

1 SEARCH FOR HIGGS BOSON PAIR PRODUCTION IN THE  $B\bar{B}WW^*$   
2 CHANNEL IN  $\sqrt{s} = 13$  TeV PROTON-PROTON COLLISIONS AT THE  
3 LARGE HADRON COLLIDER USING THE ATLAS DETECTOR

4 by

# JOHNCHRISTOPHER STEPHEN MYERS

# A DISSERTATION

12 June 2019

DISSERTATION APPROVAL PAGE

13 Student: JOHNCHRISTOPHER STEPHEN MYERS

14 Title: SEARCH FOR HIGGS BOSON PAIR PRODUCTION IN THE  $b\bar{b}WW^*$   
15 CHANNEL IN  $\sqrt{s} = 13$  TeV PROTON-PROTON COLLISIONS AT THE LARGE  
16 HADRON COLLIDER USING THE ATLAS DETECTOR

17 This dissertation has been accepted and approved in partial fulfillment of the  
18 requirements for the Doctor of Philosophy degree in the Department of Physics by:

19	Stephanie Majewski	Chair
	Eric Torrence	Advisor
	Tim Cohen	Member
	Hank Childs	Institutional Representative

and

Dean Livelybrooks	Dean of the Graduate School
-------------------	-----------------------------

20 Original approval signatures are on file with the University of Oregon Graduate  
21 School.

22 Degree awarded June 2019

© 2019 JOHNCHRISTOPHER STEPHEN MYERS

This work is licensed under a Creative Commons

**Attribution-NonCommercial-NoDerivs (United States) License.**

## DISSERTATION ABSTRACT

<sup>23</sup> JOHNCHRISTOPHER STEPHEN MYERS

<sup>24</sup> Doctor of Philosophy

<sup>25</sup> Department of Physics

<sup>26</sup> June 2019

<sup>27</sup> Title: SEARCH FOR HIGGS BOSON PAIR PRODUCTION IN THE  $b\bar{b}WW^*$   
<sup>28</sup> CHANNEL IN  $\sqrt{s} = 13$  TeV PROTON-PROTON COLLISIONS AT THE LARGE  
<sup>29</sup> HADRON COLLIDER USING THE ATLAS DETECTOR

<sup>30</sup>

<sup>31</sup>

<sup>32</sup> This dissertation presents a search for double Higgs production in the  $b\bar{b}WW^*$   
<sup>33</sup> final state in proton-proton collisions at the ATLAS detector at the Large Hadron  
<sup>34</sup> Collider. Double Higgs production is predicted in the Standard Model with a cross  
<sup>35</sup> section of  $\sigma_{HH} = 33.53^{+4.3\%}_{-6.0\%}$  fb. Many Beyond the Standard Model theories  
<sup>36</sup> predict enhancements to the production cross section through resonant production.

<sup>37</sup> The 2015-2016 ATLAS dataset has an integrated luminosity of 36.1 fb<sup>-1</sup> with  
<sup>38</sup> a center of mass energy of  $\sqrt{s} = 13$  TeV. Candidate events are broken into two  
<sup>39</sup> kinematic regions: a resolved selection containing one lepton (either an electron or  
<sup>40</sup> muon), 2 b-tagged calorimeter jets, 2 light-flavor jets, and missing transverse energy;  
<sup>41</sup> and a boosted analysis containing one lepton (electron or muon), two large radius jets,  
<sup>42</sup> one with two ghost associated, b-tagged track-jets and missing transverse energy. No  
<sup>43</sup> significant deviation from background was observed a cross section upper limit was  
<sup>44</sup> set for the SM double Higgs production of 10 pb and for resonant production as a  
<sup>45</sup> function of HH invariant mass from 500 GeV to 3000 GeV.

## CURRICULUM VITAE

46 NAME OF AUTHOR: JOHNCHRISTOPHER STEPHEN MYERS

47 GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

48

49 University of Oregon, Eugene, Oregon

50 The Ohio State University, Columbus, Ohio

51 DEGREES AWARDED:

52

53 Doctor of Philosophy, Physics, 2018 University of Oregon

54 Bachelor of Science, Physics, 2013, The Ohio State University

55 AREAS OF SPECIAL INTEREST:

56

57 Boosted object reconstruction

58 Trigger and detector operations

59 PROFESSIONAL EXPERIENCE:

60 Graduate Research Assistant, University Of Oregon: ATLAS Collaboration,  
61 2015-Present

62 Graduate Teaching Assistant, University Of Oregon: Department of Physics,  
63 2013-2015

64 PUBLICATIONS:

65 List of publications with significant contributions:

66 Morad Aaboud et al. “Search for Higgs boson pair production in the  $b\bar{b}WW^*$   
67 decay mode at  $\sqrt{s} = 13$  TeV with the ATLAS detector”. In: (2018). arXiv:  
68 1811.04671 [hep-ex]

69 Co-authored on 124 publications with minor contributions can be found at Inspire  
70 (search parameters: exactauthor:John.Myers.1)

## ACKNOWLEDGEMENTS

71 I would like to thank my advisor Eric Torrence for giving me this opportunity  
72 and guiding me throughout my time at Oregon. Thank you to the remaining members  
73 of my committee; especially Stephanie Majewski for your mentorship over the years.  
74 Thank you to David Strom for your help during my time at CERN. And to Anne  
75 McGinley for all of your help. Thank you to all of the post-docs who watched out for  
76 me at CERN, both with physics and life in general: Aparajita Dattagupta, Walter  
77 Hopkins, Jochen Heinrich, and Kate Whalen. Thank you to all my colleges who  
78 suffered with me: Ben Allen, Jason Barkeloo, Johan Bonilla, Liza Brost, Tyler Burch,  
79 Elliot Parrish, Ian Snyder, Bri Stamas, Amanda Steinhebel , and Chaowaroj ‘Max’  
80 Wanotayaroj.

81

82 Thank you to my family. You have always believed in me and your  
83 encouragement got me through many tough times. And finally and most importantly,  
84 thank you to my wife, Alison Jelden, and her family. Alison, without your support  
85 and love, I would not be where I am today.

## TABLE OF CONTENTS

Chapter	Page
86	87
88	Contents . . . . .
89	I. INTRODUCTION . . . . .
90	1.1. The Standard Model . . . . .
91	II. DI-HIGGS PRODUCTION . . . . .
92	2.1. Standard Model . . . . .
93	2.2. Resonant Production . . . . .
94	2.3. Summary . . . . .
95	III. EXPERIMENTAL SETUP . . . . .
96	3.1. Hadronic Colliders . . . . .
97	3.2. The Large Hadron Collider . . . . .
98	3.3. Detector Overview . . . . .
99	IV. SIMULATION AND EVENT RECONSTRUCTION . . . . .
100	4.1. Simulation . . . . .

Chapter		Page
101	4.2. Particle Identification . . . . .	43
102	V. ANALYSIS . . . . .	52
103	5.1. Analysis Overview . . . . .	53
104	5.2. Data and Monte Carlo Samples . . . . .	53
105	5.3. Object Reconstruction . . . . .	56
106	5.4. Resolved Analysis . . . . .	64
107	5.5. Boosted Analysis . . . . .	99
108	5.6. Results . . . . .	133
109	5.7. Conclusion . . . . .	145
110	VI. IMPROVEMENTS TO THE BOOSTED ANALYSIS . . . . .	147
111	6.1. Motivation . . . . .	147
112	6.2. Event Reconstruction . . . . .	148
113	6.3. Event selection . . . . .	150
114	6.4. Results . . . . .	156
115	6.5. Conclusion . . . . .	160
116	VII. CONCLUSION . . . . .	162
117	APPENDICES	
118	A. DERIVATION (HIGG5D2) . . . . .	163

Chapter		Page
<sup>119</sup> B.	COMPLETE LIST OF MC SAMPLES . . . . .	<sup>164</sup>
<sup>120</sup>	B.1. Lepton selection optimisation . . . . .	<sup>168</sup>
<sup>121</sup> C.	LEPTON AND B-JET OVERLAP REMOVAL: IMPACT ON SIGNAL EFFICIENCY . . . . .	<sup>169</sup>
<sup>122</sup>		
<sup>123</sup> D.	SOLVING FOR NEUTRINO LONGITUDINAL MOMENTUM . . .	<sup>171</sup>
<sup>124</sup>		
E.	FREEZING B AND D REGIONS IN QCD ESTIMATE . . . . .	<sup>175</sup>
<sup>125</sup>		
<sup>126</sup>	E.1. QCD Lepton Flavour Composition after Preselection Criteria for $\sigma_{d0}$ distribution . . . . .	<sup>179</sup>
<sup>127</sup>		
F.	STABILITY OF $T\bar{T}$ NORMALIZATION IN QCD ESTIMATE . . .	<sup>180</sup>
<sup>128</sup>		
G.	FULLY BOOSTED SYSTEMATIC UNCERTAINTIES . . . . .	<sup>183</sup>

## LIST OF FIGURES

Figure	Page
129	130
131	1.1. Electron-Positron Annihilation . . . . . 5
132	1.2. Muon emitting a neutrino . . . . . 6
133	1.3. The Higgs potential . . . . . 7
134	1.4. Higgs boson mass peak . . . . . 10
135	2.1. Di-Higgs production diagrams . . . . . 12
136	2.2. Normalized di-Higgs cross section . . . . . 13
137	2.3. BSM di-Higgs production diagrams . . . . . 14
138	2.4. Allowed regions of parameter space with enhanced di-Higgs production . . . . . 15
139	2.5. Allowed regions of parameter space with enhanced trilinear coupling . . . . . 16
140	2.6. Resonant di-Higgs production diagram . . . . . 17
141	2.7. Allowed branching ratios for resonant di-Higgs production . . . . . 18
142	3.1. CERN accelerator complex . . . . . 20
143	3.2. Cross-section of cryodipole . . . . . 22
144	3.3. Magnetic field configuration of the dipole magnets . . . . . 22
145	3.4. Integrated Luminosity . . . . . 24
146	3.5. Number of interactions per bunch crossing . . . . . 25
147	3.6. Detector coordinate system . . . . . 26
148	3.7. The ATLAS detector . . . . . 26
149	3.8. Cross section of the Inner Detector . . . . . 27
150	3.9. Sketch of the accordion structure of the EM calorimeter . . . . . 31

Figure	Page
151 3.10. Sketch of the matrix and rods in the forward calorimeter . . . . .	32
152 3.11. Tile readout system . . . . .	33
153 3.12. ATLAS muon system . . . . .	34
154 3.13. Geometry of magnet windings and tile calorimeter steel . . . . .	36
155 3.14. Schematic layout of the ATLAS TDAQ system in Run-2. . . . .	37
156 4.1. Event generation cartoon . . . . .	41
157 4.2. Parton distribution function . . . . .	42
158 4.3. Cartoon showing particle interacting in ATLAS detector subsystems . . .	44
159 4.4. Illustration of Infrared and collinear safety . . . . .	47
160 4.5. A sample parton-level event . . . . .	48
161 4.6. Schmatic view of the tracks in a b-jet . . . . .	49
162 5.1. Schematic diagram of $HH \rightarrow b\bar{b}WW^* \rightarrow b\bar{b}l\nu qq$ . . . . .	52
163 5.2. Comparison signal and $t\bar{t}$ events . . . . .	69
164 5.3. $m_{bb}$ resolution for signal and sum of backgrounds. . . . .	73
165 5.4. $m_{bb}$ and $m_{hh}$ in CR1, CR2 and CR3. . . . .	78
166 5.5. A pictoral representation of the four regions used in the ABCD 167 calculation. . . . .	80
168 5.6. The normalized shapes of the full $m_{b\bar{b}}$ distributions . . . . .	84
169 5.7. The $m_T$ distribution . . . . .	86
170 5.8. The $m_{b\bar{b}}$ distribution . . . . .	87
171 5.9. The $m_{b\bar{b}}$ distribution . . . . .	88
172 5.10. $d_0^{\text{sig}}$ distributions for data and background MC samples, identifying the 173 lepton channel. . . . .	94
174 5.11. $d_0^{\text{sig}}$ distributions and the relative ratio for data and background MC 175 samples. . . . .	95

Figure	Page
<sup>176</sup> 5.12. $p_T(bb)$ for DR and DS schemes for single top modelling. . . . .	<sup>98</sup>
<sup>177</sup> 5.13. Diagram of the reconstructed event topology . . . . .	<sup>100</sup>
<sup>178</sup> 5.14. Regions defined in the ABCD method . . . . .	<sup>104</sup>
<sup>179</sup> 5.15. Shape comparison of the $m_{HH}$ distribution in region C of the 1-tag and <sup>180</sup> 2-tag region . . . . .	<sup>111</sup>
<sup>181</sup> 5.16. The invariant mass of the reconstructed di-Higgs (hh) system, $\cancel{E}_T$ and <sup>182</sup> transverse mass of the $W \rightarrow l\nu$ system distributions of events in the mBB <sup>183</sup> control region (mBBcr). . . . .	<sup>115</sup>
<sup>184</sup> 5.17. Kinematic distributions of the reconstructed large- $R$ jet in the mBB <sup>185</sup> control region (mBBcr). . . . .	<sup>117</sup>
<sup>186</sup> 5.18. Kinematic distributions of the reconstructed $h \rightarrow WW$ system in the <sup>187</sup> mBB control region (mBBcr). . . . .	<sup>118</sup>
<sup>188</sup> 5.19. Kinematic distributions of the reconstructed $W \rightarrow q\bar{q}$ system in the mBB <sup>189</sup> control region (mBBcr). . . . .	<sup>119</sup>
<sup>190</sup> 5.20. Kinematic distributions of the reconstructed $W \rightarrow l\nu$ system in the mBB <sup>191</sup> control region (mBBcr). . . . .	<sup>120</sup>
<sup>192</sup> 5.21. Kinematic distributions of the selected lepton in the mBB control region <sup>193</sup> (mBBcr). . . . .	<sup>121</sup>
<sup>194</sup> 5.22. Kinematic distributions of the leading and sub-leading small- $R$ jets (of <sup>195</sup> the reconstructed hadronic W) in the mBB control region (mBBcr). . . . .	<sup>122</sup>
<sup>196</sup> 5.23. $\Delta R$ distribution between the selected lepton and the large- $R$ jet and $\Delta R$ <sup>197</sup> distribution between the track-jets inside the large- $R$ jet in the mBB <sup>198</sup> control region (mBBcr). . . . .	<sup>123</sup>
<sup>199</sup> 5.24. Electron and muon $d_0^{\text{sig}}$ distributions in the top control region without the <sup>200</sup> large- $R$ jet mass cut. . . . .	<sup>123</sup>
<sup>201</sup> 5.25. $m_{hh}$ distribution shape comparison between nominal $t\bar{t}$ sample and <sup>202</sup> alternative samples . . . . .	<sup>128</sup>
<sup>203</sup> 5.26. $m_{hh}$ distribution shape comparison between nominal W+jets sample and <sup>204</sup> scale variation ( $\mu_R=0.5$ , $\mu_F=0.5$ ) sample . . . . .	<sup>131</sup>

Figure	Page
205 5.27. $m_{HH}$ distributions for non-resonant and $m500$ selections in the resolved analysis . . . . .	134
207 5.28. $m_{HH}$ distributions for non-resonant and $m500$ selections in the resolved analysis for the low and high mass selections . . . . .	134
209 5.29. Expected and observed upper limit at 95% CL on the cross section of resonant pair production for the resolved analysis in the heavy scalar boson $S$ model . . . . .	138
212 5.30. $m_{HH}$ distributions after the global likelihood fit for the boosted analysis . . . . .	144
214 5.31. Expected and observed upper limits at 95% CL on the cross section of resonant pair production for the heavy scalar boson $S$ model in the boosted analysis . . . . .	145
217 6.1. Distance between parton objects . . . . .	148
218 6.2. Diagram of the fully-boosted event topology . . . . .	150
219 6.3. Comparison of $H \rightarrow WW^*$ mass, $H \rightarrow WW^*$ $p_T$ , $\cancel{E}_T$ , and $HH$ mass for the electron channel . . . . .	151
221 6.4. Comparison of $H \rightarrow WW^*$ mass, $H \rightarrow WW^*$ $p_T$ , $\cancel{E}_T$ , and $HH$ mass for the muon channel . . . . .	152
223 6.5. $M_{HH}$ Distribution in the mBB control region for the electron and muon channels . . . . .	157
225 6.6. The $m_{HH}$ distribution for data and background in the final signal region. . . . .	160
227 6.7. Expected and observed upper limits at 95% CL on the cross section of the resonant pair production for heavy scalar boson S model . . . . .	161
229 D.1. $m_W$ Distribution and the solution using $W$ mass method. . . . .	172
230 D.2. Single $W$ $\eta_l$ vs Correct Sign Choice . . . . .	173
231 D.3. $\eta_l$ and the solution using $\eta_l$ method in 700 GeV resonant sample. . . . .	173
232 D.4. Higgs pair $\eta_{lq\bar{q}}$ vs Correct Sign Choice . . . . .	173
233 D.5. $\eta_l$ and the solution using $\eta_l$ method in 700 GeV resonant sample. . . . .	174

Figure	Page
--------	------

234      D.6. $\Delta R(l, \nu)$ and the solution using $\Delta R(l, \nu)$ method in 700 GeV resonant	174
235      sample. . . . .	

## LIST OF TABLES

		Page
236	Table	Page
237		
238	1.1. Particles of the Standard Model . . . . .	3
239	1.2. Fermion doublets and singlets . . . . .	5
240	5.1. SM MC samples used for background estimation. . . . .	54
241	5.2. Di-Higgs signal samples used in the analysis. . . . .	56
242	5.3. Electron selection requirements. . . . .	58
243	5.4. Muon selection requirements. . . . .	60
244	5.5. Selection for jets with distance parameter $R = 0.4$ . . . . .	63
245	5.6. A summary of the overlap removal procedure. . . . .	64
246	5.7. List of triggers used . . . . .	66
247	5.8. Criteria for non-resonant, <i>m500</i> , <i>low-mass</i> and <i>high-mass</i> selection . . .	71
248	5.9. Window cuts on $m_H H$ . . . . .	72
249	5.10. S/B ratios . . . . .	73
250	5.11. Definition of the kinematic regions used to normalise the Top background . . . . .	74
252	5.12. The number of events in the $m_{bb}$ sideband for the non-res selection . . .	74
253	5.13. Events in $m_{bb}$ side band for the m500 selection . . . . .	75
254	5.14. Events in $m_{bb}$ side band for the m700 selection . . . . .	75
255	5.15. Events in $m_{bb}$ side band for the high-mass selection . . . . .	76
256	5.16. Normalization factors for the two CRs . . . . .	76
257	5.17. Values of R at each selection . . . . .	81
258	5.18. Cross check for coorelation with R . . . . .	82

Table	Page
259 5.19. Ratio of events in SR over SR+CR . . . . .	83
260 5.20. $R_{non-prompt}$ . . . . .	85
261 5.21. Number of SR events in the non-res selection . . . . .	88
262 5.22. Number of SR events in the m500 selection . . . . .	89
263 5.23. Number of SR events in the m700 selection . . . . .	89
264 5.24. Number of SR events in the m2000 selection . . . . .	90
265 5.25. Object systematic uncertainties . . . . .	91
266 5.26. Extrapolation uncertainties . . . . .	96
267 5.27. Theoretical uncertainties in percentage on $W/Z+jets$ event yield . . . . .	98
268 5.28. Theoretical uncertainties in percentage on the signal acceptance. . . . .	99
269 5.29. MC predicted prompt lepton backgrounds, observed data and calculated 270 multijet yields in Region B and D . . . . .	103
271 5.30. MC predicted prompt lepton backgrounds, observed data and calculated 272 multijet yields in Region C mBBcr and SR . . . . .	103
273 5.31. Multijet yields in region B and region D and also the ratio of the yields 274 for each lepton channel . . . . .	105
275 5.32. Multijet yield in region C and predicted yield in region A in the SR . . .	105
276 5.33. The acceptance of the large- $R$ jet mass signal region selection on the 277 multijet 1-tag and 2-tag region C . . . . .	107
278 5.34. The uncertainty on the multijet yield prediction in region A . . . . .	108
279 5.35. The total uncertainty on the multijet yield prediction in region A due to 280 the detector modeling uncertainties of the prompt backgrounds in region 281 B, C and D. . . . .	109
282 5.36. The total uncertainty on the multijet yield prediction in region A due to 283 the $ d_0^{\text{sig}} $ cut efficiency modeling uncertainties of the prompt backgrounds 284 in region B, D and C. . . . .	109
285 5.37. Summary of systematic uncertainties on the QCD multijet yield in the 286 signal region and mBB control region for each lepton channel . . . . .	110

Table	Page
287 5.38. Predicted multijet yield in the SR and mBBcr . . . . .	111
288 5.39. Predicted and observed yields in the mBB control region . . . . .	114
289 5.40. Summary of the (nuisance parameter) names and meanings of the detector 290 modeling systematic uncertainties. . . . .	116
291 5.41. Efficiency of the $ d_0^{\text{sig}} $ cut for electrons and muons in Data and MC . . .	124
292 5.42. The normalization uncertainty for the $t\bar{t}$ background in the SR from 293 different sources . . . . .	127
294 5.43. The normalization uncertainty for the $t\bar{t}$ background in the SR from 295 different sources . . . . .	129
296 5.44. The normalization uncertainty for the W+Jets background in the SR from 297 different sources . . . . .	131
298 5.45. The normalization uncertainty for the Z+Jets background in the SR from 299 different sources . . . . .	132
300 5.46. Theoretical uncertainties (in percentage) on the acceptance of several 301 signal mass points. . . . .	133
302 5.47. Data event yields, and post-fit signal and background event yields in 303 the final signal region for the non-resonant analysis and the resonant 304 analysis . . . . .	136
305 5.48. Statistical contribution (in percentage) to the total error in the scaling 306 factor $\alpha_{\text{sig}}$ for the non-resonant signal and three scalar-signal mass 307 hypotheses, 500 GeV, 1000 GeV and 2000 GeV, in the resolved analysis . .	139
308 5.49. Systematic contributions (in percentage) to the total error in the scaling 309 factor $\alpha_{\text{sig}}$ for the non-resonant signal and three scalar-signal mass 310 hypotheses, 500 GeV, 1000 GeV and 2000 GeV, in the resolved analysis . .	140
311 5.50. Data event yields, and post-fit signal and background event yields in 312 the final signal region for the boosted analysis and the scalar $S$ particle 313 hypothesis . . . . .	141
314 5.51. Statistical and systematic contributions (in percentage) to the total error 315 in the scaling factor $\alpha_{\text{sig}}$ in the boosted analysis for four mass hypotheses: 316 1500 GeV, 2000 GeV, 2500 GeV and 3000 GeV . . . . .	142
317 5.52. Predicted and observed yields in the signal region . . . . .	143

Table	Page
318 6.1. Alternative $p_T$ definition for the electron and muon channels . . . . .	150
319 6.2. MC predicted prompt lepton backgrounds, observed data and calculated 320 multijet yields in Region B and D . . . . .	154
321 6.3. MC predicted prompt lepton backgrounds, observed data and calculated 322 multijet yields in Region C mBBcr and SR . . . . .	154
323 6.4. Multijet yields in region B and region D and also the ratio of the yields 324 for each lepton channel . . . . .	155
325 6.5. Multijet yield in region C and predicted yield in region A in the SR . . .	155
326 6.6. The acceptance of the large- $R$ jet mass signal region selection on the 327 multijet 1-tag and 2-tag region C . . . . .	156
328 6.7. Data event yields, and signal and background event yields in the final 329 signal region . . . . .	157
330 6.8. Predicted and observed yields in the mBB control region . . . . .	158
331 6.9. List of dominant systematic uncertainties for the fully boosted analysis. 332 The full list of systematic uncertainties is listed in Appendix F . . . . .	159
333 E.1. $B/D$ at each stage in the non-resonant, low mass, and high mass 334 selections . . . . .	175
335 E.2. QCD and total background yields for the non-resonant selection without 336 freezing . . . . .	176
337 E.3. QCD and total background yields for the non-resonant selection without 338 freezing . . . . .	176
339 E.4. QCD and total background yields for the high mass selection without 340 freezing . . . . .	176
341 E.5. QCD and total background yields for the non-resonant selection after 342 freezing . . . . .	177
343 E.6. QCD and total background yields for the low mass (m700) selection after 344 freezing . . . . .	177
345 E.7. QCD and total background yields for the low mass selection after 346 freezing . . . . .	177

Table	Page
347 E.8. Lepton flavor composition in QCD sample . . . . .	179
348 F.1. Iteration of $t\bar{t}$ normalization and QCD estimate for the non-resonant 349 selection . . . . .	181
350 F.2. Iteration of $t\bar{t}$ normalization and QCD estimate for the low mass 351 selection . . . . .	181
352 F.3. Iteration of $t\bar{t}$ normalization and QCD estimate for the low mass 353 selection . . . . .	182
354 G.1. List of systematic uncertainties for the fully boosted analysis . . . . .	190

## Contents

357        The Standard Model (SM) is the culmination of more than a century of work.  
 358        The first piece added to the puzzle was the electron, discovered in 1891. Since then,  
 359        24 other particles have been discovered, with the final piece, the Higgs Boson, being  
 360        added in 2012. Since it was theorized, the SM has held up to rigorous experimentation  
 361        and remains an unbeaten theory of fundamental matter and forces. Even though the  
 362        SM is widely successful, it fails to explain all observed phenomena. Gravity, neutrino  
 363        masses, dark matter, along with other observations, all lack explanation within the  
 364        SM. The remaining task is to probe the extremes of the SM to either more precisely  
 365        measure the parameters or to find its limit.

366        **1.1. The Standard Model**

367        The Standard Model defines the basic building blocks of matter and force and  
 368        the interactions between them. Everyday matter is made of protons, neutrons, and  
 369        electrons. Electrons are a fundamental particle, called a lepton, meaning they are not  
 370        made of smaller constituents. However, protons and neutrons are not fundamental  
 371        particles. They are a composite of up and down quarks, two more fundamental  
 372        particles. The protons are made of two ups and a down and the neutrons are made  
 373        of two downs and an up. The up and down quarks, along with the electron, are types  
 374        of fermions.

375        Fermions are spin- $\frac{1}{2}$  particles that make up all matter in the SM. The fermions  
 376        can be broken down into 3 “generations”. Where a generation contains two quarks,  
 377        one with electric charge  $+\frac{2}{3}$  and one with electric charge  $-\frac{1}{3}$ , one electrically charged

378 lepton, charge -1, and one electrically neutral lepton. The quarks have an additional  
 379 color charge, of which there are 3 charges. This is additional quantum number  
 380 associated with the strong force. In all, this gives 12 fermions.

381

	Particle	Spin	Charge	Mass
<b>Quarks</b>				
u type	u			$2.4^{+0.6}_{-0.4} MeV$
	c	$\frac{1}{2}$	$\frac{2}{3}$	$1.28 \pm 0.03 GeV$
	t			$173.1 \pm 0.6 GeV$
d type	d			$4.7^{+0.5}_{-0.4} MeV$
	s	$\frac{1}{2}$	$-\frac{1}{3}$	$96^{+8}_{-4} MeV$
	b			$4.18^{+0.04}_{-0.03} GeV$
<b>Leptons</b>				
e family	e	$\frac{1}{2}$	-1	$0.5109989461 \pm 0.000000003 MeV$
	$\nu_e$		0	$< 2 eV$
$\mu$ family	$\mu$	$\frac{1}{2}$	-1	$105.6583745 \pm 0.0000024 MeV$
	$\nu_\mu$		0	$< 2 eV$
$\tau$ family	$\tau$	$\frac{1}{2}$	-1	$1776.86 \pm 0.12 GeV$
	$\nu_\tau$		0	$< 2 eV$
<b>Bosons</b>				
Vector	$\gamma$	1	0	$< 10^{-18} eV$
	g	1	0	0
	W	1	$\pm$	$80.385 \pm 0.0015 GeV$
	Z	1	0	$91.1876 \pm 0.0021 GeV$
Scalar	H	0	0	$125.09 \pm 0.21 \pm 0.11 GeV$

TABLE 1.1. Particles of the Standard Model [2]

382 Gauge bosons are spin-1 particles responsible for carrying the fundamental  
 383 forces in the standard model. There are 12 physical gauge bosons. The photon  
 384  $\gamma$  is a massless, charge neutral force carrier for the electromagnetic force. The  
 385 nuclear forces are carried by 3 massive gauge bosons. A chargeless Z boson and two  
 386 charged W bosons,  $Charge = \pm 1$ . Together, these four bosons control the electroweak  
 387 interactions in the standard model. The remaining 8 bosons are the gluons, the force  
 388 carriers for the strong nuclear interaction. Gluons are massless, electrically neutral  
 389 particles that have two color charges. There is a gluon for each linearly independent

390 combination of the three color, anti-color, charges, giving the 8 colored gluons. The  
391 color singlet state does not exist in SU(3).

392 The remaining piece of the standard is the Higgs Boson. The Higgs boson is a  
393 massive scalar, spin-0, chargeless boson. The Higgs boson is responsible for giving  
394 mass to the massive fundamental particles. The full list of SM particles and their  
395 properties are in Table 1.1

396 . Along with all of the fundamental particles, there are also anti-particles. Every  
397 particle in the SM has a partner of identical mass but opposite charge. These anti-  
398 particles, when they interact with their matching particle, annihilate. For example,  
399 the anti-particle to the electron is the positron; a particles with electric charge of -1.  
400 For the W boson, the positive and negative Ws are particle-antiparticle pairs. Some  
401 particles are their own anti particle, known as Majorana particles. For example, the  
402 anti-particle of a photon is a photon, the same is true for the Z boson.

403

#### 404 **1.1.1. Interactions**

405 The SM is governed by three different types of interactions. For leptons, the  
406 overarching theory is the electroweak theory. This can be broken down further into  
407 the electromagnetic interaction, Quantum Electrodynamics (QED) and the weak  
408 interaction. The electromagnetic interaction defines the interaction of electrically  
409 charged particles with photons. The fundamental diagram for the electromagnetic  
410 interaction is electron-positron annihilation figure 1.1, where an electron and a  
411 positron collide and produce two photons. This can also be reversed, two photons  
412 interact and produce an electron-positron pair. The strength of this interaction is the  
413 electrical charge e.

414

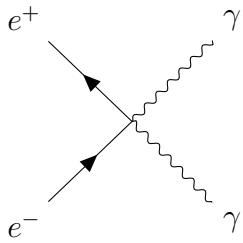


FIGURE 1.1. Electron-Positron Annihilation

415        The weak interaction defines the interaction of particles under the weak isospin  
 416        quantum number. In the SM, every fermion is a mix of a left and a right-handed  
 417        chirality. Particle with a right-handed chirality have a weak isospin  $T = 0$ . These  
 418        particles exist as singlets and do not interact with the weak force. Left-handed  
 419        particles have a weak isospin  $T = \frac{1}{2}$ . These particles live as doublets as illustrated in  
 420        table 1.2. For these particles, the third component of the weak isospin  $T_3$ ,  $+\frac{1}{2}$  for up-  
 421        type quarks and charged leptons and  $-\frac{1}{2}$  for down-type quarks and neutral leptons.  
 422        Under weak interactions, particles with  $T_3 = +\frac{1}{2}$  always transform into particles with  
 423         $T_3 = -\frac{1}{2}$ , or vice versa.

424

Left Handed Fermions, $T = \frac{1}{2}, T_3 = \pm \frac{1}{2}$	Right Handed Fermions, $T = 0, T_3 = 0$
$(\begin{smallmatrix} u \\ d \end{smallmatrix}), (\begin{smallmatrix} c \\ s \end{smallmatrix}), (\begin{smallmatrix} t \\ b \end{smallmatrix}), (\begin{smallmatrix} e \\ \nu_e \end{smallmatrix}), (\begin{smallmatrix} \mu \\ \nu_\mu \end{smallmatrix}), (\begin{smallmatrix} \tau \\ \nu_\tau \end{smallmatrix})$	u, d, c, s, t, b, e, $\nu_e$ , $\mu$ , $\nu_\mu$ , $\tau$ , $\nu_\tau$

TABLE 1.2. Fermion doublets and singlets [3]

425        The remaining piece of the weak interaction are the W and Z bosons. The W  
 426        has an isospin of  $T = 1$ . This gives three option for the third component of isospin,  
 427         $T_3 = +1, 0, -1$  which give the  $W^+$ , the  $W^0$ , and the  $W^-$ .  $W^0$  will be discussed more  
 428        in section 1.1.2. The  $W^\pm$  either raises or lowers the  $T_3$  of the fermions. The Z  
 429        boson has a weak isospin of 0 meaning it does not change the isospin of the fermions.

430 Instead, the Z boson transfers momentum, energy and spin in interactions that do not  
431 change electric charge or weak isospin. Figure 1.2 is an example of a weak interaction.

432

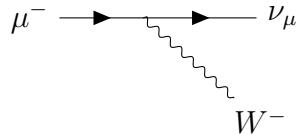


FIGURE 1.2. Muon emitting an muon neutrino and a W Boson

433 The last fundamental forces in the SM is quantum chromodynamics (QCD).  
434 QCD is governed by the quantum number known as color, of which there are three  
435 options. Colloquially, referred to as red, green, and blue, and their anti-states. Unlike  
436 the electroweak force, as the distance between a pair of colored particles increases, the  
437 force between them increases. A consequence of this is known as color confinement.  
438 As the distance between the particles becomes greater, the energy stored grows until  
439 a new quark-antiquark pair in the vacuum between them to neutralize the color of  
440 the original quarks. This process results in showers of color neutral composite mesons  
441 and bosons whenever a quark gains enough momentum to be ejected from an object.  
442 These showers are reconstructed in detectors as objects called jets. The same thing  
443 happens to the 8 gluons, as they too cannot exist by themselves. This is also the  
444 reason the color singlet gluon cannot exist. The color singlet would not be confined  
445 and could travel infinite distances, leading to long range strong interactions. Which  
446 we know do not exist.

#### 447 1.1.2. The Higgs Mechanism and Higgs Boson

448 Electroweak theory is a gauge invariant theory. This means the Lagrangian that  
449 describes the system is invariant under local gauge transformations. To satisfy this

450 symmetry, the bosons must be massless. However, the physical electroweak bosons in  
 451 the standard model, the  $W^\pm$  the  $Z$  and the  $\gamma$  are not all massless. This means that  
 452 they must not be the fundamental bosons of Electroweak theory.

453

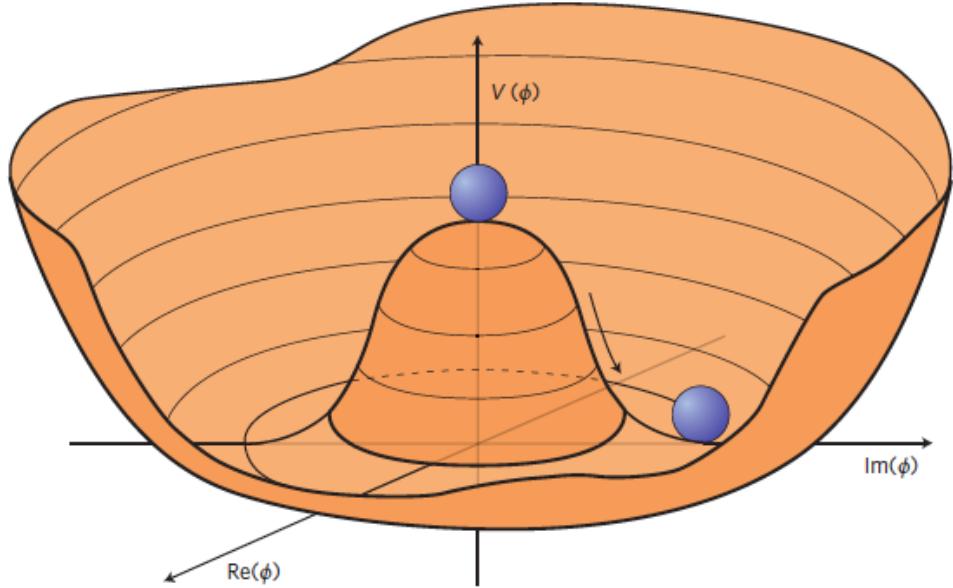


FIGURE 1.3. The Higgs potential.[4]

In QED, the four fundamental gauge bosons are  $W_\mu^i, i = 1, 2, 3$  and  $B_\mu$ . These bosons couple to a complex scalar doublet,  $\Phi \equiv (\begin{smallmatrix} \phi^+ \\ \phi^0 \end{smallmatrix})$ . This doublet has a scalar potential.

$$V(\Phi) = \mu^2 |\Phi^\dagger \Phi| + \lambda (|\Phi^\dagger \Phi|)^2 \quad (1.1)$$

Where  $\mu^2 < 0$ . This gives the Mexican hat shaped potential seen in figure 1.3, with a minimum energy at

$$\langle \phi \rangle = \sqrt{-\frac{\mu^2}{2\lambda}} \equiv \frac{\nu}{\sqrt{2}} \quad (1.2)$$

<sup>454</sup> called the vacuum expectation value (VEV) of  $\phi$ . The choice of the direction of  
<sup>455</sup> fluctuation is arbitrary but can be chosen such that

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu \end{pmatrix} \quad (1.3)$$

After the direction is chosen, the only remaining degree of freedom is the scalar field  $h(x)$ , giving

$$\phi(x) = \phi_0 + h(x) \quad (1.4)$$

The doublet can now be described by

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad (1.5)$$

This field couples to the gauge bosons as

$$(\frac{g}{2} \vec{\tau} \cdot \vec{W} + \frac{g'}{2} B) \phi_0 \quad (1.6)$$

Where  $\vec{\tau}$  are the Pauli matrices,  $\vec{W}$  are  $W_{1,2,3}$  and  $g, g'$  are the coupling constants.  
The result of the coupling is the acquisition of mass by three eigenstates of the bosons,

$$\begin{aligned} W^\pm &= \frac{1}{\sqrt{2}} (W_\mu^1 \mp i W_\mu^2) \\ Z^\mu &= \frac{-g' B_\mu + g W_\mu^3}{\sqrt{g^2 + g'^2}} \\ A^\mu &= \frac{g B_\mu + g' W_\mu^3}{\sqrt{g^2 + g'^2}} \end{aligned} \quad (1.7)$$

These four eigenstates are the bosons we observe in the standard model. With Masses

$$\begin{aligned} M_W^2 &= \frac{1}{4}g^2\nu^2 \\ M_Z^2 &= \frac{1}{4}(g^2 + g'^2)\nu^2 \\ M_A &= 0 \end{aligned} \tag{1.8}$$

456 Through the mixing that occurs in the spontaneous electroweak symmetry breaking  
457 gives mass to the standard model Gauge Bosons while leaving the photon, A, massless.  
458 However, for this to occur, an additional scalar field, the Higgs Field, is required.

459 The Higgs Boson is an excitation in the scalar Higgs field predicted in 1964 [5].  
460 The decay mechanism of this massive boson was predicted by Peter Higgs, allowing  
461 for the decay products to be measured. Giving a way to prove the existence of the  
462 scalar Higgs field. In 2012, a Higgs like scalar boson was discovered at the LHC by  
463 the ATLAS and CMS experiments with a mass of  $125\text{GeV}/c^2$  [6].

464 Since the discovery, many measurements have been made of this Higgs Boson  
465 to compare it to the standard model Higgs Boson. So far, the Higgs Boson has held  
466 up to these tests. The Higgs Boson has spin-parity  $J^P = 0^+$  [7], decays to  $bb$  [8],  
467  $\gamma\gamma, \tau\tau$  [9],  $WW$  and  $ZZ$  have been measured with appropriate signal strengths, and  
468 no significant deviations have been observed in any Run 2 analyses. However, there  
469 are still many parameters of the Higgs Boson that still need measured. One of which  
470 is the triple Higgs coupling.

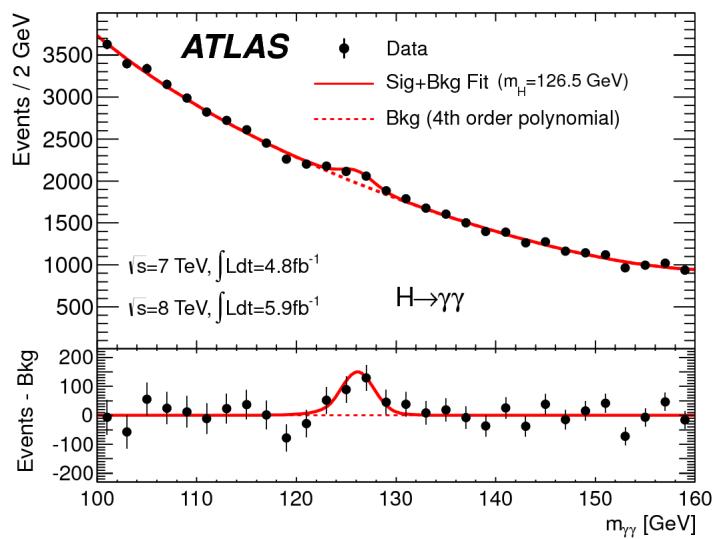


FIGURE 1.4. Invariant mass distribution of di-photon candidates for the combined root  $s = 7\text{TeV}$  and root  $s = 8\text{TeV}$  data samples. The result of a fit to the data of the sum of a signal component fixed to  $m_H = 126.5\text{GeV}$  and a background component described by a fourth-order Bernstein polynomial is superimposed. The bottom inset displays the residuals of the data with respect to the fitted background component.

## CHAPTER II

### DI-HIGGS PRODUCTION

473 Di-Higgs production rate is a necessary measurement to confirm the SM with the  
 474 observed Higgs boson. Additionally, the di-Higgs production rate gives a handle to  
 475 more accurately measure the Higgs potential. Additionally, the di-Higgs production  
 476 channel is a discovery channel for physics beyond the SM (BSM). The small SM  
 477 production rate makes the channel an important place to look for new physics. This  
 478 dissertation looks at both the measurement of the SM di-Higgs production rate and  
 479 to search for new physics through resonant di-Higgs production.

480 **2.1. Standard Model**

The Higgs self coupling potential in the SM is

$$V_{\text{self-coupling}} = \lambda(|\Phi^\dagger \Phi|)^2 \quad (2.1)$$

When  $\Phi$  is expanded around the VEV, the self coupling term can be written as

$$V_{\text{self-coupling}} = \frac{1}{2}(2\lambda\nu^2)H^2 + \lambda\nu H^3 + \frac{\lambda}{4}H^4 \quad (2.2)$$

481 Where the second term,  $\lambda\nu H^3$  is the coupling between three Higgs bosons with  
 482 strength  $\lambda\nu \equiv \lambda_{HHH}$ [10]. The trilinear Higgs coupling can be probed at the LHC  
 483 by measuring the cross section of events with two Higgs Bosons in the final state.

Currently at the LHC, there are two dominant ways to produce diHiggs events, the trilinear Higgs coupling gluon-gluon fusion (ggF) diagram and a box diagram,

figure 2.1. These diagrams interfere destructively, resulting in a low theoretical production cross section,

$$\sigma_{HH} = 33.53^{+4.3\%}_{-6.0\%} (\text{QCD} \pm 5.9\% \text{(other)}) \text{fb} \quad (2.3)$$

in pp collisions at 13 TeV [11]. This makes the trilinear coupling extremely hard to measure at the LHC but still accessible. Additionally, since the cross section is so small, it is an promising place to look for deviations from the SM, since any enhancement to the cross section would be indicative of new physics.

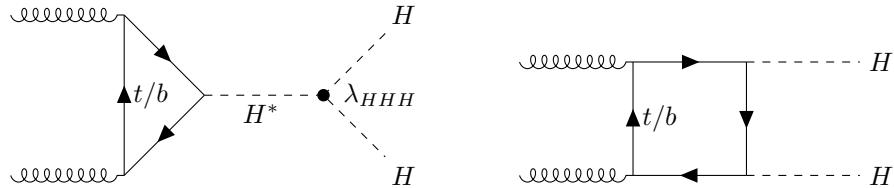


FIGURE 2.1. The dominate production method for di-Higgs events at the LHC with  $\sqrt{s} = 13\text{TeV}$ , with the trilinear Higgs coupling on the left

The SM di-Higgs production is a continuum production, with a turn-on at two times the Higgs mass, 250 GeV. Figure 2.2 shows the continuum distribution, as expected, it is a falling exponential distribution peaked around 400 GeV.

## 2.2. Resonant Production

There are several BSM models that may enhance the rate of diHiggs production at the LHC. This section will give an overview of a few of these models

### 2.2.1. Complex Higgs Singlet

The addition of a complex scalar singlet to the SM results in three neutral scalar particles after spontaneous symmetry breaking, which mix to give mass eigenstates,

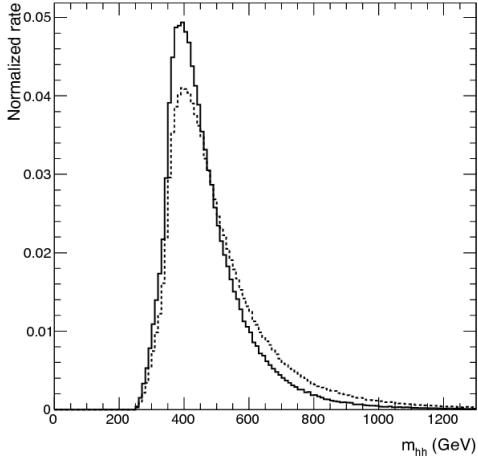


FIGURE 2.2. Normalized differential cross section for  $pp \rightarrow hh$  in the SM as a function of the invariant mass of the two Higgs bosons. The solid and dotted lines correspond respectively to  $\sqrt{s} = 14$  and  $100$  TeV.[12]

including the observed 125 GeV scalar [13].

After spontaneous symmetry breaking, the fields are defined as:

$$\Phi = \begin{pmatrix} 0 \\ \frac{h+\nu}{\sqrt{2}} \end{pmatrix}; S_c = \frac{1}{\sqrt{2}}(S + \nu_s + i(A + \nu_A)). \quad (2.4)$$

The normalizable scalar potential is

$$\begin{aligned} V(\Phi, S_c) = & \frac{\mu^2}{2}\Phi^\dagger\Phi + \frac{\lambda}{4}(\Phi^\dagger\Phi)^2 \\ & + \left(\frac{1}{4}\delta_1\Phi^\dagger\Phi S_c + \frac{1}{4}\delta_3\Phi^\dagger\Phi S_c^2\right. \\ & + a_1S_c + \frac{1}{4}b_1S_c^2 + \frac{1}{6}e_1S_c^3 + \frac{1}{6}e_2S_c|S_c|^2 \\ & \left. + \frac{1}{8}d_1S^4 + \frac{1}{8}d_3S_c^2|S_c|^2 + \text{H.C.}\right) \\ & + \frac{1}{4}d_2(|S_c|^2)^2 + \frac{\delta_2}{2}\Phi^\dagger\Phi|S_c|^2 + \frac{1}{2}b_2|S_c|^2 \end{aligned} \quad (2.5)$$

The mass eigenstate fields are given by:

$$\begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} = \begin{pmatrix} c_1 & -s_1 & 0 \\ s_1c_2 & c_1c_2 & s_2 \\ s_1s_2 & c_1s_2 & -c_2 \end{pmatrix} \begin{pmatrix} h \\ S \\ A \end{pmatrix} \quad (2.6)$$

where  $c_i = \cos \theta_i$ ,  $h$  is the SU(2) doublet field, and  $S$  and  $A$  are the real and imaginary components of the complex scalar  $S_c$ . Assigning the SM-like Higgs boson as  $h_1$  with  $m_1 = 125\text{GeV}$  and  $\nu = 246\text{GeV}$ ,  $h_2$  and  $h_3$  are physical heavy Higgs bosons with  $m_2, m_3 > m_1$ . The coupling of  $h_1$  to SM particles is dominant, suppressed by a factor of  $c_1$  from the SM rate, with the  $h_2$  couplings suppressed by  $s_1c_2$  and  $h_3$  couplings suppressed by  $s_1s_2$ . ATLAS has set limits on  $c_1 > 0.94$  at 95% CL in RUN-1. As the  $h_1$  couplings become more SM-like ( $\theta_1 \rightarrow 0$ ), the allowed  $h_2$  couplings become suppressed.

504

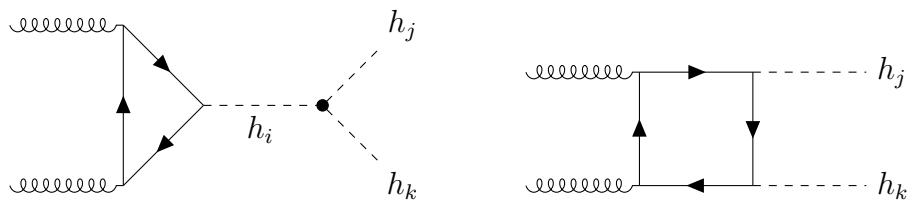


FIGURE 2.3. Feynman diagrams for the production of  $h_j h_k$ ,  $i, j, k = 1, 2, 3$ .

In the limit of  $\theta_2 \rightarrow 0$ , which is in agreement with the single Higgs rates,  $h_3$  does not directly couple to SM fermions or vector bosons. The only way to produce  $h_3$  is through  $h_1$  or  $h_2$ , with the largest production rate from  $gg \rightarrow h_2 \rightarrow h_1 h_3$ , figure 2.3. For a range of masses  $m_2$  and  $m_3$  the rate of production of  $h_1 h_3 \gg h_1 h_1$ , figure 2.4.

509

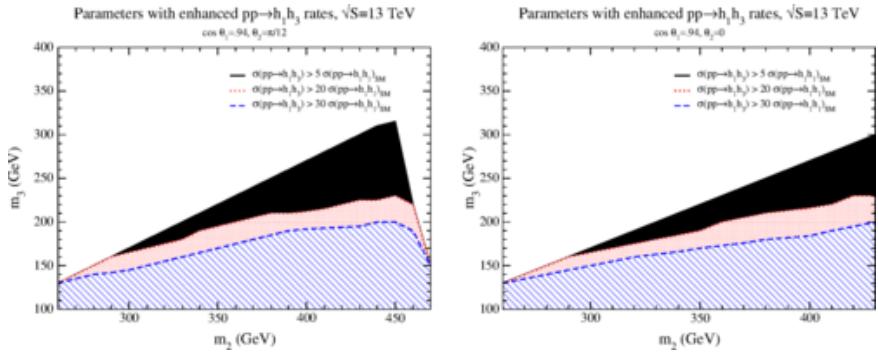


FIGURE 2.4. Regions of parameter space allowed by limits on oblique parameters, perturbative unitarity and the minimization of the potential where the rate for  $h_1h_3$  production is significantly larger than the SM  $h_1h_1$  rate at  $\sqrt{S} = 13\text{TeV}$ .

The enhancement can be seen in the potential where

$$V \rightarrow \frac{1}{2}\lambda_{211}h_1^2h_2 + \frac{1}{2}\lambda_{311}h_1^2h_3 + \frac{1}{2}\lambda_{331}h_1h_3^2 + \frac{1}{2}\lambda_{321}h_1h_2h_3 + \dots \quad (2.7)$$

510 So while the SM trilinear Higgs coupling is determined by  $m_h$ , with this extension,  
 511 the coupling is much less constrained. This leads to enhanced values seen in figure  
 512 2.5. So while this model would definitely show up in SM di-Higgs production, through  
 513 the first two terms in equation 2.7, a search for one SM Higgs and a heavy Higgs  
 514 would be more sensitive. This is a promising search moving forward but is not the  
 515 focus of this dissertation.

516 **2.2.2. Real Higgs Singlet Extension**

One simple explanation of an enhanced diHiggs production rate at the LHC is the addition of a real scalar Higgs singlet, S [14]. In this model, S can only interact with the SM through the Higgs field. In the case where there is no  $Z_2$  symmetry, where  $S \rightarrow -S$  the scalar field S mixes with the SM Higgs boson. If the mass is large enough, it is possible for S to decay to two on-shell SM Higgs Bosons, significantly

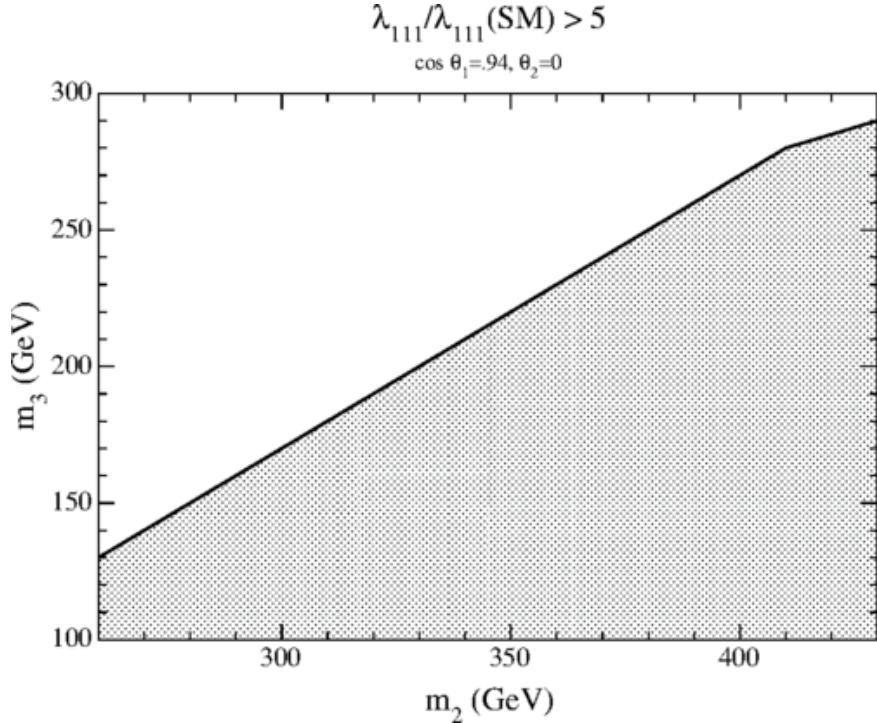


FIGURE 2.5. Region of parameter space allowed by limits on oblique parameters, perturbative unitarity and the minimization of the potential where the  $h_1 h_1 h_1$  trilinear coupling is greater than 5 times the SM value.

enhancing the diHiggs production rate.

The most general potential that can be added is

$$V(H, S) = -\mu^2 H^\dagger H + \lambda(H^\dagger H)^2 + \frac{a_1}{2} H^\dagger H S + \frac{a_2}{2} H^\dagger H S + b_1 S + \frac{b_2}{2} S^2 + \frac{b_3}{3} S^3 + \frac{b_4}{4} S^4. \quad (2.8)$$

Where  $H$  is  $\phi_0 = \frac{(h+\nu)}{\sqrt{2}}$  and  $\langle \phi_0 \rangle = \frac{\nu}{2}$ , while  $S = s + x$  where  $x$  is the vev of  $S$ . By shifting the field, it is possible to set  $x = 0$ . After electroweak symmetry breaking the fields mix to give the two mass eigenstates

$$\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} h \\ s \end{pmatrix} \quad (2.9)$$

With  $m_1 = 125\text{GeV}$ , the free parameters are  $m_2, \theta, a_2, b_2$  and  $b_4$ . For diHiggs production, in the case of  $m_2 > 2m_1$ , the important piece of the potential is

$$V(h_1 h_2) \supset \frac{\lambda_{111}}{3!} h_1^3 + \frac{\lambda_{211}}{3!} h_2 h_1^2 \quad (2.10)$$

<sup>517</sup> This give an additional resonant double Higgs production diagram , figure 2.6, for  $250\text{GeV} \leq m_2$ .

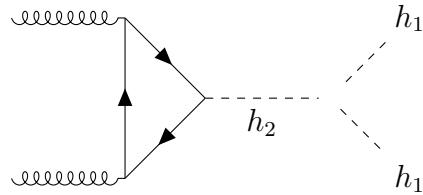


FIGURE 2.6. Feynman diagram for  $h_2 \rightarrow h_1 h_1$ .

<sup>518</sup>

<sup>519</sup> Varying the values of  $b_4$  and  $\sin^2 \theta$ , it is found that the maximum branching  
<sup>520</sup> ratio (BR) for  $h_2 \rightarrow h_1 h_1$  if obtained with  $b_4 = 4.2, \sin^2 \theta = 0.12$ . Figure 2.7, shows  
<sup>521</sup> the minimum and maximum BR as a function of  $m_2$ . The largest BR is when  
<sup>522</sup>  $m \approx 280\text{GeV}$  at  $BR(h_2 \rightarrow h_1 h_1) = 0.76$ . This corresponds to an enhancement in  
<sup>523</sup> diHiggs production of approximately 30 times the SM cross section.

### <sup>524</sup> 2.3. Summary

<sup>525</sup> The SM di-Higgs production rate is an important and achievable measurement  
<sup>526</sup> for the LHC and HL-LHC. It gives insight to the shape of the Higgs potential through  
<sup>527</sup> measurement of the trilinear Higgs coupling. It is also a valuable discovery channel for  
<sup>528</sup> BSM physics, especially for models with an extended Higgs sector, through resonant  
<sup>529</sup> di-Higgs production. This dissertation will present results for both SM and resonant  
<sup>530</sup> di-Higgs production.

