

A SEARCH FOR NON-STANDARD MODEL HIGGS BOSONS
PRODUCED IN ASSOCIATION WITH B -QUARKS AT THE
ATLAS EXPERIMENT

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF PHYSICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Caitlin Malone

June 2015

CERN-THESIS-2015-080
01/05/2015



Abstract

With the discovery of a Standard-Model-like Higgs boson at the LHC in 2012, the particle content of the Standard Model has been fully verified experimentally, and the focus turns even more fully to searches for physics beyond the Standard Model. One interesting class of theories, supersymmetry, predicts (at least) 4 additional Higgs bosons. This is a search for two of those SUSY Higgs particles, A^0 and H^0 , in production states including additional b -quarks and in final states where A^0 and H^0 are decaying to a $b\bar{b}$ final state. Using data-driven estimations to understand and model the large multi- b -quark QCD backgrounds at the LHC, we set limits on the production cross section times $b\bar{b}$ branching ratio of 1.79-0.25 pb in the range $450 \text{ GeV} < m_{A^0} < 800 \text{ GeV}$. No statistically significant excess is observed in the data relative to the background-only expectations. This is the first resonance search in this final state at ATLAS, and the highest-mass search for A^0 and H^0 decaying to $b\bar{b}$ that has yet been performed.

Acknowledgements

I would like to acknowledge the many friends, family members and colleagues who made this project possible. Most broadly, I owe great thanks to my colleagues at the LHC and ATLAS, who put in great effort day after day to make it possible to collect and study the data that this thesis is based upon. The SLAC ATLAS group in particular has been like a family to me; I'd like especially to thank John Backus Mayes, Pascal Nef, Francesco Rubbo, Bryan Fulsom, Pelle Hansson, Michael Kagan, and Till Eifert for being a great pack of postdocs, always challenging and supporting me. My fellow ATLAS students also serve as a daily inspiration, and I'd like to thank Max Swiatlowski, Qi Zeng, Ben Nachmann, Aviv Cukierman, David Miller, Jim Black, Bart Butler, and Dan Silverstein for being awesome people. Josh Cogan especially deserves recognition for putting up with me as an office mate for so many years; I don't have a brother, but if I had to name an honorary brother, it would be Josh.

Thanks to the senior scientists and professors of the group, including Rainer Bartoldus, Ariel Schwartzman, Phillipe Grenier, Tim Nelson, and especially Charlie Young for all the hard work they do in running the group and generally being such friendly, helpful, great physicists.

Pat Burchat and Lance Dixon deserve great credit for their extremely helpful comments on this thesis, and I would like to thank them as well as Sarah Church and Mike Tomz for serving on my defense committee.

This analysis in particular would not be what it is without the extremely dedicated work of a small team, and I owe them a great deal. Many thanks to Emanuel Strauss for his incredibly important work in the early days of this analysis, getting the entire project moving. Giacinto Piacquadio was instrumental in writing the fit, which is the heart of the background estimation step and makes the entire analysis possible. Tim Barklow has been the best teammate I could have asked for in the latter half of this analysis, taking the lead especially in engineering many important background modeling studies as well as having the original idea (and writing the implementation) for the rotation. Finally, many thanks to Su Dong for suggesting this search in the first place, and for always offering thoughtful and supportive guidance as we were putting this analysis together.

Most importantly I'd like to thank my family: my parents for their unconditional love and support, and my sisters Gen and Amy for being the best sajas I could have. Last but not least, I'd like to thank my husband Bobby, for his love and support over the last twelve years.

Contents

Abstract	iv
Acknowledgements	v
1 Theory of Higgs Physics	1
1.1 Introduction	1
1.2 The Standard Model	2
1.2.1 Quarks and Leptons	2
1.2.2 Bosons and Forces	4
1.3 Electroweak Symmetry Breaking and the Higgs Mechanism in the Standard Model	5
1.4 Supersymmetry	7
1.5 Higgs Physics in Supersymmetry	11
1.5.1 H/A Searches in the $b\bar{b}$ Final State	13
1.5.2 Constraints on H/A Existence and Kinematics	20
1.6 From Beautiful Theory to Messy Experiment	21
1.7 Previous Searches	24
2 ATLAS Detector	25
2.1 Introduction	25
2.2 CERN and the Large Hadron Collider	26
2.3 The ATLAS Coordinate System	31

2.4	The ATLAS Detector	32
2.4.1	Inner Detector	33
2.4.2	Calorimeters	38
2.4.3	Muon System	41
2.5	The ATLAS Trigger and Data Processing	42
3	Reconstruction and Performance	46
3.1	Track Reconstruction	47
3.2	Jet Reconstruction	49
3.3	Pileup Calibration and Removal	50
3.4	Primary Vertex Identification	52
3.5	b -Tagging	52
3.5.1	Online b -Tagging	53
3.5.2	Offline b -Tagging	56
4	Monte Carlo Simulation	58
4.1	MC Creation Procedure	58
4.2	Signal Monte Carlo Simulations	59
4.2.1	Generator Comparisons	61
4.3	Background	69
4.3.1	QCD Background	69
5	Trigger and Cuts	82
5.1	Trigger	83
5.2	Offline Cuts	84
6	Background Estimation	91
6.1	Background Estimation Strategy	91
6.2	Background Estimation Method Based on Parameterized Histogram Fitting	93
6.3	Modeling of Background Shape	95

6.3.1	Selection of Parameterized Histogram	95
6.3.2	Background Shape Variations	100
6.3.3	Non-QCD Background	101
7	Signal Kinematics	103
7.1	Signal Jet p_T Distributions	104
7.2	Signal Jet Combinatorics	106
7.3	Mass Resolution	109
7.3.1	Validation of Qualitative Shape Differences	111
7.3.2	Jet Reconstruction Modifications	112
7.3.3	Topology-Based Energy Recovery Algorithm	113
7.3.4	p_T -Based Energy Recovery Algorithm	113
7.3.5	n_{jets} Dependence	114
7.4	Eigenvector Rotation	116
8	Fit Strategy and Results	133
8.0.1	Fit Model and Categories	134
8.0.2	Signal Yield Tests in Signal+Background Scenario	135
8.1	Search Results	137
8.2	Limit Extraction	139
9	Systematic Uncertainties	141
9.1	Background Shape Variation Systematics	142
9.2	Trigger p_T Efficiency	146
9.3	Jet Energy Uncertainties	149
9.4	b -Tagging Scale Factors	151
9.4.1	Offline b -Tagging	152
9.4.2	b -Tagging in the Trigger	153
10	Interpretation	155
10.1	m_h^{max} Scenario	156

10.2 m_h^{mod} Scenarios	156
Bibliography	159

List of Tables

1.1	A summary of the fermions. In the “SM interactions” column, S stands for the strong nuclear force, W is the weak nuclear force, and EM is the electromagnetic force. As is customary in particle physics, the mass is measured in units of energy divided by the speed of light squared (since $E = mc^2$), although in practice the c^2 is often dropped and masses expressed simply in units of energy, typically electron-volts (eV) or multiples thereof[1].	3
1.2	The bosons of the Standard Model: their masses and interactions. The interactions between the Higgs field and the fermions are not usually characterized as a force, per se, but rather it is the interaction between the Higgs field and the W boson, Z boson, and fermions that gives mass to the particles.	4
2.1	The rates and latency of the three layers of the ATLAS trigger. Each layer applies increasingly stringent requirements while working with progressively more data, in order to diagnose whether an event is “interesting” enough to pass one of the trigger chains [2].	45

4.1	The cross sections and generation times for MadGraph signal MC simulations as a function of number of additional partons allowed, for 0-3 partons. There is approx. 2.5% difference between the 2-parton and 3-parton cross sections, which comes at a cost of hours of CPU time per batch of 10,000 events. The times quoted here are for generating 10,000 events per sample, before hadronization. The generation process imposes a p_T cut of 5 GeV on all jets, no further p_T cut on b -jets, a minimum dR of 0.4 between jets (no minimum between b -jets), and a maximum $ \eta $ of 5.0.	60
4.2	The signal MC samples and their parameters. These widths were generated with a $\tan \beta$ value of 30. The dataset ID is a reference for internal ATLAS bookkeeping.	60
4.3	Hard subprocesses simulated in the inclusive QCD MC event sample, along with their cross sections. Here $j = u, \bar{u}, d, \bar{d}, s, \bar{s}, c, \bar{c}, g$. The QCD slices are defined by the p_T cut placed on the leading jet in the event, which is documented in the “ p_T range” column. The dataset ID is noted as a reference for internal ATLAS bookkeeping.	70
4.4	In addition to examining the shapes of the kinematic distributions in data vs. in bb QCD MC simulation events, the relative normalization of the different categories is computed and compiled. The data and bb QCD MC events are passed through the analysis framework, applying the trigger and cuts, and then a scale factor of 2.07 is applied to the bb QCD MC simulations to bring the overall number of events to the same as the number of events in data. Then the bb QCD MC (data) events are categorized by the tag status and number of jets in the event.	72
4.5	Hard subprocesses simulated in the bb QCD MC event sample, along with their cross sections. Here $j = u, \bar{u}, d, \bar{d}, s, \bar{s}, c, \bar{c}, g$	81

5.1	The acceptance of a cut on variable X given that the events have already passed (sig-like) or failed (bkgd-like) a cut on variable Y.	84
7.1	The width of the mass distribution in the left and right shoulders of the peak, as explained in the text. The left shoulder is dominated by radiation off the Higgs and/or the Higgs daughter jets in addition to combinatorial mis-reconstructions, leading to a larger RMS than the right shoulder, which is dominated by combinatorics only.	110
7.2	The major axis eigenvector elements, and their squares. The transformation into the eigenbasis makes use of the raw elements, but the squares of the elements sum to 1.0 and can provide some physical intuition for the relative contributions of $p_{T,1}$, $p_{T,2}$ and m_{bb}	118
7.3	The search windows in m'_{bb} , based on the signal mass point under investigation. The background gets pushed to lower values of m'_{bb} by the rotation for higher signal mass values, which allows for wider mass windows for higher signal mass points.	120
8.1	The mean and RMS of the pull histograms from running 50 toy MC simulations per point of two mass points (450 and 700 GeV) and four signal cross sections (0, 0.1, 1.0, and 10.0 pb). To within the statistical error on the toy studies, which is about 15%, the toy MC studies do not point toward any major biases in the fitting process.	136
8.2	The signal cross sections found by the fit to maximize the overall likelihood across all tag and n_{jets} categories, and the statistical error on the cross section.	137
8.3	The bbb category linear variation parameters returned by the fit, where a larger parameter indicates a further deviation from the same shape in the bbb and $bbanti$ categories. In general, the variation is more significant in the 5+ jet bin than in the 3-jet or 4-jet bins.	138

8.4	The <i>bbloose</i> category linear variation parameters returned by the fit, where a larger parameter indicates a further deviation from the same shape in the <i>bbloose</i> and <i>bbanti</i> categories.	138
8.5	The expected and observed production cross sections times branching ratios to be excluded at 95% CL, along with the 1-sigma and 2-sigma statistical uncertainties	140

List of Figures

1.1	The radiative corrections to the Higgs mass can be seen in these Feynman diagrams, with a fermionic correction to the mass shown on the left and the correction from a scalar particle on the right. One problem with the bare Standard Model is that if it is valid up to the Planck scale, 10^{17} GeV, corrections such as these need to cancel to an incredible degree of precision, one part in 10^{30} , to allow the Higgs to have a mass of 126 GeV. This suggests to us that new physics at sub-Planck-scale energies might be well-motivated.[3]	7
1.2	The coupling constants for the three fundamental forces of the Standard Model, as a function of energy. On the left, the coupling constants almost unify (but don't quite meet) in the SM-only framework, whereas on the right there is unification when SUSY effects are allowed to enter the picture at around 1 TeV[4].	10
1.3	Feynman diagrams for leading-order production of h in association with b -quarks.	13
1.4	The cross section for b -quark associated production for A , H and h in proton-proton collisions at $\sqrt{s}=8$ TeV as a function of m_A for several different values of $\tan\beta$ (denoted "tb" in the legend). Figures made using FeynHiggs [5, 6, 7, 8, 9]. There are no additional cuts on the associated b -quark p_T beyond the default cuts in FeynHiggs.	15

1.5	The branching ratio to $b\bar{b}$ of A , H and h as a function of m_A for several different values of $\tan\beta$ (tb). The trend of higher branching ratios for higher values of $\tan\beta$ comes from the Higgs-to- b -quark decay vertex carrying an enhancement factor that scales with $\tan\beta$. Figures made using FeynHiggs [5, 6, 7, 8, 9].	16
1.6	The intrinsic width of A , H , and h as a function of m_A for $\tan\beta=20$ in the m_h^{max} scenario. In practice, the intrinsic width of the particle is dwarfed by the experimental resolution, which can be up to hundreds of GeV. Figures made using FeynHiggs [5, 6, 7, 8, 9].	17
1.7	The intrinsic width of A and H as a function of $\tan\beta$ in the m_h^{max} scenario when $m_A=500$ GeV. Figure made using FeynHiggs [5, 6, 7, 8, 9].	18
1.8	The favored (green) and excluded (red/blue) regions in $m_A/\tan\beta$ in the m_h^{mod} scenario, where the green region shows the parameter space that is consistent with $m_h=126$ GeV. After the discovery of h at 126 GeV, the m_h^{mod} scenario is favored because it allows for m_h to have this relatively low value (m_h^{max} , on the other hand, pushes m_h more toward 140 GeV). The left and right plots show the parameter space for m_h^{mod+} and m_h^{mod-} , respectively, the details of which can be found in [10]. Clearly in this scenario, many values of m_A are open, potentially with large cross-sections.	19
1.9	The mass evolution of the 5 Higgs bosons of the MSSM, with the mass of the Higgs-like indicated by dashed line (126 GeV). For a large range of h masses, the masses of H , A , and H^\pm evolve very closely together, so searches for H^\pm also can provide some indirect constraints on the likely masses of H and A [11].	22

2.1	Cross sections for various processes as a function of center-of-mass energy of the original collision. The “skip” at 4 TeV occurs because below that value, the cross sections are for proton-antiproton collisions like those made at the Tevatron, while higher energy proton-proton collisions are being produced at the LHC.	28
2.2	The luminosity collected in 2010, 2011 and 2012 at ATLAS.	29
2.3	The major elements in CERN’s accelerator complex, which turns ordinary hydrogen atoms into the highest-energy man-made collisions ever made. The four interaction points, labeled for their respective experiments (ALICE, ATLAS, LHCb and CMS) can be seen around the perimeter of the LHC.	30
2.4	A cutaway picture showing the main components of the ATLAS inner detector. The ATLAS coordinate system is sketched on the right-hand side of the figure.	34
2.5	A picture showing the main layers of the ATLAS pixel detector, including the disk layers on the end. Starting in 2015, a 4th layer, the insertable b layer (IBL), will be inserted inside the current layer 0.	35
2.6	The material budget of the inner detector, in units of radiation lengths, as a function of η , showing how the material in the central part of the detector (especially in the pixel subsystem) is kept at a minimum to mitigate multiple scattering.	37
2.7	A depiction of a section of the LAr calorimeter. The accordion structure of the towers is visible, as well as the three sampling layers.	39
2.8	A depiction of the scintillator-and-steel hadronic calorimeter, also known as the tile calorimeter, and a close-up view of one of its cells.	42
2.9	A diagram of the ATLAS detector as a whole, with major subsystems labeled. A couple of human figures are shown standing on and near the detector, for scale.	43

3.1	The efficiency of the track reconstruction, projected in p_T and η , as computed in Monte Carlo simulations.	48
3.2	The track resolution in mm in the x and z directions, comparing minimum bias simulation to data taken with a random trigger. . . .	48
3.3	A graphical representation of the position and energy of calorimeter deposits, and their clustering and reconstruction into jets by the anti- k_t jet clustering algorithm. A feature of anti- k_t which is immediately apparent is the circular jet shapes that it tends to make [12].	50
3.4	Distribution of the mean number of interactions per crossing, also called μ , seen in 2012 data-taking. Larger μ values are characterized by more challenging reconstruction.	51
3.5	A data event taken in 2012, showing the occupancy typical of high pileup. This particular event contains a $Z \rightarrow \mu\mu$ physics signature, mixed in with 25 primary vertices.	53
3.6	An event from early ATLAS data-taking, with a number of tracks (teal, in the zoomed-in view) leading back to the primary vertex, and a secondary vertex reconstructed with the associated tracks highlighted in purple [13].	54
3.7	The online signed impact parameter significance, as used in the trigger b -tagging algorithms.	55
3.8	The jet weight distribution for jets undergoing online b -tagging at L2 (left) and EF (right), with a comparison between MC predictions of various truth flavors (colored histograms) and data (points). As the b -tagging weight increases, the relative enrichment of b -jets grows, although the overall efficiency for b -jets goes down as tighter online weights are applied. [14]	55
3.9	The receiver-operator curve (ROC curve) for the MV1 algorithm, showing the tradeoff between light-flavor rejection (vertical axis) and b -jet efficiency (horizontal axis) [15].	57

4.1	p_T distributions for the leading two jets for Pythia (red), Sherpa (green) and MadGraph (black) for a physical Higgs mass of 380 GeV. Pythia and MadGraph show good agreement, while Sherpa jets are systematically harder.	62
4.2	m_{bb} distributions for the leading two jets for Pythia (red), Sherpa (green) and MadGraph (black) for a physical Higgs mass of 380 GeV. Pythia and MadGraph show good agreement, while the Sherpa m_{bb} distribution is higher and more sharply peaked.	63
4.3	p_T distributions for the two τ leptons (i.e. the Higgs p_T) for Sherpa (red) and Madgraph (black) for zero (upper left), one (upper right), two (lower left), and three (lower right) additional jets per event. . .	65
4.4	η distributions for the 2 τ leptons (i.e. the Higgs η) for Sherpa (red) and Madgraph (black) for zero (upper left), one (upper right), two (lower left), and three (lower right) additional jets per event. The Higgs is more forward (larger $ \eta $) in Sherpa than in MadGraph. . . .	66
4.5	$m_{\tau\tau}$ distributions for the two τ leptons for Sherpa (red) and Madgraph (black) for a 380 GeV Higgs particle. Although the agreement is not perfect in the tails of the distribution, the agreement is much better than in the m_{bb} distributions, suggesting that the challenges of correctly identifying the Higgs daughters in Sherpa is a major factor.	67
4.6	The $m_{\tau\tau}$ ($m_A=380$ GeV) distribution (left) and η of the τ leptons (right) for $bH \rightarrow b\tau\tau$ events in Sherpa, with and without gluon fusion diagrams being included. The ratio plots in the lower part of the figures show that the inclusion of the gluon fusion diagrams does not change the kinematics of the Higgs and its daughters.	68

4.7	The p_T distributions for the leading jet, comparing bb QCD MC events to data. The MC has been normalized to the same total number of events as the data (over the entire sample, not on a category-by-category basis) and the MC/data ratio is plotted in the lower subplots. The errors on the MC are statistical only.	73
4.8	The η distributions for the leading jet, comparing bb QCD MC events to data. The MC has been normalized to the same total number of events as the data (over the entire sample, not on a category-by-category basis) and the MC/data ratio is plotted in the lower subplots. The errors on the MC are statistical only.	74
4.9	The p_T distributions for the second-leading jet, comparing bb QCD MC events to data. The MC has been normalized to the same total number of events as the data (over the entire sample, not on a category-by-category basis) and the MC/data ratio is plotted in the lower subplots. The errors on the MC are statistical only.	75
4.10	The η distributions for the second-leading jet, comparing bb QCD MC events to data. The MC has been normalized to the same total number of events as the data (over the entire sample, not on a category-by-category basis) and the MC/data ratio is plotted in the lower subplots. The errors on the MC are statistical only.	76
4.11	The p_T distributions for the third-leading jet, comparing bb QCD MC events to data. The MC has been normalized to the same total number of events as the data (over the entire sample, not on a category-by-category basis) and the MC/data ratio is plotted in the lower subplots. The errors on the MC are statistical only.	77

4.12	The η distributions for the third-leading jet, comparing bb QCD MC events to data. The MC has been normalized to the same total number of events as the data (over the entire sample, not on a category-by-category basis) and the MC/data ratio is plotted in the lower subplots. The errors on the MC are statistical only.	78
4.13	The m_{bb} distributions in the 3 jet category, where the third jet in the event has been truth-matched as bottom versus charm (left) or light (right). The distributions have been normalized to the same integral, and the lower subplots show the ratio of charm (light) to bottom. Errors on the MC are statistical only. There is no evidence from these distributions that the m_{bb} of the leading 2 jets in a QCD event is affected by the truth flavor of the third jet.	79
4.14	The m_{bb} distributions in the 4 jet category, where the third jet in the event has been truth-matched as bottom versus charm (left) or light (right). The distributions have been normalized to the same integral, and the lower subplots show the ratio of charm (light) to bottom. Errors on the MC are statistical only. There is no evidence from these distributions that the m_{bb} of the leading 2 jets in a QCD event is affected by the truth flavor of the third jet.	79
4.15	The m_{bb} distributions in the 5 or more jet category, where the third jet in the event has been truth-matched as bottom versus charm (left) or light (right). The distributions have been normalized to the same integral, and the lower subplots show the ratio of charm (light) to bottom. Errors on the MC are statistical only. There is no evidence from these distributions that the m_{bb} of the leading 2 jets in a QCD event is affected by the truth flavor of the third jet.	80

5.1	The b -tagging in Level 2 and the Event Filter (in the trigger) and the MV1 b -tagging algorithm (offline) are exploited in the cut flow by requiring that at least 2 jets in the event be “triple tagged”, or b -tagged in both stages of the triggers as well as offline. The figure on the left shows the L2 and EF b -tag distributions, on the right are the EF and MV1 distributions. The cut points are noted with vertical and horizontal lines. Most of the events (more than 80%) fall in regions where the b -tagging cuts agree on whether a jet is a b -jet or not, but a significant subsample of jets are tagged by one b -tagging level/algorithm but not the other(s).	85
5.2	The cut efficiencies in signal, cut by cut. Each line is a different signal mass point, and the overall trend is for higher efficiencies for the higher-mass particles than for the lower-mass particles.	88
5.3	The cut efficiencies in background, cut by cut. The background component that has three real b -jets from QCD has the greatest survival rate through the cut flow, followed by events that have two b -jets and one c -jet.	89
5.4	The final efficiency, after all cuts, of the signal as a function of the physical mass of the generated Higgs particle. The signal efficiency grows as a function of m_A , although at the high mass values, the efficiency levels off at around 6%.	90
6.1	The distributions of the final discriminating variable, m'_{bb} , in background MC for the $bbanti$ and bbb tag categories. The three different plots compare the 3-jet, 4-jet, and 5-or-more jet categories. No signs of major shape differences are seen.	101

6.2	The m_{bb} distributions for all-hadronic $t\bar{t}$ MC after the trigger and all offline cuts are applied (linear Y axis on the left, logarithmic scale on the right). In addition to the overall cross-section, we also want to probe any shape differences that arise when the tag status changes on the third jet in the event. No significant shape differences are seen.	102
7.1	The p_T of the two jets from the Higgs (classified as “first” and “second” based on which one has more p_T in the event), and all other b -tagged jet(s) in the event. The peaks at about 155 GeV and 55 GeV is a result of the trigger turn-on points and associated cut(s). Each plot corresponds to a different Higgs mass point, which are indicated in the titles immediately above the plots.	105
7.2	The percentage probability of the Higgs decaying to a given pair of b -tagged jets in signal MC, as a function of the generated Higgs mass. The assumption being checked is that the majority of the time, the Higgs decays to the leading two b -tagged jets (blue line), so that reconstructing with these jets will yield the correct combinatorics. That assumption is correct 60-70% of the time for $m_A > 350$ GeV, with the remaining 30-40% of events being spread out over many other combinations of jets.	107
7.3	After making the requirement that the leading two jets in the event are b -tagged, the signal MC can be split into events that pass this requirement and those events that fail. For each sample, we ask how often the Higgs decays to the leading two b -jets in the event, or in other words, how likely it is to get the combinatorics of the Higgs reconstruction correct. The combinatorics for the events passing the requirement are correct about 80% of the time for high m_H , while for events failing the requirement the combinatorics success rate is only about 10%.	108

7.4	A schematic diagram showing how the shoulders are defined, for the purposes of parameterizing the width of the signal distribution. The left shoulder especially is affected when FSR is present, as the Higgs reconstructs to a lower mass than its “true” mass.	109
7.5	The normalized m_{bb} distributions for several signal mass points. The mass resolution is qualitatively worse for higher-mass Higgs bosons.	112
7.6	Comparisons of the m_{bb} distributions out-of-the-box, and after attempting the p_T -based and topology-based FSR recovery algorithms, for several signal mass points.	114
7.7	The m_{bb} distributions at a few representative mass points in signal MC simulations, as a function of the number of jets in the event. The 5+ jets bin contains all events with 5 jets or more.	115
7.8	Comparisons of the signal and background m_{bb} distributions before and after the rotation. The left two plots show 450 GeV and 800 GeV Higgs mass points before the rotation. The signal is extremely small compared to the background, whereas the same distributions (including the same signal normalization) after the rotation are on the right, for the same two mass points. The rotation pushes the background events to lower values of m'_{bb} than the signal, yielding better signal/background separation, especially for higher mass Higgs candidates.	121
7.9	The signal and background distributions for each signal mass point after the rotation, binned in n_{jets} category. The effect of the rotation on the background can be clearly seen in these plots, pushing the background further to the left (lower m'_{bb} values) for higher mass signal points.	122
7.10	The signal distributions for each signal mass point before the rotation, binned in n_{jets} category. Upstream cuts effectively cut off the distributions below 300 GeV.	123

7.11	Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the <i>bbb</i> category, for events with 3 jets, for different H/A masses after the rotation has been applied.	124
7.12	Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the <i>bbb</i> category, for events with 4 jets, for different H/A masses after the rotation has been applied.	125
7.13	Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the <i>bbb</i> category, for events with 5 jets, for different H/A masses after the rotation has been applied.	126
7.14	Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the <i>bbloose</i> category, for events with 3 jets, for different H/A masses after the rotation has been applied.	127
7.15	Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the <i>bbloose</i> category, for events with 4 jets, for different H/A masses after the rotation has been applied.	128
7.16	Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the <i>bbloose</i> category, for events with 5 jets, for different H/A masses after the rotation has been applied.	129
7.17	Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the <i>bbanti</i> category, for events with 3 jets, for different H/A masses after the rotation has been applied.	130
7.18	Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the <i>bbanti</i> category, for events with 4 jets, for different H/A masses after the rotation has been applied.	131
7.19	Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the <i>bbanti</i> category, for events with 5 jets, for different H/A masses after the rotation has been applied.	132
8.1	The expected 95% excluded $(H + A) \sigma \times \text{BR}$ as a function of m_A , for $450 < m_A < 800$ GeV.	140

9.1	The ratio of the m'_{bb} distributions in QCD MC simulations of <i>bbanti</i> and <i>bbb</i> events, fitted with a linear function, for the 3-jet, 4-jet and 5-jet categories for $m_A=450, 500,$ and 550 GeV. The slope and intercept of the best fit line are printed on the plots, along with their error and the χ^2 and number of degrees of freedom of the fit.	143
9.2	The ratio of the m'_{bb} distributions in QCD MC simulations of <i>bbanti</i> and <i>bbb</i> events, fitted with a linear function, for the 3-jet, 4-jet and 5-jet categories for $m_A=600-800$ GeV. The slope and intercept of the best fit line are printed on the plots, along with their error and the χ^2 and number of degrees of freedom of the fit.	144
9.3	The <i>bbanti</i> m'_{bb} distribution in data, for the nominal case (i.e. no adjustments made) and with a systematic variation applied based on the shape differences between <i>bbanti</i> and <i>bbb</i> , as quantified in Figures 9.1 and 9.2.	145
9.4	The p_T turn-on curves for the trigger for signal MC and data. Although this search uses a multi-object trigger, in which several conditions must simultaneously be true for a trigger acceptance, tight offline cuts can be used to isolate the effect of a single jet's p_T on the efficiency. The signal curve is fit with a logistic function, which is evaluated at a given cut point to extract the trigger efficiency at that point, and to adjust for any residual effects of other trigger objects that affect the efficiency.	150
9.5	The online <i>b</i> -tagging scale factors as a function of jet p_T for jets categorized by how tightly they are tagged offline. For jets with a p_T higher than the maximum calibrated p_T point, we apply the scale factor for the highest p_T bin that has a calibration.	154

10.1 The observed and expected exclusion limits for the $bH/A \rightarrow b\bar{b}$ search, translated into exclusions in $m_A/\tan\beta$ under the m_h^{max} and $m_h^{mod,\pm}$ scenarios. The behavior of the observed exclusion at $m_A=650$ GeV is an artifact of the weaker limits that we set on that mass point, relative to the trend in its neighbors.	157
--	-----

Chapter 1

Higgs Boson Physics, in the Standard Model and in Supersymmetry

1.1 Introduction

The goal of particle physics is to understand the fundamental particles of the universe and their interactions. It is a field that is simultaneously impressively advanced but with tantalizingly unresolved aspects; in general, progress in the field is a team effort between theorists who propose new physics possibilities and the experimentalists who build large accelerators and detectors and analyze the resulting data for hints of new physics. This thesis details an experimental search for two new particles, the Supersymmetric Higgs bosons H and A , at the ATLAS detector at the Large Hadron Collider.

In order to understand the relevance of this search, and how it is performed, we must first understand the Standard Model of particle physics, including the Higgs

mechanism, and its extensions in Supersymmetry. The Standard Model (SM) is the outcome of decades of experimental and theoretical work in particle physics; it describes all the known particles and their interactions. It is one of the most thoroughly tested theories in all of science, and it has yet to give a prediction that is not experimentally borne out—an impressive feat. At the same time, there are known blind spots in the Standard Model, since it does not include gravity, explain dark matter or dark energy, or account for the origin of the baryon asymmetry of the universe¹. The shortcomings of the Standard Model motivate searches for Beyond-Standard-Model (BSM) physics, including Supersymmetry (SUSY).

1.2 The Standard Model

The Standard Model of particle physics was painstakingly constructed over the 20th century and stands as one of the most thoroughly verified theories in science. The Standard Model (SM) is a quantum field theory that incorporates two different types of matter particles, the quarks and the leptons, as well as three fundamental forces and their corresponding particles. However, as we will see, it has several notable shortcomings that attract considerable attention from both theorists and experimentalists.

1.2.1 Quarks and Leptons

The quarks and the leptons are perhaps the most familiar subatomic particles, as they are the particles that make up matter. For example, a hydrogen atom is composed of a proton (three quarks) and an electron (a lepton). There are six quarks total, three

¹in other words, why the universe contains more matter than antimatter—the question is not why the universe *isn't* made of antimatter, but why matter and antimatter do not exist in equal proportion (if they did, there would not be much of either, as they annihilate when brought into contact)

Table 1.1: A summary of the fermions. In the “SM interactions” column, S stands for the strong nuclear force, W is the weak nuclear force, and EM is the electromagnetic force. As is customary in particle physics, the mass is measured in units of energy divided by the speed of light squared (since $E = mc^2$), although in practice the c^2 is often dropped and masses expressed simply in units of energy, typically electron-volts (eV) or multiples thereof[1].

Generation	Flavor	Electric Charge	Mass (MeV/ c^2)	SM Interactions
1	up quark (u)	+2/3	2.3	S, W, EM
	down quark (d)	-1/3	4.8	S, W, EM
	electron (e)	-1	0.511	W, EM
	electron neutrino (ν_e)	0	$<2.2 \times 10^{-6}$	W
2	charm quark (c)	+2/3	1290	S, W, EM
	strange quark (s)	-1/3	95	S, W, EM
	muon (μ)	-1	105.7	W, EM
	muon neutrino (ν_μ)	0	<0.170	W
3	top quark (t)	+2/3	173,340	S, W, EM
	bottom quark (b)	-1/3	4180	S, W, EM
	tau (τ)	-1	1776	W, EM
	tau neutrino (ν_τ)	0	<15.5	W

“up-type” with an electric charge of +2/3 and three “down-type” with charge of -1/3. There are also three leptons (the electron, muon and tau) that are electrically charged, and three neutrinos (also technically leptons) that are electrically neutral and nearly massless (the electron, muon and tau neutrinos). We can classify the quarks and leptons according to “generation”, where each generation is composed of one up-type quark, one down-type quark, one lepton, and one neutrino. The quarks and leptons are summarized in Table 1.1.

All of the quarks and leptons are fermions, meaning they have half-integer spin.

Table 1.2: The bosons of the Standard Model: their masses and interactions. The interactions between the Higgs field and the fermions are not usually characterized as a force, per se, but rather it is the interaction between the Higgs field and the W boson, Z boson, and fermions that gives mass to the particles.

Particle	Associated Force	Mass
gluon	strong	massless
photon	electromagnetic	massless
W^\pm	weak	80.4 GeV
Z	weak	91.2 GeV
Higgs boson	Higgs field	126 GeV

1.2.2 Bosons and Forces

The forces between fermions are carried by bosons, which are integer spin particles. There are three forces described in the Standard Model: electromagnetic, weak, and strong. The electromagnetic force is carried by the photon and describes, for example, electric forces between particles. Photons are massless and as a result, the electromagnetic field can extend infinitely far. The weak force is carried by W^+ , W^- and Z^0 bosons. These particles are massive, which means that they are limited in how far they can travel and thus the weak force is confined to distance scales approximately the size of an atomic nucleus. The weak force is involved when one type of fermion changes into another type of fermion; for example, when a neutron decays or a nucleus fissions. The strong force is carried by gluons, which are massless but because of confinement, the strong force is restricted to the nuclear scale. The strong force is responsible for holding quarks together into protons, neutrons and other hadrons. Last, there is the gravitational force, which we will neglect as it is many orders of magnitude weaker than the other forces under discussion.

The electromagnetic and weak forces, as it turns out, can be unified into a single “electroweak” force, as discovered in the middle part of the 20th century [16]. The vector bosons acquire mass, which is known as electroweak symmetry breaking, via the Higgs mechanism [17] [18] [19]. The Higgs mechanism, and the particle which

conveys the Higgs field (the Higgs boson), are explained in more detail in further sections. Further unification of forces, between the electroweak and strong forces, remains an unfinished project in physics but a topic of much research.

1.3 Electroweak Symmetry Breaking and the Higgs Mechanism in the Standard Model

The Standard Model is defined by its Lagrangian, which is a mathematical formula that encodes all the Standard Model particles and their interactions. The SM Lagrangian was built piece by piece over many decades, starting with classical field theory and later being generalized to account for relativity, electromagnetism, the strong and weak forces, and the unification of the weak and EM forces (among many other things).

Different terms in the SM Lagrangian account for different types of particles. The fermions, which have spin $1/2$, are governed by the Dirac Lagrangian:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi^2 \quad (1.1)$$

(Similarly, the Klein-Gordon Lagrangian is used for spin-0 particles, and the Proca Lagrangian for spin-1.) When the Dirac Lagrangian is operated upon by the Euler-Lagrange equations, the equation that results is a quantum field theory equation describing a particle of mass m and spin $\frac{1}{2}$.

One problem with simply using the Dirac Lagrangian as-is arises because the Dirac Lagrangian is not invariant under local phase transformations. In other words, if the field ψ is multiplied by an exponential term with a space-dependent (“local”)

²we use the physics notation custom of $\hbar = c = 1$ in this section

phase, $\psi \rightarrow e^{iq\theta(x)}\psi$, then substituting the new ψ into the Euler-Lagrange equations will result in an extra term because of the derivative of $\theta(x)$. To make it locally invariant, we replace the ordinary derivative with the covariant derivative:

$$\mathcal{D}_\mu \equiv \partial_\mu - iqA_\mu \quad (1.2)$$

and A_μ ³ transforms according to

$$A_\mu \rightarrow A_\mu + \partial_\mu\theta \quad (1.3)$$

Similarly, we can study the Lagrangian for a charged scalar field coupled to electromagnetism. Using $F^{\mu\nu} \equiv \partial_\nu A_\mu - \partial_\mu A_\nu$, the Lagrangian⁴ to analyze is

$$\mathcal{L} = [(\partial_\mu - iqA_\mu)\phi^\dagger][(\partial^\mu + iqA^\mu)\phi] - V(\phi^\dagger\phi) - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \quad (1.4)$$

The functional form of the potential $V(\phi)$ is now worth exploring. Consider the “Mexican Hat” potential of the form $V(\phi^\dagger\phi) = \mu^2(\phi^\dagger\phi) + \lambda(\phi^\dagger\phi)^2$, with $\mu^2 < 0$ and $\lambda > 0$. In this case, the minimum of the potential does not occur when $\phi=0$, but rather when $|\phi|^2 = \mu^2/2\lambda$. There are a continua of solutions at this minimum potential, analogous to the way that a marble in the brim of a Mexican hat can roll in a circle; this degree of freedom corresponds to a (massless) Goldstone boson [20]. In the generalization of Equation 1.4 to the electroweak theory with four gauge fields (gauge group $SU(2)\times U(1)$), three degrees of freedom are “eaten” by the massive W^\pm and Z upon gauge transformations on ϕ and the gauge fields of the Lagrangian[21]⁵. The photon remains massless.

³ A_μ is the field associated with the $U(1)_Y$ gauge field of the Standard Model

⁴this is the Lagrangian of quantum electrodynamics, a spin-0 field that is covered by the Klein-Gordon Lagrangian rather than the Dirac Lagrangian

⁵for a nice explanation of this process, see e.g.[11]



Figure 1.1: The radiative corrections to the Higgs mass can be seen in these Feynman diagrams, with a fermionic correction to the mass shown on the left and the correction from a scalar particle on the right. One problem with the bare Standard Model is that if it is valid up to the Planck scale, 10^{17} GeV, corrections such as these need to cancel to an incredible degree of precision, one part in 10^{30} , to allow the Higgs to have a mass of 126 GeV. This suggests to us that new physics at sub-Planck-scale energies might be well-motivated.^[3]

The gauge transformation that leads to masses for W^\pm and Z also “rotates away” the Goldstone boson associated with the extra degree of freedom, which is a nice feature since experimentally we see no such particle. What is left in its place, however, is a massive scalar field, and we call the associated particle the Higgs boson. The Higgs field also endows mass upon the fermions, which it interacts with via the Yukawa couplings.

1.4 Supersymmetry

Despite its robustness in the face of experimental scrutiny, the Standard Model has several important shortcomings. One of the most important is the hierarchy problem, which refers to the divergence of the Higgs mass via contributions from quantum loop corrections. On one hand, it is true at tree level (and now experimentally verified) that the Higgs mass is of the same order of magnitude as the masses of the electroweak bosons. On the other hand, the Higgs mass can receive large corrections

via its coupling to fermion-antifermion pairs⁶ and W and Z bosons, and even self-coupling of the Higgs to itself, all of which introduce correction terms to the Higgs mass. For example, the diagram shown in Figure 1.1 (a) contributes the following correction term to the mass:

$$\delta m_H^2 = -\frac{|\lambda_f|^2}{8\pi^2} [\Lambda_{UV}^2 + \dots] \quad (1.5)$$

Λ_{UV} is the ultraviolet cutoff scale, the energy at which the Standard Model breaks down and new physics must enter the picture, and λ_f is the Yukawa coupling to the fermion in question. The exact value of Λ_{UV} is not known, but a reasonable a priori guess would be the Planck scale, about 10^{19} GeV. However, since the experimentally measured mass of the Higgs is about 126 GeV, there needs to be fine-tuning on the scale of one part in 10^{30} for the numbers to come out correctly.

Supersymmetry solves this problem by introducing a mirror set of particles to the Standard Model particles, where each SM boson has a corresponding supersymmetric fermion and each SM fermion has a SUSY boson. These SUSY particles would also couple to the Higgs, and introduce additional terms to the mass, namely

$$\delta m_H^2 = \frac{|\lambda_S|^2}{16\pi^2} [\Lambda_{UV}^2 - 2m_s^2 \ln(\Lambda_{UV}/m_s) + \dots] \quad (1.6)$$

The ultraviolet cutoff Λ_{UV} enters again here, this time with the opposite sign. λ_S is the quartic coupling of the Higgs boson to a generic scalar S . If $\lambda_S = \sqrt{2}\lambda_f$, which is the relationship imposed by SUSY, then the terms proportional to Λ_{UV} cancel and the mass divergence is averted[11, 3].

If the scalar particle in Figure 1.1(b) is the superpartner of the fermion in 1.1(a), the Yukawa couplings λ_s and λ_f could quite elegantly have very similar values and the

⁶most notably the top quark, since the size of the correction is proportional to the mass of the particle in the loop correction

large correction terms cancel, leaving us with a Higgs mass much closer to the value seen experimentally.

Taken at face value, SUSY requires that the supersymmetric partners have the same masses as their SM counterparts. If this were the case, though, we would have seen the SUSY particles already—they would have been produced at experiments at the Tevatron and LEP, to say nothing of the LHC. The lack of SUSY particles with the same masses as their SM counterparts leads us to say that supersymmetry is “broken” (in the sense that it is not a symmetry of the vacuum, not in the sense that it does not hold theoretically). When we consider broken supersymmetric theories, the masses of the SUSY particles can become large, and indeed the mass limits for SUSY particles now require that most of them are at least several hundred GeV, up to several TeV. One problem with this trend is that the supersymmetric solution to the hierarchy problem assumes SUSY partners that are light, less than a TeV or so, so that the SUSY contributions to the Higgs mass are roughly the same size as the SM contributions. As direct and indirect searches exclude much of the phase space below a few TeV, it becomes harder to find SUSY scenarios that elegantly solve the hierarchy problem and SUSY becomes less appealing in that sense.

SUSY has additional motivations, however. Another appealing feature of SUSY is that it changes the logarithmic evolution of the coupling constants of the three SM forces. At very high energies the coupling constants come close to the same value, but do not quite match up, a theoretically unsatisfying fact that SUSY addresses. When SUSY enters the picture, the running values of the coupling constants change such that they unify at a high energy scale.

A third appealing feature of SUSY is that it provides a natural candidate for dark matter. From studies of galactic rotational curves, and other astrophysical investigations, it seems clear that there is a significant amount of matter (“dark matter”) floating around the galaxy that does not interact via the electromagnetic or strong nuclear forces, but does interact gravitationally (it is not known whether it interacts

Unification of the Coupling Constants in the SM and MSSM

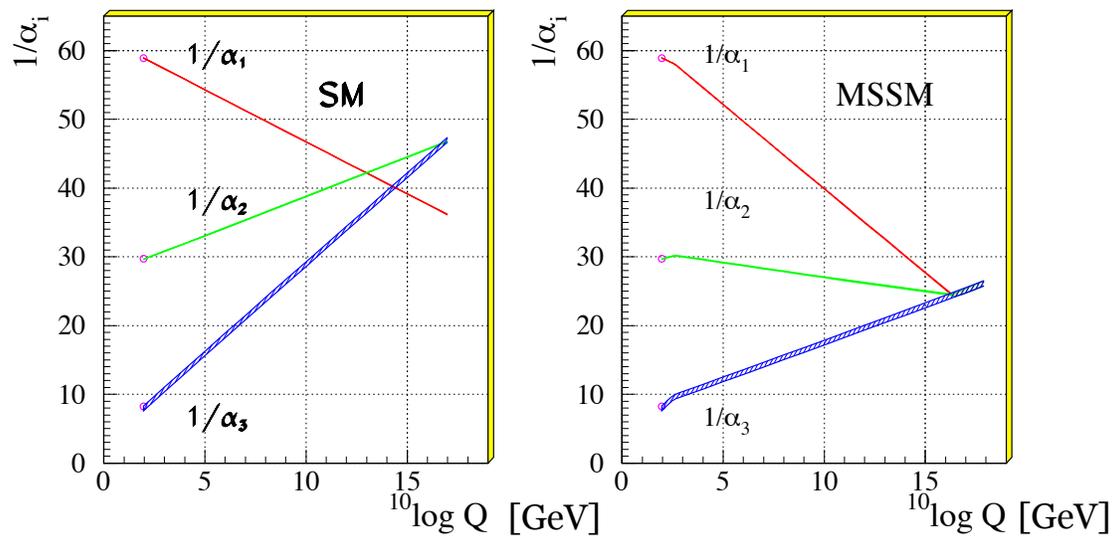


Figure 1.2: The coupling constants for the three fundamental forces of the Standard Model, as a function of energy. On the left, the coupling constants almost unify (but don't quite meet) in the SM-only framework, whereas on the right there is unification when SUSY effects are allowed to enter the picture at around 1 TeV[4].

via the weak nuclear force; a number of experiments are attempting to detect it via the weak force but so far have not produced clear and unequivocal evidence). Among the supersymmetric particles would be dark matter candidates—particles such as the (fermionic) neutralinos which can be heavy, and thus provide the gravitational interaction of dark matter. At the same time, most SUSY scenarios preserve a quantity called R-parity, which effectively means that the lightest supersymmetric particle is stable and thus would not decay away to Standard Model particles, nor interact via the strong or electromagnetic force. Additionally, the supersymmetric dark matter particles can be “thermal relics,” meaning their annihilation in the early universe stops at about the right point to give them the mass density that we see in the astronomical dark matter. In short, the neutralinos of supersymmetry are candidates for dark matter, which serves as an attractive feature of the theory.

