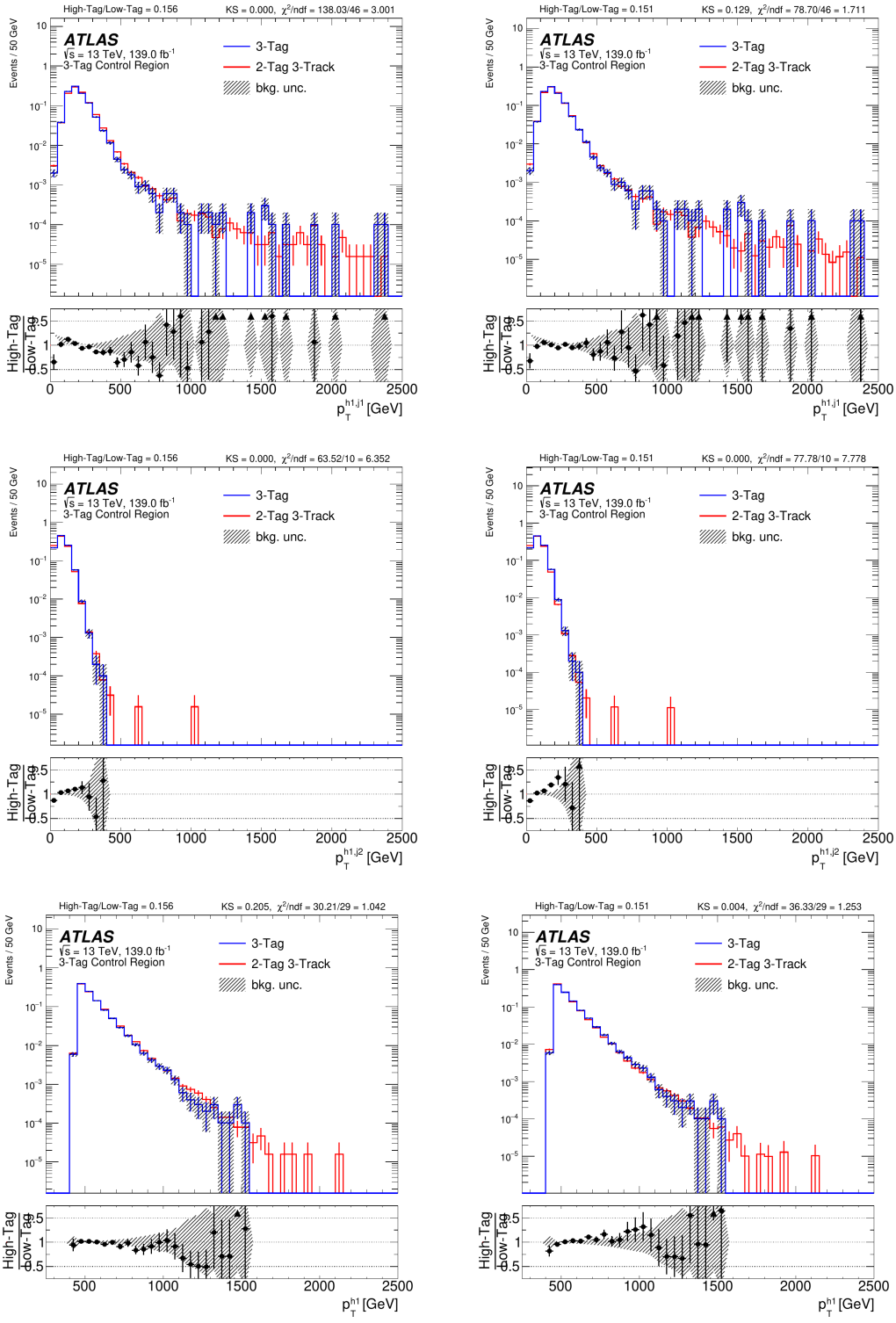


Figure E.5. High-tag and low-tag leading Higgs candidate p_T distributions in the 3 tag region before and after reweighting on the left and right respectively.

