

¹ Search for Dark Matter Produced in pp Collisions with the ATLAS Detector

² by

³ Danika MacDonell

⁴ B.Sc., University of Victoria, 2016

⁵ M.Sc., University British Columbia, 2018

⁶ A Dissertation Submitted in Partial Fulfillment of the
⁷ Requirements for the Degree of

⁸ DOCTOR OF PHILOSOPHY

⁹ in the Department of Physics and Astronomy

¹⁰ © Danika MacDonell, 2022
¹¹ University of Victoria

¹² All rights reserved. This Dissertation may not be reproduced in whole or in part, by
¹³ photocopying or other means, without the permission of the author.

¹⁴ Search for Dark Matter Produced in pp Collisions with the ATLAS Detector

¹⁵ by

¹⁶ Danika MacDonell

¹⁷ B.Sc., University of Victoria, 2016

¹⁸ M.Sc., University British Columbia, 2018

¹⁹ Supervisory Committee

²⁰ _____
²¹ Dr. R. Kowalewski, Supervisor

²² (Department of Physics and Astronomy)

²³ _____
²⁴ Dr. R. Sobie, Supervisor

²⁵ (Department of Physics and Astronomy)

²⁶ _____
²⁷ Dr. M. Roney, Departmental Member

²⁸ (Department of Physics and Astronomy)

²⁹ _____
³⁰ Dr. I. Paci, Outside Member

³¹ (Department of Chemistry)

³²

³³ **Supervisory Committee**

³⁴

³⁵ Dr. R. Kowalewski, Supervisor
³⁶ (Department of Physics and Astronomy)

³⁷

³⁸ Dr. R. Sobie, Supervisor
³⁹ (Department of Physics and Astronomy)

⁴⁰

⁴¹ Dr. M. Roney, Departmental Member
⁴² (Department of Physics and Astronomy)

⁴³

⁴⁴ Dr. I. Paci, Outside Member
⁴⁵ (Department of Chemistry)

⁴⁶

47

ABSTRACT

Longstanding evidence from observational astronomy indicates that non-luminous “dark matter” constitutes the majority of all matter in the universe, yet this mysterious form of matter continues to elude experimental detection. This dissertation presents a search for dark matter at the Large Hadron Collider using 139 fb^{-1} of proton-proton collision data at a centre-of-mass energy of $\sqrt{s} = 13 \text{ TeV}$, recorded with the ATLAS detector from 2015 to 2018. The search targets a final state topology in which dark matter is produced from the proton-proton collisions in association with a pair of W bosons, one of which decays to a pair of quarks and the other to a lepton-neutrino pair. The dark matter is expected to pass invisibly through the detector, resulting in an imbalance of momentum in the plane transverse to the beam line. The search is optimized to test the Dark Higgs model, which predicts a signature of dark matter production in association with the emission of a hypothesized new particle referred to as the Dark Higgs boson. The Dark Higgs boson is predicted to decay to a W boson pair via a small mixing with the Standard Model Higgs boson discovered in 2012. Collisions that exhibit the targeted final state topology are selected for the search, and an approximate mass of the hypothetical Dark Higgs boson is reconstructed from the particles in each collision. A search is performed by looking for a deviation between distributions of the reconstructed Dark Higgs boson masses and Standard Model predictions for the selected collisions. The data is found to be consistent with the Standard Model prediction, and the results are used to constrain the parameters of the Dark Higgs model. This search complements and extends the reach of existing searches for the Dark Higgs model by the ATLAS and CMS collaborations.

71

DECLARATION

72 **Analysis contributions**

- 73 ● Initiated and served as analysis co-contact for the mono-s(WW) semileptonic
74 analysis presented in this thesis (2020-present).
 - 75 – Defined, delegated, documented and ensured timely completion of studies and
76 tasks for the analysis.
 - 77 – Liaised with conveners, experts and performance groups to ensure that analysis
78 choices were widely understood and endorsed.
 - 79 – Organized and facilitated regular analysis group meetings.
 - 80 – Delivered regular updates and presentations to communicate the progress of the
81 analysis to ATLAS collaborators.
 - 82 – Facilitated and contributed to all stages of analysis approval within the ATLAS
83 collaboration by presenting at approval meetings, interfacing with conveners,
84 and addressing feedback from reviewers.
- 85 ● Editor and primary developer of the supporting documentation for the publi-
86 cation of the mono-s(WW) semileptonic analysis.
- 87 ● Primary analyst for the mono-s(WW) semileptonic analysis.
 - 88 – Collaborated on the design and optimization of selections used to define signal
89 and control regions for the analysis.
 - 90 – Optimized and implemented the binning strategy in the signal region.
 - 91 – Developed a framework to evaluate and implement all systematic uncertainties
92 considered in the analysis.
 - 93 – Designed and optimized the statistical interpretation of the analysis within the
94 HistFitter [1] framework.
 - 95 – Contributed to the development and release of ATLAS-wide recommendations
96 for the calibration of small-radius ($R = 0.2$) jets, and for the evaluation of
97 their associated systematic uncertainties. $R = 0.2$ jets are used to construct
98 large-radius TAR jets [2] used in the analysis (see Section 5.1.4 for details).
 - 99 – Prepared requests for Monte Carlo samples used by the analysis to model signal
100 and Standard Model background processes, and processed all data and Monte
101 Carlo samples through the full analysis chain.

- 102 ● Contributed to control region studies and the evaluation of theoretical system-
 103 atic uncertainties for the mono-s(WW) analysis in the fully hadronic channel.

104 **Computing infrastructure contributions**

- 105 ● Developed infrastructure to automate the creation and deployment of a kuber-
 106 netes cluster as an ATLAS computing site using cloud computing infrastructure
 107 at the University of Victoria.
 108 ● Implemented and tested tools for cluster federation and resource monitoring.

109 **Analysis preservation contributions**

- 110 ● Analysis preservation contact person for all searches for new physics within the
 111 ATLAS collaboration (Feb. 2020 to Oct. 2021).
 112 – Provided technical assistance and liaison to support analysis teams with the de-
 113 velopment of automated analysis preservation workflows in the RECAST frame-
 114 work [3].
 115 – Maintained central documentation of RECAST tools.
 116 – Reviewed RECAST workflows for completeness, and ensured centralized storage
 117 of workflows.
 118 – Organized hands-on training events to familiarize analysts with the tools in-
 119 volved with analysis preservation.
 120 – Facilitated the integration of the REANA data analysis platform [4] as a cen-
 121 tral tool for the development and execution of RECAST analysis preservation
 122 workflows.
 123 ● Contributed to workshop planning and ran tutorials on analysis preservation
 124 and related tools for numerous ATLAS workshops.
 125 ● Developed and currently maintain the RECAST framework for the ATLAS
 126 E_T^{miss} +jets search published in 2021 [5], which continues to be used regularly
 127 within the collaboration to constrain new models of physics beyond the Stan-
 128 dard Model.
 129 ● Developed the RECAST frameworks for the mono-s(WW) analysis in both the
 130 fully hadronic and semileptonic channels.

¹³¹ Contents

¹³²	Supervisory Committee	ⁱⁱ
¹³³	Abstract	ⁱⁱⁱ
¹³⁴	Table of Contents	^{vii}
¹³⁵	List of Tables	^{xi}
¹³⁶	List of Figures	^{xiv}
¹³⁷	1 Introduction	¹
¹³⁸	1.1 Introduction to the Standard Model	²
¹³⁹	1.1.1 Fermions	³
¹⁴⁰	1.1.2 Bosons	³
¹⁴¹	1.1.3 Collision and Decay Processes at Colliders	⁴
¹⁴²	1.1.4 Unstable Particles	⁵
¹⁴³	1.1.5 Feynman Diagrams	⁶
¹⁴⁴	1.1.6 Mathematical Formulation of the Standard Model	⁷
¹⁴⁵	1.2 Evidence for Dark Matter from Observational Astronomy	¹⁰
¹⁴⁶	1.3 Dark Matter Composition Hypotheses	¹³
¹⁴⁷	1.3.1 Origin and Interactions of Particle Dark Matter	¹⁴
¹⁴⁸	1.4 Dark Matter Search Strategies	¹⁵
¹⁴⁹	1.4.1 Direct Detection	¹⁶
¹⁵⁰	1.4.2 Indirect Detection	¹⁷
¹⁵¹	1.4.3 Searches for DM at Colliders	¹⁷
¹⁵²	1.4.4 Searching for Dark Matter at Particle Accelerators	¹⁹
¹⁵³	1.5 Summary of the Thesis	²²
¹⁵⁴	2 The Dark Higgs Model	²³

155	2.1	Theoretical Motivation for the Dark Higgs Model	23
156	2.1.1	Constraints on Generic Z' Mediator Models	25
157	2.1.2	Implications of a Dark Higgs Portal	26
158	2.2	Model Description	29
159	2.2.1	Free Parameters in the Model	31
160	2.3	Search for the Dark Higgs Model at the LHC	32
161	2.3.1	Dark Higgs Decay Channels	33
162	2.3.2	Completed and Ongoing Searches for the Dark Higgs Model .	33
163	3	Introduction to the LHC and the ATLAS Detector	40
164	3.1	The Parton Model	41
165	3.2	Decay Processes from Parton Collisions	42
166	3.2.1	Branching Fractions and W Boson Decays	42
167	3.3	Detectors at the LHC	43
168	3.4	Introduction to the ATLAS detector	44
169	3.4.1	The Inner Detector	47
170	3.4.2	The Calorimeters	47
171	3.4.3	The Muon Spectrometer	51
172	3.4.4	Missing Transverse Momentum	52
173	3.4.5	The Trigger System	53
174	4	Modelling of Signal and Background Processes	56
175	4.1	Introduction to the Monte Carlo Method	56
176	4.2	Monte Carlo Simulation of Events in the ATLAS Detector .	57
177	4.2.1	Use of Alternative MC Generators	58
178	4.2.2	Weighting and Normalization of MC Simulated Processes .	58
179	4.2.3	Comparing Data and MC Simulation to Search for New Physics	62
180	4.3	Simulation of the DH Signal Model	62
181	4.4	Simulation of SM Background Processes	63
182	4.4.1	Dominant Background Processes	64
183	4.4.2	Sub-dominant Background Processes	66
184	5	Data Reconstruction, Selection and Triggers	68
185	5.1	Object Definitions	68
186	5.1.1	Charged Leptons	68
187	5.1.2	Small-radius anti- k_t ($R = 0.4$) jets	71

188	5.1.3	Resolved W Candidate	73
189	5.1.4	Track-Assisted Reclustered Jets (Merged W Candidate)	74
190	5.1.5	E_T^{miss}	78
191	5.1.6	Overlap Removal	80
192	5.1.7	Dark Higgs Candidate Mass	81
193	5.1.8	Transverse Mass	85
194	5.2	Event Selections	86
195	5.2.1	Kinematic Categories	87
196	5.2.2	Baseline selection	89
197	5.2.3	Signal Region Definition	90
198	5.2.4	Control regions	95
199	5.2.5	Background Yields	109
200	5.3	Triggers	110
201	6	Systematic Uncertainties	118
202	6.1	Experimental Systematics	119
203	6.2	Theoretical Systematics	124
204	6.2.1	Use of Acceptance for Evaluating Theoretical Systematics of the DH Signal Model	125
205	6.2.2	Modelling the Parton Distribution Function	127
206	6.2.3	Strong Coupling Constant α_s	127
207	6.2.4	Renormalization and Factorization Scales (μ_R and μ_F) in QCD	128
208	6.2.5	Matrix Element (ME) Generator Comparison	129
209	6.2.6	Parton Showering (PS)	129
210	6.2.7	Interference Between single-top Wt and $t\bar{t}$ Processes at NLO	133
211	6.2.8	Shifts in Predicted Yields due to Theoretical Systematic Un- certainties	133
212	6.3	Relative Impact of Statistical and Systematic Uncertainties in Analysis Regions	133
213	7	Statistical Framework	141
214	7.1	Likelihood Function	141
215	7.1.1	Poisson Expectation Function	143
216	7.1.2	Nuisance Parameter Nomenclature and Correlations	144
217	7.2	Binning Strategy	144

221	7.3	Fit Setup	147
222	7.3.1	Pruning of Systematics	148
223	7.3.2	Background-only Fit and Signal Region Extrapolation	148
224	7.3.3	Exclusion Hypothesis Test and Limit Setting	151
225	8	Results	154
226	8.1	Background-only Fit	154
227	8.1.1	Pre- and Post-Fit Yields of MC Simulated Background Events	154
228	8.1.2	Nuisance Parameter Pulls and Correlations	155
229	8.2	Comparison of SM Background Expectation and Data in the Signal Region	156
230	8.3	Exclusion of the Dark Higgs Signal Model	157
231	8.3.1	Signal+Background Fit	158
232	8.3.2	Hypothesis Testing and Model Exclusion	161
233	8.3.3	Dependence of Sensitivity on Signal Strength	161
235	9	Conclusion	174
236	A	Kinematic Distributions (N-1) in Signal Regions	176
237	B	Kinematic Distributions in Signal and Control Regions	179
238	B.1	Signal region vs. $W+jets$ control region	179
239	B.2	Signal region vs. $t\bar{t}$ control region	179
240	C	Nuisance Parameter Descriptions for Systematics	188
241	D	Dependence of Sensitivity on Signal Strength	192
242	Bibliography		195