One less appealing feature of SUSY is that it is a very unconstrained set of theories—depending on the details of the SUSY version involved, there can be dozens of new supersymmetric particles that might enter the picture. Similarly, SUSY Lagrangians can have many free parameters governing the masses, interactions, etc. of the SUSY sector and it is very difficult, perhaps impossible, to probe all the SUSY phase space. Phenomenologists address this problem in a number of ways, the most important of which for this thesis is the proposal of the MSSM, or Minimally Supersymmetric Standard Model. The MSSM makes a number of assumptions about the SUSY parameters and their relationships so as to constrain the number of free parameters to 119⁷.

1.5 Higgs Physics in Supersymmetry

Once the constraints of the MSSM have restricted the SUSY phase space to a more tractable set of parameters, we can see the impact of SUSY on the Higgs sector. The

⁷in contrast, the Standard Model has 19 free parameters

MSSM also contains (at least) 5 Higgs bosons on account of the two complex Higgs doublets in the theory (these models are special cases of a more general framework called Two Higgs Doublet Models, or 2HDM).⁸

Both Higgs doublets acquire a vacuum expectation value (VEV) with values v_1 and v_2 respectively. The interaction strength associated with muon decay $G_F=1.16639 \times 10^{-5} \text{ GeV}^{-2}$ provides an important constraint on the value of v_1 and v_2 , namely that

$$v^2 = v_1^2 + v_2^2 = \frac{1}{\sqrt{2}G_F} = (246 \text{ GeV})^2 \quad (1.7)$$

There is one additional complication to the 2HDM formalism. In its most general form, the Higgs system has CP-violating couplings and flavor-changing neutral currents (FCNC), the latter of which in particular is tightly constrained by experimental evidence. The Glashow-Weinberg condition explains that if only one Higgs doublet couples to fermions of a given electric charge, there is no Higgs-induced CP violation or FCNC. There are four ways that the Glashow-Weinberg condition can be met, and the MSSM is consistent with the so-called “type II” Higgs doublet model, where one doublet couples exclusively to up-type quarks and the other couples exclusively to down-type quarks. Then v_1 is the VEV of the Higgs field coupling to up-type quarks and v_2 is the VEV to down-type quarks. While equation (1.7) constrains their sum, the ratio of the two values is a free parameter of the system and is denoted by $\tan \beta$:

$$\tan(\beta) = \frac{v_u}{v_d} = \frac{v_1}{v_2} \quad (1.8)$$

The 2HDM models imply the existence not of one Higgs boson, but of five. The 5 particles include two CP-even particles, h and H , one CP-odd particle A , and two

⁸the 2HDM of the MSSM has fewer free parameters than the most general 2HDM

electrically charged particles H^\pm . This analysis is a search for both the CP-even H and the CP-odd A ; the CP-even h is assumed to be the Higgs particle found at 126 GeV [1]. The masses of H/A are not known. While there are several free parameters in the MSSM, the SUSY Higgs sector is (at first order) governed by only two: m_A and $\tan\beta$. As a direct result of this, most interpretations of limits (or signal) are presented in terms of the $m_A/\tan\beta$ phase space favored or excluded.

1.5.1 H/A Searches in the $b\bar{b}$ Final State

The H/A search being performed in this thesis is based on two important experimental choices:

- The Higgs boson H or A is produced in association with one or more b -quarks
- H/A decay to a pair of b -quarks

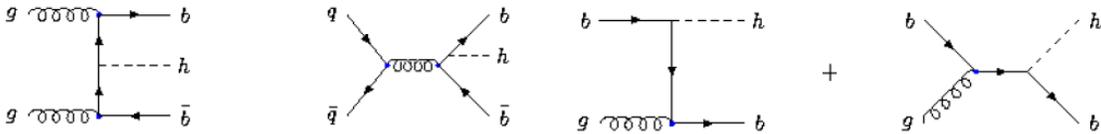


Figure 1.3: Feynman diagrams for leading-order production of h in association with b -quarks.

Technically speaking, there is not a Feynman diagram for H/A production that occurs with exactly additional one b -quark from associated production; even in the third and fourth diagrams in Figure 1.3 there is a second b -quark that comes from the PDF (parton distribution function) of the second proton. In practice, the fourth b -quark tends to be both soft in p_T and far forward in the detector, making it difficult to see experimentally. For that reason, we only require the presence of one b -quark in addition to the $b\bar{b}$ pair coming from the H/A decay, which makes for a $b\bar{b}b$ final state.

While the behavior of the Higgs system can be complicated to fully map out, there are several general trends that emerge when one examines the parameters m_A and $\tan \beta$:

- The cross section for H/A production in association with b -quarks increases for higher values of $\tan \beta$ (Figure 1.4)
- The branching fraction of H/A to b -quarks increases with $\tan \beta$ (Figure 1.5)
- At high $\tan \beta$ ⁹, the $H/A \rightarrow b\bar{b}$ branching fraction is nearly constant across a wide range of m_A (Figure 1.5)
- For a given $\tan \beta$, the production cross section falls for high m_A (Figure 1.4)
- The masses of H and A , their kinematics, and the $H/A \rightarrow b\bar{b}$ branching fractions are nearly the same, so that for search purposes, one can treat them as one particle with $\sigma \times BR$ twice that of H or A individually (Figures 1.4, 1.5, 1.6, 1.7)
- The intrinsic width of H and A increase as with m_A and $\tan \beta$ (Figures 1.6, 1.7)

As might be apparent, this analysis is most advantaged in a regime when $\tan \beta$ is large, since both the production cross-section with associated b -quarks and the branching fraction to $b\bar{b}$ benefit from the enhanced coupling. A low $\tan \beta$ search would be more suited for a $H/A \rightarrow t\bar{t}$ final state, or for lower masses, $H \rightarrow hh$ or $A \rightarrow Zh$.

An important and subtle point worth highlighting is that, while A , H , and H^\pm are often called the ‘‘SUSY Higgs bosons’’, they do not carry R-parity and experimentally decay to SM particles¹⁰ The Feynman diagrams for the production of H and A in association with b -quarks can also be drawn for b -quark associated production of the

⁹in this context, 30 (and above) would be an example of a high $\tan \beta$ value

¹⁰in contrast, direct SUSY searches are often characterized by missing energy in the final state, as SUSY particles in the final state will fail to interact with the detector

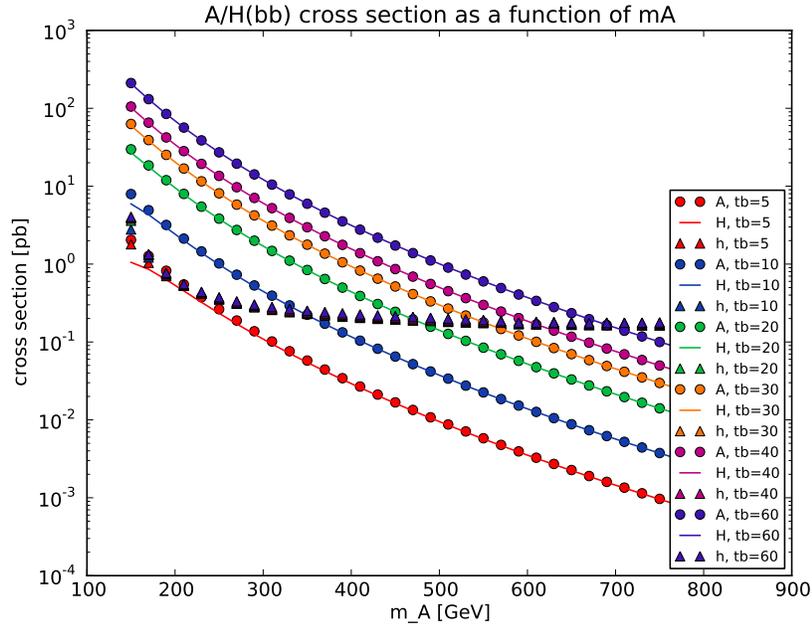


Figure 1.4: The cross section for b -quark associated production for A , H and h in proton-proton collisions at $\sqrt{s} = 8$ TeV as a function of m_A for several different values of $\tan\beta$ (denoted “tb” in the legend). Figures made using FeynHiggs [5, 6, 7, 8, 9]. There are no additional cuts on the associated b -quark p_T beyond the default cuts in FeynHiggs.

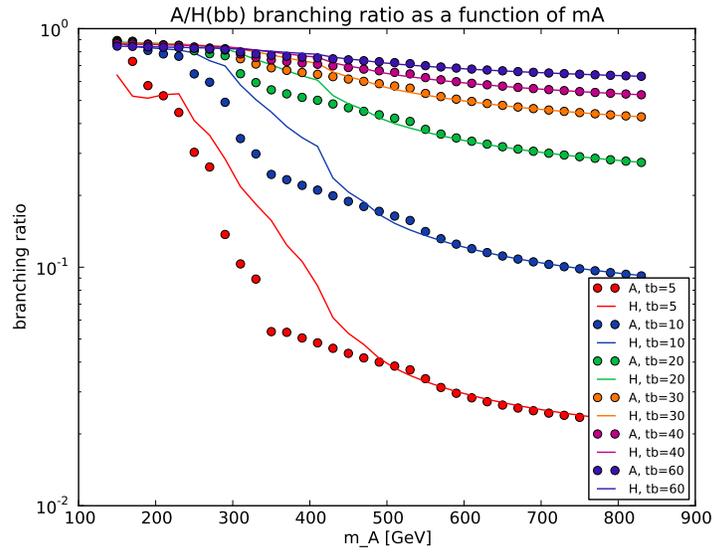


Figure 1.5: The branching ratio to $b\bar{b}$ of A , H and h as a function of m_A for several different values of $\tan\beta$ (tb). The trend of higher branching ratios for higher values of $\tan\beta$ comes from the Higgs-to- b -quark decay vertex carrying an enhancement factor that scales with $\tan\beta$. Figures made using FeynHiggs [5, 6, 7, 8, 9].

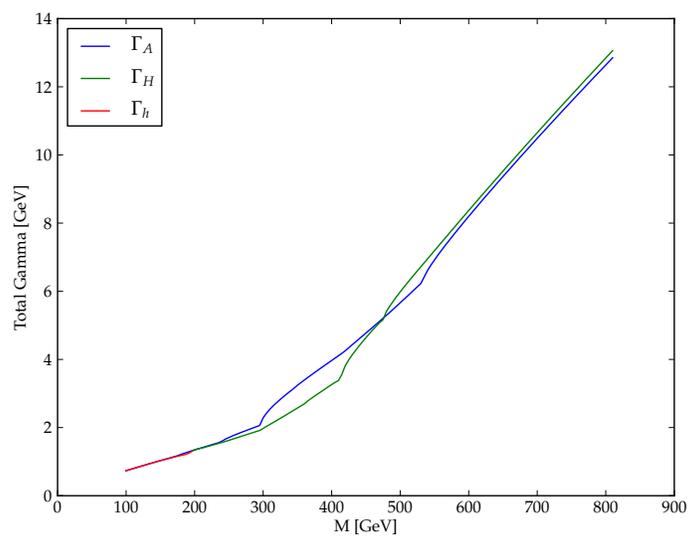


Figure 1.6: The intrinsic width of A , H , and h as a function of m_A for $\tan\beta=20$ in the m_h^{max} scenario. In practice, the intrinsic width of the particle is dwarfed by the experimental resolution, which can be up to hundreds of GeV. Figures made using FeynHiggs [5, 6, 7, 8, 9].

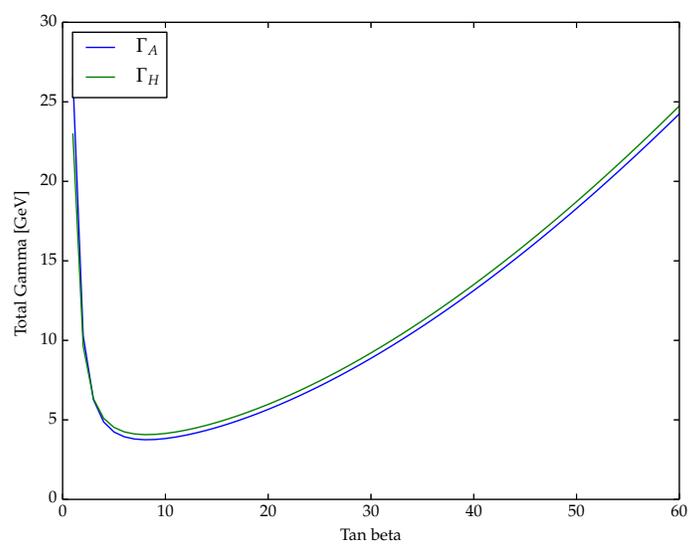


Figure 1.7: The intrinsic width of A and H as a function of $\tan\beta$ in the m_h^{max} scenario when $m_A=500$ GeV. Figure made using FeynHiggs [5, 6, 7, 8, 9].

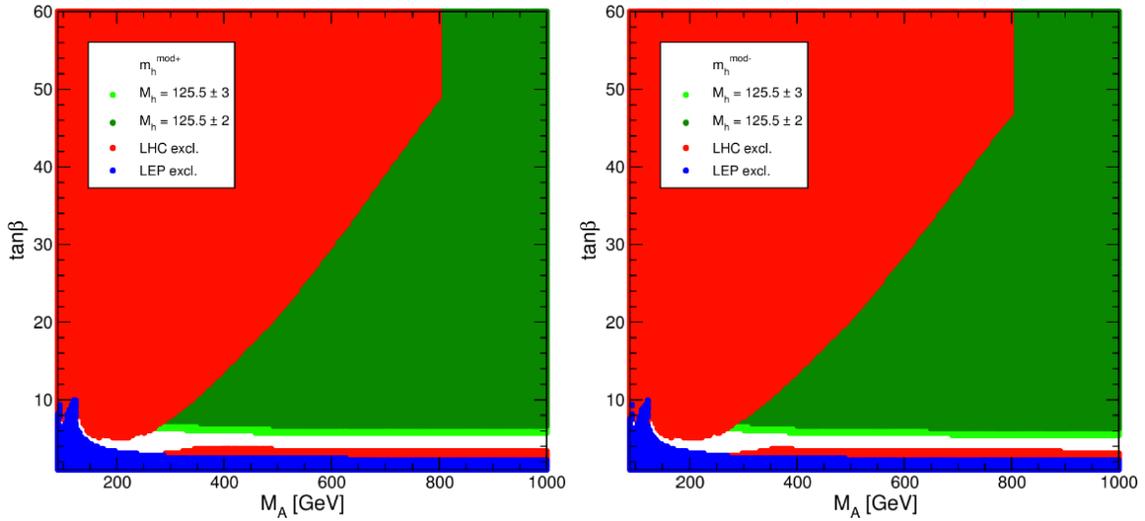


Figure 1.8: The favored (green) and excluded (red/blue) regions in $m_A/\tan\beta$ in the m_h^{mod} scenario, where the green region shows the parameter space that is consistent with $m_h=126$ GeV. After the discovery of h at 126 GeV, the m_h^{mod} scenario is favored because it allows for m_h to have this relatively low value (m_h^{max} , on the other hand, pushes m_h more toward 140 GeV). The left and right plots show the parameter space for m_h^{mod+} and m_h^{mod-} , respectively, the details of which can be found in [10]. Clearly in this scenario, many values of m_A are open, potentially with large cross-sections.

SM Higgs boson h ; the most important contribution of SUSY is that, when $\tan\beta$ is large, it provides an enhancement factor to the Hbb and Abb vertices that drives the production cross section up to a magnitude that could be large enough to see in 20 fb^{-1} of proton-proton collisions at $\sqrt{s}=8$ TeV.

However, that does not mean that the SUSY Higgs sector is independent of the details of the supersymmetric parameters. In particular, it is possible that H/A decay to a pair of SUSY particles, such as charginos or neutralinos. Depending on the scenario, the branching fraction to SUSY particles could be substantial. With the large number of free parameters, it is impractical to perform a complete scan of the MSSM parameter space, so a few benchmark scenarios are chosen where only m_A and $\tan\beta$ are allowed to vary, and the other SUSY parameters are fixed. The leading benchmark scenario is the so-called m_h^{mod} scenario, which is an update of the m_h^{max} scenario that was used for many years. In the m_h^{max} scenario, the fixed SUSY parameters are assigned values that maximize m_h , the mass of the light CP-even Higgs. However, in this scenario, m_h is predicted to be around 140 GeV, while the experimentally found value is known to be around 126 GeV. In order to reconcile these numbers, the updated m_h^{mod} scenario has become the more useful benchmark, since it does allow for $m_h=126$ GeV. There are two types of m_h^{mod} scenarios, called m_h^{mod+} and m_h^{mod-} , that differ by the sign of higher-order correction terms.

Other non-MSSM scenarios also have final states with a Higgs boson decaying to b -quarks and produced in association with b -quarks, with cross-sections that could be seen in a few tens of inverse femtobarns at the LHC [22].

1.5.2 Constraints on H/A Existence and Kinematics

The discovery in July 2012 of an SM-like Higgs boson provides important constraints of the MSSM Higgs sector, but leaves other aspects of the theory tantalizingly unconstrained. On the one hand, for many values of m_h , measuring the exact value of

m_h places a strong constraint on m_A . An important exception is when m_A is much larger than the mass of the SM Higgs boson m_h , which is called the decoupling limit, where the properties of h are unaffected by the existence of H/A . This means that for a light m_h below 140 GeV or so (recall that the mass of the discovered particle is about 126 GeV), it is very difficult to ascertain constraints on m_A by studying the properties of h . As LHC searches advance, though, they constrain our understanding of the BSM Higgs sector from a number of directions:

- The chargino and neutralino mass limits suggest that the higgsino mass parameter μ is large, above several hundred GeV
- $B_s^0 \rightarrow \mu\mu$ appears consistent with the Standard Model predictions, placing strong constraints on the possibility of nonstandard Higgs contributions to that decay at loop level [23]
- In the MSSM, H^+ and H^- are very close to H and A in mass, so searches for H^\pm can constrain the likely MSSM phase space available for H and A (Figure 1.9). In practice, the mass regime probed by this analysis (450-800 GeV) is not constrained by searches such as e.g. [24]

1.6 From Beautiful Theory to Messy Experiment

It is the experimentalist's job to understand what all this theoretical physics actually looks like in the detector. This is a deep topic, and one could go into great detail, but a few key trends are outlined here.

First, as we've implied throughout this section, many of the particles for which a physicist searches cannot be seen directly—they live for a fraction of a second, and then decay into other particles. Those daughter particles often decay themselves, and so on, in a multi-step decay chain. Only the particles from the last stage in this

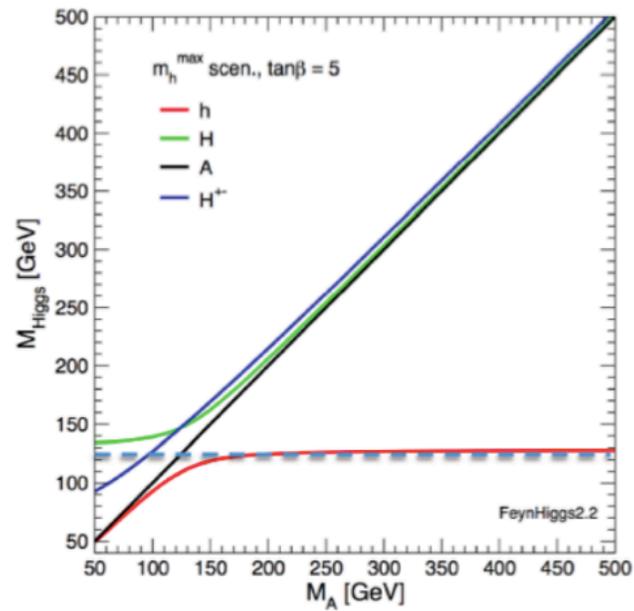


Figure 1.9: The mass evolution of the 5 Higgs bosons of the MSSM, with the mass of the Higgs-like indicated by dashed line (126 GeV). For a large range of h masses, the masses of H , A , and H^\pm evolve very closely together, so searches for H^\pm also can provide some indirect constraints on the likely masses of H and A [11].

chain actually get detected, so part of the physicist's job is to reconstruct the original particle(s) from the daughter particles. Once the daughter particles are identified, all that is needed to reconstruct the parent particle is the transverse momentum (see below), energy, polar angle, and azimuthal angle. These components go into a Lorentz 4- vector for each particle, and then the Lorentz-vectors can be added together in a relativistically invariant way to get the mass and flight direction of the parent.

Second, certain types of particles leave distinctive signatures. One important example is that particles associated with QCD, quarks and gluons, are subject to the laws of QCD confinement. Once they are produced or excited in the hard scatter, quarks and gluons hadronize and shower. A spray of particles, largely pions, show up in the detector in the general direction of the original parent quark or gluon –if all the shower particles could be collected, they could be recombined into the parent particle. These showers of particles are called jets, and are often named after the parent particle, such as b -jets, light-quark jets and gluon jets. The topic of jet reconstruction is detailed further in Section 3.2. Photons and electrons shower as well, but they generate much narrower electromagnetic showers that are easy to identify and reconstruct, when compared to most QCD showers. Muons have the cleanest signature of all; they are relatively long-lived and do not decay before exiting the detector, so they generally make a single charged track that is distinctive in the dedicated muon layers of the detector.

A third important feature of particles as they appear in the detector is their p_T , or transverse momentum. Since the initial state of each collision involves two protons, which are composite particles, it is impossible to know the exact longitudinal momentum of two quarks when they collide. However, the transverse momentum is zero, so in the final state, the transverse momentum must also sum to zero. When heavy particles are created and then decay, or have other particles deflecting off them, the p_T of the resulting particles can be tens to hundreds of GeV, which creates

a striking signal in the detector and provides an experimental handle when trying to understand the event.

The main challenges of this particular search arise from several kinematic features. First, the final state is all-hadronic, where reconstruction is difficult and the QCD background is large. Second, since the mass of H and A are relatively unconstrained, a search has to span a large range in m_A in order to be sensitive to discovery. Third, the combinatorics of selecting the correct two b -jets for reconstruction (from the three or more available in an event) require dedicated study.

1.7 Previous Searches

Searches for A and H in this final state have been performed at the Tevatron as well as CMS. The searches at the Tevatron were particularly tantalizing because both CDF [25] and D0 [26] spotted slight excesses of events, around 2σ standard deviations each, between approximately 100 and 150 GeV. The LHC experiments (ATLAS and CMS) present an opportunity to re-examine the same region of phase space, albeit with higher integrated luminosity and center-of-mass energy. CMS does not observe an excess in the 100-150 GeV range when searching over approximately $5 fb^{-1}$ of data at $\sqrt{s}=7$ TeV [27].

This analysis serves as an important update and extension of the searches done in other experiments. Due to differences in the triggers for these analyses, the ATLAS search does not have access to m_A values below 450 GeV, but it does provide the first limits between 450-800 GeV. This analysis is also the only search at $\sqrt{s}=8$ TeV, and searches over an integrated luminosity ($19.5 fb^{-1}$) that is about 4 times higher than the next largest search, done at CMS.

Chapter 2

CERN, the LHC Accelerator and the ATLAS Detector

2.1 Introduction

Experimental particle physics research has many prominent features, but the most prominent might be the scope and collaborative nature of the work. This thesis is one of many that have been produced at the ATLAS detector at CERN, and makes use of the dedicated work of many physicists to build and maintain the experimental apparatus and data processing pipeline. The Large Hadron Collider is the first major piece of this process, where beams of protons are accelerated to tremendous energies and the smashed into each other inside the cathedral-size detectors, including ATLAS. It is the collisions between individual pairs of protons that underpin the entire scientific mission: in this environment, interesting and rare physics can result from the interactions between the protons. It is the mission of the LHC to generate large numbers of high-energy proton-proton collisions, and then the mission of the detectors to record and analyze the data that result from the collisions.

2.2 CERN and the Large Hadron Collider

The European Center for Nuclear Research (CERN) is a scientific complex in Geneva, Switzerland that was established in the years following World War II. Its focus is experimental particle physics, in particular the Large Hadron Collider (LHC), and it has a long and storied history of producing particle discoveries and other scientific advances.

The LHC is the particle accelerator that serves as the centerpiece of CERN's scientific program. The accelerator is an RF-driven circular synchrotron located in an underground tunnel with a circumference of 27 km, and when it is running, it accelerates and collides two proton beams (sometimes two lead ion beams, but we will focus on the protons in this thesis) at four interaction points around the accelerator ring. Each of the proton beams contains thousands of bunches of protons which are very deliberately spaced so that every 50 ns, two of those bunches going opposite directions will overlap at one of the interaction points. In 2012, the energy of the proton beams was 4 TeV each, making for a total of 8 TeV of energy when the beams are collided head-on. These are the highest-energy proton-proton collisions ever created by man-made acceleration, and reliably making lots of collisions is a scientific and engineering endeavor completely in its own right, even before considering the physics insights that we get as a result.

There are two major figures of merit for any particle accelerator: energy and luminosity. In a proton accelerator¹, energy is typically limited by the circumference of the accelerator and by the magnetic field strength of the accelerator's bending magnets. The record-breaking energy of the LHC is a function of its tremendous size (27 km in circumference) as well as 1232 8.4 tesla superconducting dipole magnets that keep the beams in place within the accelerator. Luminosity is a measure of how many collisions the accelerator can produce, with higher luminosities meaning

¹energy in circular electron-electron or electron-positron colliders is limited by synchrotron radiation losses

that more interesting events (and non-interesting events!) can be generated. Since different processes happen at different rates, the rate at which a given process occurs is usually stated in terms of its production cross section. Cross sections are measured in units of barns (or fractions thereof), with one barn being equal to 10^{-28} m². Luminosity is measured in inverse barns, so that the cross section can be multiplied by the luminosity to give a total number of events for that type of event. Luminosity is usually limited by the intensity and focus of the beam, and the bunch crossing rate.

Energy and luminosity go hand-in-hand to push back the frontiers of physics. One of the major goals of the LHC is to discover new particles, and by conservation of energy the higher the LHC can go in energy, the heavier the new particles it can potentially create. Another important fact is that, while cross sections for all types of processes typically go up as the collision energy increases, many new physics models have production cross sections that go up faster than the backgrounds as a function of center-of-mass energy. In other words, for a given process, the signal-to-background ratio tends to go up as energy increases so new physics can be much easier to spot at higher energies than at lower energies.

Discovery reach can also be extended by increasing the luminosity of a machine, since sensitivity to a signal is typically defined as s/\sqrt{b} —for example, a 10-fold increase in luminosity means a $\sqrt{10}$ -fold increase in a search’s sensitivity. Luminosity can be increased by either tuning the accelerator to have a higher instantaneous luminosity, with a more focused beam or higher beam current for example, or by simply running the machine for a longer time. In practice, increasing instantaneous luminosity at the LHC also means increasing the number of simultaneous lower-energy collisions that, while not of physical interest, still get picked up in the detector (pileup). The removal of pileup is an intense and dedicated effort that will be addressed in Section 3.3.

The LHC is only the last in a chain of accelerators, which all work together to create the high energy and luminosity that we associate with the LHC. The first step in

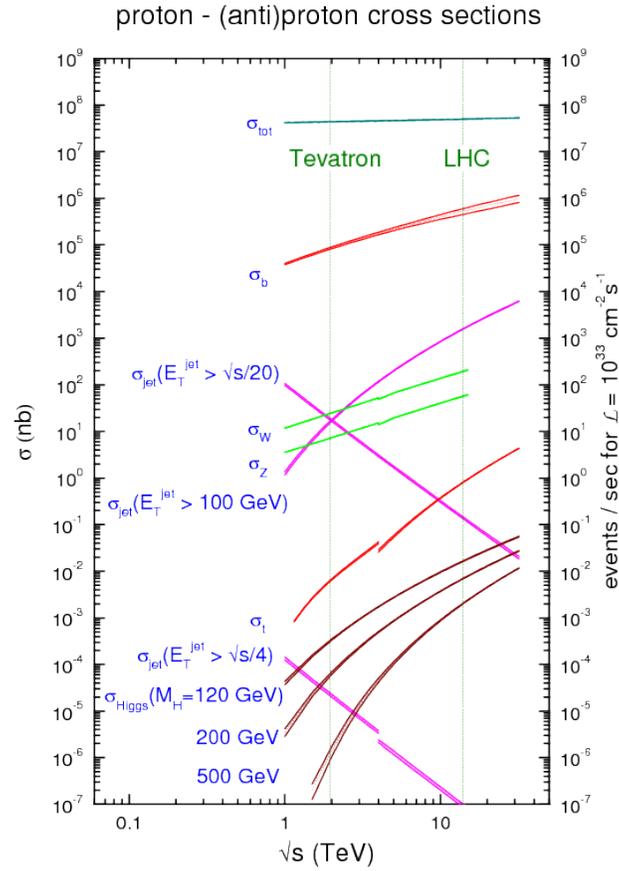


Figure 2.1: Cross sections for various processes as a function of center-of-mass energy of the original collision. The “skip” at 4 TeV occurs because below that value, the cross sections are for proton-antiproton collisions like those made at the Tevatron, while higher energy proton-proton collisions are being produced at the LHC.

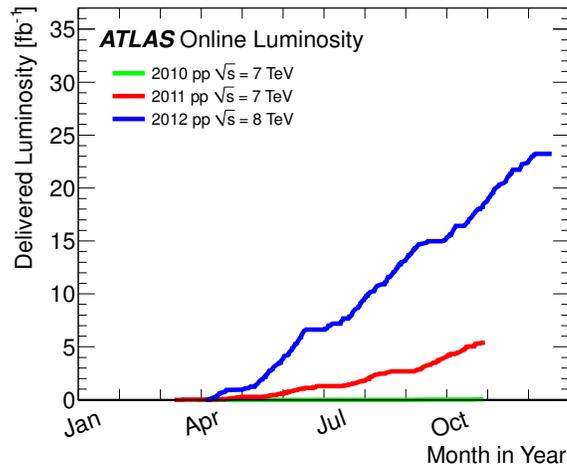


Figure 2.2: The luminosity collected in 2010, 2011 and 2012 at ATLAS.

the process is the proton source, where hydrogen atoms are subjected to an electric field that separates the protons from the electrons. Then the protons are accelerated by a 90kV electric field and focused, and fed into a linear accelerator. From this point forward, all the accelerators use radio frequency (RF) electromagnetic waves to impart energy to the protons.

The linear accelerator brings the protons up to an energy of 50 MeV, then they go to the Proton Synchrotron Booster (up to 1.4 GeV), and then the Proton Synchrotron (25 GeV). Then the protons are passed to the Super Proton Synchrotron, where they reach an energy of 450 GeV, and finally the beam is split into half and each half is injected into the LHC going a different direction. Figure 2.3 shows all the major accelerators at the CERN site.

Once the beams are in the LHC, they are accelerated up to the full collision energy of 4 TeV per beam. Then the beams are focused and carefully steered into position so that head-on collisions commence. The protons are arranged in bunches within the beams, with each bunch 50 ns away from its neighbors, so that the collision frequency within the detectors is 20 MHz. The data-taking is organized into runs,

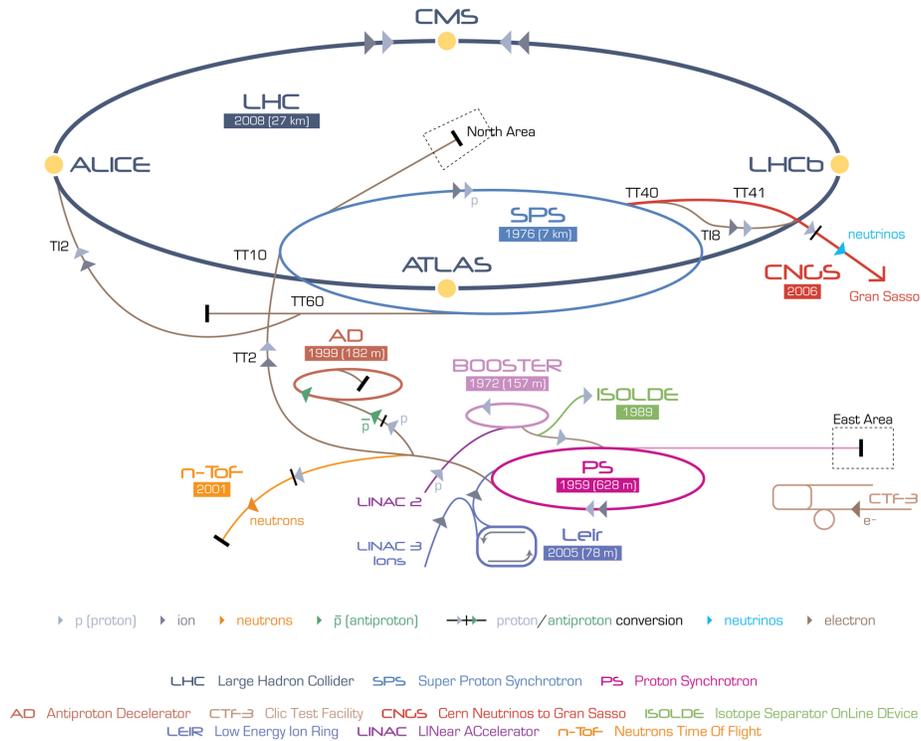


Figure 2.3: The major elements in CERN’s accelerator complex, which turns ordinary hydrogen atoms into the highest-energy man-made collisions ever made. The four interaction points, labeled for their respective experiments (ALICE, ATLAS, LHCb and CMS) can be seen around the perimeter of the LHC.

where collisions are created continuously for up to 16 or so hours at a time. A run typically ends either when enough of the beam gets burned away by the collisions that the instantaneous luminosity drops and it becomes worthwhile to refill, or when a technical glitch causes the beams to be dumped.

The history of the LHC is a dramatic one. The LHC first turned on in 2008, but within days had to shut down following a catastrophic failure of the bus bar between superconducting magnets. It restarted in 2009 at 7 TeV, half the design energy of 14 TeV (a repercussion of the accident), and the first run for physics was performed in 2010, with a total ATLAS dataset size of 45 inverse picobarns. The progress in 2011 was dramatic, as the beam intensity was steadily increased over the course of the run to create an exponential increase of the luminosity recorded, and at the end of 2011 the total dataset size was about five inverse femtobarns (a factor of more than 100 greater than the data recorded in 2010). The progress continued in 2012, with a slightly increased energy of 8 TeV and about 20 fb^{-1} collected, culminating in the 4 July 2012 announcement of the discovery of a new particle, later confirmed to be the long-sought Higgs boson. The total collected luminosity is illustrated in Figure 2.2. At the end of 2012, the LHC commenced a 2-year shutdown to complete upgrades and repairs, with the intention of starting up again in 2015 at 13 TeV and 25 ns spacing between bunch crossings. When the LHC starts up again, the search for new particles will begin again as both energy and luminosity records are broken by the upgraded machine.

2.3 The ATLAS Coordinate System

The ATLAS detector is situated at Point 1 of the LHC ring. Since the detector has a cylindrical shape, polar coordinates are the most natural basis for describing its geometry. The radial direction (usually notated “ r ”) is defined as the vector originating at the interaction point, in the middle of the detector, and pointing

transverse to the beamline. The azimuthal angle ϕ and the polar angle θ complete the coordinate system, although in practice it is more common to use pseudorapidity η instead of θ .

The definition of pseudorapidity η is

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right) \quad (2.1)$$

Pseudorapidity has the advantage of shifting additively under changes in the longitudinal boosts with velocity β :²

$$\eta \approx \eta + \ln \sqrt{\frac{1-\beta}{1+\beta}} \quad (2.2)$$

Since the proton is a composite particle, made of three valence quarks and an indeterminate number of gluons, and sea quarks, even if the proton has a known energy, its constituent partons can have a wide range of energies. In general, then, the exact partonic collision energy can vary from event to event, and outgoing particles can be boosted forward in η . However, energy is conserved in the ϕ dimension, so transverse momentum (\vec{p}_T) should add up to zero in events where there are no particles escaping detection.

2.4 The ATLAS Detector

The ATLAS detector is like an onion: it has layers, and it makes you cry. There are several main subsystems, each of which is designed to measure a certain characteristic (charge, momentum, energy, etc.) of a particular type of particle (quark, electron,

²This statement is an equality when referring to the rapidity y , and an approximation when referring to the pseudorapidity η . In highly relativistic regimes, like those experienced by particles in the LHC, $y \simeq \eta$.

muon, etc.). The subsystems work hand-in-hand with the trigger and data acquisition system, which handles the decision-making of which events to record and goes about writing those events to disk.

2.4.1 Inner Detector

The innermost layer of the ATLAS detector is the tracker, which provides precision measurements of the trajectories of charged particles. As a charged particle traverses the layers of the tracker, it ionizes the detector material which creates small electrical signals that can be amplified and read out by the system. These so-called “hits” are combined during reconstruction into tracks, which represent the paths of particles like electrons, muons, and charged pions in the detector.

The information from the tracker is used in particular to determine the transverse momentum (p_T) of charged particles, and to perform particle identification. The tracker consists of three primary subsystems: the pixels, semiconductor tracker (SCT), and transition radiation tracker (TRT). A charged particle created at or near the interaction point would typically travel through all three subsystems, creating some number of hits in each one. On average, a track has 3 pixel hits, 8 hits in the SCT (4 double-sided hits) and about 34 hits in the TRT. All the inner detector subsystems are enclosed by a 2 tesla solenoidal magnetic field.

Pixel System

The pixel system sits physically closest to the beam line and interaction point. It is built of silicon pixels that measure 50×400 micrometers each, which are organized into sensors. Each sensor contains 46,080 pixels, and then there are 16 sensors organized onto a pixel module. The pixel system as a whole contain 1744 modules, organized into 3 layers each in the barrel and the two endcaps. The pixel system in

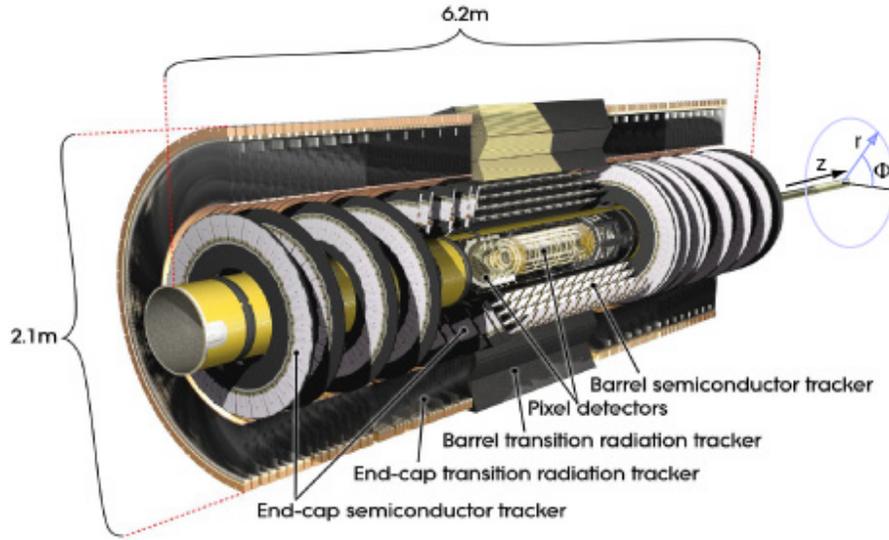


Figure 2.4: A cutaway picture showing the main components of the ATLAS inner detector. The ATLAS coordinate system is sketched on the right-hand side of the figure.

aggregate contains 80 million channels and measures about 1.4 meters long by 0.4 meters in diameter.

The pixel technology is designed to give high-precision measurements of the location and momentum of charged tracks. A pixel module has two main components, the silicon sensor and the front-end chip, which are bump-bonded together. When a charged particle traverses the sensor it ionizes silicon atoms, creating electron-hole pairs. A bias voltage applied across the sensor causes the electrons and holes to drift to opposite sides of the sensor, where they can be read out by the front-end chips.

The pixel readout is based on detecting and quantifying this ionization current. When the ionization current reads out, it creates an electric pulse in the pixel electronics that serves as the input to a discriminator. The discriminator is set with a tunable threshold number of electrons, typically 3000, and the signal metric is then

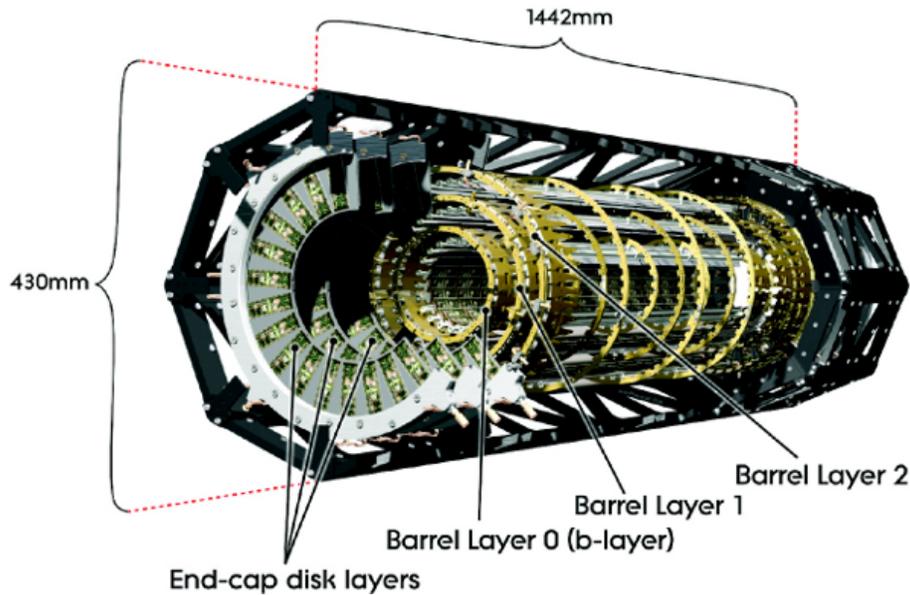


Figure 2.5: A picture showing the main layers of the ATLAS pixel detector, including the disk layers on the end. Starting in 2015, a 4th layer, the insertable b layer (IBL), will be inserted inside the current layer 0.

the length of time for which the pulse was above that value (called “time over threshold”, or TOT). Having a threshold in place helps distinguish ionization current from leakage current, which occurs when the silicon does not have robust insulator characteristics and current begins to flow across the sensor even when there is no ionizing particle present. One important effect of radiation damage to the pixel detector is that it damages the silicon, allowing for increases in leakage current up to 100 nA, so that over the course of large radiation doses to the detector the thresholds sometimes have to be raised to compensate for the damaged material.