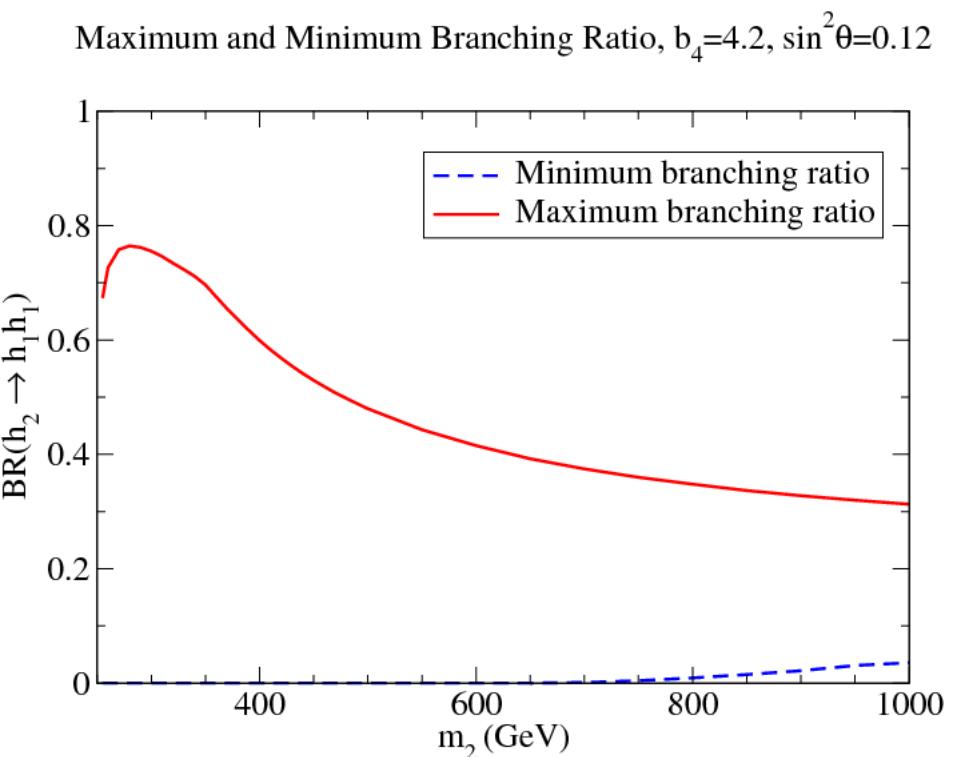


FIGURE 2.7. Maximum and minimum allowed  $BR(h_2 \rightarrow h_1 h_1)$  as a function of  $m_2$  for  $b_4 = 4.2$  and  $\sin^2 \theta = 0.12$ .

531

## CHAPTER III

532

### EXPERIMENTAL SETUP

#### 533 3.1. Hadronic Colliders

534 Hadron colliders use two beams of hadrons, particles made of three quarks,  
535 typically proton-proton or proton-antiproton. The large mass of the hadrons results  
536 in smaller synchrotron radiation in circular accelerators, when compared to leptonic  
537 colliders with equivalent energies, as the radiated power scales as  $\frac{1}{m^4}$ . This allows  
538 hadronic colliders to reach much larger center of mass energies with the same size  
539 circular ring.

540

541 While hadronic colliders typically have larger collision energies, they also have  
542 significantly “messier” collisions. In leptonic colliders, the only final state particles  
543 come from the colliding particles. In hadronic colliders, not all of the constituent  
544 partons of the hadrons interact in the hard collision. This means the exact collision  
545 energy is not known. Each parton only carries a portion of the hadron momentum,  
546 making it impossible to know the longitudinal center of mass.

#### 547 3.2. The Large Hadron Collider

548 The Large Hadron Collider (LHC) is a 27 kilometer ring underneath the  
549 Franco-Swiss border. The LHC accelerates beams of protons (or ions) to a  
550 center of mass energy of up to 13 TeV(5 TeV) in two antiparallel beams  
551 around the ring. The particles are then collided at 4 primary interaction points  
552 each of which has a dedicated detector: ATLAS, CMS, ALICE, and LHCb.

553 In order to get the protons up to the collision energy, the LHC uses a series

554 of smaller accelerators in the injection chain. The start of the chain, and the source  
 555 of protons for the LHC, is the Linear Accelerator 2 (Linac 2) [15]. Hydrogen gas  
 556 is taken from a bottle and the electrons are stripped off by an electric field. When  
 557 only the protons remain, they pass through radiofrequency (RF) cavities. These  
 558 RF cavities are shaped, such that the electromagnetic waves are resonant inside of  
 559 the cavity and build up energy. When charged particles pass through the cavity,  
 560 they feel the force of the electric field and are pushed forward. The field in the RF  
 561 cavities oscillates at a frequency specific to the distance from the previous cavity,  
 562 giving a specific energy to a passing particle depending on the momentum. When  
 563 a particle arrives early to the cavity, the field removes some of the energy from the  
 564 particle, when it arrives late it gets a boost from the cavity to get it back to the  
 565 target energy. When a proton reaches the end of Linac 2, it has an energy of 50 GeV.

566

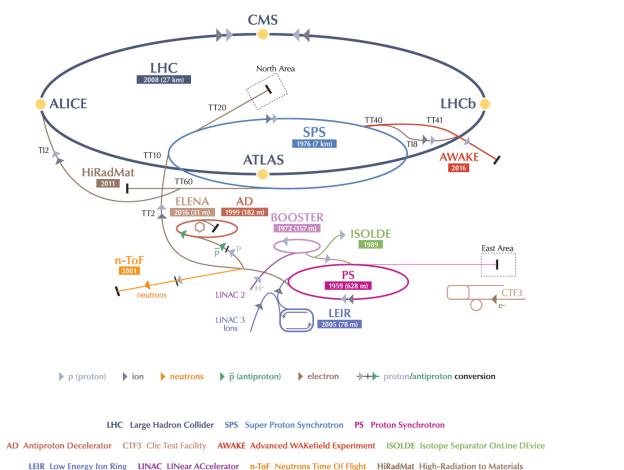


FIGURE 3.1. CERN accelerator complex [16]

567 After the protons leave the Linac 2, they enter the Proton Synchrotron Booster  
568 (PSB)[17]. The PSB accelerates the protons to an energy of 1.4 GeV and are passed  
569 to the Proton Synchrotron (PS) in two batches, 1.2 seconds apart. The PS accelerates  
570 protons to 26 GeV and delivers them to the next step in the injection chain, the Super  
571 Proton Synchrotron (SPS) in a series of 4 batches, 3.6 seconds apart. The SPS is  
572 the second largest accelerator at CERN. As protons pass through, they are boosted to  
573 an energy of 450 GeV [18]. Once the particles reach this energy, they are split and  
574 injected into the LHC in two opposite directions. Once in the LHC, the protons are  
575 brought up to their target energy of 6.5 TeV per beam.

576 **3.2.1. LHC Magnets**

577 Accelerators depend on powerful magnets to bend and focus the colliding  
578 particles. The LHC is the most powerful particle accelerator that has  
579 ever existed. In order to make this possible, the LHC was constructed  
580 with the most advanced magnet technology available at the time. These  
581 magnets are cooled to below 2 K and produce a field to over 8 T.  
582

583 The LHC uses 1104 superconducting, dipole magnets to bend the beam of  
584 particles around the ring, and another 128 in the beam dump. Each dipole  
585 is 15 m long and weigh 35 tonnes. A current of 11,000 Amps pass through  
586 an octant in series to produce the magnetic field. Each octant is powered  
587 independently. A cross section of a dipole can be seen in Figure 3.2. Inside  
588 of these dipoles, there are sextupole, octopole and decapole magnets to correct  
589 for small imperfections in the magnetic field at the outside of the dipole [19].  
590 A key feature of the dipole magnets is the 2-in-1 configuration. Where each  
591 dipole generates a magnetic field in the opposite direction in the two pipes, figure

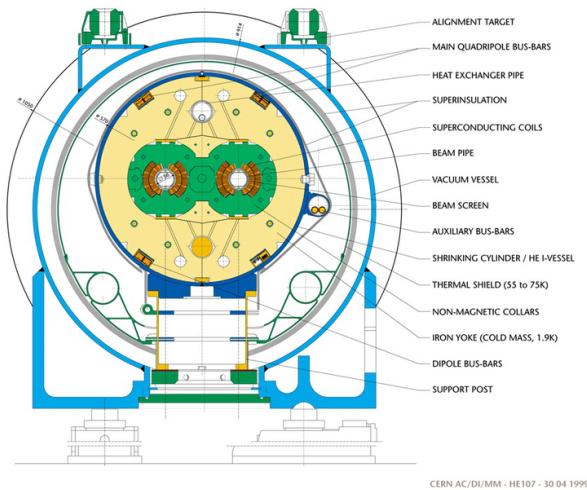


FIGURE 3.2. Cross-section of cryodipole (lengths in mm)

592 3.3. Allowing a single dipole to bend the two beams in opposite directions.

593

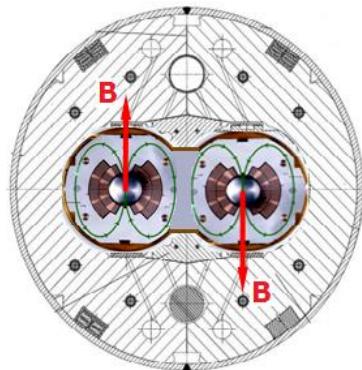


FIGURE 3.3. Magnetic field configuration of the dipole magnets [20]

594 It is also important to focus the beam to a small area, in order to maximize the  
 595 collisions. This is achieved by pair of quadrupole magnets. One magnet focuses in  
 596 the horizontal plane, defocusing in the vertical, and the other focuses in the vertical  
 597 plane, defocusing in the horizontal. These magnets work together to focus the beam  
 598 in both planes.

599 **3.2.2. Luminosity and Pileup**

In a collider, a key statistic is the rate of events produced. This value is called luminosity ( $\mathcal{L}$ ) and is defined by equation 3.1 [21], where  $\sigma$  is the cross section, or probability of collision.

$$\mathcal{L} = \frac{1}{\sigma} \frac{dN}{dt} \quad (3.1)$$

In order to maximize this number, the LHC collides bunches of protons. These bunches are made up of  $1.15 \times 10^{11}$  protons. In the LHC during run 2, the LHC beams had around 2500 bunches organized into a train of bunches, separated by small gaps with a larger gap at the end of the train. The luminosity per proton pair crossing, equation 3.2

$$\mathcal{L} = \frac{1}{4\pi\sigma_x\sigma_y} \quad (3.2)$$

is then multiplied by the number of protons per bunch for the two beams,  $N_1$  and  $N_2$ , and by the number of bunches  $N_b$  and the frequency  $f$ . The use of bunches changes the luminosity calculation to equation 3.3)

$$\mathcal{L} = \frac{N_1 N_2 f N_b}{4\pi\sigma_x\sigma_y} \quad (3.3)$$

600 By integrating the luminosity over the running time of the LHC, you  
601 obtain the total delivered luminosity, Fig 3.4. By multiplying the integrated  
602 luminosity by the probably of a particular final state, or the cross section,  
603 it is possible to obtain the number of times a final state is produced.

604

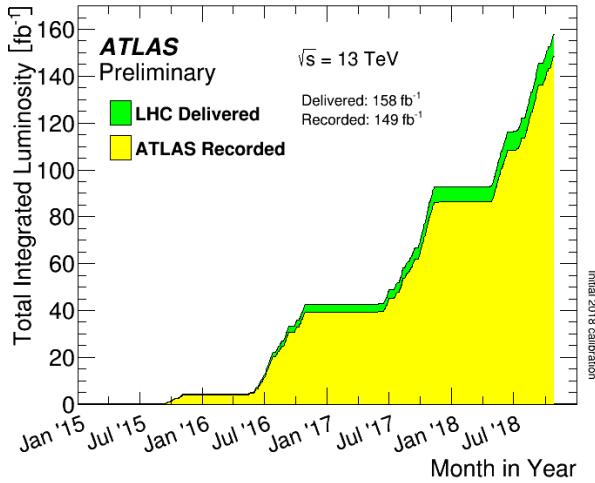


FIGURE 3.4. Cumulative luminosity versus time delivered to ATLAS (green) and recorded by ATLAS (yellow) during stable beams for pp collisions at 13 TeV centre-of-mass energy in LHC Run 2. (figure from the ATLAS Collaboration)

### 605 3.2.3. Pileup

606 By interacting bunches of protons at a time, it is possible for more than one  
 607 pair of protons to undergo an inelastic collision. In fact, during Run 2, the average  
 608 number of interactions per bunch crossing, or pileup, was roughly 32. The profile of  
 609 the pileup for all of Run 2 can be seen in Figure 3.5

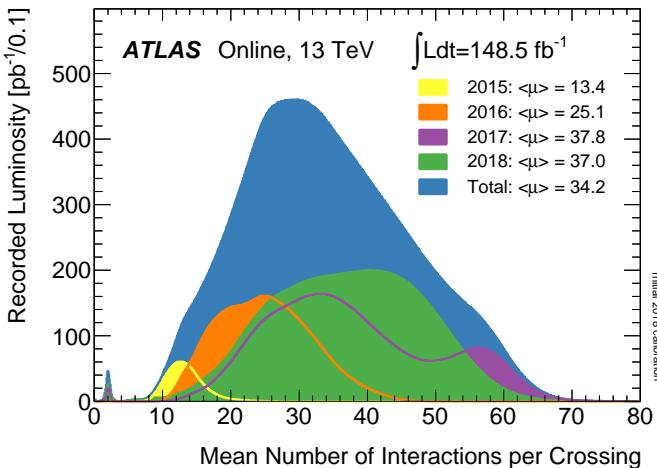


FIGURE 3.5. Shown is the luminosity-weighted distribution of the mean number of interactions per crossing for the 2015-2018 pp collision data at 13 TeV centre-of-mass energy. All data recorded by ATLAS during stable beams is shown, and the integrated luminosity and the mean mu value are given in the figure. (figure from the ATLAS Collaboration)

#### 610 3.2.4. Detector Coordinates

611 Within the ATLAS detector, the interaction point defines the origin of the  
 612 coordinate system. The z-axis, the longitudinal axis, runs along the beam line, the  
 613 positive x-axis points toward the center of the LHC ring, and the positive y-axis points  
 614 toward the surface. The detector is also described in  $r$ ,  $\eta$ ,  $\phi$  coordinates. With the  
 615 transverse plane, the plane perpendicular to the beam line, being described by  $r$  and  
 616  $\phi$ . The radial coordinate,  $r$ , describes the distance from the beam line. The azimuthal  
 617 angle,  $\phi$ , is the angle from the x-axis around the beam line. The final coordinate,  
 618  $\eta$ , is referred to pseudorapidity and is defined as  $\eta = -\ln(\tan(\frac{\theta}{2}))$ . With  $\theta$  being  
 619 the angle from the y-axis. Differences in  $\eta$  are Lorentz invariant under longitudinal  
 620 boosts. Meaning difference in the rest frame of colliding particles are not important  
 621 for massless particles. In ATLAS, the large particle boost allows for all particles

622 to be considered massless when considering  $\eta$ . Additionally, particle are produced  
 623 uniformly in  $eta$ . For these reasons,  $\eta$  is preferred over  $\theta$ . A pictoral representation  
 624 can be seen in Figure 3.6. The variable  $\Delta R = \sqrt{\eta^2 + \phi^2}$  is often used to describe  
 625 the distance between detector objects.

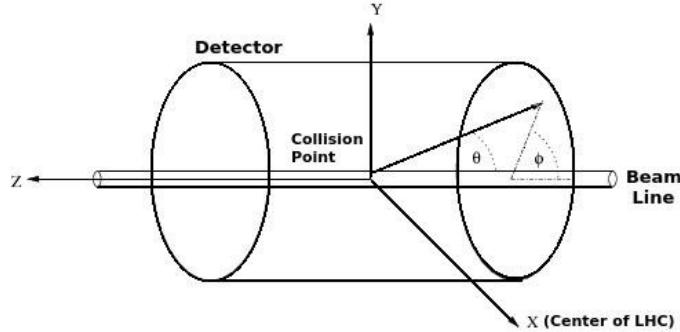


FIGURE 3.6. Detector coordinate system [22]

### 626 3.3. Detector Overview

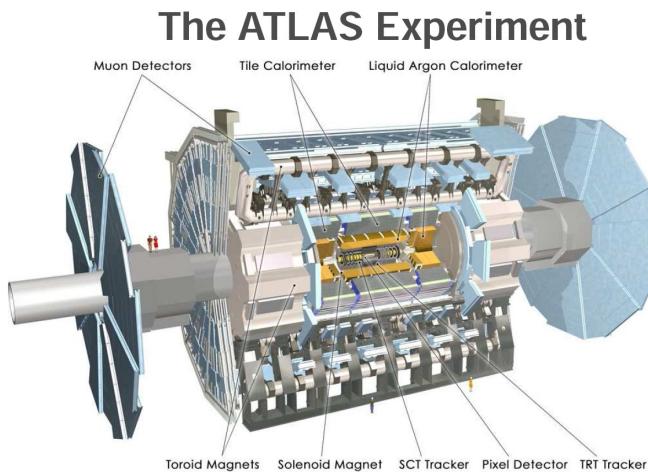


FIGURE 3.7. The ATLAS detector [23]

627 The ATLAS detector, Figure 3.7, is a general purpose detector and the largest  
 628 on the LHC. It is made up of concentric subsystems, each with a specialized task:

629 the inner detector, which is responsible for measuring the charge and momentum of  
 630 charged particles; the calorimeters, which are responsible for measuring the energy  
 631 of different electromagnetic and hadronic particles; the muon spectrometer, which  
 632 measures the momentum of minimum ionizing particles (MIP), like muons; and the  
 633 magnet system, which is responsible for bending the charged particles in the detector,  
 634 allowing their charge and momentum to be measured. The subdetectors feed into a  
 635 vast Trigger and Data Acquisition (TDAQ) system that is responsible for selecting  
 636 collision events with interesting characteristics and reading out detector elements.

### 637 **3.3.1. The Inner Detector**

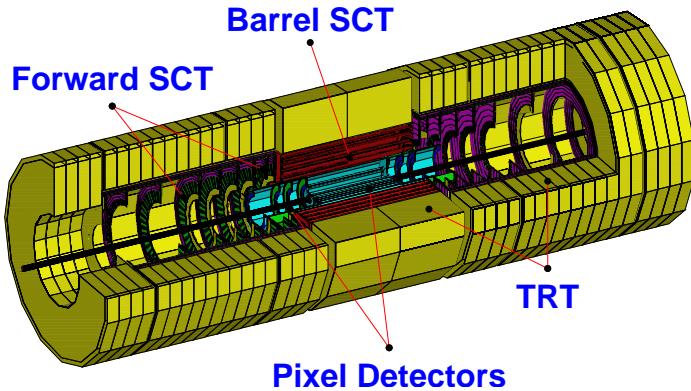


FIGURE 3.8. Cross section of the Inner Detector.

638 The inner detector (ID), Figure 3.8, is the closest system to the beam pipe  
 639 and is made of 4 separate pieces. In order of distance from the beam pipe: The  
 640 Insertable B-Layer (IBL) [24], the Pixel Detectors, the Semiconductor Tracker (SCT),  
 641 and the Transition Radiation Tracker (TRT) [25]. These subsystems work together  
 642 to give charged particle tracking within the pseudorapidity range of  $|\eta| < 2.5$ . When  
 643 a charged particle passes through the silicon semiconductor detector in the IBL,  
 644 Pixel and SCT, an electron-hole pair is created in the silicon. These electron-hole

645 pairs drift toward the charged readout electrons on the surface of the detector.  
646 This gives a "hit" in the detector. The inner detector is immersed by a 2 T  
647 solenoid magnet, section 3.3.4. The magnetic field causes charged particles to  
648 curve as they pass through the ID, leaving hits along the way. These hits are  
649 connected together into "tracks" that show the trajectory of the particle as it  
650 passes through. The radius of curvature and direction of the track gives the sign  
651 of the charge, positive or negative, along with a momentum measurement of the  
652 particle. The other task of the ID is vertexing, or determining if the transient  
653 particle came from the interaction point of a slightly displaced point. This is  
654 used to identify long lived particles, like bottom or charm quarks, which travel  
655 a small distance before decaying. This is discussed further in section 4.2.4.

656 The IBL is the newest addition to the ID, being installed during the 2016  
657 shutdown. It is placed directly outside the beam pipe in order to maintain good  
658 vertexing and b tagging in increased pileup environments. In order to facilitate the  
659 insertion of the IBL, the beam pipe inner radius was decreased by 4 mm (from 29 mm  
660 to 25 mm). The IBL utilizes planar sensors, similar to the Pixel Detector, and 3D  
661 sensors, allowing electrons to interact with the bulk of the sensor as opposed to just  
662 the surface, and functions as a fourth pixel layer of the Pixel Detector. The addition  
663 of the IBL significantly improves the vertexing of the ID

664 The Pixel Detector is a network of high granularity, silicon pixels which measure  
665 the 2D position of passing charged particles. The silicon pixels are n-doped silicon  
666 wafers. A high voltage is applied to the wafer and when a charged particle  
667 passes through the silicon an electron hole pair is created. The electron drifts  
668 to the electrode and creates a signal that is read out by the electronics. The  
669 Pixel detector central barrel is divided into 3 cylindrical layers, the innermost  
670 layer is the B-layer, followed by Layer 1 and Layer 2. Each is covered in

671  $50\mu m$  x  $400\mu m$  silicon pixels. In order to ensure complete coverage, an end  
672 cap module is placed on each side of the barrel. The end-caps consist of 4  
673 wheels, each with an inner and outer ring of trapezoid shaped silicon detectors.

674 The SCT is made of a barrel detector and two end-cap detectors.  
675 The barrel SCT has 4 cylindrical layers made up square modules covered  
676 with silicon microstrip detectors. The end-cap SCT tracker is made of  
677 rings of SCT modules with either silicon or gallium arsenide. These rings  
678 are arranged into 9 wheels on each side of the barrel. In total, the  
679 SCT contains  $61m^2$  of silicon detectors with 6.2 million readout channels.

680 Outside of the silicon detectors lies the TRT. The TRT is a straw detector  
681 comprised of 50,000, 4mm diameter straws in the barrel and 32,0000 radial straws in  
682 the end-caps. There are 42,0000 electronic channels, which give a spatial resolution  
683 of  $170\mu m$  per straw. The straws are filled with various mixtures of xenon argon,  
684 carbon dioxide, tetrafluoromethane and nitrogen gas. When a charged particle  
685 passes through the radiator between the straws, made of polypropylene, they emmit  
686 transition radiation photons. These photons ionize the gas in the straws and the  
687 free electrons are attracted to the positively charged wire and produce a signal that  
688 is later amplified and read out. The xenon in the gas mixture allows for accurate  
689 particle identification from the transition radiation photon detection. Transition  
690 radiation is emitted when a particle moves between two materials with different  
691 dielectric constants and is proportional to the Lorentz factor of the particle [26]. This  
692 gives a good discrimination between electrons and charged pions. This entire system  
693 is enclosed within a solenoid magnet to bend the charged particles inside the ID.

694

695    **3.3.2. Calorimeters**

696       Outside of the solenoid magnet lies the calorimetry system. The  
697       calorimeters are responsible for measuring the energy of both charged and  
698       neutral particles, with the exception of MIPs and non-interacting particles  
699       such as neutrinos. The calorimeters can be broken into two distinct  
700       pieces, the liquid Argon calorimeter (LAr)[27] and the tile calorimeter [28].

701       The Liquid Argon (LAr) calorimeter is a sampling calorimeter that is used  
702       for electromagnetic calorimetry for the entire range of acceptance ( $|\eta| < 4.8$ ). It  
703       is also used for hadronic calorimetry for higher pseudorapidity  $1.4 < |\eta| < 4.8$ . The  
704       central barrel of the calorimeter ( $|\eta| < 1.4$ ) is made up of 1024 lead-stainless-steel  
705       converters with copper-polyimde multilayer readout boards. The plates and readouts  
706       are arranged in an “accordion-shaped” geometry. This allows for complete azimuthal  
707       coverage with no gaps, giving an electromagnetic energy resolution that is uniform  
708       in azimuth. In between the accordion layers, liquid argon is used as the active  
709       medium. The system is enclosed in a cryostat to maintain the temperature of  
710       the detector. The LAr barrel is divided radially into 4 sampling layers. The  
711       granularity of the layers can be found in Figure 3.9. The layer closest to the  
712       beamline is the Presampler. This layer sits inside of the cryostat and is responsible  
713       for correcting for the energy loss in front of the calorimeter (the same is done in  
714       the endcap). Inside the cryostat, there are 3 additional layers. The thickness of  
715       the layers is often described in terms of radiation lengths  $\chi_0$ . Where a radiation  
716       length is the distance a electron travels before it loses approximately 1/2 of it’s  
717       energy to photon emission. The front layer has a thickness of  $4.3\chi_0$ , followed by  
718       the middle layer with a thickness of  $16\chi_0$  and the back layer of thickness  $2\chi_0$ . The  
719       shower maximum is contained in the second layer of the calorimeter, resulting in

720 the bulk of the energy being absorbed in that layer. The design of the calorimeter  
 721 allows for an energy resolution for electrons of  $\sigma_E/E \sim 10\%/\sqrt{\frac{E}{GeV}} \oplus 0.7\%$  [29].

722

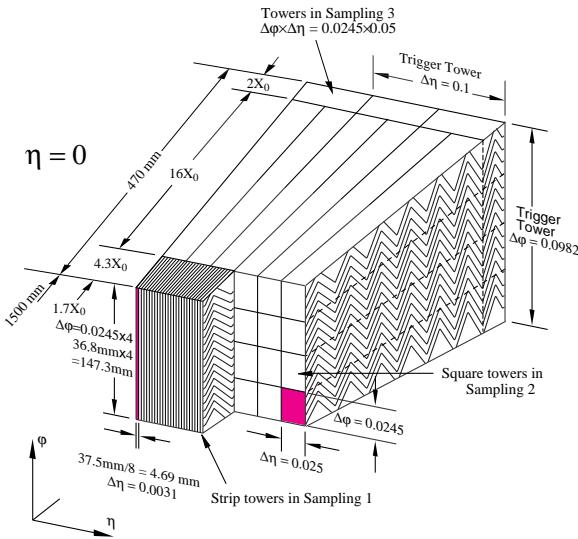


FIGURE 3.9. Sketch of the accordion structure of the EM calorimeter

723 Forward from the barrel, there are two electromagnetic endcap (EMEC) wheels  
 724 with a similar accordion structure to the barrel. One covering  $1.4 < |\eta| < 2.5$  and one  
 725 from  $2.5 < |\eta| < 3.2$ . Outside of the EMEC is the Hadronic endcap (HEC). This is also  
 726 a copper-LAr sampling calorimeter. It has a simpler parallel plate design. Finishing  
 727 out the LAr calorimeter is the Forward Calorimeter (FCal), which is contained in

728 the endcap cryostat. This calorimeter is in the very forward region of the detector.  
729 In this region, the particle flux is very high, so a dense calorimeter is necessary  
730 to avoid energy leaking into other pieces of the detector. There are 3 layers in  
731 the FCAL, the first is made of copper and the other two are made of tungsten.  
732 They are matrices of metal with concentric tubes filled with Argon, see Figure 3.10

733

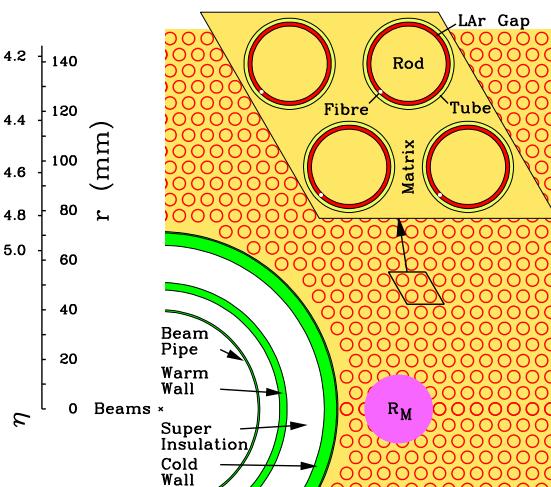


FIGURE 3.10. Sketch of the matrix and rods in the forward calorimeter

734 In the central region  $|\eta| < 1.7$ , the tile calorimeter (TileCal) is responsible for  
735 the hadronic calorimetry. The TileCal is a sampling calorimeter with alternating

736 iron plate absorbers and plastic scintillating tiles, the orientation can be seen in Fig  
 737 3.11. The scintillating tiles are placed perpendicular to the beamline and are read  
 738 out by wave-length shifting fibers on both ends of the module. The light is passed to the  
 739 photomultiplier tubes (PMTs) on the outside of the system, and then passed to the  
 740 front-end electronics. It has a fixed central barrel and 2 extended barrel sections that  
 741 can be moved. The TileCal has a depth of  $7.4\lambda$ , where  $\lambda$  is the nuclear interaction  
 742 length, the mean distance a hadronic particle travels before it undergoes an inelastic  
 743 interaction. The readout has a granularity of  $0.1 \times 0.1(\eta \times \phi)$

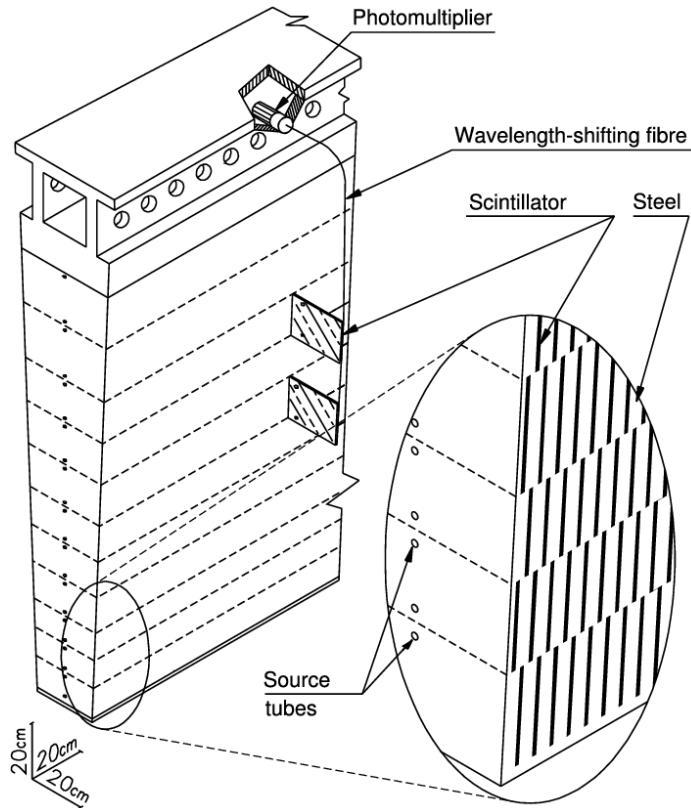


FIGURE 3.11. Schematic showing the mechanical assembly and the optical readout of the Tile Calorimeter, corresponding to a  $\phi$  wedge. The various components of the optical readout, namely the tiles, the fibers and the photomultipliers, are shown. The trapezoidal scintillating tiles are oriented perpendicular to the colliding beam axis and are read out by fibers coupled to their non-parallel sides [30]

744 **3.3.3. Muon Detectors**

745 To detect muons, ATLAS uses four different technologies [31]. For precision  
746 energy and position measurements, monitored drift tubes (MDT) and cathode strip  
747 chambers (CSC) are used. The CSCs are used in regions of high flux, where the  
748 MDTs are not suitable. For the muon trigger system, a fast detector is needed  
749 to keep up with the timing requirements and high collision rate of the LHC. In  
750 the central region, resistive plate chambers (RPC) are used, while in the forward  
751 region, where flux is higher, thin gap chambers (TGC) are used. The muon system,  
752 much like the ID utilize a magnetic field to determine the charge and momentum  
753 of passing particles. The magnet system is further discussed in Section 3.3.4.

754

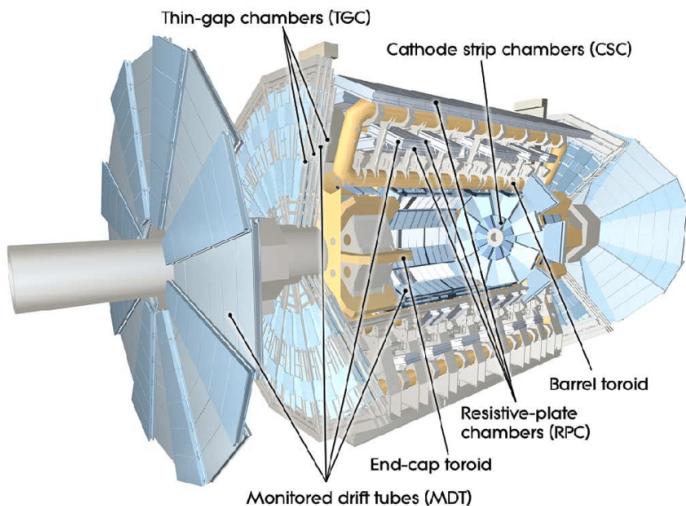


FIGURE 3.12. Cut-away view of the ATLAS muon system [32]

755 The MDTs are made up of 6 parallel layers of cylindrical aluminum drift  
756 tubes with a tungsten-rhenium wires. The drift tubes are filled with a mixture  
757 of argon, nitrogen and methane. The tubes are assembled on a support  
758 spacer and are monitored for deformation by a built-in optical system, hence

759 the **monitored** drift tubes. The monitoring ensures a high accuracy in the  
760 position of the measurement points. Allowing the MDTs to achieve a sagitta  
761 precision of  $50\mu m$  and thus a momentum precision at 1 TeV of  $\frac{\Delta p_T}{p_T} = 10\%$

762 While the MDTs are very good at precision measurements. However, they  
763 are not appropriate in areas with high rate counts ( $> 200 Hz/cm^2$ ) due to their  
764 large diameter and high operating pressure. This is the case for the first layer of  
765 muon measurement with pseudorapidities of  $|\eta| > 2.0$ . For this region, CSCs are the  
766 spectrometer of choice. CSCs are multiwire proportional chambers with a cathode  
767 strip readout. They give good single and two track resolution in this high rate region.

768 In the barrel region, the muon trigger system employs RPCs, a low  
769 occupancy chamber with fast response. RPCs are gaseous parallel-plate  
770 detectors of Bakelite with a coating of linseed oil based paint. The  
771 system can operate in two modes, avalanche and streamer. In streamer  
772 mode, a large potential across the plates generates a discharge around the  
773 ionizing particle. For avalanche mode, a smaller potential difference and  
774 large signal amplification in the electronics allows for increased rate capability.

775 Finally, in the end-cap of ATLAS, TGCs provide two important components.  
776 For the trigger system, TGCs have good timing resolution compared to the MDTs  
777 and can deal with a rate of up to  $100 KHz/cm^2$ . For measurement, TGCs provide  
778 the azimuthal coordinate to compliment the bending coordinate from the MDTs.  
779 The TGCs are made up of anode wires and graphite cathodes in between layers of  
780 fiberglass laminate.

781 **3.3.4. Magnet System**

782 The signature piece of the ATLAS detector is the Large Toroid magnet system.  
783 The Toroid has eight coils in the barrel and two endcaps, with eight coils each

[33]. Figure 3.13, 1st ref fig 2-1. The Toroid system provides a magentic field of 3.9 T(4.1 T) in the barrel (end-cap) to the muon system. The coils of the three toroids are assembled radially and symmetriclly around the beam axis. A toroid has two advantages over a solenoid. The first is the field at the edges of the detector remains perpendicular to the outgoing particles, allowing for a better measurement at high pseudorapidity. The other advantage is cost. It takes much less material to build a large toroid than an equivalently sized solenoid. This allows ATLAS to have a very large volume for particle bending in the muon system.

792

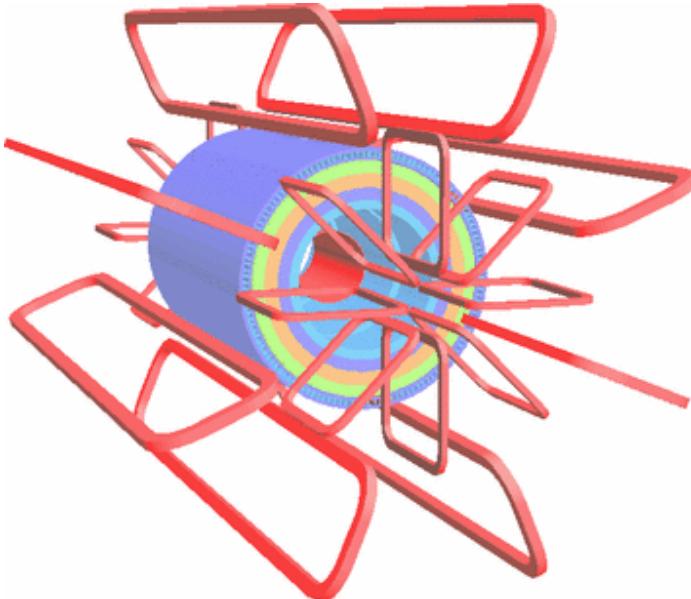


FIGURE 3.13. Geometry of magnet windings and tile calorimeter steel [32]

Along with the toroid, ATLAS has a solenoid magnet inside of the calorimeter. This solenoid provides a 2 T magnetic field to the ID for bending of charged particles. The solenoid is a single layer coil in a suporting cylinder. It is supported by the LAr cryostat. It is very important the solenoid is thin, in order to minimize the amount

797 of material in front of the calorimeters. To achieve this, the vacuum of the solenoid  
 798 and LAr are combined into one and the coil is designed to be as thin as possible.

799 **3.3.5. Trigger System**

800 The LHC delivers collisions at a rate of 40MHz. With each raw event being  
 801 about 1.6MB [34], this would give an output rate of 64TB/s. This rate of output is  
 802 beyond what can be handled by the computing resources available. In order to reduce  
 803 this rate to a manageable level, ATLAS employs a two level trigger system. The first  
 804 level trigger is a hardware based trigger, referred to as the L1 trigger. This trigger  
 805 lowers the rate to around 75kHz. This is sent to the second level, software based,  
 806 trigger, the High Level Trigger (HLT), where the rate is further reduced to below  
 807 2kHz for full event readout. When combined with the partial event readout, the total  
 808 bandwidth around 3GB/s Fig 3.14 illustrates the ATLAS trigger system data flow.

809

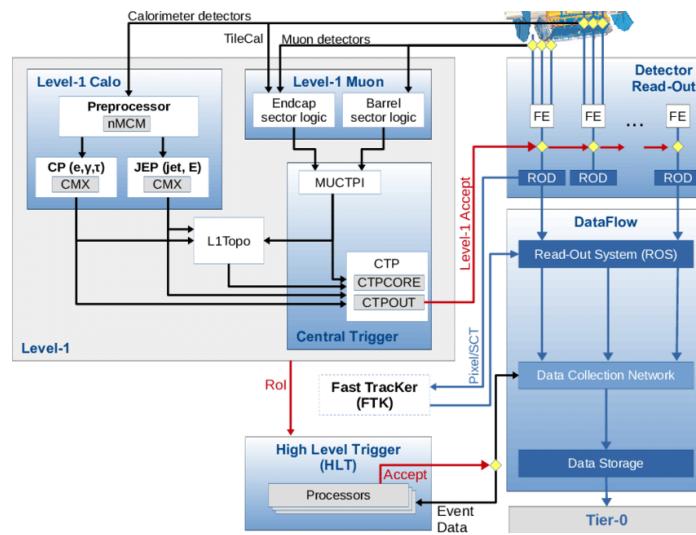


FIGURE 3.14. Schematic layout of the ATLAS trigger and data acquisition system in Run-2.[35]

810        The L1 trigger begins with signals from either the calorimeters of the muon  
811 detectors. A signal from the calorimeter is sent to Level-1 Calo (L1 Calo)  
812 system. The L1 Calo system uses low granularity calorimeter information to  
813 identify Regions of Interest (RoIs) for electrons, photons, taus, jets, as well as  
814 high total energy and missing transverse energy ( $\cancel{E}_T$ ). L1 Calo received and  
815 upgrade in Run 2 in the form of a new Multi-Chip Module (nMCM). This  
816 module allows for L1 Calo to suppress the effects of pile-up on the system.

817        Signals from the muon system are fed into the L1 Muon system. L1  
818 Muon uses information from the RPC and TGC in the barrel and end-  
819 caps. In Run 2, the muon end-cap triggers required a coincident hit in the  
820 innermost muon chamber to reduce the fake muon rate in the forward regions.

821        The signals from L1 Calo and L1 Muon are passed to the L1 Topological trigger  
822 processor (L1 Topo) and the Central Trigger Processor (CTP) simultaneously. In the  
823 case of L1 Muon, the signals from the barrel and the end-cap are merged in the muon  
824 Central Trigger Processor Interface (MuCTPi). In L1 Topo, a new system for Run 2,  
825 kinematic information from L1 Calo and L1 Muon to make topological selections,  
826 like angular separation, invariant mass requirements, and total  $\cancel{E}_T$ , at level 1.

827        Information from L1 Topo, and directly from L1 Calo and L1 Muon,  
828 are sent to the CTP where the L1 trigger Accept and the LHC timing  
829 information are fed to the sub-detector readout system. The CTP also  
830 sends the RoIs to the HLT to use as inputs for higher level algorithms.

831        The HLT trigger is made up of 40000 processing cores which run around  
832 2500 independent trigger chains. A chain is an offline-like algorithm run over the  
833 RoIs from L1. Partial event reconstruction, and even full event reconstruction  
834 can be done within the HLT depending on the event stream. There are 4  
835 types of data streams depending on the purpose of the data, physics analysis,

836 trigger level analysis, monitoring, or detector calibration. For the physics analysis  
837 stream, the full event is written out. For the other streams, only partial  
838 event information is written out, saving bandwidth to allow for higher rates.

839 The configuration of the trigger system is defined by the trigger menu.  
840 The trigger menu is made from the primary physics analysis triggers along  
841 with support triggers for efficiency and monitoring, alternative triggers, backup  
842 triggers and calibration triggers [36]. These triggers, L1 items and HLT triggers,  
843 are regulated by prescale factors. A prescale factor is used to reduce the  
844 rate of selected events. For a prescale of  $N$ , only  $1/N$  events that meet  
845 the trigger requirement are accepted. Prescales can be set to individual L1  
846 items and to specific HLT chains independently and as a function of luminosity.  
847 This allows triggers to maintain a specific rate regardless of the luminosity.  
848 Typically, the L1 primary physics triggers are unprescaled, this means, the  
849 other triggers are prescaled to a low rate to give priority to the physics  
850 triggers. As a result, most of the bandwidth is occupied by the physics chain.

851

854 **4.1. Simulation**

855 In order to draw conclusion from ATLAS data, it is necessary to compare to  
856 theoretical predictions. For particle collisions, it is not practical to create exact  
857 predictions, especially including detector effects such as resolution. To get the best  
858 estimate of these effects, ATLAS uses the Monte Carlo (MC) method to simulate  
859 data and detector response to the incident particles. This is done in multiple steps  
860 as illustrated by figure 4.1. These steps are the simulation of the hard process,  
861 where the deep inelastic collision simulated using the initial state (Parton Distribution  
862 Functions) and interaction amplitudes; the parton shower; the hadronization; the  
863 detector simulation; and finally the reconstruction. These steps together form the  
864 complete MC simulation of ATLAS data.

865

866 **4.1.1. Parton Distribution Functions**

867 At the energies at the LHC, collisions usually do not involve entire protons.  
868 Instead, they involve constituents known as partons. Protons, while often described  
869 as two up quarks and a down quark, contain a sea of gluons. This sea of gluons  
870 also creates many virtual quark-antiquark pairs. The up and down quarks are the  
871 outer, or valance, quarks. These valance quarks are the primary role players in  
872 shallow inelastic interactions. At the LHC, the collision energies are sufficient for  
873 deep inelastic scattering, where the affects of the internal quarks and gluons are non-  
874 trivial. For di-Higgs events, the dominant form of production is gluon-gluon fusion

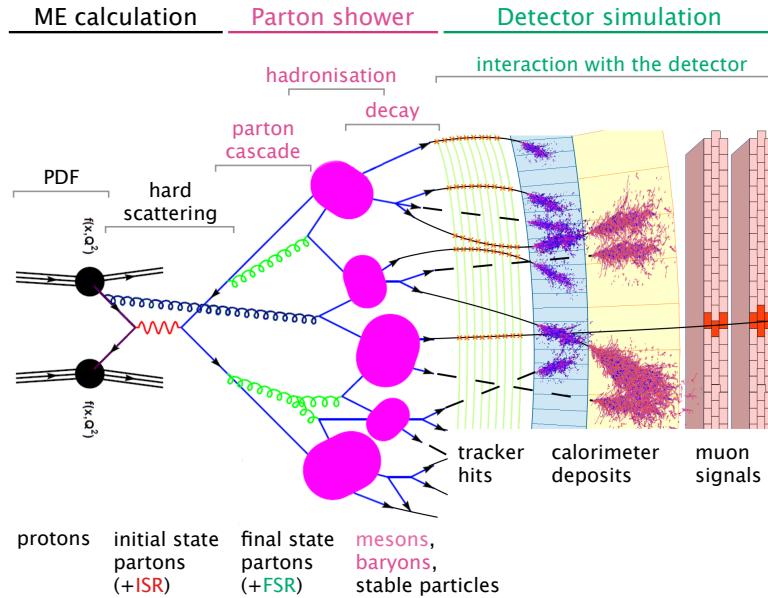


FIGURE 4.1. Pictorial representation of how an event is generated [37]

875 (ggF) This internal structure of the proton is described by a Parton Distribution  
 876 Function (PDF), figure 4.2. A PDF shows the probability density of finding a parton  
 877 carrying a momentum fraction  $x$  at a squared energy scale.

878

#### 879 4.1.2. Hard Scattering

880 The hard scattering process can be described using Feynman diagrams. These  
 881 diagrams are a pictorial representation of amplitudes. These amplitudes go into  
 882 calculating the matrix elements (ME) of various interactions. In the event generation,  
 883 these MEs are calculated to a specified order in perturbation theory. Common  
 884 examples are leading order(LO), next-to-leading order(NLO) , and so on. The higher  
 885 the order of the calculation, the more accurate the predictions. However, higher  
 886 orders can be extremely hard to theoretically calculate, often restricting the level of

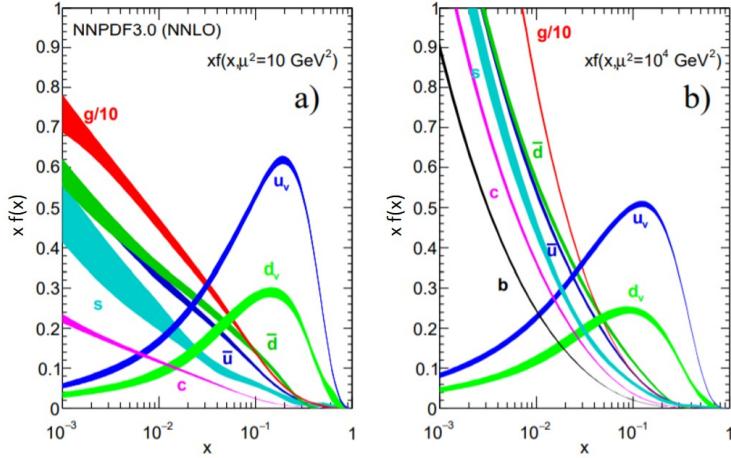


FIGURE 4.2. The bands are  $x$  times the unpolarized parton distributions  $f(x)$  (where  $f = u_v, d_v, \bar{u}, \bar{d}, s \simeq \bar{s}, c = \bar{c}, b = \bar{b}, g$ ) obtained in NNLO NNPDF3.0 global analysis at scales  $\mu^2 = 10\text{GeV}^2$  (a) and  $\mu^2 = 100\text{GeV}^2$  (b), with  $\alpha_s(M_Z^2) = 0.118$ .

887 the event generator.

888

#### 889 4.1.3. Parton Shower Calculation

890 After the ME generator, the hard partons are used as the inputs to the Parton  
 891 Shower (PS) calculation. The PS calculation takes the hard scattering process from  
 892 the event generator and calculates the parton shower. In addition to calculating  
 893 the parton shower, the PS also calculates additional hard radiation processes not  
 894 included in the base interaction. Colored particles can spontaneously emit gluons.  
 895 These gluons, in turn, create either more gluons, or quark-antiquark pairs. This can  
 896 happen either before (ISR) or after (FSR) the hard scattering process. Along with the  
 897 ISR and FSR emittance, the PS generator can also describe the hadronization and  
 898 subsequent decay of the hadrons into the final state particles. The precision of the  
 899 PS generators are described similarly to the ME, with their contributions coming in

900 as leading log (LL), next-to-leading-log (NLL), etc. for the parton showering process.

901

#### 902 **4.1.4. Detector Simulation**

903 The MC simulation up to this point can be done with generators that are written  
904 outside of the ATLAS collaboration. These generators are used by ATLAS to simulate  
905 the underlying processes which are fed into the detector simulation software. ATLAS  
906 uses GEANT4 to handle this propagation[38] through the detector. GEANT4 uses a  
907 detailed geometric description of the ATLAS detector to simulate particle interactions  
908 with the matter in the detector.

909 The final result of the MC event generation is a set of simulated data that  
910 resembles actual data from the p-p collisions in the ATLAS detector.

#### 911 **4.1.5. Reconstruction**

912 Once the data has been simulated, it is necessary to transform it into meaningful  
913 objects through reconstruction. This involves taking the raw energy deposits in each  
914 of the subdetectors and turning them into tracks or calibrated energy deposits. These  
915 groomed objects are then used to build a picture of the physics event through particle  
916 identification and event reconstruction.

### 917 **4.2. Particle Identification**

918 For all events, either MC or actual collision data, it is important to be able  
919 to identify and reconstruct the underlying physics event. In particle collisions, the  
920 energy from the final state particles is deposited in the various subdetectors within  
921 ATLAS. These energy deposits must be translated to physically meaningful objects.  
922 This is the task of the event reconstruction, to use the ATLAS detector to recreate

923 the final state particles for any given interaction. For this analysis, the final state  
 924 particles present in the signal events are a lepton, either an electron or a muon; a  
 925 neutrino, in the form of missing transverse energy; two light flavor quarks; and two  
 926 b quarks. Each of these particles has a particular signal in each of the subdetectors,  
 927 figure 4.3.

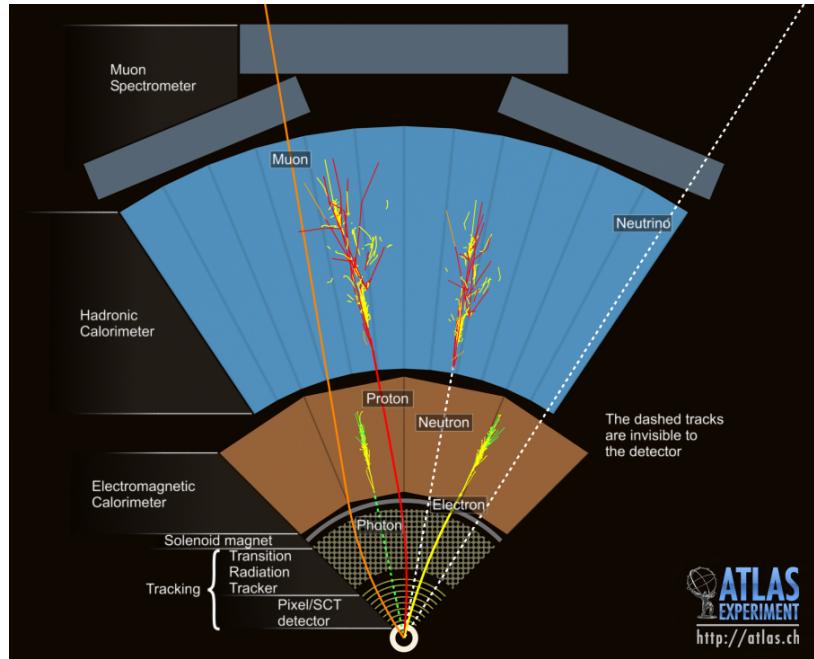


FIGURE 4.3. Event Cross Section in a computer generated image of the ATLAS detector [39]

#### 928 4.2.1. Electrons

929 Electrons are reconstructed by fitting a track using the Inner Detector and  
 930 matching this track to an energy cluster in the EM calorimeter[40]. As an electron  
 931 passes through the EM calorimeter, it produces Bremsstahlung radiation photons.  
 932 These photons then convert back to electron-positron pairs and the process repeats.  
 933 This shower of electrons, positrons, and photons give the signature energy cluster in

934 the calorimeter. Particles with the required Inner Detector track and matching EM  
935 energy cluster are selected as electron candidates.

936 Electron identification algorithms are applied to these electron candidates. These  
937 algorithms separate prompt, isolated electron candidates from backgrounds such as  
938 converted photons and misidentified jets. The electron identification algorithm uses  
939 the energy-momentum ratio, shower shape, track and track-to-cluster matching to  
940 identify electron candidates, with  $E/p$  being the most important discriminant. There  
941 are three identification working points for electron identification: Loose, Medium,  
942 and Tight. Where the operating points with higher background rejection are a subset  
943 of electron candidates with lower background rejection.

944

#### 945 4.2.2. Muons

946 The Muon Spectrometer specializes in muon detection and precision momentum  
947 measurement. Unsurprisingly, this makes the Muon Spectrometer (MS) a vital part  
948 of muon identification, but it is not the only subdetector used. The Inner Detector is  
949 also has an important part in reconstructing muons. In ATLAS, muon reconstruction  
950 is performed independently in the Inner Detector and the MS. The information is then  
951 combined to form the muon tracks. In the Inner Detector, the muons are reconstructed  
952 similarly to any other charged particle.

953 In the MS, the reconstruction looks for a hit pattern within each chamber to form  
954 segments [41]. The MDT segments are combined using a straight-line fit. Segments  
955 in the CSCs are combined using a combinatorial search in the  $\eta$  and  $\phi$  planes.

956 Muon candidates are built by fitting together hits from segments in different  
957 layers. A combinatorial search, using segments in the middle layer as seeds, is  
958 performed. The inner and outer layers are then used as seeds as the search is extended.

959 A minimum of 2 segments are required to build a track. It is possible for a segment to  
960 be included in multiple tracks, an overlap removal algorithm selects the best assigned  
961 track or can allow for a segment to be shared between two tracks. A global  $\chi^2$  fit is  
962 performed on the hits of each track. If the  $\chi^2$  of the fit passes a selection criteria, the  
963 track is accepted.

964 The information from the Inner Detector and the MS are then combined to give a  
965 muon signature. The combination method depends on the information available. The  
966 main method used is the Combined Muon reconstruction, where track reconstruction  
967 is performed in the Inner Detector and MS independently. Most of these muons are  
968 reconstructed using an "outside-in" reconstruction. This means tracks in the MS are  
969 extrapolated inward and matched to an Inner Detector track.

970

#### 971 **4.2.3. Jets**

972 Quarks very quickly undergo hadronization, with only the top quark decaying  
973 before hadronizing. This makes it impossible to measure a singular quark. Instead,  
974 collections of hadrons are formed and these are what deposit energy in the ATLAS  
975 detector. These collections of energy are called jets and can be made from various  
976 detector objects. In this analysis in particular, two different types of jets are used:  
977 calo-jets, jets constructed from energy deposited in the calorimeters; and track-jets,  
978 jets constructed from tracks in the Inner Detector.

979 Since a jet is not a physical object, rather a collection of energy deposits, there  
980 are many ways to define a jet. A jet algorithm groups objects together, based on  
981 some criteria, into a collection of jet objects. Two important characteristics of any  
982 jet algorithm are Infrared (IR) Safety and Collinear (CL) Safety. For a jet algorithm  
983 to be IR Safe, the addition or subtraction of small energy deposits will not change

984 the jet collection. A jet algorithm is CL Safe if splitting or merging high transverse  
 985 momentum particles does not change the jet collection. Figure 4.4 illustrates both  
 986 IR and CL Safety.

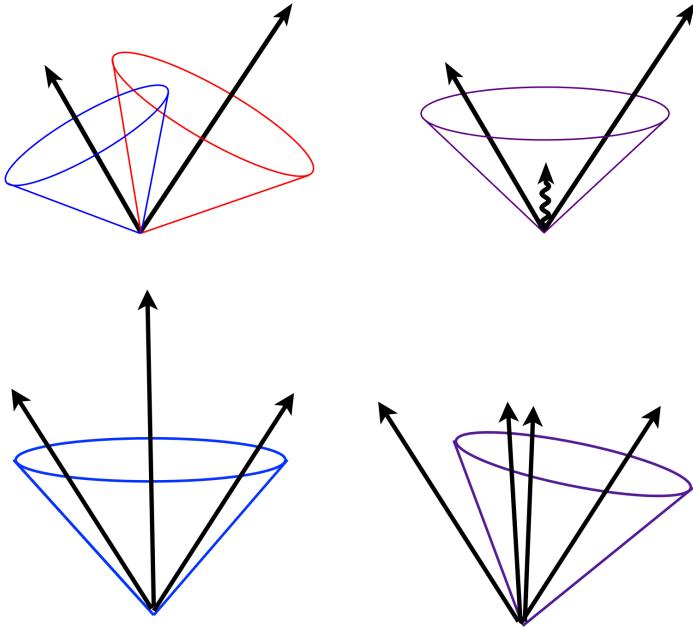


FIGURE 4.4. Illustration of the infrared sensitivity of a cursory designed jet algorithm (top). Illustration of the product of a collinear unsafe jet algorithm. A collinear splitting changes the number of jets (bottom). [42].