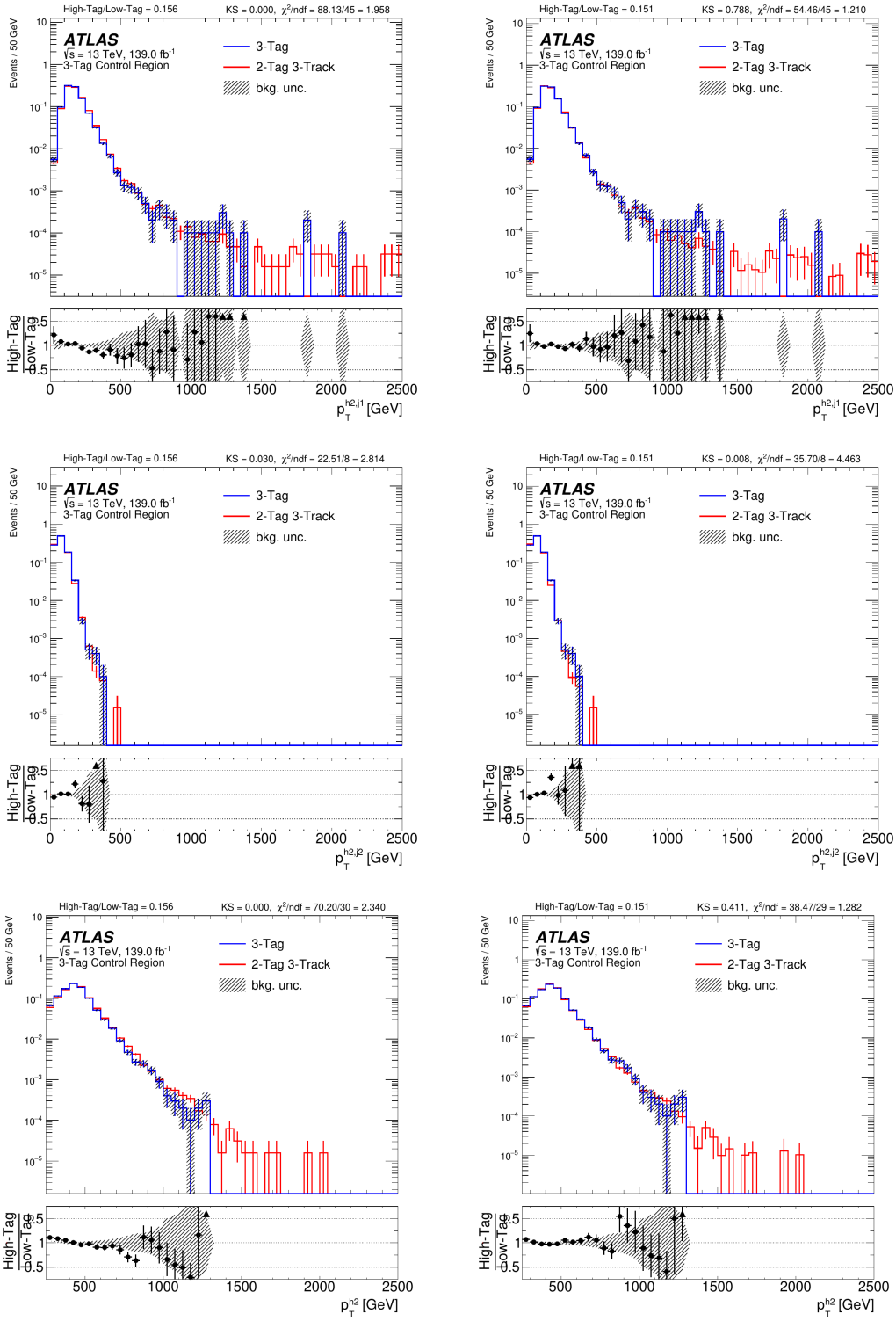


Figure E.6. High-tag and low-tag subleading Higgs candidate p_T distributions in the 3 tag region before and after reweighting on the left and right respectively.