Another important consideration when designing and constructing the pixel detector is the material budget of the system. When a charged particle traverses the pixels, interacting with the silicon, its trajectory can change by virtue of these interactions as it undergoes multiple scattering or secondary interactions. This can be a problem, for example, when reconstructing tracks—if a track has a kink from where it scattered

off detector material, it will be more challenging to reconstruct the track or correctly measure its momentum. These material effects are particularly important for the pixel detector, because they are the first layer of detector traversed by particles after they leave the collision point. One way to mitigate these effects is to place high-material components, such as power supplies and readout electronics, in more forward regions so that they are not in the path of central tracks. There are two figures of merit when evaluating the material budget of a system: the radiation length is the distance an electromagnetically interacting particle (such as an electron, positron or photon) travels before losing $1/e$ of its energy to bremsstrahlung, while the interaction length is the mean distance traveled by a hadronically interacting particle (such as a proton, neutron or pion) before undergoing an inelastic nuclear interaction. Figure 2.6 shows the material volume, in units of radiation length vs. η ; clearly, the material budget (especially pixels) is minimized in the most central regions ($|\eta| < 1.4$).

The pixel system all together provides fine resolution of the beamspot and surrounding area, which serves several important purposes. First, since there are typically many hard p-p collisions in each bunch crossing, the pixel system's longitudinal resolution of $z_0 \sin \theta \approx 0.05\text{-}0.3$ mm enables the reconstruction of multiple primary vertices, which are typically separated by a few millimeters [28]. This is critical for controlling pileup in the high-luminosity LHC environment. Second, the transverse resolution of 0.01-0.1 mm enables the track resolution required to allow precision b -tagging for identification of bottom quarks. More details on both the track reconstruction and the b -tagging methodology can be found in Sections 3.1 and 3.5.

Semiconductor Tracker (SCT)

Like the pixels, the silicon microstrip tracker (SCT) is a silicon detector, although the geometry is distinctly different from the pixel geometry. Since the SCT is at a farther radius from the interaction point than the pixels, it experiences a lower

occupancy. This allows for substantially larger detector elements, at a lower cost and using less material in the detector than if the same coverage were implemented in pixels, while maintaining $16 \mu\text{m}$ resolution of tracks in $r\phi$ and $580 \mu\text{m}$ resolution in z [29].

The SCT geometry has some notable features. The SCT consists of 8 layers of microstrips organized into four double-sided pairs, where the two members of each pair have an offset angle of 40 mrad . The SCT has four barrel layers and nine endcap disks, and the barrel modules are oriented with a tilt angle of about 11° angle relative to being perfectly tangential to a circle of constant r in the $r\phi$ plane [30]. Whereas the pixel reads out the TOT of a hit, the SCT has a binary readout: a hit is either recorded or not.

Transition Radiation Tracker

The outermost inner detector layer is the transition radiation tracker, or TRT.

The TRT uses gold-plated tungsten wires embedded in straw tubes of 4mm diameter filled with an $\text{Xe}/\text{O}_2/\text{CO}_2$ gas mixture, with a total of about 350,000 readout channels covering a pseudorapidity range out to $|\eta| < 2.0$. For a typical particle, the TRT will have about 30-35 hits with a hit precision of about $130\mu\text{m}$ [31].

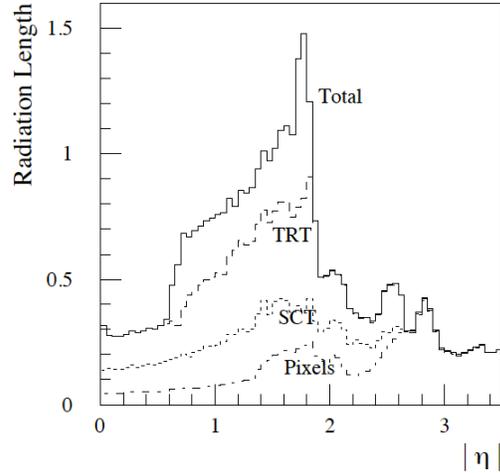


Figure 2.6: The material budget of the inner detector, in units of radiation lengths, as a function of η , showing how the material in the central part of the detector (especially in the pixel subsystem) is kept at a minimum to mitigate multiple scattering.

One important objective of the TRT is to identify tracks from electrons by detecting transition radiation (hence the name transition radiation tracker). Transition radiation occurs when an electron passes between regions with different dielectric constants; at the boundary between those regions, the electron can emit a photon which is then absorbed by the Xe gas and translates into a high-threshold hit in the detector. Electrons can be distinguished from hadrons by the presence of many high threshold hits along the track.

2.4.2 Calorimeters

The ATLAS calorimeters measure the energy of particles. There are two main calorimeter subsystems, one for electromagnetic particles such as photons and electrons, and the other for hadrons, jets, τ leptons, and missing transverse energy. In addition to energy measurements, though, information from the calorimeters also is used for particle identification, finding the direction of electromagnetic and hadronic jets, identifying and measuring missing transverse energy or MET, and selecting jets as part of the trigger.

Electromagnetic Calorimeter

The electromagnetic (EM) calorimeter measures electrons and photons after they exit the tracking system. It was designed with the discovery of the SM Higgs boson in mind, since for many possible values of m_h (including the value where the particle was actually found, $m_h=126$ GeV), the most sensitive search channels contain electromagnetic particles (electrons and photons) in the final state.

Nicknamed the LAr calorimeter for its liquid argon technology, the calorimeter is a sampling calorimeter, with the passive showering material (lead) interleaved with

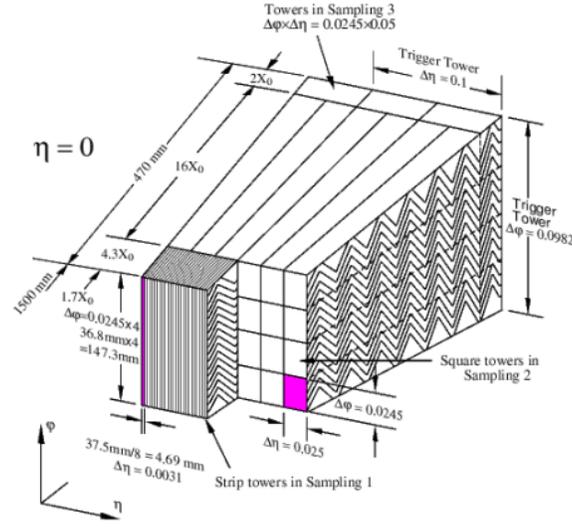


Figure 2.7: A depiction of a section of the LAr calorimeter. The accordion structure of the towers is visible, as well as the three sampling layers.

active energy measurement material (liquid argon). An electron or photon will interact with the lead as it travels through it, creating an electromagnetic shower, which then propagates to the adjacent LAr layer, where it is measured and read out.

A notable feature of the EM calorimeter is the accordion geometry, which has several key characteristics. First, the geometry enables complete coverage in ϕ without azimuthal cracks. Second, the LAr sampling layer between lead layers is constant throughout the calorimeter barrel. Third, a particle traveling through the calorimeter will generate approximately the same number of sampling instances (i.e. measurements) regardless of the direction in which it travels. These pieces add up to a very uniform coverage of electromagnetic calorimetry. The calorimeter consists of two major parts, the barrel and the endcaps; the barrel measures particles with $|\eta| < 1.475$ and the endcaps measure particles with $1.375 < |\eta| < 3.2$. The transverse segmentation of the calorimeter is $\Delta\eta \times \Delta\phi < 0.03 \times 0.03$ over the pseudorapidity region $|\eta| < 2.5$, to allow for the particle identification and energy resolution needed

[32], while the calorimeter is up to 24 radiation lengths thick to minimize punch-through.

The performance requirement for the energy resolution of the EM calorimeter is $\sigma_E/E = \frac{10\%}{\sqrt{E}} + 0.7\%$, which has the nice feature that the resolution improves as the energy of an electromagnetic jet increases. A crucial part of reaching this resolution is precisely understanding the shape of the readout pulse. The traversing particle produces an electromagnetic shower where the drift time of the particles in the shower causes a readout pulse that is roughly triangular in shape and typically 400 ns long. This pulse is shaped by the readout electronics and the signal shape is simulated with Monte Carlo and calibrated using precisely known test pulses deposited into the readout chain. However, as detailed below, the presence of multiple interactions per bunch crossing, known as pileup, has a significant effect in the calorimeters and is an important issue for understanding jet energy.

Hadronic Calorimeter

Like the EM calorimeter, the hadronic calorimeter was designed with an eye toward Higgs discovery, that it would have the resolution needed to find Higgs decays to $b\bar{b}$ and $\tau\tau$ pairs. This is important and complementary to the EM calorimeter because, although the most sensitive decay channels of the Higgs are to photons and electrons, the decay modes with the highest branching fractions ($b\bar{b}$ and a large subset of W^+W^- ³) are hadronic. For the physics in this thesis, with three b -jets in the final state, the hadronic calorimeter does the crucial jet energy detection and measurement.

Accurate measurement of hadronic energy is crucial for accurately reconstructing b -jets, and the hadronic calorimeter also measures the energy from hadrons, jets, taus

³in particular, decay modes where one W boson decays to a lepton and a neutrino, and the other decays to quarks

and allows for a measurement of missing transverse energy. The hadronic calorimeter is a sampling calorimeter, nicknamed the tile calorimeter because of its composition of scintillating tiles (active material) interleaved with steel plates (showering material) in the barrel. When a strongly interacting particle hits the steel it creates a spray of particles, called a shower, which then enters the scintillator and creates a light signal that is proportional to the energy. This process repeats until the full energy of the original particle is measured, and usually the calorimeter cells are reconstructed into an object called a jet which aims to cluster the energy together in a way that accurately represents the energy of the original particle. Much more on hadronic jet reconstruction will follow in Section 3.2.

The hadronic calorimeter is partitioned into four major subsections, two barrel sections and two extended barrel sections, allowing for measurements out to $|\eta| < 1.7$. In the forward regions, hadronic coverage is provided by the liquid argon system and the high-density forward calorimeter [32]. In order to stop all remaining particles from the collision, with the notable exception of muons, the hadronic calorimeter is about 7.4 interaction lengths thick.

2.4.3 Muon System

Muons are very similar to electrons in their interactions, but they are about 200 times heavier and as a result, they can travel largely unaffected through the inner detector and calorimeters and emerge in the outer layers of the detector, where the muon system is situated to make dedicated measurements of muons. The muon system is often used in conjunction with the tracking of the inner detector, since a muon interacts electromagnetically and would be expected to create a track in both systems. These tracks in the two different subsystems can then be combined into a single measurement of the p_T of the muon. Muons are largely not used in this analysis, so we will be brief in explaining the muon system, although they may

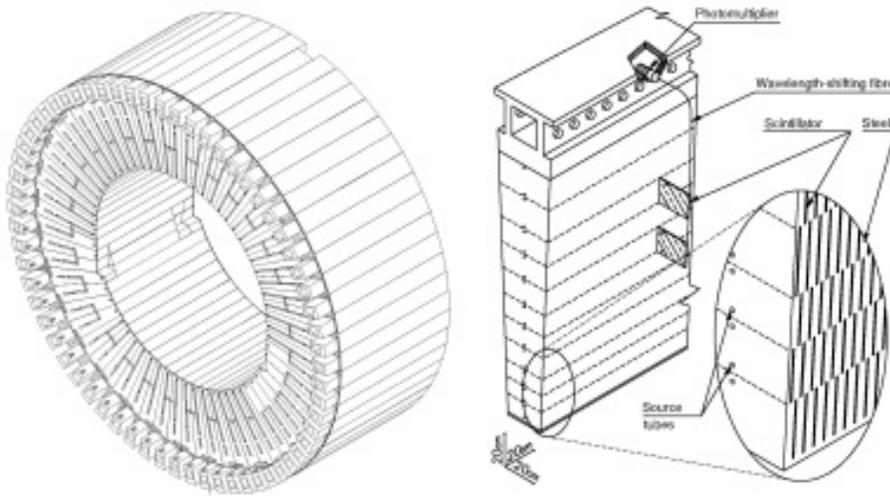


Figure 2.8: A depiction of the scintillator-and-steel hadronic calorimeter, also known as the tile calorimeter, and a close-up view of one of its cells.

play an important role in the future since muons can be used to help trigger on and identify b -jets.

The ATLAS muon system is composed of four different detector systems located within and around an air-core toroid magnet with a field of 1 Tesla. Precision tracking in the barrel is done by Monitored Drift Tubes (MDT) and in the endcap by Cathode Strip Chambers (CSC). Quick-readout triggering is done in the barrel by Resistive Plate Chambers (RPC) and in the endcap by Thin Gap Chambers (TGC). The system is designed to measure the p_T of muons with $p_T > 3$ GeV, with 3% resolution up to $p_T < 250$ GeV and 10% resolution up to 1 TeV.

2.5 The ATLAS Trigger and Data Processing

Although the LHC delivers 20 million bunch crossings per second to the ATLAS detector, the detector does not have the capacity in either storage space or readout

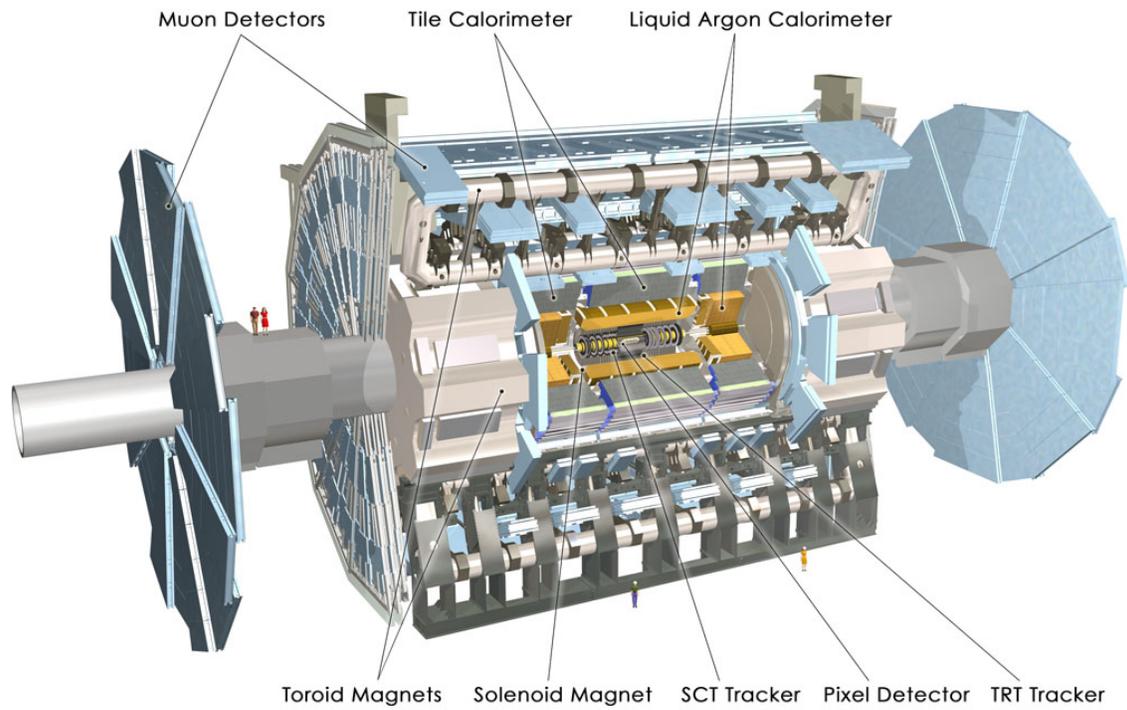


Figure 2.9: A diagram of the ATLAS detector as a whole, with major subsystems labeled. A couple of human figures are shown standing on and near the detector, for scale.

bandwidth to record all these collisions. The trigger has the task of selecting the most interesting 200 or so events per second, which are then fully reconstructed and recorded. The trigger is a three-layer system, with a first level (L1) implemented solely in hardware, a second level (L2) that reconstructs “regions of interest” (RoIs) with special fast algorithms, and an event filter (EF) that reconstructs the full event with offline algorithms. L2 and EF together are called the high level trigger, or HLT.

The trigger has a step-type structure, where progressively smaller numbers of events are processed with progressively more detailed and computationally intensive algorithms. Many physics analyses (including this one) look for signatures that have more than one physics object in them (for instance, multiple electrons, or a lepton plus jets), so multiple physics objects are often required in a single trigger. There are also single-item triggers, although typically these have higher p_T thresholds than triggers that look for multiple objects. If all of the objects in a given trigger item are seen, the event is accepted for the current level and the event moves forward in the data collection process. This process repeats three times, once for each of the levels of the trigger—only those events which pass L1 move onto L2, those which pass L2 move on to EF, and those that pass EF are written to disk. As the machine settings for the LHC changed over the course of 2012, resulting in different conditions for data-taking, the trigger menu was adjusted periodically to keep rates under control (in practice, this generally means either raising thresholds or adding prescales). The end result of this process is approximately 200 events per second (320 MB/s of data) being written to disk; the trigger rates and latency information can be found in Table 2.1.

Of particular interest to this analysis are the jet p_T thresholds and b -tagging applied in the trigger. This will be outlined in greater detail in Sections 3.5 and 5.1, but we will introduce the ideas here. As only 200 events per second get written to disk, the bandwidth has to be carefully allocated across triggers, and it is very expensive to

Trigger Level	Rate in Hz	Latency	Data Rate
None (Event rate)	20 MHz	no decision applied yet	1600 TB/s
L1	75 kHz	$\sim 1\mu s$	120 GB/s
L2	3 kHz	~ 10 ms	5 GB/s
EF	200 Hz	~ 1 s	320 MB/s

Table 2.1: The rates and latency of the three layers of the ATLAS trigger. Each layer applies increasingly stringent requirements while working with progressively more data, in order to diagnose whether an event is “interesting” enough to pass one of the trigger chains [2].

keep a trigger that allows high rates of events to be written to disk. In an analysis such as this one, the physics objects coming from the signal (b-jets from a Higgs decay) are generally higher in p_T than the background (continuum QCD processes) so one way to keep event rates reasonable in the face of rising luminosity is to place higher p_T thresholds on the jets that fire the trigger.

The thresholds increase with each trigger level, as more detector subsystems read out and more computationally intensive reconstruction algorithms are applied, so that a jet which is reconstructed with a minimum of 75 GeV of p_T at L1 might be required to have 145 GeV at EF in order to pass the full trigger chain. Once the event is written to disk, it is subject to the full offline reconstruction and calibration so that the final p_T of the jet might not be exactly the number measured at the trigger EF. There is therefore a range of p_T values, called the turn-on curve, where the trigger goes from rejecting all events to accepting all events. Within the turn-on curve even a small change in the p_T of a jet can have a dramatic difference in whether the jet fires a trigger accept. This instability is mitigated by placing p_T cuts on the trigger jets that require that their respective p_T values are above the turn-on curve. Details on the associated systematic errors in this analysis can be found in Section 9.2.

The ATLAS trigger also allows for b -tagging of jets at L2 and EF. For analyses that have b -jets in the final state, b -tagging in the trigger provides a tool for keeping rates low without pushing up jet p_T thresholds. More details are in Section 3.5.

Chapter 3

Event Reconstruction and Performance

Once the physics data has been created by the LHC, detected by the ATLAS detector, and written to disk by the ATLAS trigger and data acquisition system, it is reconstructed into physics events. This process has many interacting steps. First, hits in the inner detector are algorithmically combined into tracks, which approximate the trajectory of charged particles, while energy clusters in the calorimeters are grouped together into jets, which seek to capture the energy of the parton that originated the jet.

Higher-order quantities and corrections also come into play. Two of the most important are the calibration and removal of pileup, other lower-energy interactions from the same or adjacent bunch crossings, and b -tagging, which primarily uses tracks to identify jets that are likely originated by b -quarks. Additionally, there are jet energy corrections that account for jet radiation that could be lost by the jet clustering algorithms, b -tagging efficiency corrections, and uncertainties on the total luminosity collected.

3.1 Track Reconstruction

Track reconstruction uses measurements taken in the inner detector to perform crucial tasks such as reconstructing the locations of the primary and secondary vertices. When a charged particle traverses the layers of the inner detector, it leaves typically 8-11 hits in silicon (counting one double-sided layer of SCT silicon as capable of seeing 2 hits) and about 35 in the TRT. The track reconstruction algorithm starts by identifying track seeds and then iterating through a Kalman filter algorithm, projecting out to further tracking layers and then checking for hits along the hypothesized trajectory of the particle [33].

Track reconstruction performance can be impacted by pileup, as more pileup (and hence more tracking hits) creates more opportunities for incorrectly assigned hits and fake tracks from random hits being falsely associated with each other. A second effect of pileup on tracking is indirect: the detector as a whole, and the silicon detectors in particular, undergo radiation damage as luminosity accrues. Radiation damage in silicon detectors manifests itself as rising leakage current, which makes channels prone to more noise hits unless the thresholds are raised (but higher thresholds lower the efficiency for legitimate hits). As time goes on, the tracking quality can be affected by the extra noise hits and/or higher thresholds. The track reconstruction efficiency and resolution can be seen in Figures 3.1 and 3.2.

Good tracking performance enters this analysis most directly in its impact on the tagging of b -jets, as detailed in Section 3.5. Well-reconstructed tracks allow for the identification of a secondary vertex, which is one of the fundamental features of the ATLAS b -jet tagging algorithms.

Measurements from the inner detector (pixel, SCT and TRT) are combined in the track reconstruction. There is a broad range of properties that tracks might have, so the inner detector has to be able to measure tracks with p_T ranging from 150 MeV to 30 GeV or more.

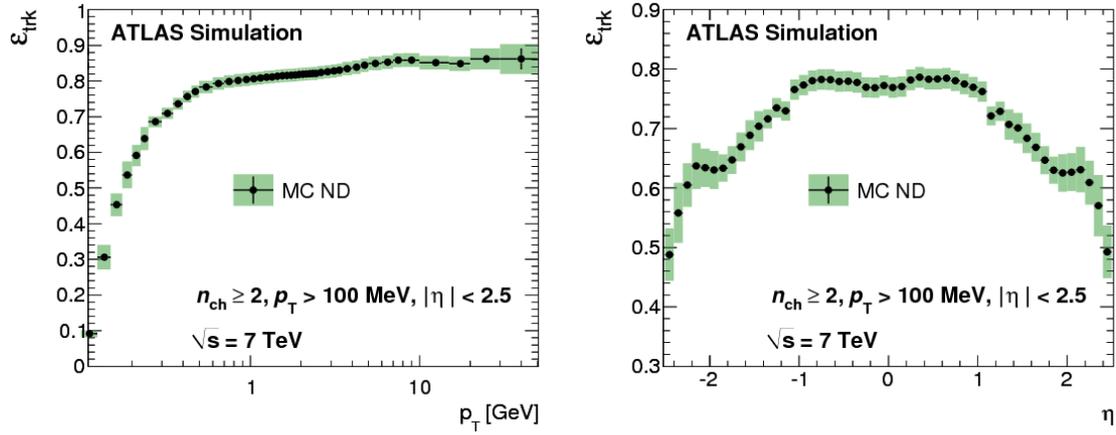


Figure 3.1: The efficiency of the track reconstruction, projected in p_T and η , as computed in Monte Carlo simulations.

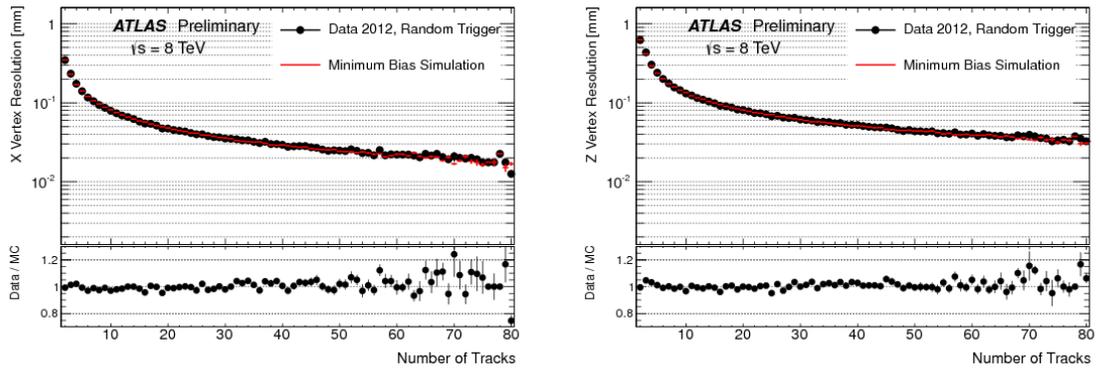


Figure 3.2: The track resolution in mm in the x and z directions, comparing minimum bias simulation to data taken with a random trigger.

3.2 Jet Reconstruction

Because of asymptotic freedom of QCD, quarks and gluons do not remain stand-alone particles once they are produced; first they undergo parton showering before hadronization, followed by hadronic showering when they reach the calorimeters. Hypothetically, all the particles of the hadronic shower can be added back up to approximate the energy, momentum, and angle of the original quark or gluon. Jet reconstruction is the process of assembling calorimeter deposits together into a physics object, called a jet, that ideally will do a good job of representing the characteristics (p_T , energy, flavor) of the quark or gluon that originated the jet. There are a number of clustering algorithms for assembling the calorimeter cells, and post-processing steps for improving the performance of jets in analyses—pileup subtraction, energy calibrations, grooming, and trimming, to name a few.

The default jet clustering algorithm in ATLAS is the anti- k_t algorithm [12] with a distance parameter of 0.4. Roughly summarized, this algorithm starts with a calorimeter cell that has an energy deposit at least 4σ ¹ higher in energy than the ambient and pileup noise, and then the surrounding cells with at least 2σ more energy than noise are grouped into the jet in a way that prioritizes high energy over close proximity. The result is that soft deposits get clustered in with hard deposits, rather than clustering amongst themselves. The distance parameter of 0.4 is roughly a cutoff as to how far away from the seed in $\eta - \phi$ space to look for additional deposits. Jets with p_T below 20 GeV or so can be difficult to distinguish from noise, so in practice, a lower limit on the p_T (in the case of this analysis, 25 GeV) is often applied. A visualization of the jet reconstruction that results from the anti- k_t algorithm can be seen in Figure 3.3.

¹in this context, σ is the spread in the noise in the calorimeter

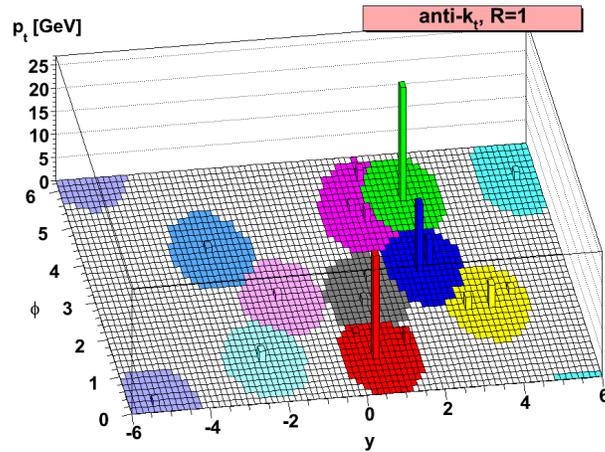


Figure 3.3: A graphical representation of the position and energy of calorimeter deposits, and their clustering and reconstruction into jets by the anti- k_t jet clustering algorithm. A feature of anti- k_t which is immediately apparent is the circular jet shapes that it tends to make [12].

3.3 Pileup Calibration and Removal

All the detector subsystems are affected by the presence of pileup, which are proton-proton collisions other than the hard scatter collision. As the LHC delivers higher luminosity for a given number of proton bunches, the luminosity increase comes at the price of many interactions per bunch crossing, and these softer interactions create extra activity in the detector that tends to make events noisier and more challenging to reconstruct accurately. In 2012, the mean number of interactions per crossing ranged from about 10 up to about 40.

The inner detector and tracking provide an important tool for understanding in-time pileup. In-time pileup is additional soft interactions in the same bunch crossing as the hard scatter. The tracking allows primary vertex reconstruction with a resolution fine enough in z_0 , for the pixels typically $z_0 \sin \theta$, to resolve separate primary vertices from each other [34]. The calorimeters cannot resolve individual primary vertices

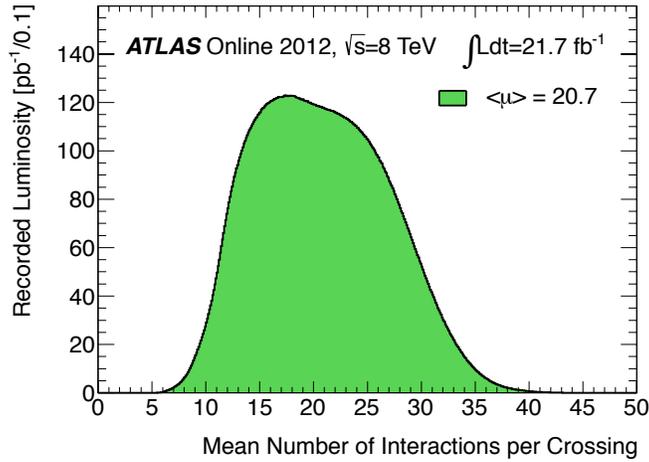


Figure 3.4: Distribution of the mean number of interactions per crossing, also called μ , seen in 2012 data-taking. Larger μ values are characterized by more challenging reconstruction.

with such precision, though, so a constant struggle in ATLAS is to measure the calorimeter deposits that come from pileup interactions, and where possible to apply corrections that subtract away pileup contributions to jets from the hard scatter. On average, each additional pileup vertex in an event adds 370(850) [35] MeV to the p_T of a jet reconstructed with the anti- k_t algorithm with $R=0.4(0.6)$.

In addition to in-time pileup, the calorimeters are prone to out-of-time pileup where the signal in a given event can be affected by the energy flow of previous collisions because of the calorimeter readout signal shapes. Out-of-time pileup has the effect of adding an average of 60 (210) MeV to central jets, and decreasing the forward jets' p_T by 350(470) MeV.

3.4 Primary Vertex Identification

Both the hard scatter collision and pileup collisions produce charged tracks; once these tracks are reconstructed, they can be traced back to the interaction region of the ATLAS detector and used to figure out where collisions took place. Each time a group of tracks can be clustered together into a presumed proton-proton collision, we refer to it as a primary vertex. The reconstruction of primary vertices is one of the reasons that tracking resolution must be so precise; primary vertices are often separated by only a few mm so imprecisely measured tracks can cause unintended merging. The consequence of mistakes in the primary vertex reconstruction can be physics objects that get grouped with the wrong primary vertex, and lead to incorrect event reconstruction and either signal loss or undesired background acceptance.

The reconstruction errors on the primary vertex are generally within $20 \mu\text{m}$ in the x/y directions (transverse to the beam direction) and $100 \mu\text{m}$ in the z direction (longitudinal to the beam) [34]. This resolution can be seen in Figure 3.5, a 2012 $Z \rightarrow \mu\mu$ event with 25 reconstructed primary vertices.

3.5 b -Tagging

Although the LHC produces huge numbers of hadronic jets, those jets that arise from b -quarks (which are called, appropriately, b -jets) are particularly important and interesting, especially for this thesis where all the final state particles are b -quarks and so the experimental signature consists entirely of b -jets. Once they are produced, b -quarks hadronize nearly instantaneously, but then the b -hadrons have lifetimes that are typically on the order of 1 ps^2 before decaying to other particles. Since b -quarks are often created in high- p_T collisions or come from the decay of heavy particles, the resulting b -hadrons and b -mesons can have considerable p_T and travel

²for a lifetime of 1 ps, $c\tau \approx 300\mu\text{m}$

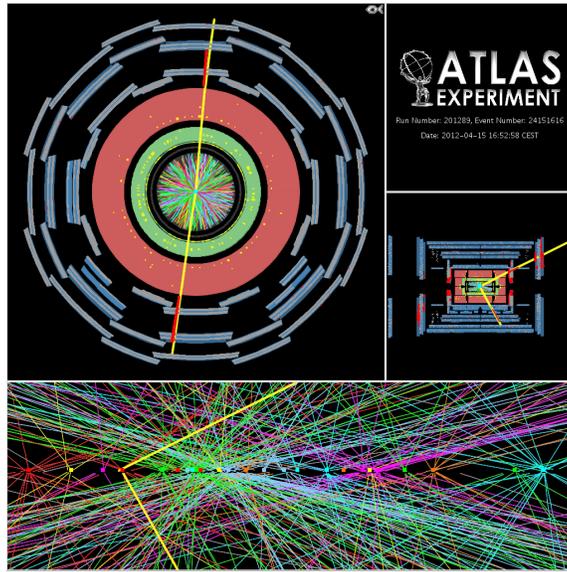


Figure 3.5: A data event taken in 2012, showing the occupancy typical of high pileup. This particular event contains a $Z \rightarrow \mu\mu$ physics signature, mixed in with 25 primary vertices.

a few millimeters to a few centimeters before decaying³. When they do decay, there is often a reconstructible secondary vertex that is offset from the primary vertex and can be used to identify, or tag, the jet as coming from a b -quark. ATLAS offers b -tagging at both the trigger level and offline; this analysis makes use of both.

3.5.1 Online b -Tagging

The first opportunity for b -tagging occurs in the trigger. The challenge of trigger (also called online) b -tagging is that it requires secondary vertex identification, which in turn requires track reconstruction, which requires the full readout of the inner detector and is also computationally expensive. The advantage, however, is that

³to reconcile the apparent contradiction between the b -hadron lifetime and these large decay lengths, recall that many of these b -hadrons are highly relativistic and undergo significant time dilation

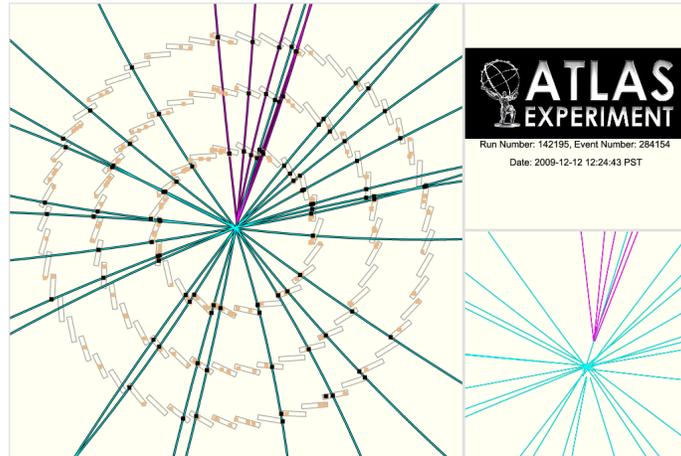


Figure 3.6: An event from early ATLAS data-taking, with a number of tracks (teal, in the zoomed-in view) leading back to the primary vertex, and a secondary vertex reconstructed with the associated tracks highlighted in purple [13].

b -jets occur at a much lower rate than light flavor jets in background, and are an important part of many search signatures, so being able to identify b -jets in the trigger allows for lower p_T thresholds than could otherwise be supported. The trigger in this analysis uses online b -tagging to keep thresholds low.

Like its offline counterpart, the online b -tagging algorithm has three major steps [36]:

- reconstructing the tracks of charged particles
- identifying the primary vertex
- calculating a discriminant weight (in 2012, a likelihood weight) based on the signed impact parameter significance (Figure 3.7) for the reconstructed tracks and the properties of secondary vertices

These steps are repeated both at L2 and EF⁴, so there are two independent layers of b -tagging applied by the trigger.

⁴for a full explication of the trigger structure, see section 2.5

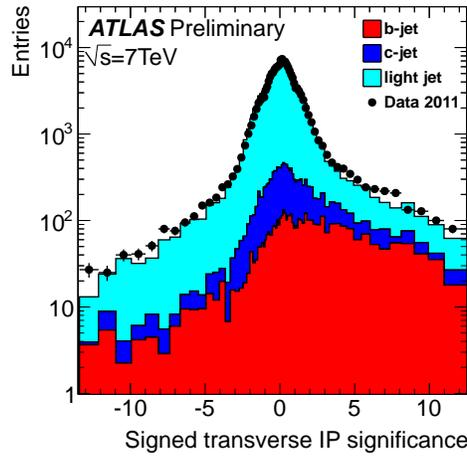


Figure 3.7: The online signed impact parameter significance, as used in the trigger b -tagging algorithms.

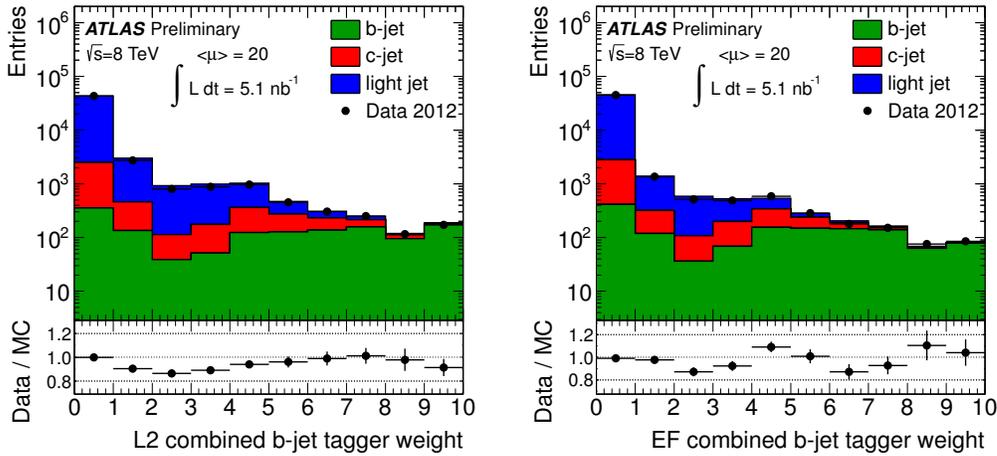


Figure 3.8: The jet weight distribution for jets undergoing online b -tagging at L2 (left) and EF (right), with a comparison between MC predictions of various truth flavors (colored histograms) and data (points). As the b -tagging weight increases, the relative enrichment of b -jets grows, although the overall efficiency for b -jets goes down as tighter online weights are applied. [14]

3.5.2 Offline b -Tagging

Offline b -tagging is very similar to online tagging, in that it uses reconstructed tracks to gather information about whether a secondary vertex might be present. There are several specialized algorithms that are each designed to exploit a specific feature of decays associated with b -jets:

- large significance of the impact parameter, d_0 (called IP3D within ATLAS [37])
- large ratio between the sum of p_T of all tracks associated with the secondary vertex to the sum of all tracks in the jet (called SV1 within ATLAS [37])
- large secondary vertex mass (SV1 [37])
- evidence of a decay chain including both a secondary vertex (from the b decay) and a tertiary vertex (from the c quark daughter of the b quark) along the same line of flight (called JetFitter within ATLAS[38])

Each of these features/algorithms can independently give discriminating power for tagging b -jets. Additionally, though, if a jet truly comes from a b -quark, then several of these features can arise in the same jet and the correlation can be used to further increase the accuracy of the tagger. That idea gives rise to the MV1⁵ tagger (where the MV stands for “multivariate”), a neural-net-based b -tagging algorithm that uses three other b -tagging algorithms (SV1, IP3D, JetFitter) as inputs. The performance curve for the MV1 algorithm can be seen in Figure 3.9; for a typical b -jet efficiency of 70% (meaning that 70% of real b -jets are positively tagged by the algorithm), the light-jet rejection is about 99.5% and the charm rejection is about 90% [15].

The importance of event reconstruction cannot be overstated. Simply put, it is the reconstruction step that takes a set of raw detector measurements and turns them into a comprehensive description of the physics going on in the event—the energy of

⁵MV1 is the internal ATLAS name for this b -tagging algorithm

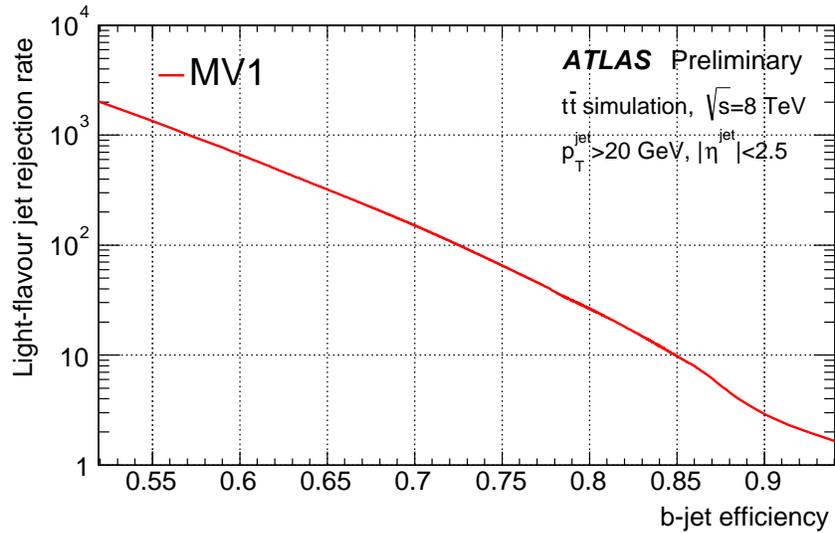


Figure 3.9: The receiver-operator curve (ROC curve) for the MV1 algorithm, showing the tradeoff between light-flavor rejection (vertical axis) and b -jet efficiency (horizontal axis) [15].

the hadronic and electromagnetic particles, the trajectories of charged particles, the identification of b -jets and other types of particle identification, and so on. However, in order to truly understand the signal and background processes, it usually takes studies on a deeper level than what we can get with purely data-driven methods; we need a “God’s eye” view of the collision and outgoing particles. We get this view by simulating physics events for ourselves using Monte Carlo methods, and using those simulated events to devise the analysis strategy. The next chapter will explain the details of that process.

Chapter 4

Monte Carlo Simulation Events

In order to understand both the signal and backgrounds better, Monte Carlo (MC) simulation datasets are computer-generated simulation events that allow a physicist to develop and validate an analysis. The generation and refinement of MC algorithms and datasets could be a thesis in its own right, but this section will touch on some of the most relevant features of the MC datasets used in this analysis.

4.1 MC Creation Procedure

MC events are generally created in four major steps.

1. First, the generator uses quantum field theory to simulate the hard scatter, generally starting with a proton and ending with all final-state particles.
2. Then those particles are handed off to a dedicated algorithm that simulates how they would shower and hadronize, where appropriate.
3. Next, the resulting particles from the first two steps are put into a detector

simulation, which uses information on the detector materials and geometry to understand how the particles evolve as they travel through the ATLAS detector.

4. The detector’s response to the particles is simulated in a digitization and reconstruction step, so that the final output is an event that has characteristics similar to those of a “real” event of that type in the ATLAS detector.

4.2 Signal Monte Carlo Simulations

MadGraph 5 [39] is used to generate the signal MC simulation events, with showering and hadronization done in Pythia6 [40]. The generation process uses a 5 flavor scheme PDF (parton distribution function), meaning that MadGraph models b -quarks in the proton as well as the more common light flavor (up, down, charm and strange quarks, and gluons). In addition to the Higgs production with the associated b -quark and the Higgs decay, there can also be extra jets in the event from initial state radiation (ISR) and/or final state radiation (FSR). We allow up to two additional partons per event in the signal MC simulations¹; the cross-section calculations tend to be more accurate for higher numbers of extra partons allowed (again, at the Feynman diagram level), but at the cost of exponentially slower generation times (Table 4.1). We found two additional partons to be a reasonable cutoff where the kinematic distributions and cross section calculations were only minimally affected by allowing for higher numbers of extra partons, but the generation time was still acceptably quick. The matching scale for knitting together the hard scatter calculations in MadGraph with the Pythia hadronization and showering was set to 20 GeV.