Some examples of jet algorithms are visualized in figure 4.5. For this analysis, the anti- $k_t$  algorithm is selected. In addition to being IR and CL safe, the anti- $k_t$  algorithm gives roughly circular jets. This makes calculating the energy density much easier than non-circular jets. The anti- $k_t$  algorithm calculates the distance between objects  $i$  and  $j$  and  $i$  and the beam  $B$ . If  $d_{ij}$  is smaller than  $d_{iB}$ , the objects are combined. If  $d_{iB}$  is smaller the object is removed and the algorithm is rerun. An important distinction between anti- $k_t$  and other jet algorithms is the definition of the

distances  $d_{ij}$  and  $d_{iB}$

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2}, \quad (4.1)$$

$$d_{iB} = k_{ti}^{2p}$$

Where  $k_{ti}$  is the transverse momentum,  $\Delta$  is the distance between objects, and  $p = -1$ . The anti- $k_t$  algorithm has a radius parameter  $R$ .  $R$  acts as a cutoff radius for energy clustering and is not strictly a radius, as objects with a  $\Delta > R$  can still be clustered together. The track-jets used in the analysis have  $R = 0.2$ , while the  $R = 0.4$  and  $R = 1.0$  calo-jets are used.

992

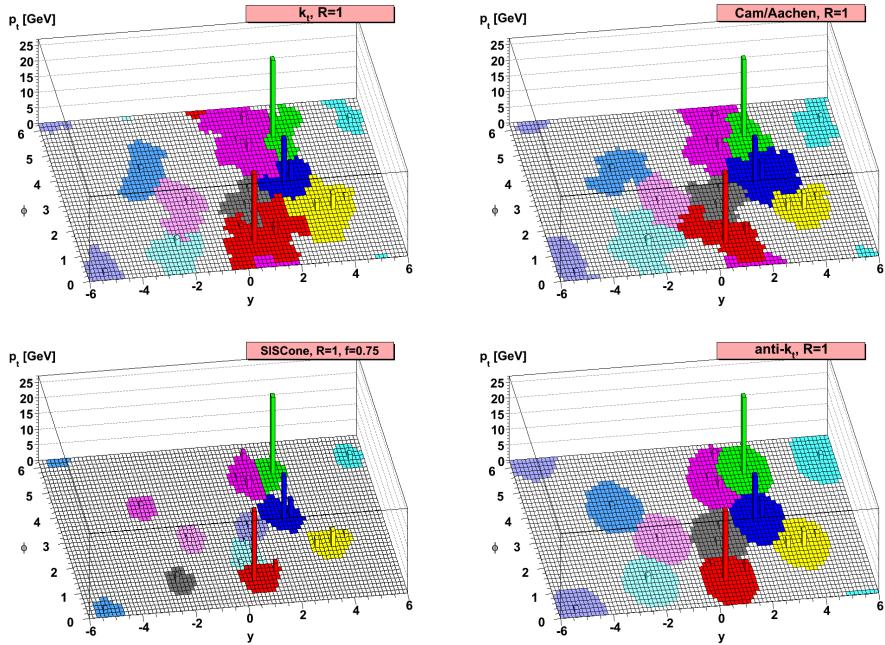


FIGURE 4.5. A sample parton-level event, together with many random soft “ghosts”, clustered with four different jets algorithms, illustrating the “active” catchment areas of the resulting hard jets[43].

993 **4.2.4. b Tagging**

994 B-Hadrons, composite particles that contain b quarks, travel a small distance  
995 before they decay. This means the track from the decay products can be traced back  
996 to the point at which the b-hadron decays, called the secondary, or displaced, vertex,  
997 figure 4.6 . This displaced vertex is used to tag jets that are likely to come from b  
998 quarks through b-tagging algorithms.

999 In this analysis, the MV2c10 is used to tag b-jets [44]. MV2 is a multivariate  
1000 discriminant that combines 3 b-tagging algorithms. The c10 signifies a 10% c-jet  
1001 fraction in the background training sample. The three algorithms that are used  
1002 as inputs to the MV2 discriminant are: an impact parameter-based algorithm, an  
1003 inclusive secondary vertex reconstruction algorithm, and a decay chain multi-vertex  
1004 reconstruction algorithm. For this analysis, the 85% fixed-cut working point is used  
1005 for b-jet identification.

1006

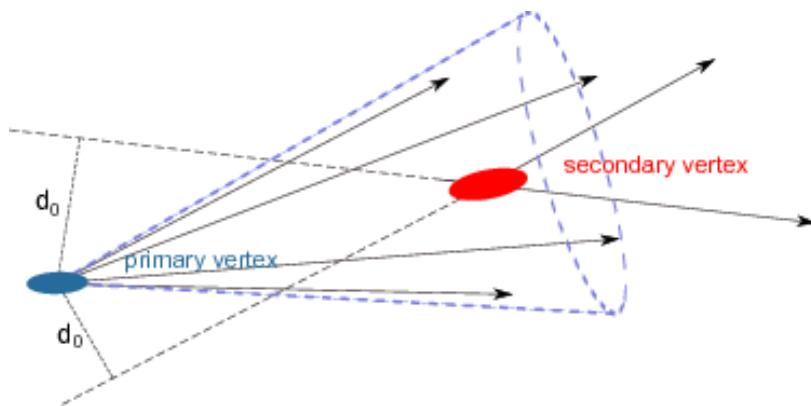


FIGURE 4.6. Schmatic view of the tracks in a b-jet [45].

1007 **4.2.5. Missing Transverse Momentum**

Neutrinos do not interact with the detector as they pass through. This means they cannot be measured like the other particles. In order to measure neutrinos, ATLAS relies on the conservation of momentum. As previously mentioned, the exact collision energy is unknown, as each partons does not carry a consistent fraction of the proton energy. However, in the transverse plane, the plane perpendicular to the beam line, the total momentum is known exactly. Before the collision, there is zero momentum in the transverse plane. After the collision, this must also be true. This implies the vector summation of all objects should have zero momentum in the transverse plane. Any imbalance in this is referred to as Missing Transverse Momentum ( $\cancel{E}_T$ ). The  $\cancel{E}_T$  is constructed as the negative vector sum of all reconstructed objects with an additional soft term reconstructed from detector signal objects not associated with any object[46].

$$E_{x(y)}^{\text{miss}} = E_{x(y)}^{\text{miss,e}} + E_{x(y)}^{\text{miss},\gamma} + E_{x(y)}^{\text{miss},\tau} + E_{x(y)}^{\text{miss,jets}} + E_{x(y)}^{\text{miss},\mu} + E_{x(y)}^{\text{miss,soft}} \quad (4.2)$$

From the  $x$  and  $y$  components of  $E^{\text{miss}}$ , the magnitude and azimuthal angle are calculated.

$$\begin{aligned} E_T^{\text{miss}} &= \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2}, \\ \phi^{\text{miss}} &= \arctan(E_y^{\text{miss}}/E_x^{\text{miss}}) \end{aligned} \quad (4.3)$$

1008 In this analysis, the  $\cancel{E}_T$  is reconstructed using VHLooseElectrons, VHLooseMuons the  
1009 analysis jets, and the track-based soft term.

1010 The  $\cancel{E}_T$  vector is a vector in the transverse plane, meaning it does not directly  
1011 correspond to a neutrino. Additional information is needed to exactly reconstruct a

<sub>1012</sub> neutrino. In this analysis, a Higgs mass constraint is used to supply the direction of  
<sub>1013</sub> the signal neutrino.

<sub>1014</sub>

## CHAPTER V

## ANALYSIS

1017 This chapter will present the results of a search for Higgs boson pair production  
 1018 where one Higgs boson decays via  $H \rightarrow b\bar{b}$  and the other via  $H \rightarrow WW^*$ . The  
 1019  $WW^*$  system decays into  $l\nu qq$  (where  $l$  is either an electron or a muon), with the  
 1020 small contamination from the leptonic  $\tau$  decays not explicitly vetoed in the analysis.  
 1021 The Higgs boson decay modes chosen for this analysis are a compromise between  
 1022 signal efficiency and background reduction. The  $H \rightarrow WW^*$  branching ratio of  
 1023 approximately 25% is the second largest after  $H \rightarrow b\bar{b}$  (approximately 58%). The  
 1024 1-lepton final state gives a strong discriminator against multijet background. The  
 1025 dominant backgrounds are  $t\bar{t}$  production, W bosons produced in association with  
 1026 jets (W+jets), and multijet events where a jet is misidentified as a lepton. There  
 1027 are smaller background contributions from single top-quark production, Z bosons  
 1028 produced in association with jets (Z+jets), and diboson production.

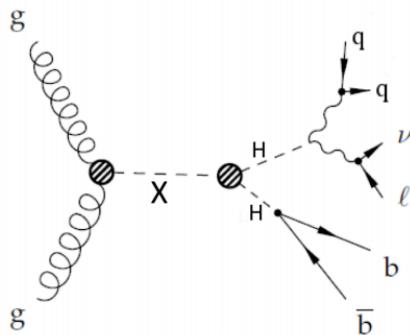


FIGURE 5.1. Schematic diagram of resonant Higgs boson pair production with the subsequent Higgs and W boson decays.

1030 This analysis sets limits on both SM Higgs boson pair production and on resonant  
1031 production. Both production methods are discussed in detail in chapter II. Figure  
1032 5.1 shows a Feynman diagram of resonant production of the Higgs boson pairs with  
1033 the subsequent decays  $H \rightarrow WW^*$  and  $H \rightarrow b\bar{b}$ .

1034 **5.1. Analysis Overview**

1035 Two complementary techniques are used to reconstruct the Higgs boson  
1036 candidates that decays into two b-quarks[1]. Both techniques use the anti- $k_t$  jet  
1037 algorithm but with different radius parameters. The first technique uses jets with  
1038 a radius parameter  $R = 0.4$  and it is used when each b-quark from the  $H \rightarrow b\bar{b}$   
1039 decay can be reconstructed as a distinct b-jet. This is referred to as the “resolved  
1040 analysis”[47]. The second technique uses jets with a radius parameter  $R = 1.0$ ,  
1041 also known as fatjets, and is used when the b-quarks cannot be reconstructed as  
1042 two distinct b-jets. Instead the Higgs boson candidate is identified as the single  
1043 fatjet. This technique is referred to as the “boosted analysis”[48]. In both analyses,  
1044 the jets from the hadronically decaying W boson are reconstructed as anti- $k_t$  jets  
1045 with radius parameter  $R = 0.4$ . The non-resonant, SM production, search uses the  
1046 resolved analysis exclusively, while the resonant analysis is performed using either the  
1047 resolved or boosted analysis technique with the most sensitive technique chosen for  
1048 each particular model and HH mass being tested.

1049 **5.2. Data and Monte Carlo Samples**

1050 **5.2.1. Data**

1051 The analysis presented uses the full proton-proton collision dataset collected in  
1052 2015 and 2016 as the center-of-mass energy of 13 TeV passing data quality checks

1053 requiring good conditions of all sub-detectors. The data that are currently used  
 1054 correspond to an integrated luminosity of  $36.1 \text{ fb}^{-1}$  ( $3.2 \text{ fb}^{-1}$  from 2015 plus  $32.8 \text{ fb}^{-1}$   
 1055 from 2016)<sup>1</sup>.  
 1056 .

### 1057 5.2.2. Monte Carlo Samples

1058 With the exception of the QCD multijet background described in 5.4.8, MC  
 1059 simulated events are used to estimate SM backgrounds and the signal acceptances.  
 1060 Table 5.1 summarizes the MC samples used for background estimation.

Process	Generator	$\sigma \times \text{BR} [\text{pb}]$
$t\bar{t} \rightarrow WWbb \rightarrow l\nu bb + X$	POWHEG+PYTHIA6	451.65
$Wt$ incl.	POWHEG+PYTHIA6	71.7
single $t$ , s-channel, $\rightarrow l\nu + X$	POWHEG+PYTHIA6	3.31
single $t$ , t-channel, $\rightarrow l\nu + X$	POWHEG+PYTHIA6	69.5
$W+\text{jets}, W \rightarrow l\nu$	SHERPA	61510
$Z+\text{jets}, Z \rightarrow ll$	SHERPA	6425
Dibosons incl.	SHERPA	47.3
$ggh$ incl.	POWHEG+PYTHIA8	48.5
$tth, \rightarrow l\nu + X$	AMC@NLO + HERWIG++	0.223

TABLE 5.1. SM MC samples used for background estimation.

1061 The  $t\bar{t}$  and single top-quark samples are generated with POWHEG-Box v2 [49]  
 1062 using CT10 parton distribution functions (PDF) interfaced to PYTHIA 6.428 [50]  
 1063 for parton shower, using the PERUGIA2012 [51] tune with CTEQ6L1 [52] PDF for  
 1064 the underlying event descriptions. EVTGEN v1.2.0 [53] is used for properties of

---

<sup>1</sup> The following GoodRunLists (GRL) are used:

data15\_13TeV.periodAllYear\_DetStatus-v79-repro20-02\_DQDefects-00-02-02\_PHYS\_StandardGRL\_All\_Good\_25ns.xml

and

data16\_13TeV.periodAllYear\_DetStatus-v88-pro20-21\_DQDefects-00-02-04\_PHYS\_StandardGRL\_All\_Good\_25ns.xml.

The GRLs were retrieved from the GoodRunListsForAnalysisRun2 twiki

the bottomed and charmed hadron decays. The mass of the top quark is set to  $m_t = 172.5$  GeV. At least one top quark in the  $t\bar{t}$  event is required to decay to a final state with a lepton. The cross section of  $t\bar{t}$  is known to NNLO in QCD including summation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms, and the reference value used in ATLAS is calculated using TOP++ 2.0 [54]. The parameter HDAMP, used to regulate the high- $p_T$  radiation in POWHEG, is set to  $m_t$  for good data/MC agreement in the high  $p_T$  region [55]. Each process of single top-quark ( $t$ -channel,  $s$ -channel and  $Wt$ -channel) is generated separately. The cross section of single-top is calculated with the prescriptions in Ref. [56, 57].

SHERPA v2.2.1 [58] with the NNPDF 3.0 [59] PDF set is used as the baseline generator for the  $(W \rightarrow \ell\nu)/(Z \rightarrow \ell\ell)$ +jets background. The diboson processes ( $WW$ ,  $WZ$  and  $ZZ$ ) are generated with SHERPA with the CT10 PDF set.

The  $ggH$  and  $VBF$  inclusive samples are generated with POWHEG using the CT10 PDF set interfaced to PYTHIA8 for parton shower, while  $tth$  is a semi-leptonic sample generated with MADGRAPH5\_AMCAtNLO interfaced to HERWIG++. The  $ggF$  cross section is normalised by using computations including up to three QCD loops (N3LO [60]).  $VBF$ ,  $Wh$  and  $Zh$  samples, with inclusive  $h$ ,  $W$  and  $Z$  decays are also generated using PYTHIA8.

Signal samples are generated with MADGRAPH5\_AMCAtNLO [61] interfaced to HERWIG++ according to the procedure defined in Ref. [62]. Events are generated with an effective Lagrangian in the infinite top-quark mass approximation, and reweighting the generated events with form factors that take into account the finite mass of the top quark. This procedure partially accounts for the finite top-quark mass effects [63]. After the full analysis chain was developed, there were also developments in the theoretical front, which took full NLO calculation and top mass into account [64, 65]. This led to a slight difference in  $m_{HH}$  shape. A re-weighting

1091 scheme was then developed to correct  $m_{HH}$  shape as described in these slides.<sup>2</sup> The  
1092 overall effect in the sensitivity is a loss of signal efficiency by about 30%, which is  
1093 also seen by other analysis such as  $HH \rightarrow bbbb$ .

1094 Table 5.2 shows the list of HH signals. They use a heavy Higgs scalar model  
1095 as the signal hypothesis. The masses of the heavy Higgs range from 260 GeV to  
1096 3000 GeV while the Higgs width is set to 10 MeV, therefore the model is valid in  
1097 the Narrow Width Approximation (NWA). The non resonant signal is normalised to  
1098  $\sigma(pp \rightarrow HH) \times Br(HH \rightarrow WWbb) = 0.590$  pb, the resonant ones are normalised to  
1099 0.044 pb for  $m_H < 2000$  GeV and to 0.041 for  $m_H \geq 2000$  GeV.

Process	Generator
HH SM	MADGRAPH5_AMCATHNLO + Herwig++ including Form Factor
$H \rightarrow HH$ ( $m_H = 260 - 3000$ ) GeV	MADGRAPH5_AMCATHNLO + Herwig++including Form Factor

TABLE 5.2. Di-Higgs signal samples used in the analysis.

1100 Additional pp collisions generated with PYTHIA 8.186 are overlaid to model the  
1101 effects of the pileup for all simulated events. All simulated events are processed  
1102 with the same reconstruction algorithm used for data. All background samples are  
1103 processed through the full ATLAS detector simulation [66] based on GEANT4 [67]  
1104 while signal samples use the Atlas Fast simulation.

1105 **5.3. Object Reconstruction**

1106 The final state of this analysis includes electrons, muons, neutrinos and  
1107 jets, including  $b$ -jets. The identification criteria and the selection applied to the  
1108 reconstructed objects are defined in the present section.

---

<sup>2</sup> <https://indico.cern.ch/event/652372/>

1109 **5.3.1. Electrons**

1110 **5.3.1.1. Electron reconstruction**

1111       Electromagnetic (EM) clusters are reconstructed with a sliding window  
1112 algorithm. EM clusters are associated with a track refitted with GSF (Gaussian  
1113 Sum Filter model) [68] to account for bremsstrahlung energy losses.

1114       Electron identification is performed using likelihood-based method. Variables  
1115 used by likelihood identification are the longitudinal and transverse shower profiles,  
1116 the track quality, the track and cluster positions to match in  $\eta$  and  $\phi$  and the presence  
1117 of high-threshold TRT hits.

1118       The isolation variables quantify the energy of the particles produced around the  
1119 electron candidate and allow to disentangle prompt electrons from other, non-isolated  
1120 electron candidates such as electrons originating from converted photons produced in  
1121 hadron decays, electrons from heavy flavor hadron decays, and light hadrons mis-  
1122 identified as electrons. The isolation variable we use for reconstructed electrons is  
1123 *Track*-based isolation,  $p_T^{\text{varcone}0.2}$ , defined as the sum of transverse momenta of all  
1124 tracks, satisfying quality requirements, within a cone of  $\Delta R = \min(0.2, 10 \text{ GeV}/E_T)$   
1125 around the candidate electron track.

1126       A more detailed discussion on the electron likelihood identification and isolation  
1127 variables and their performance with Run 2 data can be found in Ref. [69]. The  
1128 electron energy scale is calibrated such that it is uniform throughout the detector and  
1129 the residual differences between data and simulation are corrected. The calibration  
1130 strategy is based on the same strategy developed in Run 1 [70] and updates to the  
1131 calibration strategy for Run 2 is documented in Ref. [71].

1132       For this analysis, two set of electron selections are defined. They are denoted as  
1133 `VHLooseElectron` and `SignalElectron`. The selections are defined as the following

1134        **VHLooseElectron:** The electron  $p_T$  is required to be greater than 7 GeV. The  
 1135        electron cluster should be in the range of  $|\eta| < 2.47$ . Loose likelihood identification is  
 1136        applied in this criteria. Impact parameter significance ( $|d_0^{\text{sig}}| = d_0/\sigma_{d_0}$ ) less than 10  
 1137        standard deviations. and  $|\Delta z_0^{\text{IBL}} \sin \theta| < 0.5$  mm are also required, where IBL refers  
 1138        to the ATLAS Insertable *B*-Layer.

1139        **SignalElectron:** The electron is required to pass the VHLooseElectron selection  
 1140        with its  $p_T$  required to be greater than 27 GeV. The electron cluster should be in  
 1141        the range of  $|\eta| < 2.47$  but excluded from the crack region ( $1.37 < |\eta| < 1.52$ ).  
 1142        Tight likelihood identification is applied in SignalElectron criteria with the impact  
 1143        parameter significance required to be less than 2. In addition, the electron is  
 1144        required to be isolated by passing the `FixedCutTightTrackOnly` isolation working  
 1145        point which corresponds to a cut on the ratio of  $p_T^{\text{varcone}0.2}$  to electron  $p_T$  of 0.06 (i.e  
 1146         $p_T^{\text{varcone}0.2}/p_T < 0.06$ ).

1147        A summary of the electron selections is shown in Table 5.3.

Electron Selection	$p_T$	$ \eta $	ID	$ d_0^{\text{sig}} $	$ \Delta z_0^{\text{IBL}} \sin \theta $	Isolation
VHLooseElectron	>7 GeV	< 2.47	LH Loose	< 10	< 0.5 mm	-
SignalElectron	>27 GeV	< 2.47 and $\notin [1.37, 1.52]$	LH Tight	< 2	< 0.5 mm	<code>FixedCutTightTrackOnly</code>

TABLE 5.3. Electron selection requirements.

### 1148        5.3.2. Muons

#### 1149        5.3.2.1. Muon reconstruction

1150        Muon candidates are identified by using the algorithm described in Ref. [72].  
 1151        Muons are selected within  $|\eta| < 2.5$  using track quality criteria based on the number  
 1152        of hits in the inner detector and in the muon spectrometer. Medium quality criteria  
 1153        are used for muon identification. The criteria include muons identified in both the

1154 inner detector and muon spectrometer with good matching of the two tracks for  
1155  $|\eta| < 2.5$ .

1156 The muon isolation variables are similar to the electron isolation variables  
1157 above which is the *track-based* isolation,  $p_T^{textrm{varcone}0.3}$ , defined as the sum of  
1158 transverse momenta of all tracks, satisfying quality requirements, within a cone of  
1159  $\Delta R = \min(0.3, 10 \text{ GeV}/p_T)$  around the candidate muon.

1160 The performance of the muon identification and isolation variables are  
1161 documented in Ref. [41].

1162 Corrections to the muon momentum scale and resolution are applied to MC  
1163 simulation using the MuonCalibrationAndSmearingTool<sup>3</sup> to correct for data/MC  
1164 differences. The correction factors were derived from data/MC simulation  
1165 comparisons with  $Z \rightarrow \mu\mu$  and  $J/\Psi \rightarrow \mu\mu$  events (the calibration procedure to  
1166 derive the factors is documented in Ref. [41]). For this analysis, two sets of muon  
1167 selections are defined. They are denoted as VHLooseMuon and SignalMuon. The  
1168 selections are defined as the following:

1169 **VHLooseMuon:** The muon  $p_T$  is required to be greater than 7 GeV. The  
1170 muon cluster should be in the range of  $|\eta| < 2.7$ . Loose identification is applied in  
1171 this criteria. Impact parameter significance ( $|d_0^{sig}|$ ) less than 6 standard deviations.  
1172 and  $|\Delta z_0^{\text{IBL}} \sin \theta| < 0.5$  mm are also required.

1173 **SignalMuon:** The muon is required to pass the VHLooseMuon selection with  
1174 its  $p_T$  required to be greater than 27 GeV and should be in the range of  $|\eta| < 2.4$ .  
1175 Medium identification is applied in SignalMuon criteria with the impact parameter  
1176 significance required to be less than 2. In addition, the muon is required to be  
1177 isolated by passing the `FixedCutTightTrackOnly` isolation working point which

---

<sup>3</sup> as prescribed in MCPAnalysisGuidelinesMC15 twiki

1178 corresponds to a cut on the ratio of  $p_T^{textrm{varcone}0.3}$  to muon  $p_T$  of 0.06 (i.e  
1179  $p_T^{textrm{varcone}0.3}/p_T < 0.06$ ).

1180 A summary of the muon selections is shown in Table 5.4.

Muon Selection	$p_T$	$ \eta $	ID	$ d_0^{\text{sig}} $	$ \Delta z_0^{\text{IBL}} \sin \theta $	Isolation
VHLooseMuon	>7 GeV	< 2.7	Loose quality	< 6	< 0.5 mm	-
SignalMuon	>27 GeV	< 2.4	Medium quality	< 2	< 0.5 mm	FixedCutTightTrackOnly

TABLE 5.4. Muon selection requirements.

### 1181 5.3.3. Jets

#### 1182 5.3.3.1. Large-R jets

1183 For signal processes with a large resonant mass the b-jets produced by the Higgs  
1184 may be too close together to be resolved by the R=0.4 calorimeter-based jets (calo-  
1185 jets). This effect is expected to be noticeable when  $p_T^H > 500\text{GeV}$ <sup>4</sup>. Our approach to  
1186 reconstructing the  $H \rightarrow b\bar{b}$  system in this “boosted” regime is to use a large radius  
1187 (large-R) jet with radius parameter R = 1.0. The large-R jets are clustered using  
1188 the anti- $k_t$  jet algorithm [73] with topological calorimeter clusters as inputs. The  
1189 clusters are calibrated to the “local hadronic cell weighting”(LCW) scale [74]. In  
1190 order to minimize the effects from pileup on the large-R jet kinematics, the large-R  
1191 jet is then groomed using the trimming algorithm. The large-R jet energy and mass  
1192 is then calibrated to the particle-level scale. The calibration factors were derived  
1193 from MC simulation of multijet events [75]. The large-R jets are required to have  
1194  $p_T > 250\text{GeV}$  and  $|\eta| < 2.0$ .

---

<sup>4</sup> Using the rule of thumb  $\Delta R = 2m/p_T$ , where  $m = m_H$  and  $\Delta R = 0.4$

1195    **5.3.3.2. Track jets**

1196       To identify a large-R jet that is consistent with decay of  $H \rightarrow b\bar{b}$ , a method  
1197       developed by ATLAS is to reconstruct subjets within the large-R jet and identify the  
1198       subjets whether it is a b-jet or not by using a b-tagging algorithm. The baseline  
1199       method is to use subjets built from tracks (track jets). Track jets are built by  
1200       clustering Inner Detector tracks using the anti- $k_t$  algorithm with a radius parameter  
1201        $R = 0.2$ . The selected tracks are required to have  $p_T$  greater than 400 MeV and  
1202       pass a loose set of cuts, as listed in reference [76]. The smaller R parameter coupled  
1203       with the fact that tracks have better angular resolution than calorimeter clusters,  
1204       mean that the decay products of highly boosted heavy objects can still be resolved.  
1205       The selected track jets are then associated to the large-R calorimeter jets via ghost  
1206       association [77] method. A b-tagging algorithm is used to identify track jets which  
1207       are likely to contain b-hadrons which consist of the b-quarks from the Higgs boson  
1208       decay. The MV2c10 algorithm exploit the relatively long lifetime of B-hadrons with  
1209       respect to lighter hadrons, as well as the kinematics of the charged particle tracks.  
1210       For the boosted analysis, track jets are required to have  $p_T > 10\text{GeV}$  and  $|\eta| < 2.5$   
1211       for them to be within the inner detector acceptance. They are also required to have  
1212       at least 2 track constituents. The MV2c10 working point for track jets is the 77%  
1213       Fixed Cut efficiency.

1214    **5.3.3.3. Small-R jets**

1215       Small-R jets are reconstructed from three-dimensional topological calorimeter  
1216       clusters [74] using the anti- $k_t$  jet algorithm [73] with a radius parameter of 0.4.  
1217       Jet energies are corrected [78] for detector inhomogeneities, the non-compensating  
1218       nature of the calorimeter, and the impact of multiple overlapping  $pp$  interactions.

1219 Correction factors are derived using test beam, cosmic ray,  $pp$  collision data, and a  
1220 detailed Geant4 detector simulation. Jet cleaning is applied to remove events with  
1221 jets built from noisy calorimeter cells or non-collision backgrounds, requiring that jets  
1222 are not of “bad” quality.<sup>5</sup>

1223 To avoid selecting jets originating from pile-up interactions a “jet vertex tagger”  
1224 (JVT) criterion [79] is applied for jets with  $p_T < 60$  GeV and  $|\eta| < 2.5$  requiring a  
1225  $JVT > 0.59$  cut. This cut corresponds to the Default working point, as described  
1226 on the JVTCalibration twiki.

1227 **Signal jets** are defined as jets which passes the jet cleaning and JVT criteria,  
1228 described in the previous section. They are further required to have  $p_T > 20$  GeV and  
1229  $|\eta| < 2.5$ .

1230 The ATLAS jet flavor tagging algorithm, here the MV2c10 algorithm [44], is used  
1231 to select signal jets and suppress multi-jet,  $W+jets$ ,  $Z+jets$  and di-boson background.  
1232 Out of the possible working points corresponding to different  $b$ -tagging efficiencies,  
1233 the 85% Fixed-Cut<sup>6</sup> working point (WP) is selected as to keep the signal efficiency  
1234 high. Signal jets are labeled **b-jets** if they pass the MV2c10 85% WP cut and labeled  
1235 as **light-jets** if they fail the cut.

1236 The difference in the efficiency of  $b$ -tagging between data and simulation is taken  
1237 into account by applying scale factors provided by the Flavour Tagging CP group, as  
1238 prescribed on the BTagCalib2015 twiki. The uncertainties associated with  $b$ -tagging  
1239 are considered for  $b$ -,  $c$ - and light-flavor-induced jets, separately.

1240 Table 5.5 summarizes the jets selection.

---

<sup>5</sup> *LooseBad* jets, defined on the HowToCleanJets2016 twiki, are removed.

<sup>6</sup> The charm quark component is suppressed by a factor 3.10 while the light quark component is suppressed by a factor 33.5. The expected performance is documented on the BTaggingBenchmarks twiki.

	Signal Jets
Algorithm	anti- $k_t$
$p_T$	20 GeV
$ \eta $	< 2.5
Quality	not “bad” jet
Pile-up jet removal	JVT > 0.59 when $ \eta  < 2.5$ and $p_T < 60$ GeV
$b$ -tagging	MV2c10, 85% fixed-cut WP, labelled as b-jets pass cut, light-jets if fail cut

TABLE 5.5. Selection for jets with distance parameter  $R = 0.4$ .

1241 **5.3.4. Missing transverse momentum ( $\cancel{E}_T$ )**

1242 The neutrino is not directly detectable and, thus, appears only as an imbalance  
1243 in transverse momentum.

1244 The missing transverse momentum (MET, or  $\cancel{E}_T$ ) [46] used in this analysis is  
1245 computed by using electrons that pass the VHLooseElectron selection, muons passing  
1246 the VHLooseMuon selection and jets of the analysis.<sup>7</sup> The track-based soft term<sup>8</sup>  
1247 (TST) is the recommended soft term component for the MET calculation. Photons  
1248 and hadronically decaying taus are included in the  $\cancel{E}_T$  calculation as jets since they  
1249 are not used explicitly in the event reconstruction.

1250 **5.3.5. Overlap removal**

1251 In order to uniquely identify objects, overlapping objects are removed according  
1252 to the overlap removal procedure defined in this section. Electrons and muons that  
1253 pass the VHLooseElectron and VHLooseMuon selections (as defined in Sec. 5.3.1.1  
1254 and 5.3.2.1) are considered for overlap removal. Calorimeter jets which pass the  
1255 JVT requirement are also considered for overlap removal. The procedure is defined  
1256 as follows.

---

<sup>7</sup> From MET\_Core\_AntiKt4EMTopo with the MissingETAssociationMap using the METMaker tool. All calibrated jets are passed to the METMaker tool as prescribed on the EtMiss subgroup twiki

<sup>8</sup> Defined on this twiki.

1257        If an electron and a muon shares a track, the muon is removed if it is *calo-*  
 1258 *tagged*. Otherwise, the electron is removed. Calorimeter jets are then removed if they  
 1259 are within  $\Delta R(\text{calo-jet}, \text{electron}) \leq 0.2$  of surviving electrons. Electrons that satisfy  
 1260  $\Delta R(\text{electron}, \text{calo-jet}) \geq \min(0.4, 0.04 + 10 \text{ GeV}/E_T^{\text{electron}})$  are removed. The surviving  
 1261 calorimeter jets are removed if they are within  $\Delta R(\text{calo-jet}, \text{muon}) \leq 0.2$  and do **not**  
 1262 pass any of the following criteria:

- 1263        – The number of tracks in the jet are more than 2.  
 1264        –  $p_T^{\text{muon}}/p_T^{\text{calo-jet}} < 0.5$  AND  $p_T^{\text{muon}}/p_T^{\text{tracks in calo-jet}} < 0.7$ .

1265        Muons that satisfy  $\Delta R(\text{muon}, \text{calo-jet}) \geq \min(0.4, 0.04 + 10 \text{ GeV}/p_T^{\text{muon}})$   
 1266 are removed. The overlap removal procedure is implemented using ASG's  
 1267 `AssociationUtils` package and summarized in Table 5.6.

Overlapping Objects	Removal Procedure
Electron - Muon	If share track, remove muon if calo-tagged. Otherwise remove electron.
Electron - Calo-jet	If $\Delta R(\text{calo-jet}, \text{electron}) \leq 0.2$ , remove calo-jet. If $\Delta R(\text{electron}, \text{calo-jet}) \geq \min(0.4, 0.04 + 10 \text{ GeV}/E_T^{\text{electron}})$ , remove electron.
Muon - Calo-jet	If $\Delta R(\text{calo-jet}, \text{muon}) \leq 0.2$ , remove calo-jet if: a) Number of tracks in calo-jet $\leq 2$ , OR b) $p_T^{\text{muon}}/p_T^{\text{calo-jet}} > 0.5$ AND $p_T^{\text{muon}}/p_T^{\text{tracks in calo-jet}} > 0.7$ . If $\Delta R(\text{muon}, \text{calo-jet}) \geq \min(0.4, 0.04 + 10 \text{ GeV}/p_T^{\text{muon}})$ , remove muon.

TABLE 5.6. A summary of the overlap removal procedure.

## 1268 5.4. Resolved Analysis

### 1269 5.4.1. Event Selection

1270        The final state of interest consists of one charged lepton, one neutrino, and four  
 1271 quarks, two of which are b-quarks. Hence the detector signature consists of one  
 1272 charged lepton ( $e/\mu$ ), large  $\cancel{E}_T$ , and four or more anti- $k_t$  jets of which two are *b*  
 1273 jets from the  $h$  decay while the other two are light jets from the hadronic decay  
 1274 of the  $W$  boson. One challenge in the event reconstruction is to correctly identify

the pair of light jets from the  $W$  boson decay. This information is also used to solve the  $z$  component of the neutrino momentum. For the HH signal there is an additional complication due to the fact that one of the  $W$  bosons is off-shell, and thus for this  $W$  there is no  $W$  mass constraint. This section details the stages of the event reconstruction and the progression towards the final selection which defines the signal region. In addition, signal depleted control regions are defined in the next section which are used to check the consistency of the SM background predictions with the data in the control regions. The search has been kept “blinded” until the comparison between data and simulation of backgrounds are well understood in the signal depleted control regions.

#### 5.4.2. Trigger requirement

Events are selected using the unprescaled single lepton triggers. The list of triggers used in this analysis is shown in Table 5.7. Events are selected with a logical OR between the triggers listed in Table 5.7.

##### 5.4.2.1. Pre-selection

The following selection cuts are applied at the pre-selection level to the recorded events:

- In order to assure good data quality, events with bad detector conditions, namely where large part of the detectors were missing from data acquisition due to problems during a run, or when the performance of the detectors were affected by large noise, have been rejected from the data analysis. A GRL selection taken from `data15_13TeV.periodAllYear_DetStatus-v79-repro20-02_DQDefects-00-02-02_PHYS_StandardGRL_All_Good_25ns.xml` and `data16_13TeV.periodAllYear_DetStatus-v88-pro20-21_DQDefects-00-02-04_PHYS`

Dataset	Trigger items
2015	mu20_iloose_L1MU15 mu50 e24_lhmedium_L1EM18VH (MC) e24_lhmedium_L1EM20VH (data) e60_lhmedium e120_lhloose
2016 - Period A	mu24_iloose_L1MU15 (MC) mu24_iloose (data) mu40 e26_lhtight_nod0_ivarloose e60_lhmedium_nod0 e60_medium e140_lhloose_nod0 e300_etcut
2016 - Period B-D3	mu24_ivarmedium mu50 e26_lhtight_nod0_ivarloose e60_lhmedium_nod0 e60_lhmedium e140_lhloose_nod0 e300_etcut
2016 - Period D4-E3	mu26_ivarmedium mu50 e26_lhtight_nod0_ivarloose e60_lhmedium_nod0 e60_lhmedium e140_lhloose_nod0 e300_etcut
2016 - Period $\geq F$	mu26_ivarmedium mu50 e26_lhtight_nod0_ivarloose e60_lhmedium_nod0 e60_lhmedium e140_lhloose_nod0 e300_etcut

TABLE 5.7. Summary of trigger items used for 2015 and 2016 data. For 2016 data, different triggers were used for different data run periods. All triggers are unprescaled.

1299        .StandardGRL\_All\_Good\_25ns.xml is applied. Moreover, incomplete events or  
1300        events with bad detector information are rejected.

- 1301        – The presence of a primary vertex with at least two tracks. Among all primary  
1302        vertices, that with the highest  $\sum p_{T,\text{trk}}^2$ , where  $p_{T,\text{trk}}$  is the transverse momentum  
1303        of tracks associated with the vertex, is retained as the primary interaction  
1304        vertex;
- 1305        – at least one SignalElectron ( $e$ ) or SignalMuon ( $\mu$ ), as defined in Sec. 5.3.1.1 and  
1306        Sec. 5.3.2.1, and it must be trigger matched to the corresponding HLT object  
1307        which fires the trigger;
- 1308        – at least 4 jets, of which 2 and only 2 are  $b$ -tagged.

1309        **5.4.3. Event Reconstruction**

1310        Events are reconstructed by first requiring exactly two  $b$ -tag jets and at least  
1311        2 light jets and at most 3 light jets. In events with 3 light jets, the pair with the  
1312        lowest  $\Delta R$  between them are selected as  $W$  jet candidates. This procedure yields  
1313        the correct jet assignment in 70% of the cases for signal events where the hadronic  
1314        daughters of the  $W$  boson can be correctly matched to reconstructed jets as described  
1315        in Appendix ??.

The event kinematics of the  $H \rightarrow WW^* \rightarrow l\nu qq$  topology can be fully reconstructed. In fact, among all four-momenta of the final state particle, only the component of the neutrino momentum along the beam axis,  $p_z$  in the following, is unknown while its transverse momentum is the measured  $\cancel{E}_T$ . Imposing the relation:

$$m_h^2 = (p^l + p^\nu + p^{j1} + p^{j2})^2 \quad (5.1)$$

where  $p^i$  is the four-momenta of particle  $i$ , the neutrino  $p_z$  can be reconstructed using the relations:

$$p_E^\nu = E^\nu = \sqrt{P_T^2 + p_z^2} \quad p_x^\nu = P_T \cos(\phi) \quad p_y^\nu = P_T \sin(\phi)$$

where  $\phi$  is the azimuthal angle of the  $E_T$ ,  $E^\nu$  the neutrino energy,  $p_x$  and  $p_y$  the two transverse spatial components of the neutrino momentum. Eq. 6.1 is a quadratic expression in  $p_z$ . It can have two real, one real or two complex solutions. In the last case only the real part of the complex solution is taken into account, therefore a single value of  $p_z$  is obtained. In the first case the solution with the neutrino direction closest to the charged lepton is retained. It has been shown that this algorithm selects the correct solution in approximately 60% of the cases (see Appendix D).

#### 5.4.4. $bb\tau\tau$ analysis overlap removal

In order to remove overlap with the  $bb\tau\tau$  analysis we reject any event containing at least one hadronic  $\tau$  candidate that could be identified by the  $bb\tau\tau$  analysis, that fullfill the following requirements:

- $p_T > 20$  GeV and  $|\eta| < 2.5$ ;
- one or three prongs;
- unit charge;
- pass the medium  $\tau$  ID BDT working point.

The rejection of such events causes a signal efficiency drop of about 3%.

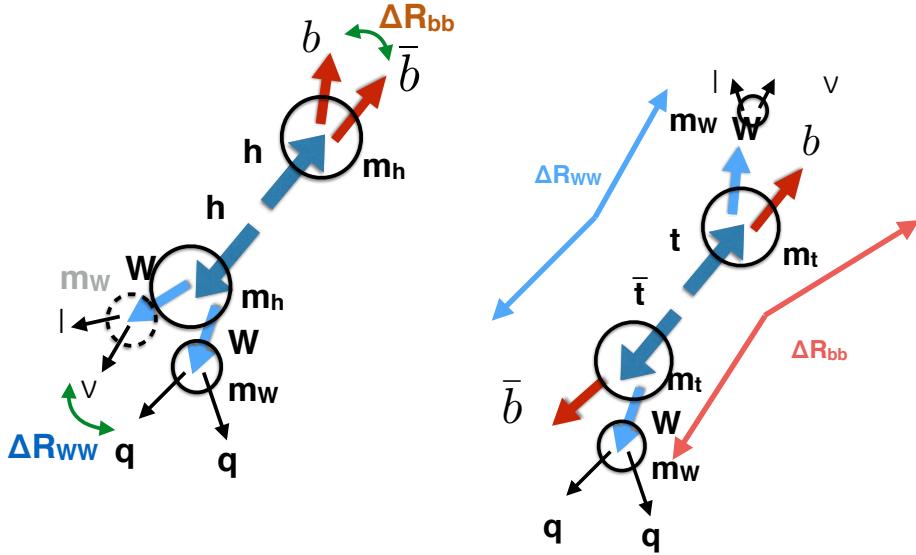


FIGURE 5.2. Schematic view of a  $HH \rightarrow WWb\bar{b}$  event compared to a  $t\bar{t} \rightarrow WWb\bar{b}$  event.

#### 1332 5.4.5. Kinematic selection

1333 Kinematic selection is used to suppress mainly  $t\bar{t}$  background while keeping high  
1334 signal efficiency. A schematic view of the  $HH \rightarrow WW*b\bar{b}$  and the  $t\bar{t} \rightarrow WWb\bar{b}$  event  
1335 topology is shown in Figure 5.2.

1336 The  $t\bar{t}$  events are typically characterised by two  $b$ -jets and two  $W$  bosons such  
1337 that the  $\Delta R$  separation between the two  $b$ -jets and between the  $W$  bosons is large.  
1338 On the contrary, in particular when the invariant mass of the  $m_{HH}$  is high, the signal  
1339 is characterised by two  $b$ -jets which are close together in  $\Delta R$  and by two  $W$  bosons  
1340 which are also relatively closer than in the  $t\bar{t}$  case. Moreover, while for the signal the  
1341 two  $b$ -jets have an invariant mass equal to  $m_h$ , this is not necessarily the case for the  
1342  $t\bar{t}$  background. Following these considerations, the typical separation variables are:

1343 – the  $p_T$  of the  $b\bar{b}$  pair ( $p_T^{bb}$ );

- 1344     – the  $\Delta R$  of the  $b\bar{b}$  pair ( $\Delta R^{bb}$ );
- 1345     – the  $p_T$  of the  $WW$  pair ( $p_T^{WW}$ );
- 1346     – the  $\Delta R$  of the  $WW$  pair ( $\Delta R^{WW}$ );
- 1347     – the mass of the  $WW$  system computed using the calculated neutrino  
1348        longitudinal momentum ( $m_{WW}$ ). This value is exactly equal to  $m_h$  if a real  
1349        solution is found, it is larger if no real solution is found;
- 1350     – the invariant mass of the di-Higgs boson candidate system ( $m_{HH}$ ).
- 1351     – the invariant mass of the 2 b-jets boson system ( $m_{bb}$ ).

1352     **5.4.6. Signal region definitions**

1353     The cuts of the selection have been optimised by maximising the Poisson  
1354        significance at the end of the selection <sup>9</sup>. The Poisson significance formula depends  
1355        on the absolute yield of expected signal and background events. For the optimisation  
1356        formula the  $t\bar{t}$  background was normalised to data with  $m_{bb} < 100$  GeV or  $m_{bb} > 140$   
1357        GeV, such region rejects, in fact, the majority of the signal. Four signal hypotheses  
1358        have been used in the optimisation:

- 1359     – a heavy Higgs with  $m_H = 500$  GeV and  $m_H = 700$  GeV (defined as low-mass  
1360        analysis),
- 1361     – a heavy Higgs with  $m_H = 2000$  GeV (defined as high-mass analysis) and
- 1362     – a non-resonant di-Higgs production (defined as non-resonant analysis).

---

<sup>9</sup> a two step procedure has been implemented. In the first step each cut is optimised, in the second step, all cuts are set to their optimal value and cuts are varied one by one to look for a different optimisation point. Correlation among variables could in fact spoil the results obtained at the first step.

variable	<i>Non-Res</i>	<i>m500</i>	<i>low-mass</i>	<i>high-mass</i>
$E_{miss}^T$ (GeV)	> 25	> 25	> 25	> 25
$m_{WW}$ (GeV)	< 130	< 130	< 130	no-cut
$p_T^{bb}$ (GeV)	> 300	> 210	> 210	> 350
$p_T^{WW}$ (GeV)	> 250	> 150	> 250	> 250
$\Delta R_{WW}$	no-cut	no-cut	no-cut	< 1.5
$m_{bb}$ (GeV)	105-135	105-135	105-135	105-135

TABLE 5.8. Criteria for non-resonant, *m500*, *low-mass* and *high-mass* selection. The  $m_{hh}$  window cut is not applied for non-resonant signal, and for resonant signals  $m_{hh}$  depends on the mass.

1363 An additional mass point with  $m_H = 1400$  GeV was also checked. The resulting  
 1364 selection and the corresponding sensitivity are very similar to the selection for  $m_H =$   
 1365 2000 GeV, and hence that selection is dropped.<sup>10</sup>

1366 The signal regions for the reference signal hypotheses are summarized in  
 1367 Table 5.8.

1368 The *non-res* and *m500* selections are exclusively used for non-resonant signal and  
 1369 resonant signal with mass 500 GeV respectively. The *low-mass* selection is used for  
 1370 signal masses from 600 to 1300 GeV, while the *high-mass* selection is used for signals  
 1371 with masses between 1400 and 3000 GeV. In addition, requirements are placed on  
 1372 the reconstructed di-Higgs invariant mass  $m_{HH}$  as a function of the signal resonance  
 1373 mass  $m_X$ , as shown in Table 5.9. The resolution of the reconstructed mH H ranges  
 1374 from 6% at 500 GeV to 10% at 3000 GeV.

#### 1375 5.4.7. Background Determination

1376 In the present analysis we expect that at the end of the event selection the  
 1377 sample will be largely dominated by  $t\bar{t}$  and multi-jet background, therefore the  $t\bar{t}$

---

<sup>10</sup> See [https://indico.cern.ch/event/641988/contributions/2604588/attachments/1465307/2265002/bbWW\\_Weekly\\_Optimization\\_Revisited\\_24May2018.pdf](https://indico.cern.ch/event/641988/contributions/2604588/attachments/1465307/2265002/bbWW_Weekly_Optimization_Revisited_24May2018.pdf)

$m_H$ (GeV)	500	600	700	750	800
$m_{hh}$ cut (GeV)	480-530	560-640	625-775	660-840	695 - 905
$m_H$ (GeV)	900	1000	1100	1200	1300
$m_{hh}$ cut (GeV)	760-970	840-1160	925-1275	1010-1390	1095-1505
$m_H$ (GeV)	1400	1500	1600	1800	2000
$m_{hh}$ cut (GeV)	1250-1550	1340-1660	1430-1770	1750-2020	1910-2170
$m_H$ (GeV)	2250	2500	2750	3000	
$m_{hh}$ cut (GeV)	2040-2460	2330-2740	2570-2950	2760-3210	

TABLE 5.9. Window cuts on  $m_H$  as a function of the resonance mass  $m_S$ .

background normalisation is derived from data while, as described in Sec. 5.4.8, the multi-jet background is derived using a data-driven ABCD method. For all the other backgrounds, e.g. di-boson, Higgs,  $W+jets$ , the MC is used appropriately normalized by using the expected cross sections and the integrated luminosity that has been collected.

#### 5.4.7.1. Top normalization and control region

The  $t\bar{t}$  background is normalised and validated using dedicated control regions (CR). Three CR's are defined, one for the SR's of the non-resonant (CR1), one for the low-mass analysis (CR2), and one for the high-mass analysis (CR3). The CRs are defined in Table 5.11.

Table 5.12 through 5.15 show the number of observed events and expected background events in the top CRs, and also in the sideband across selections that serve as validation regions. The final signal region is defined by  $m_{bb}$  cut of 105 GeV  $< m_{bb} <$  135 GeV based on optimization. The sidebands are orthogonal to the SR by virtue of having the  $m_{bb}$  cut reversed.  $m_{bb} < 100$  GeV or  $m_{bb} > 140$  GeV defines the sidebands in which the control regions are defined. The 5 GeV buffer region is kept on both sides so as to be less affected by systematic effects at the edge.

Selection	non-res	m500	m700	m2000
Buffer/SR	1.85	1.95	1.90	1.63
Sidebands/SR	20.5	12.6	13.4	5.6

TABLE 5.10. The ratios of S/B in the buffer zone and sidebands compared to the S/B in the final SR.

Fig. 5.3 shows  $m_{bb}$  for various signal mass points. A comparative study of signal over background in these three regions shows that S/B in the final SR is 5 (20) times higher than in the sidebands for m2000 (non-resonance) while S/B in the final SR is approximate twice as high as in the buffer zones. Table 5.10 shows the ratios of S/B in the buffer zones and sidebands compared to the S/B in the final SR.

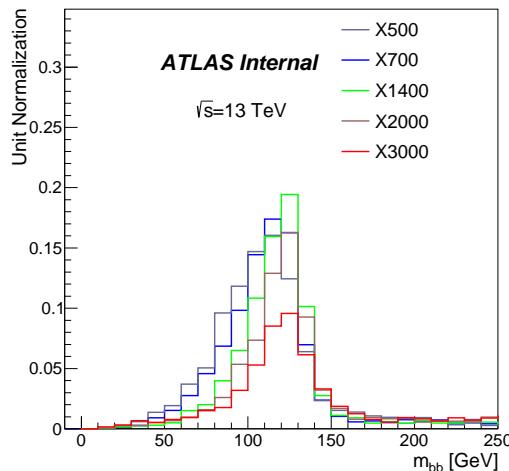


FIGURE 5.3.  $m_{bb}$  resolution for signal samples.

variable	CR1	CR2	CR3
$m_{bb}$ (GeV)	$m_{bb} < 100$ or $m_{bb} > 140$	$m_{bb} < 100$ or $m_{bb} > 140$	$m_{bb} < 100$ or $m_{bb} > 140$
$m_{WW}$ (GeV)	$< 130$	$< 130$	no-cut
$p_T^{bb}$ (GeV)	$> 300$	$> 210$	$> 350$

TABLE 5.11. Definition of the kinematic regions used to normalise the Top background.  $m_{bb} < 100$  GeV or  $m_{bb} > 140$  GeV defines the sidebands in which the control regions are defined. Expected SM backgrounds are then checked against data at each subsequent selection.

CR1: $m_{b\bar{b}}$ Sideband				
Sample	mww	bbpt210	bbpt300	wwpt250
$t\bar{t}$	$23776.6 \pm 87.2$	$531.7 \pm 13.1$	$109.9 \pm 5.9$	$63.9 \pm 4.6$
QCD	$13310.5 \pm 500.3$	$250.2 \pm 30.6$	$33.7 \pm 4.1$	$21.4 \pm 2.6$
W+jets	$3938.9 \pm 31.1$	$124.7 \pm 3.5$	$29.3 \pm 1.4$	$17.1 \pm 1.1$
SingleTop	$1605.4 \pm 18.0$	$76.0 \pm 3.8$	$20.1 \pm 2.0$	$13.5 \pm 1.7$
Dibosons	$109.9 \pm 2.7$	$8.3 \pm 0.8$	$2.2 \pm 0.4$	$1.5 \pm 0.4$
Z+jets	$1107.6 \pm 8.4$	$27.1 \pm 0.8$	$6.7 \pm 0.4$	$2.4 \pm 0.2$
Background Sum	$43849.0 \pm 509.2$	$1017.9 \pm 33.7$	$201.9 \pm 7.6$	$119.8 \pm 5.7$
XhhSM	$44.6 \pm 2.2$	$9.1 \pm 0.7$	$1.5 \pm 0.2$	$1.1 \pm 0.1$
Data	43902.0	1069.0	206.0	138.0

TABLE 5.12. The number of observed events and expected background events in the  $m_{bb}$  side-bands for the non-resonant selection. The top CR1 is defined at the bbpt300 selection. No NF has been applied to the background yields to show the level of data/expectation agreement before normalizing  $t\bar{t}$ . Only statistical uncertainties are shown.

1400

1401

1402

1403      Table 5.12 through 5.15 show the number of observed events and expected  
 1404 background events in the top CRs before the normalisation factors have been applied  
 1405 to the top background sample.

CR2: $m_{b\bar{b}}$ Sideband				
Sample	mww	bbpt210	wwpt150	hh500
$t\bar{t}$	$23776.6 \pm 87.2$	$531.7 \pm 13.1$	$432.7 \pm 11.8$	$35.5 \pm 3.2$
QCD	$13310.5 \pm 500.3$	$250.2 \pm 30.6$	$206.3 \pm 25.3$	$16.9 \pm 2.1$
W+jets	$3938.9 \pm 31.1$	$124.7 \pm 3.5$	$105.9 \pm 3.3$	$4.9 \pm 0.6$
SingleTop	$1605.4 \pm 18.0$	$76.0 \pm 3.8$	$64.9 \pm 3.5$	$2.8 \pm 0.6$
Dibosons	$109.9 \pm 2.7$	$8.3 \pm 0.8$	$6.7 \pm 0.8$	$0.9 \pm 0.2$
Z+jets	$1107.6 \pm 8.4$	$27.1 \pm 0.8$	$19.0 \pm 0.7$	$1.5 \pm 0.2$
Background Sum	$43849.0 \pm 509.2$	$1017.9 \pm 33.7$	$835.5 \pm 28.3$	$62.5 \pm 3.9$
Xhh500	$3.2 \pm 0.1$	$0.6 \pm 0.1$	$0.6 \pm 0.1$	$0.2 \pm 0.1$
Data	43902.0	1069.0	898.0	73.0

TABLE 5.13. The number of observed events and expected background events in the  $m_{b\bar{b}}$  side-bands for the low-mass selection, m500. The top CR2 is defined at the bbpt210 selection. To show how well the prediction matches data, no NF has been applied to any background. Only statistical uncertainties are shown.

CR2: $m_{b\bar{b}}$ Sideband				
Sample	mww	bbpt210	wwpt250	hh700
$t\bar{t}$	$23776.6 \pm 87.2$	$531.7 \pm 13.1$	$175.6 \pm 7.5$	$49.9 \pm 3.9$
QCD	$13310.5 \pm 500.3$	$250.2 \pm 30.6$	$72.4 \pm 8.9$	$28.4 \pm 3.5$
W+jets	$3938.9 \pm 31.1$	$124.7 \pm 3.5$	$45.7 \pm 2.1$	$13.7 \pm 1.4$
SingleTop	$1605.4 \pm 18.0$	$76.0 \pm 3.8$	$28.4 \pm 2.4$	$6.9 \pm 1.1$
Diboson	$109.9 \pm 2.7$	$8.3 \pm 0.8$	$2.8 \pm 0.5$	$0.7 \pm 0.2$
Z+jets	$1107.6 \pm 8.4$	$27.1 \pm 0.8$	$5.8 \pm 0.4$	$2.0 \pm 0.3$
Background Sum	$43849.0 \pm 509.2$	$1017.9 \pm 33.7$	$330.7 \pm 12.1$	$101.5 \pm 5.5$
Xhh700	$4.2 \pm 0.2$	$2.2 \pm 0.1$	$1.5 \pm 0.1$	$1.0 \pm 0.1$
Data	43902.0	1069.0	367.0	124.0

TABLE 5.14. The number of observed events and expected background events in the  $m_{b\bar{b}}$  side-bands for the low-mass selection, m700. The top CR2 is defined at the bbpt210 selection. To show how well the prediction matches data, no NF has been applied to any background. Only statistical uncertainties are shown.

CR3: $m_{b\bar{b}}$ Sideband				
Sample	bbpt350	wwpt250	drww15	hh2000
$t\bar{t}$	$8568.7 \pm 52.1$	$7095.6 \pm 47.5$	$1940.5 \pm 25.1$	$122.3 \pm 6.5$
QCD	$1538.7 \pm 252.7$	$1359.5 \pm 75.9$	$392.7 \pm 21.9$	$20.7 \pm 1.2$
W+jets	$2259.5 \pm 7.9$	$1952.1 \pm 7.4$	$696.6 \pm 4.6$	$55.5 \pm 1.1$
SingleTop	$1778.1 \pm 19.4$	$1601.6 \pm 18.4$	$405.4 \pm 9.2$	$29.6 \pm 2.6$
Dibosons	$170.6 \pm 3.9$	$147.1 \pm 3.7$	$46.8 \pm 2.1$	$3.4 \pm 0.6$
Z+jets	$403.6 \pm 2.1$	$307.6 \pm 1.8$	$95.6 \pm 1.1$	$7.5 \pm 0.3$
Background Sum	$14719.1 \pm 258.9$	$12463.5 \pm 91.8$	$3577.5 \pm 35.0$	$238.9 \pm 7.2$
Xhh2000	$25.7 \pm 0.4$	$24.0 \pm 0.4$	$9.6 \pm 0.3$	$2.9 \pm 0.1$
Data	14862.0	12450.0	3761.0	250.0

TABLE 5.15. The number of observed events and expected background events in the  $m_{bb}$  side-bands for the high-mass selection. The top CR3 is defined at the bbpt350 selection. No NF has been applied to the background yields. Only statistical uncertainties are shown.