²⁴³ List of Tables

²⁴⁴ Table 2.1	Summary of Dark Higgs decay channels to Standard Model par-	
²⁴⁵	ticles.	³⁴
²⁴⁶ Table 2.2	Summary of decay channels for WW and ZZ pairs and their	
²⁴⁷	branching fractions.	³⁹
²⁴⁸ Table 5.1	TAR jet reconstruction parameters.	⁷⁹
²⁴⁹ Table 5.2	Object priorities and overlap removal criteria for each pair of	
²⁵⁰	physics objects considered in the overlap removal procedure. . .	⁸¹
²⁵¹ Table 5.3	List of selection variables for which the placements of cuts were	
²⁵²	optimized when designing the merged and resolved signal regions.	⁹⁴
²⁵³ Table 5.4	Optimized selection criteria for the signal region in the merged	
²⁵⁴	and resolved categories.	⁹⁵
²⁵⁵ Table 5.5	Cutflow yields after application of preselection cuts for the dom-	
²⁵⁶	inant SM backgrounds in the merged SR.	⁹⁸
²⁵⁷ Table 5.6	Cutflow yields after application of preselection cuts for the dom-	
²⁵⁸	inant SM backgrounds in the resolved SR.	⁹⁹
²⁵⁹ Table 5.7	Cutflow yields for three sample signal points in the merged SR. .	¹⁰⁰
²⁶⁰ Table 5.8	Cutflow yields for three sample signal points in the resolved SR. .	¹⁰¹
²⁶¹ Table 5.9	Summary of control regions.	¹⁰²
²⁶² Table 5.10	Comparison of the $W + \text{jets}$ background yield, total background	
²⁶³	yield, and the composition of the $W + \text{jets}$ background relative to	
²⁶⁴	the total background.	¹⁰⁷
²⁶⁵ Table 5.11	Summary of differences in selections on $N(b\text{-jet})$ and \mathcal{S} between	
²⁶⁶	signal regions and $t\bar{t}$ control regions.	¹⁰⁹
²⁶⁷ Table 5.12	Comparison of the $t\bar{t}$ background yield, total background yield,	
²⁶⁸	and the composition of the $t\bar{t}$ background relative to the total	
²⁶⁹	background.	¹⁰⁹

270	Table 5.13 Background yields after application of all analysis cuts in the merged analysis regions.	110
271		
272	Table 5.14 Background yields after application of all analysis cuts in the resolved analysis regions.	110
273		
274	Table 5.15 Summary of E_T^{miss} triggers from the ATLAS trigger menu used for the search.	112
275		
276	Table 5.16 Summary of single muon triggers from the ATLAS trigger menu used for the search.	113
277		
278	Table 5.17 Summary of single electron triggers from the ATLAS trigger menu used for the study presented in Section 5.3.	115
279		
280	Table 6.1 Sources of experimental systematic uncertainty considered for all physics objects used in the search.	121
281		
282	Table 6.2 Symmetrized uncertainties in the total predicted yields of all SM background processes in the signal regions associated with jet energy resolution.	124
283		
284		
285	Table 6.3 Symmetrized uncertainties in the total predicted yields of all SM background processes in the control regions associated with jet energy resolution.	125
286		
287		
288	Table 6.4 Summary of theoretical uncertainties evaluated for each process considered in the search.	126
289		
290	Table 6.5 Theoretical systematic uncertainties associated with predicted yields of dominant backgrounds in the signal regions.	137
291		
292	Table 6.6 Theoretical systematic uncertainties associated with predicted yields of dominant backgrounds in the control regions.	138
293		
294	Table 6.7 Theoretical systematic uncertainties associated with predicted yields of the DH signal model at several m_s and $m_{Z'}$ in the merged signal region.	139
295		
296		
297	Table 6.8 Theoretical systematic uncertainties associated with predicted yields of the DH signal model at several m_s and $m_{Z'}$ in the re- solved signal region.	139
298		
299		
300	Table 6.9 Comparison of the statistical uncertainties associated with the total predicted yield of SM background events in each analysis region with the total combined statistical+systematic uncertainties.	139
301		
302		

303	Table 6.10 Comparison of the statistical uncertainties associated with the predicted yield of the DH signal process at $(m_s, m_{Z'}) = (210, 2100)$ GeV in each analysis region with the total combined statistical+systematic uncertainties.	140
307	Table 7.1 Summary of naming scheme and correlation information for nuisance parameters in the likelihood function.	145
309	Table 7.2 Binning used in the analysis regions.	147
310	Table 7.3 Comparison of total relative systematic uncertainty associated with the predicted yield of the $W + \text{jets}$ process in signal and control regions, evaluated using either the predicted yield or the transfer factor.	151
314	Table 8.1 Comparison of the total observed yields of events in the signal regions with the predicted yield of Standard Model background processes, either before or after extrapolation of constraints from the background-only fit in the control regions.	157
318	Table C.1 Qualitative description of nuisance parameters associated with sources of experimental systematic uncertainty considered in the search.	188
321	Table C.2 Qualitative description of nuisance parameters associated with sources of theoretical systematic uncertainty considered in the search.	190
323		

³²⁴ List of Figures

325 Figure 1.1 Names and properties of particles in the Standard Model.	2
326 Figure 1.2 Feynman diagram for the Drell Yan process.	6
327 Figure 1.3 Rotation speed of the nearby dwarf galaxy M33.	11
328 Figure 1.4 Summary of upper bounds from direct detection searches on 329 the interaction cross section for spin-independent WIMP-nucleon 330 scattering.	17
331 Figure 1.5 Comparison of the E_T^{miss} + jets process between the EFT and 332 simplified model frameworks.	21
333 Figure 2.1 Leading Feynman diagrams for hypothetical DM production at 334 the LHC by means of the DH model.	24
335 Figure 2.2 Signatures for DM production and detection via a Z' vector boson 336 mediator at the LHC.	25
337 Figure 2.3 Constraints on the masses of the Z' vector boson mediator and 338 DM from combined ATLAS and CMS dijet and monojet searches.	27
339 Figure 2.4 Global scans of the DH model, with Z' and DH as portal mediators, 340 performed by Duerr et al.	28
341 Figure 2.5 Upper bounds on the g_q coupling between the Z' and quarks 342 from recent dijet resonance searches by ATLAS and CMS.	32
343 Figure 2.6 m_s dependence of the branching fraction for decays of the DH 344 boson to SM particles.	34
345 Figure 2.7 Feynman diagrams for the re-interpreted dark matter search to 346 probe the Dark Higgs model in the $s \rightarrow bb$ channel.	37
347 Figure 2.8 Range of $(m_s, m_{Z'})$ in the DH model excluded by searches per- 348 formed by ATLAS and CMS.	39
349 Figure 3.1 Parton distribution functions modelled with MSHT20.	41
350 Figure 3.2 Summary of Standard Model cross sections for particle produc- 351 tion processes measured by the ATLAS detector.	42

352	Figure 3.3 W boson decay mechanisms.	43
353	Figure 3.5 Schematic diagram of the ATLAS detector and its standard coordinate system.	46
355	Figure 4.1 Grid of signal samples produced for the search.	63
356	Figure 4.2 Relative contributions of all SM background processes considered in the signal regions.	64
358	Figure 4.3 Feynman diagrams for $W + \text{jets}$ and $t\bar{t}$ SM background processes.	65
359	Figure 5.1 Decay mechanisms for muons and taus.	69
360	Figure 5.2 Distributions of the reconstructed W candidate mass for MC simulated events produced with the DH signal model over a range of m_s and $m_{Z'}$	74
363	Figure 5.3 Illustration of a final state with TAR-lepton overlap.	76
364	Figure 5.4 Distributions of $m^{\text{TAR Jet}}$ at several representative m_s and $m_{Z'}$, with and without application of the lepton disentanglement pre-selection.	77
367	Figure 5.5 TAR jet reconstruction algorithm depicted without lepton disentanglement.	78
369	Figure 5.6 Coordinate system used to evaluate the DH candidate mass $\min(m_s)$	82
370	Figure 5.7 Distributions of the DH candidate mass $\min(m_s)$ for MC simulated events produced for the DH signal model over a range of m_s and $m_{Z'}$	85
373	Figure 5.8 Transverse mass distribution for SM background and several signal points,	86
375	Figure 5.9 A graphical representation of the two kinematic categories used in the search based on the characteristics of the final state.	88
377	Figure 5.10N-1 Distributions for the $m^{\text{TAR Jet}}$ and $\Delta R(\text{TAR Jet}, \ell)$ variables used in the merged signal region definition.	96
379	Figure 5.11N-1 Distributions for the $m(W_{\text{Cand}})$ and $\Delta R(W_{\text{Cand}}, \ell)$ variables used in the resolved signal region definition.	97
381	Figure 5.12Comparison of typical event topologies between the signal region and the $W + \text{jets}$ control region.	103
383	Figure 5.13Comparison of N-1 distributions between the merged signal region and the merged $W + \text{jets}$ control region, for different lower bounds on \mathcal{S} in the merged signal region.	104

386	Figure 5.14 Predicted yields of MC simulated events for all signal points in the merged $W + \text{jets}$ control region.	106
387		
388	Figure 5.15 Predicted yields of MC simulated events for all signal points in the resolved $W + \text{jets}$ control region.	106
389		
390	Figure 5.16 Predicted yields of MC simulated events for all signal points in the merged $t\bar{t}$ control region.	108
391		
392	Figure 5.17 Predicted yields of MC simulated events for all signal points in the resolved $t\bar{t}$ control region.	108
393		
394	Figure 5.18 Comparison of the E_T^{miss} trigger efficiency between MC simulated events and ATLAS data.	114
395		
396	Figure 5.19 E_T^{miss} trigger efficiency in the muon channel, with muons treated as invisible in the calculation of E_T^{miss}	116
397		
398	Figure 5.20 Efficiency of the E_T^{miss} OR single muon trigger in the muon channel.	117
399		
400	Figure 6.1 Envelope of shifts in the total predicted yield of Standard Model background processes in the signal regions due to experimental systematics.	122
401		
402	Figure 6.2 Envelope of shifts in the total predicted yield of Dark Higgs signal process at $(m_s, m_{Z'}) = (210, 2100)$ in the signal regions due to experimental systematics.	123
403		
404	Figure 6.3 Shifts in predicted yields in the merged signal regions due to theoretical systematic uncertainties.	134
405		
406		
407	Figure 7.1 Predicted yields of SM background processes and the Dark Higgs signal model at several mass points in the signal regions, binned in $\min(m_s)$	146
408		
409		
410	Figure 7.2 Predicted yield of SM background processes and the Dark Higgs signal model at several mass points in the signal regions, binned in $\min(m_s)$ with optimized bin edges.	147
411		
412		
413	Figure 8.1 Comparison between the predicted yields of Standard Model background processes and observed yields of ATLAS collision data in the control regions, before and after the background-only fit.	155
414		
415		
416		
417	Figure 8.2 Post-fit values and uncertainties of all NPs in the background- only fit.	163
418		

419	Figure 8.3 Correlation matrix for nuisance parameters considered in the background-only fit.	164
420		
421	Figure 8.4 Comparison between predicted yields of Standard Model back-	
422	ground processes and observed yields in the signal regions, before	
423	and after the background-only fit and extrapolation to the SR.	165
424		
425	Figure 8.5 Comparison between predicted yields of SM background pro-	
426	cesses, the Dark Higgs signal process at $(m_s, m_{Z'}) = (160, 1000)$ GeV	
427	and observed yields in the signal regions.	166
428		
429	Figure 8.6 Comparison between predicted yields of SM background pro-	
430	cesses, the Dark Higgs signal process at $(m_s, m_{Z'}) = (210, 2100)$ GeV	
431	and observed yields in the signal regions.	167
432		
433	Figure 8.7 Comparison between predicted yields of SM background pro-	
434	cesses, the Dark Higgs signal process at $(m_s, m_{Z'}) = (310, 2900)$ GeV	
435	and observed yields in the signal regions.	168
436		
437	Figure 8.8 Post-fit value and uncertainty of the signal strength parameter	
438	in the signal+background fit for each m_s and $m_{Z'}$ in the Dark	
439	Higgs signal model considered in the search.	169
440		
441	Figure 8.9 Post-fit values of all nuisance parameters in the signal+background	
442	fit using the Dark Higgs signal model at $(m_s, m_{Z'}) = (210, 2100)$ GeV.	170
443		
444	Figure 8.10 Correlation matrix for nuisance parameters considered in the	
445	signal+background fit at a sample signal point of $(m_s, m_{Z'}) =$	
446	$(210, 2100)$ GeV	171
447		
448	Figure 8.11 Leading 30 pre-and post-fit impacts on the signal strength for	
449	nuisance parameters associated with experimental and theoretical	
450	uncertainties in the signal+background fit at $(m_s, m_{Z'}) = (210,$	
451	$2100)$ GeV.	172
452		
453	Figure 8.12 Expected and observed range of m_s and $m_{Z'}$ in the Dark Higgs	
454	model excluded by this search.	173
455		
456	Figure 8.13 Summary of m_s and $m_{Z'}$ parameters in the DH model excluded	
457	by all searches for the model by ATLAS and CMS.	173
458		
459	Figure A.1 N-1 Distributions for variables used in the merged signal region	
460	definition.	177
461		
462	Figure A.2 N-1 Distributions for variables used in the resolved signal region	
463	definition.	178
464		

453	Figure B.1 Comparison of N-1 distributions for kinematic variables of interest between the signal region and the $W+jets$ control region in the merged category.	180
454		
455		
456	Figure B.3 Comparison of N-1 distributions for kinematic variables of interest between the signal region and the $t\bar{t}$ control region in the merged category.	184
457		
458		
459	Figure B.4 Comparison of N-1 distributions for kinematic variables of interest between the signal region and the $t\bar{t}$ control region in the resolved category.	186
460		
461		
462	Figure D.1 Range of m_s and $m_{Z'}$ in the Dark Higgs signal model excluded by the search for various choices of the signal strength μ	193
463		

464

ACKNOWLEDGEMENTS

465 I would like to thank:

466 **Dr. Robert Kowalewski** for giving me the opportunity to work on the ATLAS
467 experiment and pursue my interests within it, vastly improving the quality of
468 this work through your advice and expertise, and being the most supportive
469 supervisor I could have asked for.