The Madgraph signal generation is done for 12 different mass points, spanning physical Higgs masses of 250-800 GeV. Below 250 GeV, the daughter b -jets from the Higgs

¹this statement refers to extra partons being generated at the level of the Feynman diagrams; the later hadronization and parton showering step can generate additional partons in the final state and events must be matched to the hard scatter diagrams to avoid double-counting

Table 4.1: The cross sections and generation times for MadGraph signal MC simulations as a function of number of additional partons allowed, for 0-3 partons. There is approx. 2.5% difference between the 2-parton and 3-parton cross sections, which comes at a cost of hours of CPU time per batch of 10,000 events. The times quoted here are for generating 10,000 events per sample, before hadronization. The generation process imposes a p_T cut of 5 GeV on all jets, no further p_T cut on b -jets, a minimum dR of 0.4 between jets (no minimum between b -jets), and a maximum $|\eta|$ of 5.0.

Process	$\sigma[pb]$	time
$pp \rightarrow h^0 b + \leq 0j$	$1.843e-5 \pm 4.2e-08$	seconds
$pp \rightarrow h^0 b + \leq 1j$	$2.555e-5 \pm 7.3e-08$	seconds
$pp \rightarrow h^0 b + \leq 2j$	$2.762e-5 \pm 8.2e-08$	minutes
$pp \rightarrow h^0 b + \leq 3j$	$2.83e-5 \pm 9e-08$	4 hours

Table 4.2: The signal MC samples and their parameters. These widths were generated with a $\tan\beta$ value of 30. The dataset ID is a reference for internal ATLAS bookkeeping.

Dataset ID	physical mass (GeV)	width (GeV)
181120	250	1.68682
181121	280	1.92318
181122	310	2.50849
181123	350	3.23507
181124	400	3.93561
181125	450	4.71906
181126	500	5.59454
181127	550	6.84368
181128	600	8.11044
181129	650	9.26688
181130	700	10.3760
181131	800	12.5049

tend to be too low- p_T to fire the trigger, while above 800 GeV it would require very high values of $\tan \beta$, above 60, to have a signal cross section large enough for experimental detection in this analysis, and perturbativity starts to break down for such large values of $\tan \beta$. Since both the production and decay of $bH \rightarrow b\bar{b}b$ happen via SM interactions, with SUSY only becoming relevant for increasing the production cross section and widening the intrinsic width of the A/H distributions, the signal generation can be done by using an SM $bH \rightarrow b\bar{b}b$ production model but with a modified $Hb\bar{b}$ coupling and Higgs width. There are 300,000 signal events generated for each mass point.

The detector simulation for the signal MC was done using AtlFast-II (AFII) [41], a modified version of Geant4 [42, 43] designed to cut down on the considerable time required for simulating particle interactions with the detector, especially showers in the calorimeter. AFII does full simulation of the inner detector and tracking, but uses frozen shower shapes taken from a large collection of pre-generated samples to speed up the simulation. After simulation, the digitization and reconstruction are handled by the ATLAS standard algorithms in the ATHENA framework, as is standard for all ATLAS MC.

4.2.1 Generator Comparisons

In addition to MadGraph, Sherpa and Pythia are some other signal MC generators that were studied. Pythia does not include Feynman diagrams that include extra partons in the final state, but instead gets events with additional partons from the showering that happens after the “base process”². Sherpa has the disadvantage that the Higgs is not stored in the truth table, which makes it impossible to positively disambiguate between Higgs daughter b -quarks and associated b -quarks. This is not

²Pythia has the additional disadvantage that Pythia6 was being phased out by ATLAS as this analysis was starting, and Pythia8 had not been validated for this signal process or any closely related ones

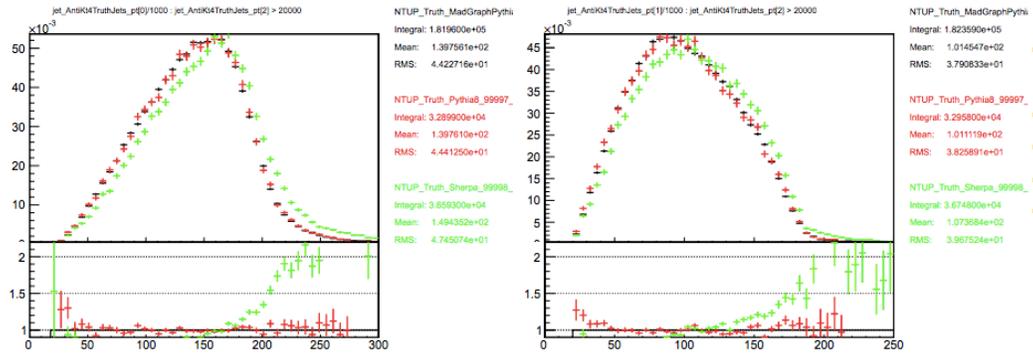


Figure 4.1: p_T distributions for the leading two jets for Pythia (red), Sherpa (green) and MadGraph (black) for a physical Higgs mass of 380 GeV. Pythia and MadGraph show good agreement, while Sherpa jets are systematically harder.

a bug per se, but rather reflects the fact that it can be indeterminate on a quantum mechanical level which b -quarks come from the Higgs decay vs. associated production. However, between the time that the signal generator studies were conducted and the writing of this thesis, a patch has been written that records the Higgs in the truth table, and might be deployed in future Sherpa releases. Sherpa also shows kinematic differences from both MadGraph and Pythia in the p_T of the truth jets (Figure 4.1), which of course then also propagates through to the $m_{b\bar{b}}$ distribution (Figure 4.2).

In order to understand if the disagreement between Sherpa and MadGraph/Pythia

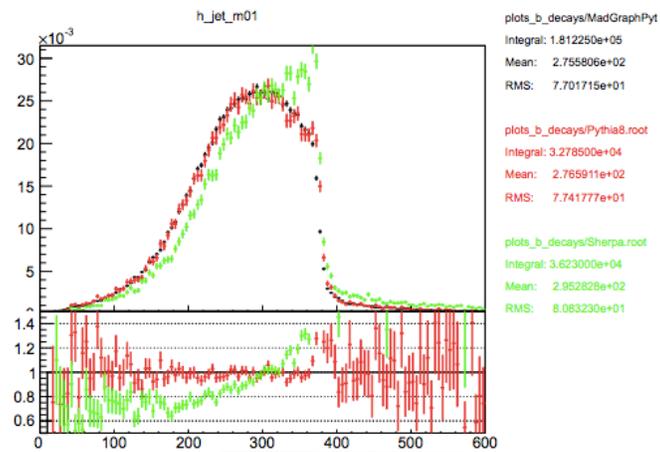


Figure 4.2: m_{bb} distributions for the leading two jets for Pythia (red), Sherpa (green) and MadGraph (black) for a physical Higgs mass of 380 GeV. Pythia and MadGraph show good agreement, while the Sherpa m_{bb} distribution is higher and more sharply peaked.

in the m_{bb} distribution was caused, at least in part, by mixing up the associated b -jets and the Higgs daughter jets, we changed the Sherpa generation process to $bH \rightarrow b\tau\tau$ and examined the properties of the Higgs itself in addition to the jets (or τ leptons) in the event. The p_T and η distributions of the di-tau (i.e. Higgs) system are in Figures 4.3 and 4.4, and the $m_{\tau\tau}$ distribution is in Figure 4.5.

Upon more careful examination, we found that Sherpa would fail with a segmentation fault when trying to produce $gg \rightarrow H$ Feynman diagrams, and so those diagrams had been excluded from the initial production, although an exchange with the Sherpa authors³ provided a patch that enabled $gg \rightarrow H$ diagrams. To understand if the absence of $gg \rightarrow H$ diagrams in Sherpa was responsible for the disagreement with MadGraph, we generated samples of MadGraph events with and without the $gg \rightarrow H$ diagrams and compared the p_T and η of the Higgs system (Figure 4.6). The Higgs p_T and η do not show a bias depending on whether the gluon diagrams are included or not, so it is not clear that including the gluon diagrams in Sherpa would resolve the disagreement that we see.

As a result of these studies, we conclude that the likely cause of the Sherpa disagreement is ambiguity in the selection of the Higgs daughter b -jets, since the reconstructed Higgs mass agreement between Sherpa and MadGraph improves substantially when looking at Higgs events decaying to $\tau\tau$ (where the combinatorics are trivial). MadGraph is also the generator of choice for the H/A searches in the $\tau\tau$ decay channel, which makes combinations or comparisons of those searches more straightforward than it would be if they use different generators.

³most notably Stefan Höche and Steffen Schumann, who we thank for their helpful advice throughout this analysis

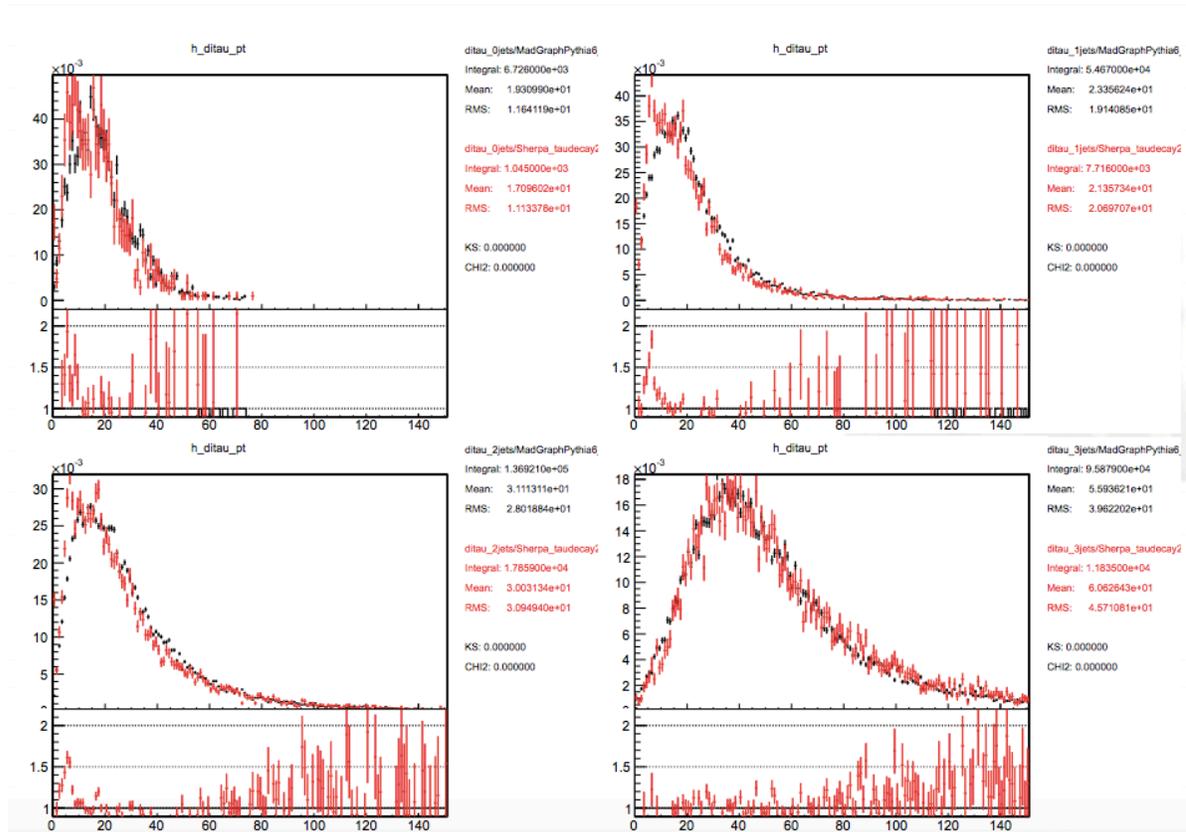


Figure 4.3: p_T distributions for the two τ leptons (i.e. the Higgs p_T) for Sherpa (red) and Madgraph (black) for zero (upper left), one (upper right), two (lower left), and three (lower right) additional jets per event.

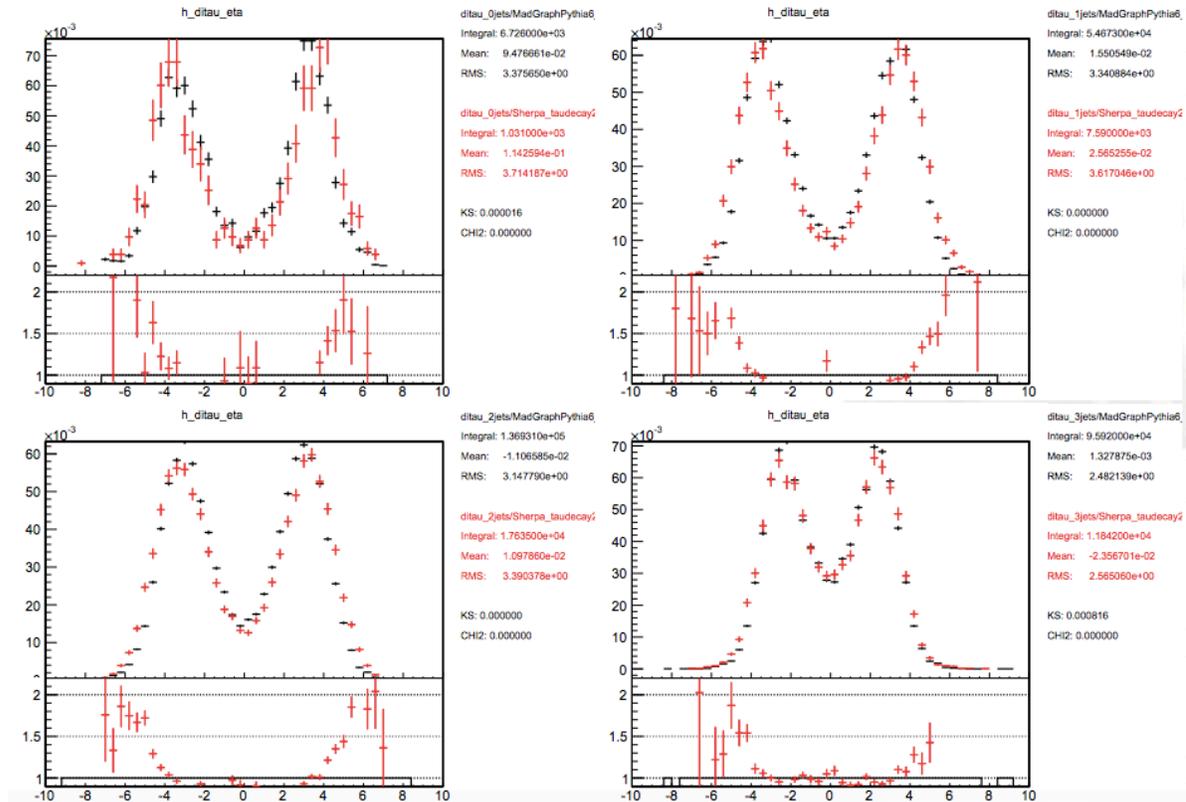


Figure 4.4: η distributions for the 2 τ leptons (i.e. the Higgs η) for Sherpa (red) and Madgraph (black) for zero (upper left), one (upper right), two (lower left), and three (lower right) additional jets per event. The Higgs is more forward (larger $|\eta|$) in Sherpa than in MadGraph.

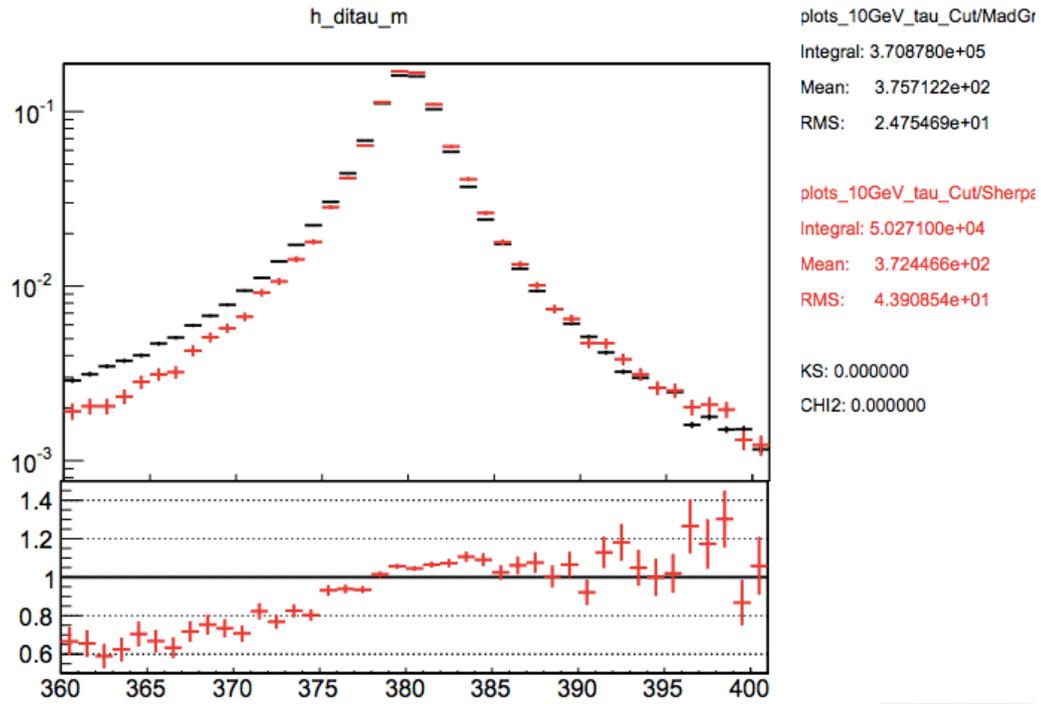


Figure 4.5: $m_{\tau\tau}$ distributions for the two τ leptons for Sherpa (red) and Madgraph (black) for a 380 GeV Higgs particle. Although the agreement is not perfect in the tails of the distribution, the agreement is much better than in the m_{bb} distributions, suggesting that the challenges of correctly identifying the Higgs daughters in Sherpa is a major factor.

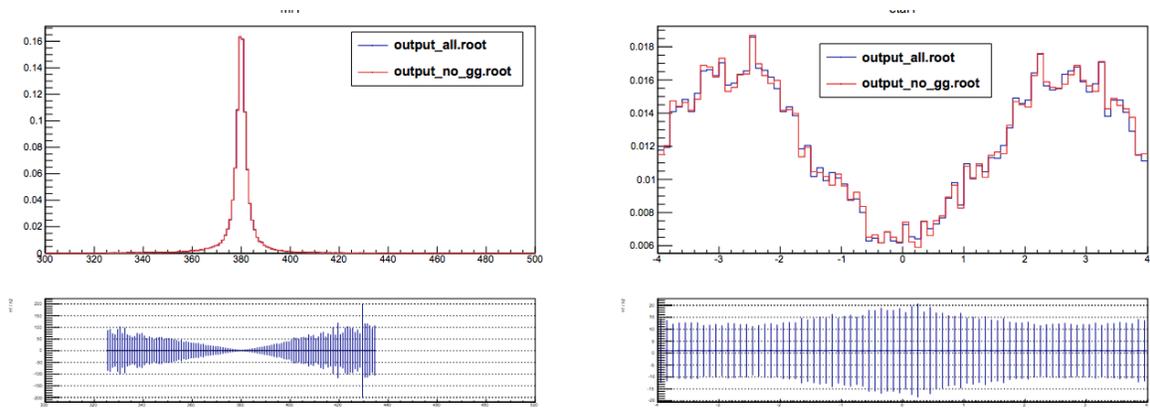


Figure 4.6: The $m_{\tau\tau}$ ($m_A=380$ GeV) distribution (left) and η of the τ leptons (right) for $bH \rightarrow b\tau\tau$ events in Sherpa, with and without gluon fusion diagrams being included. The ratio plots in the lower part of the figures show that the inclusion of the gluon fusion diagrams does not change the kinematics of the Higgs and its daughters.

4.3 Background

4.3.1 QCD Background

Although QCD is the largest and most important background in this analysis, fully modeling it with MC simulation has some important drawbacks, which motivates our decision to use a mostly data-driven background estimation method. That having been said, MC can still be a very valuable tool for making basic estimates and validating assumptions.

There are two major types of QCD background in this analysis: first, when there is one or more mistakenly b -tagged light flavor or charm jets, which we call *reducible* because, at least in theory, it could be identified and isolated/removed; and second, the *irreducible* QCD background in which three real b -jets are present in the final state but without the intermediate resonance of the Higgs. Both sources of background are expected to be significant but generally require different MC generation strategies.

QCD Multijet

The ATLAS QCD multijet MC dataset is generated using Pythia8 [44]. One of the major challenges of a truly inclusive sample like this one is that low- p_T jets dominate the production cross section but generally have very low efficiency through the triggers or analysis cut flows, which is addressed by generating events several times with filtering applied based on the p_T of the leading jet. This leads to a distinctive “slice” structure to the sample, where 8 different slices are generated, each with a different p_T range for the leading jet, and then the slices are knitted together with different relative weights to produce an inclusive spectrum with high statistics at all p_T values.

Table 4.3: Hard subprocesses simulated in the inclusive QCD MC event sample, along with their cross sections. Here $j = u, \bar{u}, d, \bar{d}, s, \bar{s}, c, \bar{c}, g$. The QCD slices are defined by the p_T cut placed on the leading jet in the event, which is documented in the “ p_T range” column. The dataset ID is noted as a reference for internal ATLAS bookkeeping.

Index	Dataset ID	p_T Range	Cross Section (nb)	Filter Eff	N Events
0	147910	0-20	7.285×10^7	0.98557	1000000
1	147911	20-80	7.285×10^7	0.00012909	999996
2	147912	80-200	2.6359×10^4	0.0039901	999893
3	147913	200-500	5.442×10^2	0.001222	914584
4	147914	500-1000	6.4435	0.00070839	999773
5	147915	1000-1500	3.9739×10^{-2}	0.0021516	957565
6	147916	1500-2000	4.161×10^{-4}	0.0046773	299929
7	147917	2000+	4.0636×10^{-5}	0.014595	299988

ATLAS has a high-statistics inclusive QCD MC sample that is used primarily in this analysis for understanding the QCD background from mistagged charm and light flavor (Table 4.3). Since the sample is inclusive, there is no filtering on the flavor of the jets that are produced (although there is dedicated effort to have high- p_T jets simulated with adequate statistics) and the vast majority of jets are light or charm jets. As a result, while we use this sample to estimate the efficiency and flavor composition of the QCD production in ATLAS as a whole, a particularly important background (QCD $b\bar{b}b$) is virtually absent from this sample.

b -Enriched QCD Multijet

Since the inclusive QCD Multijet samples are inadequate for understanding the $b\bar{b}$ and $b\bar{b}b$ QCD backgrounds, we generate a dedicated sample that undergoes filtering to enrich it in heavy flavor. This sample is generated using Sherpa 1.4.3 [45] and detector simulation is done with AFII. This sample starts with a 5-flavor PDF that assumes massive b -quarks and 2, 3 or 4 final state partons; then a filter is applied that requires that the leading 2 partons in the event be true b -quarks, as well as

requiring that the leading parton have a p_T of at least 10 GeV and the leading (subleading) jet have a p_T of 130 (45) GeV; additionally, the b -hadrons and jets must have $|\eta| < 2.9$ ⁴. We further enhance the $b\bar{b}$ contribution to the generation process by setting a flag (called `enhance_factor` in the Sherpa job options) that enhances $g \rightarrow b\bar{b}$ splitting.

A scale factor of 2.07 was applied to the MC to reach the same normalization as the full 2012 data (after the trigger and cuts are applied). This scale factor is based on the calculated total cross section and filter efficiency of the $b\bar{b}$ QCD MC (Table 4.5), as well as the number of MC events that were generated and the 2012 luminosity. In addition, the MC has pileup reweighting and b -tagging scale factors applied.

As is detailed much more thoroughly in Sections 6.1 and 7.3.5, the final analysis is categorized by the b -tag status of the third jet in each event, and the number of jets in each event. The b -tag categorization splits the sample based on whether the third jet in the event has a third tight-tagged b -jet (called the *bbb* category), a third jet that passes a loose b -tag but not a tight one (called *bbloose*) or there is no third b -tagged jet present (*bbanti*). Events are also categorized based on whether they have three, four, or five or more jets.

First we check the relative normalization of the nine categories. The number of weighted MC (data) events in each of the jet and tag categories can be seen in Table 4.4. The MC generally underestimates the number of data events in each category, with the notable exception of the *bbanti* 3-jet category, where the data has fewer events than the MC predicts. The discrepancy is about 3.2%, but as this category has the highest statistics of any category, that is still a difference of about 22,000 events.

The kinematics of the leading three jets are compared between 19.5 fb^{-1} of 8 TeV ATLAS data and $b\bar{b}$ QCD MC simulations (scaled to the same luminosity as data) in

⁴These kinematic requirements help keep an acceptable efficiency when the sample is passed through the trigger and offline cuts

Table 4.4: In addition to examining the shapes of the kinematic distributions in data vs. in bb QCD MC simulation events, the relative normalization of the different categories is computed and compiled. The data and bb QCD MC events are passed through the analysis framework, applying the trigger and cuts, and then a scale factor of 2.07 is applied to the bb QCD MC simulations to bring the overall number of events to the same as the number of events in data. Then the bb QCD MC (data) events are categorized by the tag status and number of jets in the event.

	3 jets	4 jets	5+ jets
<i>bbb</i>	77,288 (92,853)	70,289 (74,060)	52,998 (54,753)
<i>bbloose</i>	59,189 (56,826)	46,521 (49,502)	30,491 (34,725)
<i>bbanti</i>	693,518 (671,412)	276,593 (285,076)	122,354 (128,712)

Figures 4.7- 4.12. The kinematics are generally in good (but not perfect) agreement, with the notable exception of the p_T distribution of the second leading jet in the *bbb* 3-jet category. We also compare the m_{bb} distributions for background events where we examine the truth flavor of the third jet in the event, to see if the irreducible *BBB* background⁵ has a different m_{bb} shape than the reducible *BBC* and/or *BBL*. There appears to be no dependence of the m_{bb} shape on the truth flavor or *b*-tag status of the third jet, which is critical to the analysis strategy of using the *bbloose* and *bbanti* m_{bb} distributions in data to model the shape of the *bbb* m_{bb} distribution.

⁵where the capital letters here indicate truth flavor of the jets, rather than whether they are *b*-tagged or not

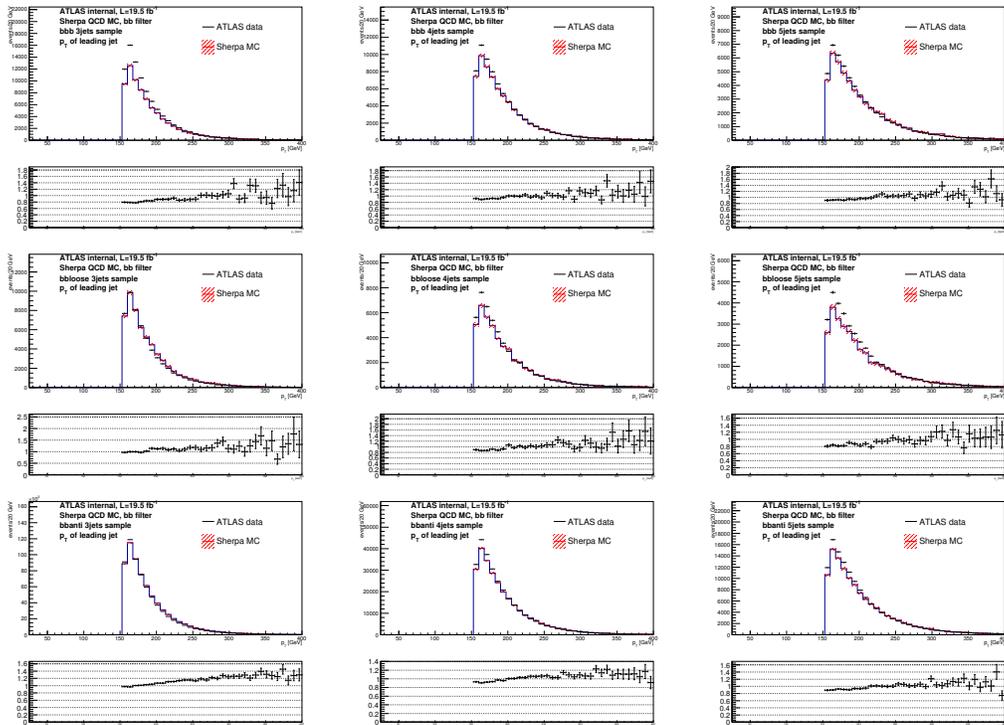


Figure 4.7: The p_T distributions for the leading jet, comparing bb QCD MC events to data. The MC has been normalized to the same total number of events as the data (over the entire sample, not on a category-by-category basis) and the MC/data ratio is plotted in the lower subplots. The errors on the MC are statistical only.

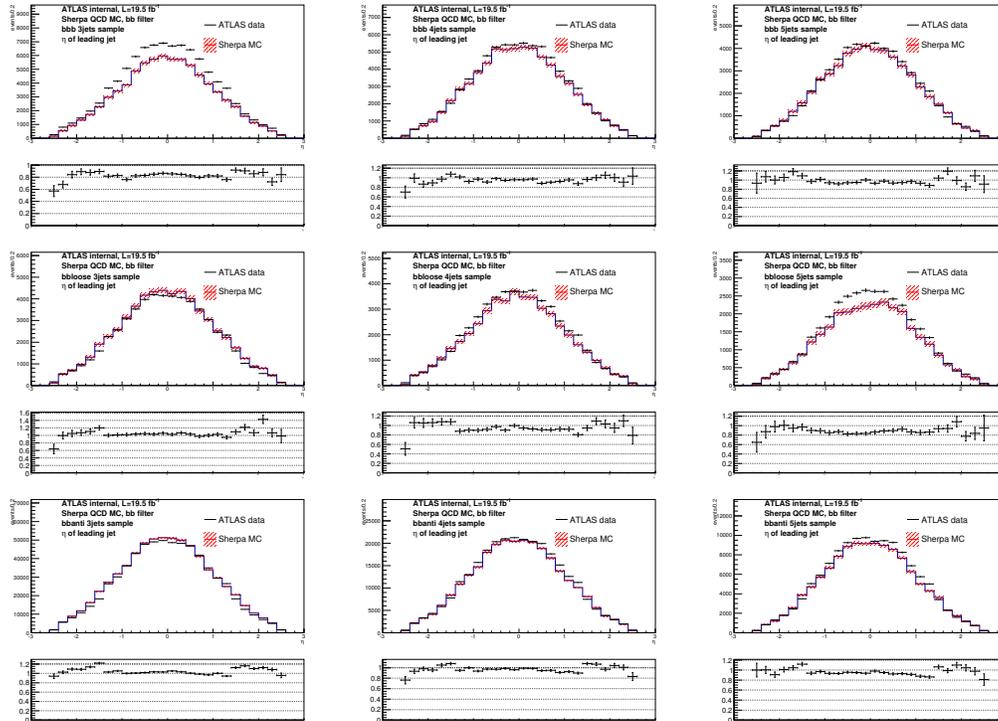


Figure 4.8: The η distributions for the leading jet, comparing bb QCD MC events to data. The MC has been normalized to the same total number of events as the data (over the entire sample, not on a category-by-category basis) and the MC/data ratio is plotted in the lower subplots. The errors on the MC are statistical only.

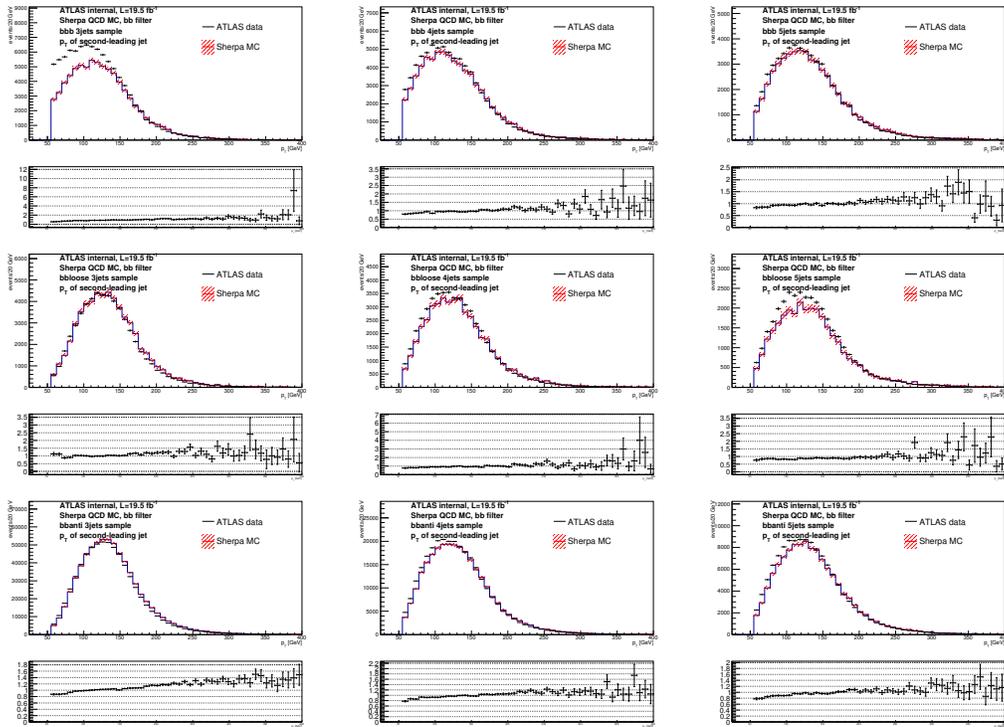


Figure 4.9: The p_T distributions for the second-leading jet, comparing bb QCD MC events to data. The MC has been normalized to the same total number of events as the data (over the entire sample, not on a category-by-category basis) and the MC/data ratio is plotted in the lower subplots. The errors on the MC are statistical only.

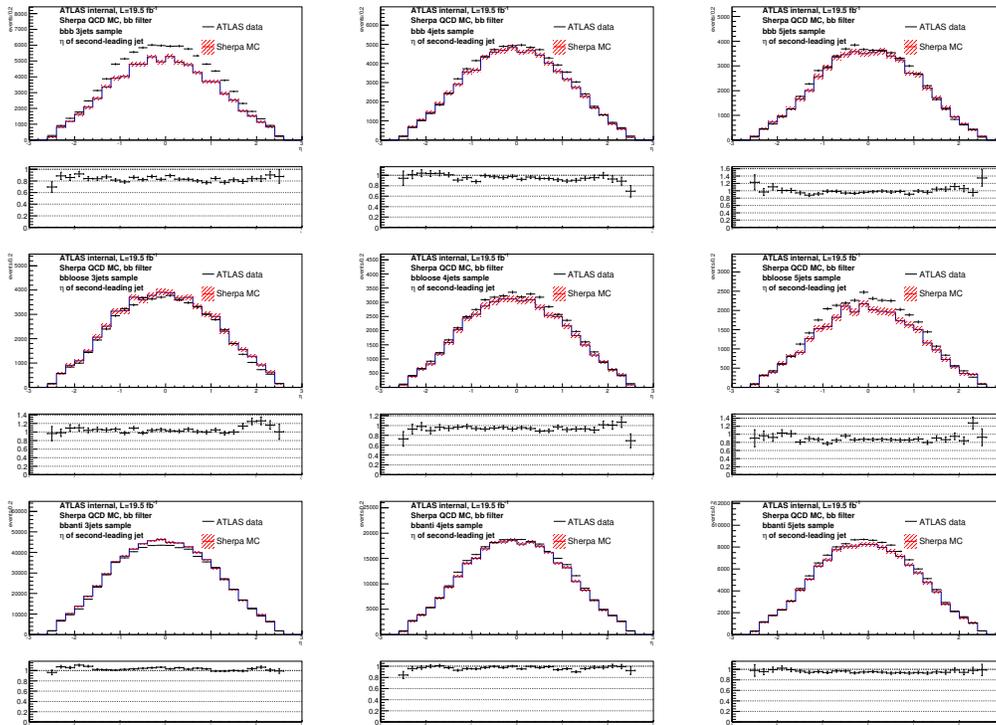


Figure 4.10: The η distributions for the second-leading jet, comparing bb QCD MC events to data. The MC has been normalized to the same total number of events as the data (over the entire sample, not on a category-by-category basis) and the MC/data ratio is plotted in the lower subplots. The errors on the MC are statistical only.

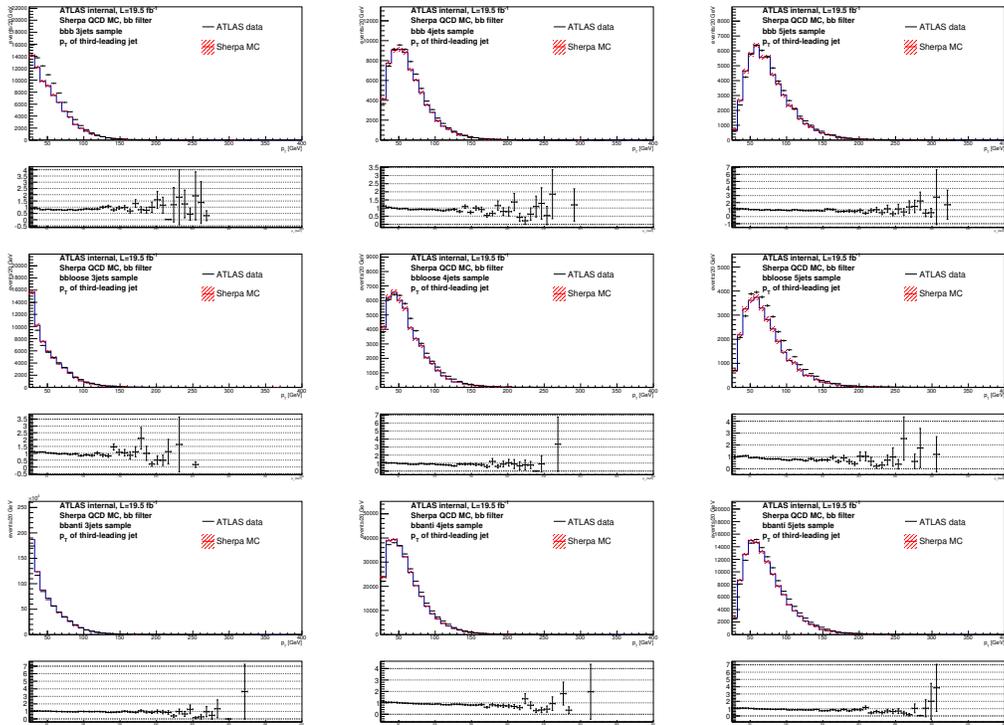


Figure 4.11: The p_T distributions for the third-leading jet, comparing bb QCD MC events to data. The MC has been normalized to the same total number of events as the data (over the entire sample, not on a category-by-category basis) and the MC/data ratio is plotted in the lower subplots. The errors on the MC are statistical only.

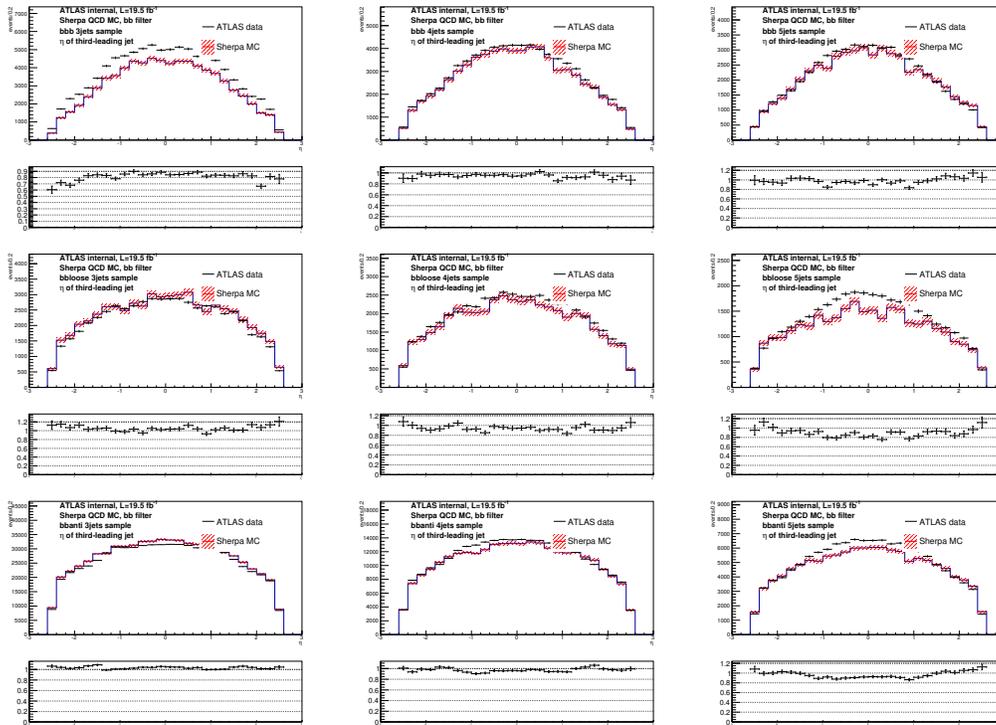


Figure 4.12: The η distributions for the third-leading jet, comparing bb QCD MC events to data. The MC has been normalized to the same total number of events as the data (over the entire sample, not on a category-by-category basis) and the MC/data ratio is plotted in the lower subplots. The errors on the MC are statistical only.

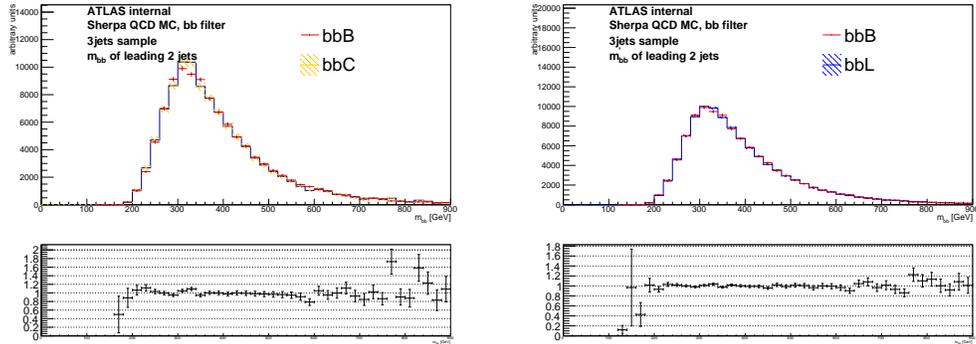


Figure 4.13: The m_{bb} distributions in the 3 jet category, where the third jet in the event has been truth-matched as bottom versus charm (left) or light (right). The distributions have been normalized to the same integral, and the lower subplots show the ratio of charm (light) to bottom. Errors on the MC are statistical only. There is no evidence from these distributions that the m_{bb} of the leading 2 jets in a QCD event is affected by the truth flavor of the third jet.