Top background normalisation factors in the two CRs.

region	NF	$\sigma_{stat.}$	$\sigma_{syst.}$
non-res	1.04	$\pm 0.20$	$\pm 0.43$
low-mass	1.14	$\pm 0.10$	$\pm 0.35$
high-mass	1.02	$\pm 0.02$	$\pm 0.07$

TABLE 5.16. Normalization factors for the two CRs, the statistical error includes only data statistics, the systematic error is obtained subtracting in quadrature the statistical error from the total error.

1406        The top normalisation factors are determined by a simultaneous fit of signal  
1407        and control regions, which include both Top CR and QCD CR 5.4.8. It also depends  
1408        slightly on the  $m_{hh}$  window due to the presence of top background in the signal region,  
1409        and it is furthermore different for the *non res*, *low mass* and *high mass* analyses. The  
1410        normalisation factors of the three top control regions are shown in Table 5.16.

1411        Fig. 5.4 shows the  $m_{bb}$  and  $m_{hh}$  distributions in the two CRs.relevant variables  
1412        used in the analysis selection in the top CR.

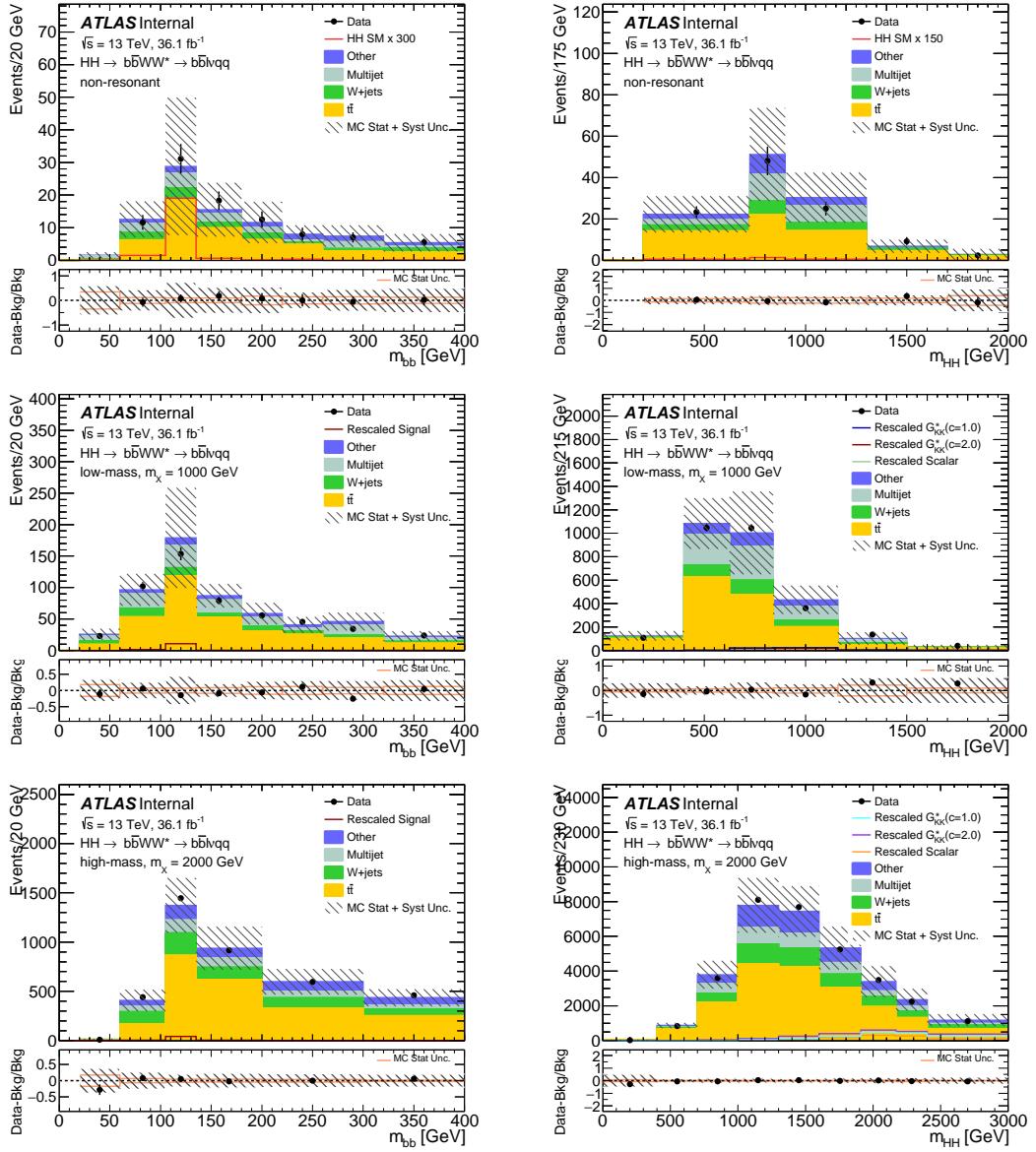


FIGURE 5.4.  $m_{bb}$  and  $m_{hh}$  in CR1, CR2 and CR3.  $t\bar{t}$  NFs as described in 5.4.7.1 have been applied. The uncertainties shown include the statistical and systematic uncertainties described in 5.4.9. Data are blinded in the region  $100 < m_{bb} < 140$ .

#### 1413 5.4.8. Multi-jet background

1414 Multi-jet backgrounds can enter in the event selection if a jet from heavy flavour  
 1415 decays is mis-identified as an electron or a muon and used as a lepton in the

1416 analysis. Such phenomena are not accurately reproduced by MC simulation, due  
 1417 to large uncertainties in the jet shower shape simulation and uncertainties in muon  
 1418 fragmentation functions and kinematics. In order to estimate the contributions of  
 1419 multi-jet processes, a data-driven ABCD method is used to estimate this background  
 1420 in the present analysis.

1421 The ABCD method uses three control regions (the B, C, and D regions) to  
 1422 estimate the contribution of a given background in the signal (A) region. Cuts on two  
 1423 ideally orthogonal variables are used to create the signal and various control regions,  
 1424 e.g. the A region passes both cuts, the B and C regions each pass one cut and fail  
 1425 the other, while the D region fails both cuts. The absolute value of the significance  
 1426 of the lepton impact parameter and the missing transverse energy (MET) are used  
 1427 as the two variables used to define the regions in the ABCD method for this analysis.

1428 The regions are thus defined

- 1429 – A region: MET > 25 GeV,  $|\sigma_{d_0}| < 2.0$
- 1430 – B region: MET < 25 GeV,  $|\sigma_{d_0}| < 2.0$
- 1431 – C region: MET > 25 GeV,  $|\sigma_{d_0}| > 2.0$
- 1432 – D region: MET < 25 GeV,  $|\sigma_{d_0}| > 2.0$

Figure 5.5 shows a pictoral representation of the four regions. Assuming that the two variables chosen to define the ABCD regions are completely uncorrelated, the yield of the process being modelled (QCD multi-jets in this case) in the A region is given by

$$N_A = N_C \frac{N_B}{N_D} \quad (5.2)$$

where the yields  $N_i$  are yields calculated from data - all Monte Carlo backgrounds ( $t\bar{t}$ , W/Z+jets, single top, diboson processes) in region  $i$  ( $N_i = N_i^{\text{data}} - N_i^{\text{MC Bkgs}}$ ).

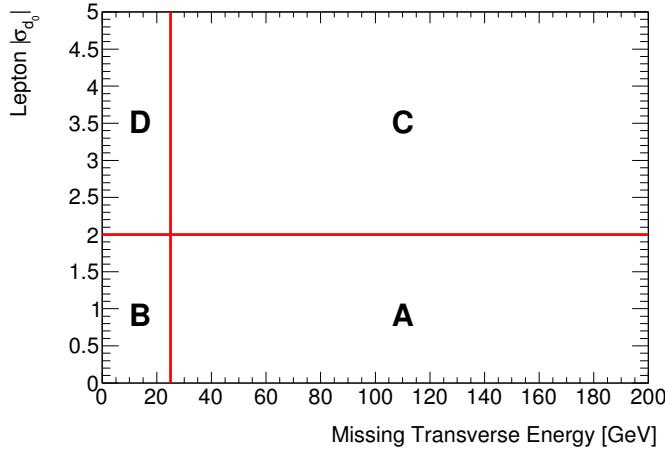


FIGURE 5.5. A pictoral representation of the four regions used in the ABCD calculation.

The assumption underlying Equation 5.2 is that the relationship between the yields in the B and D regions is the same as the relationship between the A and C regions, i.e.

$$\frac{N_A}{N_C} = \frac{N_B}{N_D} \quad (5.3)$$

Using equation 5.3, the quantity  $R = \frac{N_C N_B}{N_A N_D}$  can be defined. In the case of two completely uncorrelated variables,  $R = 1$  and the ABCD estimation reduces to Equation 5.2. If the two variables are not completely uncorrelated, the  $R$  factor enters as a correction to Equation 5.2 for the multi-jet estimation in the A region, and the expression can be rewritten as

$$N_A = R \frac{N_C N_B}{N_D} \quad (5.4)$$

<sup>1433</sup> The  $R$  factor is calculated for each selection (non-resonant, low mass, and high mass) individually, and the results at each cut in each selection is provided in Table 5.17.

QCD $R$ Values, Non-resonant Selection			
mww	bbpt210	bbpt300	wwpt250
$0.74 \pm 0.04$	$0.79 \pm 0.23$	$1.07 \pm 1.18$	—
QCD $R$ Values, Low Mass Selection (m500)			
mww	bbpt210	wwpt150	hh500
$0.74 \pm 0.04$	$0.79 \pm 0.23$	—	—
QCD $R$ Values, Low Mass Selection (m700)			
mww	bbpt210	wwpt250	hh700
$0.74 \pm 0.04$	$0.79 \pm 0.23$	$0.09 \pm 0.14$	—
QCD $R$ Values, High Mass Selection			
bbpt350	wwpt250	drww15	hh2000
$0.48 \pm 0.09$	$0.43 \pm 0.08$	$0.50 \pm 0.16$	$4.28 \pm 5.30$

TABLE 5.17. Values calculated for  $R$  at each stage in the non-resonant, low mass, and high mass selections. The estimate of multi-jet contribution in the A region uses the  $R$  value calculated after the first cut of each selection.

1435 In order to minimize the statistical error on  $R$ , the  $R$  value calculated after the  
 1436 first cut of each selection (0.74 and 0.48) is used in Equation 5.4 to estimate the multi-  
 1437 jet background after each subsequent cut. This can be done given the compatibility  
 1438 of the  $R$  value at the end of the cut flow with that at the point where  $R$  is evaluated.  
 1439 In order to check such compatibility with higher statistics, the  $R$  value has been  
 1440 calculated applying each selection cut just after the cut where  $R$  is evaluated, in  
 1441 order to check that  $R$  is not correlated with each of the selection cuts. The result is  
 1442 shown in Table 5.18.

1443 Once the normalization of the multi-jet background in the A region is calculated  
 1444 using Equation 5.4, the shape of the multi-jet template is taken from the data - Monte  
 1445 Carlo distribution in the C region since the two are kinematically identical except for  
 1446 the cut on  $|\sigma_{d_0}|$ .

1447 The uncertainty due to the limited statistics in the B and D regions is the main  
 1448 source of the multi-jet estimation method systematics. In order to minimise such

QCD $R$ Values, Non-resonant Selection			
mww	mww + bbpt210	mww + bbpt300	mww + wwpt250
$0.74 \pm 0.04$	$0.79 \pm 0.23$	$1.12 \pm 1.22$	$0.25 \pm 0.20$
QCD $R$ Values, Low Mass (m500) Selection			
mww	mww + bbpt210	mww + wwpt150	mww + hh500
$0.74 \pm 0.04$	$0.79 \pm 0.23$	$0.50 \pm 0.08$	$0.52 \pm 0.09$
QCD $R$ Values, Low Mass (m700) Selection			
mww	mww + bbpt210	mww + wwpt250	mww + hh700
$0.74 \pm 0.04$	$0.79 \pm 0.23$	$0.25 \pm 0.20$	$0.63 \pm 0.13$
QCD $R$ Values, High Mass Selection			
bbpt350	bbpt350 + wwpt250	bbpt350 + drww15	bbpt350 + hh2000
$0.47 \pm 0.06$	$0.44 \pm 0.08$	$0.52 \pm 0.17$	$1.07 \pm 0.67$

TABLE 5.18. Values of  $R$  obtained applying a single selection cut after the cut stage where  $R$  is nominally evaluated.

error, the yields from the B and D regions used in the ABCD calculation are frozen at a level of the cutflow to minimise statistical fluctuations. The B and D region yields are frozen after the  $p_T^{bb} > 210$  GeV cut for the non-resonant and low mass selection, and after the  $p_T(WW) > 250$  GeV for the high mass selection. Appendix E details the study carried out to select the stage at which to freeze the B and D regions.

To further reduce the error coming from the C region, the shape of the data - Monte Carlo, i.e. non-prompt,  $m_{b\bar{b}}$  distribution was studied as a function of each individual cut for each selection. If the  $m_{b\bar{b}}$  shape is unchanged by the additions of cuts later in a given selection, the C region shape can be taken from an earlier stage in the cutflow, reducing the shape uncertainty and overall statistical error on the QCD yield. To determine the stability of the  $m_{b\bar{b}}$  shape, the ratio of events in the  $m_{b\bar{b}}$  signal region ([100, 140] GeV) over the numbers of events in the full  $m_{b\bar{b}}$  spectrum was computed for each individual cut using the C region of the ABCD method. This ratio was found to be stable across each selection, and the results of this calculation are provided in Table 5.19.

reOptNonRes: $m_{b\bar{b}}$ SR/Total Ratios For Individual Cuts			
mww	bbpt210	bbpt300	wwpt250
$0.17 \pm 0.02$	$0.15 \pm 0.03$	$0.13 \pm 0.05$	$0.16 \pm 0.05$
reOpt500: $m_{b\bar{b}}$ SR/CR Ratios For Individual Cuts			
mww	bbpt210	wwpt150	hh500
$0.17 \pm 0.02$	$0.15 \pm 0.03$	$0.18 \pm 0.02$	$0.22 \pm 0.03$
reOpt700: $m_{b\bar{b}}$ SR/Total Ratios For Individual Cuts			
mww	bbpt210	wwpt250	hh700
$0.17 \pm 0.02$	$0.15 \pm 0.03$	$0.16 \pm 0.05$	$0.16 \pm 0.02$
reOpt2000: $m_{b\bar{b}}$ SR/Total Ratios For Individual Cuts			
bbpt350	wwpt250	drww15	hh2000
$0.12 \pm 0.08$	$0.16 \pm 0.05$	$0.19 \pm 0.02$	$0.11 \pm 0.14$

TABLE 5.19. The ratio of events in the  $m_{b\bar{b}}$  signal region ([100, 140] GeV) over the numbers of events in the full  $m_{b\bar{b}}$  spectrum for each individual cut. Both event yields are calculated in the C region of the ABCD method. The ratios are found to be stable across each selection.

1464        The normalized shapes of the full  $m_{b\bar{b}}$  distributions after each cut are shown in  
 1465        Figure 5.6. The earliest selection cut with a consistent shape was chosen as the shape  
 1466        for each cutflow and this corresponds to the shape obtained after the  $p_T^{bb} > 210$  GeV  
 1467        cut for the non-resonant the m500 and m700 selections, and after the  $p_T^{WW} > 250$   
 1468        GeV for the m2000 selection.

1469        Since  $t\bar{t}$  and multi-jet contaminate the control regions used for their estimation,  
 1470        additional studies were performed via an iterative procedure to ensure that the  $t\bar{t}$  and  
 1471        QCD yields converge to stable values and that the estimation technique is able to  
 1472        disentangle between the two backgrounds. Appendix F details the results of the study,  
 1473        and the yields were found to converge (stable within  $< 5\%$ ) after a few iterations for  
 1474        each cutflow.

1475        To evaluate the systematic error of the estimation, a Sherpa multi-jet  $bb$  sample is  
 1476        used to compare the ABCD prediction to the Monte Carlo expectation using events

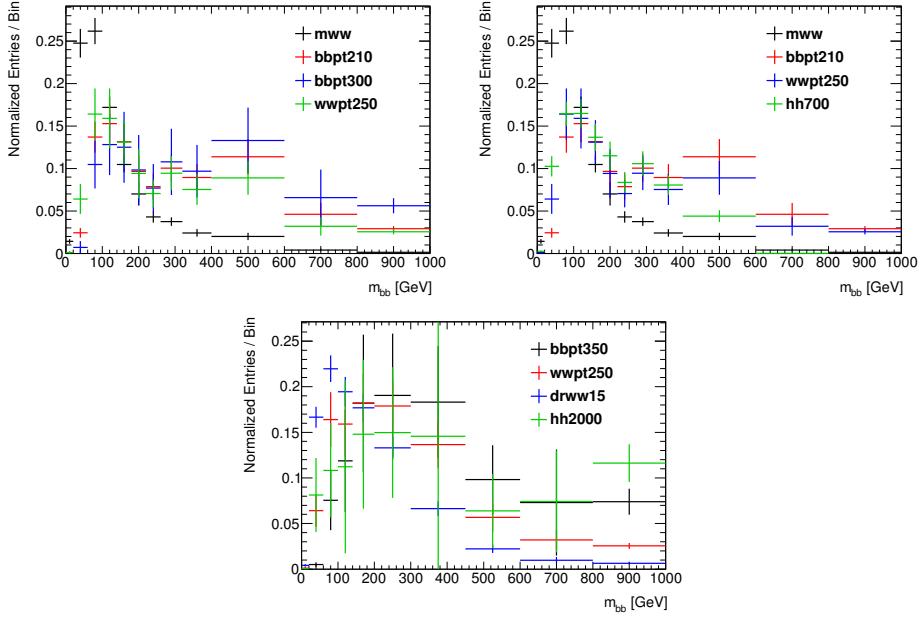


FIGURE 5.6. The normalized shapes of the full  $m_{b\bar{b}}$  distributions after each selection. Shapes from the non-resonant cutflow are shown at the top left, shapes from the low mass selection are shown at the top right, and shapes from the high mass selection are shown at the bottom. The earliest selection cut with a consistent shape was chosen as the shape for each cutflow and this corresponds to the shape obtained after the  $m_{WW}$  and  $p_T^{bb} > 210$  GeV cuts for the non-resonant and low mass selections and after the  $p_T^{bb} > 350$  GeV and  $p_T^{WW} > 250$  GeV for the high mass selection.

with exactly one lepton and four jets. No  $b$ -tagging requirements or other event selection cuts are applied. Pseudo-data is produced using the multi-jet Monte Carlo and events from the nominal  $t\bar{t}$  Monte Carlo. The  $R$  factor is calculated using the inclusive  $b$ -tag (0, 1, or 2  $b$ -tags) and used to estimate the QCD contribution in the two  $b$ -tag exclusive region. The percent difference between the number of events from the ABCD estimation and the number of multi-jet events from Monte Carlo used to produce the pseudo-data is taken as the systematic uncertainty. This uncertainty is calculated to be 36%. In addition, the  $bb$  MC is used to calculate the  $R$  factor for each selection (non-resonant, low mass, and high mass) and each lepton channel, and the results at each cut in each selection are provided in the Table 5.20.

Non-resonant: QCD $R_{non-prompt}$ Values				
	mww	bbpt210	bbpt300	wwpt250
Combined	$0.74 \pm 0.04$	$0.80 \pm 0.05$	$0.64 \pm 0.08$	$0.62 \pm 0.05$
Electrons	$0.72 \pm 0.05$	$0.75 \pm 0.06$	$0.68 \pm 0.10$	$0.64 \pm 0.06$
Muons	$0.81 \pm 0.06$	$0.92 \pm 0.11$	$0.55 \pm 0.15$	$0.59 \pm 0.10$
Low-mass (m500): QCD $R_{non-prompt}$ Values				
	mww	bbpt210	wwpt150	hh500
Combined	$0.74 \pm 0.04$	$0.80 \pm 0.05$	$0.68 \pm 0.03$	$0.50 \pm 0.03$
Electrons	$0.72 \pm 0.05$	$0.75 \pm 0.06$	$0.65 \pm 0.03$	$0.44 \pm 0.04$
Muons	$0.81 \pm 0.06$	$0.92 \pm 0.11$	$0.75 \pm 0.06$	$0.58 \pm 0.07$
Low-mass (m700): QCD $R_{non-prompt}$ Values				
	mww	bbpt210	wwpt250	hh700
Combined	$0.74 \pm 0.04$	$0.80 \pm 0.05$	$0.62 \pm 0.25$	$0.60 \pm 0.03$
Electrons	$0.72 \pm 0.05$	$0.75 \pm 0.06$	$0.64 \pm 0.06$	$0.52 \pm 0.04$
Muons	$0.81 \pm 0.06$	$0.92 \pm 0.11$	$0.59 \pm 0.10$	$0.74 \pm 0.08$
High-mass: QCD $R_{non-prompt}$ Values				
	bbpt350	wwpt250	drww15	hh2000
Combined	$0.47 \pm 0.09$	$0.62 \pm 0.05$	$0.68 \pm 0.03$	$0.58 \pm 0.13$
Electrons	$0.48 \pm 0.10$	$0.64 \pm 0.06$	$0.60 \pm 0.04$	$0.57 \pm 0.14$
Muons	$0.44 \pm 0.20$	$0.59 \pm 0.10$	$0.82 \pm 0.07$	$0.65 \pm 0.39$

TABLE 5.20.  $R_{non-prompt}$  calculated for the QCD Monte Carlo sample at using each cut individually from all selections. The lepton channels are shown separated and combined.

#### 1487 5.4.9. Background Shape and Cutflow

The modelling of the background was checked at all selection stages and, in general, shows good agreement with data. Figure 5.7 shows the  $m_T$  distribution of the leptonic  $W$  boson candidate in the three top control regions. The  $m_T$  variable is defined as:

$$m_T = \sqrt{2p_T^\ell E_T \cdot (1 - \cos\Delta\phi)} ,$$

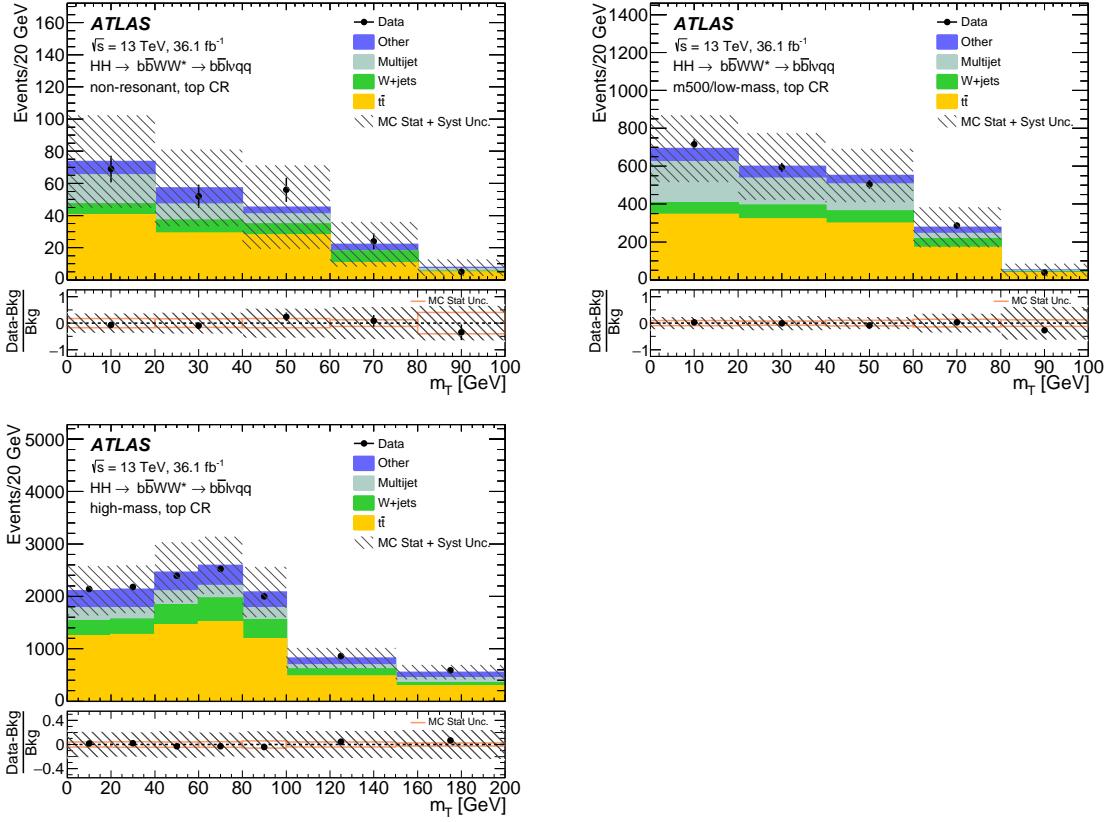


FIGURE 5.7. The  $m_T$  distribution in the three top-background control regions for the *non-res*, *low-mass*, and the *high-mass* selections of the resolved analyses. The signal contamination is negligible, and hence not shown. The lower panel shows the fractional difference between the data and the total expected background with the corresponding statistical and total uncertainty.

where  $\Delta\phi$  is the azimuthal angle between  $p_T^\ell$  and  $\cancel{E}_T$ . The multijet background populates the low values of the  $m_T$  distribution, so any mis-modelling of the multijet background would be clearly visible in the  $m_T$  distribution.

Figures 5.8 and 5.9 show the  $m_{b\bar{b}}$  distributions at the selection stage where all requirements, including the  $m_{HH}$  cut, are applied except the one on  $m_{b\bar{b}}$  itself. The expected background is in agreement with the data over the entire distribution, and close to the signal region in particular. All simulated backgrounds are normalised

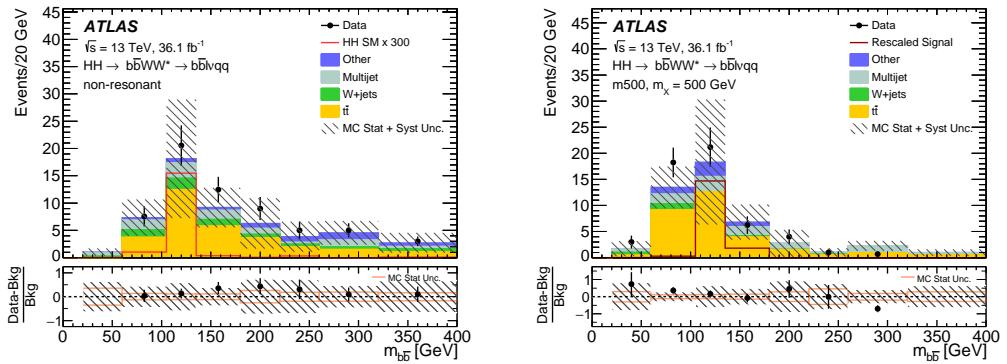


FIGURE 5.8. The  $m_{b\bar{b}}$  distribution in the resolved analysis for the *non-res* and  $m500$  selections at the end of the selection sequence, before applying the  $m_{b\bar{b}}$  requirement. The signals shown are from SM non-resonant  $HH$  production scaled up by a factor of 300 (left) and from a scalar resonance with mass 500 GeV scaled to the expected upper-limit cross section reported in Section ?? (right). The lower panel shows the fractional difference between data and the total expected background with the corresponding statistical and total uncertainty.

according to their theoretical cross-sections, except  $t\bar{t}$ , which is normalised in the top CRs.

1497

1498

1499

#### 1500 5.4.10. Systematic Uncertainties

1501 This section describes the sources of systematic uncertainties considered in  
 1502 the analysis. These uncertainties are divided into four categories: experimental  
 1503 uncertainties, uncertainties on the data driven background estimation, uncertainties  
 1504 on the modelling of background processes estimated from simulation, theoretical  
 1505 uncertainties on the signal processes. In the statistical analysis each systematic  
 1506 uncertainty is treated as a nuisance parameter the names of which are defined below.

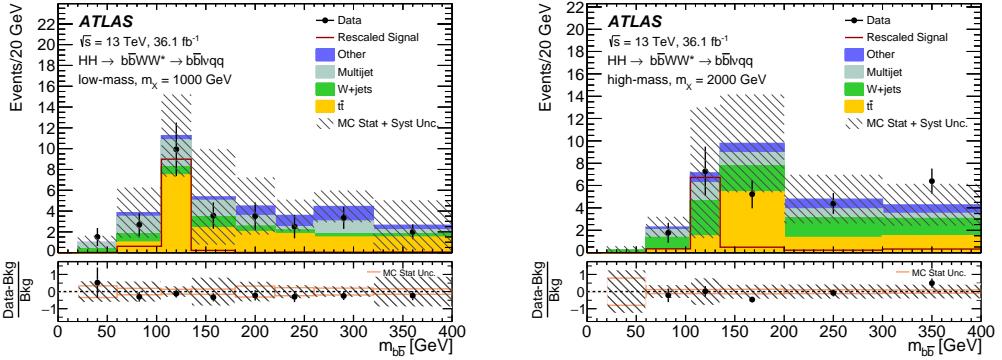


FIGURE 5.9. The  $m_{b\bar{b}}$  distribution in the resolved analysis for the *low-mass* and *high-mass* selections at the end of the selection sequence, before applying the  $m_{b\bar{b}}$  requirement. The signals shown are from scalar resonances with mass 1000 GeV (left) and 2000 GeV (right) scaled to the expected upper-limit cross section reported in Section ???. The lower panel shows the fractional difference between data and the total expected background with the corresponding statistical and total uncertainty.

SR: $100 < m_{b\bar{b}} < 140$ GeV					
Sample	mww	bbpt210	bbpt300	wwpt250	mbb
$t\bar{t}$	$7461.0 \pm 48.6$	$162.9 \pm 7.3$	$27.9 \pm 2.9$	$18.4 \pm 2.4$	$15.4 \pm 2.2$
QCD	$2756.2 \pm 210.5$	$48.7 \pm 14.2$	$6.6 \pm 1.9$	$4.2 \pm 1.2$	$3.6 \pm 1.6$
Wv221	$640.8 \pm 12.7$	$19.1 \pm 1.4$	$5.0 \pm 0.6$	$3.1 \pm 0.5$	$2.3 \pm 0.4$
SingleTop	$452.2 \pm 9.6$	$14.3 \pm 1.7$	$1.7 \pm 0.5$	$1.0 \pm 0.4$	$0.6 \pm 0.3$
Dibosonsv221	$21.6 \pm 1.3$	$0.6 \pm 0.2$	$0.4 \pm 0.2$	$0.0 \pm 0.0$	$0.0 \pm 0.0$
Zv221	$262.8 \pm 4.4$	$3.1 \pm 0.3$	$1.0 \pm 0.2$	$0.2 \pm 0.1$	$0.2 \pm 0.1$
Background Sum	$11594.7 \pm 216.7$	$248.6 \pm 16.1$	$42.6 \pm 3.6$	$27.0 \pm 2.8$	$22.1 \pm 2.8$
XhhSM	$68.3 \pm 2.4$	$20.7 \pm 0.9$	$6.7 \pm 0.4$	$5.5 \pm 0.3$	$4.8 \pm 0.3$
Data	11450.0	232.0	47.0	31.0	22.0

TABLE 5.21. The number of expected background and signal events in the  $m_{b\bar{b}}$  SR for the non-resonant selection. Only statistical uncertainties are shown. No NF has been applied.

SR: $100 < m_{b\bar{b}} < 140$ GeV					
Sample	mww	bbpt210	wwpt150	hh500	mbb
$t\bar{t}$	$7461.0 \pm 48.6$	$162.9 \pm 7.3$	$141.7 \pm 6.8$	$17.3 \pm 2.2$	$12.6 \pm 1.9$
QCD	$2756.2 \pm 210.5$	$48.7 \pm 14.2$	$40.2 \pm 11.7$	$3.3 \pm 1.0$	$2.9 \pm 1.3$
Wv221	$640.8 \pm 12.7$	$19.1 \pm 1.4$	$15.3 \pm 1.3$	$0.1 \pm 0.0$	$-0.2 \pm 0.1$
SingleTop	$452.2 \pm 9.6$	$14.3 \pm 1.7$	$12.2 \pm 1.6$	$3.6 \pm 0.8$	$2.8 \pm 0.7$
Dibosonsv221	$21.6 \pm 1.3$	$0.6 \pm 0.2$	$0.5 \pm 0.2$	$0.1 \pm 0.0$	$0.1 \pm 0.0$
Zv221	$262.8 \pm 4.4$	$3.1 \pm 0.3$	$1.9 \pm 0.2$	$0.5 \pm 0.1$	$0.4 \pm 0.1$
Background Sum	$11594.7 \pm 216.7$	$248.6 \pm 16.1$	$211.8 \pm 13.7$	$24.9 \pm 2.5$	$18.6 \pm 2.4$
Xhh500	$6.6 \pm 0.2$	$1.9 \pm 0.1$	$1.7 \pm 0.1$	$0.9 \pm 0.1$	$0.8 \pm 0.1$
Data	11450.0	232.0	194.0	32.0	26.0

TABLE 5.22. The number of expected background and signal events in the  $m_{b\bar{b}}$  SR for the low-mass selection, m500. Only statistical uncertainties are shown. No NF has been applied.

CR: $m_{b\bar{b}}$ Sideband					
SR: $100 < m_{b\bar{b}} < 140$ GeV					
Sample	mww	bbpt210	wwpt250	hh700	mbb
$t\bar{t}$	$7461.0 \pm 48.6$	$162.9 \pm 7.3$	$61.5 \pm 4.7$	$21.9 \pm 2.7$	$15.3 \pm 2.2$
QCD	$2756.2 \pm 210.5$	$48.7 \pm 14.2$	$14.1 \pm 4.1$	$5.5 \pm 1.6$	$4.8 \pm 2.2$
Wv221	$640.8 \pm 12.7$	$19.1 \pm 1.4$	$9.7 \pm 1.1$	$4.1 \pm 0.8$	$2.6 \pm 0.6$
SingleTop	$452.2 \pm 9.6$	$14.3 \pm 1.7$	$2.6 \pm 0.7$	$0.5 \pm 0.2$	$0.3 \pm 0.2$
Dibosonsv221	$21.6 \pm 1.3$	$0.6 \pm 0.2$	$0.2 \pm 0.1$	$0.2 \pm 0.1$	$0.2 \pm 0.1$
Zv221	$262.8 \pm 4.4$	$3.1 \pm 0.3$	$0.6 \pm 0.1$	$0.1 \pm 0.0$	$0.1 \pm 0.0$
Background Sum	$11594.7 \pm 216.7$	$248.6 \pm 16.1$	$88.7 \pm 6.4$	$32.3 \pm 3.2$	$23.3 \pm 3.1$
Xhh700	$9.2 \pm 0.3$	$7.8 \pm 0.2$	$5.9 \pm 0.2$	$5.0 \pm 0.2$	$4.4 \pm 0.2$
Data	11450.0	232.0	75.0	25.0	22.0

TABLE 5.23. The number of expected background and signal events in the  $m_{b\bar{b}}$  SR for the low-mass selection, m700. Only statistical uncertainties are shown. No NF has been applied.

SR: $100 < m_{b\bar{b}} < 140$ GeV					
Sample	bbpt350	wwpt250	drww15	hh2000	mbb
$t\bar{t}$	$1307.8 \pm 20.2$	$1024.9 \pm 17.7$	$287.5 \pm 9.4$	$2.2 \pm 0.8$	$1.4 \pm 0.6$
QCD	$207.2 \pm 99.5$	$191.2 \pm 29.0$	$55.2 \pm 8.4$	$2.9 \pm 0.4$	$2.2 \pm 0.5$
Wv221	$341.3 \pm 3.4$	$291.5 \pm 3.2$	$110.7 \pm 2.1$	$4.8 \pm 0.3$	$3.4 \pm 0.3$
SingleTop	$144.1 \pm 5.6$	$126.6 \pm 5.3$	$29.2 \pm 2.6$	$0.5 \pm 0.3$	$0.5 \pm 0.3$
Dibosonsv221	$25.9 \pm 1.5$	$21.8 \pm 1.3$	$6.6 \pm 0.7$	$0.0 \pm 0.0$	$0.0 \pm 0.0$
Zv221	$53.8 \pm 0.8$	$40.4 \pm 0.7$	$13.2 \pm 0.4$	$0.8 \pm 0.1$	$0.7 \pm 0.1$
Background Sum	$2080.1 \pm 101.8$	$1696.5 \pm 34.6$	$502.5 \pm 13.1$	$11.2 \pm 1.0$	$8.2 \pm 0.8$
Xhh2000	$21.0 \pm 0.4$	$19.3 \pm 0.4$	$8.4 \pm 0.2$	$3.4 \pm 0.1$	$2.9 \pm 0.1$
Data	2182.0	1830.0	587.0	11.0	9.0

TABLE 5.24. The number of expected background and signal events in the  $m_{b\bar{b}}$  SR for the high-mass selection, m2000. Only statistical uncertainties are shown. No NF has been applied.

1507 These systematic variations are estimated on the final expected yield in the signal  
 1508 regions.

#### 1509 5.4.11. Experimental uncertainties

1510 Each reconstructed object has several sources of uncertainties, each of which  
 1511 are evaluated separately. Wherever possible, we follow the latest available  
 1512 recommendations from the combined performance (CP) groups. The leading  
 1513 instrumental uncertainties are the uncertainty on the  $b$ -tagging efficiency and the  
 1514 jet energy scale (JES). The summary of experimental uncertainties is presented in  
 1515 Table 5.25.

Source	Description	Analysis Name
Muons	$p_T$ resolution MS	MUON_MS
Muons	$p_T$ resolution ID	MUON_ID
Muons	$p_T$ scale	MUON_SCALE
Muons	Isolation efficiency SF	MUON_ISO_SYS
Muons	Isolation efficiency SF	MUON_ISO_STAT
Muons	Reconstruction efficiency SF	MUON_EFF_SYS
Muons	Reconstruction efficiency SF	MUON_EFF_STAT
Muons	Trigger efficiency SF	MUON_EFF_TrigStatUncertainty
Muons	Trigger efficiency SF	MUON_EFF_TrigSystUncertainty
Electrons	$p_T$ resolution	EG_RESOLUTION_ALL
Electrons	$p_T$ scale	EG_SCALE_ALL
Electrons	Isolation efficiency SF	EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR
Electrons	Reconstruction efficiency SF	EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR
Electrons	Trigger efficiency SF	EL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR
Electrons	Identification efficiency SF	EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR
Tau	Energy scale model	TAUS_TRUEHADTAU_SME_TES_MODEL
Tau	Energy scale detector	TAUS_TRUEHADTAU_SME_TES_DETECTOR
Tau	In-situ energy calibration	TAUS_TRUEHADTAU_SME_TES_INSITU
MET	Soft term	MET_SoftTrk_ResoPerp
MET	Soft term	MET_SoftTrk_ResoPara
MET	Soft term	MET_SoftTrk_Scale
Small-R Jets	JES strongly reduced	JET_SR1_JET_GroupedNP_1
Small-R Jets	JES strongly reduced	JET_SR1_JET_GroupedNP_2
Small-R Jets	JES strongly reduced	JET_SR1_JET_GroupedNP_3
Small-R Jets	JES strongly reduced	JET_SR1_JET_EtaIntercalibration_NonClosure
Small-R Jets	Energy resolution	JET_JER_SINGLE_NP
Small-R Jets	JVT efficiency SF	JET_JvtEfficiency
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_0
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_1
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_2
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_3
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_4
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_5
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_6
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_7
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_8
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_9
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_10
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_11
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_12
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_Light_13
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_B_0
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_B_1
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_B_2
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_B_3
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_B_4
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_C_0
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_C_1
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_C_2
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_C_3
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_C_4
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_extrapolation
$b$ -tagging	Flavor tagging scale factors	FT_EFF_Eigen_extrapolation_from_charm

TABLE 5.25. Qualitative summary of the object systematic uncertainties included in this analysis.

1516    **5.4.11.1. Luminosity**

1517       The uncertainty in the combined 2015+2016 integrated luminosity is 2.1%<sup>11</sup>.  
1518       It is derived, following a methodology similar to that detailed in [21], from a  
1519       preliminary calibration of the luminosity scale using x-y beam-separation scans  
1520       performed in August 2015 and May 2016. The luminosity uncertainty is applied  
1521       to those backgrounds estimated from simulation and the signal samples.

1522    **5.4.11.2. Trigger**

1523       Systematic uncertainties on the efficiency of electron and muon triggers are  
1524       evaluated as recommended by the corresponding combined performance groups as  
1525       documented here.<sup>12</sup>

1526    **5.4.11.3. Muons**

1527       The following systematic uncertainties are applied to muons in estimations based  
1528       on the simulation:

- 1529       – Identification efficiency: The efficiencies are measured with the tag and probe  
1530       method using the  $Z$  mass peak.
- 1531       – Energy and Momentum scales: These are also measured with  $Z$  mass line shape,  
1532       and provided by the CP groups.

---

<sup>11</sup> <https://twiki.cern.ch/twiki/bin/view/Atlas/LuminosityForPhysics>

<sup>12</sup> el: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ElectronEfficiencyRun2>  
mu:<https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCPAnalysisGuidelinesMC15>

1533    **5.4.11.4. Electrons**

1534        The following systematic uncertainties are applied to electron in estimations  
1535        based on the simulation:

- 1536        – Identification efficiency: The efficiencies are measured with the tag and probe  
1537           method using the  $Z$  mass peak. They include contributions from reconstruction,  
1538           identification and isolation;
- 1539        – Energy and Momentum scales: These are also measured with  $Z$  mass line shape,  
1540           and provided by the CP groups.

1541    **5.4.11.5. Jet uncertainties**

1542        The jet energy uncertainties are derived based on in situ measurements performed  
1543        during Run 2 conditions [80]. The jet energy resolution uncertainty is evaluated  
1544        by smearing jet energies according to the systematic uncertainties of the resolution  
1545        measurement [81]. The uncertainty in the  $b$ -tagging efficiency is evaluated by  
1546        propagating the systematic uncertainty in the measured tagging efficiency for  $b$ -  
1547        jets [82]. The “Loose” reduction scheme is used.

1548    **5.4.11.6. Missing transverse energy**

1549        The systematic uncertainties related to the missing transverse energy are  
1550        obtained by the propagation of the systematic uncertainty on the objects that  
1551        build the MET, in particular the muon, electron and jets energy resolution and  
1552        scale. The resolution and scale of the MET soft-term is broken down into its  
1553        components: METScale, METResoPara, METResoPerp, and full uncertainties from  
1554        each component is taken into account in the final fit.

1555    **5.4.11.7.  $d_0^{\text{sig}}$  uncertainties**

1556       The uncertainty due to the  $d_0^{\text{sig}}$  cut has been evaluated by making the ratio  
 1557       between the efficiency of the cut for data and the efficiency of the cut for the MC  
 1558       background samples. We selected di-electron or di-muon event, requiring an invariant  
 1559       di-leptons mass within 80-100 GeV Z Mass window. To be as similar to our signal  
 1560       region as possible but to keep high statistics, loose pre-selection cuts are applied  
 1561       in selecting the events. The leading lepton is required to have  $p_T > 27$  GeV and  
 1562        $\cancel{E}_T > 25$  GeV. At least 4 resolved jets are required of which exactly 2 are  $b$ -jets.  
 1563       The  $d_0^{\text{sig}}$  distributions for data and MC samples for each lepton channel are shown in  
 1564       the Figure 5.10, in Figure 5.11 the relative ratio of the total distributions is shown.  
 1565       The ratio of the efficiency of the  $d_0^{\text{sig}}$  cut for data over MC samples is about 96%,  
 1566       this is equivalent if the ratio is estimated by using only muons or only electrons.  
 1567       The difference of this ratio from one is the fractional uncertainty due to the  $d_0^{\text{sig}}$  cut  
 1568       efficiency.

1569       This results in about 4% for the  $d_0^{\text{sig}}$  uncertainty independent from the lepton  
 flavour.

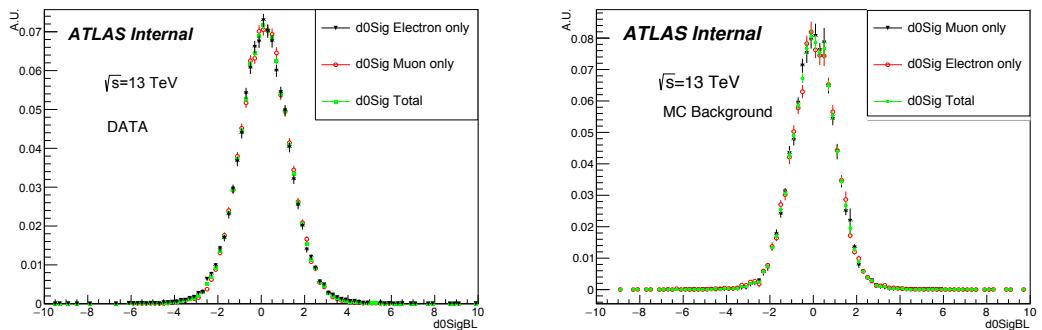


FIGURE 5.10.  $d_0^{\text{sig}}$  distributions for data and background MC samples, identifying the lepton channel.

1570

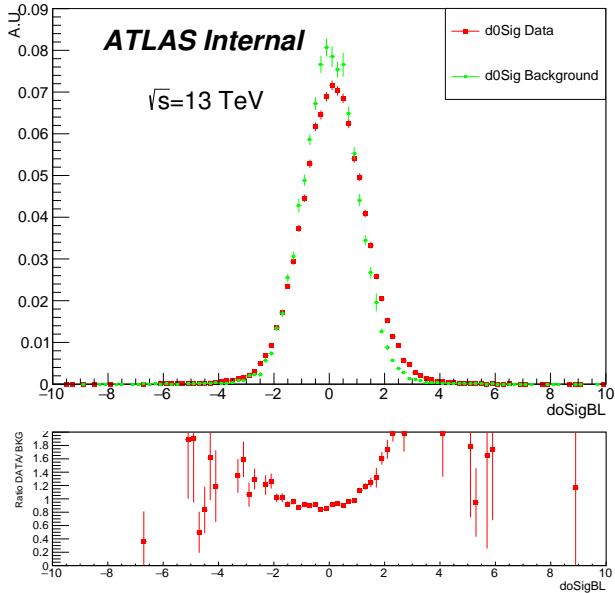


FIGURE 5.11.  $d_0^{\text{sig}}$  distributions and the relative ratio for data and background MC samples.

#### 5.4.12. Background modelling uncertainties

Several systematics have been evaluated to take into account the uncertainties on the modelling of backgrounds.

##### 5.4.12.1. Uncertainties from the modelling of $t\bar{t}$

The dominant background  $t\bar{t}$  is normalised in dedicated CRs. MC is used to extrapolate the shapes from the control regions to the signal region, so theoretical uncertainties are related to such extrapolation. PDF and scale uncertainties are evaluated by applying event selection at truth level. The resulting uncertainties are approximately 5 to 6% and are included in the final fit.

Additional uncertainties in  $t\bar{t}$  modelling stems from the difference in the matrix element (ME) implementation across generators, hadronisation and fragmentation

Source	Non-resonant	Low-Mass	High-Mass
Matrix Element	7.2	0.5	4.1
Parton Shower	3.7	16.4	9.5
ISR/FSR	14.7	4.9	8.2
PDF	5.2	3.5	6.1
Scale	3.3	2.2	3.7

TABLE 5.26. Extrapolation uncertainties in percentage from the CR to the SR for all selections, provided to the fit for the  $t\bar{t}$  modelling systematics.

modelling (called parton shower, PS), and the amount of initial and final state radiation (ISR/FSR). The ME uncertainty is computed by comparing the events generated by AMC@NLO with the events generated by POWHEG-Box v2, both interfaced to HERWIG++ for parton shower. The difference computed close to the signal region with enough statistics is used. The PS uncertainty is computed by comparing the the nominal POWHEG+PYTHIA6 sample with the PS variation POWHEG+HERWIG++ sample in a region close to the SR but with enough statistics. For ISR/FSR, the dedicated RADHI and RADLO samples with modified HDAMP parameter are compared. The sample with the higher impact on the fit is kept as the uncertainty due to ISR/FSR. Table 5.26 shows the numbers provided to the fit for the various  $t\bar{t}$  modelling systematics for the low and high mass selections. Since the uncertainties are computed based on extrapolation from the CRs, these uncertainties are input to the statistical fit as rate uncertainties in the SR only.

#### 5.4.12.2. Single top uncertainty

Theoretical cross section uncertainties of 5.3% is assigned to the associated  $Wt$  production, 3.9% to the s-channel and 4.2% to t-channel single top production. The dominant production for this analysis is the  $Wt$  channel. The single top modelling systematic uncertainties have been calculated employing the difference between the

nominal (DR scheme) and the (available) systematic variation sample (DS scheme) in SR for the  $Wt$  channel. The uncertainties are computed to be 48%, 48%, and 84% for non-resonant, low-mass and high-mass analyses respectively. Single top is a very small background in the analysis. However, the large difference between the two schemes warranted getting feedback from the Top Group. We narrowed the huge difference down to very tight  $p_T(bb)$  cut. See Figure 5.12. The Top Group is working on a better prescription but in the meantime this is the recommended method and the huge difference is what we keep. Additional uncertainties on single top have been calculated employing the difference between the nominal and the (available) systematic variation samples. The recommendation is taken from the Top Twiki.<sup>13</sup> The comparison is done between the nominal sample and the Powheg+Herwig for fragmentation, MC@NLO for Matrix Element, and RadHi/RadLo for ISR/FSR uncertainties. The comparison is done at the selection where MC statistical uncertainties are small. This leads to  $p_T(bb) < 120$  for non-resonance and low-mass selections, and  $\Delta R(WW) < 1.5$  for high-mass selection. The uncertainties vary across selections being 3.5% at the smallest to 23% at the largest.

#### 5.4.12.3. $W/Z+jets$ modelling uncertainty

Uncertainties on the modelling of  $W+jets$  background were computed in each SR and top CR. Three sources of uncertainties were considered: scale variation uncertainties, PDF uncertainties and PS/modelling uncertainties. Scale uncertainties were computed by varying by a factor 2 the nominal renormalisation and factorisation scales, PDF uncertainties were computed according to the NNPDF recipe, that is computing the standard deviation of the 100 PDF eigenset, while modelling uncertainties were computed comparing Sherpa with Alpgen+Pythia6. The values

---

<sup>13</sup> <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopSystematics2015>

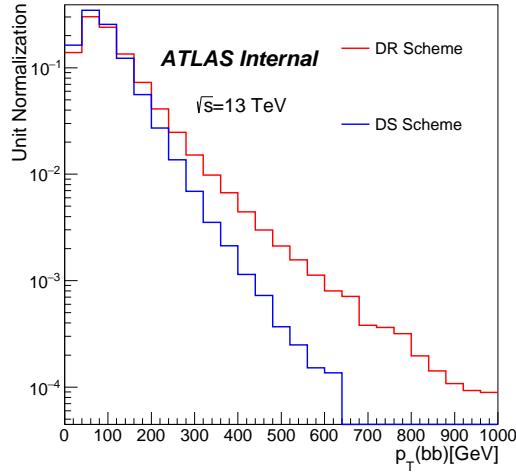


FIGURE 5.12.  $p_T(bb)$  for DR and DS schemes for single top modelling.

Source	Non-resonant		Low-Mass		High-Mass	
	SR	CR	SR	CR	SR	CR
Modelling/PS	37	37	37	37	18	18
PDF	7	30	10	36	17	31
Scale	25	35	17	31	29	28

TABLE 5.27. Theoretical uncertainties in percentage on  $W/Z+\text{jets}$  event yield computed in the CR and the SR of all selections, provided to the fit for the  $W/Z$  jets modelling systematics.

obtained in each region are summarised in Table 5.27. Same uncertainties are used for the subleading  $Z+\text{jets}$  background.

#### 5.4.12.4. QCD uncertainty

An overall 36% uncertainty is assigned to the QCD multijet background. The systematic uncertainty was calculated following the steps described in Section 5.4.8.

Signal model	Non-resonant	$m_H 500 - 1300 \text{ GeV}$	$m_H > 1300 \text{ GeV}$
scale	$\pm 1.1\%$	$\pm 0.8 \%$	$\pm 0.7 \%$
PDF	$\pm 1.3\%$	$\pm 1.3 \%$	$\pm 1.3 \%$

TABLE 5.28. Theoretical uncertainties in percentage on the signal acceptance.

1629 **5.4.13. Model uncertainties on the signal**

1630 Systematics on signal acceptance are computed generating multiple-weight  
 1631 samples that include weights corresponding to variation of the normalisation and  
 1632 factorisation scales by a factor  $(\xi_r, \xi_f) = 2$ . The envelope is built excluding the  
 1633 cases where  $\xi_r/\xi_f = 4$  and  $\xi_r/\xi_f = 1/4$ . The fractional uncertainty is obtained  
 1634 dividing 1/2 of the envelop by the central value. PDF uncertainties are compute  
 1635 using PD4LHC\_mc\_30 pdf sets, that include the envelope of three PDF sets, namely  
 1636 the CT14, MMHT'14, NNPDF3. The CTEQ error formula is used to compute the  
 1637 uncertainties. Results are summarised in table 5.28 for each signal hypothesis.

1638 **5.5. Boosted Analysis**

1639 **5.5.1. Event Reconstruction**

1640 For the boosted analysis, events are reconstructed by requiring at least one  
 1641 reconstructed lepton. This lepton will be referred to as the selected lepton. To  
 1642 reconstruct the  $H \rightarrow b\bar{b}$  candidate, there should be at least one large-R jet with  
 1643  $\Delta R > 1.0$  from the selected lepton. The highest  $p_T$  large-R jet is chosen as the  
 1644  $H \rightarrow b\bar{b}$  candidate. The large-R jet is then required to have two at least two  
 1645 track jets associated to it. Events with the large-R jet mass to be in the range  
 1646 of  $30\text{GeV} < m_{\text{large-R jet}} < 300\text{GeV}$  are retained for further analysis.

1647 In order to reconstruct  $W \rightarrow q\bar{q}$  candidate, events are required have at least two  
 1648 signal small-R jets with  $\Delta R > 1.4$  of the  $H \rightarrow b\bar{b}$  candidate. The  $\Delta R$  requirement

ensures that the small-R jets used for  $W \rightarrow q\bar{q}$  reconstruction do not overlap with the  $H \rightarrow b\bar{b}$  candidate large-R jet. If there are exactly two signal small-R jets, they are used to reconstruct the  $W \rightarrow q\bar{q}$  candidate. If there are at least three signal small-R jets, the  $W \rightarrow q\bar{q}$  candidate is reconstructed from the combination of pairs of the three highest  $p_T$  small-R jets with the smallest  $\Delta R$  between the small-R jets.

The full  $h \rightarrow WW^* \rightarrow l\nu qq$  is reconstructed identically to the resolved analysis as described in section 5.4.3. With the  $H \rightarrow b\bar{b}$  candidate identified and  $h \rightarrow WW^* \rightarrow l\nu qq$  system fully reconstructed, the di-Higgs (HH) system is reconstructed by the sum of four-momenta of the  $H \rightarrow b\bar{b}$  candidate large-R jet and the 552 reconstructed  $h \rightarrow WW^* \rightarrow l\nu qq$  system. Figure 5.13 shows a diagram of the event topology after the event reconstruction.

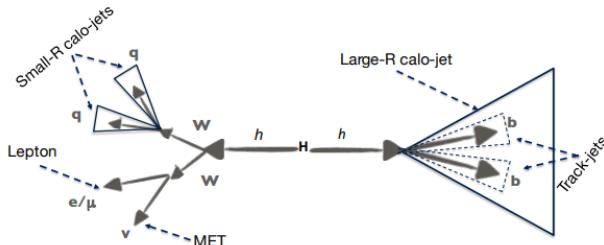


FIGURE 5.13. Diagram of the reconstructed event topology

### 5.5.2. Event Selection

After the event is reconstructed, a b-jet veto is applied on the event by requiring all signal small-R jets do not pass b-tagging requirement to reject  $t\bar{t}$  events. The two highest  $p_T$  track-jets in the large-R jet are required to be b-tagged and events that passes this requirement is considered to be in the "2-tag" region. In addition, the  $E_T$  is required to be more than 50 GeV to reject events from QCD multijet background.

1666 **5.5.3. Kinematic Selection**

1667 **5.5.3.1. Signal Region**

1668 In order to enhance the sensitivity to a resonant HH signal, similarly to the  
1669 resolved analysis, it is required that the  $H \rightarrow b\bar{b}$  candidate large-R jet has a mass  
1670 consistent with the Standard Model Higgs boson mass. Events which have the  
1671  $H \rightarrow b\bar{b}$  candidate large-R jet mass in a window of  $90\text{GeV} < m_{\text{Large-Rjet}} < 140\text{GeV}$   
1672 is considered to be in the signal region (SR). The signal region is blinded and the  
1673 blinding strategy is implemented by removing any events in data that passes the  
1674 signal region requirement on the large-R jet mass.

1675 **5.5.3.2. mBB Control Region**

1676 In order to asses the modeling of the backgrounds, a control region is defined to  
1677 be any events which fails the large-R jet signal region mass window requirement. Any  
1678 event which has a large-R jet mass  $90\text{GeV} < m_{\text{Large-Rjet}}$  or  $m_{\text{Large-Rjet}} > 140\text{GeV}$  falls  
1679 in the mBB control region (mBBcr). By construction, this region is orthogonal to  
1680 the signal region.

1681 **5.5.4. Multijet Background**

1682 As with the resolved analysis, the QCD multijet background is estimated using  
1683 the same data-driven method with slight modifications.