470 **Dr. Randall Sobie** for the interesting and very rewarding work on ATLAS com-
471 puting, and the support throughout.

472 **Dr. Kenji Hamano** for the guidance and advice, which was invaluable to the mono-
473 s analysis.

474 **Dr. Philipp Mogg** for graciously teaming up with me as analysis co-contact, and
475 sharing your impressive expertise and insight throughout.

476 **Dr. Rolf Seuster** for supporting my ATLAS computing work, and the many de-
477 lightful conversations.

478 **Dr. Richard Keeler** for your generosity in offering me an inspiring first exposure to
479 physics research as an undergraduate student at a time when I wasn't really sure
480 I was cut out for the field, and for your continued support and encouragement
481 since then.

482 **My parents and sisters** for encouraging me to carve my own path, and for your
483 unwavering love and support at every twist and turn.

484 **Struan Eamer** for being crazy enough to hike the Juan de Fuca trail in the middle
485 of winter so I could meet you, and for changing my life ever since.

486 **My friends** for reminding me that there's more to life than my computer.

487

DEDICATION

488

To Paul

Chapter 1

Introduction

With the discovery of the Higgs boson in 2012 [6, 7], the final piece of the Standard Model fell into place. Developed during the latter half of the 20th century, the Standard Model describes all known particles and their interactions. The model has demonstrated remarkable predictive power, and can account for nearly all phenomena observed in particle physics detectors to date, with no firmly established deviations from its predictions. Yet, it is known to be incomplete. Among the observed phenomena that it fails to explain [8, 9, 10] are multiple lines of astronomical observation that collectively point to the existence of a new form of non-luminous matter in the universe known as “dark matter” (DM).

While the nature of DM remains a mystery, the observational data strongly suggest that it will take the form of one or more new particles beyond the Standard Model. In addition to its gravitational interactions with massive particles, theoretical considerations give good reason to expect that the new particle(s) could experience additional, albeit weak, couplings to particles of the Standard Model by mechanisms yet to be determined. The study presented in this thesis is part of a multi-pronged international effort to search for evidence of particle DM by means of its non-gravitational interactions in particle physics experiments. In particular, this study probes new ground in the search for DM production in high-energy particle collisions at the Large Hadron Collider (LHC) at CERN.

This chapter introduces the Standard Model, focusing on aspects that are particularly relevant to the presented search, and discusses the astronomical evidence for the existence of DM in the universe. It also introduces the range of experimental strategies that are currently employed to search for evidence of DM in particle physics detectors, and how this search fits into the wider DM search programme at the LHC.

515 1.1 Introduction to the Standard Model

516 The Standard Model (SM) describes all known elementary particles and three of
 517 the four known forces by which they interact with one another - the strong, the
 518 electromagnetic (EM), and weak forces. The theory of general relativity [8], which
 519 describes the gravitational force, has yet to be incorporated into the SM.

520 The known particles, illustrated in Figure 1.1 are divided into two classes known
 521 as “fermions” and “bosons” on the basis of an intrinsic form of angular momentum
 522 known as “spin”. Fermions carry spin $\frac{1}{2}$, and bosons carry integer spin.

523 The specific forces by which particles in the SM interact with one another are
 524 determined by the charge(s) that they carry. Particles that carry electric charge
 525 interact with other particles carrying this charge via the EM force. Similarly, particles
 526 that carry weak and colour charge interact via the weak and strong forces, respectively.

527 Each fermion has a corresponding anti-particle with the same mass, but with
 528 opposite values of the charges it carries - for example, the electron e^- carries negative
 529 electric charge and its antiparticle, the positron e^+ , carries positive electric charge.

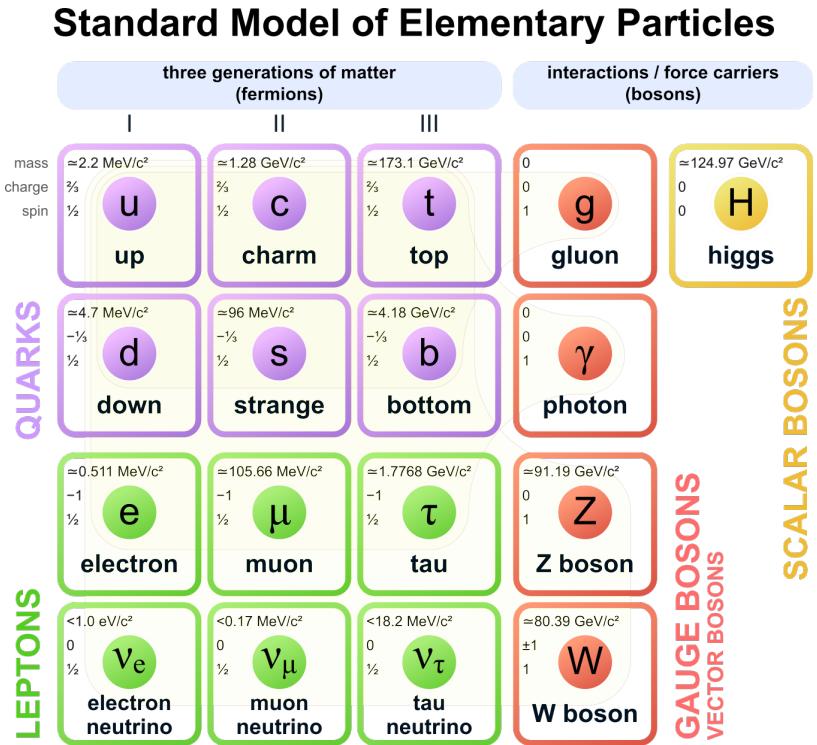


Figure 1.1: Names and fundamental properties of particles in the Standard Model.
 Figure from © [11].

530 **1.1.1 Fermions**

531 Fermions are further sub-divided into leptons and quarks, depending on the charges
 532 they carry, and hence by the forces with which they interact. There are three known
 533 generations of fermions, labelled I, II and III in Figure 1.1, each with significantly
 534 larger mass than the last. Each generation contains a pair of quarks and a pair of
 535 leptons, along with their associated antiparticles. The quark pair consists of one “up-
 536 type” quark with positive electric charge and one “down-type” with negative charge.
 537 The lepton pair consists of one charged lepton and one charge-neutral “neutrino”.

538 Leptons carry electric charge and weak isospin, and as a result interact with one
 539 another and with other particles carrying these charges via the EM and weak forces,
 540 respectively.

541 Like leptons, quarks also carry electric charge and weak isospin, and additionally
 542 carry colour charge. The colour charge allows quarks to interact via the strong force.
 543 As a result, quarks interact by all three forces described by the SM. Unlike charged
 544 leptons, which carry an electric charge of ± 1 , quarks carry fractional electric charge;
 545 up-type (down-type) quarks carry a charge of $+\frac{2}{3}$ ($-\frac{1}{3}$).

546 Due to an effect known as “colour confinement”, quarks cannot exist as stable par-
 547 ticles in isolation, and must instead combine with other quarks to form stable “colour-
 548 neutral” states called “hadrons”. The two major forms of hadrons are “mesons”
 549 formed by a quark-antiquark pair and “baryons” formed by three quarks. Although
 550 the intrinsic strength (i.e. probability) with which particles couple via the strong
 551 interaction is 3 (14) orders of magnitude greater than via the EM (weak) interaction
 552 [12], the range of strong force interactions is limited by colour confinement to the
 553 approximate size of the proton (10^{-15}m). The LHC collides protons with sufficient
 554 energy to probe interactions between their constituent quarks and gluons at length
 555 scales smaller than the size of the proton. Due to the large strong force coupling at
 556 this scale, there is a relatively high probability that the interactions initiated by these
 557 collisions will proceed via the strong force compared with other forces.

558 **1.1.2 Bosons**

559 Bosons in the SM are divided into “gauge bosons” and “scalar bosons”. The gauge
 560 bosons are spin 1 force carriers that mediate interactions between particles. The
 561 photon mediates EM interactions between electrically charged particles. The gluon
 562 mediates the strong interaction between quarks. Unlike charge-neutral photons, the

563 gluon itself carries colour charge, which allows it to self-interact via the strong force.
 564 The weak force is mediated by three particles: the electrically neutral Z boson, and
 565 two W bosons (W^\pm) with opposite electric charges of ± 1 .

566 Scalar bosons are defined as spin 0 particles. There is only one scalar boson in the
 567 SM, namely the Higgs boson (or, simply, the “Higgs”) [13, 14, 15]. Particles in the
 568 SM acquire mass via their interaction with the Higgs field. As such, the Higgs only
 569 interacts with massive SM particles, which includes all particles except the photon
 570 and the gluon. The more massive the particle, the stronger its coupling with the
 571 Higgs. Neutrinos are a possible exception; there is at present no mechanism in the
 572 SM by which neutrinos could interact with the Higgs field, so the origin of their tiny
 573 masses remains an open question.

574 1.1.3 Collision and Decay Processes at Colliders

575 The high-energy counter-rotating proton beams at the LHC are brought into head-on
 576 collisions at four interaction points around the ring, each of which is surrounded by
 577 a detector¹. At each interaction point, constituents of the colliding protons known as
 578 “partons”² can pair annihilate to form observable collision products via one or more
 579 “virtual mediators”³, and the collision products are subsequently measured by the
 580 detector.

581 Each process that describes a physically allowed mechanism by which partons
 582 may annihilate to form observable particles has a certain probability of taking place
 583 relative to other possible annihilation and production processes. The probability
 584 that a given process will take place is quantified by its “cross section” σ . The beam
 585 luminosity \mathcal{L} relates the rate of collisions $\frac{dN}{dt}$ that proceed via a given process to the
 586 cross section of the process:

$$\frac{dN}{dt} = \mathcal{L}\sigma \quad (1.1)$$

587 The luminosity can be integrated over a period of time t_1 to t_2 , such that the total
 588 number of events expected to be produced via a process with cross section σ over the
 589 given period is related to the “integrated luminosity” \mathcal{L}_{int} by:

¹See Chapter 3 for a detailed discussion of the LHC and the detectors that surround the four interaction points.

²See Section 3.1 for an introduction to the parton model.

³See Section 1.1.5 for a discussion of virtual particles.

$$N = \sigma \int_{t_1}^{t_2} \mathcal{L}(t) dt = \sigma \mathcal{L}_{\text{int}} \quad (1.2)$$

590 **1.1.4 Unstable Particles**

591 The lowest-mass “first-generation” quarks and charged leptons located in column I
 592 in Figure 1.1 are, along with neutrinos and the massless photons and gluons, the
 593 only stable particles in the SM. All other particles are unstable, and will decay to
 594 less-massive particles after they are produced.

595 The decay of an unstable particle occurs randomly with respect to the time elapsed
 596 since the particle was produced, with the elapsed time governed by Poisson statistics.
 597 The probability $P(t)$ that an unstable particle will have decayed after a time t has
 598 elapsed in its rest frame is given by the following cumulative exponential distribution:

$$P(t) = 1 - e^{-\frac{t}{\tau}} \quad (1.3)$$

599 where τ is the mean lifetime of the particle.

600 **Unstable Resonance and the Breit-Wigner Formula**

601 Due to their finite lifetime, the Heisenberg uncertainty principle implies that unstable
 602 particles will not be produced with a well-defined mass, but rather with a mass
 603 distribution peaked at a central value m_0 . The probability density function $p(m)$
 604 associated with measuring an unstable particle to have a mass m is given by the
 605 Breit-Wigner formula [16]:

$$p(m) \propto \frac{1}{(m - m_0)^2 + \frac{\Gamma_E^2}{4}} \quad (1.4)$$

606 which exhibits a resonant peak of width Γ_E centred at m_0 . The lifetime of the unstable
 607 particle is related to the width of its Breit-Wigner resonance by $\tau = \frac{\hbar}{\Gamma_E}$.

608 For unstable mediators produced in high-energy colliders which decay to a pair
 609 of detectable particles, the mediator mass can be reconstructed as the combined
 610 invariant mass⁴ of its measured decay products. Neglecting detector resolution effects,
 611 the cross section $\sigma(m)$ associated with producing an unstable particle with a given
 612 reconstructed mass is expected to be proportional to its Breit-Wigner distribution:

⁴The invariant mass m of two particles with four-momenta p_1 and p_2 is in general given by:
 $m^2 = (p_1 + p_2)^2 = (E_1 + E_2)^2 - (\mathbf{p}_1 + \mathbf{p}_2)^2$

613 $\sigma(m) \propto p(m)$. This results in a characteristic resonant peak in the reconstructed
614 invariant mass distribution of the mediator's decay products, which can be used to
615 identify the unstable mediator, and to measure its mass and lifetime.