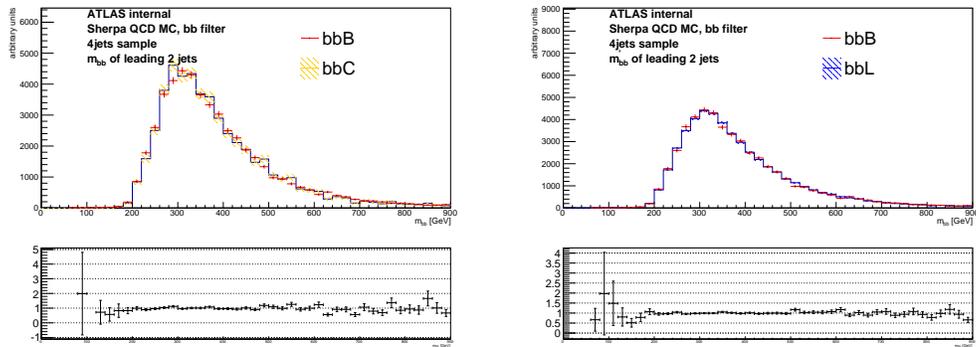


Figure 4.14: The m_{bb} distributions in the 4 jet category, where the third jet in the event has been truth-matched as bottom versus charm (left) or light (right). The distributions have been normalized to the same integral, and the lower subplots show the ratio of charm (light) to bottom. Errors on the MC are statistical only. There is no evidence from these distributions that the m_{bb} of the leading 2 jets in a QCD event is affected by the truth flavor of the third jet.

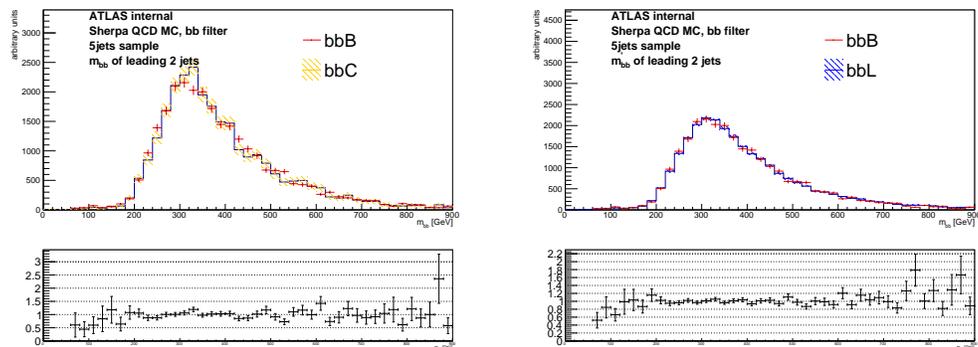


Figure 4.15: The m_{bb} distributions in the 5 or more jet category, where the third jet in the event has been truth-matched as bottom versus charm (left) or light (right). The distributions have been normalized to the same integral, and the lower subplots show the ratio of charm (light) to bottom. Errors on the MC are statistical only. There is no evidence from these distributions that the m_{bb} of the leading 2 jets in a QCD event is affected by the truth flavor of the third jet.

Table 4.5: Hard subprocesses simulated in the bb QCD MC event sample, along with their cross sections. Here $j = u, \bar{u}, d, \bar{d}, s, \bar{s}, c, \bar{c}, g$

. Kinematic cuts on these samples are detailed in Section 4.3.1.

subprocess	cross section (pb)
$jj \rightarrow b\bar{b}jj$	58531
$jj \rightarrow b\bar{b}j$	33411
$bj \rightarrow bj\bar{j} + \bar{b}j \rightarrow \bar{b}j\bar{j}$	22147
$bj \rightarrow bj\bar{j}j + \bar{b}j \rightarrow \bar{b}j\bar{j}j$	16282
$bj \rightarrow bj + \bar{b}j \rightarrow \bar{b}j$	12135
$jj \rightarrow b\bar{b}$	1672
$cj \rightarrow cb\bar{b}j + \bar{c}j \rightarrow \bar{c}b\bar{b}j$	1602
$bj \rightarrow bb\bar{b}j + \bar{b}j \rightarrow \bar{b}b\bar{b}j$	997
$jj \rightarrow b\bar{b}c\bar{c}$	776
$cj \rightarrow cb\bar{b} + \bar{c}j \rightarrow \bar{c}b\bar{b}$	681
$bj \rightarrow bb\bar{b} + \bar{b}j \rightarrow \bar{b}b\bar{b}$	387
$jj \rightarrow b\bar{b}b\bar{b}$	376
$b\bar{c} \rightarrow b\bar{c}j + \bar{b}c \rightarrow \bar{b}cj$	206
$bc \rightarrow bcj + \bar{b}\bar{c} \rightarrow \bar{b}\bar{c}j$	194
$b\bar{c} \rightarrow b\bar{c}j\bar{j} + \bar{b}c \rightarrow \bar{b}cj\bar{j}$	143
$bc \rightarrow bcj\bar{j} + \bar{b}\bar{c} \rightarrow \bar{b}\bar{c}j\bar{j}$	136
$bc \rightarrow bc + \bar{b}\bar{c} \rightarrow \bar{b}\bar{c}$	122
$b\bar{c} \rightarrow b\bar{c} + \bar{b}c \rightarrow \bar{b}c$	121
$b\bar{b} \rightarrow b\bar{b}j$	62
$bb \rightarrow bbj + \bar{b}\bar{b} \rightarrow \bar{b}\bar{b}j$	53
$b\bar{b} \rightarrow b\bar{b}j\bar{j}$	44
$bb \rightarrow bbj\bar{j} + \bar{b}\bar{b} \rightarrow \bar{b}\bar{b}j\bar{j}$	39
$b\bar{b} \rightarrow b\bar{b}$	37
$bb \rightarrow bb + \bar{b}\bar{b} \rightarrow \bar{b}\bar{b}$	30

Chapter 5

Trigger and Cuts

The very first steps of any analysis consist of gathering the data, and then reducing the background as much as possible. For an experiment like ATLAS, with a dataset that approaches 3 PB of raw data collected per year, picking a subset of that data for performing the search is both nontrivial and crucial to the overall success of an analysis.

The data selection starts with the trigger, which determines which events will even be recorded in the first place. The trigger must balance the competing forces of signal processes that can be very rare (which would suggest a permissive trigger, so as to not lose already-rare signal events) and the huge background rates at the LHC (which would suggest a strict trigger). Since the trigger must make decisions about event acceptance in near real-time, it can generally only require basic physics objects like jets and b -tags, and more sophisticated background rejection gets implemented as offline event selection criteria, also called cuts.

5.1 Trigger

The purpose and structure of the ATLAS trigger are explained in Section 2.5. Most importantly for this section, recall that the ATLAS trigger has three levels, called L1, L2, and EF. The trigger chain for this analysis is as follows:

- L1: At least one J75 RoI¹, implying an L1 jet with $|\eta| < 2.8$ and a p_T of at least 75 GeV
- L2: At least two 30 GeV jets, of which one must be at least 140 GeV. Additionally, at least two jets must satisfy a medium b -tag (L2 xComb² > 1.276)
- EF: At least two 35 GeV jets, of which one must be at least 145 GeV. Additionally, at least two jets must satisfy a medium b -tag (EF xComb³ > 1.099). There is no explicit requirement that the b -tagged jets from L2 correspond with the b -tagged jets from EF. However, as detailed later, this requirement is added later as an offline cut.

In the trigger, the L2 and EF jets are b -tagged using xComb, a likelihood ratio of IP3D⁴ (significance of z_0 and d_0 impact parameters), SV1 (mass of the secondary vertex), NVTX (the number of vertices with two tracks), and EVTX (the energy fraction of the secondary vertex), as well as the number of vertices with 2 tracks. The online b -tagging cuts used in the trigger for this analysis operate at the so-called “loose” working point, where the b -jet efficiency is 70%.

¹RoI stands for region of interest, an ATLAS acronym for a physical area in the calorimeter which has a large number of adjacent firing calorimeter cells, which collectively serve as a jumping-off point for a jet clustering algorithm

²xComb is the name of the b -tagging algorithm that is used in the trigger

³xComb is used in both L2 and EF of the trigger; the values of the inputs to the algorithm can change between L2 and EF as more computationally intensive reconstruction techniques are used in EF compared to L2

⁴IP3D, like the other algorithms described here, uses statistical learning algorithms trained on a large sample of Monte Carlo simulation events to predict whether a given jet is a b -jet or not. The various b -taggers vary in the algorithm and input features, which are very briefly summarized in this list.

Table 5.1 below shows the correlations between the two levels of online b -tagging, as well as the offline (MV1) b -tagging for three offline working points (60%, 70%, 80%).

Table 5.1: The acceptance of a cut on variable X given that the events have already passed (sig-like) or failed (bkgd-like) a cut on variable Y.

		Signal MC		Unbiased Data	
X	Y	sig-like	bkgd-like	sig-like	bkgd-like
EF	L2	0.917	0.248	0.424	0.035
L2	MV1 (80)	0.675	0.031	0.510	0.077
EF	MV1 (80)	0.824	0.090	0.515	0.050
L2 and EF	MV1 (80)	0.676	0.036	0.414	0.031
MV1 (80)	L2 and EF	0.977	0.434	0.640	0.076
L2	MV1 (70)	0.716	0.060	0.652	0.086
EF	MV1 (70)	0.867	0.131	0.665	0.060
L2 and EF	MV1 (70)	0.721	0.060	0.563	0.038
MV1 (70)	L2 and EF	0.954	0.340	0.538	0.034
L2	MV1 (60)	0.757	0.101	0.762	0.095
EF	MV1 (60)	0.903	0.193	0.779	0.070
L2 and EF	MV1 (60)	0.765	0.103	0.692	0.045
MV1 (60)	L2 and EF	0.902	0.250	0.429	0.015

5.2 Offline Cuts

Once an event has fulfilled the trigger requirements and been written to disk, it must pass a series of offline cuts before being used in the final fit and search.

- **p_T cuts of 155 and 55 GeV on leading and sub-leading jets, respectively.** In order to get out of the p_T turn-on curves of the trigger, cuts of 155 GeV and 55 GeV are applied to the leading and second jets in the event.
- **p_T cut of 25 GeV on third jet.** A third jet is not required by the trigger, but is required by the offline cuts, and it must have a p_T greater than 25 GeV.

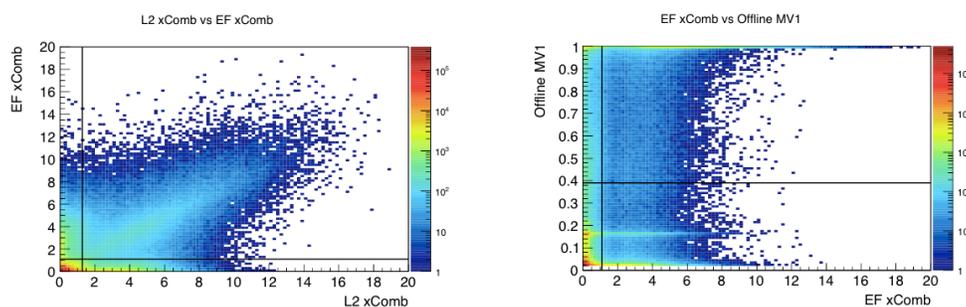


Figure 5.1: The b -tagging in Level 2 and the Event Filter (in the trigger) and the MV1 b -tagging algorithm (offline) are exploited in the cut flow by requiring that at least 2 jets in the event be “triple tagged”, or b -tagged in both stages of the triggers as well as offline. The figure on the left shows the L2 and EF b -tag distributions, on the right are the EF and MV1 distributions. The cut points are noted with vertical and horizontal lines. Most of the events (more than 80%) fall in regions where the b -tagging cuts agree on whether a jet is a b -jet or not, but a significant subsample of jets are tagged by one b -tagging level/algorithm but not the other(s).

There is no veto on additional jets. Further details on the jet cuts can be found in Section 3.2.

- **η cuts on the leading 3 jets** The leading three jets must be central, with $|\eta| < 2.5$. This is strongly motivated by the b -tagging requirements downstream; the inner tracker only provides precision tracking out to $|\eta| < 2.5$, which propagates through to where b -tags can be computed and applied.
- **Two jets passing online (L2 and EF in trigger) and offline (MV1) b -tags** As part of the trigger, two jets are required to be b -tagged at L2 and at EF. It is not explicitly required by the trigger that the same jets pass b -tagging at the two trigger levels, or that those jets be tagged offline by the MV1 algorithm, so we make an offline cut that imposes that requirement. There is further discussion of the online and offline b -tagging correlations in Section 3.5.1, where one of the important conclusions is that there is not perfect correspondence between jets passing the three levels of b -tagging (L2, EF, MV1). For example, a jet with a large xComb value at L2 (and passing the b -tagging requirement at L2) might have a small xComb value at EF (and fail the b -tagging requirement at EF). Likewise with L2 and MV1, or EF and MV1. As a result, we make an explicit requirement offline that the same two jets be b -tagged in L2, EF and MV1. Jets that are tagged in both L2 and EF of the trigger are referred to as “trigger-tagged” jets, and jets that pass L2, EF and MV1 are called “triple-tagged” jets. The offline cut on the trigger-matched b -jets is set to the 60% efficiency operating point. Due to the bias introduced by the trigger (see Table 5.1), this provides additional rejection of light jets for a relatively small decrease in signal efficiency.
- **Leading 2 jets must pass tight MV1 b -tag** We also require that the two jets in the event with the highest p_T be b -tagged offline by the MV1 algorithm. There is no explicit requirement that the leading two jets be trigger-tagged or triple-tagged. In addition to preferentially keeping signal events at a higher

rate than background events, this cut considerably improves the combinatorics of reconstructing the Higgs and we see markedly better signal mass peak resolution when this cut is in place. Further details can be found in Section 7.2.

- Categorization based on b -tagging of a 3rd jet** After two b -jets have been triple-tagged, the event is categorized based on whether a third b -tagged jet is present. The most signal-enriched region is where a third jet passes a tight (60% efficiency) MV1 b -tag (bbb category); the next most signal-enriched region is when no jets in the event pass a 60% b -tag but there is at least one jet passing a loose (80% efficiency) b -tag ($bbloose$ category); the least sensitive region has no additional jets passing either 60% or 80% b -tags (besides the triple-tagged jets, $bbanti$ category). These tag categories are mutually exclusive and all have different signal enrichments, kinematics, and provide varying levels of sensitivity when used in the fit (Section 6.1). If there are more than three jets in an event, the event is categorized by the most b -like (highest b -tag-weighted) jet in the event⁵). If there are more than five jets in an event, only the leading five are considered when assigning that event to a b -tag category.
- Categorization based on the number of jets in the event** Similarly, the events are categorized into 3-jet, 4-jet, and 5 (or more) jet categories, primarily because the varying signal resolution in the different jet bins leads to different signal to background ratios, and separating the categories allows for extra sensitivity to be extracted (Section 7.3.5). Only jets with $p_T > 25$ GeV and $|\eta| < 2.5$ are considered when counting jets.
- Rotation into Eigenbasis, and p'_T cuts** After all the preceding precuts and categorizations are applied, the events are rotated into a new basis based on the eigenvectors of the matrix composed of the signal m_{bb} and the p_T s of the leading 2 jets in the event. Then cuts based on the new (rotated) $p_{T\prime}$ s of the jets are

⁵for example, an event with four jets total, three of which pass tight b -tags, will be categorized as bbb

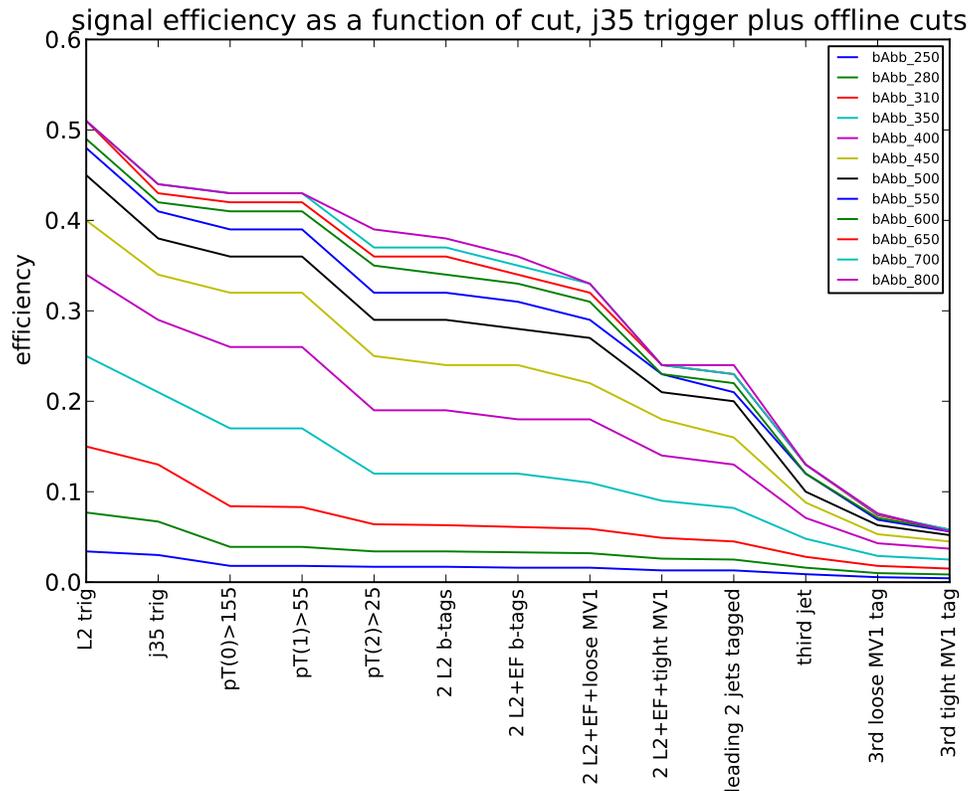


Figure 5.2: The cut efficiencies in signal, cut by cut. Each line is a different signal mass point, and the overall trend is for higher efficiencies for the higher-mass particles than for the lower-mass particles.

applied, and the rotated m_{bb} is used as the final discriminating variable. Much more detail about this procedure, its motivation and results, can be found in Section 7.4).

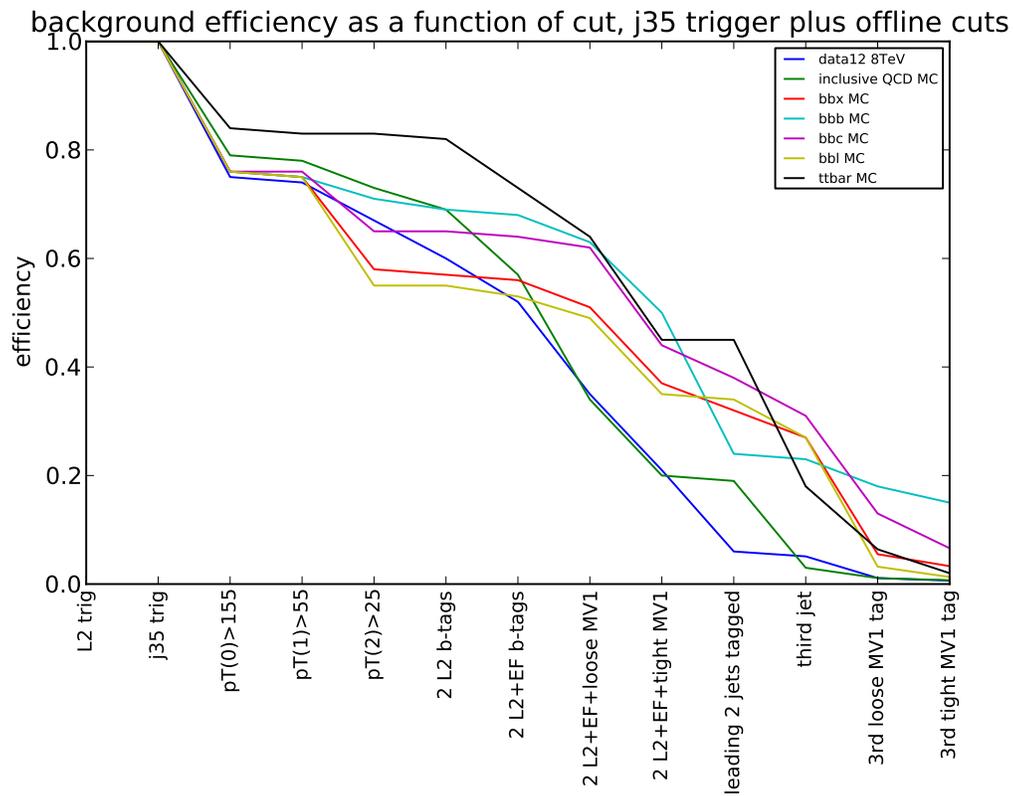


Figure 5.3: The cut efficiencies in background, cut by cut. The background component that has three real b -jets from QCD has the greatest survival rate through the cut flow, followed by events that have two b -jets and one c -jet.

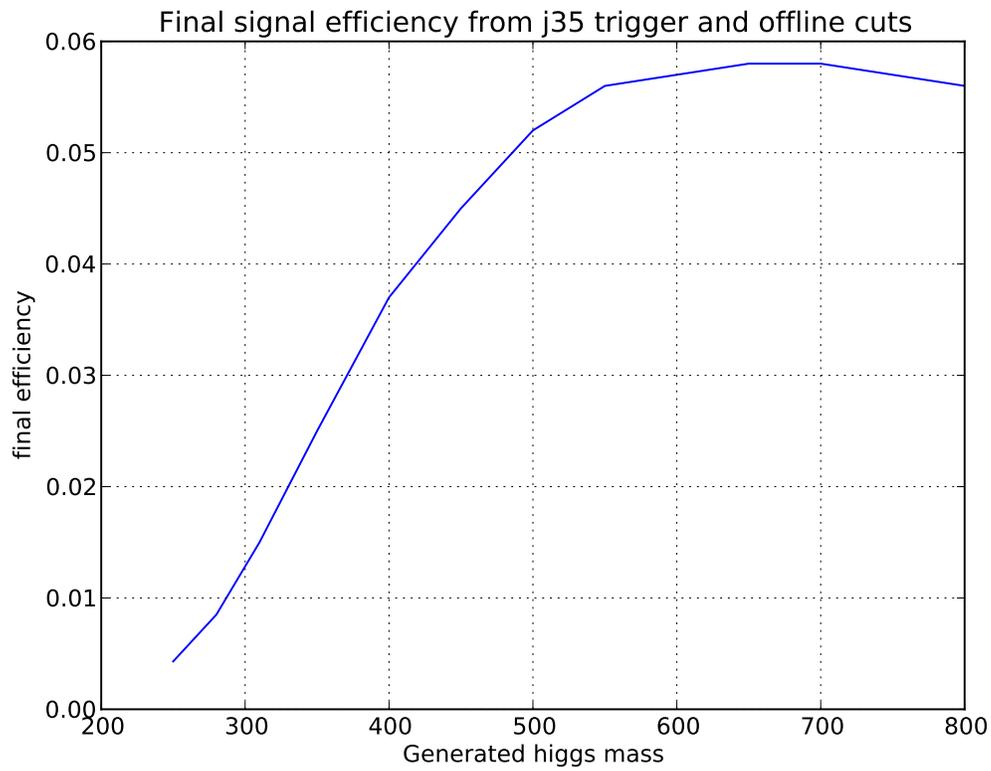


Figure 5.4: The final efficiency, after all cuts, of the signal as a function of the physical mass of the generated Higgs particle. The signal efficiency grows as a function of m_A , although at the high mass values, the efficiency levels off at around 6%.

Chapter 6

Background Estimation

Once the background has been reduced as much as practical using the trigger and kinematic cuts, a reliable estimate of the shape and size of the remaining background is critical to optimizing the exclusion limits. As detailed in Section 7.3, since the reconstructed mass peak of the leading two b -jets is so broad in signal, a mis-estimation of the background shape can lead to systematic errors that could wash out any possible signal (or worse, be mistaken for a signal where there is none).

At the same time, the backgrounds in this analysis are challenging to estimate, either in Monte Carlo simulations or using data-driven methods. Therefore, much of the work of this analysis is dedicated to validating the background, especially its shape in jet p_T , m_{bb} , and so on.

6.1 Background Estimation Strategy

As QCD is the dominant background in this analysis, it is important to understand what flavors of QCD jets compose the population of events that survive the cut chain.

In order to do this, we apply the trigger and cut chain detailed in Chapter 5.

Once the events have been passed through the cut chain, the signal and background are split into 3 exclusive regions based on the b -tag weight of the third jet in the event¹:

- *bbb*: one or more jets (in addition to the two triple-tagged jets) passing a tight (60% b -jet efficiency working point) MV1² cut
- *bbloose*: events failing the *bbb* classification but which have one or more jets passing a loose (80% efficiency working point) MV1 cut
- *bbanti*: events that have no jets passing an 80% MV1 cut-effectively a veto on the presence of any b -tagged jets other than those firing the trigger

When assigning events to one of these categories, we only allow b -tags on the leading five jets to count toward the *bbb* or *bbloose* categories. In other words, if the third b -tagged jet in an event is the 6th jet (in p_T ordering) overall, the event will be classified as *bbanti*. This requirement is motivated by our physics awareness that the b -jets coming from signal events should be fairly high p_T , so this should have a minimal effect of rejecting signal that would otherwise be accepted. On the other hand, with high-jet-multiplicity QCD events, there is the combinatorial effect of looking across more jets for b -tags that increases the likelihood of an event being classified as signal-like (*bbb* or *bbloose*) as there are more jets in the event. Only looking at the leading 5 jets for b -tags keeps this effect under control.

Once all the events have been categorized, the search proceeds by fitting the background in the *bbanti* region, which we sometimes refer to as the *bbanti* control region.

¹as noted below, the “third jet” moniker does not strictly mean the third-highest p_T jet in the event; rather, it is the jet with the highest b -tag weight that is *not* one of the leading two jets in p_T , which are already b -tagged coming out of the cut chain

²recall that MV1 is the name of the offline b -tagging algorithm used in most ATLAS analyses

The events in this category will be mostly QCD background, especially $b\bar{b}$ QCD, although about 20% of the signal might fall into this category. The bbb category will be enriched in $b\bar{b}b$ events, either from signal or QCD background, although 60% of the signal falls into this category and the QCD backgrounds are much lower when a third b -jet is required. The background estimation strategy for this analysis is to fit the m_{bb} distribution in the background-dominated $bbanti$ category, and use that shape to predict the background shape in the bbb signal region. The extrapolation from $bbanti$ to bbb is validated using the bb QCD MC sample described in Section 4.3.1.

6.2 Background Estimation Method Based on Parameterized Histogram Fitting

The background consists almost exclusively of QCD with 2 or more real b -jets, which fortunately has a m'_{bb} ³ spectrum that does not have any peaks or other difficult structure above about 350 GeV. Below that mass, the trigger turn-on curve becomes a major feature of the spectrum. Above that mass, there is a smoothly falling distribution that we fit to an `RhhBinnedPdf`, which is a parameterized histogram PDF (probability distribution function) that is also used for fitting the signal.

In a few words, the background fit strategy proceeds in the following way:

- Start with a signal mass point
- Apply a mass-point-specific rotation based on the the eigenvectors of the signal MC sample of that mass point, as calculated using m_{bb} , $p_{T,1}$ and $p_{T,2}$; we call the components of this rotated basis m'_{bb} , p_T^1 and p_T^2 (more details in Section 7.4). Briefly, the m'_{bb} variable is a linear combination of m_{bb} , $p_{T,1}$ and $p_{T,2}$, which has a shape similar to m_{bb} but exploits correlations in the signal to suppress the

³the m'_{bb} variable is described in more detail immediately below, and in Section 7.4

backgrounds.

- Apply cuts to $p'_{T,1}$ and $p'_{T,2}$, and use m'_{bb} as the final discriminating variable
- Fit the m'_{bb} distributions within the m'_{bb} ranges specified in Table 7.3 with a parameterized histogram, separately for each b -tag and n_{jets} category in signal MC.
- Fix the signal shapes and relative category normalizations to the fits found in MC
- Create a composite PDF that is the sum of the signal PDF and a background PDF; the background PDF is not yet known (either shape or normalization) but the model is a parameterized histogram, with the same range and binning as the signal fits.
- Fit the data in the same m'_{bb} range with the signal+background PDF. The fit is a simultaneous fit to the data in the bbb , $bbloose$, and $bbanti$ tag categories⁴. For example, all 3-jet events (whether in the bbb , $bbloose$, or $bbanti$ tag category) are fit to have the same background shape, modulo a single degree of freedom that allows for a linear shape variation for the different tag categories (more details on this linear variation below). This has the practical effect of $bbanti$ dominating the background fit result because of its higher statistics, so the resulting histogram is dominated by the m'_{bb} distribution in background tag categories.
- Allow the overall signal normalization to float in the final fit (not, however, the relative normalization between categories or the signal shapes) so that, while the background shape will be dominated by the signal-depleted $bbanti$ region, the bbb region in particular can have contributions from signal.

⁴a simultaneous fit will find the parameters that maximize the likelihood across all three tag categories together

- Fit the m'_{bb} distribution in all tag and n_{jets} categories, solving for the signal cross section, background shape, background linear variation between tag categories, and background normalization that maximizes the total likelihood across all tag categories simultaneously. There is no requirement that the background shape be the same across the n_{jets} categories.
- Extract the signal cross section that maximizes the fit likelihood.
- Repeat for the next signal mass point

Although we have signal MC with m_A values as low as 250 GeV, this fit strategy only begins to work above $m_A = 400$ GeV, because of the trigger turn-on curve. The signal mass points to which we would have sensitivity range from 450 to 800 GeV (potentially higher, but 800 GeV is our highest signal MC point).

While the analysis was in development, it was blinded to avoid bias. This was done by removing bbb events from the data, and replacing them with a sample of events with the same normalization drawn from the $bbanti$ distribution in data. For the final search, the bbb events were substituted back in.

6.3 Modeling of Background Shape

6.3.1 Selection of Parameterized Histogram

The fitting of both signal and background is done in RooFit. There are a number of candidate models for the background fit, which can be evaluated both on their goodness-of-fit (as measured by a metric like χ^2/DOF) and their ease of convergence.

A number of mathematical functions were attempted in fitting the background, but in the final analysis, a parameterized histogram was used in part because of its easier

convergence on background, and in part it allowed us to avoid an unstable functional fit to the signal.

The compromise of fitting a histogram to the background removes some of the constraints that a functional fit provides (for example, a decaying exponential imposes the constraint that the background is always decreasing as m'_{bb} increases, but a histogram could fluctuate up in the case of an excess). That means that we must rely more on our background-dominated *bbanti* region to understand the shape of the background (instead of e.g. sidebands).

The PDFs for m'_{bb} are based on a parameterized step function, which is equivalent to parameterizing a normalized histogram in terms of the content of each bin. The number of needed parameters correspond to the number of bins minus one (because of the normalization). Using directly the bin content in each bin b_i would induce an unstable behavior, since the last bin content would need to be parameterized as $1 - \sum_{i=1}^{(N_{bins}-1)} b_i$ and the fit would easily converge to configurations where the last bin is negative, and at that point convergence is spoiled. Instead, the values of p_i are re-parameterized in terms of another set of parameters b_i ⁵.

- $b_1 = p_1$
- $b_2 = p_2 (1 - p_1)$
- $b_3 = p_3 (1 - p_2) (1 - p_1)$
-

With this choice all parameters p_i with i from 1 to $N - 1$ can be limited to $[0, 1]$, without loss of generality, and the fit will always converge reliably, provided the dataset the fit is applied to has at least one event in each bin of the PDF.

If this PDF choice were fit simultaneously in all the b -tag categories as described

⁵This nice trick was introduced for the first time by Aaron Roodman (SLAC) in the BaBar experiment

above, it would introduce the assumption of the same background shape in the *bbb*, *bbloose* and *bbanti* tag categories. However, from studies in background MC simulations, we suspect that there may be (relatively small) systematic shape differences in the different tag categories, and those shape differences are linear in the ratio of the m'_{bb} distributions in *bbb* and *bbanti* (also *bbloose* and *bbanti*; for more details on these background MC studies, see Section 9.1). Although these shape differences are small compared to the overall shape of the m'_{bb} distributions, the difference can be comparable to the effect from a possible signal.

In order to account for this, we add an additional parameter to the fit, which allows for a linear variation on the ratio of m'_{bb} between the different tag categories. We call this the linear variation parameter and it allows the *bbb*, *bbloose* and *bbanti* PDFs to vary by a linear factor:

$$PDF_{bbb}^{i \text{ jets}} = (\alpha m'_{bb}) PDF_{bbanti}^{i \text{ jets}} \quad (6.1)$$

$$PDF_{bbloose}^{i \text{ jets}} = (\beta m'_{bb}) PDF_{bbanti}^{i \text{ jets}} \quad (6.2)$$

α and β are (independent) parameters found during the fit to the data and i is the n_{jets} category. In order to constrain the PDFs to always be positive-definite, we reparameterize the α and β terms so they are constrained to $[-1, 1]$ where -1 corresponds to reducing the contents of the first bin to approximately zero and doubling the contents of the last bin, while $+1$ doubles the contents of the first bin and reduces the contents of the last bin to approximately zero.

In the final fit to the data sample the following parameters are extracted:

- The signal strength μ ;
- The background normalizations $N_{bkg,l}$, separately in each category;

- The background $PDF_{bkg}(m'_{bb})$, which is constrained to the same shape in all tag categories aside from a possible linear variation allowed between various tag categories; and
- The size and sign of linear variation in the m'_{bb} distribution from tag category to tag category (the α and β parameters)

Although the final fit model was a histogram, the following functions were also considered and we discuss below the behavior they exhibited:

- Bernstein Polynomial
- Power Decay Series
- Decaying Exponential with 1 parameter
- Decaying Exponential with 2 parameters

Bernstein Polynomials

The Bernstein polynomials are a family of polynomials that are characterized particularly by their attractive feature of being positive-definite, or never predicting a negative value for a PDF composed of them. However, we find that it takes a high degree of polynomial (5 parameters or more) to fit the background over the full mass range, and a polynomial with a degree this high struggles (and often fails) to converge. Moreover, a drawback of high-degree polynomials is their capability to “wiggle” and potentially absorb a signal within the background model.

Power Decay Series

A power decay series of the form $\frac{a}{x} + \frac{b}{x^3} + \frac{c}{x^5} + \dots$ is another possibility; if all the powers of x in the denominators are odd, this series cannot wiggle like the Bernstein

polynomials. However, a simple series with a few terms does not fit the background shape well over the full mass range, and when many terms are added to the series then the fit struggles to find a global maximum of the likelihood function, leading to non-convergence.

Decaying Exponential with 1 Parameter

Like a decaying power series, a decaying exponential function is monotonically decreasing; however, unlike a power series, an exponential with a single parameter (i.e. a model of the form $f(x) = e^{-x/\tau}$) provides a relatively good fit to the background distribution in most tag and jet categories, with χ^2/DOF values below 2 for all *bbb* and *bbloose* distributions⁶. Additionally, the simplicity of a 1-parameter exponential means that the fits converge quickly and reliably over the full mass range.

Decaying Exponential with 2 Parameters

A decaying exponential with two parameters (i.e. a model of the form $f(x) = e^{-x/\tau + \omega x^2}$) has some of the same nice features of a single-parameter exponential (simple functional form, no possibility of signal-spoofing wiggles) while offering more flexibility than the 1-parameter exponential. The χ^2/DOF and pulls for the 2-parameter exponential fits reflect this flexibility to better fit the data, every jet/tag category has a fit that is as good or better for the 2-parameter exponential as for the 1-parameter exponential. In particular, the *bbanti* categories that have higher χ^2/DOF values in the 1-parameter fit show χ^2/DOF results of 1.0-1.2 with the 2-parameter fits. However, the convergence of this model (especially when it is used in a composite model that can also include signal) can be tricky, and we found that in practice the parameters need to be initialized and constrained very precisely to

⁶the *bbanti* distributions prove harder to fit with a 1-parameter exponential, with χ^2/DOF values between 2.8 and 3.9

get good results. We also found that this model was prone to introducing spurious signal when used in combination with our signal, so that even when running in configurations where no signal was present, the fit was prone to returning a nonzero signal cross section.

6.3.2 Background Shape Variations

A critical assumption of this fit method is that the shape of the m_{bb} distribution varies only linearly based on the flavor (or, relatedly, b -tag status) of the third jet in the event. We validate this assumption in the b -jet enriched QCD MC sample detailed in Section 4.3.1

As detailed in Section 7.4, our final discriminating variable is actually not m_{bb} but a variable we call m'_{bb} , which is a linear combination of m_{bb} , $p_{T,1}$ and $p_{T,2}$. So while we do check that the m_{bb} distributions are not biased by the flavor of the third jet, the real figure of merit lies in showing that m'_{bb} is not biased by the flavor of the third jet either. The plots in Figure 6.1 show m'_{bb} in the bbb and $bbanti$ tag categories for the 700 GeV signal mass point, with a ratio plot drawn below to see any shape differences more clearly.

A fuller suite of plots can be found in the Systematics Chapter, Section 9.1. In these plots, we fit linear functions to the ratios, and find that especially in the 5+ jets category, the shape can vary between categories by a statistically significant amount.

We use the background MC simulations to verify that the fit, as written, can correctly account for linear shape differences from tag category to tag category. We manually apply linear shape variations to the bbb tag category, relative to $bbanti$, and send the resulting distributions through the fitting infrastructure to verify that the parameters reflect the presence of the variations. We find that the fit correctly accounts for the

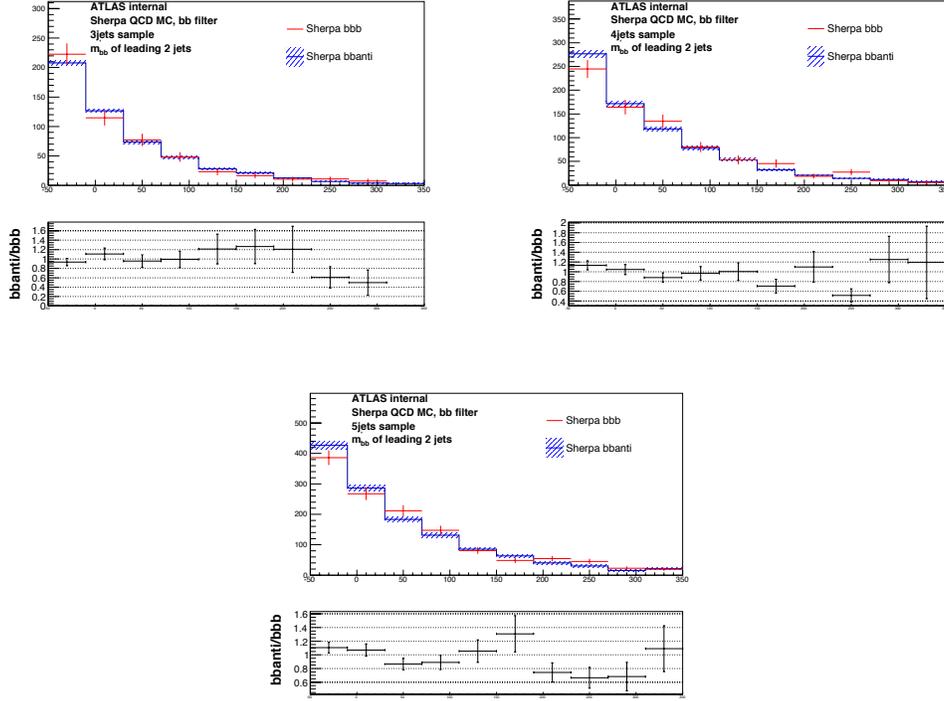


Figure 6.1: The distributions of the final discriminating variable, m'_{bb} , in background MC for the *bbanti* and *bbb* tag categories. The three different plots compare the 3-jet, 4-jet, and 5-or-more jet categories. No signs of major shape differences are seen.

variations, and that the background shapes found by the fit agree perfectly with the distributions that were fed into the fit.

6.3.3 Non-QCD Background

All-Hadronic $t\bar{t}$ Background

When $t\bar{t}$ decays all-hadronically, it can create events with several high- p_T jets and two or more b -tagged jets (where the b -tags come from both real b -quarks and from mistagged light flavor). We anticipate that, because it has a production cross section

that is much smaller than QCD⁷, $t\bar{t}$ will not be a major background.

We find that in Pythia MC simulations, all-hadronic $t\bar{t}$ has an efficiency of 7.5% after the EF_2j35_loose_j145_j35_a4tchad trigger, and approximately 2% efficiency in the offline cuts relative to the trigger. Estimating the $t\bar{t}$ cross section as 165 pb, and a 44% branching ratio in the all-hadronic decay channel, this gives a 0.11 pb $t\bar{t}$ cross section expected after the trigger and offline cuts. In the full 2012 dataset, this amounts to about 2400 events. While this is not a negligible cross section compared to the signal, it is more than an order of magnitude smaller than the QCD background.

In addition to checking the magnitude of the $t\bar{t}$ background, we check the shape for any shape differences in the m_{bb} distribution depending on the tag status of the third jet, and do not find any major discrepancies that point toward $t\bar{t}$ as a potential peaking background in the signal region. The m_{bb} distributions in the bbb , $bbloose$ and $bbanti$ bins for the all-hadronic $t\bar{t}$ can be found in Figure 6.2.

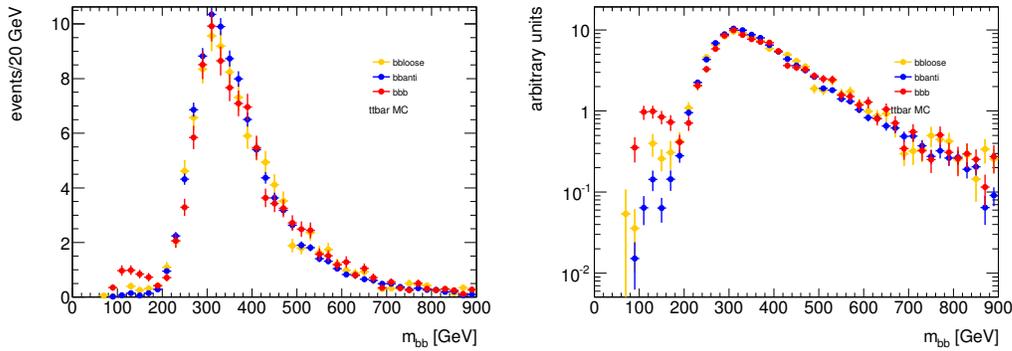


Figure 6.2: The m_{bb} distributions for all-hadronic $t\bar{t}$ MC after the trigger and all offline cuts are applied (linear Y axis on the left, logarithmic scale on the right). In addition to the overall cross-section, we also want to probe any shape differences that arise when the tag status changes on the third jet in the event. No significant shape differences are seen.

⁷this statement depends somewhat on cuts, b -tagging, etc., but as explained below it does turn out to be the case in the relevant region in phase space for this analysis

Chapter 7

Signal Kinematics

Once the trigger and cuts have been applied to the data, the events that remain can be signal, background, or a mixture of the two. It then falls to understanding the kinematics of the signal and background to look for features that can provide enough discrimination to detect the presence of the signal. In this case, that means reconstructing the invariant mass of the leading two jets in the event, which will create a smoothly falling power law distribution in QCD background but a resonance structure in signal. The kinematics of the signal distribution then come into play in a few major ways:

- **Signal Jet p_T Distributions:** By examining the p_T distributions of the Higgs daughters, and the associated b -jet(s), we can get a better understanding of what might be optimal jet p_T cuts for future iterations of this analysis.
- **Signal Jet Combinatorics:** A perfectly good signal event can be inadvertently turned into a type of background if the associated b -jet is chosen when reconstructing the Higgs; unfortunately these combinatorial mistakes can arise easily when there are 3 b -jets but only two jets being used for reconstruction.

- **Mass Resolution:** The Higgs resonance width depends on $m_A/\tan\beta$, with intrinsic widths of up to a few tens of GeV; however, if energy is lost to FSR, the resonance can be much wider and this makes the signal much more difficult to distinguish from the background.
- **Eigenvector Rotation:** In signal events, especially those with low FSR and good mass resolution, there is a correlation between m_{bb} , $p_{T,1}$ and $p_{T,2}$ that is not present in the resonance-less background. By constructing a 3x3 tensor out of these three variables in signal, and rotating the analysis so the discriminating variable falls along the leading eigenvector of that tensor, we can take advantage of the correlations in signal and boost the signal-to-background ratio.