1684 The ABCD method uses three control regions (the B, C, and D regions) to  
1685 estimate the contribution of a given background in the signal (A) region. Cuts on  
1686 two ideally orthogonal variables are used to create the signal and various control  
1687 regions, e.g. the A region passes both cuts, the B and C regions each pass one cut  
1688 and fail the other, while the D region fails both cuts.

1689 For the boosted analysis, the ABCD regions are defined by using the same  
 1690 variables as in the resolved analysis, which are the significance of the lepton impact  
 1691 parameter ( $d_0^{\text{sig}}$ ) and the missing transverse momentum ( $\cancel{E}_T$ ). A small difference  
 1692 would be the higher cut value for the  $\cancel{E}_T$  with respect to the resolved analysis. The  
 1693 regions are defined as follow:

- 1694 – Region A:  $\cancel{E}_T > 50 \text{ GeV}$ ,  $|d_0^{\text{sig}}| < 2.0$
- 1695 – Region B:  $\cancel{E}_T < 50 \text{ GeV}$ ,  $|d_0^{\text{sig}}| < 2.0$
- 1696 – Region C:  $\cancel{E}_T > 50 \text{ GeV}$ ,  $|d_0^{\text{sig}}| > 2.0$
- 1697 – Region D:  $\cancel{E}_T < 50 \text{ GeV}$ ,  $|d_0^{\text{sig}}| > 2.0$

1698 Figure 5.14 shows the four regions represented on on the lepton  $d_0$  significance  
 1699 vs  $\cancel{E}_T$  plane. Assuming that the two variables chosen to define the ABCD regions are  
 1700 completely uncorrelated, the QCD multijet yield in region A can be predicted. The  
 1701 The correlation between  $|d_0^{\text{sig}}|$  vs  $\cancel{E}_T$  is estimated in multiple MC background samples  
 1702 and also in data, and they are found to negligible.

1703 The ABCD method is done separately between the muon and electron channel  
 1704 as it is expected that the  $\frac{N_B^{\text{QCD}}}{N_D^{\text{QCD}}}$  ratio and QCD multijet contribution to the total  
 1705 predicted background will be significantly different between the channels.

#### 1706 5.5.4.1. Yield prediction

1707 Table 5.29 lists the MC predicted prompt lepton backgrounds, observed data and  
 1708 calculated multijet yields in Region B and D before the large- $R$  jet mass is applied  
 1709 and Table 5.30 shows the yields in Region C mBB control region and signal region.  
 1710 The  $\frac{N_B^{\text{QCD}}}{N_D^{\text{QCD}}}$  ratio are calculated inclusively in the large- $R$  jet mass distribution.  
 1711 In other words, the ratio is calculated without the jet mass window selection, which

Samples	Region B		Region D	
	Electron	Muon	Electron	Muon
$t\bar{t}$	$307.7 \pm 11.5$	$279.8 \pm 10.6$	$21.3 \pm 2.7$	$18.3 \pm 2.6$
W+Jets	$173.2 \pm 5.2$	$179.3 \pm 5.6$	$11.6 \pm 1.4$	$10.6 \pm 1.1$
Single-top	$42.9 \pm 3.4$	$33.5 \pm 3.6$	$3.0 \pm 0.9$	$0.8 \pm 0.5$
Z+Jets	$78.5 \pm 1.9$	$72.5 \pm 1.7$	$6.4 \pm 0.6$	$5.5 \pm 0.5$
Dibosons	$19.1 \pm 1.5$	$17.7 \pm 1.5$	$1.6 \pm 0.4$	$2.2 \pm 0.8$
Total Prompt	$621.3 \pm 13.2$	$582.7 \pm 12.7$	$44.0 \pm 3.3$	$37.4 \pm 3.0$
Data	$1003 \pm 31.7$	$711 \pm 26.7$	$144 \pm 12.0$	$98 \pm 9.9$
QCD	$381.7 \pm 34.3$	$128.3 \pm 29.5$	$100.0 \pm 12.4$	$60.6 \pm 10.4$

TABLE 5.29. MC predicted prompt lepton backgrounds, observed data and calculated multijet yields in Region B and D. The multijet yield is calculated by subtracting the estimated total prompt lepton backgrounds from the observed data. The statistical uncertainty on the yields is shown.

Samples	mBBcr		SR	
	Electron	Muon	Electron	Muon
$t\bar{t}$	$38.7 \pm 4.2$	$46.8 \pm 7.9$	$28.5 \pm 3.1$	$22.0 \pm 2.7$
W+Jets	$22.3 \pm 2.0$	$20.0 \pm 1.7$	$9.6 \pm 1.3$	$10.0 \pm 1.8$
Single-top	$7.6 \pm 2.1$	$6.5 \pm 1.3$	$7.1 \pm 1.5$	$2.7 \pm 0.8$
Z+Jets	$4.6 \pm 0.8$	$3.8 \pm 0.5$	$1.6 \pm 0.3$	$1.9 \pm 0.6$
Dibosons	$2.2 \pm 0.6$	$1.2 \pm 0.4$	$0.8 \pm 0.3$	$1.7 \pm 0.4$
Total Prompt	$75.4 \pm 5.2$	$78.4 \pm 8.2$	$47.5 \pm 3.7$	$38.3 \pm 3.4$
Data	$148 \pm 12.2$	$126 \pm 11.2$	$91 \pm 9.5$	$71 \pm 8.4$
QCD	$72.6 \pm 13.2$	$47.6 \pm 13.9$	$43.5 \pm 10.2$	$32.7 \pm 9.1$

TABLE 5.30. MC predicted prompt lepton backgrounds, observed data and calculated multijet yields in Region C mBBcr and SR. The multijet yield is calculated by subtracting the estimated total prompt lepton backgrounds from the observed data. The statistical uncertainty on the yields is shown.

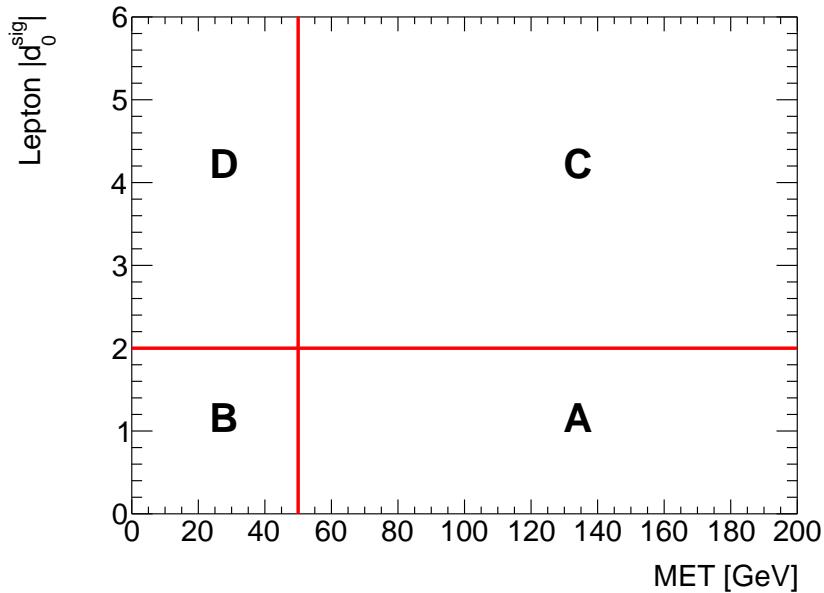


FIGURE 5.14. Regions defined in the ABCD method based on the lepton  $d_0$  significance vs  $\cancel{E}_T$  plane. Region A is the signal enriched region which we want to estimate the multijet background. Region C is where the shape template is derived from and used as shape prediction of the multijet background in region A. The ratio of the multijet yields in region B to region D is used to scale the multijet yield in region C to predict the multijet background yield in region A.

<sub>1712</sub> defines the SR and mBBcr, applied. The ratio is then used to scale the QCD multijet  
<sub>1713</sub> yield in the SR and mBBcr of region C to predict the QCD multijet yield in region  
<sub>1714</sub> A.

<sub>1715</sub> Table 5.31 shows the ratio in the electron channel and muon channel. As  
<sub>1716</sub> expected, the electron channel ratio is larger than the muon channel. The muon  
<sub>1717</sub> channel ratio has a larger uncertainty due to the more limited statistics in region B  
<sub>1718</sub> and region D compared to the electron channel. The predicted yields of the QCD  
<sub>1719</sub> multijet background in the mBB control region and signal region are presented in  
<sub>1720</sub> Table 5.32. The QCD multijet background is estimated to be 19% of the total  
<sub>1721</sub> background in the signal region (Table 5.52).

Multijet yield in region	Electron	Muon
$N_B^{\text{QCD}}$	$381.7 \pm 34.3$	$128.3 \pm 29.5$
$N_D^{\text{QCD}}$	$100.0 \pm 12.4$	$60.6 \pm 10.4$
$N_B^{\text{QCD}}/N_D^{\text{QCD}}$	$3.8 \pm 0.6$ (15.3%)	$2.1 \pm 0.6$ (28.7%)

TABLE 5.31. Multijet yields in region B and region D and also the ratio of the yields for each lepton channel. The error on the  $\frac{N_B^{\text{QCD}}}{N_D^{\text{QCD}}}$  ratio is propagated from the statistical uncertainties on the multijet yields in each region.

Multijet yield in region	Electron	Muon
SR		
$N_C^{\text{QCD}}$	$43.4 \pm 10.2$	$32.7 \pm 9.1$
$N_A^{\text{QCD}}$	$165.9 \pm 46.6$ (28.1%)	$69.3 \pm 27.7$ (39.9%)
mBBcr		
$N_C^{\text{QCD}}$	$72.6 \pm 13.2$	$47.6 \pm 13.9$
$N_A^{\text{QCD}}$	$277.1 \pm 66.0$ (23.8%)	$100.8 \pm 41.3$ (41.0%)

TABLE 5.32. Multijet yield in region C and predicted yield in region A in the SR. The error on  $N_A^{\text{QCD}}$  are propagated from the error on the  $N_B^{\text{QCD}}/N_D^{\text{QCD}}$  ratio and statistical uncertainty on  $N_C^{\text{QCD}}$  yield. The numbers in brackets are the relative uncertainty in percentage.

1722    **5.5.4.2. Shape prediction**

1723       In order to predict the shape of the HH mass distribution (and also other  
1724       kinematic distribution) of the QCD multijet background, the shape template of all  
1725       kinematic distributions are obtained by subtracting all the MC backgrounds from  
1726       data in region C.

1727       It was found that distributions in region C suffer from lack of statistics due the  
1728       low number of data events which results in shape templates with severe statistical  
1729       fluctuations. To overcome this, the shape templates are derived from a sample of 1  
1730        $b$ -tagged (1-tag) events in region C. This sample requires that one of the two leading  
1731       track-jet is  $b$ -tagged but not both at the same time.

1732    **5.5.4.3. Multijet yield uncertainties**

1733       **5.5.4.3.1. Statistical**    The uncertainty on the predicted yield of the multijet  
1734       background is determined by propagating the statistical uncertainty of the  $\frac{N_B^{\text{QCD}}}{N_D^{\text{QCD}}}$   
1735       ratio, as shown in Table 5.31, and the statistical uncertainty on the multijet yield in  
1736       region C ( $N_C^{\text{QCD}}$ ), as in Table 5.32.

1737       **5.5.4.3.2. 1-tag/2-tag jet mass acceptance**    Another source of uncertainty  
1738       on the multijet yield is the the difference of acceptance of the large- $R$  jet mass cut  
1739       between 1-tag and 2-tag. This uncertainty is included since the template for or the  
1740       multijet shape prediction uses the multijet shape from the 1-tag region C. Table 6.6  
1741       shows the acceptance of the large- $R$  jet mass signal and mBB control region selection  
1742       in the multijet 1-tag region C and 2-tag region C yields. The relative difference  
1743       between the acceptance in 1-tag region C and in 2-tag region C is considered as an  
1744       uncertainty on the normalization of the QCD multijet prediction.

Region	Electron	Muon
SR		
1-tag $\frac{N_{\text{SR}}}{N_{\text{Inc}}}$	$31.6 \pm 2.7 \%$	$27.9 \pm 2.7 \%$
2-tag $\frac{N_{\text{SR}}}{N_{\text{Inc}}}$	$37.5 \pm 9.7 \%$	$40.7 \pm 9.7 \%$
Rel. difference between 1-tag and 2-tag	15.6 %	31.5 %
mBBcr		
1-tag $\frac{N_{\text{mBBcr}}}{N_{\text{Inc}}}$	$68.4 \pm 2.7 \%$	$72.1 \pm 2.7 \%$
2-tag $\frac{N_{\text{mBBcr}}}{N_{\text{Inc}}}$	$62.5 \pm 7.0 \%$	$59.3 \pm 7.0 \%$
Rel. difference between 1-tag and 2-tag	9.4 %	21.6 %

TABLE 5.33. The acceptance of the large- $R$  jet mass signal region selection on the multijet 1-tag and 2-tag region C.  $N_{\text{SR}}(N_{\text{Inc}})$  is the multijet yield with (without) the signal region large- $R$  jet mass selection.

1745 **5.5.4.3.3. mBB control region fit** A likelihood fit of the large- $R$  jet mass in the  
 1746 mBB control region is performed with the normalization of the multijet background  
 1747 to be unconstrained in the fit. In the muon channel, the post-fit normalization factor  
 1748 is consistent with unity and in the electron channel, the normalization factor is 0.676  
 1749  $\pm 0.130$ . Due to this significant deviation from unity for the electron channel, we  
 1750 assign a normalization uncertainty of 32.4% for the multijet background in both the  
 1751 mBB control region and signal region.

1752 **5.5.4.3.4.  $t\bar{t}$  and W+jets MC modeling** The uncertainties on the MC modeling  
 1753 of the  $t\bar{t}$  and W+jets, the two largest prompt background predicted by MC in region  
 1754 B,D and C are taken as a systematic on the predicted multijet background in region  
 1755 A since the multijet background is calculated by subtracting the prompt background  
 1756 from observed data. The uncertainties on the normalization of  $t\bar{t}$  and W+jets in each  
 1757 region are calculated by comparing the yields between the nominal  $t\bar{t}$  and W+jets  
 1758 samples with their alternative samples (see Section 5.5.10.2.2 and 5.5.10.2.3).

1759 The uncertainty on the multijet yield prediction in region A is then calculated  
 1760 by recalculating the multijet yield in each region with the  $t\bar{t}$  and W+jets yields be

varied up and down, simultaneously in all regions, by the uncertainty due to the MC modeling of the background. The resulting multijet yield prediction in region A for each background uncertainty is then compared to the nominal prediction in region A and the difference is then taken as the uncertainty on the multijet yield prediction in region A. Table 5.34 shows the uncertainty on the multijet yield prediction in region A signal and mBB control region due to the the uncertainty on the  $t\bar{t}$  and W+jets MC modeling.

	Electron	Muon
SR		
$t\bar{t}$	26.5 %	60.1 %
W+jets	24.7 %	70.4 %
mBBcr		
$t\bar{t}$	37.4 %	101.0%
W+jets	29.5 %	77.6%

TABLE 5.34. The uncertainty on the multijet yield prediction in region A due to the normalization uncertainty of the  $t\bar{t}$  and W+jets backgrounds in region C.

5.5.4.3.5. **Detector modeling of prompt backgrounds** The detector modeling systematic uncertainties on the prompt background in regions B, D and C are propagated through the ABCD method to estimate the uncertainty on the multijet yield prediction region A. Table 5.35 shows the uncertainties on the predicted multijet yield in both lepton channels.

5.5.4.3.6.  **$|d_0^{\text{sig}}|$  cut efficiency modeling** The  $|d_0^{\text{sig}}|$  cut efficiency modeling uncertainty for the prompt MC backgrounds are also taken into account in the ABCD method due to the tighther cut on the leptons'  $|d_0^{\text{sig}}|$ , than the recommended value. The determination of the  $|d_0^{\text{sig}}|$  cut efficiency modeling uncertainty is discussed in Section 5.5.10.1.1. This uncertainty is propagated through the ABCD method

	Electron	Muon
SR		
Total Uncertainty	46.0%	105.6%
mBBcr		
Total Uncertainty	45.5%	127.3 %

TABLE 5.35. The total uncertainty on the multijet yield prediction in region A due to the detector modeling uncertainties of the prompt backgrounds in region B, C and D.

1778 by varying the normalization of the prompt backgrounds in regions B, D and C  
 1779 simultaneously to estimate the uncertainty on the multijet yield prediction region A  
 1780 and is treated as anti-correlated between regions B and regions D,C. Table 5.36 shows  
 1781 the uncertainties on the predicted multijet yield due  $|d_0^{\text{sig}}|$  cut efficiency modeling  
 1782 uncertainty on prompt MC backgrounds in both lepton channels.

	Electron	Muon
SR		
Total Uncertainty	46.4%	50.9%
mBBcr		
Total Uncertainty	42.0%	110.6%

TABLE 5.36. The total uncertainty on the multijet yield prediction in region A due to the  $|d_0^{\text{sig}}|$  cut efficiency modeling uncertainties of the prompt backgrounds in region B, D and C.

1783 **5.5.4.3.7. Total uncertainty on the yield** The total uncertainty on the multijet  
 1784 prediction is calculated as the sum in quadrature of the uncertainties from the two  
 1785 sources explained above. Table 5.37 summarizes the systematic uncertainties on the  
 1786 predicted multijet yield in the electron channel and the muon channel.

Source of uncertainty	Electron	Muon
SR		
Statistical	28.1 %	39.9 %
2-tag/1-tag jet mass acceptance	15.6 %	31.5 %
mBB control region fit	32.3 %	-
$t\bar{t}$ MC modeling	26.5 %	60.1%
W+jets modeling	24.7 %	70.4%
Detector modeling of prompt backgrounds	46.0 %	105.6%
$ d_0^{\text{sig}} $ cut efficiency	46.4 %	50.9%
Total	87.5 %	157.8 %
mBBcr		
Statistical	23.8 %	41.0 %
2-tag/1-tag jet mass acceptance	9.4 %	21.6 %
mBB control region fit	32.3 %	-
$t\bar{t}$ MC modeling	37.4 %	101.0%
W+jets MC modeling	29.5 %	77.6%
Detector modeling of prompt backgrounds	45.5 %	127.3%
$ d_0^{\text{sig}} $ cut efficiency	42.0 %	110.6%
Total	88.3 %	216.4 %

TABLE 5.37. Summary of systematic uncertainties on the QCD multijet yield in the signal region and mBB control region for each lepton channel. The total uncertainty calculated by adding in quadrature the uncertainties from all sources.

1787 **5.5.4.4. Multijet shape uncertainties**

1788 The uncertainty on the  $m_{HH}$  distribution prediction for the QCD multijet  
 1789 background is estimated by comparing the  $m_{HH}$  distribution shape in the 2-tag region  
 1790 C and 1-tag region C (Figure 5.15).

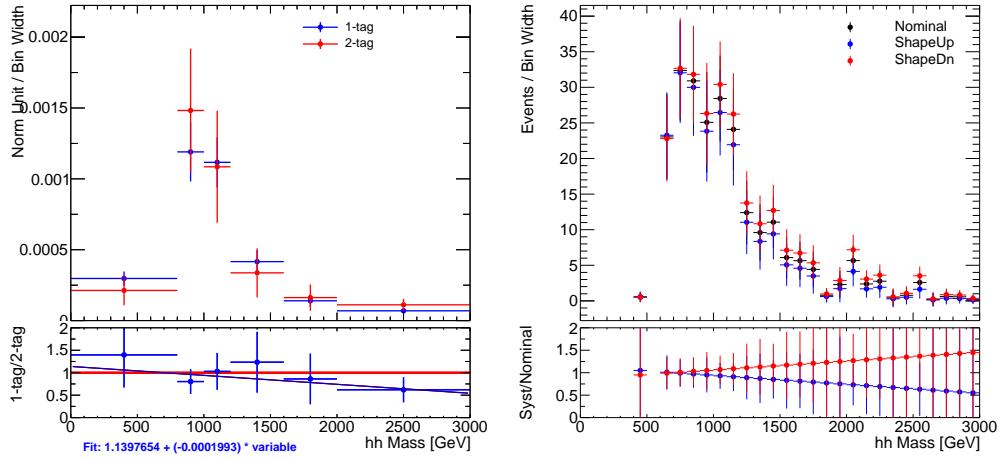


FIGURE 5.15. (a) Shape comparison of the  $m_{HH}$  distribution in region C of the 1-tag and 2-tag region. The linear fit to the ratio is used as the shape systematic of QCD background prediction. (b) The up and down QCD shape systematic variation for the predicted  $m_{HH}$  distribution.

1791 **5.5.4.5. Final prediction and validation**

1792 The predicted multijet yield and its uncertainty in the signal region and mBB  
 1793 control region are shown in Table 5.38.

Region	Electron	Muon	Combined
SR	$165.9 \pm 145.2(87.5\%)$	$69.3 \pm 109.3(157.8\%)$	$235.2 \pm 181.8 (77.3\%)$
mBBcr	$277.1 \pm 244.8(88.3\%)$	$100.8 \pm 218.1(216.4\%)$	$377.9 \pm 327.8 (86.8\%)$

TABLE 5.38. Predicted multijet yield with uncertainties in the signal region (SR) and mBB control region (mBBcr) for each lepton channel.

1794        The ABCD method is validated by assesing the agreement of data and total  
1795        background prediction with the QCD multijet background included in the mBB  
1796        control region. This is dicussed in Section 5.5.9.

1797        **5.5.5.  $t\bar{t}$**

1798        MC simulation is used to model the shape of the  $t\bar{t}$  background. The background  
1799        is predicted to be the largest ( $\sim 52\%$ ) of the total background. A top-enriched control  
1800        region is used to validate the modeling of the  $t\bar{t}$  background.

1801        **5.5.6. V+jets**

1802        MC simulation is used to model the shape and predict the yield of the W+jets  
1803        and Z+jets background. The W+jets background is predicted to be the third largest  
1804        background ( $\sim 18\%$ ) of the total background while the Z+jets background is expected  
1805        to be  $\sim 2\%$  in the signal region (Table 5.52).

1806        **5.5.7. Single Top**

1807        MC simulation is used to model the shape and predict the yield of background  
1808        from single-top processes. This background is predicted to be  $\sim 9\%$  of the total  
1809        background (Table 5.52). The single-top background is predominantly consist of  
1810        Wt production process ( $\sim 90\%$ ), followed by the t-channel production ( $\sim 9\%$ ) and the  
1811        s-channel production ( $\sim 1\%$ ).

1812        **5.5.8. Diboson**

1813        MC simulation is used to model the shape and predict the yield of the diboson  
1814        background. This background is predicted to be  $2\%$  of the total background in the  
1815        signal region (Table 5.52).

1816 **5.5.9. Data/Prediction comparisons in control regions**

1817 To asses the modeling of the background, the total predicted background is asses  
1818 with data in the mBB control region. The electron channel and the muon channel  
1819 are combined into a single channel.

1820 Table 5.39 shows the predicted yield of each background. The data, collected in  
1821 2015+2016, corresponding to the integrated luminosity of  $36.1 \text{ fb}^{-1}$  are used. The  
1822 MC backgrounds  $t\bar{t}$ , W+jets, Single-top, Z+jets and Dibosons are normalized to  
1823 luminosity. The QCD multijet prediction is estimated from the ABCD method.  
1824 Statistical errors are shown for the individual backgrounds and the total predicted  
1825 background while the error due to systematic uncertainties are shown only for the total  
1826 predicted background. Detector modeling uncertainties, MC background modeling  
1827 uncertainties and uncertainties from the ABCD method for QCD multijet background  
1828 are considered. The predicted total background yield has an error of about 27% due  
1829 to systematic uncertainties and the observed yield in data is in good agreement with  
1830 the total background yield.

1831 Figure 5.16, 5.17, 5.18, 5.19, 5.20, 5.21, 5.22 and 5.23 show the distributions  
1832 of kinematic variables for events which fall into the mBB control regions. The  
1833 observed data (black circle) corresponds to an integrated luminosity of  $36.1 \text{ fb}^{-1}$ .  
1834 The MC backgrounds  $t\bar{t}$  (orange), W+jets (blue), Single-top (red), Z+jets (green) and  
1835 Dibosons (yellow) are normalized to cross-section prediction scaled to luminosity of  
1836  $36.1 \text{ fb}^{-1}$ . The QCD multijet background (grey) is predicted from the ABCD method.  
1837 The hashed grey band is the statistical uncertainty on the predicted background  
1838 and the red box band is the statistical+systematics uncertainty on the predicted  
1839 background. The systematics uncertainty on the predicted background consists of the

Sample	Yield	Stats Unc	Systs Unc
$t\bar{t}$	1005.6	$\pm 20.6$	+283.6(+28.2%) -288.8(-28.7%)
W+Jets	565.6	$\pm 10.3$	+277.9(+49.1%) -270.0(-47.7%)
QCD	377.9	$\pm 19.6$	+328.0(+86.8%) -328.0(-86.8%)
Single-top	161.3	$\pm 7.2$	+114.4(+70.9%) -114.4(-70.9%)
Z+Jets	55.9	$\pm 1.6$	+27.7(+49.5%) -27.2(-48.6%)
Dibosons	39.7	$\pm 2.6$	+23.4(+58.9%) -23.3(-58.7%)
Prediction	2206.0	$\pm 31.2$	+593.7(+26.9%) -586.1(-26.6%)
Data	2179	-	-
Data/Pred	0.99	-	-

TABLE 5.39. Predicted and observed yields in the mBB control region. Detector modeling uncertainties, MC background modeling uncertainties and QCD background modeling uncertainties from ABCD method are considered for the systematic uncertainties.

1840 detector modeling systematic uncertainties, MC background modeling uncertainties  
 1841 and uncertainties from the ABCD method for QCD multijet background.

1842 Figure 5.16 is the invariant mass of the reconstructed di-Higgs (hh)system  
 1843 distribution, the  $\cancel{E}_T$ , and  $W \rightarrow l\nu$  system transverse mass distributions and as it  
 1844 can be seen that it is reasonably modelled. The good modeling observed of the  
 1845 distributions gives confidence to the QCD multijet prediction as the events from the  
 1846 background tend to have low values of  $\cancel{E}_T$  and transverse mass.

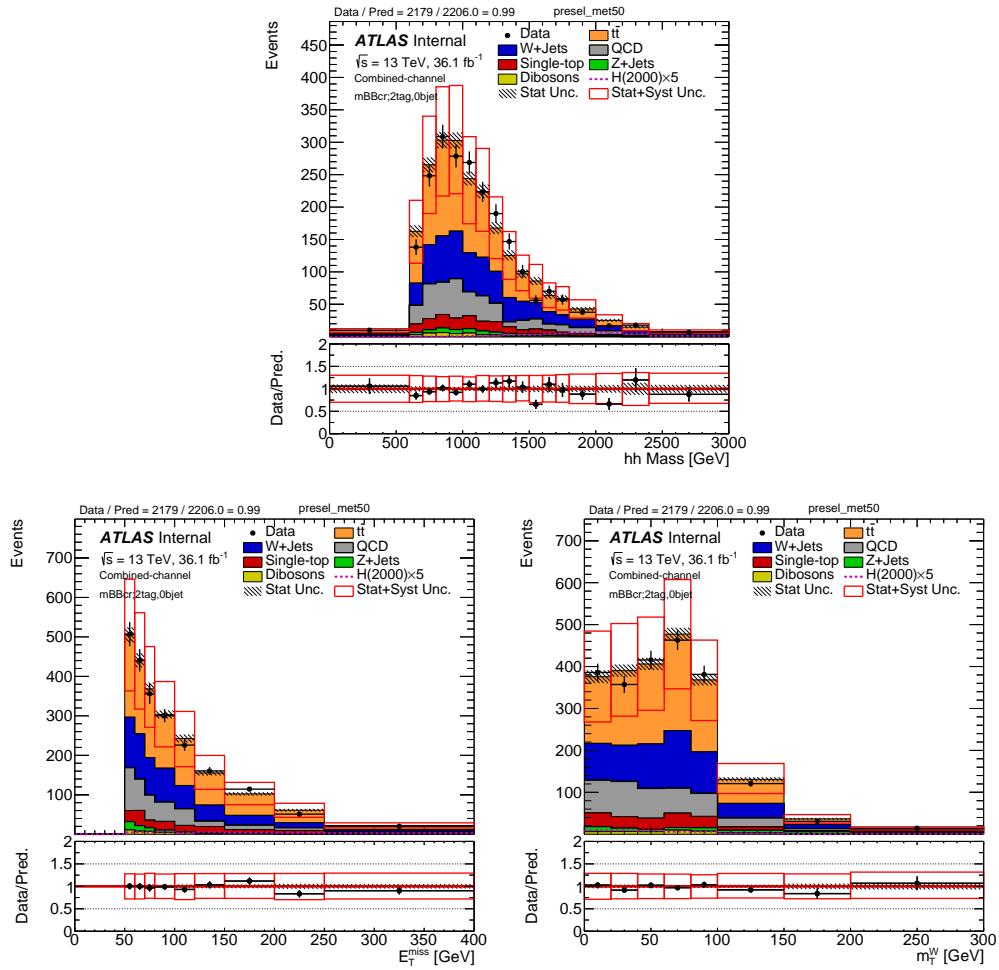


FIGURE 5.16. The invariant mass of the reconstructed di-Higgs (hh) system,  $\cancel{E}_T$  and transverse mass of the  $W \rightarrow l\nu$  system distributions of events in the mBB control region (mBBcr).

### 1847 5.5.10. Systematics

#### 1848 5.5.10.1. Detector modeling uncertainties

1849 The experimental uncertainties considered in the analysis are listed in Table 5.40.

1850 The uncertainties are applicable to signal and background processes that are modelled  
1851 using MC simulation.

Systematic uncertainty	Short description
Event	
ATLAS_LUMI_2015_2016	uncertainty on total integrated luminosity
PRW_DATASF	pile-up reweighting uncertainty
Electrons	
EL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR	trigger efficiency uncertainty
EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	reconstruction efficiency uncertainty
EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	ID efficiency uncertainty
EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	isolation efficiency uncertainty
EG_SCALE_ALL	energy scale uncertainty
EG_RESOLUTION_ALL	energy resolution uncertainty
Muons	
MUON_EFF_TrigStatUncertainty	trigger efficiency uncertainty
MUON_EFF_TrigSystUncertainty	
MUON_EFF_STAT	reconstruction and ID efficiency uncertainty for muons
MUON_EFF_SYS	
MUON_ISO_STAT	isolation efficiency uncertainty
MUON_ISO_SYS	energy scale uncertainty
MUONS_SCALE	energy resolution uncertainty from inner detector
MUONS_ID	energy resolution uncertainty from muon system
MUONS_MS	
Small- $R$ jets	
JET_SR1_JET_GroupedNP_1	
JET_SR1_JET_GroupedNP_2	
JET_SR1_JET_GroupedNP_3	energy scale uncertainties strongly-reduced to 4 components.
JET_SR1_JET_EtaIntercalibration_NonClosure	
JET_JER_SINGLE_NP	energy resolution uncertainty
JET_JvtEfficiency	JVT efficiency uncertainty
Large- $R$ jets	
FATJET_Medium_JET_Comb_Baseline_Kin	
FATJET_Medium_JET_Comb_modeling_Kin	energy scale uncertainties ( $p_T$ and mass scales are fully correlated)
FATJET_Medium_JET_Comb_TotalStat_Kin	
FATJET_Medium_JET_Comb_Tracking_Kin	
FATJET_JER	energy resolution uncertainty
FATJET_JMR	mass resolution uncertainty
Track-jets and Small- $R$ jets	
FT_EFF_Eigen_B	$b$ -tagging efficiency uncertainties (“BTAG_MEDIUM”): 3
FT_EFF_Eigen_C	components for $b$ jets, 4 for $c$ jets and 5 for light jets
FT_EFF_Eigen_L	
FT_EFF_Eigen_extrapolation	$b$ -tagging efficiency uncertainty on the extrapolation to high $p_T$ jets
FT_EFF_Eigen_extrapolation.from_charm	$b$ -tagging efficiency uncertainty on tau jets
$\cancel{E}_T$	
MET_SoftTrk_ResoPara	track-based soft term related longitudinal resolution uncertainty
MET_SoftTrk_ResoPerp	track-based soft term related transverse resolution uncertainty
MET_SoftTrk_Scale	track-based soft term related longitudinal scale uncertainty

TABLE 5.40. Summary of the (nuisance parameter) names and meanings of the detector modeling systematic uncertainties.

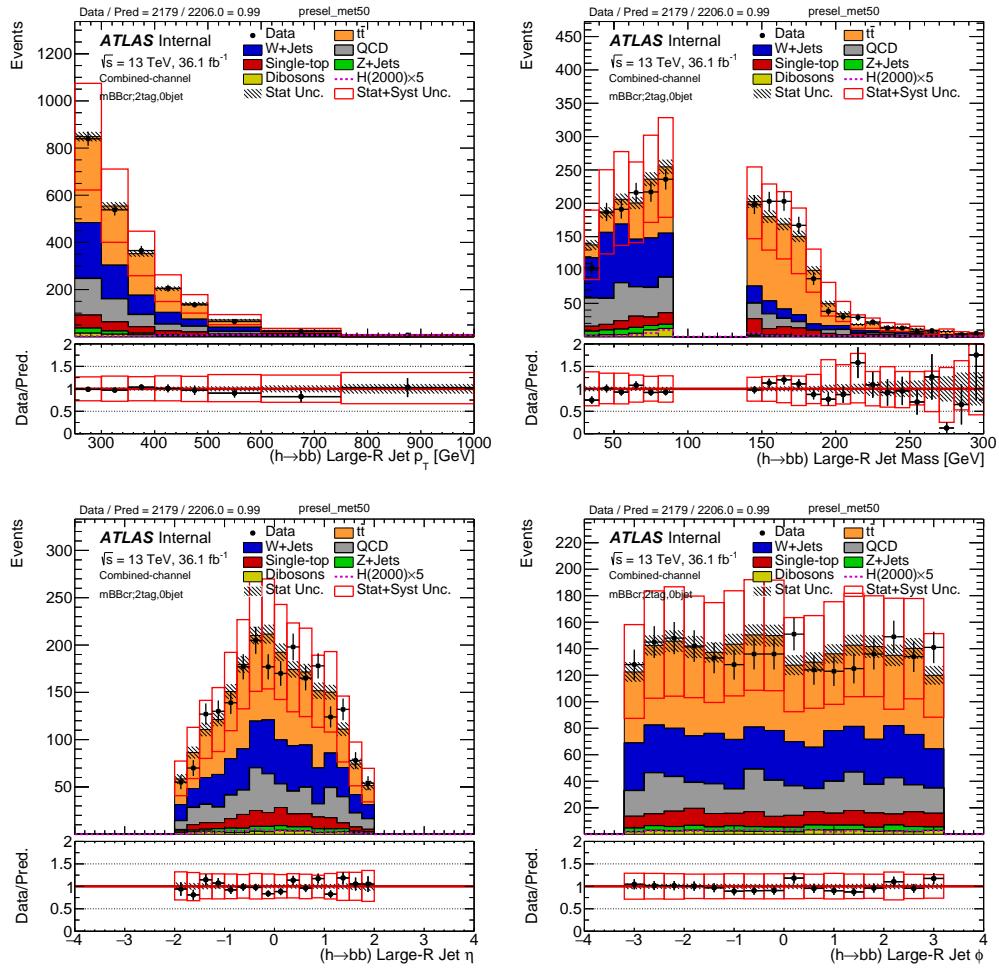


FIGURE 5.17. Kinematic distributions of the reconstructed large- $R$  jet in the mBB control region (mBBcr).

1852    **5.5.10.1.1.  $|d_0^{\text{sig}}|$  cut efficiency modeling**    In this analysis, the  $|d_0^{\text{sig}}|$  cut value  
 1853    used for electrons and muons are not the recommended value by CP groups. The  
 1854    recommended value for electrons is 5 while for muons it is 3. The modeling of the  $|d_0^{\text{sig}}|$   
 1855    significance is assesed in a top-enriched control region. The event reconstruction and  
 1856    selection criteria for the top-enriched control region are exactly the same as outlined  
 1857    in Section 5.5.1 to 5.5.3.1 with the exception that the  $b$ -tagging requirement on the  
 1858    event is different. For this control region, each event is required to have either the  
 1859    leading or sub-leading track-jet to be  $b$ -tagged but not both track-jets to  $b$ -tagged.

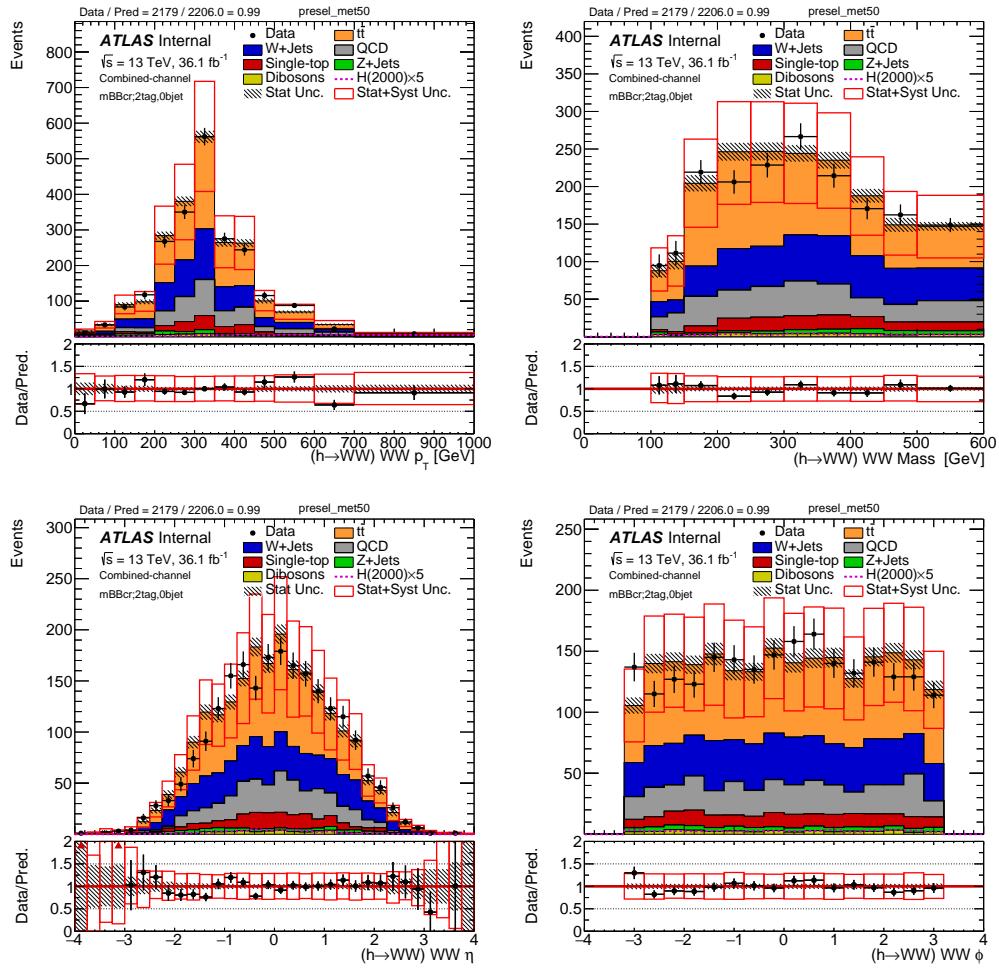


FIGURE 5.18. Kinematic distributions of the reconstructed  $h \rightarrow WW$  system in the mBB control region (mBBcr).

1860 The event is also required to have at least one  $b$ -tagged signal small- $R$  jets, which is in  
 1861 other words the  $b$ -jet veto is reversed. For the purposes of studying the modeling the  
 1862 of the  $|d_0^{\text{sig}}|$  distribution, no  $|d_0^{\text{sig}}|$  requirement is applied on the reconstructed leptons.

1863 Figure 5.24 shows the  $|d_0^{\text{sig}}|$  distribution for the electron and the muon in the top-  
 1864 enriched control region. A clear bias in Data can be observed and this is consistent  
 1865 with other studies through ATLAS. In order to take into account the effect of the  
 1866 mismodeling of the  $d_0^{\text{sig}}$  bias on the tight  $d_0^{\text{sig}}$  cut used in this analysis, the efficiency  
 1867 of the  $|d_0^{\text{sig}}|$  cut is evaluated and compared between Data and MC. Table 5.41 shows

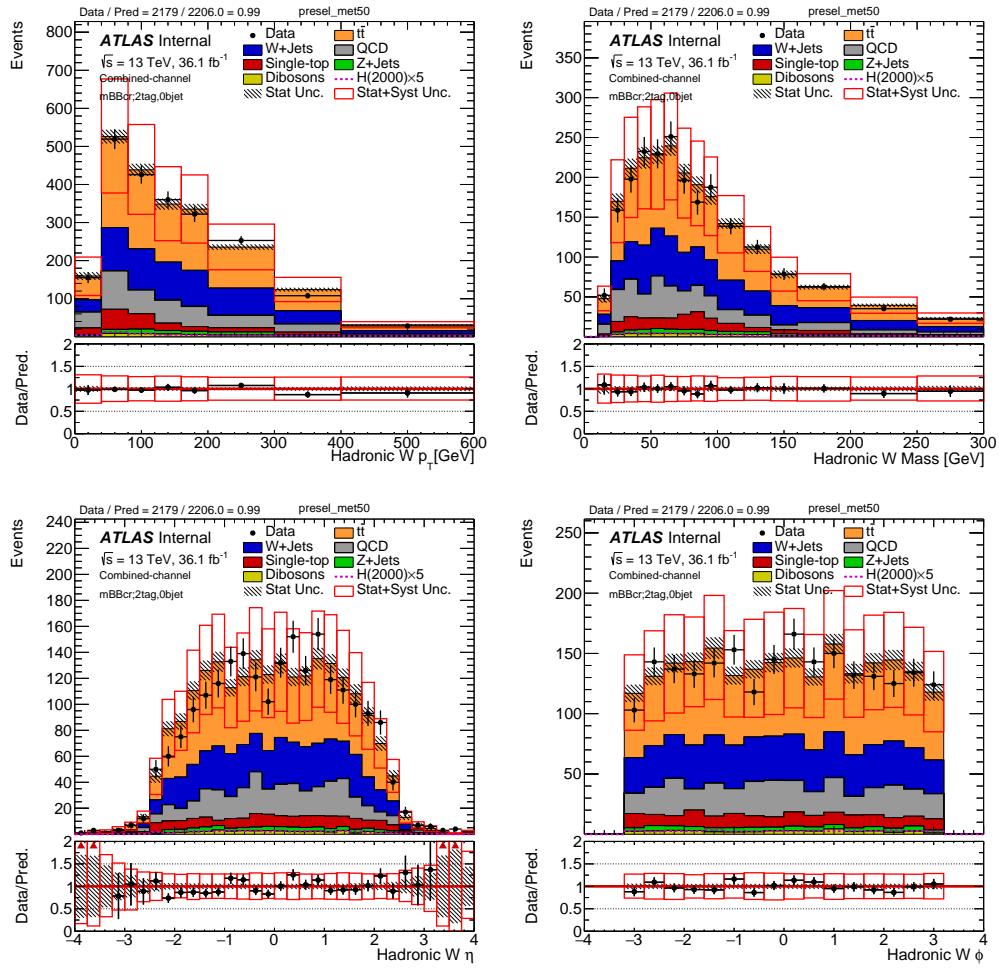


FIGURE 5.19. Kinematic distributions of the reconstructed  $W \rightarrow q\bar{q}$  system in the mBBcr control region (mBBcr).

1868 the efficiency of the  $|d_0^{\text{sig}}|$  cut in Data and MC and the ratio of efficiency between  
 1869 Data and MC in each lepton channel. The relative difference between Data/MC ratio  
 1870 and unity is taken as the uncertainty on the  $|d_0^{\text{sig}}|$  significance cut efficiency and it  
 1871 is assigned for all processes with prompt leptons modelled by MC simulation. The  
 1872 nuisance parameter name associated to this uncertainty is LEP\_d0\_CutEff

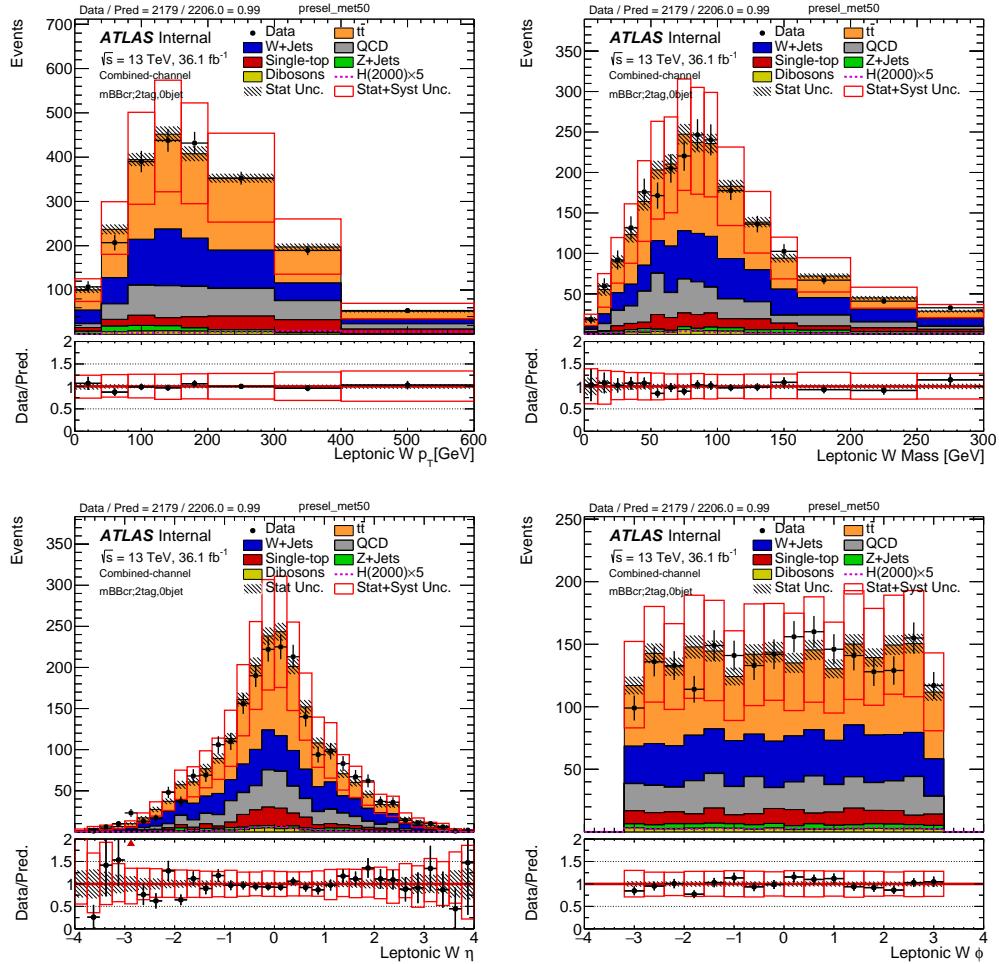


FIGURE 5.20. Kinematic distributions of the reconstructed  $W \rightarrow l\nu$  system in the mBB control region (mBBcr).

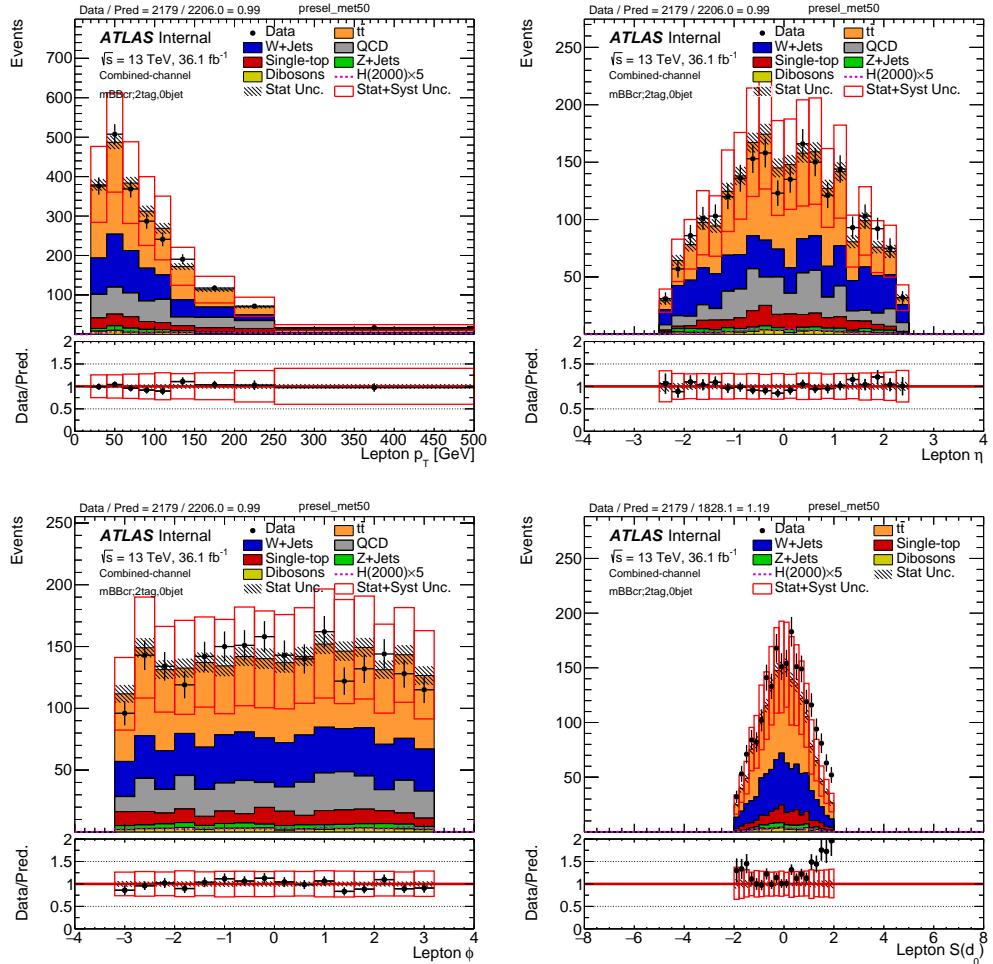


FIGURE 5.21. Kinematic distributions of the selected lepton in the mBB control region (mBBcr).

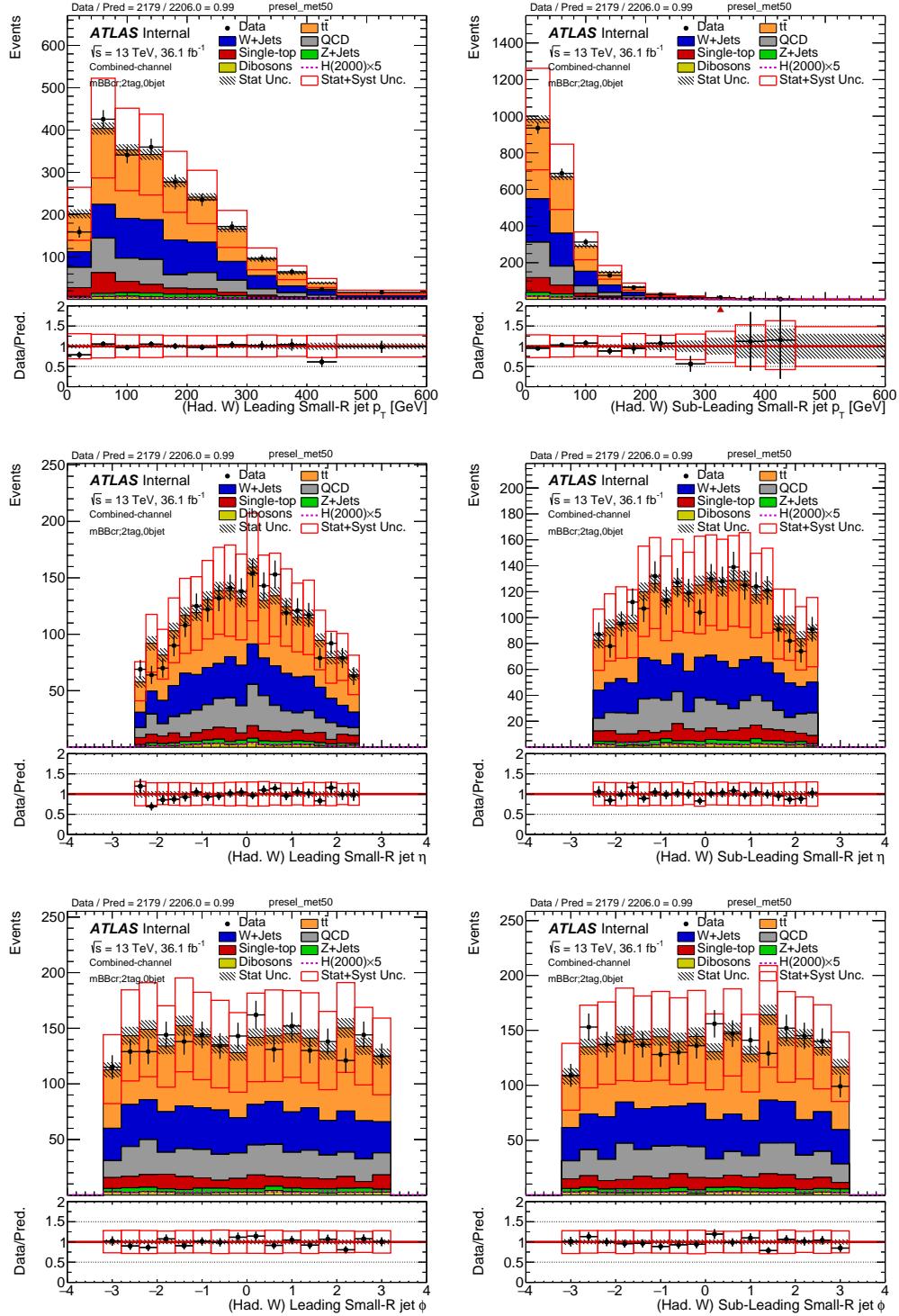


FIGURE 5.22. Kinematic distributions of the leading and sub-leading small- $R$  jets (of the reconstructed hadronic  $W$ ) in the mBB control region (mBBcr).

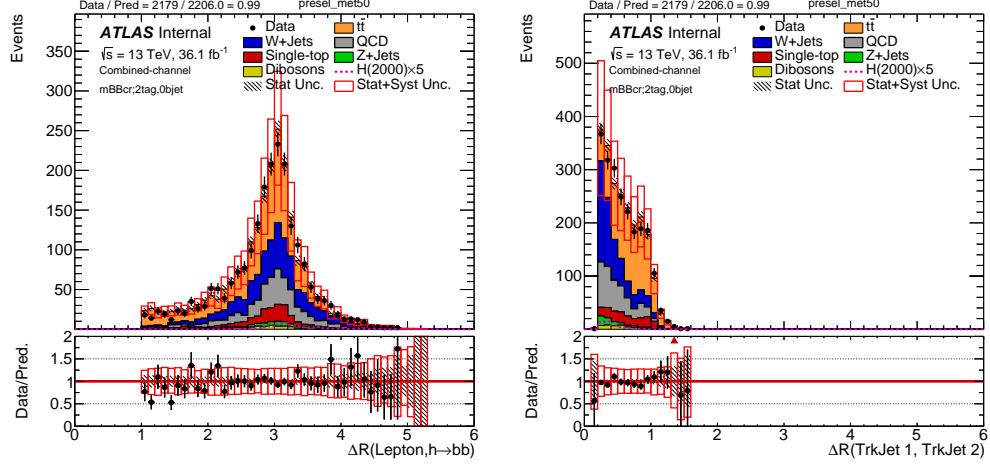


FIGURE 5.23.  $\Delta R$  distribution between the selected lepton and the large- $R$  jet and  $\Delta R$  distribution between the track-jets inside the large- $R$  jet in the mBB control region (mBBcr).

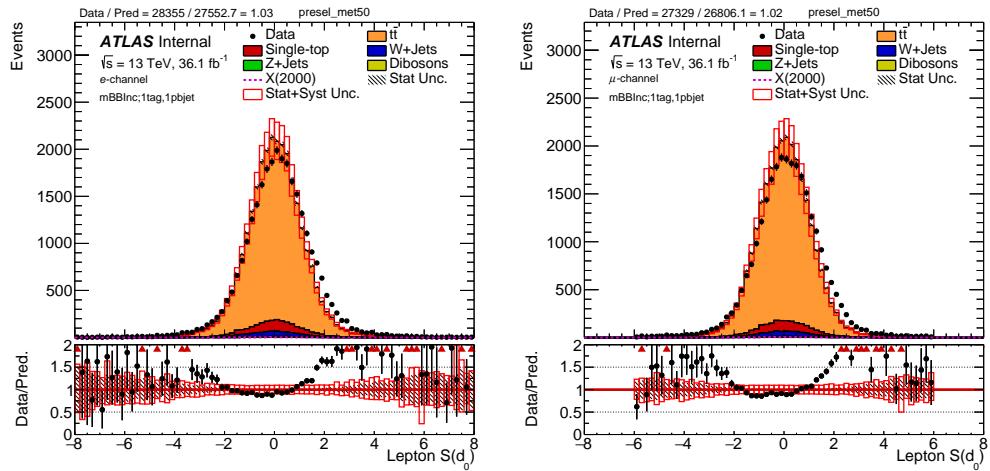


FIGURE 5.24. Electron and muon  $d_0^{\text{sig}}$  distributions in the top control region without the large- $R$  jet mass cut.