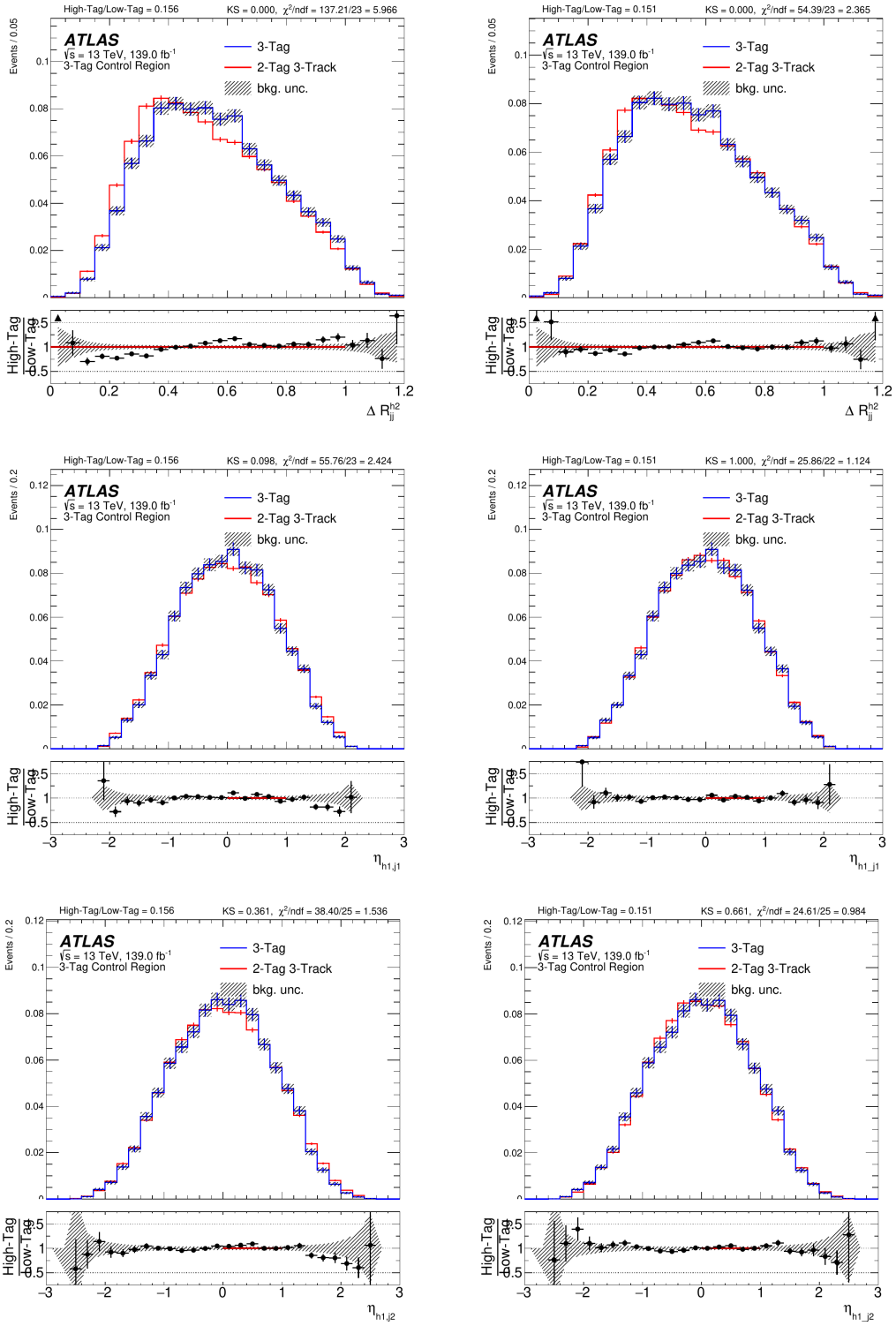


Figure E.7. High-tag and low-tag leading Higgs candidate angular distributions in the 3 tag region before and after reweighting on the left and right respectively.