616 **1.1.5 Feynman Diagrams**

617 The interaction mechanisms by which observable collision products are produced from
618 the annihilation of two partons can be represented by Feynman diagrams, which are
619 described in detail in Chapter 2 of Ref. [12] and summarized here. As an example,
620 the Feynman diagram for the Drell Yan process in which a $q\bar{q}$ pair annihilate to form
621 a lepton pair $\ell\bar{\ell}$ via the exchange of a virtual photon (γ^*)⁵ or Z^* boson mediator is
622 shown in Figure 1.2.

623 The particles involved in the interactions are represented as lines in a Feynman
624 diagram, with different particle types represented by different line styles - fermions
625 are generally represented by solid straight lines, and bosons (with the exception of
626 gluons) are generally represented by wavy lines. Particle interactions are represented
627 by vertices at which lines in the diagram intersect. The annihilation of a $q\bar{q}$ pair to
628 form the virtual γ^*/Z^* mediator is represented in Figure 1.2 by the vertex at which
629 the q and \bar{q} fermion lines meet the γ^*/Z^* boson line, and the subsequent decay of the
630 γ^*/Z^* to $\ell\bar{\ell}$ is represented by the vertex to the right at which the γ^*/Z^* line meets
631 the ℓ and $\bar{\ell}$ lines. Note that time flows horizontally from left to right in Feynman
632 diagrams, so the colliding $q\bar{q}$ pair are shown on the left and the observable decay
633 products $\ell\bar{\ell}$ on the right.

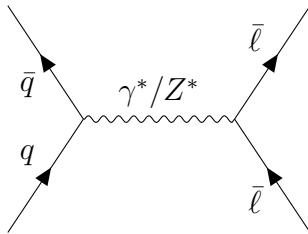


Figure 1.2: Feynman diagram for the Drell Yan process.

5The * indicates that mass of the virtual particle may be off-shell (see discussion of virtual particles below).

634 **Virtual Particles**

635 In general, ingoing and outgoing lines in a Feynman diagram represent real observable
 636 particles, and internal lines represent so-called virtual particles. Virtual particles
 637 are not observable, but are rather a representation of the mechanism involved with
 638 producing the observable final-state particles. Importantly, virtual particles can take
 639 on any mass needed to satisfy energy and momentum conservation at each interaction
 640 vertex.

641 **Matrix Element**

642 The final-state observables (momenta of outgoing particles) associated with a process
 643 of particle production from pp collisions at the LHC, such as the Drell-Yan process
 644 shown in Figure 1.2, depend in general on both the phase space and dynamics of the
 645 process. The phase space represents the full space of available kinematics (masses
 646 and momenta) for the incoming and outgoing particles. The dynamics are encoded
 647 in the matrix element \mathcal{M} , which is calculated on the basis of the internal structure
 648 of the process as represented by the Feynman diagram.

649 **1.1.6 Mathematical Formulation of the Standard Model**

650 The SM is formulated mathematically as a quantum field theory, in which particles of
 651 the SM are represented as excitations of quantum fields. The mathematical formula-
 652 tion of the SM is presented in detail in standard texts [12, 17], and briefly summarized
 653 in this section, with focus placed on aspects that are relevant to later discussions in
 654 this thesis.

655 **Lagrangian Densities**

656 As in classical field theories, the quantum fields of the SM and their interactions are
 657 powerfully described by the formalism of Lagrangian densities, which are functions of
 658 the quantum fields and their derivatives. For example, interactions between photons
 659 and electrically charged fermions are described in quantum electrodynamics (QED)
 660 by the following Lagrangian density term:

$$\mathcal{L}_{\text{QED, interaction}} = -q\psi^\dagger(x)\gamma^0\gamma^\mu\psi(x)A_\mu(x) \quad (1.5)$$

661 where $\psi(x)$, a function of the four spacetime coordinates represented by x , is the
 662 quantum spinor field of the spin- $\frac{1}{2}$ fermions in the SM. $A_\mu(x)$ represents the vector
 663 field of the massless spin-1 photon and γ^μ are the Dirac matrices [12]. The index μ
 664 runs over the four spacetime coordinates. The factor q represents the electric charge
 665 of the fermion involved in the interaction and its value⁶ determines the strength of
 666 the interaction.

667 Symmetries and Group Theory Description

668 Symmetries in the Lagrangian density function associated with each fundamental
 669 interaction are described in the language of group theory by classifying the funda-
 670 mental interactions into gauge groups that describe their symmetries. For example,
 671 QED exhibits a symmetry under local phase transformations, described by the uni-
 672 tary local gauge group $U(1)$. This means that the Lagrangian \mathcal{L}_{QED} is invariant
 673 under the multiplication of the fermion spinor $\psi(x)$ by a unitary function $U = e^{i\theta(x)}$
 674 (unitarity implies that $U^\dagger U = 1$), where $\theta(x)$ can be any function of the spacetime
 675 coordinates x . It is found that this symmetry can be ensured by the inclusion of
 676 the vector field $A_\mu(x)$ in the QED Lagrangian, which is identified with the physical
 677 photon. Because it ensures invariance under the $U(1)$ gauge group, the vector field
 678 $A_\mu(x)$ is referred to as a “gauge field”, and the corresponding boson (the photon) as
 679 a “gauge boson”. The symmetries in the SM are described by the direct product⁷ of
 680 the $U(1) \times SU(2) \times SU(3)$ gauge groups.

681 Quantum Chromodynamics

682 The theory of quantum chromodynamics (QCD), presented in standard texts such
 683 as Ref. [19], describes the strong interactions mediated by gluons between particles
 684 with colour charge (quarks and gluons). Its symmetries are described by the $SU(3)$
 685 gauge group. The quarks are represented by a three-component vector of spinors:
 686 $\psi_c = \{\psi_r, \psi_g, \psi_b\}$, where the subscripts refer to the three colours - red, green and blue
 687 - that constitute the charges of the strong interaction. The QCD Lagrangian (see eg.
 688 Eq. 10.88 in Ref. [12]) is symmetric under a transformation of the vector of quark
 689 spinors ψ_c by a 3×3 $SU(3)$ matrix, which is unitary with determinant 1. The $SU(3)$
 690 symmetry is ensured by the inclusion of an eight-component set \mathbf{A}_μ of vector gauge

⁶The electric charges of all fermions in the SM are as follows: ± 1 for charged leptons, $+\frac{2}{3}$ ($-\frac{1}{3}$) for up (down) type quarks and 0 for neutrinos.

⁷General definitions in group theory can be found, for example, in Section 1.1 of Ref. [18].

691 fields in the QCD Lagrangian. The associated gauge bosons are identified as the eight
 692 physical gluons, each of which possesses a unique superposition of *rgb* colour states
 693 [12].

694 Electroweak Theory and the Higgs Mechanism

695 The mathematical descriptions of the weak and EM forces are unified into a single
 696 “electroweak” theory (for a review, see [20]) whose symmetries are described by the
 697 $SU(2) \times U(1)$ product of gauge groups.

698 Developed in 1954, Yang-Mills theory [21] showed that a set of three massless
 699 vector gauge bosons, referred to as the “isospin triplet” \mathbf{W} are needed to satisfy the
 700 $SU(2)$ symmetry, and a fourth massless vector gauge boson B is needed to satisfy the
 701 $U(1)$ symmetry. Despite satisfying the $SU(2)$ symmetry of the weak interaction, the
 702 weak isospin triplet predicted by Yang-Mills theory falls short of fully describing the
 703 physical W^\pm and Z bosons that mediate the weak interaction, which are known to
 704 be massive [22].

705 The $U(1) \times SU(2) \times SU(3)$ symmetry of the SM Lagrangian does not admit mass
 706 terms of the form $m_X^2 X^\dagger X$, where X is an arbitrary field. Proposed in 1964, the
 707 “Higgs mechanism” [13, 14, 15] provides a means of generating the masses of the
 708 physical W^\pm and Z bosons, as well as all other massive particles (with the exception
 709 of neutrinos), by adding the following term to the SM Lagrangian:

$$\mathcal{L}_{\text{Higgs}} = (D^\mu H)^\dagger (D_\mu H) - V(H) \quad (1.6)$$

710 where

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ h + v \end{pmatrix} \quad (1.7)$$

711 $h(x)$ is interpreted as the scalar field of the physical Higgs boson, and v as the
 712 so-called “vacuum expectation value”. With H in this form, $\mathcal{L}_{\text{Higgs}}$ is described by
 713 the $U(1)$ symmetry group but not the $SU(2)$ group, and is thus said to “break” the
 714 electroweak symmetry $SU(2) \times U(1)$ to the QED gauge symmetry $U(1)$.

715 The covariant derivative $D_\mu H$ in Eq. 1.6 takes the form:

$$D_\mu H = \left(\partial_\mu + i \frac{1}{2} g \sigma_k W_\mu^k + i \frac{1}{2} g' B_\mu \right) H \quad (1.8)$$

716 where σ_k are the Pauli matrices, and g and g' are the coupling constants between the
 717 Higgs field and the \mathbf{W} and \mathbf{B} fields, respectively.

718 The Higgs potential $V(H)$ takes the form:

$$V(H) = -\mu^2 H^\dagger H + \lambda(H^\dagger H)^2 \quad (1.9)$$

719 where the second term describes quartic self-interactions of the Higgs field.

720 The emergence of the massive physical W^\pm , Z bosons and the massless photon
 721 comes from the interactions of the electroweak \mathbf{W} and B fields with the Higgs field to
 722 produce “mass” terms in the Lagrangian of the form $m_X^2 X^\dagger X$. This can be seen by
 723 expanding Equations 1.6 and 1.8, considering only the terms involving the vacuum
 724 expectation value v :

$$\mathcal{L}_{\text{Higgs}} = \frac{v^2}{8} \left[g^2 ((W_\mu^1)^2 + (W_\mu^2)^2) + (gW_\mu^3 - g'B_\mu)^2 \right] + [...] \quad (1.10)$$

725 with the physical vector boson fields and masses defined as:

$$\begin{aligned} W_\mu^\pm &\equiv \frac{1}{2}(W_\mu^1 \mp W_\mu^2) & \text{with mass } m_W &= \frac{gv}{2} \\ Z_\mu &\equiv \frac{1}{\sqrt{g^2 + g'^2}}(gW_\mu^3 - g'B_\mu) & \text{with mass } m_Z &= \frac{v}{2}\sqrt{g^2 + g'^2} \\ A_\mu &\equiv \frac{1}{\sqrt{g^2 + g'^2}}(g'W_\mu^3 + gB_\mu) & \text{with mass } m_A &= 0 \end{aligned} \quad (1.11)$$

726 Inserting the definitions of the physical vector boson fields from Eq. 1.11 back
 727 into Eq. 1.10, it can be readily confirmed that Eq. 1.10 takes the form $\mathcal{L}_{\text{Higgs}} =$
 728 $[m_W^2(W^\pm)_\mu^\dagger(W^\pm)^\mu + m_Z^2 Z_\mu^\dagger Z^\mu + m_A^2 A_\mu^\dagger A^\mu] + [...]$. Masses of fermions are likewise
 729 generated by so-called Yukawa couplings [23] between the fermion and Higgs fields.
 730 The dark Higgs model used to optimize and interpret the DM search presented in
 731 this thesis postulates that DM, as well as any hypothetical new bosons that mediate
 732 its interactions with SM particles, would acquire their masses by means of their
 733 interaction with the dark Higgs field S , as discussed in Chapter 2.

734 1.2 Evidence for Dark Matter from Observational Astronomy

735 Looking beyond the SM, many independent astronomical observations collectively
 736 provide compelling evidence for the presence and abundance of a new form of matter

in the universe that is not directly observable because it neither emits nor absorbs light. Some of the earliest and clearest evidence for this so-called “dark matter” (DM) came in 1978, when Rubin et al. [24] reported systematic anomalies in measured rotation speeds of spiral galaxies. In particular, distributions of the rotation speed as a function of the radial distance from the galactic centre differed in shape from what would be expected on the basis of the distribution of galactic mass measured from the observed luminosity profile.

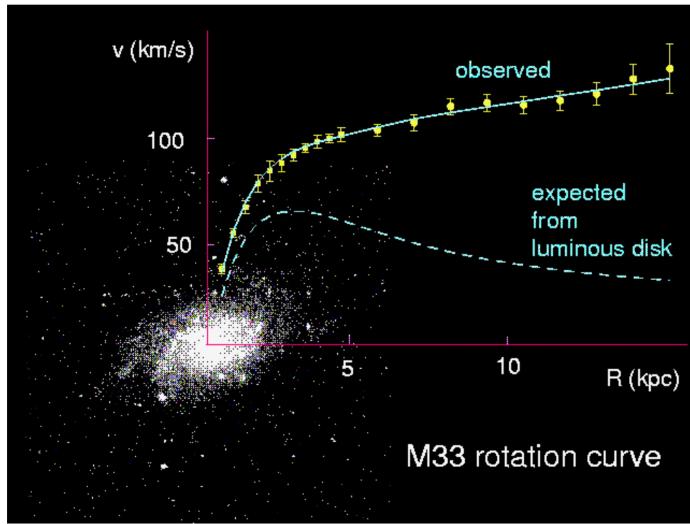


Figure 1.3: Observed rotation speed of the nearby dwarf galaxy M33, overlaid on an optical image of the galaxy. Yellow data points show observed rotational speed of the galaxy as a function of the radial distance from the galactic centre (in kpc). Dashed line shows the expected rotational speed on the basis of the calculated mass of the luminous stellar disk. Figure from © [25].