7.1 Signal Jet p_T Distributions

An important place to start is with the p_T distributions of the jets after the trigger and cuts have been applied. Using truth-matching in the Monte Carlo, we can identify which jets are daughters of the Higgs and which are associated b -jet(s). Since the Higgs is a heavy particle, its decay jets tend to have higher p_T than the associated b -jet(s), which can be seen in Figure 7.1. However, in this plot, b -tagging has been applied so an important potential component is missing: how does the event topology change when we also include non- b -jets, which can potentially be high p_T as well.

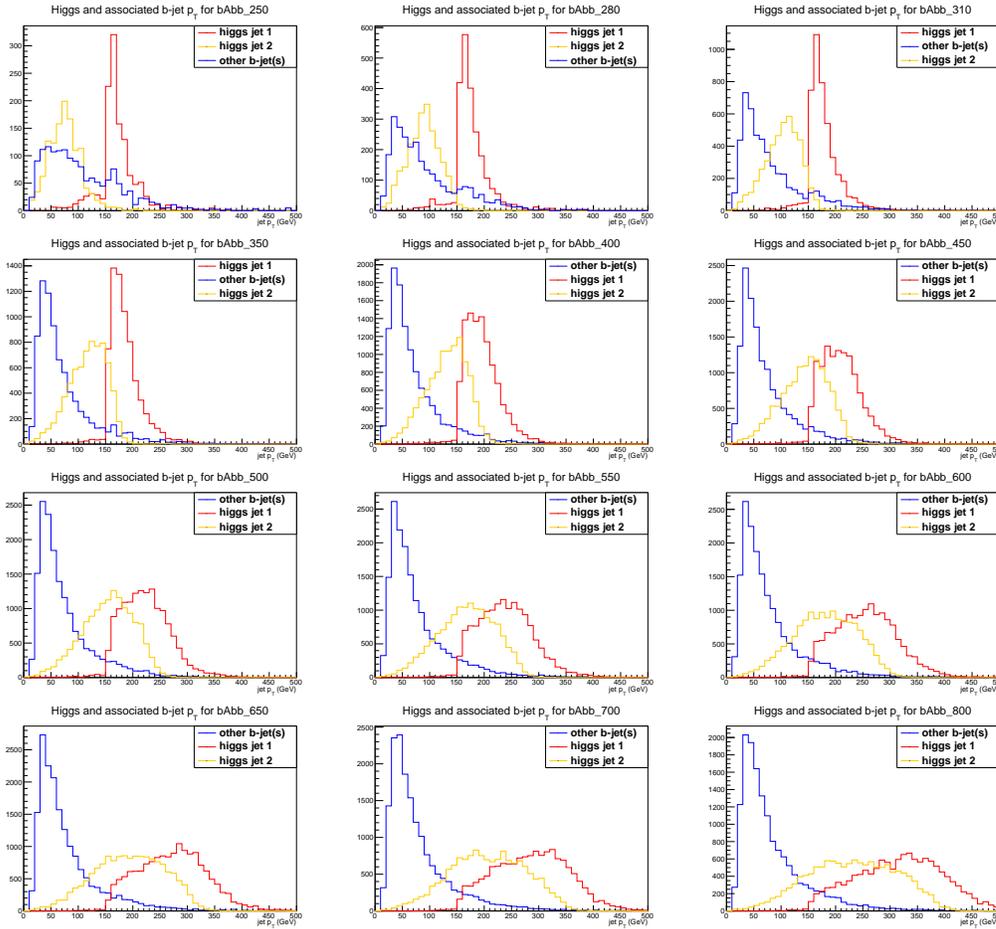


Figure 7.1: The p_T of the two jets from the Higgs (classified as “first” and “second” based on which one has more p_T in the event), and all other b -tagged jet(s) in the event. The peaks at about 155 GeV and 55 GeV is a result of the trigger turn-on points and associated cut(s). Each plot corresponds to a different Higgs mass point, which are indicated in the titles immediately above the plots.

As it turns out, the presence of additional non- b jets can have a significant effect in signal. Since the hard scatter has only b -jets in the final state, the extra jet(s) must come from other sources like initial state radiation (ISR) or final state radiation (FSR). This can smear out the signal, as explained in more detail in Section 7.3, and combating the effects became a significant part of the analysis.

7.2 Signal Jet Combinatorics

An important point when trying to reconstruct the Higgs is which two b -tagged jets (of the three or more available) should be used in reconstruction. If the associated b -jet is accidentally selected, then the Higgs will be mis-reconstructed and the sensitivity will suffer. At the same time, since the mass of the Higgs is not known, tools like a kinematic fit are not available to help with the combinatorics. In Figure 7.2, we use the MC truth information in the signal MC to plot how often the Higgs decays to various pairs of jets within the event; we find that especially for masses above 350 GeV or so it usually decays to the leading and second b -tagged jets in the event (with about a 70% probability).

If we require only that the leading two b -tagged jets be used in reconstruction, there can still be events in our sample where the leading two b -tagged jets in the event are not the leading two jets overall—for example, there can be a very hard light jet, and the b -tagged jets are the second and third jets in p_T . When we introduce a cut that eliminates these events, by requiring that the leading two jets in the event be b -tagged, both the signal to background ratio improves (this cut is about 25% efficient in background, and over 90% efficient in signal) and the mass resolution is improved in signal as the events that are rejected have hard FSR present and the reconstructed m_{bb} is far from the generated m_A .

Quantifying the effect of this cut on the mass resolution is tricky because of the m_{bb} shape in signal, which is highly asymmetric. In order to make some progress, though, we define the “left” and “right” shoulders of the distribution, where the left shoulder corresponds to the low-mass side of the m_{bb} distribution (i.e. where the reconstructed mass is below the peak for a given signal mass point), and the right shoulder is the part of the distribution above the peak. Then we can define a threshold corresponding to 20% of the peak height, and ask for the left-shoulder and right-shoulder windows where the distribution is above the threshold (a schematic

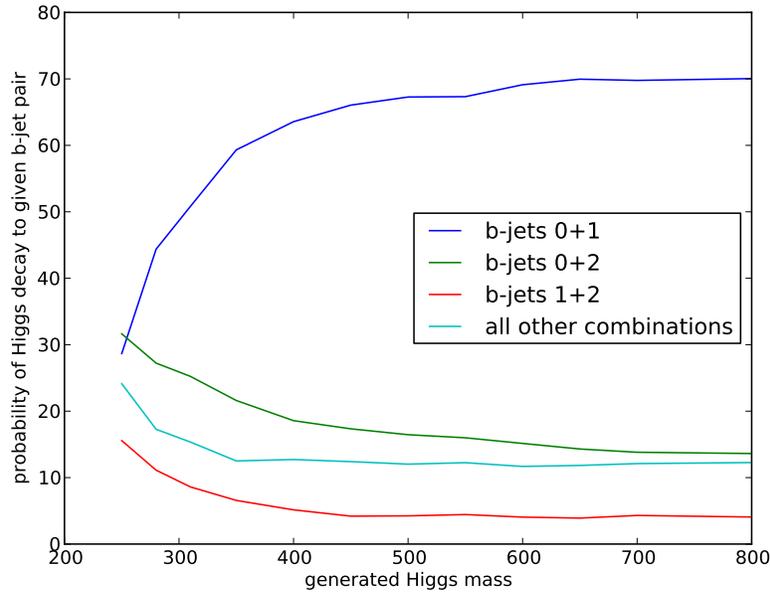


Figure 7.2: The percentage probability of the Higgs decaying to a given pair of b -tagged jets in signal MC, as a function of the generated Higgs mass. The assumption being checked is that the majority of the time, the Higgs decays to the leading two b -tagged jets (blue line), so that reconstructing with these jets will yield the correct combinatorics. That assumption is correct 60-70% of the time for $m_A > 350$ GeV, with the remaining 30-40% of events being spread out over many other combinations of jets.

illustrating the “shoulder” definition can be seen in Figure 7.4, for a low-mass and high-mass example). The right shoulder is much narrower, and any events in the high-mass tail (above the generated value for m_A) arise from combinatorial misreconstructions. The left shoulder, where FSR and combinatorics are at play, is significantly wider. Using the width of the shoulder (in GeV) as our metric, we see an improvement of 40-300% when we add the requirement that the leading two jets in the event be b -tagged relative to when this cut is not in place. The 40% improvement comes in the left shoulder of the 800 GeV signal mass point, pointing toward a reduction of events that show high rates of FSR; the 300% improvement

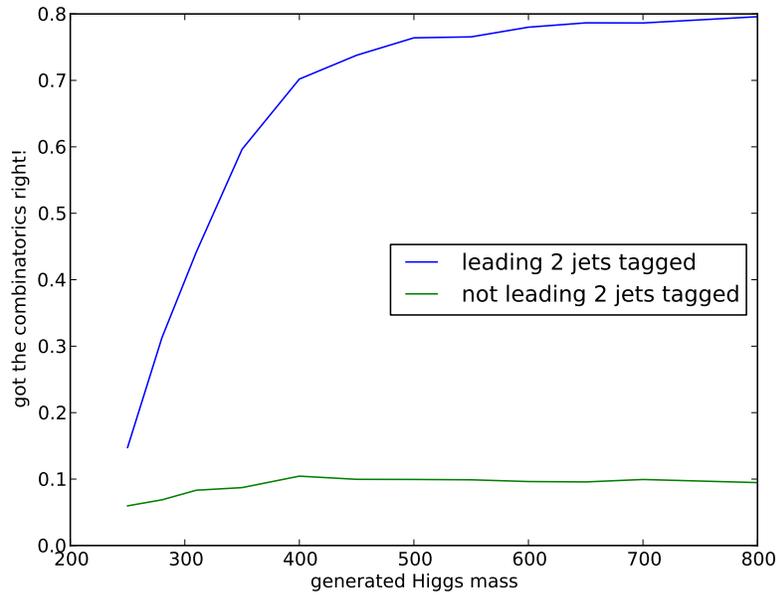


Figure 7.3: After making the requirement that the leading two jets in the event are b -tagged, the signal MC can be split into events that pass this requirement and those events that fail. For each sample, we ask how often the Higgs decays to the leading two b -jets in the event, or in other words, how likely it is to get the combinatorics of the Higgs reconstruction correct. The combinatorics for the events passing the requirement are correct about 80% of the time for high m_H , while for events failing the requirement the combinatorics success rate is only about 10%.

is seen at our 250 GeV mass point and again can be attributed to lower FSR rates. Since the higher-mass signal distributions are much wider to begin with, though, the absolute effect of this cut is greatest at high m_A . The widths of the shoulders for all mass points, before and after this cut is applied, can be seen in Table 7.1.

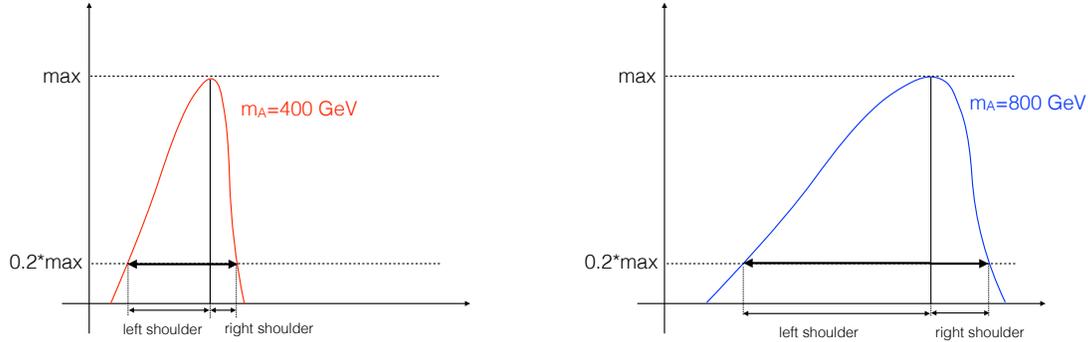


Figure 7.4: A schematic diagram showing how the shoulders are defined, for the purposes of parameterizing the width of the signal distribution. The left shoulder especially is affected when FSR is present, as the Higgs reconstructs to a lower mass than its “true” mass.

7.3 Mass Resolution

Although the cut on the b -tag status of the leading two jets improves the combinatorics and mass resolution, the width of the m_{bb} distributions after this cut still remain very broad, which makes them challenging to distinguish from background. In this section, we document a number of investigations into the mass resolution, and some attempts to recover FSR or mitigate its effects. In brief summary, this section includes discussion of the following mass resolution topics:

- validation that FSR has a relatively larger effect for high-mass Higgs events than for low-mass ones

Table 7.1: The width of the mass distribution in the left and right shoulders of the peak, as explained in the text. The left shoulder is dominated by radiation off the Higgs and/or the Higgs daughter jets in addition to combinatorial mis-reconstructions, leading to a larger RMS than the right shoulder, which is dominated by combinatorics only.

m_A (GeV)	b -tags on leading jets (GeV)		no b -tags on leading jets (GeV)	
	left	right	left	right
250	13.5	89.2	36.7	87.7
280	19.2	68.8	27.5	71.4
310	19.2	58.3	31.8	58.1
350	28.3	34.5	33.9	36.2
400	33.4	26.1	45.1	26.5
450	39.5	25.7	62.2	25.9
500	48.6	29.3	81.1	29.8
550	54.9	31.1	93.1	33.3
600	73.0	26.0	117.6	30.5
650	75.4	31.7	130.8	31.8
700	75.7	40.3	143.4	40.7
800	100.9	36.2	177.4	36.3

- investigation of modifications to the jet reconstruction parameters, and the effect on the mass resolution
- an FSR recovery algorithm that looks for additional jets near ($\Delta R < 1.0$) the b -jets in signal candidate events, with the goal including the additional jets in the reconstruction
- an FSR recovery algorithm that looks for additional hard ($p_T > 80$ GeV) non- b -tagged jets in signal candidate events, for possible inclusion when reconstructing the Higgs
- a change of basis where the discriminating variable is changed from the reconstructed mass, m_{bb} , to a linear combination of the reconstructed mass and the p_{TS} of the leading two jets ($p_{T,1}$ and $p_{T,2}$), which takes advantage of correlations between m_{bb} , $p_{T,1}$ and $p_{T,2}$ that are present for signal (and especially signal events with good mass resolution) to suppress the backgrounds

7.3.1 Validation of Qualitative Shape Differences

The first question to ask is whether the high-mass m_{bb} distributions are qualitatively lower-resolution than the low-mass distributions—in other words, whether the ratio of the width to the generated mass changes as a function of mass. We probe this by “normalizing” the signal MC mass, dividing each entry in the m_{bb} histogram by the nominal mass which was generated; for example, a signal MC event that was generated with a Higgs mass of 350 GeV and reconstructed to have 300 GeV would be entered into the histogram as (300/350). Repeating this normalization process for several mass points allows us to compare the shapes for different m_A values. We find that even when the m_{bb} distributions are renormalized, the high-mass distributions remain wider than the low-mass distributions. That points toward FSR as a likely culprit, and in the following sections we will detail several attempts at FSR recovery.

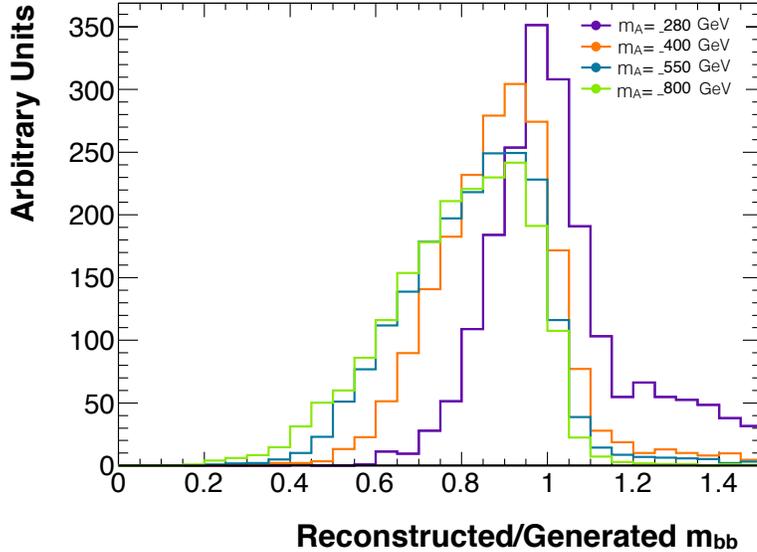


Figure 7.5: The normalized m_{bb} distributions for several signal mass points. The mass resolution is qualitatively worse for higher-mass Higgs bosons.

7.3.2 Jet Reconstruction Modifications

As noted in Section 3.2, this analysis uses anti- k_t jets with a distance parameter of $R=0.4$. It is possible that, for a significant number of jets, some of the radiation associated with the jet is falling outside of the jet radius and is therefore lost during clustering and reconstruction. A way to test this hypothesis is to reconstruct the signal events with $R=0.6$ jets, and see if the mass resolution improves.

We find that the mass resolution does get marginally better (narrower distribution and higher mean for the m_{bb} distribution) for high-mass signal MC events, but only

by a few percent. Practical concerns prevent $R=0.6$ jets from being a serious consideration: first, the background would also be pushed toward higher masses as QCD jets are reconstructed with larger radii; second, pileup calibration and removal would be a significantly larger concern; and third, the b -tagging calibrations are created only for $R=0.4$ jets so the b -tagging systematics would be much more difficult to quantify.

7.3.3 Topology-Based Energy Recovery Algorithm

One hypothesis for the distribution of the FSR is that, for the high- p_T jets that result from the Higgs daughter b -quarks, they might radiate away partons that are hard enough to be clustered into their own jets ($p_T > 25$ GeV) but end up nearby the leading jets in the detector (where nearby is defined as within $\Delta R < 1.0$). In other words, we look for a topology where there are one or more jets near one of the leading two jets, and in cases where such jets are found, we add them back in when reconstructing m_{bb} . This is found to have little effect on the mass resolution. In practice, only about 20% of events have the topology where additional jet(s) are found near the leading two jets, and the mass resolution looks virtually unchanged before and after adding this correction. This suggests a couple of possible conditions: either the FSR is too diffuse and/or low-energy to be clustered into jets, and/or it is farther than the $\Delta R < 1.0$ search zone allows for recovery.

7.3.4 p_T -Based Energy Recovery Algorithm

The p_T -based energy recovery algorithm is based on the idea that perhaps the FSR takes the form of one or two very high-energy radiated partons that can end up far away from the leading two jets but can be identified by their high p_T . In practice, what this means is looking for events with a non- b -tagged jet that has $p_T > 80$ GeV,

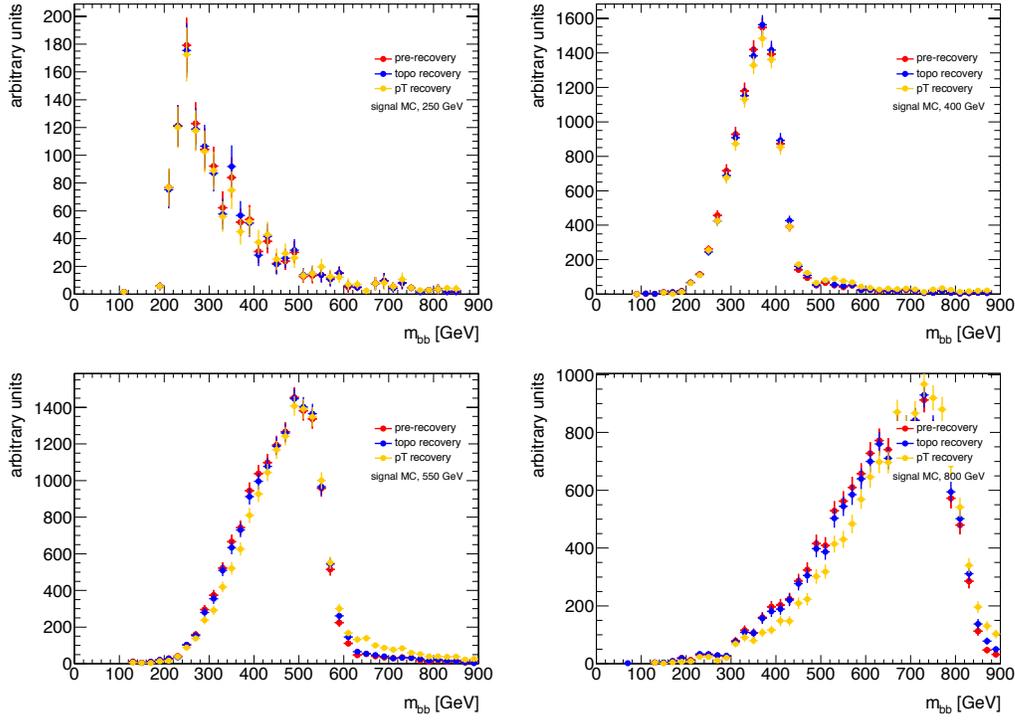


Figure 7.6: Comparisons of the m_{bb} distributions out-of-the-box, and after attempting the p_T -based and topology-based FSR recovery algorithms, for several signal mass points.

and when one is found, adding this back in when reconstructing m_{bb} . Unfortunately the effect on the mass resolution is negligible. A few events move out of the low-mass tail, and the high-mass tail gets a few more events as jets can be mistakenly added back in even when no FSR was lost.

7.3.5 n_{jets} Dependence

The signal mass resolution also displays a large dependence on n_{jets} , the number of central¹ jets over 25 GeV in the event. If the event has significant FSR, that

¹central is defined here as $|\eta| < 2.5$

radiation can end up clustered into extra jets in the event, and the mass resolution suffers at the same time. In events where there are only three jets, we can say rather confidently that two of the jets come from the Higgs decay and the third is from the associated production; as long as we pick the jets properly, the mass resolution should not show major effects of FSR. That better resolution propagates through to allow the fit an easier time in distinguishing signal from background, and generally gives better sensitivity than when the signal has poor mass resolution, all other things being equal. That motivates our decision in the fitting step of the analysis to treat the 3-jet, 4-jet, and 5 or more jet events separately, and only combine at the end when computing the net sensitivity. More details on that procedure can be found in Section 8.0.1.

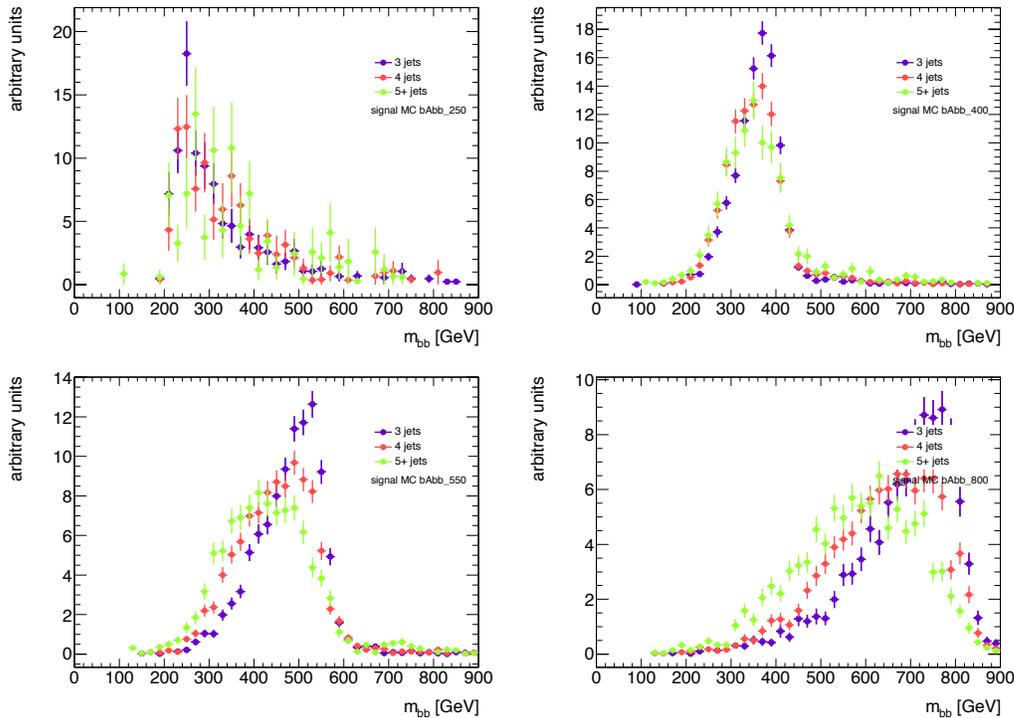


Figure 7.7: The m_{bb} distributions at a few representative mass points in signal MC simulations, as a function of the number of jets in the event. The 5+ jets bin contains all events with 5 jets or more.

7.4 Eigenvector Rotation

Once the events have been categorized based on b -tag status and the number of jets in the events, but before doing the final fit, we apply a change of variables based on p_T s of the leading two jets and m_{bb} in each event. These three kinematic variables as the discriminating variable (rather than m_{bb}). This section details the motivation and procedure of the variable change, which helps separate the signal and background in the final search and results in an analysis sensitivity that is 15-50% better than an analysis without these steps.

The motivation for this variable change strategy has its origins in trying to control the effect of FSR on the analysis sensitivity. At high m_A , the jets from the Higgs (and also the associated b -quark) have large p_T , which they then tend to radiate away as FSR. This causes the m_{bb} peak to smear out and be more difficult to distinguish from background than in the lower- m_A cases. However, since FSR occurs stochastically on an event-by-event basis, a subset of the signal events will have little or no FSR and will reconstruct to (nearly) m_A ²—if these events can be isolated from the others, they offer a chance to improve the sensitivity via the improved mass resolution and signal-to-background ratio.

A simple strategy might be to apply a cut to the p_T of the leading and/or second jet in the event, and only accept events where the p_T is above some threshold optimized by comparing signal MC to *bbanti* data. That would isolate events with poor mass resolution (i.e. lots of FSR) from those with little FSR and better reconstruction properties. However, since there is a correlation between $p_{T,1}$, $p_{T,2}$ and m_{bb} , a cut like this ends up sculpting the background as well. In addition, the $p_{T,1}$ and $p_{T,2}$ cut thresholds that looked best on signal MC are “too good” on background-dominated *bbanti* events and there is no background above a certain m_{bb} value. That makes modeling the background extremely difficult, and any fit difficult to validate.

²modulo any systematic errors in the jet reconstruction, calorimeter noise, etc.

The rotation is performed as follows:

- For each variable m_{bb} , $p_{T,1}$ and $p_{T,2}$, calculate the mean of all signal MC simulation events, and subtract it from each event (i.e. center the m_{bb} , $p_{T,1}$ and $p_{T,2}$ distributions at zero)
- For each event, fill a 3×3 tensor with the following components

$$\begin{bmatrix} m_{bb}m_{bb} & m_{bb}p_{T,1} & m_{bb}p_{T,2} \\ & p_{T,1}p_{T,1} & p_{T,1}p_{T,2} \\ & & p_{T,2}p_{T,2} \end{bmatrix} \quad (7.1)$$

(the matrix is symmetric, so we write only the elements above the diagonal for brevity and clarity)

- Sum these tensor matrices for all signal events in a given n_{jets} category; do not differentiate based on tag category (the resulting tensor will also be 3×3)
- Find the leading eigenvector³ of the tensor matrix
- Define the m'_{bb} variable as follows:

$$m'_{bb} = [m_{bb} \ p_{T,1} \ p_{T,2}] \cdot [e_1 \ e_2 \ e_3] \quad (7.2)$$

where the latter vector is the leading eigenvector of the tensor matrix

- perform this transformation for each signal mass point separately

After this rotation, m'_{bb} is an admixture of m_{bb} , $p_{T,1}$, and $p_{T,2}$. The exact contribution from each of these variables changes on a mass-point-by-mass-point basis, as the correlations between them change. At lower masses, m_{bb} provides most of the basis for m'_{bb} , with 76% of the m'_{bb} coming from m_{bb} . At higher masses, where FSR is more of an issue, the correlations between m_{bb} and the leading jets' p_{T} s are more

³where the leading eigenvector is the eigenvector associated with the largest eigenvalue

Table 7.2: The major axis eigenvector elements, and their squares. The transformation into the eigenbasis makes use of the raw elements, but the squares of the elements sum to 1.0 and can provide some physical intuition for the relative contributions of $p_{T,1}$, $p_{T,2}$ and m_{bb} .

m_A	e_1	e_2	e_3	e_1^2	e_2^2	e_3^2
450	0.87	0.35	0.35	0.76	0.12	0.12
500	0.84	0.38	0.39	0.71	0.14	0.15
550	0.80	0.40	0.46	0.64	0.16	0.21
600	0.78	0.41	0.47	0.61	0.17	0.22
650	0.72	0.46	0.52	0.52	0.21	0.27
700	0.71	0.47	0.54	0.50	0.22	0.29
800	0.67	0.50	0.55	0.45	0.25	0.30

important and for $m_A=800$ GeV, over half of m'_{bb} (55%) comes from $p_{T,1}$ and $p_{T,2}$. The full set of m'_{bb} components can be found in Table 7.2.

The effect of this transformation on 450 and 800 GeV signal and background-dominated data can be seen in Figure 7.8. These figures show the difference in the m_{bb} distributions before and after the transformation, with the mass distribution being pushed to lower values in the background distribution while the signal mass maintains a similar shape. This background reshaping affects the signal to background ratio in the neighborhood of the signal peak; before the rotation, significant background remains in the neighborhood of the signal peak, but after the rotation the signal is virtually unchanged while the background has been shifted lower.

While the m'_{bb} variable is derived using the leading eigenvector of the tensor matrix, the second and third eigenvectors can also be used in an analogous way and we refer to the resulting transformed variables as $p'_{T,1}$ and $p'_{T,2}$.

There is one more step that optimizes the signal to background discrimination further, which is placing cuts on $p'_{T,1}$ and $p'_{T,2}$. If we require that $p'_{T,1} > -10$ GeV, and $p'_{T,2} > -50$ GeV, it excludes a region of m'_{bb} phase space where the modeling is difficult and little sensitivity is gained.

Since the rotation is calculated on each signal mass point separately, and the mass points are characterized by different correlations between m'_{bb} , $p_{T,1}$ and $p_{T,2}$, in particular because the mass resolution varies significantly based on the mass of H/A . As a result, the background will change as a function of the signal mass point. The signal and background shapes after the rotation can be seen in Figure 7.9. The signal before the rotation is drawn in Figure 7.10.

Signal-Mass-Point-Specific Search Windows

The signal shape is fit using an RhhBinnedPdf in RooFit, a form of parameterized histogram where the parameterization helps the fit converge more reliably and allows a single parameter to capture the integral (i.e. signal yield) of the distribution. Once the full analysis cut chain has been applied, including the rotation, the signal m'_{bb} distributions are fit in signal MC and then the shapes are fixed. The fitting procedure is described in detail in Chapter 8; the rotation has different effects on the background depending on the signal mass point under inspection, so we allow the search range in m'_{bb} to vary by mass point. As Figure 7.9 shows, the background can shift significantly from left to right depending on the details of the (mass-point-specific) rotation, so in particular the low edge of the search window can be moved to more negative values of m'_{bb} when one is looking for a high-mass H/A , than when searching for an H/A with a lower mass. The mass windows for each search point are listed in Table 7.3.

For all fits, the histogram bins are 40 GeV wide. 20 GeV binning was also investigated, as well as variable-width binning, but there were only negligible differences in the final sensitivity based on with bin widths.

The choice of a parameterized histogram for the signal model was taken after it was found that the signal distribution is a difficult one to fit to a function, since the shape can vary widely by jet bin and b -tag category. Attempts to fit the distributions with

Table 7.3: The search windows in m'_{bb} , based on the signal mass point under investigation. The background gets pushed to lower values of m'_{bb} by the rotation for higher signal mass values, which allows for wider mass windows for higher signal mass points.

signal mass point	low edge m'_{bb}	high edge m'_{bb}
450 GeV	-50	350
500 GeV	-90	350
550 GeV	-130	350
600 GeV	-130	350
650 GeV	-170	350
700 GeV	-210	350
800 GeV	-250	350

e.g. a Cruijff function (a bifurcated Gaussian function with non-Gaussian tails) would work well for some categories and signal mass points, but fail to give good shape descriptions for other categories and signal mass points. By fitting to a histogram, we are able to fit more difficult shapes than could be reliably fit with a function, and the fits are more stable under the addition of perturbations from systematic errors.

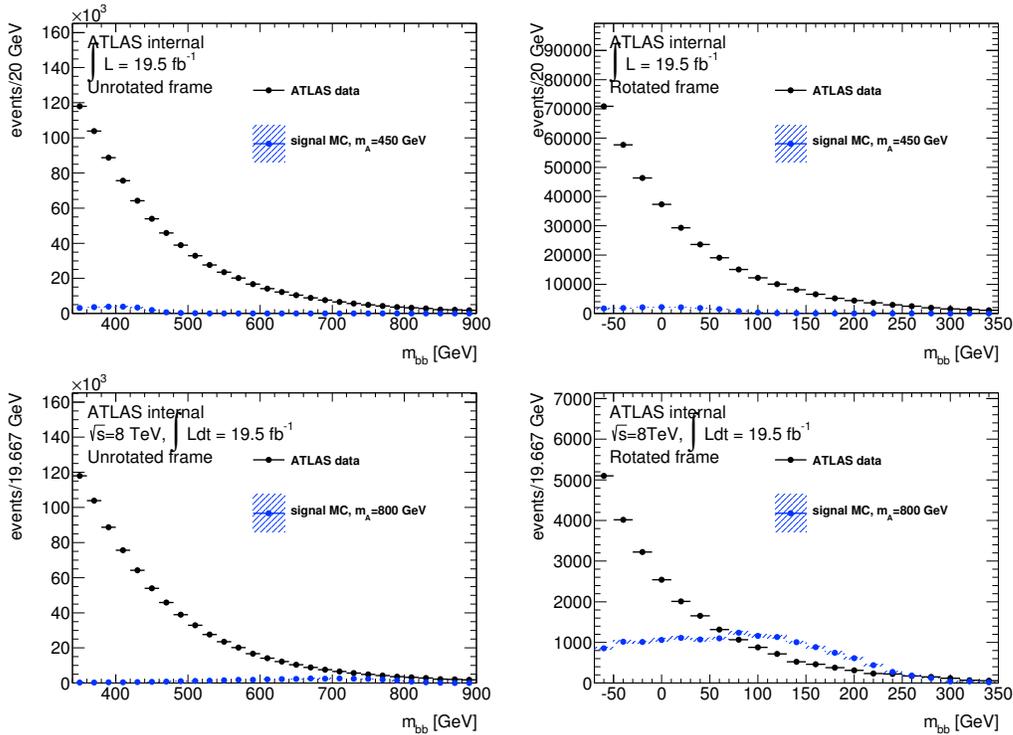


Figure 7.8: Comparisons of the signal and background m_{bb} distributions before and after the rotation. The left two plots show 450 GeV and 800 GeV Higgs mass points before the rotation. The signal is extremely small compared to the background, whereas the same distributions (including the same signal normalization) after the rotation are on the right, for the same two mass points. The rotation pushes the background events to lower values of m'_{bb} than the signal, yielding better signal/background separation, especially for higher mass Higgs candidates.

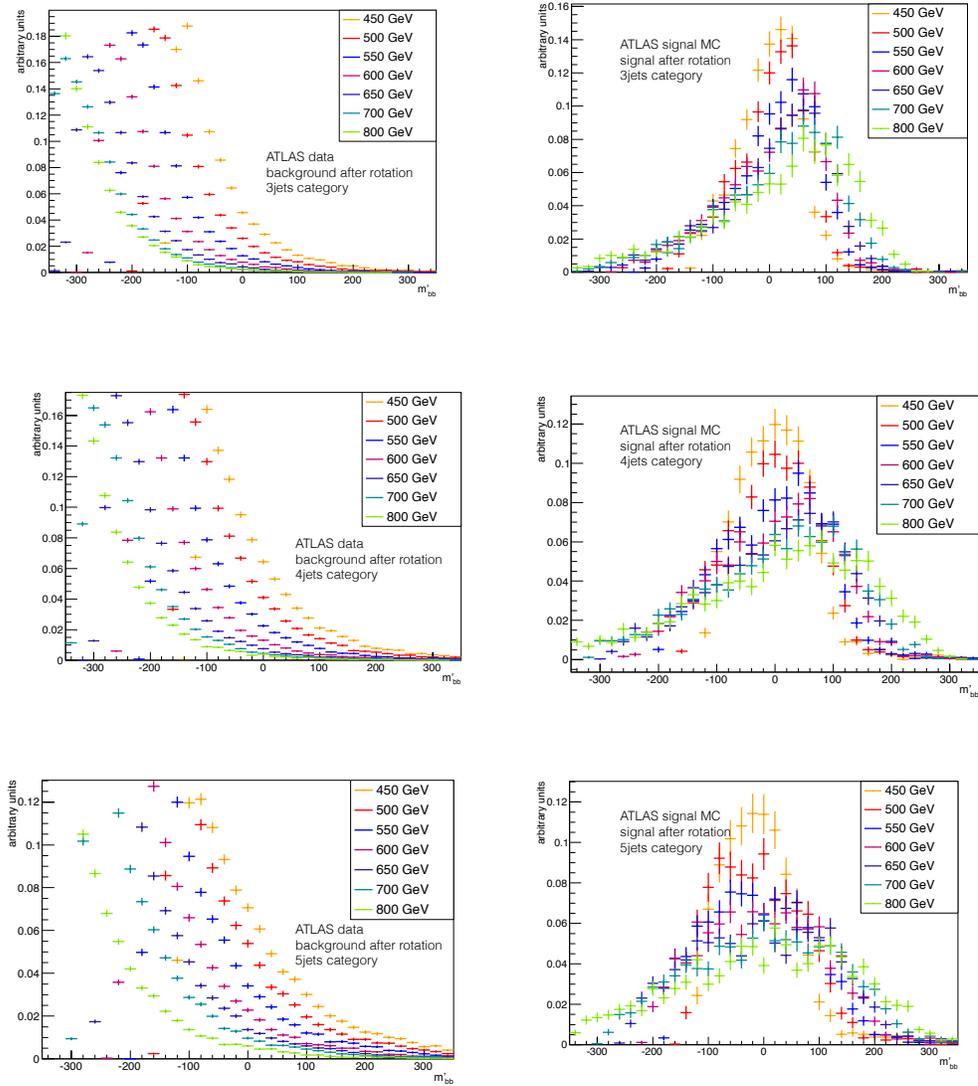


Figure 7.9: The signal and background distributions for each signal mass point after the rotation, binned in n_{jets} category. The effect of the rotation on the background can be clearly seen in these plots, pushing the background further to the left (lower m'_{bb} values) for higher mass signal points.

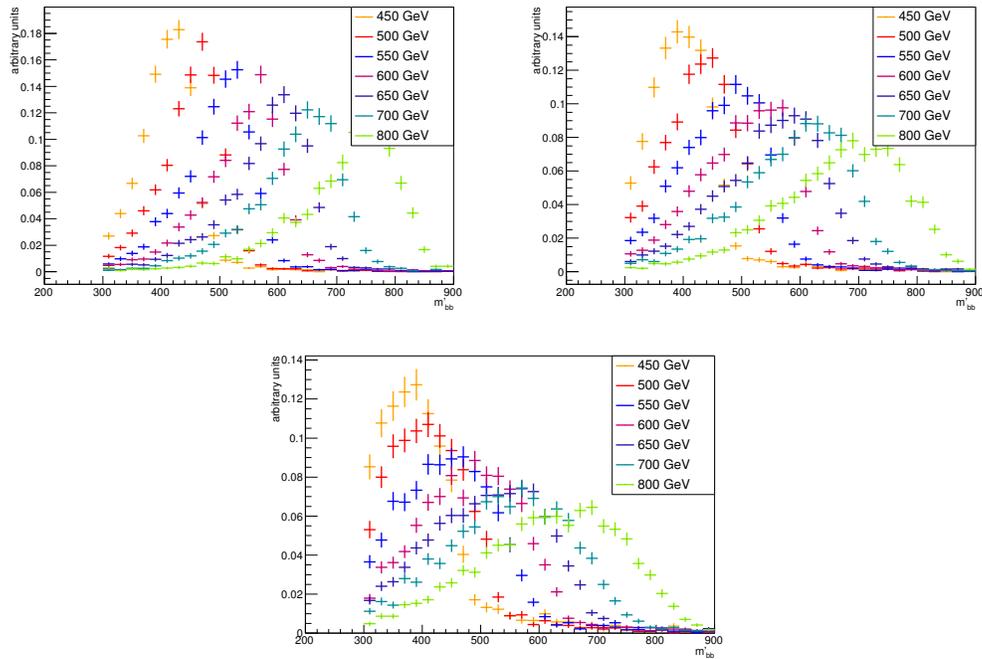


Figure 7.10: The signal distributions for each signal mass point before the rotation, binned in n_{jets} category. Upstream cuts effectively cut off the distributions below 300 GeV.

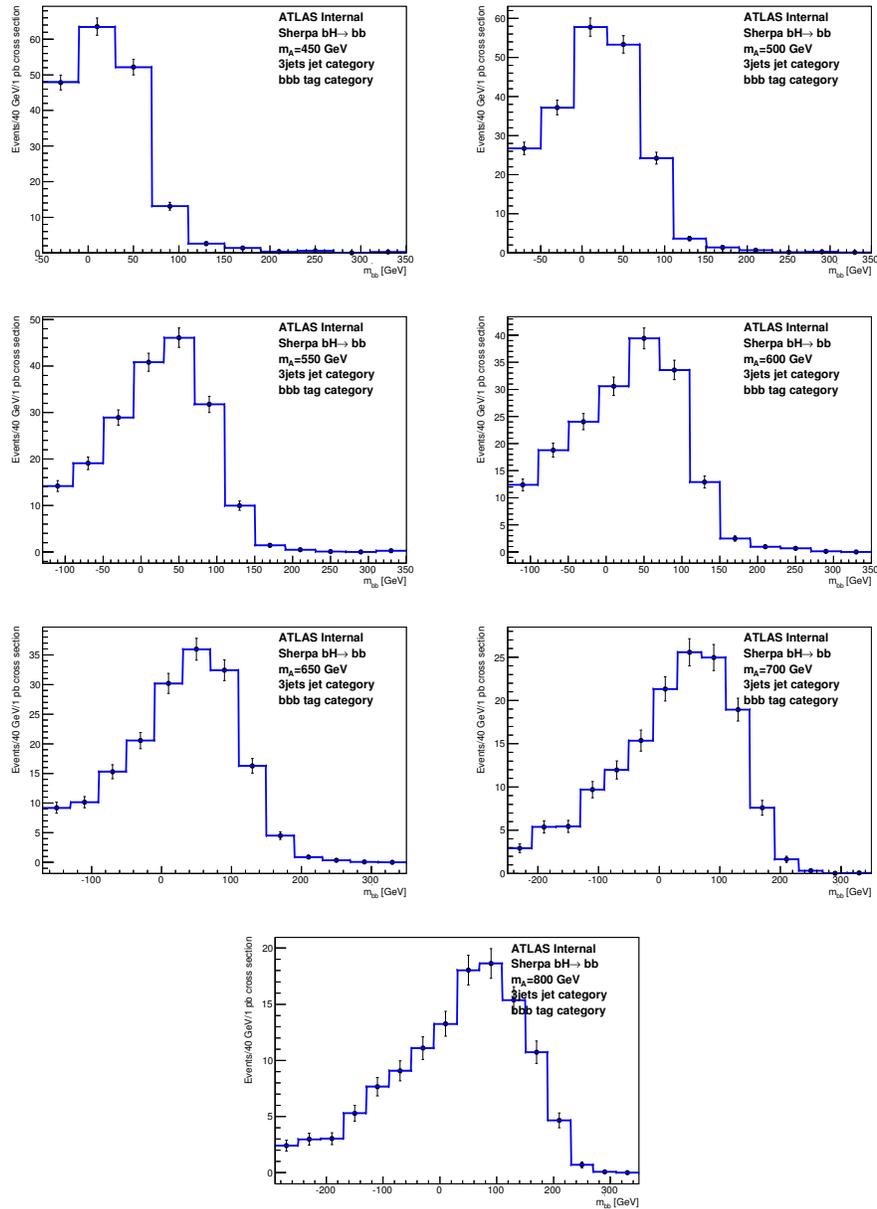


Figure 7.11: Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the bbb category, for events with 3 jets, for different H/A masses after the rotation has been applied.