	Electron	Muon
$ d_0^{\text{sig}}  \geq 2.0$		
Data	0.884	0.897
MC	0.926	0.937
Data/MC	0.955	0.957
$ d_0^{\text{sig}}  < 2.0$		
Data	0.115	0.104
MC	0.074	0.063
Data/MC	1.572	1.640

TABLE 5.41. Efficiency of the  $|d_0^{\text{sig}}|$  cut for electrons and muons in Data and MC. The Data/MC ratio is also calculated and the difference between the ratio and unity is taken as the systematic uncertainty on the  $|d_0^{\text{sig}}|$  efficiency modeling for events with leptons that pass ( $\geq 2.0$ ) or fail ( $< 2.0$ ) the  $|d_0^{\text{sig}}|$  cut.

1873 **5.5.10.2. Background and Signal modeling Systematics**

1874 **5.5.10.2.1. Methodology** Uncertainties in the  $m_{hh}$  distributions are assigned  
1875 to the dominant backgrounds,  $t\bar{t}$  and V+Jets and single top by comparing the  
1876 nominal MC samples to a number of alternative MC generators at the reconstruction  
1877 level. The comparisons are performed in with the same event selection in Sect 5.5.  
1878 Each uncertainty contains 2 components, a shape systematic and a normalization  
1879 systematic due to acceptance.

1880 The shape systematic corresponds to a reweighting function derived by fitting a  
1881 1<sup>st</sup> order polynomial to the ratio of  $m_{hh}$  distribution of the variation sample over the  
1882 nominal sample. The  $m_{hh}$  distribution for the the variation sample is normalized to  
1883 the same number of events of the norminal sample.

1884 **5.5.10.2.2. Top-quark processes ( $t\bar{t}$  & single top)** Four alternative MC  $t\bar{t}$   
1885 samples are used to assess 3 aspects of the MC modeling, whilst five alternative MC  
1886 samples are used to assess 4 aspects of the MC modeling for single top quark pair  
1887 production. The alternative samples considered are:

- 1888 – **Powheg +Herwig++:** The ME POWHEG generator uses the same setup as  
1889 that used for the nominal POWHEG +PYTHIA 6 configuration, but the parton  
1890 shower (PS) generator is swapped out for Herwig++ version 2.7.1 using the UE-  
1891 EE-5 tune and CTEQ6L1 PDF set. The purpose therefore of this comparison  
1892 is to test the PS, hadronisation, underlying event (UE) and Multiple Parton  
1893 Interation (MPI) models whilst maintaining the same hard scattering model  
1894 given by POWHEG.
- 1895 – **aMC@NLO+Herwig++:** The ME generator is swapped out for aMC@NLO  
1896 using the CT10 PDF set, interfaced with Herwig++ using the CTEQ6L1-UE-

1897       EE-5 tune and CTEQ6LI PDF set. This sample is compared to the previous  
1898       POWHEG +Herwig++ sample. This fixes the PS generator component, but  
1899       alters the hard scattering generator, making this variation sensitive to the hard  
1900       scatter model.

1901       – **Powheg +Pythia 6 Radhi/RadLo:** Using the same setup as that used  
1902       for the nominal POWHEG +PYTHIA 6 sample, the RadHi and RadLo samples  
1903       correspond to either the enhancement (high) or reduction (low) of initial/final  
1904       state radiation (IFSR). The two samples are compared to the nominal sample  
1905       setup, and so are sensitive to variations of IFSR models.

1906           \* RadHi: The renormalisation ( $\mu_R$ ) and factorisation scale ( $\mu_F$ ) scales are  
1907           decreased by a factor of 0.5, the POWHEG `hdamp` parameter is doubled  
1908           ( $2 \times m_{top}$ ), and the high radiation PERUGIA2012 tune is used.

1909           \* RadLo: The renormalisation ( $\mu_R$ ) and factorisation scale ( $\mu_F$ ) scales are  
1910           increased by a factor of 2, the POWHEG `hdamp` parameter is kept at  $m_{top}$ ,  
1911           and the low radiation PERUGIA2012 tune is used.

1912       – **Powheg +Pythia 6 Diagram Subtraction:** For the production of a  
1913       single top quark in association with a W-boson ( $Wt$ ) the interference with  
1914       the  $t\bar{t}$  production process at NLO in QCD is removed by subtracting the  
1915       cross-section associated with the  $t\bar{t}$  double resonance amplitude terms, rather  
1916       than subtracting the same terms from the amplitude prior to the calculation  
1917       (Diagram Removal).

1918       Table 5.42 shows the estimated uncertainty on the normalization of the  $t\bar{t}$   
1919       background in the signal region from the comparison of the nominal  $t\bar{t}$  sample to  
1920       the alternative samples. The largest uncertainty comes the RadLo variation, which

1921 is about  $\sim 8.4\%$  with similar level of uncertainties from alternative ME generator  
1922 choice and alternative PS generator choice. The normalization of  $t\bar{t}$  background is  
1923 assinged with a single nuisance parameter with the total uncertainty set as the prior  
1924 uncertainty.

1925 Shape comparisons of the  $m_{hh}$  distribution between the nominal  $t\bar{t}$  sample to the  
1926 alternative samples were made and they are shown in Figure 5.25.

Variation	Uncertainty (%)
RadHi	1.4
RadLo	8.4
aMC@NLO	7.1
Herwig++	7.8
PDF	1.9
Scale	5.0
Total	13.5

TABLE 5.42. The normalization uncertainty for the  $t\bar{t}$  background in the signal region from different sources. The total uncertainty is calculated as the sum of quadrature from all the sources.

1927 For the single-top background, only the uncertainties on the modeling of the Wt  
1928 production process are considered since it is the dominant single-top process in the  
1929 signal region. Table 5.43 shows the estimated uncertainty on the normalization of  
1930 the single-top background. The biggest uncertainties comes from the PS generator  
1931 choice and comparisons to the DS sample. The total uncertainty is abnormally large,  
1932 which is larger than 100%.

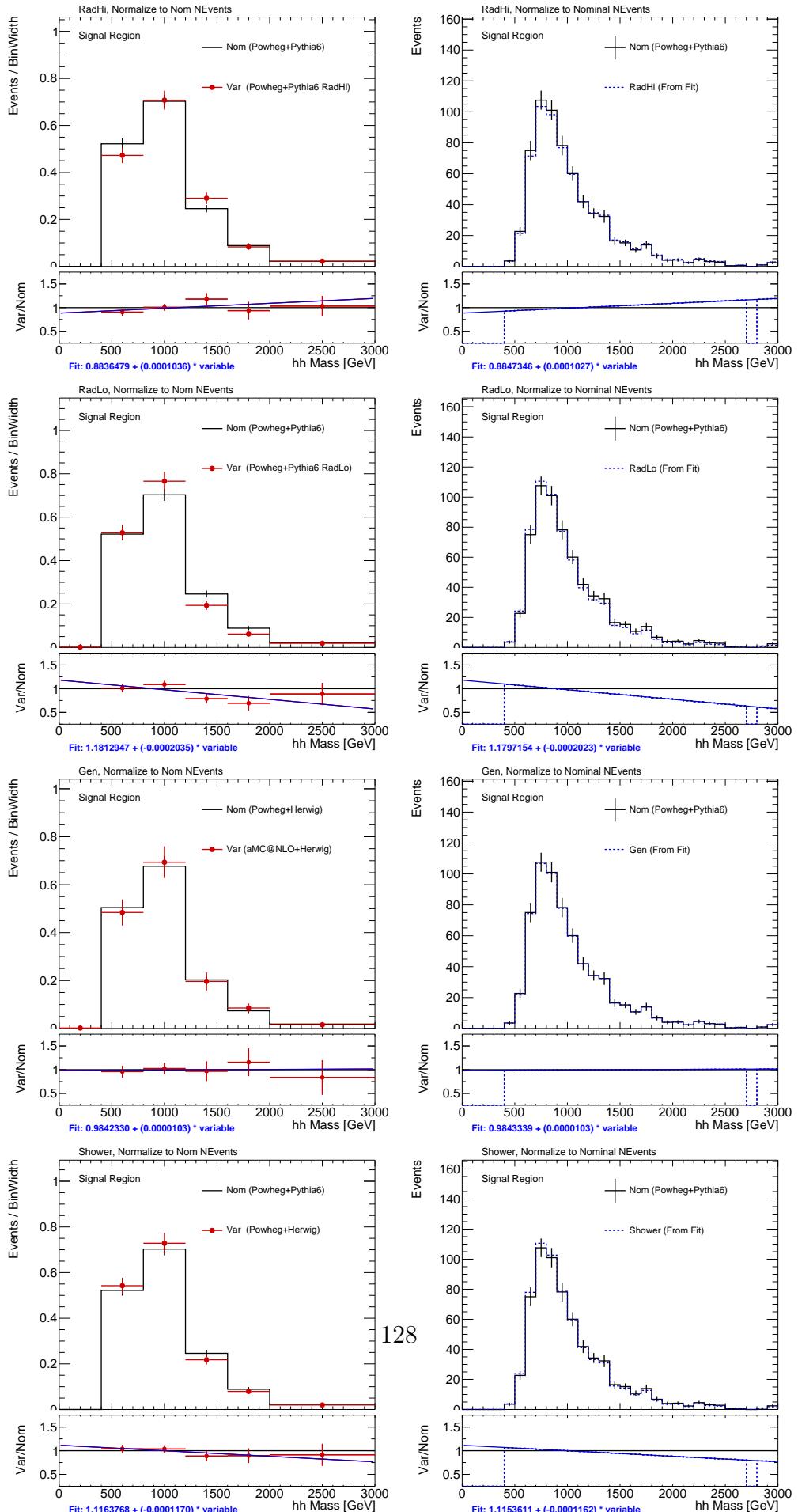


FIGURE 5.25.  $m_{hh}$  distribution shape comparison between nominal  $t\bar{t}$  sample and

Variation	Uncertainty (%)
RadHi	15.1
RadLo	19.0
Herwig++	33.5
DR	72.5
aMC@NLO	25.4
Total	85.9

TABLE 5.43. The normalization uncertainty for the  $t\bar{t}$  background in the signal region from different sources. The total uncertainty is calculated as the sum of quadrature from all the sources.

1933 **5.5.10.2.3. V+Jets processes** The nominal V+Jets prediction, uses the ME+PS  
1934 generator SHERPA 2.2.1 interfaced with the NNPDF 3.0 NNLO PDF set. This default  
1935 configuration provides a prediction for vector boson production plus associated jets  
1936 at NLO accuracy at the ME level for up to 2 extra partons, and LO accuracy for 3  
1937 and 4 extra partons in QCD. The merging of additional parton multiplicities arising  
1938 from the internal SHERPA PS, is regulated by the MEPS@NLO merging technique.

1939 The alternative samples used to assess the modeling uncertainties are:

1940 – **MadGraph5+Pythia 8.186** : The LO ME generator MadGraph5 using the  
1941 NNPDF3.0(2.3) NLO(LO) PDF set interfaced with PYTHIA 8 version 8.186  
1942 using the A14 tune, offers a LO+NLL accurate prediction for vector boson  
1943 production in association with jets for up to 4 extra partons from the ME  
1944 and 4+ parton from PYTHIA 8 at NLL accuracy. The comparison between the  
1945 nominal SHERPA 2.2.1 sample with this sample convolves the ME and PS model  
1946 variation. Due to the unavailability of this sample at reconstruction level, the  
1947 comparison is made at (particle) truth level.

1948 – **Sherpa 2.2.1 scale variations:** Configured in the same manner as the nominal  
1949 V+Jets sample, the renormalisation  $\mu_R$  and resummation  $\mu_F$  scales are varied  
1950 up/down by a factor of 2.

1951 – **Sherpa 2.2.1 PDF variations:** Configured in the same manner as the nominal  
1952 V+Jets sample. The 100 NNPDF3.0NNLO replicas variations are available.  
1953 The central values of 2 alternative PDF sets, MMHT2014NNLO 68%CL and  
1954 CT14NNLO are also available.

1955 – **Sherpa 2.2.1  $\alpha_s(PDF)$  variations:** Configured in the same manner as the  
1956 nominal V+Jets sample, the  $\alpha_s$  value used by the nominal NNPDF 3.0 NNLO  
1957 PDF is varied up and down according to a variation of the  $\mu_R$  scale by a factor  
1958 of 2.

1959 Table 5.44 shows the estimated uncertainty on the normalization of the W+jets  
1960 background in the signal region from the comparison of the nominal W+jets sample  
1961 to the alternative samples. The largest uncertainty comes from the renormalization  
1962 and resummation scale, which is about ~42% and dominates the total uncertainty on  
1963 W+jets background. The normalization of W+jets background is assinged with a  
1964 single nuisance parameter with the total uncertainty set as the prior uncertainty.

1965 Shape comparisons of the  $m_{hh}$  distribution between the nominal W+jets sample  
1966 to the alternative samples were made and only one variation was found to have  
1967 noticeable difference: the scale variation of  $\mu_R=0.5$ ,  $\mu_F=0.5$  as in Figure 5.26.

1968 Table 5.45 shows the estimated uncertainty on the normalization of the Z+jets  
1969 background in the signal region from the comparison of the nominal Z+jets sample  
1970 to the alternative samples. The largest uncertainty comes from the renormalization  
1971 and resummation scale, which is about ~48% and dominates the total uncertainty  
1972 on Z+jets background. The normalization of Z+jets background is assinged with a

Variation	Uncertainty (%)
Scale	41.9
$\alpha_S$ (PDF)	8.4
PDF alternative set	1.6
NNPDF replicas	5.6
Madgraph+Pythia8	11.0
Total	44.5

TABLE 5.44. The normalization uncertainty for the W+jets background in the signal region from different sources. The total uncertainty is calculated as the sum of quadrature from all the sources.

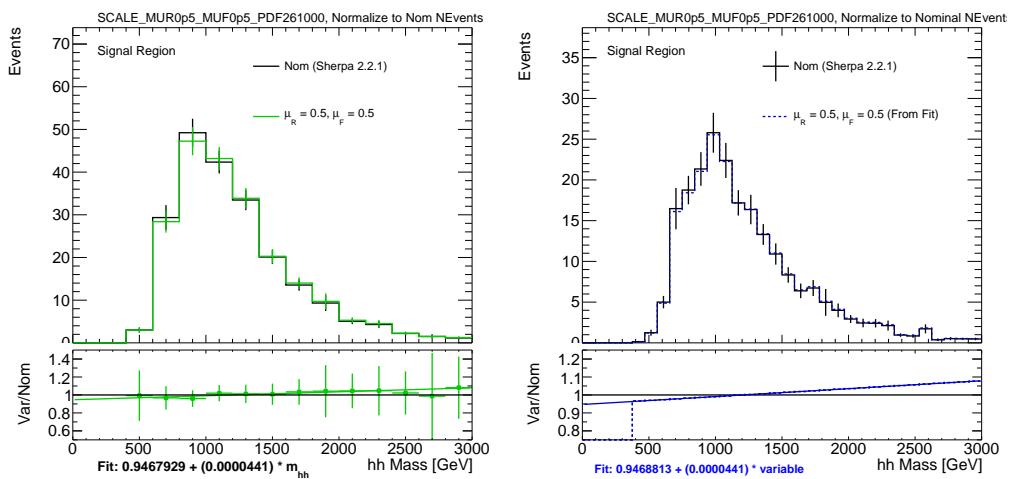


FIGURE 5.26.  $m_{hh}$  distribution shape comparison between nominal W+jets sample and scale variation ( $\mu_R=0.5$ ,  $\mu_F=0.5$ ) sample. Plot on the left is a direct comparison between the nominal and variation sample while on the right, the variation comes from the reweighted function applied to the nominal W+jets sample. The linear fit in the ratio of the left plot is used as the reweighted function.

1973 single nuisance parameter with the total uncertainty set as the prior uncertainty. The  
1974 uncertainty on the  $m_{hh}$  shape is found to be negligible and therefore ignored.

Variation	Uncertainty (%)
Scale	48.3
$\alpha_S$ (PDF)	1.6
PDF alternative set	2.7
NNPDF replicas	1.4
Total	48.4

TABLE 5.45. The normalization uncertainty for the Z+jets background in the signal region from different sources. The total uncertainty is calculated as the sum of quadrature from all the sources.

1975 **5.5.10.2.4. Diboson processes** The systematic uncertainty on the normalization  
1976 of the Diboson background is assigned to be 40%. This uncertainty is taken from the  
1977 resolved analysis. As the background is small, the uncertainty is considered to be  
1978 conservative.

1979 **5.5.10.2.5. production** Systematic uncertainties on the acceptance of signal  
1980 processes are computed by generating alternative variation signal samples and then  
1981 compare their acceptance with respect to the nominal signal samples. The difference  
1982 sources of uncertainty considered are:

1983 – **Scale variations:** Configured in the same manner as the nominal signal  
1984 samples but the renormalisation and resummation scales are varied up/down  
1985 by a factor of 2.

1986 – **Parton shower choice:** Configured in the same manner as the nominal signal  
1987 samples but with Pythia 8 chosen as the shower generator instead of Herwig++.

1988 Table 5.46 lists the systematic uncertainties for 4 different scalar signal sample  
1989 mass points.

Variation	Xhh1000	Xhh1500	Xhh2000	Xhh2500	Xhh3000
Scale	0.2	0.2	0.4	0.4	0.4
PDF	0.4	0.2	0.4	0.2	0.1
Shower	0.4	0.8	1.6	3.4	4.1

TABLE 5.46. Theoretical uncertainties (in percentage) on the acceptance of several signal mass points.

1990 **5.5.10.3. QCD multijet modeling**

1991 Systematic uncertainties related to the modeling of the multijet background were  
1992 discussed in Section 6.3.4.3 for the predicted yield and in Section 6.3.4.3 for the  
1993 predicted  $m_{hh}$  distribution.

1994 **5.6. Results**

1995 **5.6.1. Resolved Analysis**

1996 The resolved analysis is described in detail in Section 5.4. The event selection  
1997 is described in Section 5.4.1 and summarised in Table 5.8. For each selected event,  
1998 the invariant mass of the  $HH$  system ( $m_{HH}$ ) is reconstructed and its distribution  
1999 is shown in Figure 5.27 for the *non-res* and the *m500* analyses, and in Figure 5.28  
2000 for the *low-mass* and the *high-mass* analyses. Data are generally in good agreement  
2001 with the expected background predictions within the total uncertainty. The signal  
2002  $m_{HH}$  distribution is shown in the figure for the non-resonant and the scalar  
2003 resonance. Because the scalar-resonance samples are simulated in the narrow-width  
2004 approximation, the reconstructed resonance width is exclusively due to the detector  
2005 resolution.

2006 The  $m_{HH}$  distribution is sampled with resonance-mass-dependent  $m_{HH}$   
2007 requirements as reported in Table 5.9. The numbers of events in the signal and

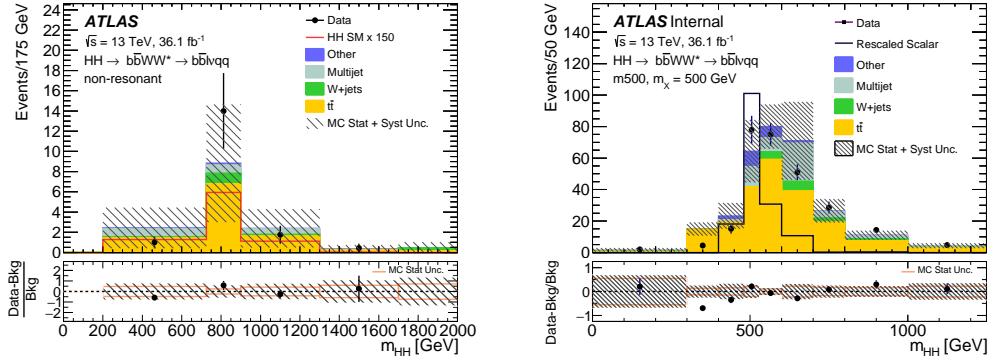


FIGURE 5.27.  $m_{HH}$  distributions for non-resonant and  $m_{500}$  selections in the resolved analysis. For each selection the corresponding signal hypothesis, non-resonant and scalar resonance, is shown. For the scalar signal, resonances with mass 500 GeV are shown. The lower panel shows the fractional difference between data and the total expected background with the corresponding statistical and total uncertainty. The non-resonant signal is multiplied by a factor of 150 with respect to the expected SM cross section. The scalar signal is multiplied by a factor of 5 with respect to the expected upper-limit cross section.)

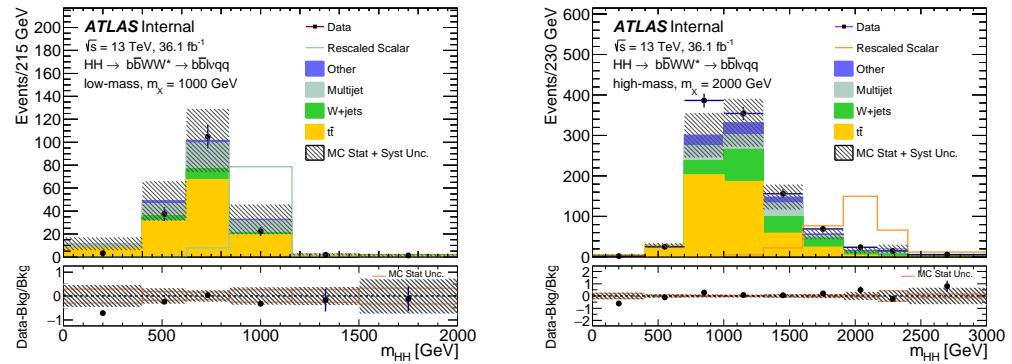


FIGURE 5.28.  $m_{HH}$  distributions in the resolved analysis selections. For each selection the corresponding signal hypothesis and mass 1000 (2000) GeV for the *low-mass* (*high-mass*) analysis, are shown. The lower panel shows the fractional difference between data and the total expected background with the corresponding statistical and total uncertainty. In the plot the scalar signal is multiplied by a factor of 8 with respect to the expected upper-limit cross section; for the plot on the right the multiplying factor is 20 for the scalar signal.

control regions (the  $t\bar{t}$  control region and the C region of the multijet estimation procedure) are simultaneously fit using a maximum-likelihood approach. The fit includes six contributions: signal,  $W$ +jets,  $Z$ +jets,  $t\bar{t}$ , single-top-quark production, diboson and multijet. The  $t\bar{t}$  and multijet normalisations are free to float, the C region of the ABCD method being directly used in the fit, while the diboson,  $W$ +jets and  $Z$ +jets backgrounds are constrained to the expected SM cross sections within their uncertainties.

The fit is performed after combining the electron and muon channels. Statistical uncertainties due to the limited sample sizes of the simulated background processes are taken into account in the fit by means of nuisance parameters, which are parameterised by Poisson priors. Systematic uncertainties are taken into account as nuisance parameters with Gaussian constraints. For each source of systematic uncertainty, the correlations across bins and between different kinematic regions, as well as those between signal and background, are taken into account. Table 5.47 shows the post-fit number of predicted backgrounds, observed data, and the signal events normalised to the expected upper limit cross sections. Expected event yields vary across mass because of varying selections. For instance, the requirement on  $p_T^{b\bar{b}}$  is higher in *non-res* selection than in *low-mass* selection. Similarly, even within *low-mass* or *high-mass* selection, the requirement on  $m_{HH}$  vary across mass.

No significant excess over the expectation is observed and the results are used to evaluate an upper limit at the 95% confidence level (CL) on the production cross section times the branching fraction for the signal hypotheses under consideration. The exclusion limits are calculated with a modified frequentist method [83], also known as  $CL_s$ , and the profile-likelihood test statistic [84]. None of the considered systematic uncertainties is significantly constrained or pulled in the likelihood fit.

Resonant analysis			
$m_X$ [GeV]	$S$	Total Bkg.	Data
500	$18 \pm 5$	$19 \pm 6$	26
600	$13 \pm 2$	$17 \pm 6$	16
700	$16 \pm 2$	$25 \pm 8$	22
750	$20 \pm 2$	$22 \pm 9$	27
800	$18.4 \pm 1.5$	$20 \pm 8$	28
900	$16.3 \pm 1.6$	$20 \pm 7$	23
1000	$12.0 \pm 1.3$	$14 \pm 5$	11
1100	$9.6 \pm 1.2$	$8 \pm 3$	8
1200	$8.1 \pm 0.9$	$6 \pm 3$	5
1300	$5.1 \pm 0.7$	$3.5 \pm 1.8$	1
1400	$4.3 \pm 0.3$	$1.1 \pm 0.2$	0
1500	$3.5 \pm 0.3$	$1.1 \pm 0.2$	0
1600	$3.1 \pm 0.3$	$0.4 \pm 0.3$	1
1800	$14.1 \pm 1.8$	$17 \pm 5$	21
2000	$8.7 \pm 1.0$	$8 \pm 3$	9
2250	$7.9 \pm 1.1$	$6 \pm 2$	7
2500	$5.5 \pm 0.8$	$3.3 \pm 1.4$	3
2750	$5.7 \pm 1.0$	$3.1 \pm 1.3$	3
3000	$4.3 \pm 0.7$	$2.1 \pm 1.0$	1

Non-resonant analysis			
Rescaled SM signal	Total Bkg.	Data	
$17 \pm 2$	$21 \pm 8$	22	

TABLE 5.47. Data event yields, and post-fit signal and background event yields in the final signal region for the non-resonant analysis and the resonant analysis in the 500–3000 GeV mass range. The errors shown are the MC statistical and systematic uncertainties described in Section 5.4.9. The yields are shown for a scalar ( $S$ ) signal model. Signal event yields are normalized to the expected upper-limit cross section.

In the non-resonant signal hypothesis the observed (expected) upper limit on the  $\sigma(pp \rightarrow HH) \times \mathcal{B}(HH \rightarrow b\bar{b}WW^*)$  at 95% CL is:

$$\sigma(pp \rightarrow HH) \cdot \mathcal{B}(HH \rightarrow b\bar{b}WW^*) < 2.5 (2.5^{+1.0}_{-0.7}) \text{ pb.}$$

The branching fraction  $\mathcal{B}(HH \rightarrow b\bar{b}WW^*) = 2 \times \mathcal{B}(H \rightarrow b\bar{b}) \times \mathcal{B}(H \rightarrow WW^*) = 0.248$  is used to obtain the following observed (expected) limit on the  $HH$  production cross section at 95% CL:

$$\sigma(pp \rightarrow HH) < 10 (10^{+4}_{-3}) \text{ pb},$$

which corresponds to  $300 (300^{+100}_{-80})$  times the SM predicted cross section.  
 Including only the statistical uncertainty, the expected upper limit for the non-resonant production is 190 times the SM prediction. This result, when compared with other  $HH$  decay channels, is not competitive. This is mainly due to the similarity of the reconstructed  $m_{HH}$  spectrum between the non-resonant SM signal and the  $t\bar{t}$  background that makes the separation between the two processes difficult.

Figure 5.29 shows the expected and observed limit curves for the production cross section of a scalar  $S$  particle. Different selections are used in different resonance mass ranges without attempting to statistically combine them. The switch from one selection to another is performed based on the best expected limit for that resonance mass. The outcome of this procedure is that the *m500* selection is used to set limits on resonances of mass of 500 GeV, the *low-mass* selection is used up to masses of 1600 GeV, while the *high-mass* selection is used in the mass range 1600-3000 GeV.

Overall, the resolved analysis is most sensitive for a mass value of 1300 GeV with an expected upper limit of 0.35 pb on  $\sigma(pp \rightarrow HH)$ . At this mass the observed exclusion limit is 0.2 pb. In both the non-resonant and resonant cases, the impact of the systematic uncertainties is observed to be large. In order to quantify the impact

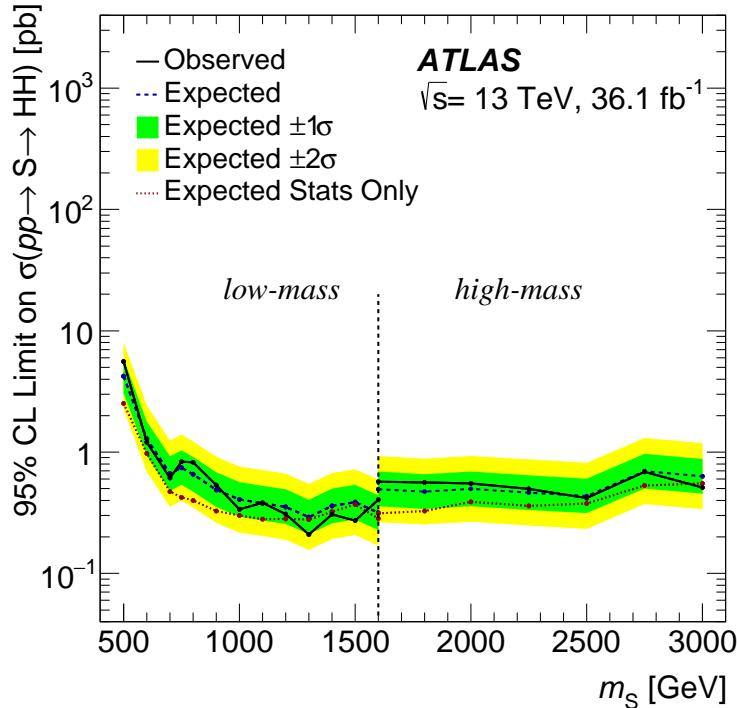


FIGURE 5.29. Expected and observed upper limit at 95% CL on the cross section of resonant pair production for the resolved analysis in the heavy scalar boson  $S$  model. The plot also shows the expected limit without including the systematic errors in order to show their impact.

of the systematic uncertainties, a fit is performed where the estimated signal yield,  
 normalized to an arbitrary cross-section value, is multiplied by a scaling factor  $\alpha_{\text{sig}}$ ,  
 which is treated as the parameter of interest in the fit. The fit is performed using  
 pseudo-data and the contribution to the uncertainty in  $\alpha_{\text{sig}}$  from several sources is  
 determined. The contribution of the statistical uncertainty to the total uncertainty  
 in  $\alpha_{\text{sig}}$ , shown in Table 5.48, is decomposed into signal region statistics, top CR  
 statistics and multijet CR statistics. The contribution of the systematic uncertainties  
 to the total uncertainty is decomposed into the dominant components and shown  
 in Table 5.49. The dominant systematic uncertainties vary across the mass range,

Statistical source	Resolved analysis			
	<i>Non-Res (%)</i>	500 GeV (%)	1000 GeV (%)	2000 GeV (%)
Signal region	+60/-40	+60/-60	+70/-60	+80/-70
Top control region	+40/-30	+28/-30	+20/-12	+13/-13
Multijet control region	+40/-30	+24/-26	+30/-30	+30/-30
Total statistical	+80/-60	+70/-70	+80/-70	+90/-80

TABLE 5.48. Statistical contribution (in percentage) to the total error in the scaling factor  $\alpha_{\text{sig}}$  for the non-resonant signal and three scalar-signal mass hypotheses, 500 GeV, 1000 GeV and 2000 GeV, in the resolved analysis. The values are extracted by calculating the difference in quadrature between the total statistical error and the error obtained after setting constant the normalisation factor of the background that dominates the region of interest.

2059 but some of the most relevant ones are due to  $t\bar{t}$  modelling,  $b$ -tagging systematic  
2060 uncertainties, and those related to jet measurements.

Systematic source	Resolved analysis			
	Non-Res (%)	500 GeV (%)	1000 GeV (%)	2000 GeV (%)
$t\bar{t}$ modelling ISR/FSR	+30/-20	+10/-5	+7 / -4	+2/-2
Multijet uncertainty	+10/-10	+20/-10	+20 / -20	+30/-30
$t\bar{t}$ Matrix Element	+10/-10	—	—	—
$W+jets$ modelling PDF	+4/-7	+10/-10	+2 / -6	+7/-5
$W+jets$ modelling scale	+9/-10	+9/-4	+9 / -2	+20/-10
$W+jets$ modelling gen.	+10/-8	+10/-10	+9 / -1	+9/-9
$t\bar{t}$ modelling PS	+3/-2	+30/-20	+20 / -20	+2/-2
$b$ tagging	+30/-20	+11/-5	+7 / -6	+30/-30
JES/JER	+13/-20	+20/-20	+50 / -50	+10/-6
$E_T^{\text{miss}}$ soft term res.	+20/-20	+8/-1	+9 / -7	+7/-7
Pile-up reweighting	+3/-10	+5/-3	+9 / -10	+6/-6
Total systematic	+60/-80	+70/-70	+60/-70	+40/-60

TABLE 5.49. Systematic contributions (in percentage) to the total error in the scaling factor  $\alpha_{\text{sig}}$  for the non-resonant signal and three scalar-signal mass hypotheses, 500 GeV, 1000 GeV and 2000 GeV, in the resolved analysis. The first column quotes the source of the systematic uncertainty. The “–” symbol indicates that the specified source is negligible. The contribution is obtained by calculating the difference in quadrature between the total error in  $\alpha_{\text{sig}}$  and that obtained by setting constant the nuisance parameter(s) relative to the contribution(s) under study.

TABLE 5.50. Data event yields, and post-fit signal and background event yields in the final signal region for the boosted analysis and the scalar  $S$  particle hypothesis. The errors shown are the MC statistical and systematic uncertainties described in Section 5.5.10. For illustration a signal mass point of 2000 GeV is reported in the table. The signal samples are normalized to the expected upper limit cross sections.

$m_X$ [GeV]	$S$	$G_{\text{KK}}^* (c = 1.0)$	$G_{\text{KK}}^* (c = 2.0)$	Total Bkg.	Data
2000	$28 \pm 0.5$	$36.4 \pm 0.8$	$43.0 \pm 0.7$	$1255 \pm 27$	1107

### 5.6.2. Boosted analysis

The boosted analysis applies the selection criteria described in Section 5.5.2. After applying the large- $R$  jet mass requirement  $90 < m_{\text{Large-}R \text{ jet}} < 140$  GeV, the  $m_{HH}$  distribution is reconstructed and its shape is fit to data using MC signal and background templates. The distribution is fit using 17 bins, with almost uniform width except at low and high  $m_{HH}$ , where the bin width is modified in order to have a MC statistical uncertainty smaller than 20%. All backgrounds, except multijet, are simulated using MC generators and normalised using the cross section of the simulated process. The multijet background is estimated using the ABCD method, and its normalisation obtained from this method is kept fixed in the fit. The bias due to possible signal contamination in the ABCD regions was studied and found to have negligible effect on the result. The integral of the  $m_{HH}$  distribution for the boosted analysis is shown in Table 5.50.

Systematic uncertainties affecting the  $m_{HH}$  shape are parameterised as linear functions of  $m_{HH}$ , and the function parameters are treated as nuisance parameters in the fit. Statistical uncertainties due to the limited sample sizes of the simulated background processes are taken into account in the fit by means of further nuisance parameters, which are parameterised by Poisson priors.

Uncertainty source	Boosted analysis			
	1500 GeV [%]	2000 GeV [%]	2500 GeV [%]	3000 GeV [%]
Data statistics	+50/-52	+59/-61	+64/-66	+70/-72
Total systematic	+87/-85	+81/-79	+76/-75	+71/-69
MC statistics	+42/-48	+42/-50	+39/-48	+39/-49
$t\bar{t}$ modelling	+29/-31	+36/-38	+40/-45	+32/-39
Multijet uncertainty	+11/-14	+19/-23	+16/-20	+11/-16
$W$ -jets modelling	+27/-30	+8/-12	+11/-10	+11/-10
Single-top modelling	+22/-26	+5/-6	+4/-5	+5/-5
$b$ tagging	+31/-19	+36/-22	+36/-17	+34/-14
JES/JER	+14/-14	+6/-6	+14/-11	+7/-9
Large- $R$ jet	+29/-10	+27/-8	+27/-7	+29/-8

TABLE 5.51. Statistical and systematic contributions (in percentage) to the total error in the scaling factor  $\alpha_{\text{sig}}$  in the boosted analysis for four mass hypotheses: 1500 GeV, 2000 GeV, 2500 GeV and 3000 GeV. The first column quotes the source of the uncertainty. The contribution is obtained by calculating the difference in quadrature between the total error in  $\alpha_{\text{sig}}$  and that obtained by setting constant the nuisance parameter(s) relative to the contribution(s) under study.

2079        The systematic uncertainties included in the fit are described in Section 5.5.10.  
 2080        The contribution of the systematic uncertainties to the total uncertainty is  
 2081        decomposed into the dominant components and summarised in Table 5.51. The most  
 2082        relevant systematic uncertainties are due to the limited size of the MC samples, the  
 2083         $t\bar{t}$  modelling and the  $b$ -tagging systematic uncertainties.

Sample	Yield	Stats Err	Systs Err
$t\bar{t}$	648.7	$\pm 16.4$	+177.3(+27.3%) -169.2(-26.1%)
W+Jets	217.0	$\pm 6.5$	+104.3(+48.1%) -100.9(-46.5%)
QCD	235.2	$\pm 18.9$	+181.8(+77.3%) -181.8(-77.3%)
Single-top	109.2	$\pm 6.0$	+86.0(+78.8%) -85.8(-78.6%)
Z+Jets	20.5	$\pm 1.1$	+11.2(+54.6%) -10.9(-52.9%)
Dibosons	24.4	$\pm 1.9$	+15.3(+62.6%) -14.7(-60.1%)
Prediction	1255.0	$\pm 26.7$	+324.3(+25.8%) -311.3(-24.8%)
Data	1107	-	-
Data/Pred	0.88	-	-

TABLE 5.52. Predicted and observed yields in the signal region. Detector modeling uncertainties, MC background modeling uncertainties are considered for the systematic uncertainties. The expected background yields are predicted from MC and no normalization factors are applied.

2084      Figure 5.30 shows the  $m_{HH}$  distribution for data and the background components  
 2085      for the boosted analysis. Data are generally in good agreement with the background  
 2086      expectations within the quoted systematic errors. The signal  $m_{HH}$  distribution is  
 2087      shown in the figure for the scalar resonance. Figure 5.31 shows the observed and the  
 2088      expected upper limit on the production cross section of the scalar  $S$  particle.

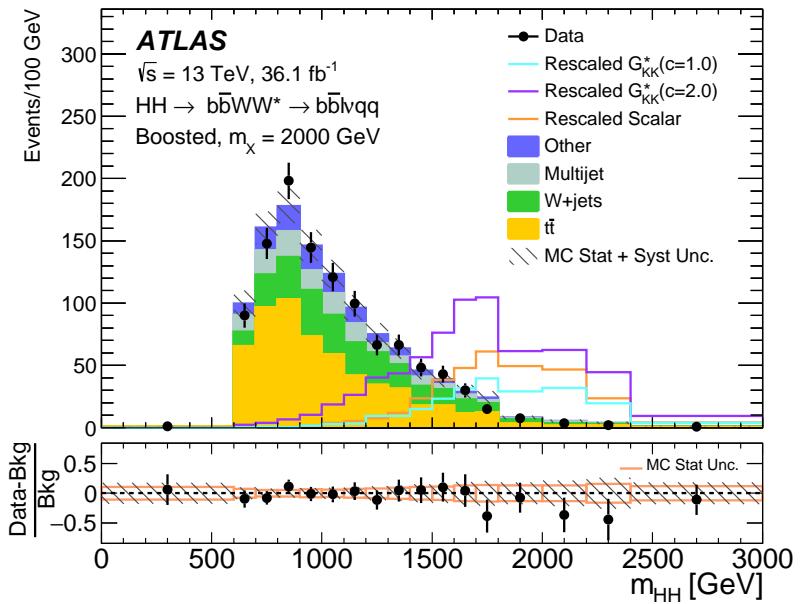


FIGURE 5.30.  $m_{HH}$  distributions after the global likelihood fit for the boosted analysis. The lower panel shows the fractional difference between data and the total expected background with the corresponding statistical and total uncertainty. The signals shown correspond to resonances of mass 2000 GeV. The scalar signal is multiplied by a factor of 4 with respect to the expected upper-limit cross section.

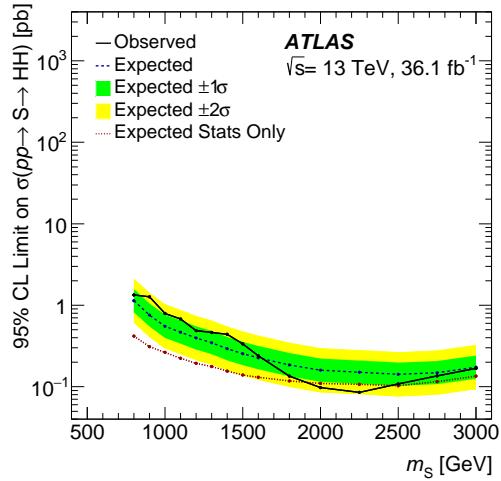


FIGURE 5.31. Expected and observed upper limits at 95% CL on the cross section of resonant pair production for the heavy scalar boson  $S$  model in the boosted analysis. The plot also shows the expected limits without including the systematic errors in order to show their impact.

## 2089 5.7. Conclusion

2090 A search for resonant and non-resonant Higgs boson pair production in the  
 2091  $b\bar{b}WW^*$  decay mode is performed in the  $b\bar{b}\ell\nu qq$  final state using  $pp$  collision data  
 2092 corresponding to an integrated luminosity of  $36.1 \text{ fb}^{-1}$ , collected at  $\sqrt{s} = 13 \text{ TeV}$  by  
 2093 the ATLAS detector at the Large Hadron Collider. No evidence of an excess of events  
 2094 over the background expectation is found. Limits are set on resonant production as a  
 2095 function of the resonance mass for a scalar resonance and for spin-2 gravitons in the  
 2096 mass range 500 to 3000 GeV. An upper limit is set on the cross section of non-resonant  
 2097 pair production  $\sigma(pp \rightarrow HH) \cdot \mathcal{B}(HH \rightarrow b\bar{b}WW^*) < 2.5 \text{ pb}$  at 95% CL corresponding  
 2098 to 300 times the predicted SM cross section. Given the result of this work, in order  
 2099 to bring relevant sensitivity improvement to the  $HH$  non-resonant SM searches in  
 2100 this channel at the LHC and at future colliders, more advanced analysis techniques,

2101 development of new methods for the normalisation of the  $t\bar{t}$  background, and a more  
2102 refined estimation of the multijet background, need to be deployed.

## IMPROVEMENTS TO THE BOOSTED ANALYSIS

2105        The boosted analysis described in chapter [chap:anal] is optimized for the case  
 2106        where the  $H \rightarrow b\bar{b}$  system is boosted and the  $H \rightarrow WW^*$  system is resolved. However,  
 2107        at high resonant masses, one would expect both the  $H \rightarrow b\bar{b}$  and the  $H \rightarrow WW^*$  to  
 2108        be boosted. However, the semileptonic decay of the W boson pair adds an additional  
 2109        complication of having a lepton within the radius of a large-R jet. This chapter will  
 2110        describe a method for reconstructing this complex topology and the improvements it  
 2111        offers to the boosted, semi-leptonic  $HH \rightarrow b\bar{b}WW^*$  analysis.

2112        **6.1. Motivation**

2113        The resonant analysis covers a large range of mass hypotheses, from 400 - 3000  
 2114        GeV. As the resonant mass increases, the two Higgs systems become more and more  
 2115        collimated. Figure 6.1 shows the average distance between final state partons as a  
 2116        function of the resonant mass for simulated HH events. The  $WW$  bosons become  
 2117        very collimated (within  $\Delta R = 0.5$ ) around 1000 GeV.

2118        On the  $H \rightarrow b\bar{b}$  side, this is accounted for through the boosted analysis selection.  
 2119        Moving toward the full Run 2 analysis, it is worth while to look at the potential gain  
 2120        from including a boosted  $H \rightarrow WW^*$  selection. This “fully-boosted” analysis can  
 2121        be used in conjunction with the current boosted and resolved analysis to increase  
 2122        sensitivity and reach.

2123        The “fully-boosted” analysis piggybacks off of the analysis presented in Chapter  
 2124        V. This means the data and Monte Carlo Samples , object reconstruction, and trigger

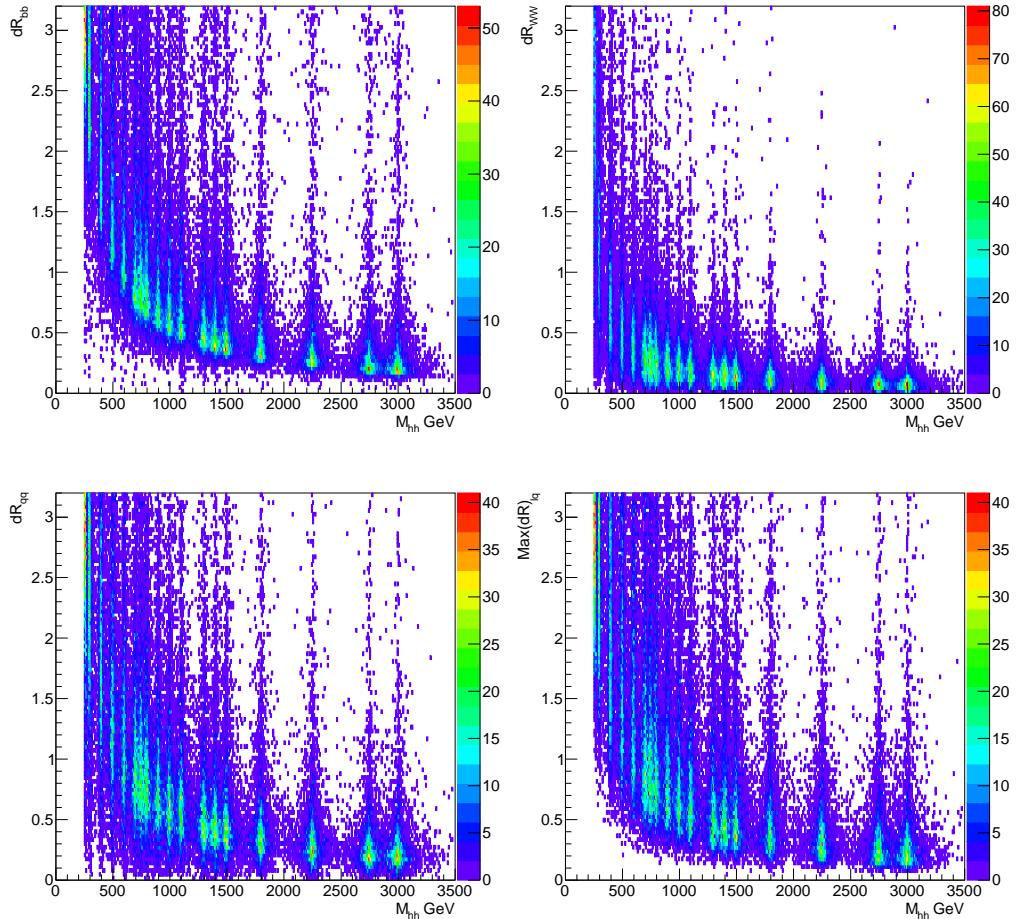


FIGURE 6.1. Distance between the two b partons (top left); the two W bosons (top right); the two light quarks (bottom left); and the lepton and the most distant light quark (bottom right) for resonant  $HH$  production as a function of resonant mass.

2125 requirements are the same as the previously presented analysis.

2126

## 2127 6.2. Event Reconstruction

Identically to the boosted analysis, section 5.5, events are reconstructed by requiring at least one reconstructed lepton. To reconstruct the  $H \rightarrow b\bar{b}$  candidate, there should be at least one large-R jet with  $\Delta R > 1.0$  from the selected lepton. The

highest of these large-R jets is selected as the  $H \rightarrow b\bar{b}$  candidate. This large-R jet is then required to have at least two track jets associated to it. Events with a  $H \rightarrow b\bar{b}$  in the range  $30\text{GeV} < m_{bb} < 300\text{GeV}$  are retained for further analysis.

To reconstruct the  $H \rightarrow WW^*$  candidate, there should be at least one large-R jet with  $\Delta R < 1.0$  from the selected lepton. This large-R jet is selected as the  $H \rightarrow WW^*$  jet candidate. Once the  $H \rightarrow b\bar{b}$  jet has been selected, they are split into either electron or muon channel for the full reconstruction.

Calorimeter jets are clusters of energy that are grouped together into an object based on distances. If an electron, which deposits the majority of its energy into the calorimeter, were to fall within the radius of a calorimeter jet, its energy should be measured as part of the jet energy. Using this, it is possible to use a single large-R jet to measure the energy of the  $W \rightarrow qq$  system and the electron. With the large-R jet and the  $\cancel{E}_T$ , it is possible to fully reconstruct the  $H \rightarrow WW^*$  system. The neutrino is reconstructed using a similar method as in Section 5.4.3. Imposing the relation:

$$m_h^2 = (p^\nu + p^{\text{large-Rjet}})^2 \quad (6.1)$$

the neutrino  $p_z$  can be reconstructed using the relations:

$$p_E^\nu = E^\nu = \sqrt{P_T^2 + p_z^2} \quad p_x^\nu = P_T \cos(\phi) \quad p_y^\nu = P_T \sin(\phi)$$

where  $\phi$  is the azimuthal angle of the  $\cancel{E}_T$ ,  $E^\nu$  the neutrino energy,  $p_x$  and  $p_y$  the two transverse spatial components of the neutrino momentum.

Muons do not deposit a significant amount of energy in the calorimeters, this means we cannot use the same reconstruction as the electrons. Instead, the muons are treated in a more traditional fashion. In the muon channel, the large-R jet contains the energy of the  $W \rightarrow qq$  system. The muon is reconstructed using the MS and ID

Lepton Channel	Alternative $p_T$ definition
Muon Channel	$p'_T = p_T^{\text{Large-Rjet}}$
Electron channel	$p'_T = \sqrt{(p_x^{\text{Large-Rjet}} - p_x^{\text{electron}})^2 + (p_y^{\text{Large-Rjet}} - p_y^{\text{electron}})^2}$

TABLE 6.1. Alternative  $p_T$  definition for the electron and muon channels

2134 information and the neutrino is reconstructed identically to 5.4.3, with the hadronic  
 2135  $W$  as a single object.

2136 Figure 6.2 shows a diagram of the event topology after the event reconstruction.

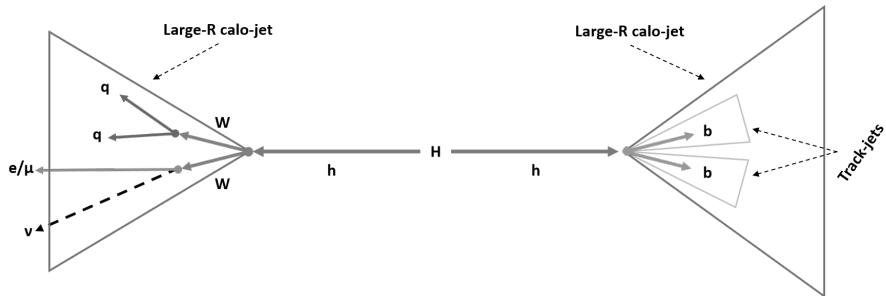


FIGURE 6.2. Diagram of the fully-boosted event topology

### 2137 6.3. Event selection

2138 After the event is reconstructed, a b-tag requirement is applied to the 2 track-  
 2139 jets in the  $H \rightarrow b\bar{b}$  candidate. The  $\cancel{E}_T$  is required to be more than 50 GeV to reject  
 2140 events from QCD background. Finally a  $p_T$  requirement is placed on the  $H \rightarrow WW^*$   
 2141 candidate. It is important to cut on the same physics objects. To accomplish this,  
 2142 an “adjusted  $p_T$ ” ( $p'_T$ ) cut of 250 GeV is applied. Table 6.1 defines the  $p'_T$  for both  
 2143 channels.

2144 **6.3.1. Comparison of reconstruction objects**

2145 Figure 6.3 and 6.4 show a comparison of the old boosted analysis reconstruction  
 2146 and the new fully-boosted analysis for the electron and muon selection respectively.  
 2147 The most

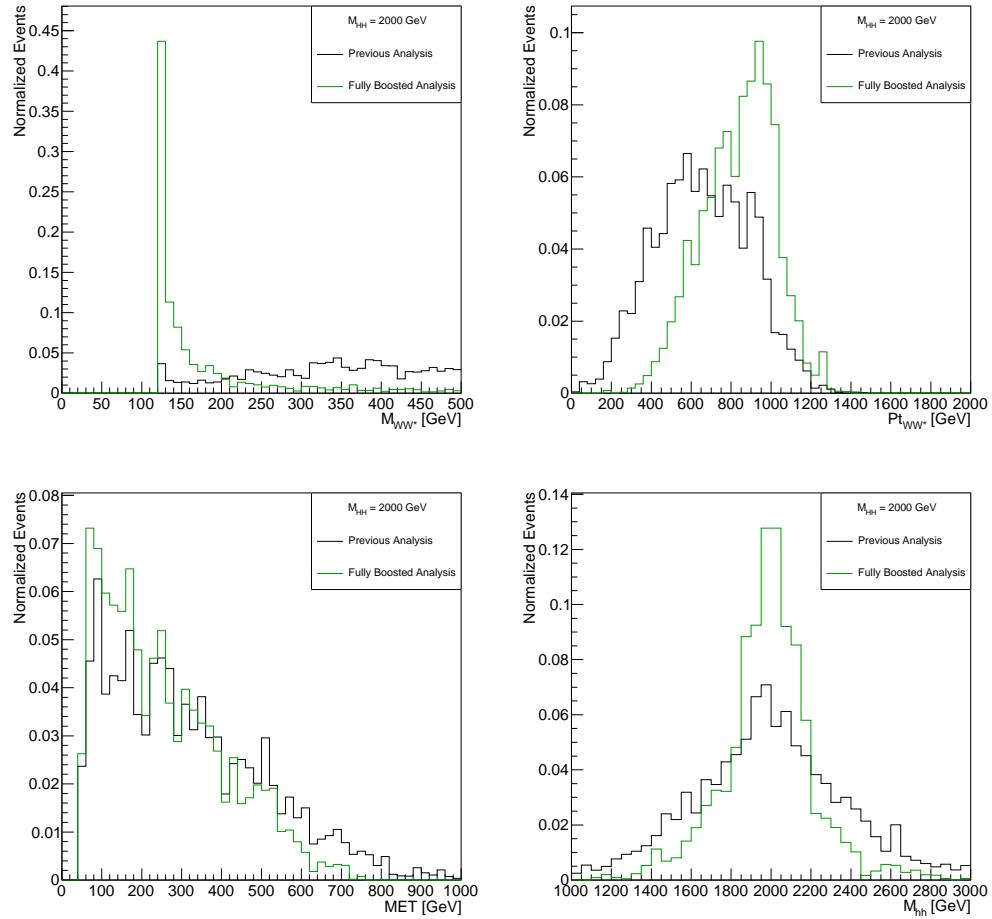


FIGURE 6.3. Comparison of  $H \rightarrow WW^*$  mass (top left),  $H \rightarrow WW^*$   $p_T$  (top right),  $\cancel{E}_T$  (bottom left), and  $HH$  mass for the previous boosted analysis reconstruction and new fully boosted selection for a resonant signal with a mass of 2000 GeV in the electron channel.

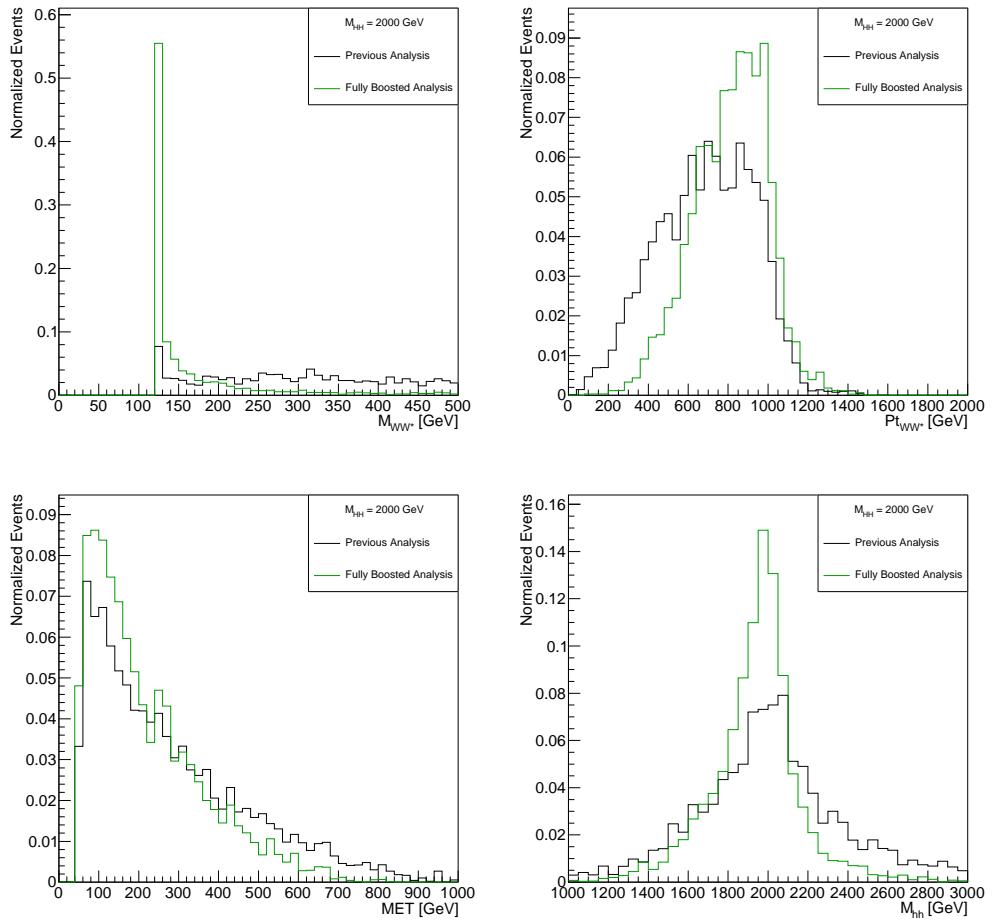


FIGURE 6.4. Comparison of  $H \rightarrow WW^*$  mass (top left),  $H \rightarrow WW^*$   $p_T$  (top right),  $\cancel{E}_T$ (bottom left), and  $HH$  mass for the previous boosted analysis reconstruction and new fully boosted selection for a resonant signal with a mass of 2000 GeV in the muon channel.