Spiral galaxies were known at the time to be comprised of a central spheroidal “galaxy bulge” that contains the majority of luminous matter in the galaxy, in addition to a “disk” extending out to larger radii from the galactic centre, within which the density of luminous matter falls off exponentially with radius. If one assumes that the distribution of mass in the spiral galaxy follows its luminosity profile, application of Newtonian gravitational mechanics⁸ would predict the rotation speed to peak near the edge of the central galaxy bulge, as illustrated in the blue dashed line in Figure 1.3 for the dwarf galaxy M33, and fall off at larger radii due to the exponentially

⁸Rotation speeds of spiral galaxies are in general non-relativistic (typically $v_{\text{rot}}/c < 1\%$ [26]). Since Newtonian gravitational mechanics represents an accurate approximation of general relativity in this macroscopic non-relativistic regime, it is generally assumed to provide an appropriate framework for the mathematical description of galactic rotation curves.

752 decaying matter density of the disk. However, the observed galactic rotation speed,
 753 shown as yellow data points in Figure 1.3, is generally observed to continue increasing
 754 well beyond the luminous galactic bulge. These anomalies in galactic rotation curves,
 755 which have since been observed in hundreds of spiral galaxies [26], can be explained
 756 by postulating an additional source of non-luminous matter density in galaxies, DM,
 757 which extends well beyond the luminous bulge and provides the necessary gravita-
 758 tional potential to prevent the rotation speeds from falling off beyond the bulge.

759 In the years following the early reports by Rubin et al., modifications to the laws of
 760 Newtonian gravity at galactic scales [27] were considered as an alternative to explain
 761 the anomalies without invoking the need for DM. However, while the proposed modi-
 762 fications to gravity were successful in describing the observed galactic rotation curves,
 763 numerous astronomical observations in other contexts have independently turned up
 764 results that indicate a need for DM in the universe, many of which cannot be easily
 765 explained by modifying Newtonian gravity. Additional evidence at galactic scales
 766 comes from significant differences between the spatial distributions of matter density,
 767 measured using gravitational lensing, and of luminous matter following collisions of
 768 galaxy clusters, such as the Bullet cluster identified in 1995 [28], which indicate that
 769 the majority of the matter density in the colliding galaxies is non-luminous. Studies
 770 of the relative contribution to the masses of galaxy clusters from luminous matter
 771 using data from the Chandra X-ray observatory [29] suggest that only 15-20% of the
 772 mass composition of the galaxies studied is comprised of luminous matter, with DM
 773 comprising the remaining 80-85%.

774 Because of their stability and EM interactions, protons and neutrons comprised
 775 of bound quarks, as well as their bound electrons - collectively known as “baryonic
 776 matter” - comprise by mass the overwhelming majority of known luminous matter
 777 in the universe. The theory of Big Bang nucleosynthesis (BBN) (for a review, see
 778 Section 24 of Ref. [22]) predicts that the production of light nuclei - D, ^2He , ^4He ,
 779 and ^7Li - took place in the early universe following the Big Bang (see Ref. [30] for a
 780 review of the Big Bang theory), as the universe expanded and cooled sufficiently to
 781 allow their formation by means of nuclear fusion reactions. BBN theory predicts that
 782 the abundances of these light nuclei in the universe were fixed during BBN, following
 783 which the rate of the nuclear fusion reactions became negligibly small due to the con-
 784 tinued expansion and cooling of the universe. Importantly, the theory also predicts
 785 that the relative abundances of the light nuclei are highly sensitive to the density
 786 of baryonic matter in the universe, which was fixed prior to their formation. As a

result, their relative abundance can be used to infer the density of baryonic matter in the universe in the context of BBN theory. Using this approach, precision measurements of the abundances of light nuclei inferred from observational data indicate that baryonic matter constitutes approximately 5% [22] of the energy density of the universe. Current measurements of anisotropies in the cosmic microwave background (CMB) [31] measured by the Planck collaboration [32], interpreted in the context of the Λ CDM model of cosmology (for a review, see Section 25 of [22]), indicate that approximately 30% of the energy density of the universe is comprised of matter, with the missing 25% identified as non-baryonic DM. This result implies that 85% of all matter in the universe is comprised of non-luminous DM, consistent with the findings discussed above from measurements of galaxy clusters.

1.3 Dark Matter Composition Hypotheses

The previous section presented a diverse range of astronomical observations that collectively point to the need for DM in the universe. While active research continues within the theoretical community (see, for example, Refs. [33, 34]) into the possibility of modifying the laws of gravitation at astronomical scales to explain these observations without the need for DM, there are significant theoretical challenges involved with designing modifications that can consistently explain the range of observational anomalies at scales ranging from individual galaxies to galaxy clusters, while simultaneously addressing the apparent need for DM at cosmological scales from the discrepancy between measurements of the baryonic mass density from BBN and the much larger total mass density inferred from anisotropies in the CMB. As a result, DM is widely considered the leading hypothesis to explain the full range of observational data.

While the astronomical observations provide a wealth of information regarding the composition of DM in the universe by means of its gravitational effects on visible matter, they provide relatively few clues as to what actually comprises the DM. Its abundance in the present day universe indicates that it must be stable on cosmological timescales (i.e. billions of years). The evidence from BBN and CMB anisotropies indicates that the DM must be non-baryonic. Its non-luminous nature further implies that it neither emits nor absorbs photons, and therefore has negligible or no charge under the EM force. Besides baryons, neutrinos - with their tiny but nonzero

819 masses⁹ - represent the only other massive stable particles currently known to the SM,
 820 and satisfy the requirement of being electrically neutral. However, the possibility of
 821 neutrinos constituting any appreciable fraction of the DM was ruled out by studies
 822 published in the 1980's [35], which demonstrated that the large scale structure of the
 823 universe would differ significantly from what is observed today if the mass density
 824 of the universe were dominated by neutrinos due to their ultra-relativistic velocity.
 825 More generally, analysis of the measured anisotropies in the CMB measured by the
 826 Planck collaboration [32] is found to strongly favour the standard Λ CDM model in
 827 which the DM is predominantly comprised of "cold" particles, so called because they
 828 travel at non-relativistic velocities.

829 With the stable particles of the SM ruled out, the current most widely accepted
 830 hypothesis is that DM is comprised of a new form of cold non-baryonic matter that
 831 is not currently described by the SM.

832 1.3.1 Origin and Interactions of Particle Dark Matter

833 Despite the observable effects of its gravitational interactions at astronomical scales,
 834 the strength of gravitational couplings between massive particles is ~ 30 orders of
 835 magnitude weaker than any of the other three known forces [12]. As a result, gravi-
 836 tational interactions between DM and SM particles are far too weak to be observable
 837 in particle detectors. Given that there have not yet been any conclusive indications
 838 of DM in particle detectors, it can be further deduced that any non-gravitational
 839 interactions between DM and particles of the SM are relatively weak compared with
 840 the strong, weak and EM couplings between SM particles. However, most theories
 841 that aim to describe the origin of the observed abundance of DM in the present day
 842 universe imply the existence of non-gravitational couplings between DM and SM par-
 843 ticles, and in many cases predict that the couplings could be strong enough to be
 844 probed by modern particle detection methods. This produces a generic class of DM
 845 candidates known as weakly-interacting¹⁰ massive particles (WIMPs).

846 A positive detection of WIMPs would not only confirm the hypothesis of particle
 847 DM, but would also allow physicists to begin to study its properties as a particle, and
 848 test theoretical extensions of the SM that incorporate particle DM.

⁹Current constraints from cosmology place an upper limit on the sum of neutrino masses from all generations of 0.17 eV, 3×10^6 times smaller than the electron mass.

¹⁰The "weak" interactions of the WIMP DM candidates are in general not necessarily associated with the weak force, but are simply too weak to have produced a measurable signature in particle detectors to date

849 **Dark Matter Origin from Thermal Freeze-out**

850 A review of the existing hypotheses for the origin of DM can be found in Section 27.3
 851 of Ref. [22]. Of these, the so-called “thermal freeze-out” scenario is a popular candi-
 852 date, because it postulates that the observed DM density in the present day universe
 853 originated from the same process of thermal decoupling that produced the primordial
 854 abundances of light nuclei in the well-tested BBN scenario discussed earlier. The
 855 hypothesis postulates that in the very early universe, matter was sufficiently dense
 856 and energetic to establish thermal equilibrium between DM and SM particles due to
 857 interactions between DM and SM particles (so-called “DM-SM interactions”). As the
 858 universe expanded and cooled, eventually the rate of DM-SM interactions became too
 859 low to maintain thermal equilibrium between the two species. At this point, known
 860 as “thermal freeze-out”, DM became decoupled from SM particles, thus fixing the
 861 relic abundance of DM observed in the present-day universe.

862 For cold DM relics ($v/c \lesssim 0.1$ at the time of freeze-out), and assuming that the
 863 relic abundance is predominantly set by direct DM-SM interactions, analysis of the
 864 observed relic abundance of DM in the context of the thermal freeze-out hypothesis
 865 (see, for example, Section B of Ref. [36]) implies that the cross section for SM-
 866 DM interactions should be $\sigma_{\text{SM-DM}} \gtrsim 1 \text{ pb}$, comparable to typical cross sections for
 867 interactions mediated by the weak force. Searches for DM in particle detectors (for
 868 a review, see [37]) have yet to turn up any hints of a DM candidate with interaction
 869 cross sections with the SM near the weak scale. However, the cross section constraint
 870 can be significantly relaxed by considering a scenario in which the relic abundance
 871 of DM is set not by direct interactions between the DM and the SM, but rather by
 872 interactions between DM and an unstable mediator, which subsequently decays to
 873 SM particles (see, for example, Ref. [38]). The DM search presented in this thesis
 874 is interpreted in the context of such a scenario, wherein the unstable mediator is the
 875 Dark Higgs boson [39, 40].

876 **1.4 Dark Matter Search Strategies**

877 There are three complementary approaches used to search for particle DM by means
 878 of its non-gravitational interactions. Direct detection searches (for a review, see [41])
 879 aim to detect evidence of a recoil induced by elastic scattering between a DM particle
 880 in the galactic halo and a target particle in the detector. Indirect searches (for a
 881 review, see [42]) use observational data to search for evidence of particles produced

882 by DM annihilation or decay in particular regions of the observable universe that
 883 are expected to have a high DM density. Searches for DM at colliders (for a review,
 884 see [43]), of which the work in this thesis is an example, study the decay products
 885 from high-energy collisions of subatomic particles to search for an above-background
 886 excess of events that could be consistent with DM having been produced in some of
 887 the collisions.

888 1.4.1 Direct Detection

889 Direct detection searches operate in very low-background environments, typically in
 890 underground facilities such as SNOLAB (for a review, see Ref. [44]), in order to
 891 minimize scattering events in the detectors from non-DM sources such as cosmic rays
 892 and radioactivity, and detailed studies are performed to determine the expected rate of
 893 events from all possible background sources. As a result, a significant excess of elastic
 894 scattering events, particularly if observed in multiple direct detection experiments,
 895 would offer a clear signature of interactions with DM in the galactic halo.

896 If no evidence of excess scattering events is found, experiments place upper bounds
 897 on DM-nucleon interaction cross section with a largely standard set of methods and
 898 assumptions (most notably the local DM density and the relative speed with which the
 899 DM passes through Earth) - see, for example, Ref. [45] - which facilitates comparison
 900 between different experiments. Figure 1.4 summarizes the current upper bounds on
 901 the spin-independent¹¹ DM-nucleon interaction cross section from all direct detection
 902 searches. The searches probe down to many orders of magnitude below the weak scale
 903 ($\sigma \sim 10^{-36} \text{cm}^{-2}$) over ~ 4 orders of magnitude of candidate DM masses. However,
 904 current direct detection strategies generally suffer practical limitations to the ranges
 905 of DM masses and interaction cross sections that can be probed. The lower bound
 906 on accessible DM masses is in general dictated by the signal to noise ratio of the
 907 detector, referred to as the “noise wall”, which is quite difficult to overcome. The
 908 range of accessible cross sections is also bounded from below for most direct detection
 909 experiments by the so-called “solar neutrino floor”, below which the measured event
 910 rate becomes dominated by the irreducible flux of solar neutrinos passing through the
 911 Earth.

¹¹Spin-dependent vs. spin-independent DM-nucleon interactions differ according to whether the coupling is sensitive to the spin state of the target nucleon [41].