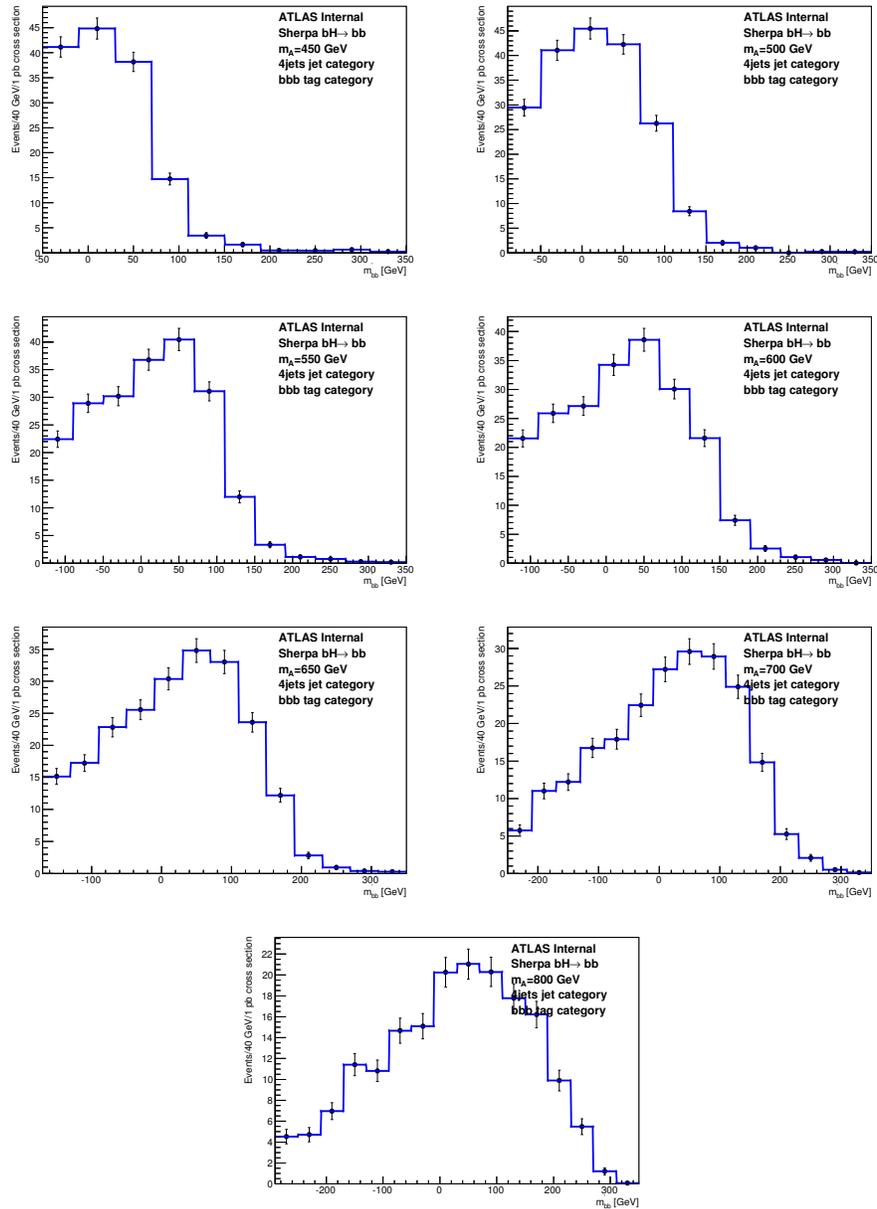


Figure 7.12: Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the bbb category, for events with 4 jets, for different H/A masses after the rotation has been applied.

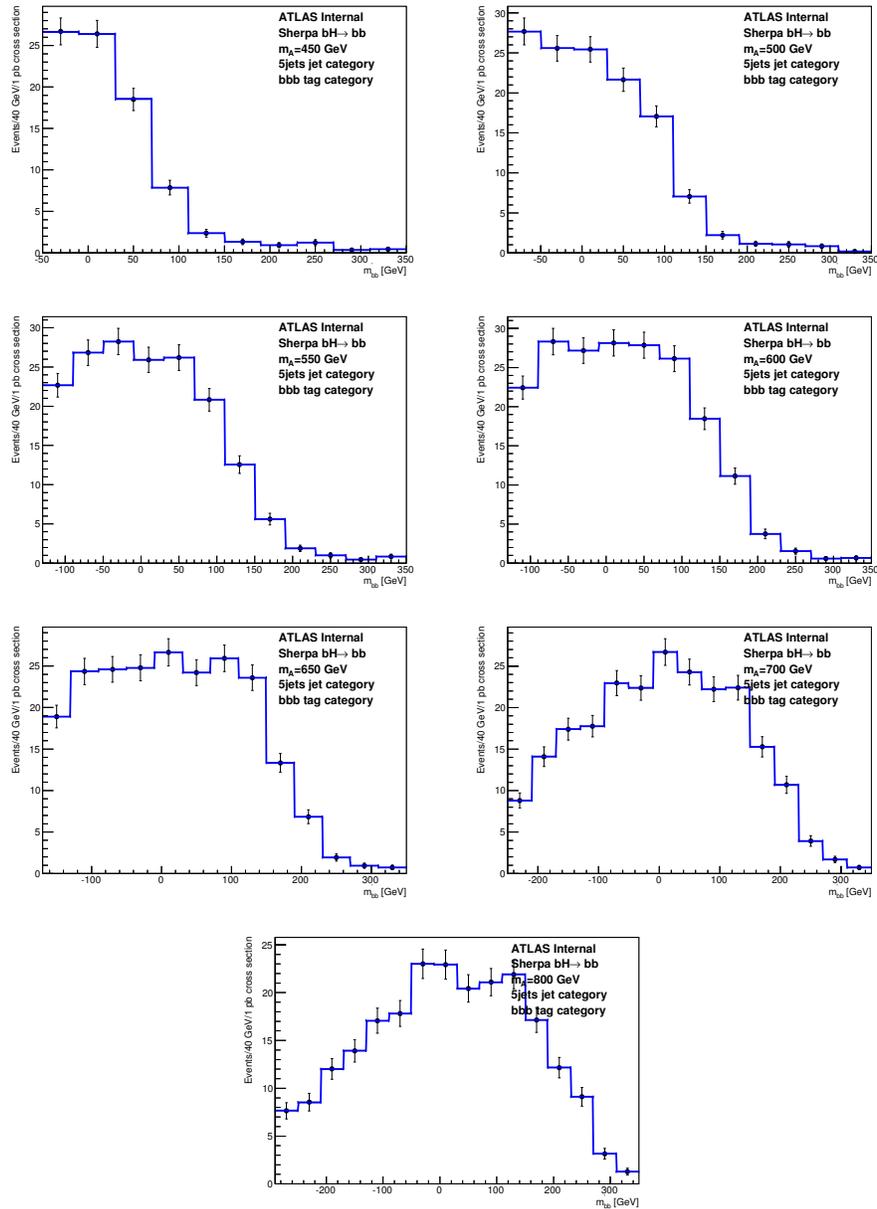


Figure 7.13: Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the bbb category, for events with 5 jets, for different H/A masses after the rotation has been applied.

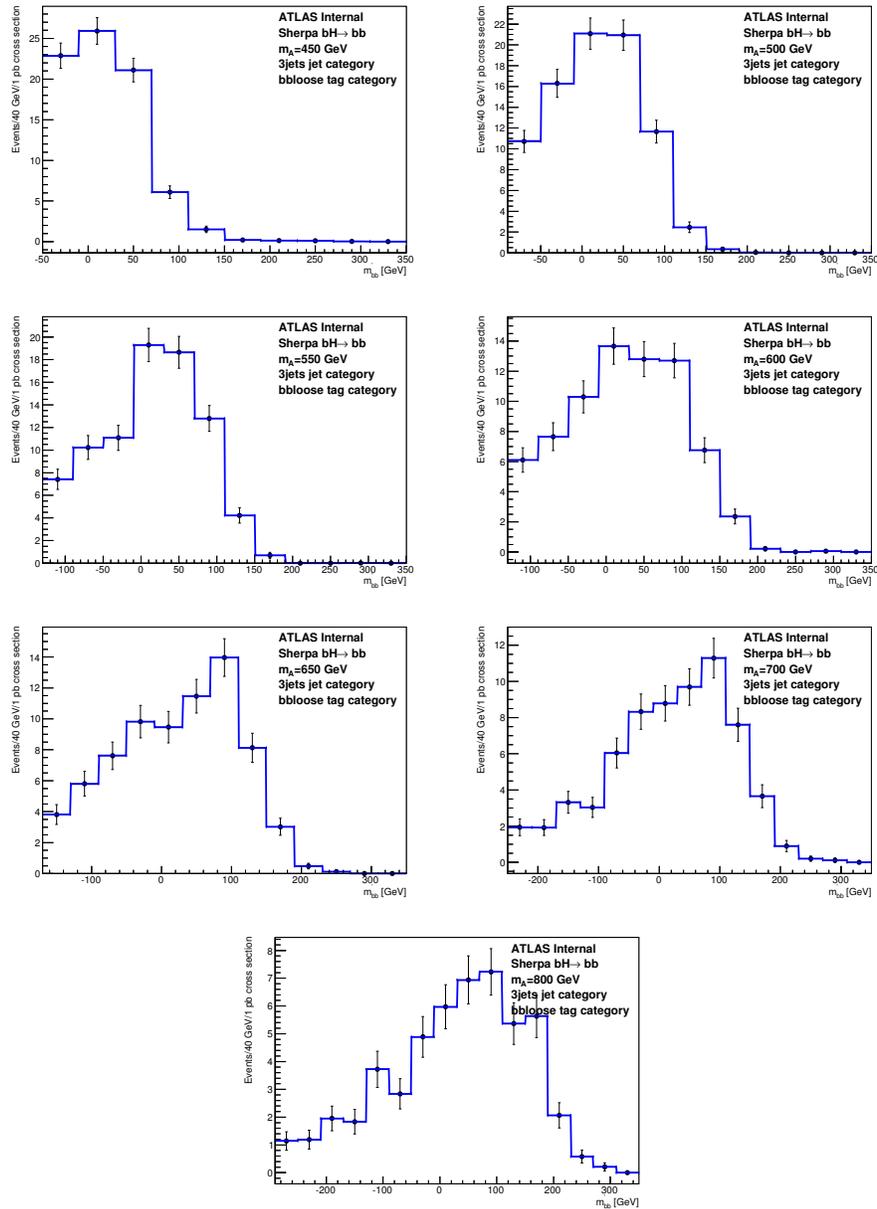


Figure 7.14: Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the *bbloose* category, for events with 3 jets, for different H/A masses after the rotation has been applied.

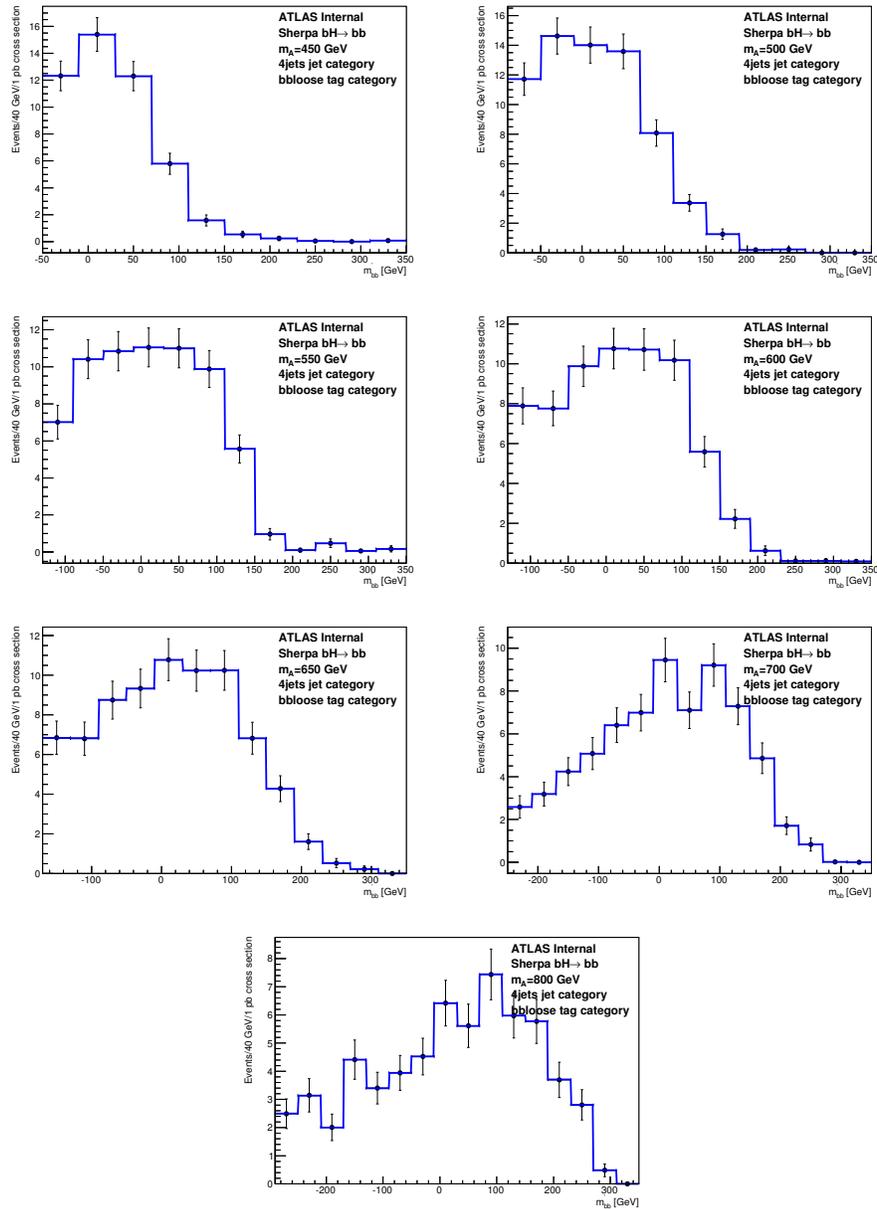


Figure 7.15: Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the *bbloose* category, for events with 4 jets, for different H/A masses after the rotation has been applied.

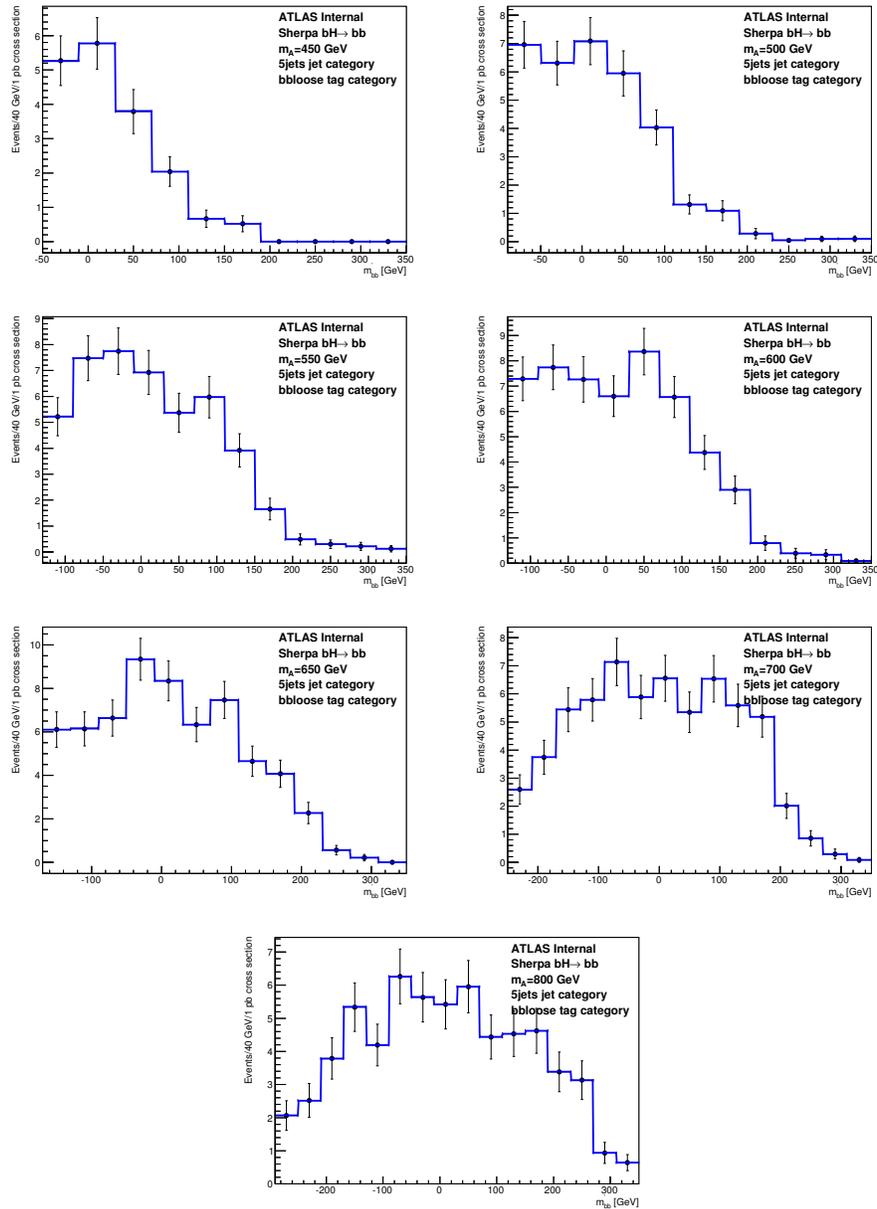


Figure 7.16: Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the *bbloose* category, for events with 5 jets, for different H/A masses after the rotation has been applied.

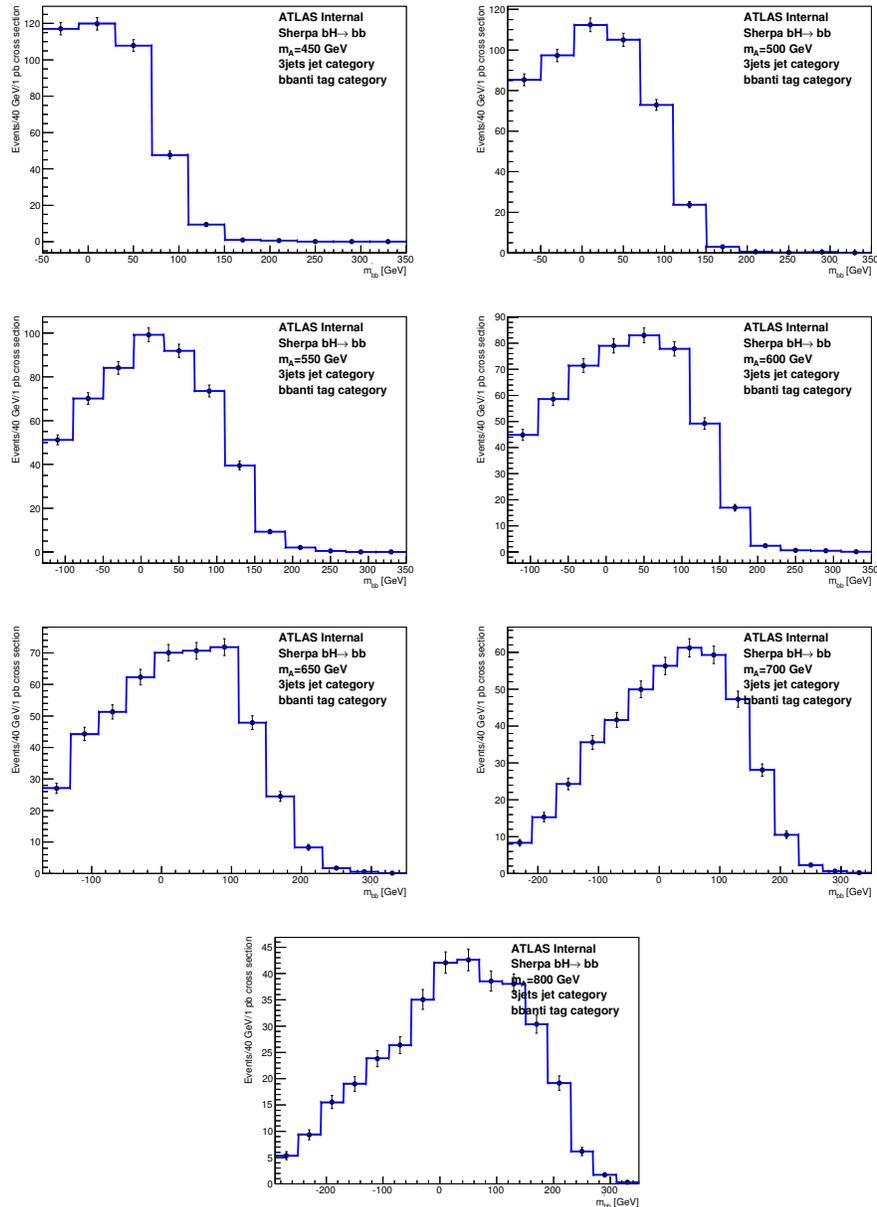


Figure 7.17: Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the $bbanti$ category, for events with 3 jets, for different H/A masses after the rotation has been applied.

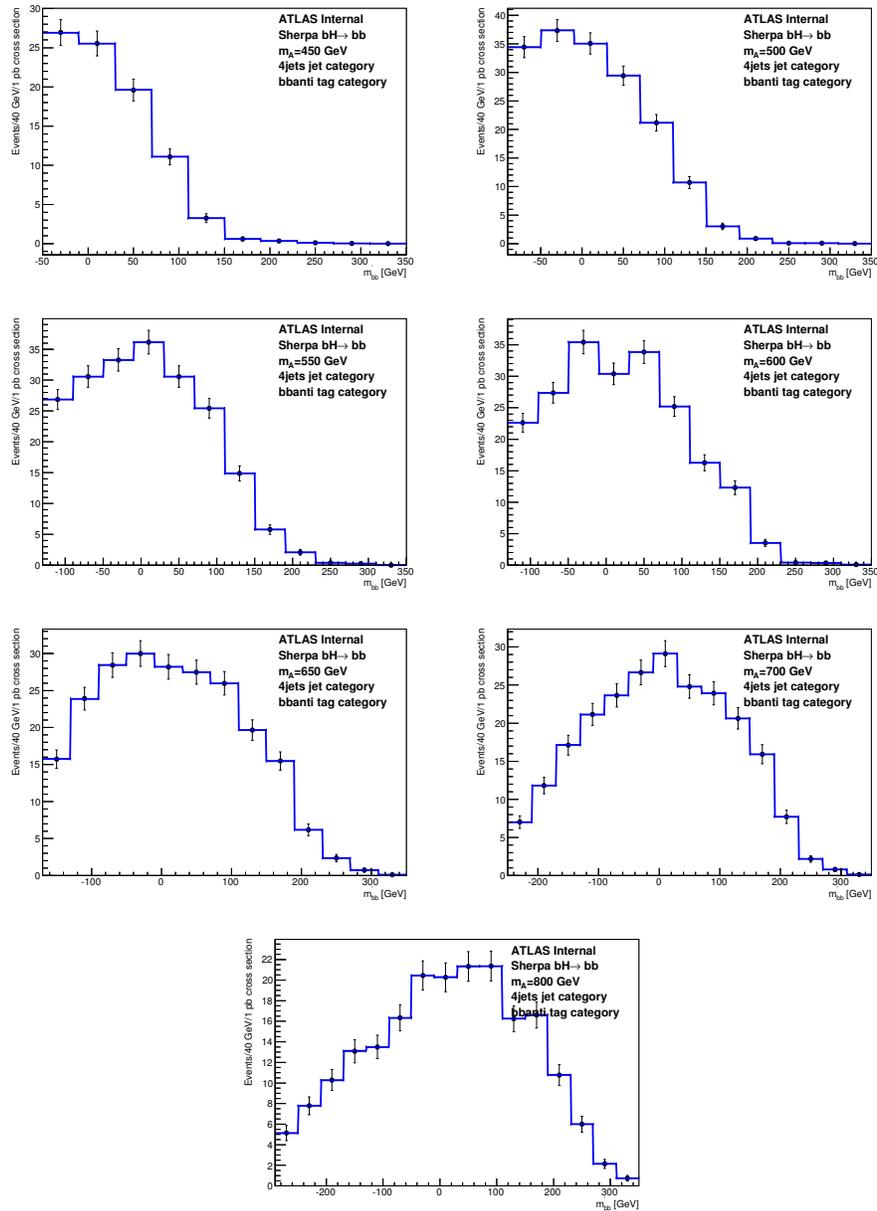


Figure 7.18: Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the $bbanti$ category, for events with 4 jets, for different H/A masses after the rotation has been applied.

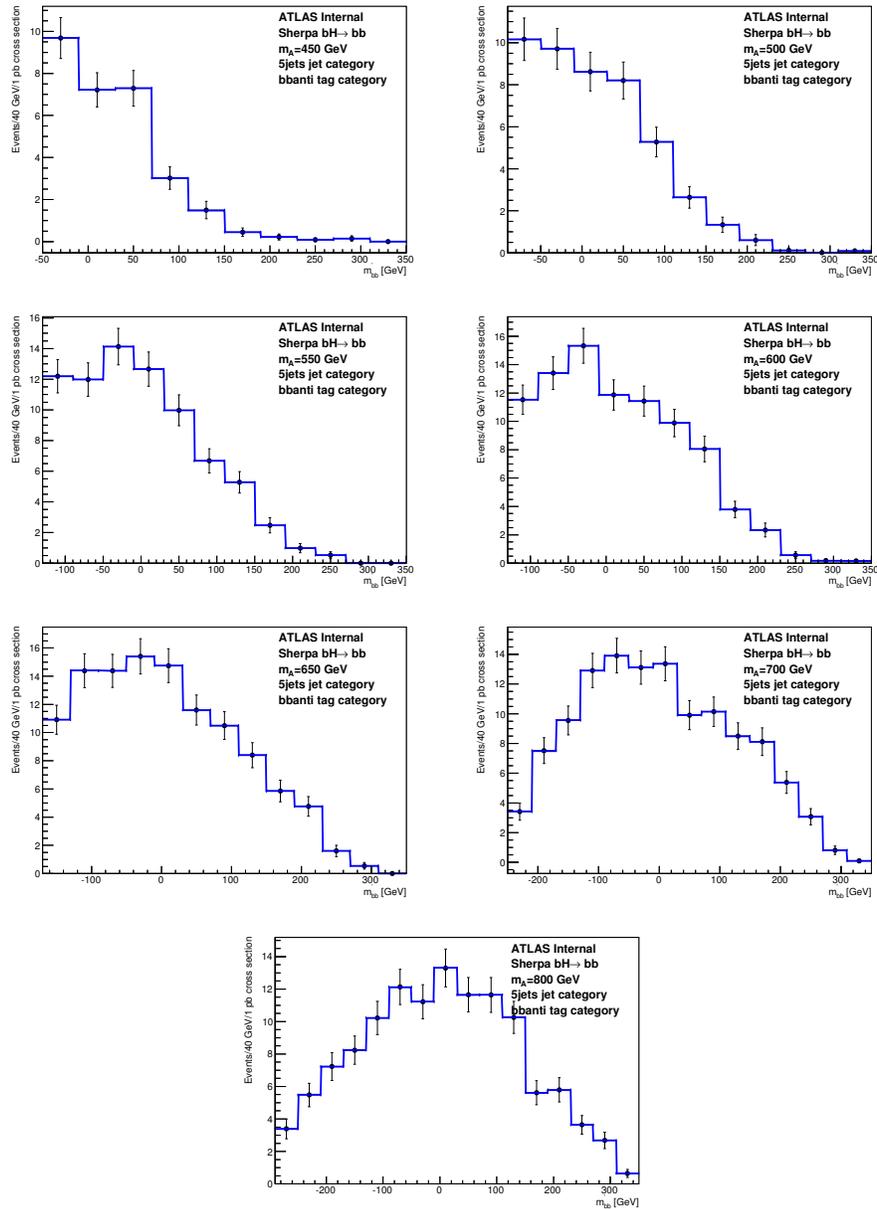


Figure 7.19: Signal MC distributions and RhhBinnedPdf PDFs for m_{bb} in the $bbanti$ category, for events with 5 jets, for different H/A masses after the rotation has been applied.

Chapter 8

Fit Strategy and Results

Once the data has been sent through the trigger and cuts, it is still expected that a large number of QCD events will remain, in which there might be a (probably comparatively small) number of signal events. The role of the fit is to mathematically describe the distribution of the m'_{bb} of all the remaining events, and to enable the possible extraction of a signal from the background.

This general strategy is possible because the signal events are coming from a resonance, the H/A particle, while the background consists of a smoothly falling spectrum owing to the fact that it comes from various QCD processes. That means that H/A , if they are present, will show up as a bump in the m'_{bb} spectrum. Using the signal MC, we define what we expect for the shape of the signal resonance (the normalization is given to us by nature, and is a free parameter that must be extracted, potentially with a value of zero indicating no sign of a signal-like excess found) while the background fit is completely data-driven. The m'_{bb} distribution is fit in the background-dominated *bbanti* control region together with the signal-enriched *bbb* and *bbloose* categories; an excess that is present in *bbb* but with no corresponding excess in *bbanti* is the hallmark of a potential signal. This fit strategy, in which the

bbanti distribution dominates the background expectation in the *bbb* signal region, is tested in background MC to verify that it does not introduce a major bias from the m'_{bb} shapes differing in the *bbb* and *bbanti* regions, and an extrapolation systematic uncertainty is added to account for any residual disagreements between the signal region and control region.

These fit results are then the ingredients for the limit-setting procedure, which is detailed in the next section.

8.0.1 Fit Model and Categories

The fit is performed in several different categories, which vary in the signal/background ratio, the signal and background shapes, and the absolute normalizations of the signal and background. In a situation like this, where there are several categories that can be defined and the search sensitivity varies depending on which category is being examined, it can benefit the overall search sensitivity to fit each category separately and then combine them at the end. While the parameteric form of the fit (for both signal and background) is the same in each category (histograms fit over the ranges shown in Table 7.3), the signal is fit separately in each n_{jets} and *b*-tag category. The background is fit separately in bins of n_{jets} , but the shape is constrained to be the same in all *b*-tag categories for a given number of jets up to an overall linear function that allows for linear variations in shape by *b*-tag/ n_{jets} category. Allowing only linear variations from category to category, and constraining the other background shape parameters to be the same in all *b*-tag categories, helps with convergence and has minimal chance of signal contamination biasing the fit, since in practice the background fit is heavily dominated by the high-statistics *bbanti* control region.

The statistical analysis of the data employs an unbinned likelihood function, defined

as:

$$\text{Pois}(N|\mu S + B) \prod_{k=1}^{n_{jet}} \prod_{l=1}^{cat} \prod_{i=1}^{N_l} [\mu N_{S,k,l} PDF_{sig,k,l}(m_{bb,i}) + N_{B,k,l} PDF_{bkg}(m_{bb,i})] \quad (8.1)$$

where:

- the product is over the b -tag categories l , the n_{jets} categories k , and over the events in each category i ;
- $N_{S,l}$ and $N_{B,l}$ are the expected signal and background yield in each category;
- $PDF_{sig,l}(m_{bb})$ and $PDF_{bkg}(m_{bb})$ are the signal and background probability density functions for the different categories;
- μ is a signal strength which multiplies the overall signal prediction.

The b -tag fit categories l are the three exclusive categories to which events are assigned based on the b -tag value of the third b -jet jet in the event (we call it the third b -jet because two b -jets have already been b -tagged in the trigger, and are assumed to be true b -jets). These are the categories outlined in Section 6.1: *bbb*, *bbloose*, and *bbanti*.

The second type of categorization is based on the number of jets in the event: 3, 4, or 5 or more jets. We find that the signal shape can change based on the number of events, as well as the overall signal and background normalizations. For more details on the effect of the number of jets on the signal distributions, see Section 7.3.5.

8.0.2 Signal Yield Tests in Signal+Background Scenario

An important cross-check is that the signal+background model returns the correct background and, especially, signal yields that are present in the dataset that is being

Table 8.1: The mean and RMS of the pull histograms from running 50 toy MC simulations per point of two mass points (450 and 700 GeV) and four signal cross sections (0, 0.1, 1.0, and 10.0 pb). To within the statistical error on the toy studies, which is about 15%, the toy MC studies do not point toward any major biases in the fitting process.

mass point	cross section injected	pull mean/RMS
450	0 pb	0.190±1.008
450	0.1 pb	0.012±0.9506
450	1 pb	0.059±0.9103
450	10 pb	0.016±1.028
700	0 pb	0.156±1.047
700	0.1 pb	0.008±1.108
700	1 pb	-0.151±0.998
700	10 pb	-0.104±1.093

fit. We need to check both the zero-signal case, verifying that the signal yield found by the fit is consistent with zero, and a number of nonzero signal cross sections, checking that the cross-section found by the fit comes back with minimal or no bias. The fit is checked by running toy MC simulations, where a signal of a given cross section is injected, a PDF is built out of the background plus signal, a dataset is generated using that PDF which is then re-fit with the signal histogram. We repeat this process 50 times for each signal mass point, and record the pull in a histogram (where pull is defined as $(sig_{injected} - sig_{fit})/\sigma_{stat,fit}$). For a bias-free fit, the pull histogram should have a mean of zero and an RMS of one.

The results of these studies are summarized in Table 8.1. We find that for larger signal cross sections, of 1 pb or more, there can be a pull bias of up to 15%. This bias varies by signal cross section and mass point, and in the scenario where a significant excess were observed, the mass point and cross section for that excess would be used as inputs into higher-statistics toy studies to precisely quantify how much of the excess might be attributable to pull bias.

Table 8.2: The signal cross sections found by the fit to maximize the overall likelihood across all tag and n_{jets} categories, and the statistical error on the cross section.

mass point	$\sigma \times \text{BR}$ returned by fit [pb]
450 GeV	-0.216 \pm 0.988
500 GeV	-0.165 \pm 0.596
550 GeV	-0.0135 \pm 0.514
600 GeV	0.324 \pm 0.426
650 GeV	-0.0394 \pm 0.356
700 GeV	-0.0832 \pm 0.216
800 GeV	-0.165 \pm 0.171

8.1 Search Results

The signal cross sections returned by the fit can be seen in Table 8.2, along with the associated statistical errors. All the signal cross sections are negative, although once the statistical errors are taken into account they are consistent with zero signal present (that is, neither an excess nor a deficit). In other words, these signal yields are more consistent with the background-only hypothesis than a signal+background hypothesis, but it remains the job of the limit-setting procedure detailed in the next section to precisely quantify that statement.

Another interesting parameter returned by the fit is the slope parameter between the bbb and the $bbanti$ categories, and the $bbloose$ vs. $bbanti$ categories, on a jet-bin-by-jet-bin basis. As detailed in Section 6.3.1, this parameter is constructed to be constrained to the range $[-1, 1]$, where -1 corresponds to the first bin having contents of approximately zero and the last bin having its contents doubled, while a slope parameter of 1 has the inverse interpretation. A slope parameter of zero indicates the same shape being found in all tag categories. The slope parameters returned are generally on the order of $1e-02$, which corresponds to a slope in the ratio of m'_{bb} distributions of approximately $1e-04$, which is the same order of magnitude found in background MC studies (Section 9.1). Those parameters and their associated

Table 8.3: The *bbb* category linear variation parameters returned by the fit, where a larger parameter indicates a further deviation from the same shape in the *bbb* and *bbanti* categories. In general, the variation is more significant in the 5+ jet bin than in the 3-jet or 4-jet bins.

mass point	3 jets <i>bbb</i>	4 jets <i>bbb</i>	5+ jets <i>bbb</i>
450 GeV	$-1.93\text{e-}2 \pm 2.75\text{e-}02$	$4.66\text{e-}02 \pm 2.45\text{e-}02$	$4.96\text{e-}02 \pm 2.53\text{e-}02$
500 GeV	$4.25\text{e-}4 \pm 2.89\text{e-}02$	$5.58\text{e-}02 \pm 2.53\text{e-}02$	$9.38\text{e-}02 \pm 2.59\text{e-}02$
550 GeV	$1.47\text{e-}02 \pm 3.17\text{e-}02$	$6.97\text{e-}02 \pm 2.66\text{e-}02$	$1.13\text{e-}01 \pm 2.57\text{e-}02$
600 GeV	$5.05\text{e-}02 \pm 3.95\text{e-}02$	$7.62\text{e-}02 \pm 3.33\text{e-}02$	$1.16\text{e-}01 \pm 3.03\text{e-}02$
650 GeV	$6.04\text{e-}02 \pm 4.14\text{e-}02$	$6.75\text{e-}02 \pm 3.52\text{e-}02$	$1.16\text{e-}01 \pm 3.03\text{e-}02$
700 GeV	$1.08\text{e-}01 \pm 3.97\text{e-}02$	$5.93\text{e-}02 \pm 3.60\text{e-}02$	$1.10\text{e-}01 \pm 3.03\text{e-}02$
800 GeV	$3.22\text{e-}01 \pm 3.79\text{e-}02$	$5.08\text{e-}01 \pm 2.19\text{e-}02$	$6.84\text{e-}01 \pm 1.41\text{e-}02$

Table 8.4: The *bbloose* category linear variation parameters returned by the fit, where a larger parameter indicates a further deviation from the same shape in the *bbloose* and *bbanti* categories.

mass point	3 jets <i>bbloose</i>	4 jets <i>bbloose</i>	5+ jets <i>bbloose</i>
450 GeV	$2.04\text{e-}03 \pm 3.13\text{e-}02$	$1.90\text{e-}02 \pm 2.98\text{e-}02$	$-4.85\text{e-}02 \pm 3.23\text{e-}02$
500 GeV	$-1.88\text{e-}02 \pm 3.40\text{e-}02$	$2.08\text{e-}02 \pm 3.05\text{e-}02$	$-5.17\text{e-}02 \pm 3.53\text{e-}02$
550 GeV	$-2.83\text{e-}02 \pm 3.78\text{e-}02$	$1.34\text{e-}02 \pm 3.30\text{e-}02$	$-4.93\text{e-}02 \pm 3.59\text{e-}02$
600 GeV	$3.50\text{e-}02 \pm 4.64\text{e-}02$	$-3.47\text{e-}02 \pm 4.34\text{e-}02$	$-1.08\text{e-}02 \pm 4.08\text{e-}02$
650 GeV	$4.50\text{e-}02 \pm 4.84\text{e-}02$	$-1.94\text{e-}02 \pm 4.54\text{e-}02$	$-1.06\text{e-}02 \pm 4.12\text{e-}02$
700 GeV	$5.97\text{e-}03 \pm 5.21\text{e-}02$	$-3.08\text{e-}02 \pm 4.67\text{e-}02$	$-1.25\text{e-}02 \pm 4.11\text{e-}02$
800 GeV	$1.10\text{e-}01 \pm 5.88\text{e-}02$	$3.34\text{e-}01 \pm 3.85\text{e-}02$	$8.62\text{e-}01 \pm 8.07\text{e-}03$

statistical errors are in Table 8.3 and 8.4; for most categories, the magnitudes of the slopes are the same order of magnitude as the statistical errors, suggesting that the shape differences between the tag categories are not statistically significant. However, especially the distributions with 5+ jets show a significant slope, which is consistent with what we observed in MC studies.

8.2 Limit Extraction

Once the signal and background have been fit, the CL_s method is used to determine the $\sigma \times BR$ that can be excluded at the 95% confidence level [46], [47], [48]. The CL_s method is a frequentist limit-setting procedure that is designed to give upper limits that are greater than the true value of $\sigma \times BR$ with a probability of at least 95%.

Once the signal and background PDF shapes are determined in the fit, the limit setting procedure proceeds as follows. The fits are used to generate an ensemble of toy MC datasets that are combined into a single representative dataset, called the “Asimov dataset” for that distribution. Then in order to obtain the nominal fit result in terms of μ^1 and σ_μ ,² the likelihood function is maximized with respect to all parameters. This is referred to as the maximized log-likelihood value MLL. The test statistic q_μ is then constructed according to the profile likelihood:

$$q_\mu = 2\ln(\mathcal{L}(\mu, \hat{\theta}_\mu)/\mathcal{L}(\hat{\mu}, \hat{\theta})), \quad (8.2)$$

where $\hat{\mu}$ and $\hat{\theta}$ are the parameters that maximize the likelihood (with the constraint $0 \leq \hat{\mu} \leq \mu$), and $\hat{\theta}_\mu$ are the nuisance parameter values that maximize the likelihood for a given μ . This test statistic is used to measure the compatibility of the background-only model with the observed data.

Expected and observed exclusion limits following this procedure are shown in Fig. 8.1 and Table 8.5. The signal cross sections that were found by the fit to maximize the global likelihood are compiled in Table 8.2. No significant excess of events is observed relative to the expected exclusion limits.

¹ μ is the parameter of interest, in this case the signal cross section

² σ_μ is the uncertainty on μ

Table 8.5: The expected and observed production cross sections times branching ratios to be excluded at 95% CL, along with the 1-sigma and 2-sigma statistical uncertainties

mass point	observed limit (pb)	expected limit (pb)	stat error [pb]	syst error [pb]
450 GeV	1.797	1.910	0.92	0.21
500 GeV	1.097	1.105	0.50	0.09
550 GeV	0.990	0.910	0.34	0.21
600 GeV	1.114	0.767	0.27	0.25
650 GeV	0.659	0.659	0.19	0.19
700 GeV	0.406	0.505	0.13	0.17
800 GeV	0.250	0.335	0.10	0.10

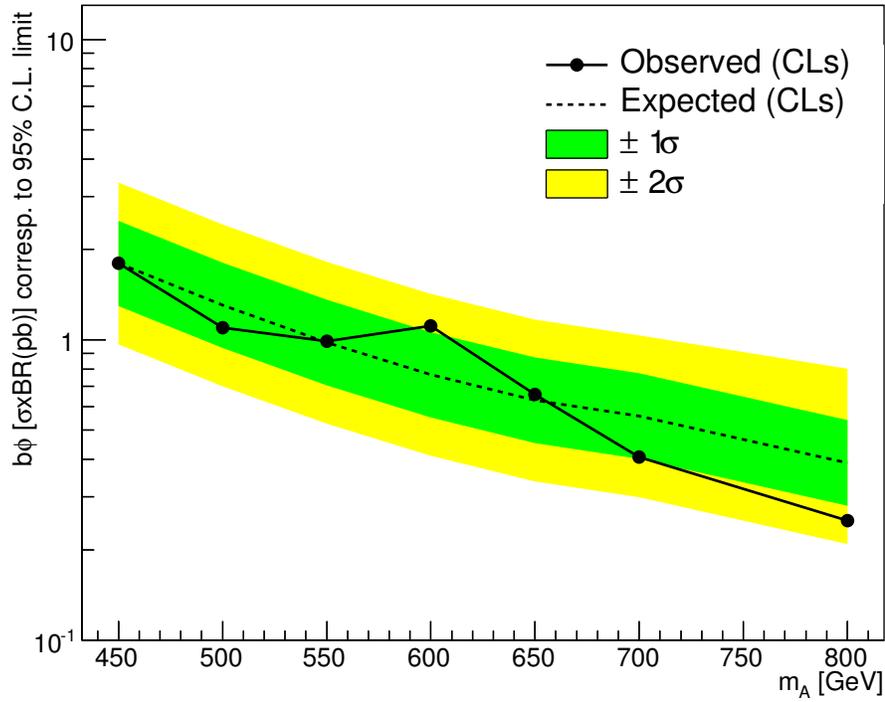


Figure 8.1: The expected 95% excluded $(H + A) \sigma \times \text{BR}$ as a function of m_A , for $450 < m_A < 800$ GeV.