### 2148 6.3.2. Signal Region Definition

2149 As with the boosted analysis in Section 5.5, the  $h \rightarrow b\bar{b}$  candidate must have a  
 2150 jet mass in the window  $90\text{GeV} < m_{bb} < 140\text{GeV}$  to be considered in the signal region  
 2151 (SR). The previous boosted analysis included a b-jet veto for all jets outside of the

2152  $H \rightarrow b\bar{b}$  candidate. In order to increase statistics for the QCD multijet estimate, this  
2153 was removed from the fully boosted analysis.

2154 **6.3.3. mBB Control Region**

2155 To check the modeling of the background, a control region is created with an  
2156 inverted  $m_{bb}$  cut.

2157 **6.3.4. Multijet Background**

2158 The QCD multijet background is estimated using the same data-driven  
2159 background as the boosted analysis. The ABCD method with the regions defined  
2160 as:

- 2161 – Region A:  $\cancel{E}_T > 50$  GeV,  $|d_0^{\text{sig}}| < 2.0$
- 2162 – Region B:  $\cancel{E}_T < 50$  GeV,  $|d_0^{\text{sig}}| < 2.0$
- 2163 – Region C:  $\cancel{E}_T > 50$  GeV,  $|d_0^{\text{sig}}| > 2.0$
- 2164 – Region D:  $\cancel{E}_T < 50$  GeV,  $|d_0^{\text{sig}}| > 2.0$

2165 **6.3.4.1. Yield Prediction**

2166 Table 6.2 lists the MC predicted prompt lepton backgrounds, observed data and  
2167 calculated multijet yields in Region B and D before the  $H \rightarrow b\bar{b}$  mass cut is applied  
2168 and Table 6.3 shows the yields in Region C mBB control region and signal region.

2169

2170 Table 6.4 shows the ratio in the electron channel and muon channel. The  
2171 predicted yields of the QCD multijet background in the mBB control region and

Samples	Region B		Region D	
	Electron	Muon	Electron	Muon
$t\bar{t}$	$138.7 \pm 7.3$	$146.1 \pm 8.3$	$14.4 \pm 2.7$	$6.3 \pm 1.5$
W+Jets	$27.3 \pm 1.7$	$27.9 \pm 1.9$	$1.1 \pm 0.3$	$2.0 \pm 0.4$
Single-top	$7.1 \pm 1.4$	$5.2 \pm 1.4$	$0.2 \pm 0.2$	$0.4 \pm 0.3$
Z+Jets	$18.3 \pm 0.8$	$9.3 \pm 0.5$	$1.9 \pm 0.3$	$0.9 \pm 0.2$
Dibosons	$3.1 \pm 0.5$	$1.3 \pm 0.3$	$0.2 \pm 0.1$	$0.2 \pm 0.1$
Total Prompt	$194.4 \pm 7.7$	$189.9 \pm 8.7$	$17.8 \pm 2.7$	$9.8 \pm 1.6$
Data	$274.0 \pm 16.6$	$218.0 \pm 14.8$	$34.0 \pm 5.8$	$30.0 \pm 5.5$
QCD	$79.6 \pm 18.3$	$28.1 \pm 17.1$	$16.2 \pm 6.4$	$20.2 \pm 5.7$

TABLE 6.2. MC predicted prompt lepton backgrounds, observed data and calculated multijet yields in Region B and D. The multijet yield is calculated by subtracting the estimated total prompt lepton backgrounds from the observed data. The statistical uncertainty on the yields is shown.

Samples	mBBcr		SR	
	Electron	Muon	Electron	Muon
$t\bar{t}$	$21.7 \pm 4.1$	$12.8 \pm 2.1$	$6.4 \pm 1.4$	$5.4 \pm 1.2$
W+Jets	$3.4 \pm 0.6$	$2.3 \pm 0.3$	$1.4 \pm 0.3$	$1.2 \pm 0.3$
Single-top	$0.7 \pm 0.5$	$0.3 \pm 0.2$	$0.4 \pm 0.3$	$0.2 \pm 0.2$
Z+Jets	$1.0 \pm 0.2$	$0.4 \pm 0.1$	$0.4 \pm 0.1$	$0.2 \pm 0.1$
Dibosons	$0.2 \pm 0.1$	$0.0 \pm 0.0$	$0.1 \pm 0.1$	$0.0 \pm 0.0$
Total Prompt	$26.9 \pm 4.1$	$15.9 \pm 2.2$	$8.7 \pm 1.5$	$7.0 \pm 1.3$
Data	$53.0 \pm 7.3$	$33.0 \pm 5.7$	$12.0 \pm 3.5$	$20.0 \pm 4.5$
QCD	$26.1 \pm 8.4$	$17.1 \pm 6.1$	$3.3 \pm 3.8$	$13.0 \pm 4.7$

TABLE 6.3. MC predicted prompt lepton backgrounds, observed data and calculated multijet yields in Region C mBBcr and SR. The multijet yield is calculated by subtracting the estimated total prompt lepton backgrounds from the observed data. The statistical uncertainty on the yields is shown.

Multijet yield in region	Electron	Muon
$N_B^{\text{QCD}}$	$79.6 \pm 18.3$	$28.1 \pm 17.1$
$N_D^{\text{QCD}}$	$16.2 \pm 6.4$	$20.2 \pm 5.7$
$N_B^{\text{QCD}}/N_D^{\text{QCD}}$	$4.9 \pm 2.62$ (46.0%)	$1.4 \pm 0.94$ (67.2%)

TABLE 6.4. Multijet yields in region B and region D and also the ratio of the yields for each lepton channel. The error on the  $\frac{N_B^{\text{QCD}}}{N_D^{\text{QCD}}}$  ratio is propagated from the statistical uncertainties on the multijet yields in each region.

2172 signal region are presented in table 6.5. The QCD multijet background is estimated  
 2173 to be 13% of the total background in the signal region (Table 6.8).

Multijet yield in region	Electron	Muon
SR		
$N_C^{\text{QCD}}$	$3.3 \pm 3.8$	$13.0 \pm 4.7$
$N_A^{\text{QCD}}$	$16.4 \pm 20.5$ (126.8%)	$18.1 \pm 13.9$ (76.8%)
mBBcr		
$N_C^{\text{QCD}}$	$26.1 \pm 8.4$	$17.1 \pm 6.1$
$N_A^{\text{QCD}}$	$128.3 \pm 79.4$ (61.9%)	$23.9 \pm 18.1$ (75.7%)

TABLE 6.5. Multijet yield in region C and predicted yield in region A in the SR. The error on  $N_A^{\text{QCD}}$  are propagated from the error on the  $N_B^{\text{QCD}}/N_D^{\text{QCD}}$  ratio and statistical uncertainty on  $N_C^{\text{QCD}}$  yield. The numbers in brackets are the relative uncertainty in percentage.

#### 2174 6.3.4.2. Shape prediction

2175 See Section 5.5.4.2.

#### 2176 6.3.4.3. Multijet yield uncertainties

2177 **6.3.4.3.1. Statistical** The uncertainty on the predicted yield of the multijet  
 2178 background is determined by propagating the statistical uncertainty of the  $\frac{N_B^{\text{QCD}}}{N_D^{\text{QCD}}}$   
 2179 ratio, as shown in Table 6.4, and the statistical uncertainty on the multijet yield in  
 2180 region C ( $N_C^{\text{QCD}}$ ), as in Table 6.5.

2181 **6.3.4.3.2. 1-tag/2-tag jet mass acceptance** Another source of uncertainty  
 2182 on the multijet yield is the the difference of acceptance of the large- $R$  jet mass cut  
 2183 between 1-tag and 2-tag. This uncertainty is included since the template for or the  
 2184 multijet shape prediction uses the multijet shape from the 1-tag region C. Table 6.6  
 2185 shows the acceptance of the large- $R$  jet mass signal and mBB control region selection  
 2186 in the multijet 1-tag region C and 2-tag region C yields. The relative difference  
 2187 between the acceptance in 1-tag region C and in 2-tag region C is considered as an  
 2188 uncertainty on the normalization of the QCD multijet prediction.

Region	Electron	Muon
SR		
1-tag $\frac{N_{\text{SR}}}{N_{\text{Inc}}}$	32%	23 %
2-tag $\frac{N_{\text{SR}}}{N_{\text{Inc}}}$	11%	43 %
Rel. difference between 1-tag and 2-tag	26.4 %	46 %
mBBcr		
1-tag $\frac{N_{\text{mBBcr}}}{N_{\text{Inc}}}$	68%	77 %
2-tag $\frac{N_{\text{mBBcr}}}{N_{\text{Inc}}}$	89%	56 %
Rel. difference between 1-tag and 2-tag	24 %	37 %

TABLE 6.6. The acceptance of the large- $R$  jet mass signal region selection on the multijet 1-tag and 2-tag region C.  $N_{\text{SR}}(N_{\text{Inc}})$  is the multijet yield with (without) the signal region large- $R$  jet mass selection.

### 2189 **6.3.5. mBBcr Plots**

2190 Plots in the mBB control region, Figure 6.5 show an agreement between data  
 2191 and background estimates.

## 2192 **6.4. Results**

2193 The fully-boosted analysis reconstructs the  $m_{HH}$  distribution and the shape is  
 2194 fit to data using MC signal and background templates. The distribution is fit using

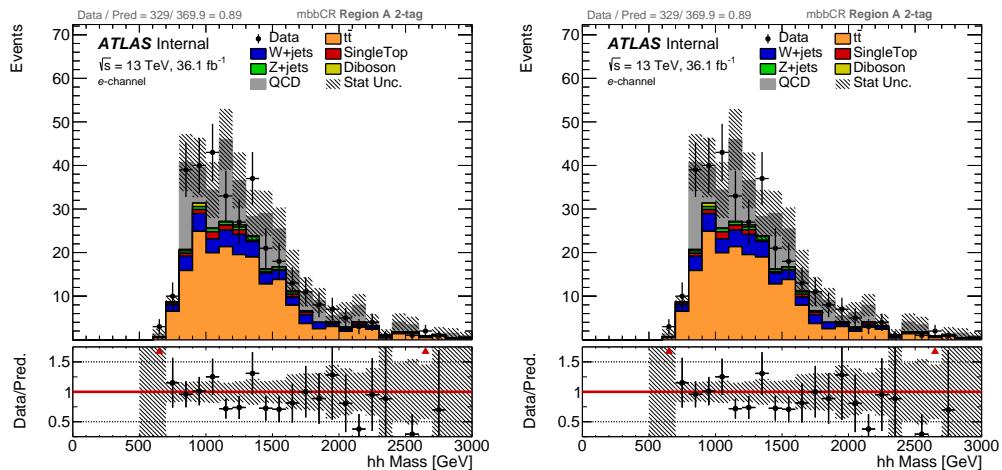


FIGURE 6.5.  $M_{HH}$  Distribution in the mBB control region for the electron (left) and muon (right) channels.

$m_X$ [GeV]	$S$	Total Bkg.	Data
2000	$26.8 \pm 0.5$	$271.0 \pm 10.7$	268

TABLE 6.7. Data event yields, and signal and background event yields in the final signal region for the boosted analysis and the scalar  $S$  particle hypothesis. The errors shown are the MC statistical uncertainties. For illustration a signal mass point of 2000 GeV is reported in the table. The signal samples are normalized to the expected upper limit cross sections.

17 bins, with almost uniform width except at low and high  $m_{HH}$ . All backgrounds, except multijet, are simulated using MC generators and normalised using the cross section of the simulated process. The multijet background is estimated using the ABCD method, and its normalisation obtained from this method is kept fixed in the fit. The bias due to possible signal contamination in the ABCD regions was studied and found to have negligible effect on the result. The integral of the  $m_{HH}$  distribution for the boosted analysis is shown in Table 6.7.

The dominant components of the systematic uncertainties are summarized in Table 6.9.

Sample	Yield	Stats Unc
$t\bar{t}$	187.7	$\pm 8.8$
W+Jets	33.7	$\pm 1.9$
QCD	34.5	$\pm 5.5$
Single-top	7.0	$\pm 1.3$
Z+Jets	4.7	$\pm 0.4$
Dibosons	3.3	$\pm 0.6$
Prediction	271.0	$\pm 10.7$
Data	268	-
Data/Pred	0.99	-

TABLE 6.8. Predicted and observed yields in the mBB control region. Detector modeling uncertainties, MC background modeling uncertainties and QCD background modeling uncertainties from ABCD method are considered for the systematic uncertainties.

2204       Figure 6.6 shows the  $m_{HH}$  distribution for data and the background components  
 2205       for the boosted analysis. Data are generally in good agreement with the background  
 2206       expectations within the quoted systematic errors. The signal  $m_{HH}$  distribution is  
 2207       shown in the figure for the scalar resonance. Figure 6.7 shows the observed and the  
 2208       expected upper limit on the production cross section of the scalar  $S$  particle.

Uncertainty	Up/Down
SysFT_EFF_Eigen_Light_0_AntiKt2PV0TrackJets_-1down	-12.9/12.5
SysFT_EFF_Eigen_C_0_AntiKt2PV0TrackJets_-1down	-12.6/12.1
SysFT_EFF_Eigen_C_0_AntiKt2PV0TrackJets_-1up	11.3/-11.9
SysFT_EFF_Eigen_Light_0_AntiKt2PV0TrackJets_-1up	11.3/-11.9
SysFATJET_Medium_JET_Comb_Baseline_Kin_-1up	-6.47/5.95
SysFATJET_Medium_JET_Comb_Baseline_Kin_-1down	5.83/-6.43
SysFT_EFF_Eigen_B_0_AntiKt2PV0TrackJets_-1down	-3.49/2.97
SysFT_EFF_Eigen_B_1_AntiKt2PV0TrackJets_-1down	-3.32/2.8
SysFT_EFF_Eigen_C_1_AntiKt2PV0TrackJets_-1down	-2.97/2.45
SysFT_EFF_Eigen_B_0_AntiKt2PV0TrackJets_-1up	2.39/-2.94
SysFT_EFF_Eigen_B_1_AntiKt2PV0TrackJets_-1up	2.23/-2.78
SysFT_EFF_Eigen_C_1_AntiKt2PV0TrackJets_-1up	1.9/-2.45
SysFATJET_Medium_JET_Comb_Tracking_Kin_-1down	1.77/-2.38
SysFATJET_Medium_JET_Comb_Tracking_Kin_-1up	-2.2/1.68
SysFT_EFF_extrapolation_AntiKt2PV0TrackJets_-1up	-2.07/1.59
SysFATJET_JMR_-1up	1.41/-1.97
SysFT_EFF_extrapolation_from_charm_AntiKt2PV0TrackJets_-1up	-1.62/1.1
SysFT_EFF_extrapolation_AntiKt2PV0TrackJets_-1down	0.963/-1.54
SysPRW_DATASF_-1down	-1.47/0.94
SysJET_SR1_JET_GroupedNP_1_-1down	0.764/-1.33
SysJET_SR1_JET_GroupedNP_1_-1up	-1.2/0.678
SysFATJET_JER_-1up	0.574/-1.16
SysFT_EFF_Eigen_Light_1_AntiKt2PV0TrackJets_-1up	-1.12/0.576
SysFT_EFF_extrapolation_from_charm_AntiKt2PV0TrackJets_-1down	0.553/-1.1
SysFATJET_Medium_JET_Comb_TotalStat_Kin_-1down	-1.03/0.498
Total Up	27.2
Total Do	27.9

TABLE 6.9. List of dominant systematic uncertainties for the fully boosted analysis. The full list of systematic uncertainties is listed in Appendix F

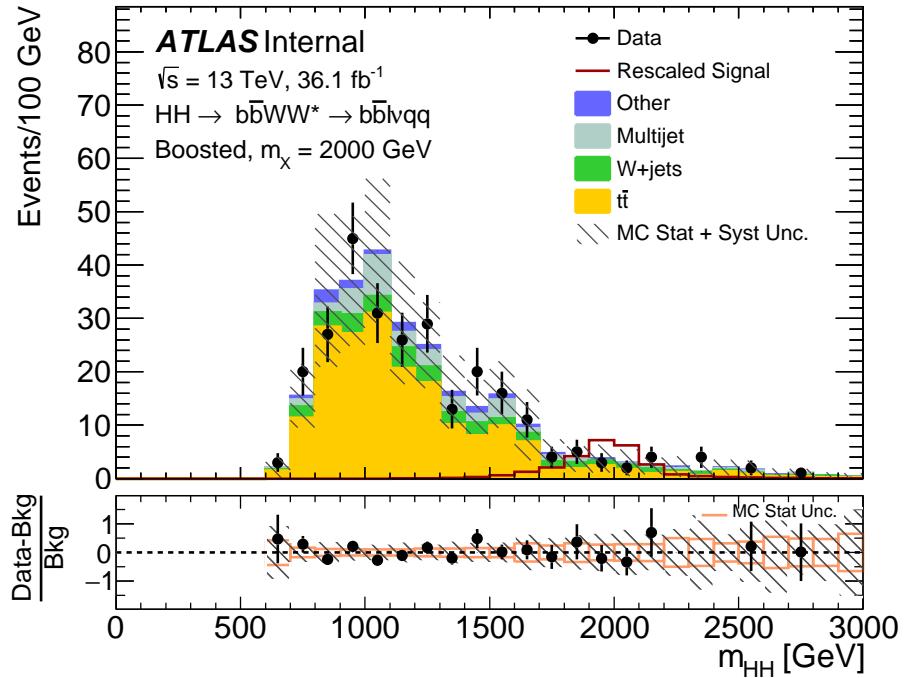


FIGURE 6.6. The  $m_{HH}$  distribution for data and background in the final signal region.

## 2209 6.5. Conclusion

2210 The fully-boosted  $HH \rightarrow b\bar{b}WW^*$  semi-leptonic analysis can strengthen the limit  
2211 on the cross section of the resonant pair production by approximately a factor of 2  
2212 over a large range of resonant masses.

2213 Several improvements can be made to the fully-boosted analysis moving toward  
2214 the full Run 2 analysis. The isolated lepton trigger that is used for these studies  
2215 is not suitable for highly collimated topologies and likely leads to a loss of signal  
2216 efficiency at high resonant masses. A large-R jet trigger would allow for leptons much  
2217 closer to other objects and could increase the number of events, especially in the high  
2218 resonant mass region. Additionally, the QCD multijet estimation suffers from a lack  
2219 of statistics in the C and D regions. This gives a large error on the QCD background

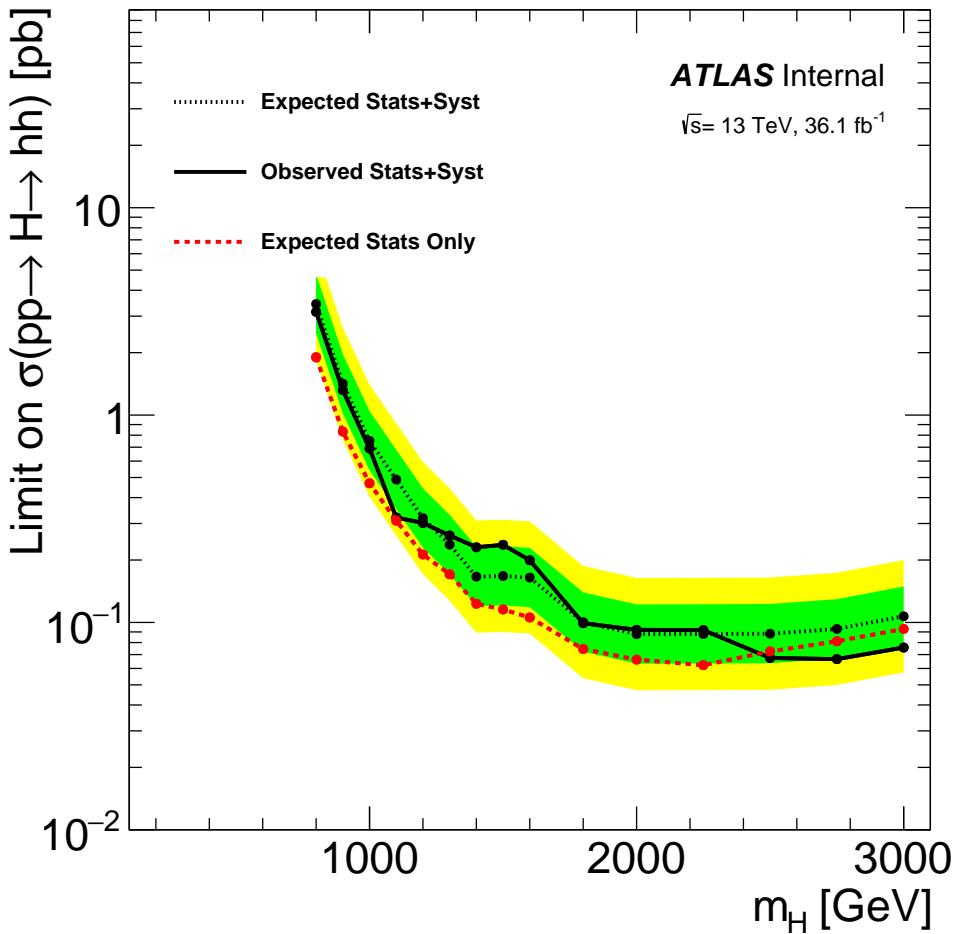


FIGURE 6.7. Expected and observed upper limits at 95% CL on the cross section of the resonant pair production for heavy scalar boson S model. The expected limits without including the systematic errors is also shown to show their impact.

estimation. A different method of QCD estimation, such as the matrix method, may be more suitable.

The fully-boosted analysis can be combined with the previous resolved and boosted analysis for the full Run-2 search. By using all three regions, it is possible to maximize the reach of the analysis across the entire mass range, even extending the range higher.

## CHAPTER VII

## CONCLUSION

2228 A search for resonant and non-resonant Higgs boson pair production in the  
2229  $b\bar{b}WW^*$  decay mode is done in the  $b\bar{b}l\nu qq$  final state using  $pp$  collision data with an  
2230 integrated luminosity of  $36.1 \text{ fb}^{-1}$  collected at  $\sqrt{s} = 13 \text{ TeV}$  by the ATLAS detector at  
2231 the LHC. No excess of events over the background only expectation is found. Limits  
2232 we set on resonant and non-resonant production.

2233 In addition to the complete analyses presented in this dissertation, an  
2234 complimentary event reconstruction is presented. This new fully boosted analysis  
2235 offers roughly a factor of 2 increase in sensitivity for the same dataset and is  
2236 a promising addition to the  $HH \rightarrow b\bar{b}WW^* \rightarrow b\bar{b}l\nu qq$  analysis for the full Run 2  
2237 analysis.

2238

## APPENDIX A

2239

### DERIVATION (HIGG5D2)

2240

Pre-selection is applied to both data and MC samples using the derivation  
framework in order to reduce the xAOD sample size. We use the HIGG5D2 derivation  
for our DxAOD sample production. More information on the derivation framework  
can be found in the Higgs group's Twiki.<sup>1</sup>

2241

2242

2243

---

<sup>1</sup> <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/HSG2xAODMigration>

## APPENDIX B

## COMPLETE LIST OF MC SAMPLES

2246        The following MC samples have been used to simulate the signal and SM  
 2247        backgrounds at the center-of-mass energy of 13 TeV. The detailed information  
 2248        including cross section, k-factor (where applicable), and the corresponding job options  
 2249        files are listed in the Twiki page<sup>1</sup>

2250        **B.0.1. Signal Samples**

```
2251 mc15_13TeV.342053.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_hh_WWbb.merge.DAOD_HIGG5D2.e4392_a766_a821_r7676_p2949
2252 mc15_13TeV.343764.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m260_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2253 mc15_13TeV.343766.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m300_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2254 mc15_13TeV.343769.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m400_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2255 mc15_13TeV.343771.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m500_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2256 mc15_13TeV.343772.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m600_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2257 mc15_13TeV.343773.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m700_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2258 mc15_13TeV.343774.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m750_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2259 mc15_13TeV.343775.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m800_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2260 mc15_13TeV.343776.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m900_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2261 mc15_13TeV.343777.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m1000_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2262 mc15_13TeV.343778.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m1100_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2263 mc15_13TeV.343779.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m1200_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2264 mc15_13TeV.343780.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m1300_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2265 mc15_13TeV.343781.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m1400_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2266 mc15_13TeV.343782.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m1500_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2267 mc15_13TeV.343783.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m1600_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2268 mc15_13TeV.343784.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m1800_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2269 mc15_13TeV.343785.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m2000_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2270 mc15_13TeV.343786.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m2250_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2271 mc15_13TeV.343787.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m2500_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2272 mc15_13TeV.343788.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m2750_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
2273 mc15_13TeV.343789.aMcAtNloHerwigppEvtGen_UEEE5_CTEQ6L1_CT10ME_Xhh_m3000_wwbb.merge.DAOD_HIGG5D2.e5153_a766_a821_r7676_p2949
```

---

<sup>1</sup> <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/CentralMC15ProductionList>

2274 **B.0.2. Background Samples**

2275 **B.0.2.1.  $t\bar{t}$**

2276 mc15\_13TeV.410000.PowhegPythiaEvtGen\_P2012\_ttbar\_hdamp172p5\_nonallhad.merge.DAOD\_HIGG5D2.e3698\_s2608\_s2183\_r7725\_r7676\_p2949

2277 **B.0.2.2. Sherpa  $W+jets$**

2278 mc15\_13TeV.364156.Sherpa\_221\_NNPDF30NNLO\_Wmumu\_MAXHPTV0\_70\_CVetoBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2279 mc15\_13TeV.364157.Sherpa\_221\_NNPDF30NNLO\_Wmumu\_MAXHPTV0\_70\_CFilterBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2280 mc15\_13TeV.364158.Sherpa\_221\_NNPDF30NNLO\_Wmumu\_MAXHPTV0\_70\_BFilter.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2281 mc15\_13TeV.364159.Sherpa\_221\_NNPDF30NNLO\_Wmumu\_MAXHPTV70\_140\_CVetoBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2282 mc15\_13TeV.364160.Sherpa\_221\_NNPDF30NNLO\_Wmumu\_MAXHPTV70\_140\_CFilterBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2283 mc15\_13TeV.364161.Sherpa\_221\_NNPDF30NNLO\_Wmumu\_MAXHPTV70\_140\_BFilter.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2284 mc15\_13TeV.364162.Sherpa\_221\_NNPDF30NNLO\_Wmumu\_MAXHPTV140\_280\_CVetoBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2285 mc15\_13TeV.364163.Sherpa\_221\_NNPDF30NNLO\_Wmumu\_MAXHPTV140\_280\_CFilterBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2286 mc15\_13TeV.364164.Sherpa\_221\_NNPDF30NNLO\_Wmumu\_MAXHPTV140\_280\_BFilter.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2287 mc15\_13TeV.364165.Sherpa\_221\_NNPDF30NNLO\_Wmumu\_MAXHPTV280\_500\_CVetoBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2288 mc15\_13TeV.364166.Sherpa\_221\_NNPDF30NNLO\_Wmumu\_MAXHPTV280\_500\_CFilterBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2289 mc15\_13TeV.364167.Sherpa\_221\_NNPDF30NNLO\_Wmumu\_MAXHPTV280\_500\_BFilter.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2290 mc15\_13TeV.364168.Sherpa\_221\_NNPDF30NNLO\_Wmumu\_MAXHPTV500\_1000.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2291 mc15\_13TeV.364169.Sherpa\_221\_NNPDF30NNLO\_Wmumu\_MAXHPTV1000\_E\_CMS.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2292 mc15\_13TeV.364170.Sherpa\_221\_NNPDF30NNLO\_Wenu\_MAXHPTV0\_70\_CVetoBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2293 mc15\_13TeV.364171.Sherpa\_221\_NNPDF30NNLO\_Wenu\_MAXHPTV0\_70\_CFilterBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2294 mc15\_13TeV.364172.Sherpa\_221\_NNPDF30NNLO\_Wenu\_MAXHPTV0\_70\_BFilter.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2295 mc15\_13TeV.364173.Sherpa\_221\_NNPDF30NNLO\_Wenu\_MAXHPTV70\_140\_CVetoBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2296 mc15\_13TeV.364174.Sherpa\_221\_NNPDF30NNLO\_Wenu\_MAXHPTV70\_140\_CFilterBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2297 mc15\_13TeV.364175.Sherpa\_221\_NNPDF30NNLO\_Wenu\_MAXHPTV70\_140\_BFilter.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2298 mc15\_13TeV.364176.Sherpa\_221\_NNPDF30NNLO\_Wenu\_MAXHPTV140\_280\_CVetoBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2299 mc15\_13TeV.364177.Sherpa\_221\_NNPDF30NNLO\_Wenu\_MAXHPTV140\_280\_CFilterBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2300 mc15\_13TeV.364178.Sherpa\_221\_NNPDF30NNLO\_Wenu\_MAXHPTV140\_280\_BFilter.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2301 mc15\_13TeV.364179.Sherpa\_221\_NNPDF30NNLO\_Wenu\_MAXHPTV280\_500\_CVetoBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2302 mc15\_13TeV.364180.Sherpa\_221\_NNPDF30NNLO\_Wenu\_MAXHPTV280\_500\_CFilterBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2303 mc15\_13TeV.364181.Sherpa\_221\_NNPDF30NNLO\_Wenu\_MAXHPTV280\_500\_BFilter.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2304 mc15\_13TeV.364182.Sherpa\_221\_NNPDF30NNLO\_Wenu\_MAXHPTV500\_1000.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2305 mc15\_13TeV.364183.Sherpa\_221\_NNPDF30NNLO\_Wenu\_MAXHPTV1000\_E\_CMS.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2306 mc15\_13TeV.364184.Sherpa\_221\_NNPDF30NNLO\_Wtaunu\_MAXHPTV0\_70\_CVetoBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2307 mc15\_13TeV.364185.Sherpa\_221\_NNPDF30NNLO\_Wtaunu\_MAXHPTV0\_70\_CFilterBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2308 mc15\_13TeV.364186.Sherpa\_221\_NNPDF30NNLO\_Wtaunu\_MAXHPTV0\_70\_BFilter.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2309 mc15\_13TeV.364187.Sherpa\_221\_NNPDF30NNLO\_Wtaunu\_MAXHPTV70\_140\_CVetoBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2310 mc15\_13TeV.364188.Sherpa\_221\_NNPDF30NNLO\_Wtaunu\_MAXHPTV70\_140\_CFilterBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2311 mc15\_13TeV.364189.Sherpa\_221\_NNPDF30NNLO\_Wtaunu\_MAXHPTV70\_140\_BFilter.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2312 mc15\_13TeV.364190.Sherpa\_221\_NNPDF30NNLO\_Wtaunu\_MAXHPTV140\_280\_CVetoBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2313 mc15\_13TeV.364191.Sherpa\_221\_NNPDF30NNLO\_Wtaunu\_MAXHPTV140\_280\_CFilterBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2314 mc15\_13TeV.364192.Sherpa\_221\_NNPDF30NNLO\_Wtaunu\_MAXHPTV140\_280\_BFilter.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2315 mc15\_13TeV.364193.Sherpa\_221\_NNPDF30NNLO\_Wtaunu\_MAXHPTV280\_500\_CVetoBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2316 mc15\_13TeV.364194.Sherpa\_221\_NNPDF30NNLO\_Wtaunu\_MAXHPTV280\_500\_CFilterBVeto.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2317 mc15\_13TeV.364195.Sherpa\_221\_NNPDF30NNLO\_Wtaunu\_MAXHPTV280\_500\_BFilter.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2318 mc15\_13TeV.364196.Sherpa\_221\_NNPDF30NNLO\_Wtaunu\_MAXHPTV500\_1000.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949  
2319 mc15\_13TeV.364197.Sherpa\_221\_NNPDF30NNLO\_Wtaunu\_MAXHPTV1000\_E\_CMS.merge.DAOD\_HIGG5D2.e5340\_s2726\_r7772\_r7676\_p2949

2320 **B.0.2.3.  $Z+jets$**

2321 mc15\_13TeV.364100.Sherpa\_221\_NNPDF30NNLO\_Zmumu\_MAXHPTV0\_70\_CVetoBVeto.merge.DAOD\_HIGG5D2.e5271\_s2726\_r7772\_r7676\_p2949

```

2322 mc15_13TeV.364101.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_CFilterBVeto.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949
2323 mc15_13TeV.364102.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_BFilter.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949
2324 mc15_13TeV.364103.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_140_CVetoBVeto.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949
2325 mc15_13TeV.364104.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_140_CFilterBVeto.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949
2326 mc15_13TeV.364105.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_140_BFilter.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949
2327 mc15_13TeV.364106.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_CVetoBVeto.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949
2328 mc15_13TeV.364107.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_CFilterBVeto.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949
2329 mc15_13TeV.364108.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_BFilter.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949
2330 mc15_13TeV.364109.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_CVetoBVeto.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949
2331 mc15_13TeV.364110.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_CFilterBVeto.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949
2332 mc15_13TeV.364111.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_BFilter.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949
2333 mc15_13TeV.364112.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV500_1000.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949
2334 mc15_13TeV.364113.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV1000_E_CMS.merge.DAOD_HIGG5D2.e5271_s2726_r7772_r7676_p2949
2335 mc15_13TeV.364114.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CVetoBVeto.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949
2336 mc15_13TeV.364115.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CFilterBVeto.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949
2337 mc15_13TeV.364116.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_BFilter.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949
2338 mc15_13TeV.364117.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_140_CVetoBVeto.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949
2339 mc15_13TeV.364118.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_140_CFilterBVeto.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949
2340 mc15_13TeV.364119.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_140_BFilter.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949
2341 mc15_13TeV.364120.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CVetoBVeto.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949
2342 mc15_13TeV.364121.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CFilterBVeto.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949
2343 mc15_13TeV.364122.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_BFilter.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949
2344 mc15_13TeV.364123.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CVetoBVeto.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949
2345 mc15_13TeV.364124.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CFilterBVeto.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949
2346 mc15_13TeV.364125.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_BFilter.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949
2347 mc15_13TeV.364126.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV500_1000.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949
2348 mc15_13TeV.364127.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV1000_E_CMS.merge.DAOD_HIGG5D2.e5299_s2726_r7772_r7676_p2949
2349 mc15_13TeV.364128.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CVetoBVeto.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949
2350 mc15_13TeV.364129.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CFilterBVeto.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949
2351 mc15_13TeV.364130.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_BFilter.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949
2352 mc15_13TeV.364131.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_140_CVetoBVeto.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949
2353 mc15_13TeV.364132.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_140_CFilterBVeto.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949
2354 mc15_13TeV.364133.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_140_BFilter.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949
2355 mc15_13TeV.364134.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CVetoBVeto.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949
2356 mc15_13TeV.364135.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CFilterBVeto.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949
2357 mc15_13TeV.364136.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_BFilter.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949
2358 mc15_13TeV.364137.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CVetoBVeto.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949
2359 mc15_13TeV.364138.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CFilterBVeto.merge.DAOD_HIGG5D2.e5313_s2726_r7772_r7676_p2949
2360 mc15_13TeV.364139.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_BFilter.merge.DAOD_HIGG5D2.e5313_s2726_r7772_r7676_p2949
2361 mc15_13TeV.364140.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV500_1000.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949
2362 mc15_13TeV.364141.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV1000_E_CMS.merge.DAOD_HIGG5D2.e5307_s2726_r7772_r7676_p2949

```

#### 2363 B.0.2.4. Dibosons

```

2364 mc15_13TeV.361091.Sherpa_CT10_WplvWmqq_SHv21_improved.merge.DAOD_HIGG5D2.e4607_s2726_r7772_r7676_p2949
2365 mc15_13TeV.361092.Sherpa_CT10_WpqqWmlv_SHv21_improved.merge.DAOD_HIGG5D2.e4607_s2726_r7772_r7676_p2949
2366 mc15_13TeV.361093.Sherpa_CT10_WlvZqq_SHv21_improved.merge.DAOD_HIGG5D2.e4607_s2726_r7772_r7676_p2949
2367 mc15_13TeV.361094.Sherpa_CT10_WqqZll_SHv21_improved.merge.DAOD_HIGG5D2.e4607_s2726_r7772_r7676_p2949
2368 mc15_13TeV.361095.Sherpa_CT10_WqqZvv_SHv21_improved.merge.DAOD_HIGG5D2.e4607_s2726_r7772_r7676_p2949
2369 mc15_13TeV.361096.Sherpa_CT10_ZqqZll_SHv21_improved.merge.DAOD_HIGG5D2.e4607_s2726_r7772_r7676_p2949
2370 mc15_13TeV.361097.Sherpa_CT10_ZqqZvv_SHv21_improved.merge.DAOD_HIGG5D2.e4607_s2726_r7772_r7676_p2949

```

2371   **B.0.2.5. Single Top**

```
2372 mc15_13TeV.410011.PowhegPythiaEvtGen_P2012_singletop_tchan_lept_top.merge.DAOD_HIGG5D2.e3824_s2608_s2183_r7725_r7676_p2949
2373 mc15_13TeV.410012.PowhegPythiaEvtGen_P2012_singletop_tchan_lept_antitop.merge.DAOD_HIGG5D2.e3824_s2608_s2183_r7725_r7676_p2949
2374 mc15_13TeV.410013.PowhegPythiaEvtGen_P2012_Wt_inclusive_top.merge.DAOD_HIGG5D2.e3753_s2608_s2183_r7725_r7676_p2949
2375 mc15_13TeV.410014.PowhegPythiaEvtGen_P2012_Wt_inclusive_antitop.merge.DAOD_HIGG5D2.e3753_s2608_s2183_r7725_r7676_p2949
2376 mc15_13TeV.410025.PowhegPythiaEvtGen_P2012_SingleTopSchan_noAllHad_top.merge.DAOD_HIGG5D2.e3998_s2608_s2183_r7725_r7676_p2949
2377 mc15_13TeV.410026.PowhegPythiaEvtGen_P2012_SingleTopSchan_noAllHad_antitop.merge.DAOD_HIGG5D2.e3998_s2608_s2183_r7725_r7676_p2949
```

2378   **B.0.2.6. Single Higgs**

```
2379 mc15_13TeV.342282.PowhegPythia8EvtGen_CT10_AZNLOCTEQ6L1_ggH125_inc.merge.DAOD_HIGG5D2.e4850_a766_a821_r7676_p2949
2380 mc15_13TeV.342285.Pythia8EvtGen_A14NNPDF23LO_ZH125_inc.merge.DAOD_HIGG5D2.e4246_s2608_s2183_r7772_r7676_p2949
```

2381   **B.0.2.7. QCD Multijet**

```
2382 mc15_13TeV.344715.Sherpa_CT10_bb_MassiveCB_2Bjets_Pt30_50.merge.DAOD_HIGG5D2.e5681_a766_a821_r7676_p2949
2383 mc15_13TeV.344716.Sherpa_CT10_bb_MassiveCB_2Bjets_Pt50_80.merge.DAOD_HIGG5D2.e5681_a766_a821_r7676_p2949
2384 mc15_13TeV.344717.Sherpa_CT10_bb_MassiveCB_2Bjets_Pt80_130.merge.DAOD_HIGG5D2.e5681_a766_a821_r7676_p2949
2385 mc15_13TeV.344718.Sherpa_CT10_bb_MassiveCB_2Bjets_Pt130_200.merge.DAOD_HIGG5D2.e5681_a766_a821_r7676_p2949
2386 mc15_13TeV.344719.Sherpa_CT10_bb_MassiveCB_2Bjets_Pt200_E_CMS.merge.DAOD_HIGG5D2.e5681_a766_a821_r7676_p2949
```

## B.1. Lepton selection optimisation

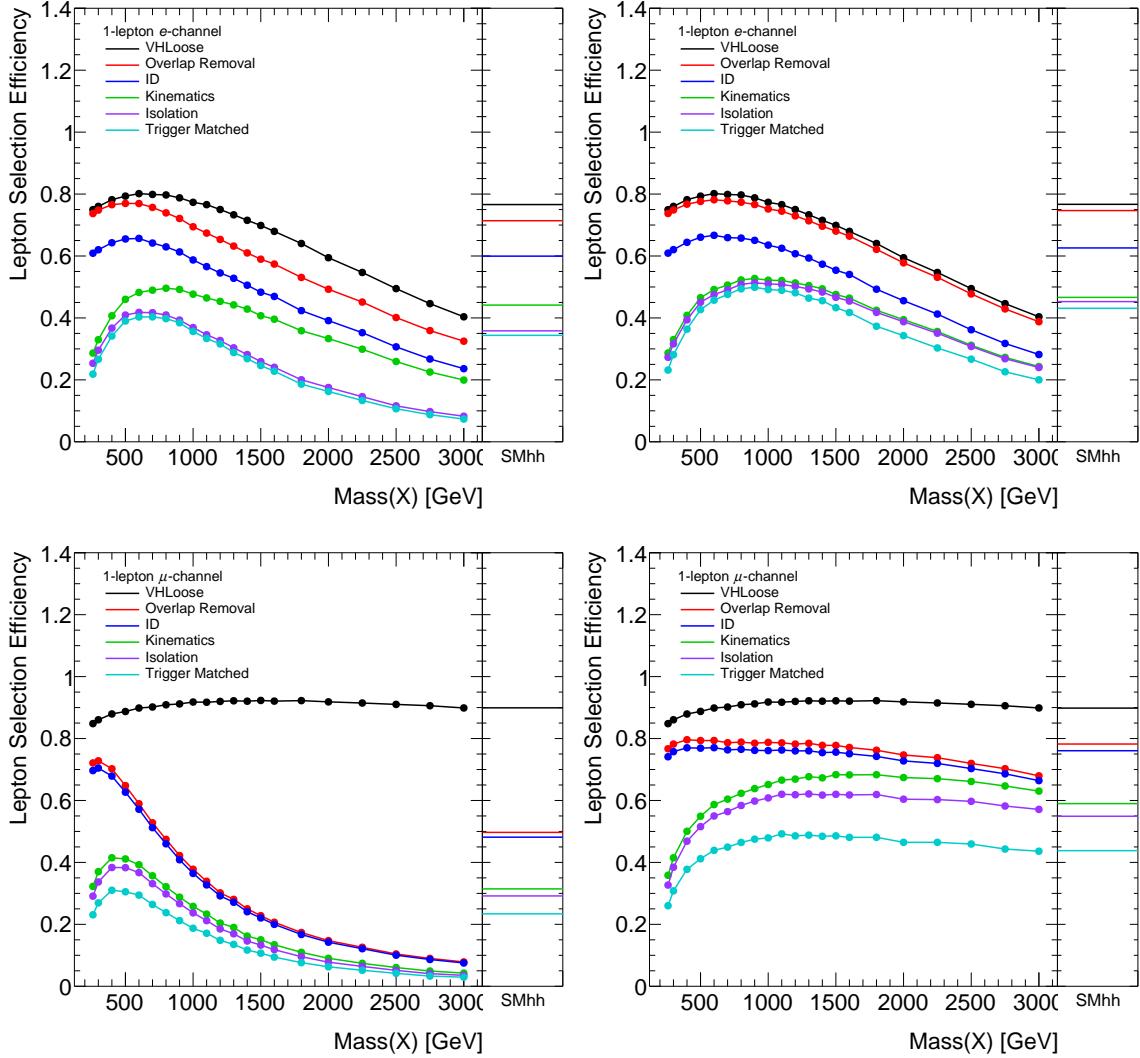


FIGURE B.1. Lepton selection efficiency as a function of signal Scalar resonance mass for the electron-channel (top) and the muon-channel (bottom). The efficiency for the SM di-Higgs signal sample (SMhh) is also shown. The plots on the left are efficiencies for the lepton selections as used in the ICHEP 2016 analysis while the plot on the right are the optimised baseline lepton selections.

The overlap removal procedure, as presented in Sec. 5.3.5 requires calorimeter-jets to be removed from the event if it is within. Due to the presence of  $b$ -jets from the  $h \rightarrow b\bar{b}$  decay, it is possible for a reconstructed muon or electron to be in close proximity of the  $b$ -jets, due to semi-leptonic decays of the  $b$ -hadron, and the  $b$ -jets fail the overlap removal criteria. Figure C.1 shows the signal efficiency to find the leading and sub-leading  $b$ -jet in the event. In this study, the  $b$ -jets are identified by requiring at least one  $b$ -hadron within the  $b$ -jets.

The efficiency loss due to electrons, which pass the VHLooseElectron selection, to be within  $\Delta R \geq 0.2$  of the b-jets are negligible as can be seen in the top figures. For muons that pass VHLooseMuon selection, the impact is slightly more pronounced but still small ( $\sim 3\%$ ). Note that in the jet-muon overlap removal procedure, additional requirements are imposed on the muon and the overlapping jet before deciding to remove the jet and this is expected to mitigate the lost of efficiency. The impact of the additional requirements are not studied here.

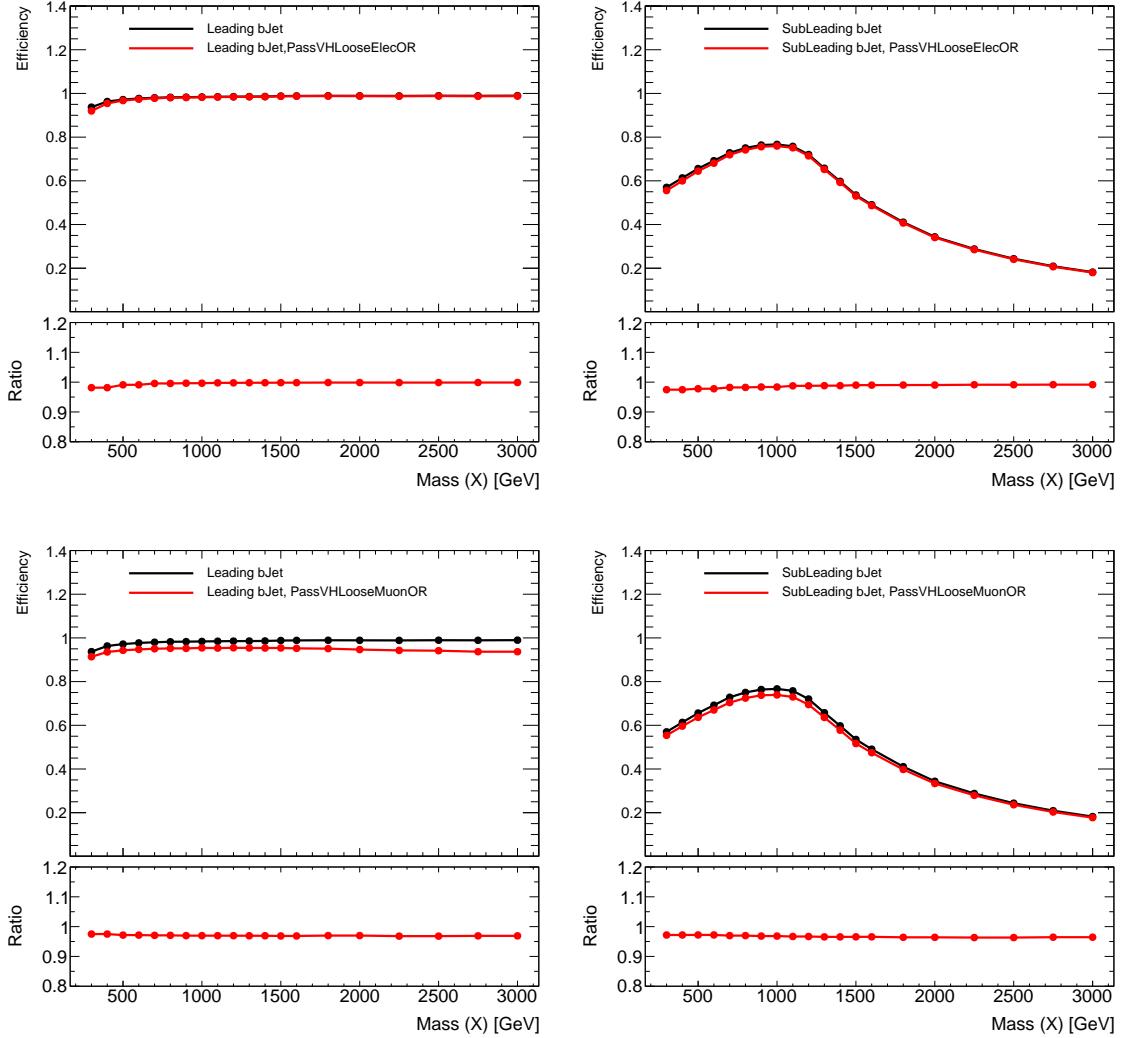


FIGURE C.1. The signal efficiencies to find the leading (left) and sub-leading (right) jets which has at least 1  $b$ -hadron within the jets. The red curves correspond to the requirement that the jets do not have a VHLoose electrons (top) and muons (bottom) within  $\Delta R \leq 0.2$  from the jet axis.

2405

## APPENDIX D

2406

### SOLVING FOR NEUTRINO LONGITUDINAL MOMENTUM

2407        Neutrinos are not detected directly using the ATLAS detector. Instead,  
 2408        their transverse momentum is determined using conservation of momentum. The  
 2409        longitudinal momentum of neutrinos can be calculated in terms of the momentum  
 2410        and masses of other particles in the event, but it is a solution to a quadratic equation,  
 2411        and hence there are two possible solutions. Please see Sec. ?? for the equations. We  
 2412        investigated the efficiencies of several different methods of picking the solution.

2413        Several methods for selecting the sign were tested on truth level using the signal  
 2414        Monte Carlo samples for resonant diHiggs production with mass of 700 GeV, 2000  
 2415        GeV and 5000 GeV. The choice of sign is said to be correct if the solution is within  
 2416        10% of the truth value if the truth value is larger than 100 GeV or within 10 GeV if the  
 2417        truth value is smaller than 100 GeV. This criterion was chosen to take into account  
 2418        the fact that we do not need to know the exact momentum, just the momentum up  
 2419        to errors from the detector resolution.

2420

#### D.0.1. $W$ mass method

2421        The mass of the  $W$  that decays leptonically can be calculated from the  
 2422        momentum of the lepton and neutrino since  $P_W = P_l + P_\nu$ . The mass of the leptonic  
 2423         $W$  we calculate with the chosen solution depends on our choice of sign. As figure D.3  
 2424        shows, the  $W$  mass distribution has peaks at 40 and 80 GeV. So for this method, we  
 2425        will use the solution that minimises  $|m_W - 80|$  or  $|m_W - 40|$ .

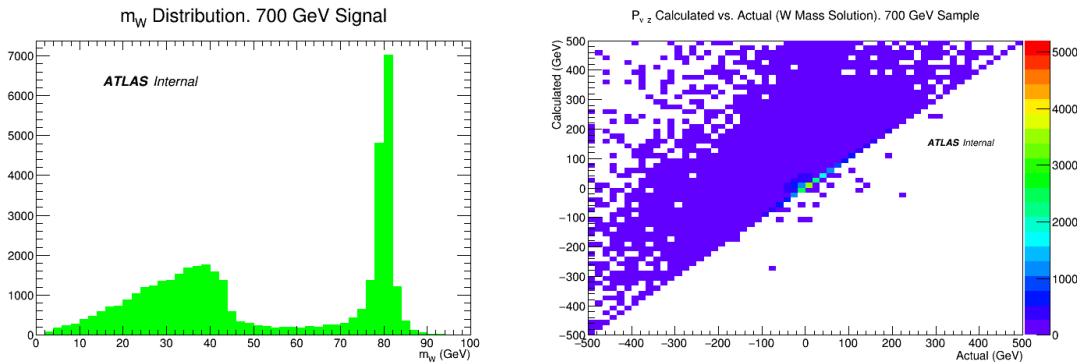


FIGURE D.1.  $m_W$  Distribution and the solution using  $W$  mass method.

### 2426 D.0.2. $\eta$ solution

2427 The next method we tested was picking the sign in the solution that is the  
 2428 opposite of the sign of  $\eta$  of the lepton. This was motivated by studying the same  
 2429 problem for single  $W$  production and noticing that there is a correlation between the  
 2430 the correct sign choice and the sign of  $\eta$ . This correlation is shown in figure D.5.  
 2431 It is interesting to note that this method is mathematically equivalent to using the  
 2432 opposite sign of the longitudinal momentum of the lepton. This can be verified by  
 2433 looking at the definition of  $\eta_l = \frac{1}{2} \log \frac{p_l + p_{l,z}}{p_l - p_{l,z}}$ . If  $p_{l,z} < 0$  then  $\frac{p_l + p_{l,z}}{p_l - p_{l,z}} < 1$  so  $\eta_l < 0$  and  
 2434 similarly if  $p_{l,z} > 0$  then  $\eta_l > 0$ .

2435 We also tested using the opposite sign of  $\eta$  of the  $lq\bar{q}$  system. Using the additional  
 2436 information from the quark antiquark pair allows us to pick the correct sign more often  
 2437 than just using  $\eta_l$ , D.4

### 2438 D.0.3. $\Delta R$ Method

2439 We also investigated minimising  $\Delta R(l, \nu)$ . This is motivated by the idea that  
 2440 the neutrino and the charged lepton produced from the same  $W$  should be near each  
 2441 if the  $W$  boson is boosted, which is often the case for heavy resonances decaying into

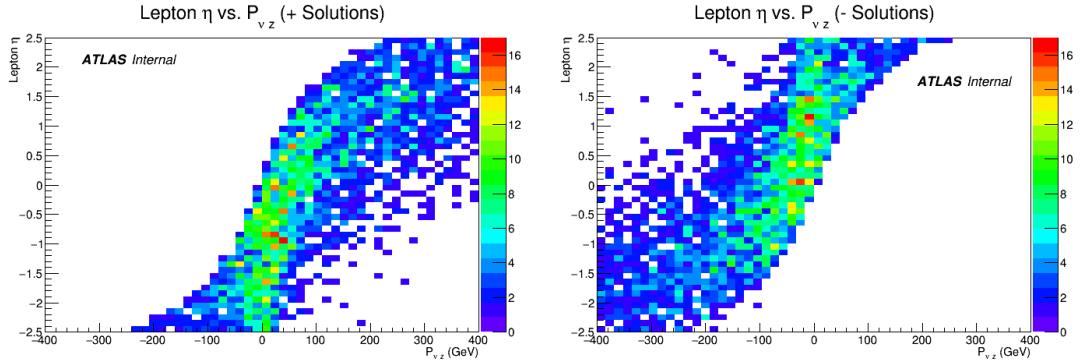


FIGURE D.2. Single  $W$   $\eta_l$  vs Correct Sign Choice

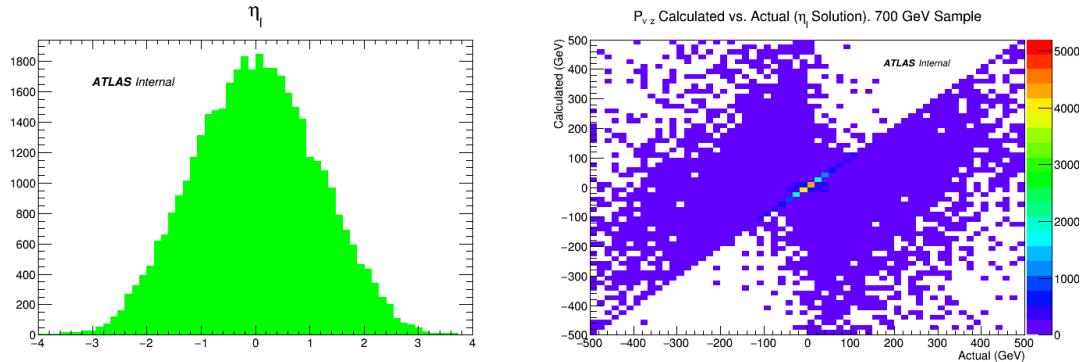


FIGURE D.3.  $\eta_l$  and the solution using  $\eta_l$  method in 700 GeV resonant sample.

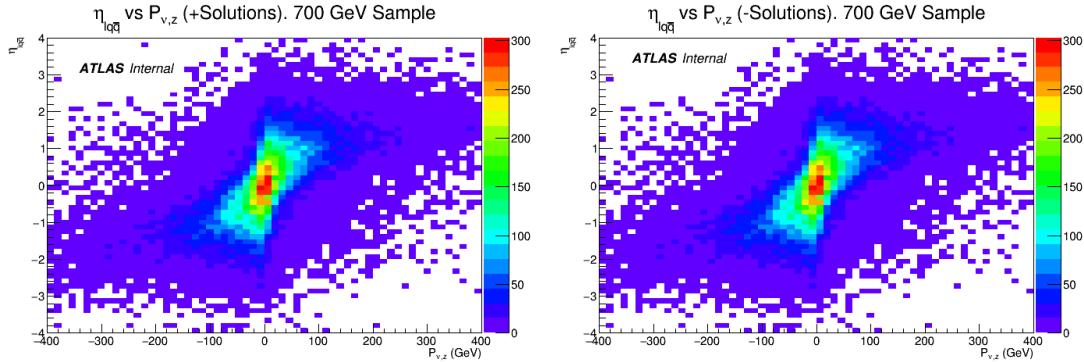


FIGURE D.4. Higgs pair  $\eta_{l\bar{q}}$  vs Correct Sign Choice

<sup>2442</sup>  $W$  boson. Fig. D.6 shows the  $\Delta R(l, \nu)$  and the solution using  $\Delta R(l, \nu)$  method in  
<sup>2443</sup> 700 GeV resonant sample.