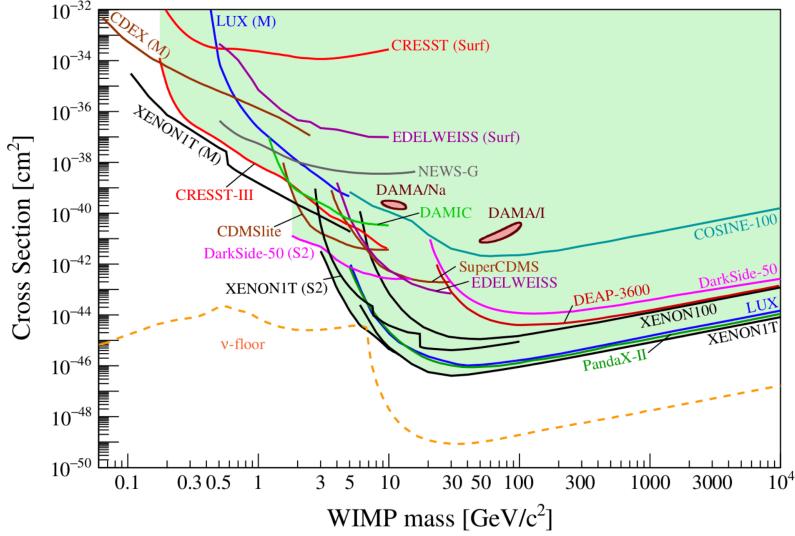


Figure 1.4: Summary of upper bounds from direct detection searches on the interaction cross section for spin-independent WIMP-nucleon scattering, over a range of hypothetical WIMP masses. Upper bounds from individual searches are shown as solid lines. Results labelled “M” were obtained assuming the Migdal effect [46]. Shaded green region shows combined exclusion from all searches, excluding results obtained assuming the Migdal effect. Yellow dashed line shows the solar neutrino floor for a Ge target, computed using the assumptions and methodology presented in Refs. [47, 48]. Figure from © [41].

912 1.4.2 Indirect Detection

913 By searching for excesses of several potential DM annihilation products in observational
 914 data - gamma rays, charged leptons and antimatter - in addition to neutrinos,
 915 indirect detection searches (reviewed in Refs. [42, 49, 22]) can avoid the limitation of
 916 the solar neutrino floor. Depending on the target species, these searches can also tar-
 917 get a wider range of candidate DM masses compared with direct detection searches.
 918 Due to the many potential processes that could produce the target particles in obser-
 919 vational data - both within and beyond the SM - indirect searches generally contend
 920 with relatively large uncertainties associated with modelling the expected flux from
 921 these background sources.

922 1.4.3 Searches for DM at Colliders

923 Rather than searching for non-gravitational interactions of relic DM on Earth or in the
 924 observable universe, searches for DM at colliders (for a review, see [43]) instead look
 925 for evidence of DM production from high-energy particle collisions. Like neutrinos,

926 DM would be expected to pass invisibly through any detector surrounding the collision
 927 point due to its very low interaction cross section, producing a momentum imbalance
 928 transverse to the beam line referred to as E_T^{miss} ¹². An excess of collision events with
 929 large final-state E_T^{miss} above the rate expected from SM processes with final-state
 930 neutrinos would be consistent with the production of DM in the high-energy collisions.
 931 Given that other hypothetical new physics processes (see, for example, Refs. [50, 51])
 932 could also produce an excess of high- E_T^{miss} events from non-DM sources, any such
 933 findings would benefit from corroborating DM detections in direct and/or indirect
 934 detection experiments.

935 Despite operating in a very high background environment, colliders offer numerous
 936 advantages that allow DM searches performed using collider data to complement and
 937 potentially extend the reach of direct and indirect searches. First, the detectors used
 938 by particle colliders are often designed to measure all final-state particles produced
 939 by the collisions and their kinematic information with high precision. This detailed
 940 final-state information allows DM searches to target specific final-state topologies,
 941 which can lead to substantial reductions in SM background processes and considerably
 942 enhance the sensitivity to hypothetical DM production processes that predict events
 943 with the targeted topology. Second, by targeting DM produced in the collisions
 944 and adopting a search strategy that does not require the DM to interact with the
 945 detector, searches at colliders are insensitive to the neutrino floor that will challenge
 946 the sensitivity of next-generation direct detection searches. Third, while the range of
 947 DM masses is bound from above by the centre of maximum centre-of-mass energy of
 948 the particle collisions (\sim TeV for proton-proton collisions at the LHC), DM searches
 949 at colliders do not suffer the noise wall that limits the sensitivity of direct detection
 950 searches below ~ 1 GeV (see Figure 1.4).

951 Even if particle DM is first discovered at a non-collider experiment, the detailed
 952 final-state information available in particle collision data will enable detailed mea-
 953 surements of its properties and interactions, provided that it can be produced at
 954 colliders.

¹²See Section 3.4.4 for a detailed introduction to missing transverse momentum in the ATLAS detector.

955 **1.4.4 Searching for Dark Matter at Particle Accelerators**

956 **Approaches used to Search for DM at Colliders**

957 The concept of searching for evidence of DM production in high-energy particle col-
 958 lisions is currently being pursued by numerous collaborations. The particular energy
 959 scales and detector technologies available to each experiment can be exploited to tar-
 960 get specific mass ranges and possible DM production mechanisms, thus allowing for
 961 a rich programme of complementary searches.

962 The proton-proton collision experiments at the Large Hadron Collider (LHC)¹³
 963 [52] - ATLAS [53], CMS [54] and LHCb [55] - benefit from the world-leading 13 TeV
 964 centre-of-mass energy of the pp collisions to probe models with massive mediators
 965 (m_{med} up to a few TeV) of the DM-SM interactions that could be produced in the
 966 collisions. The hermetic¹⁴ coverage and precise event reconstruction available with the
 967 ATLAS and CMS general-purpose detectors make it possible to probe a wide range
 968 of hypothetical DM production models and final-state signatures, typically targeting
 969 DM candidates with masses in the \sim GeV-TeV range (see Ref. [56] for a review
 970 of DM searches performed with the ATLAS and CMS detectors). Meanwhile, DM
 971 searches at LHCb (for a review, see [57]) take advantage of the detector's excellent
 972 forward-angle coverage and vertex resolution to target signatures with lower-mass
 973 DM (\sim MeV-GeV) and displaced vertices.

974 In addition to proton-proton collisions at the LHC, production of DM in electron-
 975 positron e^+e^- collisions has been probed with the BABAR experiment [58] at the
 976 Stanford Linear Accelerator Centre (SLAC), as well as the Belle experiment [59] and
 977 its recent Belle II upgrade [60] at the SuperKEKB collider [61]. With a 10.6 GeV
 978 centre-of-mass collision energy, the Belle II experiment is particularly well suited
 979 to study DM with masses in the range of a few MeV to \sim 10 GeV. Searches at
 980 e^+e^- colliders benefit in precision both from the well-defined initial state afforded
 981 by colliding fundamental particles, and from the vastly reduced background of QCD
 982 activity¹⁵ in the final state compared with pp collision events. Searches performed with
 983 early Belle II data, reviewed in Ref. [62], are already showing promising sensitivity
 984 to a number of low-mass DM candidates, with significant sensitivity improvements

¹³See Section 3 for a general introduction to the LHC and its major detectors.

¹⁴Hermetic detectors, of which ATLAS and CMS are examples, are designed to detect all SM decay products from a collision with the exception of neutrinos.

¹⁵See Section 3.4.5 for a more detailed discussion of the QCD background in the context of the ATLAS triggering system.

985 expected as more data is collected in the coming years.

986 The most direct way to search for DM at pp and e^+e^- colliders is to look for
 987 evidence of the so-called “ $E_T^{\text{miss}} + X$ ” events introduced above, in which the DM
 988 is produced along with detectable SM particles, thus producing a signature of SM
 989 particles recoiling against E_T^{miss} in the final state. An alternative and less direct
 990 approach, known as a “resonance search”, is to look for evidence of the production
 991 of new massive mediators which could potentially mediate DM-SM interactions by
 992 looking for resonant peaks in the invariant mass distribution of one or more pairs of
 993 final-state particles. Such a peak, if not associated with any known SM mediators,
 994 would be indicative of the production and subsequent decay of a new massive mediator
 995 to a pair of SM particles.

996 DM in the \sim MeV-GeV mass range can also be probed with competitive sensitivity
 997 at fixed-target experiments in which a beam of energetic electrons or protons is di-
 998 rected at a fixed target, and downstream detectors search for evidence of DM produced
 999 from the electron-nucleon or proton-nucleon collisions. A variety of approaches are
 1000 employed by different experiments to search for signatures of DM production. Some
 1001 searches re-purpose neutrino detectors, such as MiniBooNE [63] and NOvA [64] at
 1002 Fermilab, to directly detect any DM that may be produced in the collisions by means
 1003 of DM-nucleon or DM-electron collisions with the detector material, placing addi-
 1004 tional shielding between the fixed target and the detector to reduce the background
 1005 flux of neutrinos (see, for example, Refs. [65, 66]). Others such as NA64 [67] at
 1006 CERN employ a fully hermetic detector to search for a $E_T^{\text{miss}} + X$ signature of DM in
 1007 the downstream collision products.

1008 Models of DM Production

1009 Models of DM production in colliders can range in complexity from an effective field
 1010 theory (EFT), where the DM production mechanism is completely unspecified, to
 1011 a complete model such as supersymmetry, which predicts viable DM candidates as
 1012 part of a hypothesized extension to the SM designed to address a range of phenomena
 1013 unexplained by the SM (see Ref. [68] for a review of supersymmetric DM candidates).

1014 In principle, complete theories of physics beyond the SM, such as the minimal
 1015 supersymmetric SM (MSSM) (for a review, see [71]) can offer theoretically motivated
 1016 and experimentally accessible models that specify the details of candidate processes
 1017 by which the colliding partons may annihilate to produce DM. However, these theories

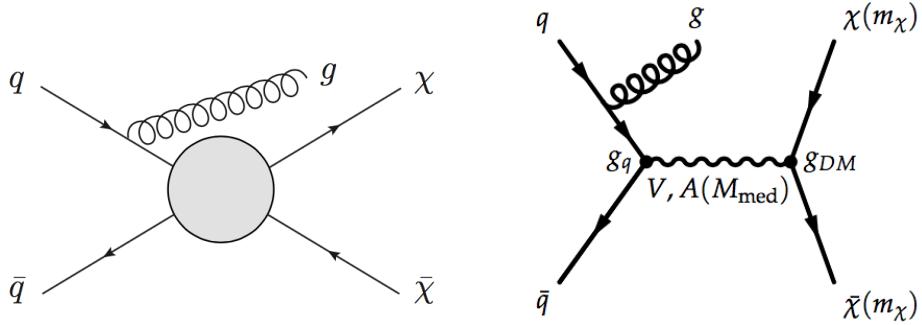


Figure 1.5: Left: $E_T^{\text{miss}} + \text{jets}$ process in the EFT framework (figure from © [69]). Right: $E_T^{\text{miss}} + \text{jets}$ process in a simplified model framework, where the pair production of DM occurs via a new vector or axial-vector (V, A) mediator of mass M_{med} , which couples to quarks and DM with coupling constants g_q and g_{DM} , respectively (figure from © [70])

1018 tend to be quite complex, with many free parameters - over 100 in the case of MSSM
 1019 - most of which need to be fixed to generate a reasonably testable model. Relying on
 1020 complete theories alone to guide experimental signatures may run the risk of missing
 1021 important parameter space of new physics for which a complete theory has not yet
 1022 been developed.

1023 Simplified models, widely used in recent and ongoing DM searches at the LHC
 1024 (see, for example, the 2015 report of the ATLAS/CMS Dark Matter Forum [70]),
 1025 are designed to bridge the gap between EFT and complete theories. They provide a
 1026 “first-order” description of theoretically motivated new physics scenarios that could
 1027 be accessible at collider energies. They provide guidance for experimental searches
 1028 without fully specifying the details of any additional new physics at energies above
 1029 the collider scale that would be needed for a complete theory. In terms of DM
 1030 production at the LHC, one or more new “portal mediators” associated with new
 1031 physics scenarios may be considered, which allow for mixing between SM particles
 1032 and DM. The process by which the mixing occurs is represented with a tree-level
 1033 diagram whose experimental signature would be accessible at LHC energies, such as
 1034 the diagram shown in Figure 1.5, which represents a DM benchmark model featured
 1035 in Ref. [70].

1036 Many simplified models predict the so-called $E_T^{\text{miss}} + X$ final-state signature dis-
 1037 cussed above, in which the DM is produced in association with detectable SM particles
 1038 (X). Depending on the details of the hypothesized DM production mechanism and
 1039 the parameter ranges considered, different models can vary widely in terms of the

1040 identity and topology of the detectable final-state particles (X) predicted in $E_T^{\text{miss}} + X$
 1041 final states. Therefore, a broad-based program has been undertaken at the LHC to
 1042 search for DM production in a variety of $E_T^{\text{miss}} + X$ final states to ensure maximal
 1043 coverage of potential DM production scenarios. Results from a selection of recent
 1044 DM searches in $E_T^{\text{miss}} + X$ final states can be found in Refs. [72, 5, 73, 74, 75, 76, 77,
 1045 78, 79].

1046 The search presented in this thesis, which targets a final state of DM produced
 1047 in association with a pair of W bosons ($E_T^{\text{miss}} + \text{WW}$), is interpreted with the “Dark
 1048 Higgs” simplified model [40] discussed in Chapter 2. Searches for DM at the LHC,
 1049 interpreted in the context of this model, are sensitive to DM with mass in the range
 1050 of ~ 100 GeV.