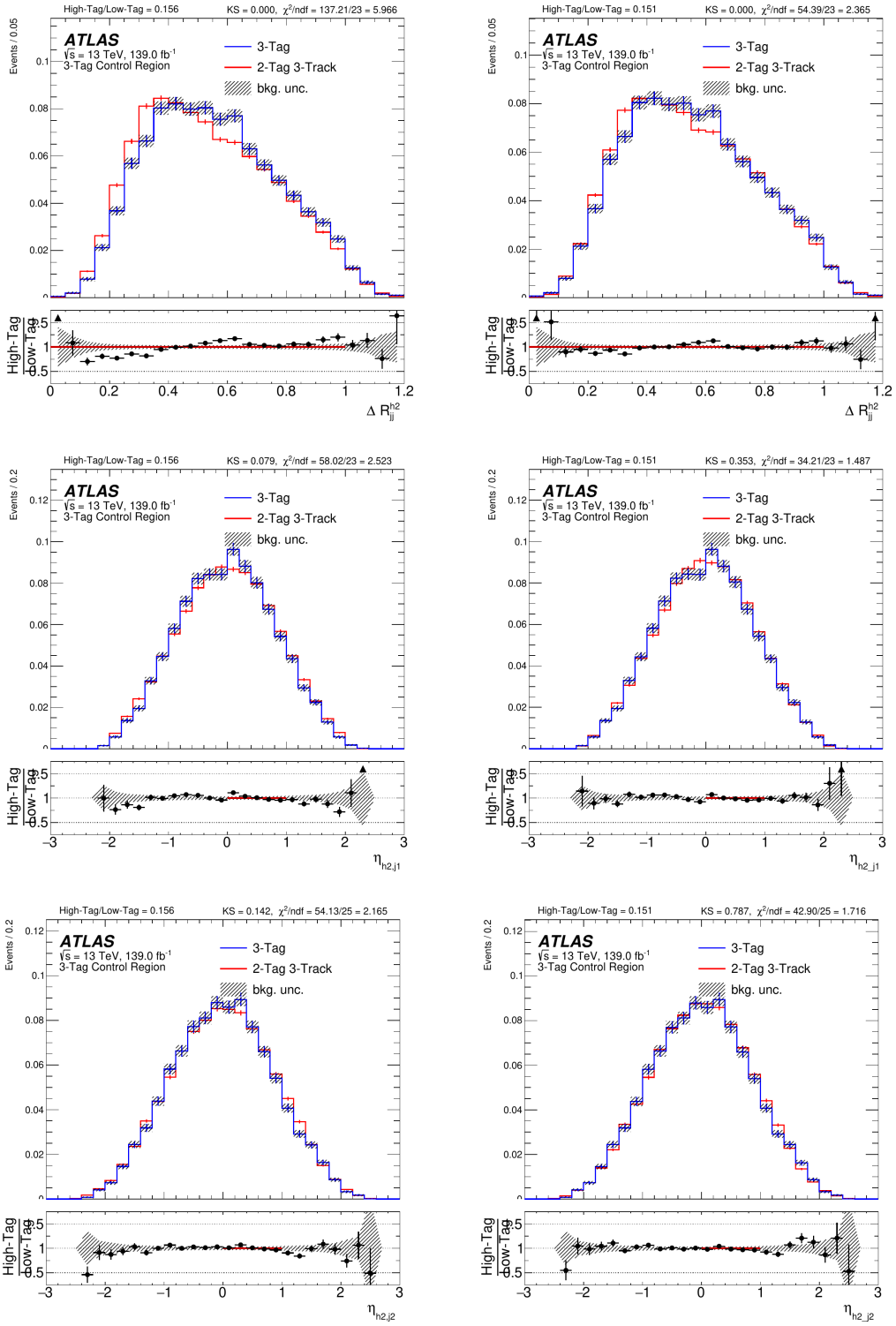


Figure E.8. High-tag and low-tag subleading Higgs candidate angular distributions in the 3 tag region before and after reweighting on the left and right respectively.