Chapter 9

Systematic Uncertainties

The last step in computing a result in experimental physics involves identifying and quantifying sources of systematic error. Systematic errors arise from calibration errors, data/MC disagreement, detector uncertainty, noise, and other sources that do not decrease as the overall dataset size increases.

Since we use a data-driven background estimate, most of the systematic errors are only relevant for the signal shape and normalization, which we study using MC simulations. For example, b -tagging systematics, jet energy systematics, and trigger turn-on systematics only apply to signal and not to background. However, the shape extrapolation from the $bbanti$ control region to the bbb signal region is one important source of background uncertainty; we validate the extrapolation and compute systematic errors for it using QCD MC simulations.

9.1 Background Shape Variation Systematics

The background shape is data-driven, and is dominated by the *bbanti* m'_{bb} distribution in data, but the assumptions about the relative shape of the *bbanti* background control region compared to the *bbb* signal region needs to be validated in MC. In general, these validation studies show clearly that the m'_{bb} shape agrees closely between the *bbanti* and *bbb* regions (validation plots can be found in Section 4.3.1), but any small differences must be quantified using the linear shape variation parameter in the fit, and the uncertainty on that parameter is then a source of systematic error.

In order to compare the shapes, we calculate the ratio between the m'_{bb} distributions in the *bbb* and *bbanti* samples in QCD MC simulation datasets. Then we fit these ratios with a linear function, looking at the slope of the line to see if it is significantly different from zero, indicating a systematic shape difference. These fitted ratio plots can be seen in Figures 9.1 and 9.2. Clearly the largest shape uncertainty arises in the 5 jet category; we believe that in the higher-jet categories (four-jet and five-jet) there is a combinatorial effect where several jets are available to provide the “third” *b*-tag required for the event to be classified in the *bbb* category and that has a slight effect on the m'_{bb} distribution relative to the three-jet category.

In order to maintain the data-driven spirit of the background estimation, these MC fits are primarily used to validate the assumption that a linear variation allows for the an appropriate amount of freedom in the fit. Put another way, these MC studies validate that a linear variation is a suitable approach, compared to another functional form like quadratic. Based on the χ^2/DOF of these fits, and studies of the sensitivity given the linear assumption vs. a quadratic assumption, we find that a linear function describes the ratio between the distributions well while constraining the background shape not to allow for background wiggles in the signal region that could cover up a possible signal. The magnitude and sign of these slopes is taken directly from the data in the fitting process, and the statistical uncertainty on that parameter is taken

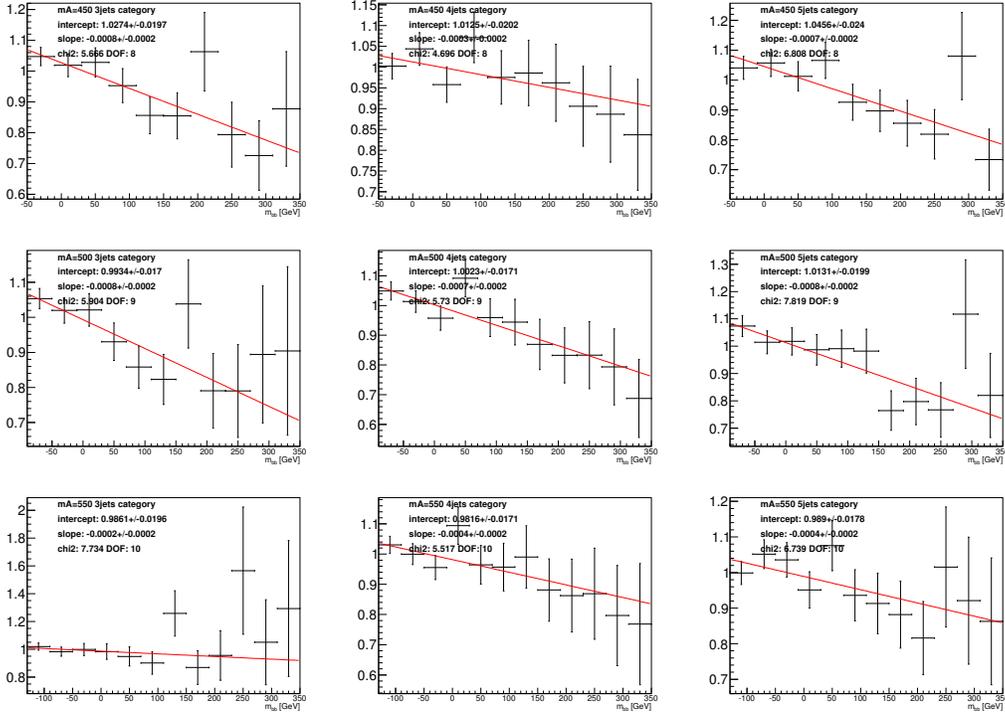


Figure 9.1: The ratio of the m'_{bb} distributions in QCD MC simulations of $bb\bar{t}$ and bbb events, fitted with a linear function, for the 3-jet, 4-jet and 5-jet categories for $m_A = 450, 500, \text{ and } 550$ GeV. The slope and intercept of the best fit line are printed on the plots, along with their error and the χ^2 and number of degrees of freedom of the fit.

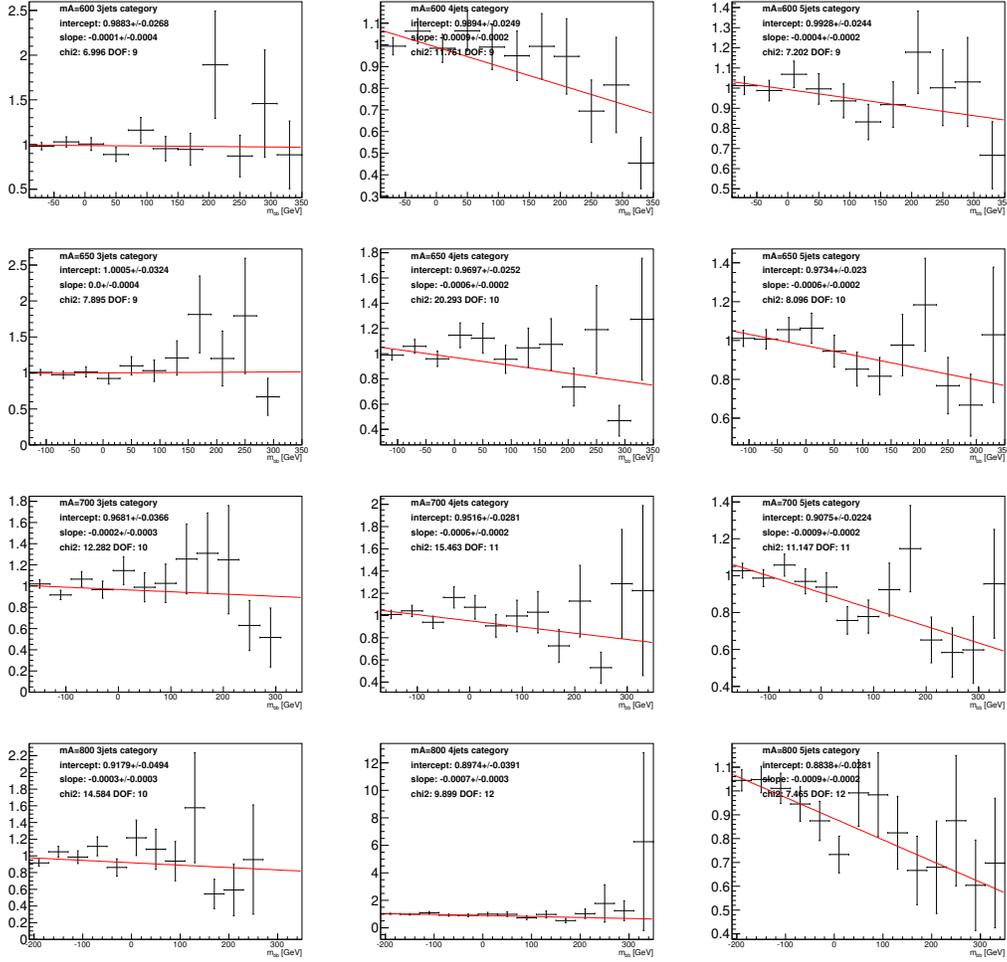


Figure 9.2: The ratio of the m'_{bb} distributions in QCD MC simulations of $bb\bar{anti}$ and bbb events, fitted with a linear function, for the 3-jet, 4-jet and 5-jet categories for $m_A=600-800$ GeV. The slope and intercept of the best fit line are printed on the plots, along with their error and the χ^2 and number of degrees of freedom of the fit.

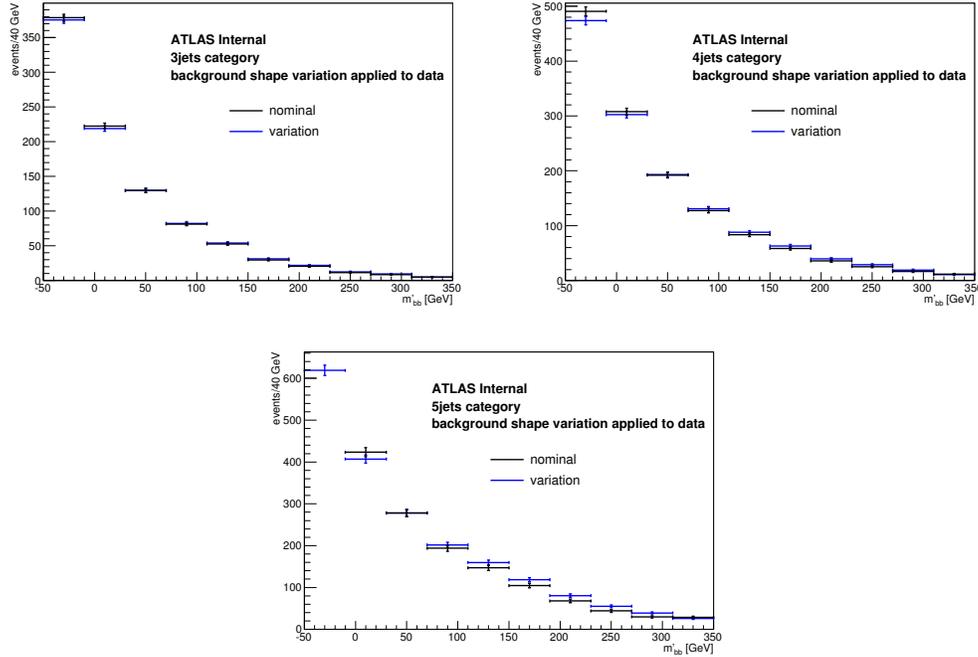


Figure 9.3: The $bbanti$ m'_{bb} distribution in data, for the nominal case (i.e. no adjustments made) and with a systematic variation applied based on the shape differences between $bbanti$ and bbb , as quantified in Figures 9.1 and 9.2.

as a systematic variation.

To study the size of this uncertainty, we use the background MC simulations as a starting point, and fit the ratio of the bbb and $bbanti$ distributions of m'_{bb} with a linear function (Figure 9.1 and 9.2). Then we take the $bbanti$ distribution, vary it by the linear function returned by the fit, rescale it to the same number of events as the bbb distribution, and use the resulting distribution in the fit to both check that the fit can return an appropriate result and to quantify the statistical errors on the shape variation parameter. Representative distributions reweighted $bbanti$ distributions can be seen in Figure 9.3, alongside the same distributions but without any shape variations applied.

Since all the other systematics are corrections to MC simulations, and MC events

are only used directly to model the signal (the background estimate is data-driven), this is the leading systematic of the analysis. The effect of this systematic on the final limits, organized by signal mass point, is listed in Table 8.5; it ranges between 0.09 and 0.25 pb. This is smaller than the statistical errors for all mass points except 700 and 800 GeV.

9.2 Trigger p_T Efficiency

The trigger simulation can introduce systematic effects that must be quantified as uncertainties; this can happen in two ways on this analysis. Since the trigger efficiency is estimated for signal using MC simulations, it is possible that detector mismodelings in the MC simulation process can over- or under-estimate the trigger efficiency as a function of the p_T of the jets in the event. We call this the trigger p_T efficiency, and we estimate and correct for it by parameterizing the turn-on curve with a logistic function, and using the ratio of the function between data and MC to calculate and apply scale factors that get applied to the MC events falling in the turn-on curve. The second source of trigger mismodeling comes from the online b -tag that is included in the trigger, but because of correlations between the online and offline tagging, we estimate this uncertainty as part of the b -tagging systematic (Section 9.4).

The general idea behind the trigger turn-on curve is that the overall trigger efficiency is more volatile if one or both of the relevant jets are near the p_T thresholds. Since there are two jets in the EF_2b35_loose_j145_j35_a4tchad trigger, with thresholds of 145 and 35 GeV respectively, events with jets near these values should be assessed for systematic errors. However, since this is a multi-object trigger (requiring b -tags in addition to the two jets), the efficiency for a given event can be a complex function of the p_T of the leading jet, the second jet, and the b -tagging characteristics of the jets in the event. In order to factorize out these effects, we place tight offline cuts

on all objects in the trigger *except* the object under examination. For example, to compute the trigger turn-on curve for the leading jet, we place the following cuts before computing any efficiencies:

- at least 60 GeV p_T for the second jet
- two jets in the event must pass L2 and EF b -tags

These cuts effectively remove inefficiencies due to the second jet or the b -tagging, since the second jet p_T cut (60 GeV) is well into the efficiency plateau for that jet, and the requirement of two jets being b -tagged online removes events where the jet p_T is high enough, but the event fails to trigger because of the b -tagging requirement. In this discussion the same algorithm is applied to quantify the trigger turn-on of both the leading and sub-leading jet, but for brevity we will refer only to the leading jet. The parameterized turn-on curves for the leading and sub-leading jets in signal MC and data can be found in Figure 9.4.

For signal MC we can then compute the efficiency as a function of the leading jet p_T by examining the leading jet p_T of events that both do and do not pass the trigger. Residual inefficiencies remaining from the second jet p_T , and the b -tagging requirements, would show up in the form of the distribution plateauing slightly below 100% efficiency; the logistic function will identify the plateau efficiency and the residual error can be corrected when computing the trigger efficiency for a given jet p_T . The logistic function used for fitting has the mathematical form $\epsilon = \frac{b}{1+e^{c-a*p_T}}$. The values found by ROOT for the parameters of this function are as follows:

```
----- turn-on for j145 jet -----
FCN=485.311 FROM MIGRAD   STATUS=CONVERGED   229 CALLS   230 TOTAL
EDM=2.12809e-08   STRATEGY= 1   ERROR MATRIX ACCURATE
EXT PARAMETER
NO.  NAME      VALUE      ERROR      STEP      FIRST
    NAME      VALUE      ERROR      SIZE      DERIVATIVE
 1  a          7.02887e+01  4.79927e-01  1.69260e-04 -1.12470e-02
 2  b          9.32043e-01  7.98882e-04  8.34460e-06  1.40194e-01
 3  c          4.76427e-01  3.29985e-03  1.16262e-06  1.61881e+00
```

For the turn-on curves in the data, the procedure is slightly different. Unlike in signal MC simulations, there is no direct way of querying which events would have passed the trigger, but failed because the leading jet p_T was not above threshold; those events are simply rejected by the trigger and never recorded. In that sense, the data collected by the trigger is biased. However, some heavily prescaled trigger items with minimum or zero bias are collected for the express purpose of trigger calibrations. The ones that are used in this analysis are:

- EF_rd0_filled_NoAlg: the L1 trigger is fired randomly, then the event proceeds through the normal L2 and EF reconstruction without any decision being taken
- EF_j145_a4tchad: the p_T requirement on the leading jet is the same as in the analysis trigger, but no second jet p_T or b -tagging requirement is in place
- EF_j35_a4tchad: the same as EF_j145_a4tchad, but with the appropriate p_T threshold to calibrate the second jet

Only the EF_rd0_filled_NoAlg trigger is truly unbiased, but because of its heavy prescale, there are no events in data that are accepted by both this trigger and the EF_2b35_loose_j145_j35_a4tchad analysis trigger. However, we can do a two-step calibration using EF_j145_a4tchad and EF_j35_a4tchad as intermediate steps, as those have plenty of events that overlap with both the analysis trigger and the zero-bias trigger. Comparing signal MC simulations to data turn-on curves derived from this bootstrap procedure, we find that while our $H \rightarrow b\bar{b}$ MC simulations do not perfectly reproduce the turn-on curves seen in data, the effect of the difference on the final signal efficiency is negligible. This is true even for the low m_A values, like 450 GeV, where the effect is expected to be the strongest (because the p_T of the jets will be lower for the lower-mass Higgs particles, leading to more jets falling in or near the turn-on curves). However, in an effort to be conservative and account for any correlation effects (between jets, since this is a multi-object trigger) that might not be completely factored out, we can look at the disagreement in the plateau region,

where the data and signal MC show more disagreement than near the cut point of (in the case of the leading jet) 155 GeV. For a 400 GeV Higgs boson, this comparison is shown in Figure 9.4 and amounts to about a 3% difference. When this is propagated through to the final m'_{bb} distribution, only a fraction of events fall near enough the leading jet turn-on curve for signal MC modeling effects to be relevant, and the effect of this systematic on the final signal shape and normalization is negligible.

In practice, only the turn-on curve for the leading jet will have any practical effect on the signal MC simulation efficiency, since there is virtually no effect on the signal efficiency arising from a p_T cut on the second jet. In effect, the leading two signal jets are well-balanced enough (in terms of their p_T) that requiring the leading jet have a p_T of at least 145 GeV nearly always means that the second jet will have a p_T well above 35 GeV; this is established by looking at the cut chain as applied to signal MC simulations.

9.3 Jet Energy Uncertainties

Hadronic particles fragment via QCD in the ATLAS detector, leaving deposits of energy in the calorimeters that must be clustered and reconstructed into jets. The observed jets then need some correction so that, on average, the reconstructed jet energy corresponds to the energy of the associated stable particles. The calibration for this correction, called jet energy scale (JES), is calculated using MC simulations and then checked in data. The residual uncertainty on the JES is then a systematic error on the analysis.

There are a number of sources of JES uncertainty, each of which has an associated error. All the sources are listed here for completeness, but some are not used for reasons as noted:

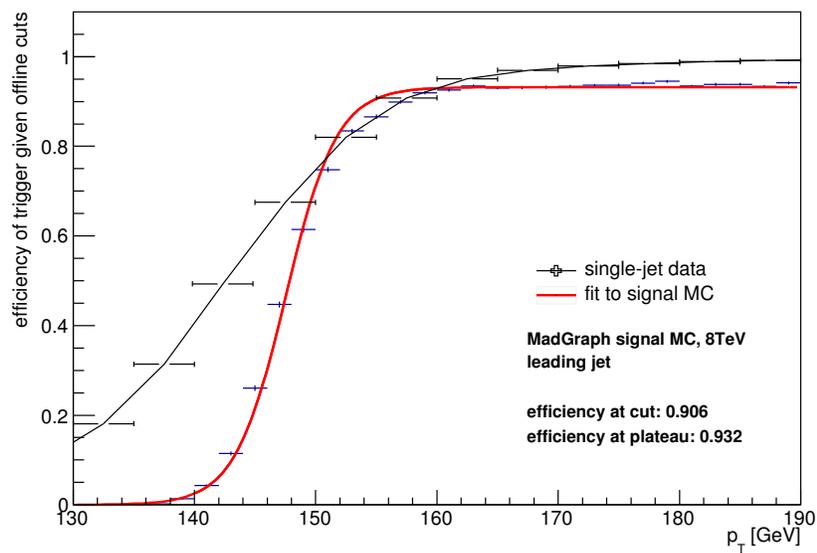


Figure 9.4: The p_T turn-on curves for the trigger for signal MC and data. Although this search uses a multi-object trigger, in which several conditions must simultaneously be true for a trigger acceptance, tight offline cuts can be used to isolate the effect of a single jet's p_T on the efficiency. The signal curve is fit with a logistic function, which is evaluated at a given cut point to extract the trigger efficiency at that point, and to adjust for any residual effects of other trigger objects that affect the efficiency.

- 6 nuisance parameters from in-situ analyses (Z+jet balance, photon+jet balance, and multi-jet balance) that are reduced from an original list of 60 that encompasses effects such as uncertainties in calorimeter response, the JES calibration method, and modelling in Monte Carlo generators
- 2 nuisance parameters from η intercalibration
- 4 nuisance parameters associated with pileup (relative offsets for NPV and μ , the p_T of the pileup, and the ρ ¹ topology)
- 1 nuisance parameter that quantifies the flavor composition and flavor response associated with b -jets

The jet energy scale and its associated uncertainties are estimated using in situ and test-beam-based measurements of isolated hadron response [49]. The systematic errors are integrated into the analysis via 14 independent variations, which are multiplicative factors applied to the jet 4-vectors and then the adjusted 4-vectors are run through the analysis framework. The effect of the jet energy scale varies depending on the signal mass point, from about 3% per variation at 450 GeV to <1% at 800 GeV. Summing over all variations gives approximately an 11% JES uncertainty at 450 GeV (the uncertainties are summed in quadrature, since they are independent) down to about 4% at 800 GeV.

9.4 b -Tagging Scale Factors

Although every effort is made to accurately simulate the production rates and kinematic properties of b -quarks, as well as the performance of the ATLAS detector, it is difficult to imagine that Monte Carlo simulation perfectly represents the b -tagging efficiencies that might be found in data. At the same time, it is challenging to derive

¹in this context, ρ refers to the average energy density from pileup throughout the calorimeter

pure data-driven samples of b , c and light jets to which the b -tagging can then be applied, and used to derive the efficiencies:

$$eff_{b,c,light} = \frac{\text{number of tagged } b, c, \text{ light}}{\text{number of truth } b, c, \text{ light}} \quad (9.1)$$

Scale factors are computed by the flavor tagging performance group to quantify the difference between the data and MC efficiencies.

$$eff_{data} = eff_{MC} \times SF_{data} \quad (9.2)$$

The scale factors are applied as event weights to the MC events, where the value of the scale factor depends on the p_T , η , b -tagging working point, and truth flavor of the jet under examination. When there are multiple jets being examined for b -tags in a single event, the net event weight is a product of all the relevant scale factors. In this analysis, since only the first 5 jets are relevant for determining whether an event passes the analysis cuts or which b -tag category (bbb , $bbloose$, or $bbanti$), no scale factors are applied for jets that are 6th or lower in the p_T hierarchy of an event.

9.4.1 Offline b -Tagging

The scale factors are calculated by comparing the b -tagging efficiency computed in MC to the efficiency measured in carefully curated data samples, where the truth flavor contents of the data are relatively well-known [15]. For the b -jet scale factors, a sample of $t\bar{t}$ events is used for the data component; for c -jets, the calibration is done with D^* events.

The scale factors are used to correct the MC b -tagging performance back to the

b -tagging behavior seen in data, but the calibration process can introduce systematic uncertainties that have to be quantified and propagated through to the final result. The b -tagging systematics are computed and applied using a method called the eigenvector method. The eigenvector method is used to derive a set of independent variations of the data-MC scale factors, in such a way to take the full covariance matrix of the calibration measurements into account, including correlations among working points and different jet p_T regions. For further explanation of the eigenvector method please see section 2.3 of [50].

The differences between the eigenvector variations and the nominal scale factors are combined in quadrature separately for each category in n_{jets} . We find that the effect on the normalization of the signal is 15-19%, depending on mass point and n_{jets} category.

9.4.2 b -Tagging in the Trigger

The presence of b -tagging in the trigger means the sample of data collected for this analysis is already enriched in jets that look more b -like, which must also be taken into account during the calibration process. The scale factors are computed separately based on how a jet is tagged (online, offline, or both). For each of these permutations, a different scale factor is retrieved (for example, jets that are tagged offline are separated from their non-tagged counterparts before the online scale factors are computed). Then the scale factors are combined for an overall event weight as usual.

When we examine the online scale factors for systematic uncertainties, the most important errors are the errors on jets that are b -tagged offline as well as online, since that describes the jets that are most important when identifying signal events. The systematic uncertainties on these scale factors are approximately 4% per jet, and there is not any significant relationship between the p_T of the jet and the online

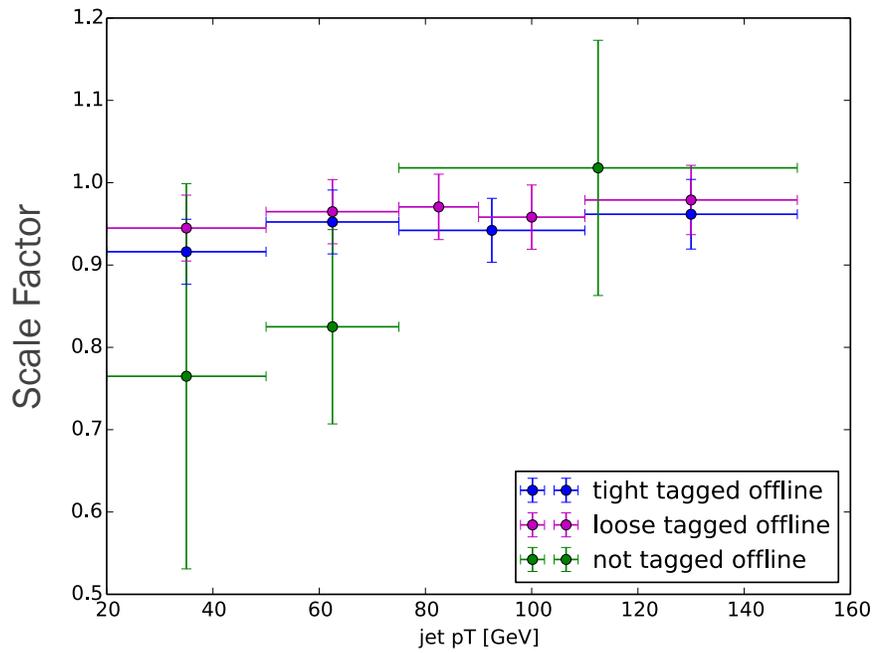


Figure 9.5: The online b -tagging scale factors as a function of jet p_T for jets categorized by how tightly they are tagged offline. For jets with a p_T higher than the maximum calibrated p_T point, we apply the scale factor for the highest p_T bin that has a calibration.

b -tagging uncertainty.

Chapter 10

Interpretation of Results

The last component of this search is asking how the sensitivity in terms of (the model-independent) quantity of production cross section times branching ratio can be interpreted as a sensitivity with respect to the MSSM parameters $\tan\beta$ and m_A . This interpretation introduces model dependence, as any interpretation does, and a number of standard benchmark scenarios are provided by the LHC Higgs Cross Section Working Group [51] which span a range of plausible MSSM scenarios including m_h^{max} and $m_h^{mod,\pm}$. These benchmark scenarios are a joint effort of ATLAS, CMS, and theorists. The benchmarks generally set the parameters of the model under inspection, and then calculate the production cross sections and branching fractions for that model; then each analysis can convert its exclusions in cross section times branching fraction to an exclusion in (usually) $m_A/\tan\beta$. The production cross sections are calculated using the Santander matching algorithm [52] and generally follow the scenario prescriptions suggested in [10].

10.1 m_h^{max} Scenario

The m_h^{max} scenario is one of the older and more “classic” scenarios for SUSY Higgs searches. It uses the value of X_t that maximizes m_h for large m_A and a given $\tan\beta$ value. Once the Higgs boson was discovered around 126 GeV, this scenario fell somewhat out of favor because it tends to have an m_h that is too high, around 140 GeV. In the era of LEP searches, of course, it was not yet known what the mass of the Higgs was, and that was as good a guess as any other. In addition, this scenario gives conservative bounds on m_A and $\tan\beta$.

In the post-Higgs discovery world, the m_h^{max} scenario still serves as a useful benchmark for its conservative limits, and for placing comparisons against legacy analyses that used this scenario in their interpretations.

10.2 m_h^{mod} Scenarios

There are straightforward modifications of the m_h^{max} scenario that allow for a larger amount of parameter space that is still consistent with $m_h=126$ GeV; one can reduce the amount of mixing in the stop sector to bring m_h down to a value that agrees better with experiment. In this scenario, the sign of X_t can be either positive or negative, with changes in the excluded region depending on which one is chosen¹.

The m_h^{mod+} and m_h^{mod-} scenarios are generated with the following setup:

- HIGLU for calculating the total Higgs production cross section due to gluon fusion [53]
- ggHNNLO for next to next to leading order gluon fusion calculations [54]

¹the scenario name encodes the sign of X_t ; $m_h^{mod,+}$ has a positive X_t while X_t is negative for $m_h^{mod,-}$

- bbHNNLO (5FS) for next to next to leading order b -quark associated production cross sections with the five-flavor scheme [55]
- bbH (4FS) scans for Higgstrahlung off bottom quarks in the 4-flavor scheme [56] [57]

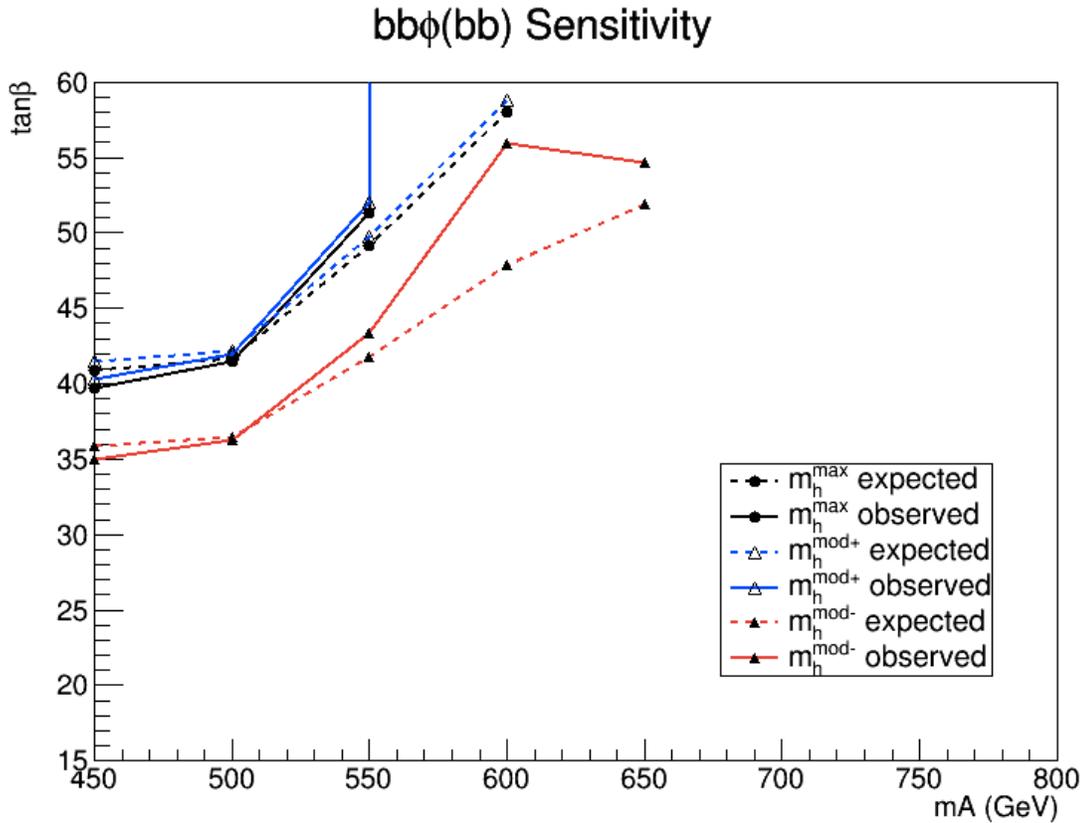


Figure 10.1: The observed and expected exclusion limits for the $bH/A \rightarrow b\bar{b}$ search, translated into exclusions in $m_A/\tan\beta$ under the m_h^{\max} and $m_h^{\text{mod},\pm}$ scenarios. The behavior of the observed exclusion at $m_A=650$ GeV is an artifact of the weaker limits that we set on that mass point, relative to the trend in its neighbors.

Our exclusions in $m_A/\tan\beta$ are plotted in Figure 10.1. The expected and observed

$\tan \beta$ exclusions range from 35-40 at 450 GeV to 60 at 600-650 GeV (depends on the exact MSSM scenario); perturbativity begins to break down above $\tan \beta \approx 60$ so interpretations are only provided up to that value. The expected and observed limits match closely for all mass points except $m_A=600$ GeV, where we see a larger deficit in the signal cross section returned by the fit, which propagates through to a weaker limit and weaker agreement between expected and observed MSSM exclusions.

In the final calculus, this search does not surpass searches for $bA/H \rightarrow b\tau\tau$ in terms of either sensitivity to production cross section times branching fraction, or exclusions in $m_A/\tan \beta$. This is for several reasons—at low m_A , the search in the $\tau\tau$ final state is advantaged by lower trigger p_T thresholds on the τ leptons than can be afforded for b -jets; at high m_A , the FSR from the b -quarks smears out the Higgs mass peak more severely than radiative effects in the $\tau\tau$ search. Additionally, this analysis has the constant challenge of combinatorial mismatches, where the jets can be incorrectly chosen when reconstructing H/A .

On the other hand, this search channel remains a very interesting one. The branching fraction to b -quarks is 5-7 times higher than the branching fraction to τ leptons, and with more clever triggering planned for the LHC Run 2², this search channel remains an interesting one to continue developing. Similarly, smart offline analysis advances such as a multivariate analysis or kinematic cuts that are more carefully tuned for the m_A values under investigation might yield more sensitive limits. Nonetheless, the search reported here (which is the first known search in this final state and mass range) reinforces the conclusions drawn by the $b\tau\tau$ analysis with respect to $m_A/\tan \beta$, covers new ground for a search in the $b\bar{b}$ final state, and lays an important piece of groundwork for studies in this channel going forward.

²specifically, topological triggers that look for the presence of muons in jets, allowing for a rough form of b -tagging at L1 of the trigger

Bibliography

- [1] Particle Data Group. Review of particle physics. *Chinese Physics C*, 38(9):161–214, 2014.
- [2] Cristobal Padilla on behalf of the ATLAS Collaboration. The atlas trigger system. *IEEE Transactions on Nuclear Science*, 57(2):650–657, April 2010.
- [3] S. P. Martin. A supersymmetry primer. *Adv.Ser.Direct.High Energy Phys.*, 21:1–153, 2010.
- [4] D. I. Kazakov. Beyond the standard model (in search of supersymmetry). *hep-ph/0012288 v2*, 2000.
- [5] T. Hahn, S. Heinemeyer, W. Hollik, H. Rzehak, G. Weiglein. High-precision predictions for the light cp-even higgs boson mass of the mssm. *Phys.Rev.Lett.*, 112(114):141801, 2014.
- [6] The Higgs Boson Masses and Mixings of the Complex MSSM in the Feynman-Diagrammatic Approach. M. Frank, T. Hahn, S. Heinemeyer, W. Hollik, H. Rzehak, G. Weiglein. *JHEP*, 2007(02):47, 2007.
- [7] G. Degrandi, S. Heinemeyer, W. Hollik, P. Slavich, G. Weiglein. Towards high-precision predictions for the mssm higgs sector. *European Physical Journal C*, 28:133–143, 2003.

- [8] S. Heinemeyer, W. Hollik, G. Weiglein. The masses of the neutral cp-even higgs bosons in the mssm: Accurate analysis at the two-loop level. *European Physical Journal C*, 9:343–366, 1998.
- [9] S. Heinemeyer, W. Hollik, G. Weiglein. Feynhiggs: a program for the calculation of the masses of the neutral cp-even higgs bosons in the mssm. *Comput.Phys.Commun.*, 124:76–89, 2000.
- [10] M. Carena et. al. Mssm higgs boson searches at the lhc: Benchmark scenarios after the discovery of a higgs-like particle. *European Physical Journal C*, 73:2552, 2013.
- [11] J. Ellis. Higgs physics. *arXiv: hep-ph/1312.5672*, 2013.
- [12] M. Cacciari, G. Salam, and G. Soyez. The anti- k_t jet clustering algorithm. *JHEP*, 0804:063, 2008.
- [13] The ATLAS Collaboration. Performance of the atlas detector using first collision data. *JHEP*, 1009:056, 2010.
- [14] The ATLAS Collaboration. b-jet tagger data to simulation comparison using 8 tev data. *ATL-COM-DAQ-2013-018*, 2013.
- [15] The ATLAS Collaboration. Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 atlas data. *ATLAS-CONF-2014-046*, 2014.
- [16] S. Weinberg. A model of leptons. *Phys.Rev.Lett.*, 19:1264–1266, 1967.
- [17] P. W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys.Lett.*, 12:132–133, 1964.
- [18] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys.Rev.Lett.*, 13:321, 1964.

- [19] G. Guralnik, C. Hagen, and T. Kibble. Global conservation laws and massless particles. *Phys.Rev.Lett.*, 13:585, 1964.
- [20] J. Goldstone, S. Weinberg and A. Salam. Broken symmetries. *Physical Review*, 127:965, 1962.
- [21] S. Glashow. Partial symmetries of weak interactions. *Nuclear Physics*, 22:579–588, 1961.
- [22] M. Carena et. al. LHC Discovery Potential for Non-Standard Higgs Bosons in the 3b Channel. *JHEP*, 1207:091, 2012.
- [23] LHCb Collaboration. Measurement of the $b_s^0 \rightarrow \mu^+\mu^-$ branching fraction and search for $b^0 \rightarrow \mu^+\mu^-$ decays at the lhcb experiment. *Phys.Rev.Lett.*, 111:101805, 2013.
- [24] The ATLAS Collaboration. Search for charged higgs bosons decaying via h to tau v in fully hadronic final states using pp collision data at sqrt(s) = 8 tev with the atlas detector. *JHEP*, 1503:088, 2015.
- [25] CDF Collaboration. Search for higgs bosons produced in association with b-quarks. *Phys.Rev.D*, 85:032005, 2012.
- [26] D0 Collaboration. Search for neutral higgs bosons in the multi-b-jet topology in 5.2 ifb of ppbar collisions at sqrt(s)=1.96 tev. *Phys.Lett.B*, 698:97–104, 2011.
- [27] The CMS Collaboration. Search for a higgs boson decaying into a b-quark pair and produced in association with b quarks in proton-proton collisions at 7 tev. *Phys.Lett.B*, 722:207–232, 2013.
- [28] T. Lari. Measurements of spatial resolution of atlas pixel detectors. *Nucl. Instrum. Meth. A*, 465:112–114, 2000.

- [29] B. M. Demirkoz. *Construction and Performance of the ATLAS SCT Barrels and Cosmic Tests*. PhD thesis, Oxford University, 2007.
- [30] The ATLAS Collaboration. The atlas experiment at the cern large hadron collider. *JINST*, 3:S08003, 2008.
- [31] The ATLAS Collaboration. Alignment of the ATLAS Inner Detector Tracking System with 2010 LHC proton-proton collisions at $\sqrt{s} = 7$ TeV. *ATLAS-CONF-2011-012*, 2011.
- [32] The ATLAS Collaboration. Atlas calorimeter performance technical design report. *CERN-LHCC-96-40*, 1996.
- [33] T. Cornelissen et al. The new atlas track reconstruction (newt). *Journal of Physics: Conference Series*, 119:032014, 2008.
- [34] The ATLAS Collaboration. Performance of the atlas inner detector track and vertex reconstruction in the high pile-up lhc environment. *ATLAS-CONF-2012-042*, 2012.
- [35] The ATLAS Collaboration. Pile-up corrections for jets from proton-proton collisions at $s=\sqrt{7}$ tev in atlas in 2011. *ATLAS-CONF-2012-064*, 2012.
- [36] Viviana Cavaliere, on behalf of the ATLAS Collaboration. b-jet triggering in atlas. *Journal of Physics: Conference Series*, 396:022008, 2012.
- [37] M. Lehmacher. b-tagging algorithms and their performance at atlas. *ATL-PHYS-PROC-2008-052*, 2008.
- [38] G. Piacquadio and C. Weiser. A new inclusive secondary vertex algorithm for b-jet tagging in atlas. *Journal of Physics: Conference Series*, 119:032032, 2008.
- [39] J. Alwall, et.al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations.

- JHEP*, 1407:079, 2014.
- [40] T. Sjostrand, S. Mrenna, and P. Z. Skands. Pythia 6.4 physics and manual. *JHEP*, 0605:026, 2006.
- [41] W. Lukas. Fast simulation for atlas: Atlfast-ii and isf. *Journal of Physics: Conference Series*, 396:022031, 2012.
- [42] S. Agostinelli et. al. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research*, 506(3):250–303, July 2003.
- [43] Allison, J. et. al. Geant4 developments and applications. *IEEE Transactions on Nuclear Science*, 53(1):270–278, February 2006.
- [44] T. Sjöstrand, S. Mrenna, and P. Skands. A brief introduction to pythia 8.1. *Comput.Phys.Commun.*, 178:852–867, 2008.
- [45] T. Gleisberg et. al. Event generation with sherpa 1.1. *JHEP*, 0902:007, 2009.
- [46] A. L. Read. Presentation of search results: The cl(s) technique. *J. Phys. G: Nucl. Part. Phys.*, 28:2693–2704, 2002.
- [47] G. Cowan, K. Cranmer, E. Gross and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. *European Physical Journal C*, 71:1554, 2011.
- [48] The ATLAS Collaboration. Expected performance of the atlas experiment—detector, trigger and physics. *arXiv:0901.0512 [hep-ex]*, 2009.
- [49] The ATLAS Collaboration. Jet energy measurement with the atlas detector in proton-proton collisions at $\sqrt{s} = 7$ tev. *European Physical Journal C*, 73(3):2304, 2013.
- [50] G. Aad et. al. b -tagging for vh , with $h \rightarrow bb$. *ATL-COM-PHYS-2013-697*, 2013 [ATLAS internal].

- [51] MSSM Cross Section Working Group, <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHX>
- [52] R. Harlander, M. Krämer, and M. Schumacher. Bottom-quark associated higgs-boson production: reconciling the four- and five-flavour scheme approach. *arXiv: 1112.3478*, 2011.
- [53] M. Spira. Higlū: A program for the calculation of the total higgs production cross section at hadron colliders via gluon fusion including qcd corrections. *arXiv: hep-ph/9510347*, 1995.
- [54] R. V. Harlander and W. B. Kilgore. Next-to-next-to-leading order higgs production at hadron colliders. *Phys.Rev.Lett.*, 88:201801, 2002.
- [55] R. V. Harlander and W. B. Kilgore. Higgs boson production in bottom quark fusion at next-to-next-to-leading order. *Phys.Rev.D*, 68:013001, 2003.
- [56] S. Dittmaier, M. Kramer, and M. Spira. Higgs radiation off bottom quarks at the fermilab tevatron and the cern lhc. *Phys.Rev.D*, 70:074010, 2004.
- [57] S. Dawson, C. B. Jackson, L. Reina, and D. Wackerroth. Exclusive higgs boson production with bottom quarks at hadron colliders. *Phys.Rev.D*, 69:074027, 2004.