1051 1.5 Summary of the Thesis

1052 Following a brief introduction to the Standard Model (SM) of particle physics, this
 1053 chapter presented multiple lines of evidence from observational astronomy for the
 1054 abundance of DM in the universe, and for its hypothesized composition as one or
 1055 more new particles beyond the SM. This was followed by a discussion of the ongoing
 1056 worldwide effort to search for evidence of particle DM using particle detectors, and
 1057 how the search presented in this thesis fits into the wider effort.

1058 The following chapter discusses the “Dark Higgs” model that is used to interpret
 1059 the search. Chapter 3 introduces the LHC machine and the ATLAS detector used to
 1060 collect the particle collision data. Chapter 4 introduces the Monte Carlo method and
 1061 its application to modelling the expected yields of events in the ATLAS detector, both
 1062 from the Dark Higgs signal process and from known Standard Model processes that
 1063 constitute a background in the search. The reconstruction and analysis of the ATLAS
 1064 collision data is discussed in Chapter 5, and Chapter 6 presents the methods used
 1065 to quantify the impacts of uncertainties from theoretical and experimental sources.
 1066 Chapter 7 discusses the statistical framework used to interpret the results of the
 1067 search. Chapter 8 presents the range of Dark Higgs model parameters excluded by
 1068 the search. Chapter 9 concludes with a discussion of the experimental strategy and
 1069 results.

1070

Chapter 2

1071

The Dark Higgs Model

1072 The dark matter (DM) search presented in this thesis is motivated by and interpreted
 1073 with the “Dark Higgs” (DH) model [40]. The DH model predicts a mechanism for
 1074 DM production from proton-proton collisions at the LHC by means of portal interac-
 1075 tions with the “dark sector”. The dark sector, which is predicted as part of various
 1076 BSM models, represents a collection of quantum fields and associated particles that
 1077 are assumed to interact gravitationally, but which do not couple via any of the other
 1078 known forces - electromagnetic, strong and weak - of the SM. Non-gravitational cou-
 1079 plings between the dark sector and the SM proceed instead via one or more so-called
 1080 “portal mediators”.

1081 In the DH model, the DM is a particle in the dark sector, and is produced from
 1082 high-energy $q\bar{q}$ collisions at the LHC via a hypothetical spin 1 vector boson portal
 1083 mediator referred to as the Z' . The model introduces an additional Higgs boson in
 1084 the dark sector called the “Dark Higgs” (DH), which acts as a portal mediator by
 1085 decaying to SM particles via a small mixing with the SM Higgs boson.

1086 Figure 2.1 shows three Feynman diagrams, which illustrate some of the dominant
 1087 modes by which the DH model could produce a measurable signature of DM produc-
 1088 tion at the LHC. In all cases, the DM pair is produced via the Z' mediator, along
 1089 with the emission of a DH boson s , which decays to a pair of SM particles.

1090 2.1 Theoretical Motivation for the Dark Higgs Model

1091 Given that the particles of the SM acquire mass via their interaction with the Higgs
 1092 field [13, 14, 15], the existence of a hypothetical “Dark Higgs” (DH) field - and its
 1093 associated particle the DH boson - is motivated by the need to likewise generate the
 1094 masses of particles in the dark sector. More generally, the existence of so-called portal

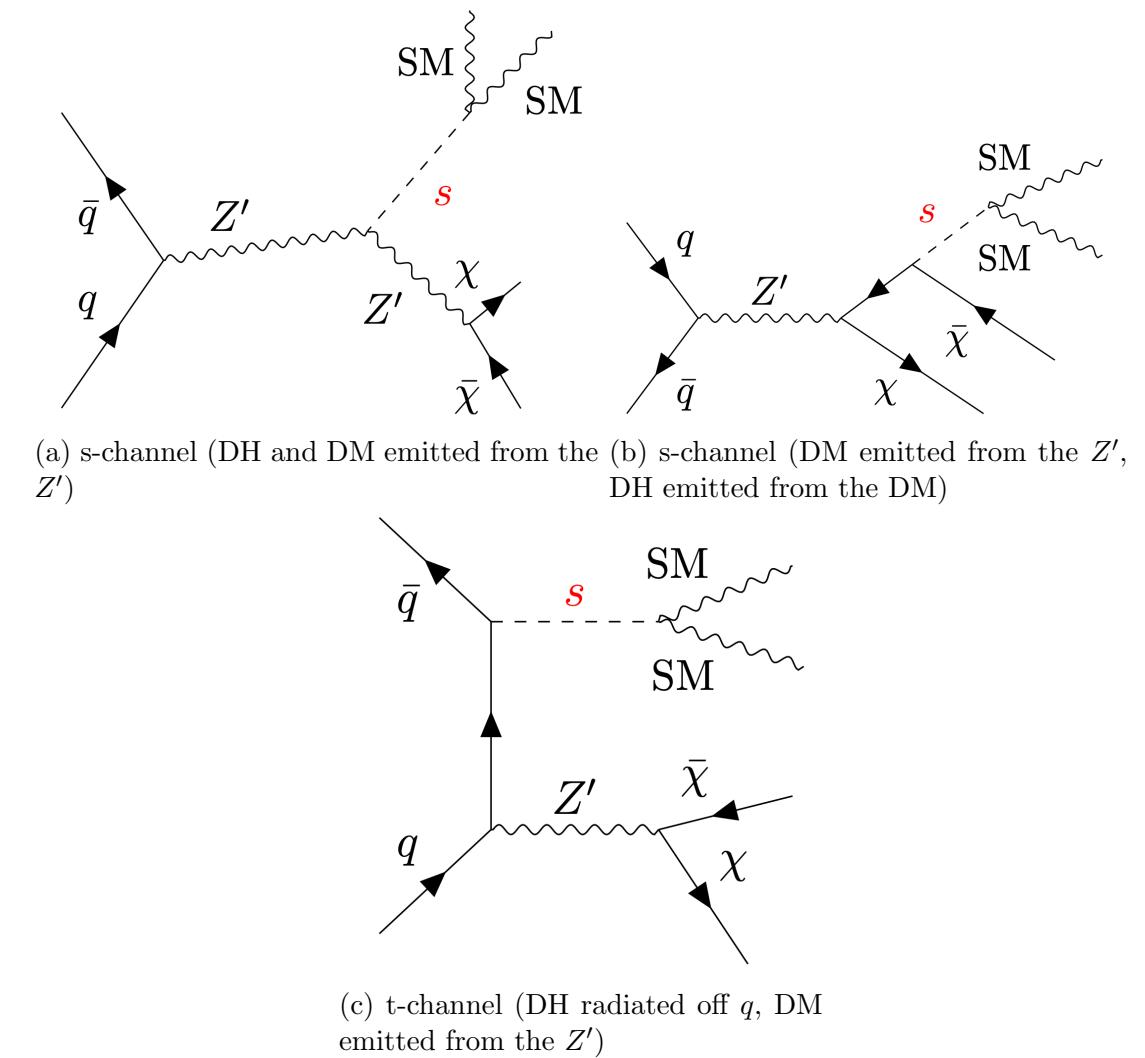


Figure 2.1: Feynman diagrams with leading contributions to the cross section of hypothetical DM production at the LHC by means of the DH model.

mediators that enable interactions between dark sector and SM particles is motivated by theoretical arguments, discussed in Section 1.3.1, for the presence of thermal equilibrium between DM and SM particles in the early Universe, which would be established by creation and annihilation processes between particles in the two sectors. The present-day relic abundance of DM, set at the time of thermal freeze-out, places constraints on the details of these creation and annihilation processes. The hypothesized DH boson would open up a new mechanism for portal interactions between DM and SM particles. As discussed in the following section, this new mechanism allows for a relaxation of constraints from the DM relic abundance compared with simpler models in which portal interactions are limited to those mediated by a vector boson mediator (the so-called Z').

2.1.1 Constraints on Generic Z' Mediator Models

In addition to providing a mechanism by which particles acquire mass in the dark sector, the introduction of a new Higgs boson in the dark sector is motivated by strong theoretical and experimental constraints on the more generic simplified model in which portal interactions between the dark sector and the DM are mediated exclusively by the Z' vector boson mediator (see Ref. [80] for a review of the Z' portal mediator model, and Refs. [81, 82] for reviews of experimental constraints on the model). Removing the emission of the DH s from the contributing Feynman diagrams of the DH model in Figure 2.1 reduces all three to the generic “s-channel” mechanism by which SM particles would pair-annihilate to form DM via the Z' mediator, shown in Figure 2.2a.

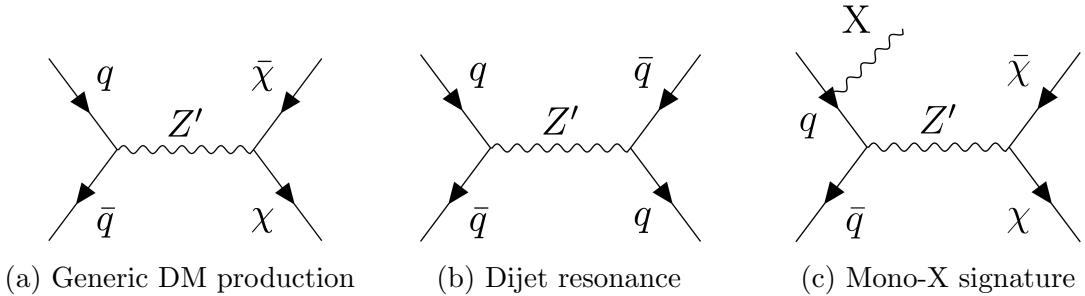


Figure 2.2: Signatures for DM production and detection via a Z' vector boson mediator at the LHC.

The Z' mediator model is probed at the LHC using either dijet resonance searches, which search for a signature of a Z' being created and subsequently decaying back

1119 into a pair of quarks as shown in Figure 2.2b, or with so-called “ $E_T^{\text{miss}} + X$ ” searches in
 1120 which a SM particle “X” is emitted as initial state radiation from one of the colliding
 1121 quarks, as shown in Figure 2.2c, to produce a signature of SM particles recoiling
 1122 against missing transverse momentum due to the undetected DM pair.

1123 Dijet resonance searches probe the model by searching for the presence of a res-
 1124 onant peak in the dijet invariant mass spectrum over the SM background of QCD-
 1125 induced dijet events, where this above-background peak would be induced by the
 1126 process in Figure 2.2b. Ref. [81] presents a statistical combination of several dijet
 1127 searches that were performed with the ATLAS and CMS detectors, as well as an in-
 1128 terpretations of the observed absence of any such above-background resonance peaks
 1129 in the context of the Z' mediator model. It is found that for typical choices of the
 1130 coupling strengths g_q (g_χ) between the Z' and quarks (DM), the model is excluded
 1131 over a wide range of Z' masses ($500 \text{ GeV} < m_{Z'} < 3 \text{ TeV}$) for nearly all DM masses
 1132 up to 2 TeV, as shown in Figure 2.3a. A statistical combination of monojet searches
 1133 performed by ATLAS and CMS, in which the radiated particle X in Figure 2.2c is
 1134 a quark or gluon, is presented in Ref. [82]. As shown in Figure 2.3b, the monojet
 1135 searches likewise exclude a large region of DM and vector boson mediator masses for
 1136 a range of choices for the coupling constants g_q and g_χ .

1137 2.1.2 Implications of a Dark Higgs Portal

1138 The implications of introducing a new portal interaction mediated by a spin 0 boson
 1139 (the DH) - which couples to the SM via a mixing with the SM Higgs boson - to the
 1140 generic Z' mediator portal model are studied in detail by Duerr et al. in Ref. [39].
 1141 In this study, it is found that within various regimes of the coupling strengths and
 1142 masses of the hypothetical particles - the Z' , DH and DM - in this two-mediator model,
 1143 referred to as the DH model, it is possible to relax or evade some of the constraints
 1144 described above that are placed on the generic Z' mediator model by a combination
 1145 of experimental results and the observed relic DM density in the Universe. As a
 1146 result, in addition to providing a mechanism by which particles acquire mass in the
 1147 dark sector, the DH also introduces new parameter space to the model that is not yet
 1148 excluded by existing constraints. In particular, for $m_s < 2m_\chi$ and provided there are
 1149 no other lighter particles in the dark sector, the only available decay route for the s
 1150 is to SM particles via mixing with the SM Higgs, regardless of the mixing strength.
 1151 In this case, the DM relic abundance is predominately set by the process $\chi\chi \rightarrow ss$

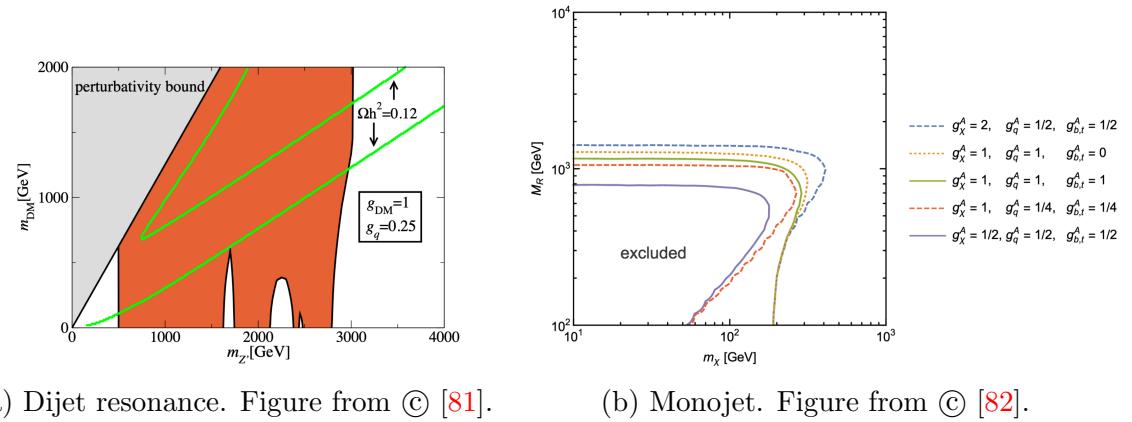


Figure 2.3: Constraints on the masses of the Z' vector boson mediator (labelled $m_{Z'}$ in 2.3a and M_R in 2.3b) and DM (labelled m_{DM} in 2.3a and m_χ in 2.3b) from combined ATLAS and CMS dijet searches [82] (left) and monojet searches [82] (right). The results are shown for typical coupling choices of the Z' to DM (g_{DM} in 2.3a and g_χ^A in 2.3b), and of the Z' to quarks (g_q in 2.3a and g_q^A or $g_{b,t}^A$ in 2.3b, where $g_{b,t}^A$ specifies the coupling to heavy quarks, which may be 0 for models that prohibit heavy quark couplings). Green lines in 2.3a contain the region of $m_{Z'}$, m_χ that reproduces the observed relic density of DM in the Universe. In the grey region perturbative unitarity is violated.

followed by decays of s into SM particles, which allows for a significant relaxation of relic density constraints.