BIBLIOGRAPHY

[1] ATLAS Trigger Performance: Status Report. Tech. rep., CERN, Geneva, Jun 1998.

[2] Search for Higgs boson pair production in the two bottom quarks plus two photons final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Tech. rep., CERN, Geneva, Mar 2021.

[3] Aaboud, M., Aad, G., Abbott, B., Abdinov, O., Abeloos, B., Abhayasinghe, D. K., Abidi, S. H., AbouZeid, O. S., Abraham, N. L., and et al. In situ calibration of large-radius jet energy and mass in 13 tev proton–proton collisions with the atlas detector. *The European Physical Journal C 79*, 2 (Feb 2019).

[4] Aad, G., Abbott, B., Abbott, D.C., Abud, A. Abed, Abeling, K., Abhayasinghe, D.K., Abidi, S.H., AbouZeid, O.S., Abraham, N.L., Abramowicz, H., and et al. Atlas data quality operations and performance for 2015–2018 data-taking. *Journal of Instrumentation 15*, 04 (Apr 2020), P04003–P04003.

[5] Aaij, Roel, Anderlini, Lucio, Benson, Sean, Cattaneo, Marco, Charpentier, Philippe, Clemencic, Marco, Falabella, Antonio, Ferrari, Fabio, Fontana, Marianna, Gligorov, Vladimir Vava, and et al. Selection and processing of calibration samples to measure the particle identification performance of the lhcb experiment in run 2. *EPJ Techniques and Instrumentation 6*, 1 (Feb 2019).

[6] Abbott, B., et al. Production and integration of the ATLAS Insertable B-Layer. *JINST 13* (2018), T05008.

[7] ATLAS and CMS Collaborations. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett. 114* (2015), 191803.

[8] ATLAS Collaboration. ATLAS Liquid-Argon Calorimeter: Technical Design Report. ATLAS-TDR-2, 1996.

[9] ATLAS Collaboration. ATLAS Tile Calorimeter: Technical Design Report. ATLAS-TDR-3, 1996.

[10] ATLAS Collaboration. ATLAS Muon Spectrometer: Technical Design Report. ATLAS-TDR-10, 1997.

[11] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST 3* (2008), S08003.

[12] ATLAS Collaboration. ATLAS Insertable B-Layer Technical Design Report. ATLAS-TDR-19, 2010.

[13] ATLAS Collaboration. Performance of large-R jets and substructure reconstruction with the ATLAS detector. ATLAS-CONF-2012-065, 2012.

[14] ATLAS Collaboration. Performance of jet substructure techniques for large-R jets in proton–proton collisions at $\sqrt{s} = 7\,\text{TeV}$ using the ATLAS detector. *JHEP 09* (2013), 076.

[15] ATLAS Collaboration. Basic ATLAS TRT performance studies of Run 1. ATL-INDET-PUB-2014-001, 2014.

[16] ATLAS Collaboration. Operation and performance of the ATLAS semiconductor tracker. *JINST 9* (2014), P08009.

[17] ATLAS Collaboration. Search for Higgs boson pair production in the $b\bar{b}b\bar{b}$ final state from pp collisions at $\sqrt{s} = 8\,\text{TeV}$ with the ATLAS detector. *Eur. Phys. J. C 75* (2015), 412.

[18] ATLAS Collaboration. Search for Higgs Boson Pair Production in the $\gamma\gamma b\bar{b}$ Final State Using pp Collision Data at $\sqrt{s} = 8\,\text{TeV}$ from the ATLAS Detector. *Phys. Rev. Lett. 114* (2015), 081802.

[19] ATLAS Collaboration. Searches for Higgs boson pair production in the $hh \to bb\tau\tau, \gamma\gamma WW^*, \gamma\gamma bb, bbbb$ channels with the ATLAS detector. *Phys. Rev. D 92* (2015), 092004.

[20] ATLAS Collaboration. Jet mass reconstruction with the ATLAS Detector in early Run 2 data. ATLAS-CONF-2016-035, 2016.