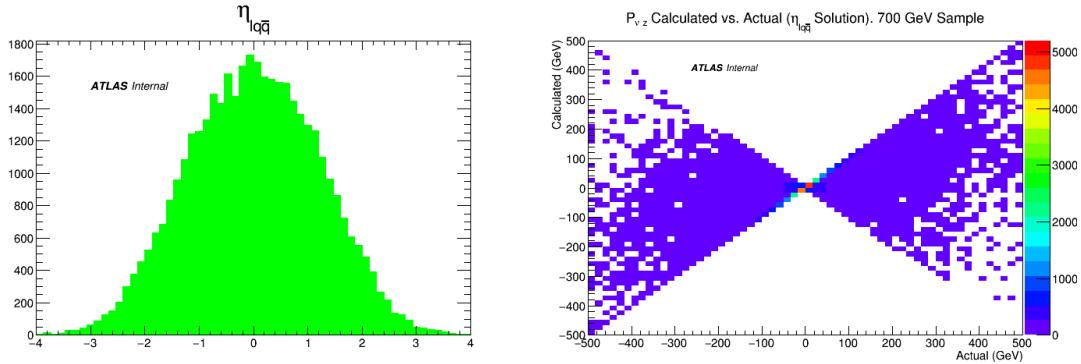


FIGURE D.5.  $\eta_l$  and the solution using  $\eta_l$  method in 700 GeV resonant sample.

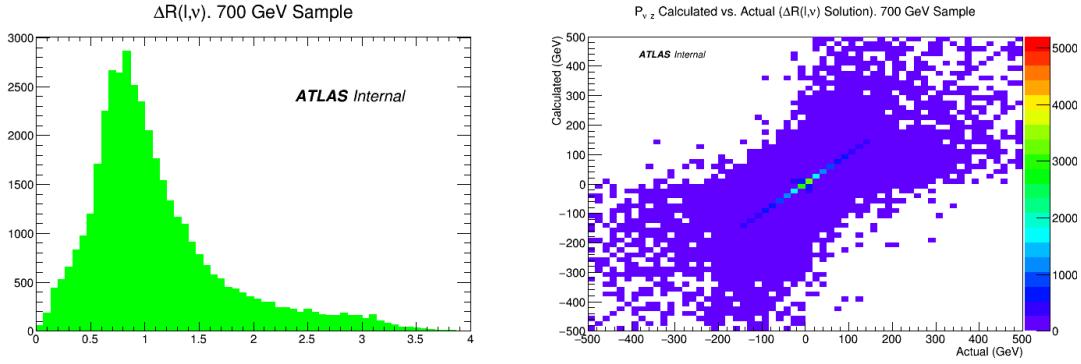


FIGURE D.6.  $\Delta R(l, \nu)$  and the solution using  $\Delta R(l, \nu)$  method in 700 GeV resonant sample.

#### D.0.4. Results

The fraction of the time each method picks the correct sign is given as a percentage in the table below.

Method	700 GeV Sample	2000 GeV Sample	5000 GeV Sample
$m_W$	55.7	56.3	58.3
$\eta_l$	52.9	47.6	47.4
$\eta_{lq\bar{q}}$	59.9	60.5	63.3
$\Delta R(l, \nu)$	57	65	72.7

## APPENDIX E

## FREEZING B AND D REGIONS IN QCD ESTIMATE

2450 This appendix summarizes the study undertaken to select regions to 'freeze' the  
 2451 cuts in the B and D regions used in the ABCD estimation for QCD background. For  
 2452 each selection, the B/D values diverge from the value calculated at the beginning of  
 2453 the selection when the statistics in the B and/or D region drops significantly. To avoid  
 2454 large statistical errors in the normalization calculated for the multi-jet contribution  
 2455 in the A region, the yields from the B and D regions used in the ABCD calculation  
 2456 are frozen, i.e. no further cuts applied after the earliest cut in the selection which  
 2457 has a B/D ratio consistent with the last statistically stable B/D ratio. Using this  
 2458 method, the multi-jet modelling is kept as close as possible to the phase space in the  
 2459 final signal region while taking advantage of higher statistics earlier in the cutflow.  
 2460 All numbers in this study were conducted with a  $t\bar{t}$  normalization factor equal to  
 2461 1.0, i.e. no data-driven normalization was applied to  $t\bar{t}$ .

QCD $B/D$ Values, Non-resonant Selection			
mww	bbpt210	bbpt300	wwpt250
$0.36 \pm 0.01$	$0.35 \pm 0.07$	$0.49 \pm 0.30$	$0.14 \pm 0.10$
QCD $B/D$ Values, Low Mass (m700) Selection			
mww	bbpt210	wwpt250	hh700
$0.36 \pm 0.01$	$0.35 \pm 0.07$	$0.04 \pm 0.02$	$0.11 \pm 0.07$
QCD $B/D$ Values, High Mass Selection			
bbpt350	wwpt250	drww15	hh2000
$0.29 \pm 0.04$	$0.28 \pm 0.04$	$0.22 \pm 0.06$	$1.24 \pm 0.86$

TABLE E.1. Values calculated for  $B/D$  at each stage in the non-resonant, low mass, and high mass selections. The estimate of multi-jet contribution in the A region uses the  $R$  value calculated after the selection described in the text.

<sup>2462</sup> The QCD and total background yields obtained in the  $m_{b\bar{b}}$  control region without  
<sup>2463</sup> freezing the B and D regions are shown in Tables E.2, E.3, and E.4.

Non-resonant Selection in $m_{b\bar{b}}$ Control Region, No B/D Freezing				
Sample	mww	bbpt210	bbpt300	wwpt250
QCD	$13310.5 \pm 500.3$	$250.2 \pm 30.6$	$24.6 \pm 3.0$	$54.8 \pm 6.7$
Background Sum	$43849.0 \pm 509.2$	$1017.9 \pm 33.7$	$192.8 \pm 7.1$	$153.2 \pm 8.4$
Data	43902.0	1069.0	206.0	138.0

TABLE E.2. QCD and total background yields for the non-resonant selection without freezing the selection cuts used in the B and D regions, i.e. the yields in the B and D region after each cut are used in the ABCD calculation up until the  $m_{b\bar{b}}$  cut. Non-monotonic QCD yields are observed.

Low Mass (m700) Selection in $m_{b\bar{b}}$ Control Region, No B/D Freezing				
Sample	mww	bbpt210	wwpt250	hh700
QCD	$13310.5 \pm 500.3$	$250.2 \pm 30.6$	$585.3 \pm 71.7$	$54.8 \pm 6.7$
Background Sum	$43849.0 \pm 509.2$	$1017.9 \pm 33.7$	$843.7 \pm 72.1$	$104.7 \pm 7.6$
Data	43902.0	1069.0	367.0	89.0

TABLE E.3. QCD and total background yields for the low mass selection without freezing the selection cuts used in the B and D regions, i.e. the yields in the B and D region after each cut are used in the ABCD calculation up until the  $m_{b\bar{b}}$  cut.

High Mass Selection in $m_{b\bar{b}}$ Control Region, No B/D Freezing				
Sample	bbpt350	wwpt250	drww15	hh2000
QCD	$1538.7 \pm 252.7$	$1359.5 \pm 75.9$	$486.4 \pm 27.1$	$4.6 \pm 0.3$
Background Sum	$14719.1 \pm 258.9$	$12463.5 \pm 91.8$	$3671.3 \pm 38.5$	$222.8 \pm 7.1$
Data	14862.0	12450.0	3761.0	250.0

TABLE E.4. QCD and total background yields for the high mass selection without freezing the selection cuts used in the B and D regions, i.e. the yields in the B and D region after each cut are used in the ABCD calculation up until the  $m_{b\bar{b}}$  cut.

<sup>2464</sup> The B/D ratio at the last cut in each selection have large errors (near or larger  
<sup>2465</sup> than 100%) and are found to be unstable. The yields after freezing the B and D  
<sup>2466</sup> regions to their yields after the earliest selection cut with a B/D ratio consistent with

<sup>2467</sup> the last statistically stable B/D ratio (after  $p_T^{bb} > 210$  GeV for the non-resonant and  
<sup>2468</sup> low-mass selections and after  $p_T^{WW} > 250$  GeV for the high mass selection) are shown  
<sup>2469</sup> in Tables E.5, E.6, and E.7.

Non-resonant Selection in $m_{b\bar{b}}$ Control Region, B/D Frozen after $p_T^{bb} > 210$				
Sample	mww	bbpt210	bbpt300	wwpt250
QCD	$13310.5 \pm 500.3$	$250.2 \pm 30.6$	$33.7 \pm 4.1$	$21.4 \pm 2.6$
Background Sum	$43849.0 \pm 509.2$	$1017.9 \pm 33.7$	$201.9 \pm 7.6$	$119.8 \pm 5.7$
Data	43902.0	1069.0	206.0	138.0

TABLE E.5. QCD and total background yields for the non-resonant selection after freezing the selection cuts used in the B and D regions after requiring  $p_T^{bb} > 210$  GeV. Monotonic QCD yields are now observed.

Low Mass (m700) Selection in $m_{b\bar{b}}$ Control Region, B/D Frozen after $p_T^{bb} > 210$				
Sample	mww	bbpt210	wwpt250	hh700
QCD	$13310.5 \pm 500.3$	$250.2 \pm 30.6$	$72.4 \pm 8.9$	$16.3 \pm 2.0$
Background Sum	$43849.0 \pm 509.2$	$1017.9 \pm 33.7$	$330.7 \pm 12.1$	$66.2 \pm 4.1$
Data	43902.0	1069.0	367.0	89.0

TABLE E.6. QCD and total background yields for the low mass (m700) selection after freezing the selection cuts used in the B and D regions after requiring  $p_T^{bb} > 210$  GeV.

High Mass Selection in $m_{b\bar{b}}$ Control Region, B/D Frozen after $p_T^{WW} > 250$ GeV				
Sample	bbpt350	wwpt250	drww15	hh2000
QCD	$1538.7 \pm 252.7$	$1359.5 \pm 75.9$	$392.7 \pm 21.9$	$20.7 \pm 1.2$
Background Sum	$14719.1 \pm 258.9$	$12463.5 \pm 91.8$	$3577.5 \pm 35.0$	$238.9 \pm 7.2$
Data	14862.0	12450.0	3761.0	250.0

TABLE E.7. QCD and total background yields for the low mass selection after freezing the selection cuts used in the B and D regions after requiring  $p_T^{WW} > 250$  GeV.

<sup>2470</sup> After freezing, the yields in the non-resonant  $m_{b\bar{b}}$  control region are monotonically  
<sup>2471</sup> decreasing as expected, and the absolute statistical error on the QCD estimate in the

<sup>2472</sup> final signal region is significantly reduced compared to the yields obtained without  
<sup>2473</sup> freezing the B and D regions.

2474 **E.1. QCD Lepton Flavour Composition after Preselection Criteria for  $\sigma_{d0}$**

2475 **distribution**

2476 We estimate the lepton flavour composition in the QCD sample and SM signal  
2477 sample just after the preselection criteria. By using the  $\sigma_{d0}$  distribution we identify  
2478 the lepton flavour and the origin of each lepton. The majority of the muons results  
2479 to be from bottom meson while most of the electrons come from photon conversion  
2480 (about 70%). The impact of the  $\sigma_{d0}$  cut on the two population is mostly independent  
2481 from the lepton flavour, namely the cut  $|\sigma_{d0}| \leq 2.0$  removes for muon leptons  $(76 \pm$   
2482 100)% for QCD sample and  $(6.9 \pm 1.1)\%$  for signal sample, while for the electrons  
2483 these fractions are  $(27.9 \pm 5.0)\%$  for QCD sample and  $(5.7 \pm 0.9)\%$  for signal. The  
2484 Table E.8 shows the complete lepton flavour composition.

Electrons	
Photon Conversion	65.3%
Not Defined	15.0%
Bottom Meson	11.9%
Dalitz Decay	4.6%
Others	$\leq 0.1\%$

Muons	
Bottom Meson	99%
Charm Barion	1%

TABLE E.8. Lepton flavor composition in QCD sample. The lepton origin is reported for each event passing the preselection criteria.

2485

## APPENDIX F

2486

### STABILITY OF $t\bar{t}$ NORMALIZATION IN QCD ESTIMATE

2487

This appendix contains tables summarizing a study performed to test the stability of the normalization factor applied to  $t\bar{t}$  events when calculated iteratively with the ABCD method for QCD estimation. Since both the ABCD method and the  $t\bar{t}$  normalization are calculated using data-driven techniques, and they contaminate each other control region, the aim is to check that the  $t\bar{t}$  normalization and the QCD estimate converge to stable values after a small number of iterations. The iterative procedure begins with the QCD estimate set to zero and continues as follows:

2494

- Calculate  $t\bar{t}$  NF at appropriate cut **before** QCD is calculated

2495

- Start at beginning of cutflow, apply  $t\bar{t}$  NF, and calculate QCD

2496

- Move through cutflow to  $t\bar{t}$  NF cut

2497

- Calculate new  $t\bar{t}$  NF

2498

- Repeat steps 2 - 4

2499

If the procedure converges, the estimate is said to be stable. The following tables show the iteration for the non-resonant, low mass, and high mass selections. Iteration 0 starts with zero QCD in the background estimate. Yields are provided for QCD and  $t\bar{t}$  after the first selection cut (noted as sample\_first) and in the final  $m_{b\bar{b}}$  signal region (noted as sample\_SR). The QCD yield at the  $t\bar{t}$  normalization cut is also provided. The final column shows the calculated normalization factor for  $t\bar{t}$ . For all selections, convergence (change in  $t\bar{t}$  normalization <5%) is observed after just a few iterations.

2506

Iteration	nQCD_first	n_QCD_nf	nQCD_SR	ttbar_first	ttbar_SR	ttbarNF
0	0.00	0.00	0.00	24637.64	65.92	$1.43 \pm 0.02$
1	4226.57	273.95	3.30	35249.05	94.31	$0.92 \pm 0.02$
2	15765.74	1381.06	19.04	22708.20	60.75	$0.94 \pm 0.03$
3	15311.64	1335.16	18.33	23145.35	61.92	$0.94 \pm 0.03$
4	15254.28	1329.37	18.24	23200.59	62.07	$0.94 \pm 0.03$
5	15247.34	1328.67	18.23	23207.28	62.09	$0.94 \pm 0.03$
6	15246.50	1328.58	18.23	23208.09	62.09	$0.94 \pm 0.03$
7	15246.40	1328.57	18.23	23208.18	62.09	$0.94 \pm 0.03$

TABLE F.1. Iteration of  $t\bar{t}$  normalization and QCD estimate for the non-resonant selection. The 'first' region is defined after the  $m_{WW} < 130$  GeV cut and the 'nf' region is defined after requiring  $p_T^{bb} > 150$  GeV. Convergence is observed after two iterations.

Iteration	nQCD_first	n_QCD_nf	nQCD_SR	ttbar_first	ttbar_SR	ttbarNF
0	0.00	0.00	0.00	24637.64	11.28	$1.43 \pm 0.02$
1	4226.57	273.95	1.13	35249.05	16.14	$0.92 \pm 0.02$
2	15765.74	1381.06	8.14	22708.20	10.40	$0.94 \pm 0.03$
3	15311.64	1335.16	7.81	23145.35	10.60	$0.94 \pm 0.03$
4	15254.28	1329.37	7.76	23200.59	10.62	$0.94 \pm 0.03$
5	15247.34	1328.67	7.76	23207.28	10.63	$0.94 \pm 0.03$
6	15246.50	1328.58	7.76	23208.09	10.63	$0.94 \pm 0.03$
7	15246.40	1328.57	7.76	23208.18	10.63	$0.94 \pm 0.03$

TABLE F.2. Iteration of  $t\bar{t}$  normalization and QCD estimate for the low mass selection. The 'first' region is defined after the  $m_{WW} < 130$  GeV cut and the 'nf' region is defined after requiring  $p_T^{bb} > 150$  GeV. Convergence is observed after two iterations.

Iteration	nQCD_first	n_QCD_nf	nQCD_SR	ttbar_first	ttbar_SR	ttbarNF
0	0.00	0.00	0.00	8810.24	6.70	$1.26 \pm 0.02$
1	92.73	40.56	0.11	11133.25	8.47	$0.99 \pm 0.02$
2	1595.54	541.28	1.40	8679.56	6.60	$1.03 \pm 0.03$
3	1238.18	432.83	1.12	9063.72	6.90	$1.03 \pm 0.02$
4	1196.43	419.79	1.09	9108.80	6.93	$1.03 \pm 0.02$
5	1190.60	417.96	1.09	9115.10	6.94	$1.03 \pm 0.02$
6	1189.77	417.70	1.09	9115.99	6.94	$1.03 \pm 0.02$
7	1189.65	417.66	1.09	9116.12	6.94	$1.03 \pm 0.02$

TABLE F.3. Iteration of  $t\bar{t}$  normalization and QCD estimate for the low mass selection. The 'first' region is defined after the  $p_T^{bb} > 350$  GeV cut and the 'nf' region is defined after requiring  $p_T^{WW} > 360$  GeV. Convergence is observed after two iterations.

## FULLY BOOSTED SYSTEMATIC UNCERTAINTIES

Uncertainty	Up/Down
SysFT_EFF_Eigen_Light_0_AntiKt2PV0TrackJets_-1down	-12.9/12.5
SysFT_EFF_Eigen_C_0_AntiKt2PV0TrackJets_-1down	-12.6/12.1
SysFT_EFF_Eigen_C_0_AntiKt2PV0TrackJets_-1up	11.3/-11.9
SysFT_EFF_Eigen_Light_0_AntiKt2PV0TrackJets_-1up	11.3/-11.9
SysFATJET_Medium_JET_Comb_Baseline_Kin_-1up	-6.47/5.95
SysFATJET_Medium_JET_Comb_Baseline_Kin_-1down	5.83/-6.43
SysFT_EFF_Eigen_B_0_AntiKt2PV0TrackJets_-1down	-3.49/2.97
SysFT_EFF_Eigen_B_1_AntiKt2PV0TrackJets_-1down	-3.32/2.8
SysFT_EFF_Eigen_C_1_AntiKt2PV0TrackJets_-1down	-2.97/2.45
SysFT_EFF_Eigen_B_0_AntiKt2PV0TrackJets_-1up	2.39/-2.94
SysFT_EFF_Eigen_B_1_AntiKt2PV0TrackJets_-1up	2.23/-2.78
SysFT_EFF_Eigen_C_1_AntiKt2PV0TrackJets_-1up	1.9/-2.45
SysFATJET_Medium_JET_Comb_Tracking_Kin_-1down	1.77/-2.38
SysFATJET_Medium_JET_Comb_Tracking_Kin_-1up	-2.2/1.68
SysFT_EFF_extrapolation_AntiKt2PV0TrackJets_-1up	-2.07/1.59
SysFATJET_JMR_-1up	1.41/-1.97
SysFT_EFF_extrapolation_from_charm_AntiKt2PV0TrackJets_-1up	-1.62/1.1
SysFT_EFF_extrapolation_AntiKt2PV0TrackJets_-1down	0.963/-1.54
SysPRW_DATASF_-1down	-1.47/0.94
SysJET_SR1_JET_GroupedNP_1_-1down	0.764/-1.33
SysJET_SR1_JET_GroupedNP_1_-1up	-1.2/0.678

SysFATJET_JER_1up	0.574/-1.16
SysFT_EFF_Eigen_Light_1_AntiKt2PV0TrackJets_1up	-1.12/0.576
SysFT_EFF_extrapolation_from_charm_AntiKt2PV0TrackJets_1down	0.553/-1.1
SysFATJET_Medium_JET_Comb_TotalStat_Kin_1down	-1.03/0.498
SysFT_EFF_Eigen_C_2_AntiKt2PV0TrackJets_1down	-0.983/0.458
SysMET_SoftTrk_Scale_1down	-0.971/0.439
SysJET_JER_SINGLE_NP_1up	-0.936/0.375
SysEL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR_1up	-0.935/0.405
SysFATJET_Medium_JET_Comb_Modelling_Kin_1up	-0.903/0.381
SysMUON_EFF_SYS_1up	-0.88/0.353
SysJET_SR1_JET_GroupedNP_3_1up	-0.871/0.346
SysJET_SR1_JET_GroupedNP_2_1up	-0.826/0.289
SysFT_EFF_Eigen_Light_2_AntiKt2PV0TrackJets_1down	-0.825/0.301
SysJET_SR1_JET_EtaIntercalibration_NonClosure_1up	-0.792/0.261
SysEL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR_1up	-0.78/0.25
SysFT_EFF_Eigen_C_3_AntiKt2PV0TrackJets_1up	-0.767/0.233
SysFT_EFF_Eigen_Light_3_AntiKt2PV0TrackJets_1up	-0.762/0.229
SysEL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR_1up	-0.71/0.179
SysFT_EFF_Eigen_B_2_AntiKt2PV0TrackJets_1up	-0.705/0.173
SysMUON_ID_1down	-0.698/0.166
SysFT_EFF_Eigen_B_2_AntiKt4EMTopoJets_1down	-0.662/0.13
SysFT_EFF_extrapolation_AntiKt4EMTopoJets_1up	-0.651/0.119
SysFT_EFF_Eigen_B_0_AntiKt4EMTopoJets_1down	-0.637/0.107
SysEL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR_1up	-0.611/0.0791
SysMUON_EFF_STAT_1up	-0.604/0.0725
SysFT_EFF_Eigen_B_1_AntiKt4EMTopoJets_1down	-0.593/0.0613

SysMUON_ID_1up	-0.588/0.0559
SysFT_EFF_Eigen_Light_2_AntiKt4EMTopoJets_1down	-0.582/0.0499
SysMET_SoftTrk_Scale_1up	0.0485/-0.582
SysMUON_ISO_SYS_1up	-0.576/0.0453
SysFT_EFF_Eigen_Light_0_AntiKt4EMTopoJets_1down	-0.572/0.0358
SysFT_EFF_Eigen_Light_1_AntiKt2PV0TrackJets_1down	0.0513/-0.571
SysFT_EFF_Eigen_Light_1_AntiKt4EMTopoJets_1up	-0.567/0.0359
SysMUON_MS_1up	-0.558/0.0258
SysFT_EFF_Eigen_C_1_AntiKt4EMTopoJets_1up	-0.556/0.024
SysFT_EFF_Eigen_Light_3_AntiKt4EMTopoJets_1down	-0.555/0.0236
SysFT_EFF_Eigen_Light_4_AntiKt2PV0TrackJets_1down	-0.551/0.0151
SysFT_EFF_Eigen_C_0_AntiKt4EMTopoJets_1up	-0.546/0.0135
SysMUON_ISO_STAT_1up	-0.545/0.0136
SysFT_EFF_Eigen_C_3_AntiKt4EMTopoJets_1up	-0.543/0.0112
SysFT_EFF_Eigen_Light_4_AntiKt4EMTopoJets_1down	-0.54/0.00864
SysFT_EFF_Eigen_C_2_AntiKt4EMTopoJets_1up	-0.539/0.00807
SysFT_EFF_extrapolation_from_charm_AntiKt4EMTopoJets_1down	-0.537/0.00495
SysLEPTON_d0Eff_Acc_1down	-0.532/0
SysLEPTON_d0Eff_Acc_1up	-0.532/0
SysMODEL_DIBOSONS_Norm_Inc_Acc_1down	-0.532/0
SysMODEL_DIBOSONS_Norm_Inc_Acc_1up	-0.532/0
SysMODEL_DIBOSONS_Norm_Inc_XS_1down	-0.532/0
SysMODEL_DIBOSONS_Norm_Inc_XS_1up	-0.532/0
SysMODEL_QCD_Norm_Inc_1down	-0.532/0
SysMODEL_QCD_Norm_Inc_1up	-0.532/0
SysMODEL_STOP_Norm_Inc_Acc_1down	-0.532/0

SysMODEL_STOP_Norm_Inc_Acc_1up	-0.532/0
SysMODEL_STOP_Norm_Inc_XS_1down	-0.532/0
SysMODEL_STOP_Norm_Inc_XS_1up	-0.532/0
SysMODEL_TTBar_Norm_Inc_Acc_1down	-0.532/0
SysMODEL_TTBar_Norm_Inc_Acc_1up	-0.532/0
SysMODEL_TTBar_Norm_Inc_XS_1down	-0.532/0
SysMODEL_TTBar_Norm_Inc_XS_1up	-0.532/0
SysMODEL_WJETS_Norm_Inc_Acc_1down	-0.532/0
SysMODEL_WJETS_Norm_Inc_Acc_1up	-0.532/0
SysMODEL_WJETS_Norm_Inc_XS_1down	-0.532/0
SysMODEL_WJETS_Norm_Inc_XS_1up	-0.532/0
SysMODEL_ZJETS_Norm_Inc_Acc_1down	-0.532/0
SysMODEL_ZJETS_Norm_Inc_Acc_1up	-0.532/0
SysMODEL_ZJETS_Norm_Inc_XS_1down	-0.532/0
SysMODEL_ZJETS_Norm_Inc_XS_1up	-0.532/0
SysMUON_SCALE_1down	-0.532/0
SysMODEL_TTBar_Shape_hhMass_SR_Shower_1up	-0.532/0
SysMODEL_TTBar_Shape_hhMass_SR_Shower_1down	-0.532/0
SysMODEL_WJETS_Norm_mbbar_XS_1up	-0.532/0
SysMODEL_WJETS_Norm_mbbar_XS_1down	-0.532/0
SysMODEL_ZJETS_Norm_SR_XS_1up	-0.532/0
SysMODEL_ZJETS_Norm_SR_XS_1down	-0.532/0
SysMODEL_ZJETS_Norm_SR_Acc_1up	-0.532/0
SysMODEL_ZJETS_Norm_SR_Acc_1down	-0.532/0
SysMODEL_WJETS_Shape_hhMass_SR_Scale_1up	-0.532/0
SysMODEL_WJETS_Shape_hhMass_SR_Scale_1down	-0.532/0

SysMODEL_DIBOSONS_Norm_mbbcr_XS_1up	-0.532/0
SysMODEL_DIBOSONS_Norm_mbbcr_XS_1down	-0.532/0
SysMODEL_DIBOSONS_Norm_mbbcr_Acc_1up	-0.532/0
SysMODEL_DIBOSONS_Norm_mbbcr_Acc_1down	-0.532/0
SysMODEL_QCD_Norm_mbbcr_1up	-0.532/0
SysMODEL_QCD_Norm_mbbcr_1down	-0.532/0
SysMODEL_STOP_Norm_mbbcr_Acc_1up	-0.532/0
SysMODEL_STOP_Norm_mbbcr_Acc_1down	-0.532/0
SysMODEL_STOP_Norm_mbbcr_XS_1up	-0.532/0
SysMODEL_STOP_Norm_mbbcr_XS_1down	-0.532/0
SysMODEL_TTBar_Norm_mbbcr_Acc_1up	-0.532/0
SysMODEL_TTBar_Norm_mbbcr_Acc_1down	-0.532/0
SysMODEL_TTBar_Norm_mbbcr_XS_1up	-0.532/0
SysMODEL_TTBar_Norm_mbbcr_XS_1down	-0.532/0
SysMODEL_WJETS_Norm_mbbcr_Acc_1up	-0.532/0
SysMODEL_WJETS_Norm_mbbcr_Acc_1down	-0.532/0
SysMODEL_ZJETS_Norm_mbbcr_Acc_1up	-0.532/0
SysMODEL_ZJETS_Norm_mbbcr_Acc_1down	-0.532/0
ZJETS_Norm_mbbcr_XS_1up	-0.532/0
ZJETS_Norm_mbbcr_XS_1down	-0.532/0
SysMODEL_DIBOSONS_Norm_SR_Acc_1up	-0.532/0
SysMODEL_DIBOSONS_Norm_SR_Acc_1down	-0.532/0
SysMODEL_DIBOSONS_Norm_SR_XS_1up	-0.532/0
SysMODEL_DIBOSONS_Norm_SR_XS_1down	-0.532/0
SysMODEL_QCD_Norm_SR_1up	-0.532/0
SysMODEL_QCD_Norm_SR_1down	-0.532/0

SysMODEL_QCD_Shape_hhMass_SR_1up	-0.532/0
SysMODEL_QCD_Shape_hhMass_SR_1down	-0.532/0
SysMODEL_STOP_Norm_SR_Acc_1up	-0.532/0
SysMODEL_STOP_Norm_SR_Acc_1down	-0.532/0
SysMODEL_STOP_Norm_SR_XS_1up	-0.532/0
SysMODEL_STOP_Norm_SR_XS_1down	-0.532/0
SysMODEL_TTBar_Norm_SR_Acc_1up	-0.532/0
SysMODEL_TTBar_Norm_SR_Acc_1down	-0.532/0
SysMODEL_TTBar_Norm_SR_XS_1up	-0.532/0
SysMODEL_TTBar_Norm_SR_XS_1down	-0.532/0
SysMODEL_TTBar_Shape_hhMass_SR_Gen_1up	-0.532/0
SysMODEL_TTBar_Shape_hhMass_SR_Gen_1down	-0.532/0
SysMODEL_TTBar_Shape_hhMass_SR_RadLo_1up	-0.532/0
SysMODEL_TTBar_Shape_hhMass_SR_RadLo_1down	-0.532/0
SysMODEL_WJETS_Norm_SR_Acc_1up	-0.532/0
SysMODEL_WJETS_Norm_SR_Acc_1down	-0.532/0
SysMODEL_WJETS_Norm_SR_XS_1up	-0.532/0
SysMODEL_WJETS_Norm_SR_XS_1down	-0.532/0
SysFT_EFF_extrapolation_from_charm_AntiKt4EMTopoJets_1up	-0.527/-0.00495
SysJET_JvtEfficiency_1up	-0.526/-0.00613
SysJET_SR1_JET_GroupedNP_2_1down	-0.0414/-0.525
SysFT_EFF_Eigen_C_2_AntiKt4EMTopoJets_1down	-0.524/-0.00814
SysFT_EFF_Eigen_Light_4_AntiKt4EMTopoJets_1up	-0.523/-0.00865
SysFT_EFF_Eigen_C_3_AntiKt4EMTopoJets_1down	-0.52/-0.0113
SysMUON_ISO_STAT_1down	-0.518/-0.0136
SysFT_EFF_Eigen_C_0_AntiKt4EMTopoJets_1down	-0.518/-0.0125

SysMUON_SAGITTA_RESBIAS_1up	-0.0125/-0.516
SysFT_EFF_Eigen_Light_4_AntiKt2PV0TrackJets_1up	-0.513/-0.015
SysFT_EFF_Eigen_Light_3_AntiKt4EMTopoJets_1up	-0.508/-0.0235
SysFT_EFF_Eigen_C_1_AntiKt4EMTopoJets_1down	-0.508/-0.0239
SysMET_SoftTrk_ResoPara_1up	-0.505/-0.0311
SysFT_EFF_Eigen_Light_0_AntiKt4EMTopoJets_1up	-0.498/-0.0297
SysFT_EFF_Eigen_Light_1_AntiKt4EMTopoJets_1down	-0.496/-0.0358
SysMUON_ISO_SYS_1down	-0.487/-0.0453
SysFATJET_Medium_JET_Comb_TotalStat_Kin_1up	-0.048/-0.484
SysFT_EFF_Eigen_Light_2_AntiKt4EMTopoJets_1up	-0.482/-0.0499
SysJET_JvtEfficiency_1down	-0.478/-0.0538
SysJET_SR1_JET_EtaIntercalibration_NonClosure_1down	-0.474/-0.0887
SysEG_SCALE_ALL_1down	-0.471/-0.0611
SysMUON_MS_1down	-0.47/-0.0619
SysFT_EFF_Eigen_B_1_AntiKt4EMTopoJets_1up	-0.468/-0.0644
SysMUON_SCALE_1up	-0.462/-0.0696
SysMUON_EFF_STAT_1down	-0.46/-0.0725
SysFT_EFF_Eigen_C_2_AntiKt2PV0TrackJets_1up	-0.0799/-0.459
SysEL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR_1down	-0.453/-0.0791
SysMUON_SAGITTA_RHO_1down	-0.429/-0.106
SysMUON_SAGITTA_RHO_1up	-0.429/-0.106
SysFT_EFF_Eigen_B_0_AntiKt4EMTopoJets_1up	-0.426/-0.107
SysFT_EFF_extrapolation_AntiKt4EMTopoJets_1down	-0.413/-0.119
SysEL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR_1down	-0.129/-0.405
SysMET_SoftTrk_ResoPerp_1up	-0.403/-0.13
SysFT_EFF_Eigen_B_2_AntiKt4EMTopoJets_1up	-0.402/-0.13

SysEG_SCALE_ALL_1up	-0.132/-0.4
SysFATJET_Medium_JET_Comb_Modelling_Kin_1down	-0.181/-0.384
SysJET_SR1_JET_GroupedNP_3_1down	-0.369/-0.198
SysFT_EFF_Eigen_B_2_AntiKt2PV0TrackJets_1down	-0.358/-0.173
SysEL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR_1down	-0.353/-0.179
SysMUON_EFF_SYS_1down	-0.185/-0.352
SysMUON_SAGITTA_RESBIAS_1down	-0.335/-0.2
SysPRW_DATASF_1up	-0.195/-0.322
SysFT_EFF_Eigen_Light_3_AntiKt2PV0TrackJets_1down	-0.301/-0.229
SysFT_EFF_Eigen_Light_2_AntiKt2PV0TrackJets_1up	-0.24/-0.3
SysFT_EFF_Eigen_C_3_AntiKt2PV0TrackJets_1down	-0.296/-0.233
SysEL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR_1down	-0.284/-0.25
Total Up	27.2
Total Do	27.9

TABLE G.1. List of systematic uncertainties for the fully boosted analysis

## Bibliography

- 2509 [1] Morad Aaboud et al. “Search for Higgs boson pair production in the  $b\bar{b}WW^*$   
2510 decay mode at  $\sqrt{s} = 13$  TeV with the ATLAS detector”. In: (2018). arXiv:  
2511 1811.04671 [hep-ex].
- 2512 [2] M. Tanabashi et al. “Review of Particle Physics”. In: *Phys. Rev. D* 98 (3 Aug.  
2513 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: <https://link.aps.org/doi/10.1103/PhysRevD.98.030001>.
- 2515 [3] Ian Snyder. “SEARCH FOR A SCALAR PARTNER OF THE TOP QUARK  
2516 IN THE JETS+Missing ET T FINAL STATE IN PROTON-PROTON  
2517 COLLISIONS AT  $Sqrt(s) = 13$  TEV WITH THE ATLAS DETECTOR”.  
2518 PhD thesis. University of Oregon, 2018.
- 2519 [4] John Ellis. “Higgs Physics”. In: arXiv:1312.5672. KCL-PH-TH-2013-49. LCTS-  
2520 2013-36. CERN-PH-TH-2013-315 (Dec. 2013). 52 pages, 45 figures, Lectures  
2521 presented at the ESHEP 2013 School of High-Energy Physics, to appear as  
2522 part of the proceedings in a CERN Yellow Report, 117–168. 52 p. URL: <http://cds.cern.ch/record/1638469>.
- 2524 [5] Peter W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In:  
2525 *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.  
2526 13.508. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.508>.
- 2527 [6] Georges Aad et al. “Observation of a new particle in the search for the Standard  
2528 Model Higgs boson with the ATLAS detector at the LHC”. In: *Phys. Lett.* B716  
2529 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214  
2530 [hep-ex].

- 2531 [7] Georges Aad et al. “Evidence for the spin-0 nature of the Higgs boson using  
2532 ATLAS data”. In: *Phys. Lett.* B726 (2013), pp. 120–144. DOI: 10.1016/j.  
2533 physletb.2013.08.026. arXiv: 1307.1432 [hep-ex].
- 2534 [8] Morad Aaboud et al. “Observation of  $H \rightarrow b\bar{b}$  decays and  $VH$  production with  
2535 the ATLAS detector”. In: *Phys. Lett.* B786 (2018), pp. 59–86. DOI: 10.1016/  
2536 j.physletb.2018.09.013. arXiv: 1808.08238 [hep-ex].
- 2537 [9] Morad Aaboud et al. “Cross-section measurements of the Higgs boson decaying  
2538 into a pair of  $\tau$ -leptons in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the  
2539 ATLAS detector”. In: *Submitted to: Phys. Rev.* (2018). arXiv: 1811.08856  
2540 [hep-ex].
- 2541 [10] R. Belusevic and G. Jikia. “Higgs selfcoupling in gamma gamma collisions”. In:  
2542 *Phys. Rev.* D70 (2004), p. 073017. DOI: 10.1103/PhysRevD.70.073017. arXiv:  
2543 hep-ph/0403303 [hep-ph].
- 2544 [11] Albert M Sirunyan et al. “Combination of searches for Higgs boson pair  
2545 production in proton-proton collisions at  $\sqrt{s} = 13$  TeV”. In: *Submitted to:*  
2546 *Phys. Rev. Lett.* (2018). arXiv: 1811.09689 [hep-ex].
- 2547 [12] Aleksandr Azatov et al. “Effective field theory analysis of double Higgs  
2548 production via gluon fusion”. In: *Physical Review D* 92 (Feb. 2015). DOI: 10.  
2549 1103/PhysRevD.92.035001.
- 2550 [13] S. Dawson and M. Sullivan. “Enhanced di-Higgs boson production in the  
2551 complex Higgs singlet model”. In: *Phys. Rev. D* 97 (1 Jan. 2018), p. 015022.  
2552 DOI: 10.1103/PhysRevD.97.015022. URL: <https://link.aps.org/doi/10.1103/PhysRevD.97.015022>.

- 2554 [14] Ian M. Lewis and Matthew Sullivan. “Benchmarks for Double Higgs Production  
2555 in the Singlet Extended Standard Model at the LHC”. In: *Phys. Rev.* D96.3  
2556 (2017), p. 035037. DOI: 10.1103/PhysRevD.96.035037. arXiv: 1701.08774  
2557 [hep-ph].
- 2558 [15] “Linear accelerator 2”. In: (Sept. 2012). URL: <https://cds.cern.ch/record/1997427>.
- 2560 [16] Cinzia De Melis. “The CERN accelerator complex. Complexe des accélérateurs  
2561 du CERN”. In: (Jan. 2016). General Photo. URL: <https://cds.cern.ch/record/2119882>.
- 2563 [17] K.H. Reich. “The CERN Proton Synchrotron Booster”. In: (Jan. 1969). Round  
2564 Table On Booster Injectors. URL: [http://accelconf.web.cern.ch/Accelconf/p69/PDF/PAC1969\\_0959.PDF](http://accelconf.web.cern.ch/Accelconf/p69/PDF/PAC1969_0959.PDF).
- 2566 [18] Giorgio Brianti. “The CERN synchrotrons”. In: (1997). URL: <https://cds.cern.ch/record/340514>.
- 2568 [19] Lyndon Evans and Philip Bryant. “LHC Machine”. In: *Journal of Instrumentation*  
2569 3.08 (Aug. 2008), S08001–S08001. DOI: 10.1088/1748-0221/3/08/s08001.  
2570 URL: <https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08001>.
- 2571 [20] Xabier Cid Vidal Ramon Cid Manzano. *Lorentz Force Taking a closer look at  
2572 LHC*. URL: [https://www.lhc-closer.es/taking\\_a\\_closer\\_look\\_at\\_lhc/0.lorentz\\_force](https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.lorentz_force).
- 2574 [21] Morad Aaboud et al. “Luminosity determination in pp collisions at  $\sqrt{s} = 8$  TeV  
2575 using the ATLAS detector at the LHC”. In: *Eur. Phys. J.* C76.12 (2016), p. 653.  
2576 DOI: 10.1140/epjc/s10052-016-4466-1. arXiv: 1608.03953 [hep-ex].
- 2577 [22] Genesis Perez. “Unitarization Models For Vector Boson Scattering at the  
2578 LHC”. PhD thesis. Jan. 2018. DOI: 10.5445/IR/1000082199.

- 2579 [23] The ATLAS Collaboration. “The ATLAS Experiment at the CERN  
2580 Large Hadron Collider”. In: *Journal of Instrumentation* 3.08 (Aug. 2008),  
2581 S08003–S08003. DOI: 10.1088/1748-0221/3/08/s08003. URL: <https://doi.org/10.1088/1748-0221/3/08/s08003>.
- 2582  
2583 [24] M Capeans et al. *ATLAS Insertable B-Layer Technical Design Report*. Tech. rep.  
2584 CERN-LHCC-2010-013. ATLAS-TDR-19. Sept. 2010. URL: <https://cds.cern.ch/record/1291633>.
- 2585  
2586 [25] *ATLAS inner detector: Technical Design Report, 1*. Technical Design Report  
2587 ATLAS. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/331063>.
- 2588 [26] Boris Dolgoshein. “Transition radiation detectors”. In: *Nuclear Instruments and*  
2589 *Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors*  
2590 *and Associated Equipment* 326.3 (1993), pp. 434–469. ISSN: 0168-9002. DOI:  
2591 [https://doi.org/10.1016/0168-9002\(93\)90846-A](https://doi.org/10.1016/0168-9002(93)90846-A). URL: <http://www.sciencedirect.com/science/article/pii/016890029390846A>.
- 2592  
2593 [27] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design  
2594 Report ATLAS. Geneva: CERN, 1996. URL: <https://cds.cern.ch/record/331061>.
- 2595  
2596 [28] *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report  
2597 ATLAS. Geneva: CERN, 1996. URL: <https://cds.cern.ch/record/331062>.
- 2598 [29] Georges Aad et al. “Electron and photon energy calibration with the ATLAS  
2599 detector using LHC Run 1 data”. In: *Eur. Phys. J.* C74.10 (2014), p. 3071. DOI:  
2600 [10.1140/epjc/s10052-014-3071-4](https://doi.org/10.1140/epjc/s10052-014-3071-4). arXiv: 1407.5063 [hep-ex].
- 2601 [30] Ana Maria Henriques Correia. *The ATLAS Tile Calorimeter*. Tech. rep. ATL-  
2602 TILECAL-PROC-2015-002. Geneva: CERN, Mar. 2015. URL: <https://cds.cern.ch/record/2004868>.

- 2604 [31] *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report  
2605 ATLAS. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/331068>.
- 2606 [32] Jeremy R Love. “A Search for Technicolor at The Large Hadron Collider”.  
2607 PhD thesis. Boston U., 2011-12-12.
- 2608 [33] *ATLAS magnet system: Technical Design Report, 1*. Technical Design Report  
2609 ATLAS. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/338080>.
- 2610 [34] ATLAS Outreach. “ATLAS Fact Sheet : To raise awareness of the ATLAS  
2611 detector and collaboration on the LHC”. 2010. URL: <https://cds.cern.ch/record/1457044>.
- 2613 [35] Aranzazu Ruiz-Martinez and ATLAS Collaboration. *The Run-2 ATLAS Trigger  
2614 System*. Tech. rep. ATL-DAQ-PROC-2016-003. Geneva: CERN, Feb. 2016. URL:  
2615 <https://cds.cern.ch/record/2133909>.
- 2616 [36] Martin zur Nedden. *The Run-2 ATLAS Trigger System: Design, Performance  
2617 and Plan*. Tech. rep. ATL-DAQ-PROC-2016-039. Geneva: CERN, Dec. 2016.  
2618 URL: <https://cds.cern.ch/record/2238679>.
- 2619 [37] Chaowaroj Wanotayaroj and Jim Brau. “Search for a Scalar Partner of the Top  
2620 Quark in the Jets+MET Final State with the ATLAS detector”. Presented 25  
2621 Oct 2016. Nov. 2016. URL: <https://cds.cern.ch/record/2242196>.
- 2622 [38] S. Agostinelli et al. “GEANT4: A Simulation toolkit”. In: *Nuclear Instruments  
2623 and Methods in Physics Research Section A: Accelerators, Spectrometers,  
2624 Detectors and Associated Equipment* (2003).
- 2625 [39] Joao Pequenao. “Event Cross Section in a computer generated image of the  
2626 ATLAS detector.” Mar. 2008. URL: <https://cds.cern.ch/record/1096081>.

- 2627 [40] Grigore Tarna. *Electron identification with the ATLAS detector*. Tech. rep. ATL-  
2628 PHYS-PROC-2017-173. Geneva: CERN, Sept. 2017. URL: <https://cds.cern.ch/record/2286383>.
- 2630 [41] Georges Aad et al. “Muon reconstruction performance of the ATLAS detector  
2631 in proton-proton collision data at  $\sqrt{s} = 13$  TeV”. In: *Eur. Phys. J.* C76.5  
2632 (2016), p. 292. DOI: [10.1140/epjc/s10052-016-4120-y](https://doi.org/10.1140/epjc/s10052-016-4120-y). arXiv: [1603.05598](https://arxiv.org/abs/1603.05598)  
2633 [hep-ex].
- 2634 [42] Bora Isildak. “Measurement of the differential dijet production cross section in  
2635 proton-proton collisions at  $\sqrt{s} = 7$  tev”. PhD thesis. Bogazici U., 2011. arXiv:  
2636 [1308.6064](https://arxiv.org/abs/1308.6064) [hep-ex].
- 2637 [43] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “The anti- $k_t$  jet clustering  
2638 algorithm”. In: *JHEP* 04 (2008), p. 063. DOI: [10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063). arXiv: [0802.1189](https://arxiv.org/abs/0802.1189) [hep-ph].
- 2640 [44] *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*.  
2641 Tech. rep. ATL-PHYS-PUB-2016-012. Geneva: CERN, June 2016. URL: <https://cds.cern.ch/record/2160731>.
- 2643 [45] Per Ola Hansson Adrian. *The ATLAS b-Jet Trigger*. Tech. rep. arXiv:1111.4190.  
2644 Comments: 4 pages, 6 figures, conference proceedings for PIC2011. Geneva:  
2645 CERN, Nov. 2011. URL: <https://cds.cern.ch/record/1397942>.
- 2646 [46] *Performance of missing transverse momentum reconstruction for the ATLAS  
2647 detector in the first proton-proton collisions at at  $\sqrt{s}= 13$  TeV*. Tech. rep. ATL-  
2648 PHYS-PUB-2015-027. Geneva: CERN, July 2015. URL: <https://cds.cern.ch/record/2037904>.

- 2650 [47] Biagio Di Micco et al. *Search for Higgs boson pair production in the bbWW final*  
2651 *state at  $\sqrt{s} = 13$  TeV with the ATLAS detector.* Tech. rep. ATL-COM-PHYS-  
2652 2016-486. Geneva: CERN, May 2016. URL: <https://cds.cern.ch/record/2151893>.
- 2654 [48] Cigdem Issever et al. *Boosted Analysis: Search for Higgs boson pair production*  
2655 *in the bbWW final state at  $\sqrt{s} = 13$  TeV with the ATLAS detector.* Tech. rep.  
2656 ATL-COM-PHYS-2017-1088. Geneva: CERN, July 2017. URL: <https://cds.cern.ch/record/2276099>.
- 2658 [49] Stefano Frixione, Paolo Nason, and Carlo Oleari. “Matching NLO QCD  
2659 computations with Parton Shower simulations: the POWHEG method”. In:  
2660 *JHEP* 11 (2007), p. 070. DOI: [10.1088/1126-6708/2007/11/070](https://doi.org/10.1088/1126-6708/2007/11/070). arXiv:  
2661 0709.2092 [hep-ph].
- 2662 [50] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. “PYTHIA 6.4  
2663 Physics and Manual”. In: *JHEP* 0605 (2006), p. 026. DOI: [10.1088/1126-6708/2006/05/026](https://doi.org/10.1088/1126-6708/2006/05/026). arXiv: [hep-ph/0603175](https://arxiv.org/abs/hep-ph/0603175) [hep-ph].
- 2665 [51] Peter Zeiler Skands. “Tuning Monte Carlo Generators: The Perugia Tunes”. In:  
2666 *Phys.Rev.* D82 (2010), p. 074018. DOI: [10.1103/PhysRevD.82.074018](https://doi.org/10.1103/PhysRevD.82.074018). arXiv:  
2667 1005.3457 [hep-ph].
- 2668 [52] J. Pumplin et al. “New generation of parton distributions with uncertainties  
2669 from global QCD analysis”. In: *JHEP* 0207 (2002), p. 012. arXiv: [hep-ph/0201195](https://arxiv.org/abs/hep-ph/0201195) [hep-ph].
- 2671 [53] D. J. Lange. “The EvtGen particle decay simulation package”. In: *Nucl.*  
2672 *Instrum. Meth.* A462 (2001), pp. 152–155. DOI: [10.1016/S0168-9002\(01\)00089-4](https://doi.org/10.1016/S0168-9002(01)00089-4).

- 2674 [54] Michal Czakon and Alexander Mitov. “Top++: A Program for the Calculation  
2675 of the Top-Pair Cross-Section at Hadron Colliders”. In: *Compt/ Phys. Commun.*  
2676 (2014), p. 2930. arXiv: 1112.5675 [hep-ph].
- 2677 [55] ATLAS Collaboration. *Comparison of Monte Carlo generator predictions for*  
2678 *gap fraction and jet multiplicity observables in top-antitop events*. Tech. rep.  
2679 ATL-PHYS-PUB-2014-005. Geneva: CERN, May 2014. URL: <https://cds.cern.ch/record/1703034>.
- 2681 [56] Nikolaos Kidonakis. “Next-to-next-to-leading-order collinear and soft gluon  
2682 corrections for t-channel single top quark production”. In: *Phys. Rev.* D83  
2683 (2011), p. 091503. DOI: 10.1103/PhysRevD.83.091503. arXiv: 1103.2792  
2684 [hep-ph].
- 2685 [57] Nikolaos Kidonakis. “Two-loop soft anomalous dimensions for single top quark  
2686 associated production with a W- or H-”. In: *Phys. Rev.* D82 (2010), p. 054018.  
2687 DOI: 10.1103/PhysRevD.82.054018. arXiv: 1005.4451 [hep-ph].
- 2688 [58] T. Gleisberg et al. “Event generation with SHERPA 1.1”. In: *JHEP* 0902 (2009),  
2689 p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].
- 2690 [59] Hung-Liang Lai et al. “New parton distributions for collider physics”. In:  
2691 *Phys. Rev.* D82 (2010), p. 074024. DOI: 10.1103/PhysRevD.82.074024. arXiv:  
2692 1007.2241 [hep-ph].
- 2693 [60] C. Anastasiou et al. “High precision determination of the gluon fusion Higgs  
2694 boson cross-section at the LHC”. In: *JHEP* 05 (2016), p. 058. arXiv: 1602.  
2695 00695 [hep-ph].
- 2696 [61] J. Alwall et al. “The automated computation of tree-level and next-to-  
2697 leading order differential cross sections, and their matching to parton shower

- 2698 simulations”. In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079.  
2699 arXiv: 1405.0301 [hep-ph].
- 2700 [62] R.Frederix et al. “Higgs pair production at the LHC with NLO and parton-  
2701 shower effects”. In: *Phys. Lett.* B732 (2014), p. 142. DOI: 10.1016/j.physletb.  
2702 2014.03.026.
- 2703 [63] G. Degrassi, P. P. Giardino, R. Groeber. “On the two-loop virtual QCD  
2704 corrections to Higgs boson pair production in the Standard Model”. In:  
2705 *Eur.Phys.J.* C76 (2016), p. 411. DOI: 10.1140/epjc/s10052-016-4256-9.
- 2706 [64] S. Borowka et al. “Full top quark mass dependence in Higgs boson pair  
2707 production at NLO”. In: *JHEP* 10 (2016), p. 107. DOI: 10.1007/JHEP10(2016)  
2708 107. arXiv: 1608.04798 [hep-ph].
- 2709 [65] S. Borowka et al. “Higgs Boson Pair Production in Gluon Fusion at Next-  
2710 to-Leading Order with Full Top-Quark Mass Dependence”. In: *Phys. Rev. Lett.*  
2711 117.1 (2016). [Erratum: *Phys. Rev. Lett.* 117,no.7,079901(2016)], p. 012001. DOI:  
2712 10.1103/PhysRevLett.117.079901 , 10.1103/PhysRevLett.117.012001.  
2713 arXiv: 1604.06447 [hep-ph].
- 2714 [66] ATLAS Collaboration. “The ATLAS Simulation Infrastructure”. In: *Eur.Phys.J.*  
2715 C70 (2010), pp. 823–874. DOI: 10.1140/epjc/s10052-010-1429-9. arXiv:  
2716 1005.4568 [physics.ins-det].
- 2717 [67] S. Agostinelli et al. “GEANT4: A Simulation toolkit”. In: *Nucl.Instrum.Meth.*  
2718 A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.
- 2719 [68] ATLAS Collaboration. *Improved electron reconstruction in ATLAS using the*  
2720 *Gaussian Sum Filter-based model for bremsstrahlung*. ATLAS-CONF-2012-047.  
2721 May 2012. URL: <https://cds.cern.ch/record/1449796>.

- 2722 [69] ATLAS Collaboration. *Electron efficiency measurements with the ATLAS*  
2723 *detector using the 2015 LHC proton-proton collision data.* ATLAS-CONF-2016-  
2724 024. June 2016. URL: <https://cds.cern.ch/record/2157687>.
- 2725 [70] ATLAS Collaboration. “Electron and photon energy calibration with the  
2726 ATLAS detector using LHC Run 1 data”. In: *Eur. Phys. J.* C74.10 (2014),  
2727 p. 3071. DOI: 10.1140/epjc/s10052-014-3071-4. arXiv: 1407.5063.
- 2728 [71] ATLAS Collaboration. *Electron and photon energy calibration with the ATLAS*  
2729 *detector using data collected in 2015 at  $\sqrt{s} = 13$  TeV.* Aug. 2016. URL: <https://cds.cern.ch/record/2203514>.
- 2731 [72] URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MuonPerformance#AnchorReconstruction>.
- 2733 [73] M. Cacciari, C. P. Salam and G. Soyez. “The anti- $k_t$  jet clustering algorithm”.  
2734 In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063.
- 2735 [74] ATLAS Collaboration. “Topological cell clustering in the ATLAS calorimeters  
2736 and its performance in LHC Run 1”. In: (2016). arXiv: 1603.02934 [hep-ex].
- 2737 [75] *Jet mass reconstruction with the ATLAS Detector in early Run 2 data.* Tech. rep.  
2738 ATLAS-CONF-2016-035. Geneva: CERN, July 2016. URL: <https://cds.cern.ch/record/2200211>.
- 2740 [76] *Expected Performance of Boosted Higgs ( $\rightarrow b\bar{b}$ ) Boson Identification with the*  
2741 *ATLAS Detector at  $\sqrt{s} = 13$  TeV.* Tech. rep. ATL-PHYS-PUB-2015-035.  
2742 Geneva: CERN, Aug. 2015. URL: <https://cds.cern.ch/record/2042155>.
- 2743 [77] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “The Catchment Area  
2744 of Jets”. In: *JHEP* 04 (2008), p. 005. DOI: 10.1088/1126-6708/2008/04/005.  
2745 arXiv: 0802.1188 [hep-ph].

- 2746 [78] ATLAS Collaboration. “ Jet energy scale measurements and their systematic  
2747 uncertainties in proton-proton collisions at  $\text{sqrt}(s) = 13$  TeV with the ATLAS  
2748 detector ”. In: (2017). arXiv: 1703.09665 [hep-ex].
- 2749 [79] ATLAS Collaboration. “Performance of pile-up mitigation techniques for jets  
2750 in  $pp$  collisions at  $\sqrt{s} = 8$  TeV using the ATLAS detector”. In: *Eur. Phys.*  
2751 *J.* C76.11 (2016), p. 581. DOI: 10.1140/epjc/s10052-016-4395-z. arXiv:  
2752 1510.03823 [hep-ex].
- 2753 [80] ATLAS Collaboration. *Jet Calibration and Systematic Uncertainties for Jets*  
2754 *Reconstructed in the ATLAS Detector at  $\sqrt{s} = 13$  TeV*. Tech. rep. ATL-PHYS-  
2755 PUB-2015-015. Geneva: CERN, July 2015. URL: <https://cds.cern.ch/record/2037613>.
- 2757 [81] ATLAS Collaboration. “Jet energy measurement and its systematic uncertainty  
2758 in proton-proton collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector”. In: *Eur.*  
2759 *Phys. J.* C75 (2015), p. 17. DOI: 10.1140/epjc/s10052-014-3190-y. arXiv:  
2760 1406.0076 [hep-ex].
- 2761 [82] ATLAS Collaboration. *Calibration of b-tagging using dileptonic top pair events*  
2762 *in a combinatorial likelihood approach with the ATLAS experiment*. Tech. rep.  
2763 ATLAS-CONF-2014-004. Geneva: CERN, Feb. 2014. URL: <http://cds.cern.ch/record/1664335>.
- 2765 [83] A.L. Read. “Presentation of search results: the  $\text{CL}_s$  technique”. In: *J. Phys.*  
2766 G28 (2002), p. 2693. DOI: 10.1088/0954-3899/28/10/313.
- 2767 [84] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new  
2768 physics”. In: *Eur. Phys. J. C* 71 (2011), p. 1554. DOI: 10.1140/epjc/s10052-  
2769 011-1554-0. arXiv: 1603.00385 [hep-ph].