[21] ATLAS Collaboration. Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13\,\text{TeV}$. *Eur. Phys. J. C 76* (2016), 292.

[22] ATLAS Collaboration. In-situ measurements of the ATLAS large-radius jet response in 13 TeV pp collisions. ATLAS-CONF-2017-063, 2017.

[23] ATLAS Collaboration. Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run. ATL-PHYS-PUB-2017-013, 2017.

[24] ATLAS Collaboration. Performance of the ATLAS trigger system in 2015. *Eur. Phys. J. C 77* (2017), 317.

[25] ATLAS Collaboration. Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC. *Eur. Phys. J. C 77* (2017), 332.

[26] ATLAS Collaboration. Search for Heavy Resonances Decaying to a W or Z Boson and a Higgs Boson in the $q\bar{q}^{(\prime)}b\bar{b}$ Final State in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector. ATLAS-CONF-2017-018, 2017.

[27] ATLAS Collaboration. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *Eur. Phys. J. C 77* (2017), 490.

[28] ATLAS Collaboration. Variable Radius, Exclusive-k_T, and Center-of-Mass Subjet Reconstruction for Higgs($\to b\bar{b}$) Tagging in ATLAS. ATL-PHYS-PUB-2017-010, 2017.

[29] ATLAS Collaboration. Measurement of the Higgs boson mass in the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels with $\sqrt{s} = 13\,\text{TeV}$ pp collisions using the ATLAS detector. *Phys. Lett. B 784* (2018), 345.

[30] ATLAS Collaboration. Performance of top-quark and W-boson tagging with ATLAS in Run 2 of the LHC. *Eur. Phys. J. C 79* (2019), 375.

[31] ATLAS Collaboration. Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector. *JHEP 01* (2019), 030.

[32] Branco, G.C., Ferreira, P.M., Lavoura, L., Rebelo, M.N., Sher, Marc, and Silva, João P. Theory and phenomenology of two-higgs-doublet models. *Physics Reports 516*, 1-2 (Jul 2012), 1–102.

[33] Cacciari, Matteo, Salam, Gavin P., and Soyez, Gregory. The anti-k_t jet clustering algorithm. *JHEP 04* (2008), 063.

[34] Carl Edward Rasmussen, Christopher K.I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2006.

[35] Collaboration, ATLAS. Reconstruction and identification of boosted di- systems in a search for higgs boson pairs using 13 tev proton−proton collision data in atlas, 2020.

[36] Cowan, Glen, Cranmer, Kyle, Gross, Eilam, and Vitells, Ofer. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C 71* (2011), 1554.

[37] Dawson, S. Introduction to electroweak symmetry breaking, 1999.

[38] et al., LHC Higgs Cross Section Working Group. Handbook of lhc higgs cross sections: 1. inclusive observables, 2011.

[39] et al., LHC Higgs Cross Section Working Group. Handbook of lhc higgs cross sections: 2. differential distributions, 2012.

[40] et al., The LHC Higgs Cross Section Working Group. Handbook of lhc higgs cross sections: 3. higgs properties, 2013.

[41] Evans, Lyndon, and Bryant, Philip. LHC Machine. *JINST 3* (2008), S08001.

[42] Frate, Meghan, Cranmer, Kyle, Kalia, Saarik, Vandenberg-Rodes, Alexander, and Whiteson, Daniel. Modeling smooth backgrounds and generic localized signals with gaussian processes, 2017.

[43] Hugging, F. The atlas pixel detector. *IEEE Transactions on Nuclear Science 53*, 3 (Jun 2006), 1732–1736.

[44] Konstantinidis, Nikos. *Lecture Notes for the 2017 School for Experimental High Energy Physics Students*. RAL Technical Report RAL-TR-2017-009, 2017.

[45] Randall, Lisa, and Sundrum, Raman. Large mass hierarchy from a small extra dimension. *Physical Review Letters 83*, 17 (Oct 1999), 3370–3373.

[46] W.N. Cottingham, D.A. Greenwood. *An Introduction to the Standard Model of Particle Physics*. Cambridge University Press, Cambridge, 2007.