For various choices of m_χ , Duerr et al. [39] consider a range of m_s , $m_{Z'}$ and m_χ , and perform global scans over all other free parameters in the DH model (see Section 2.2 below), as shown in Figure 2.4, to identify regions in which the model has not yet been excluded by existing constraints. Particularly for $m_\chi \geq 200$ GeV, it is found that the model evades all existing constraints for a large region of m_s and $m_{Z'}$ (up to ~ 1000 GeV in m_s and up to ~ 2500 GeV in $m_{Z'}$).

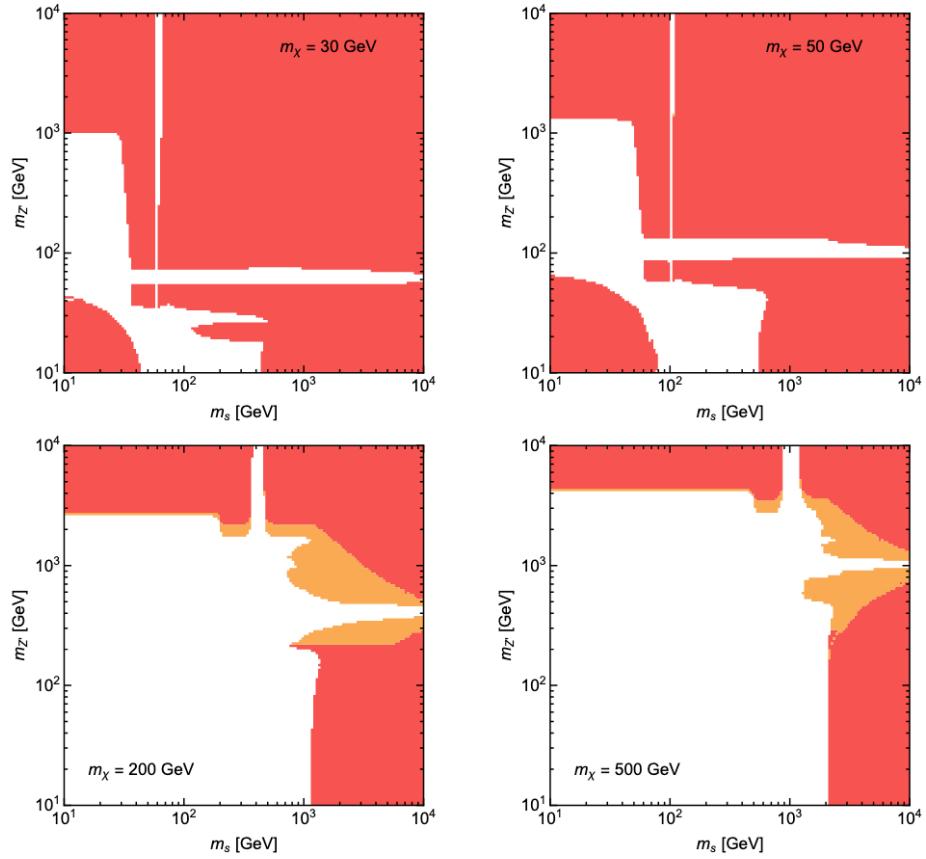


Figure 2.4: Global scans of the DH model, with Z' and DH as portal mediators, performed by Duerr et al. in Ref. [39]. The red shaded region is excluded for all possible combinations of couplings, while in the white region all constraints can be evaded. In the orange shaded region it is not possible to exclude large values of g_q corresponding to $\Gamma'_Z/m_{Z'} > 0.3$, where Γ'_Z is the decay width of the Z' mediator. Figure from © [39].

1160 **2.2 Model Description**

1161 The Dark Higgs (DH) model presented in Refs. [40] and [39] belongs to a wider class of
 1162 simplified dark sector models, which hypothesize that DM interacts with particles of
 1163 the SM via the exchange of one or more new mediators that act as so-called “portals”
 1164 between the SM and the dark sector. Candidate portals are broadly categorized
 1165 according to the portal mediator into vector-, neutrino-, Higgs- and axion-mediated
 1166 portals (for reviews, see Refs. [83, 84], [85], [86] and [87], respectively). In the DH
 1167 model, DM is assumed to be a Majorana fermion, which means that - like the photon,
 1168 for example - it is its own antiparticle. The model postulates that DM interacts with
 1169 the SM via both vector-mediated and Higgs-mediated portals.

1170 A new gauge group called the $U(1)'$, with an associated vector gauge boson referred
 1171 to as the Z' , is introduced as an extension of the SM gauge group presented in Chapter
 1172 1. Both the DM and the Z' are assumed to acquire their mass from a new Higgs field
 1173 with vacuum expectation value w , which gives rise to a new physical Higgs boson,
 1174 referred to as the DH boson s . The DM acquires an axial coupling to the Z' , such
 1175 that all three dark sector particles interact with one another.

1176 The interactions of the $U(1)'$ gauge group within the dark sector are expressed
 1177 by the interaction Lagrangian (of which more details can be found in Appendix A of
 1178 Ref. [39]):

$$\mathcal{L}_{U(1)'} = -\frac{1}{2}g_\chi Z'^\mu \bar{\chi} \gamma^5 \gamma_\mu \chi - g_\chi \frac{m_\chi}{m_{Z'}} s \bar{\chi} \chi + 2g_\chi Z'^\mu Z'_\mu (g_\chi s^2 + m_{Z'} s) \quad (2.1)$$

1179 Considering each term in $\mathcal{L}_{U(1)'}$ individually:

$$\mathcal{L}_{\chi, Z'} = -\frac{1}{2}g_\chi Z'^\mu \bar{\chi} \gamma^5 \gamma_\mu \chi \quad (2.2)$$

1180 describes the axial coupling between the DM χ and the Z' , with coupling strength
 1181 g_χ .

$$\mathcal{L}_{\chi, \text{DH}} = -g_\chi \frac{m_\chi}{m_{Z'}} s \bar{\chi} \chi = -\frac{y_\chi}{2\sqrt{2}} s \bar{\chi} \chi \quad (2.3)$$

1182 describes the coupling between the DM χ and the DH field S , where the associated
 1183 coupling strength y_χ on the right-hand side of Eq. 2.3 has the following dependence
 1184 on the masses and coupling strength of the DM and the Z' : $y_\chi = 2\sqrt{2}g_\chi \frac{m_\chi}{m_{Z'}}$.

$$\mathcal{L}_{Z', \text{DH}} = 2g_\chi Z'^\mu Z'_\mu (g_\chi s^2 + m_{Z'} s) \quad (2.4)$$

describes the interaction between the Z' and the DH field.

Motivated by models of gauged baryon number [40, 39, 88], the SM quarks are charged under the $U(1)'$ gauge group, and as a result have vector couplings with the Z' , which are described by the following interaction Lagrangian:

$$\mathcal{L}_{q,Z'} = -g_q Z'^\mu \bar{q} \gamma_\mu q \quad (2.5)$$

The quark- Z' coupling provides a mechanism for vector-mediated portal interactions between the SM and the dark sector. Axial vector couplings of the Z' to quarks would also be expected to produce a signature at the LHC, but are neglected in the model for simplicity [40].

The most general Lagrangian describing the scalar couplings of the SM Higgs h and the DH s is given by (see, for example, Ref. [89]):

$$\mathcal{L}_{\text{scalar}} = -\lambda_h \left(H^\dagger H - \frac{v^2}{2} \right) - \lambda_s \left(S^\dagger S - \frac{w^2}{2} \right) - \lambda_{hs} \left(H^\dagger H - \frac{v^2}{2} \right) \left(S^\dagger S - \frac{w^2}{2} \right) \quad (2.6)$$

where $H = \frac{1}{\sqrt{2}}(0, h + v)$ represents the SM Higgs field with vacuum expectation value v , and $S = \frac{1}{\sqrt{2}}(s + w)$ represents the DH field. The third term with coupling λ_{hs} mixes the DH and SM Higgs fields, such that the physical mass eigenstates h' and s' are a superposition of the scalars h and s :

$$\begin{pmatrix} h' \\ s' \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} h \\ s \end{pmatrix} \quad (2.7)$$

where the “mixing angle” θ is related to SM and DH field couplings λ_h and λ_s and the vacuum expectation values v and w by:

$$\tan 2\theta = \frac{\lambda_{hs}vw}{\lambda_h v^2 - \lambda_s w^2} \quad (2.8)$$

Therefore, if the mixing between the SM and DH fields is nonzero ($\sin \theta > 0$), the physical Higgs eigenstate in the SM Lagrangian becomes $h \cos \theta + s \sin \theta$. This opens up the possibility for interactions between the physical DH and SM particles, with coupling strengths scaled by $\sin \theta$. For example, the decays of the DH to SM fermions f and vector bosons V are described by the following Lagrangians [39, 22]:

$$\mathcal{L}_{h f \bar{f}} = -g_{h f \bar{f}} \bar{f} f h \sin \theta \quad (2.9)$$

1206 and

$$\mathcal{L}_{hVV} = -\delta_V g_{hVV} V_\mu V^\mu h \sin \theta \quad (2.10)$$

1207 where (from SM physics):

$$g_{h\bar{f}\bar{f}} = \frac{f}{v}, \quad g_{hVV} = \frac{2m_V^2}{v}, \quad \delta_W = 1, \quad \delta_Z = \frac{1}{2} \quad (2.11)$$

1208 2.2.1 Free Parameters in the Model

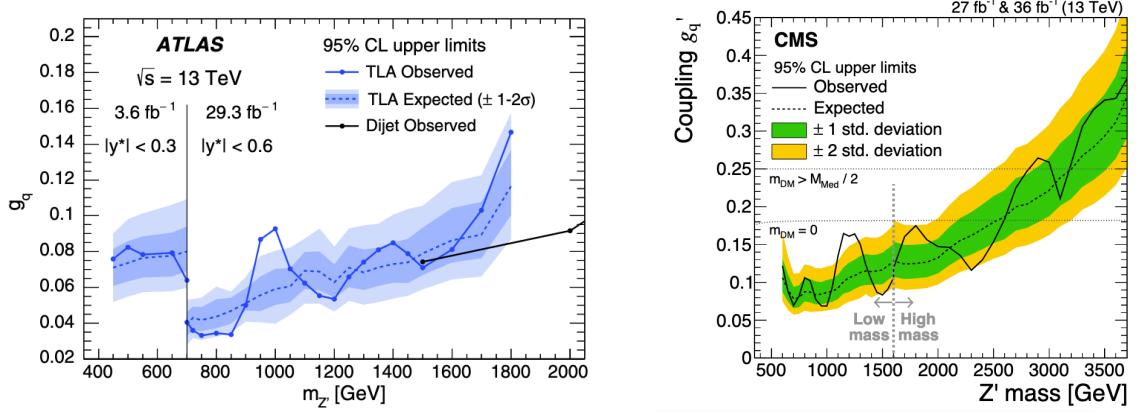
1209 The free parameters in the DH model are as follows:

- 1210 • **\mathbf{m}_χ :** The mass of the DM particle χ .
- 1211 • **$\mathbf{m}_{Z'}$:** The mass of the hypothetical Z' boson in the dark sector, which mediates
1212 interactions between SM quarks q and the DM χ .
- 1213 • **\mathbf{m}_s :** The mass of the hypothetical DH boson, whose associated Higgs field
1214 provides a mechanism for particles in the dark sector to acquire mass, and
1215 which can decay to SM particles via a small mixing with the SM Higgs boson.
- 1216 • **\mathbf{g}_q :** Strength of the vector coupling between SM quarks and the Z' boson.
- 1217 • **\mathbf{g}_χ :** The coupling strength between DM and the Z' boson.
- 1218 • **θ :** The mixing angle between the DH and the SM Higgs boson. The coupling
1219 strength associated with the decay of the DH to SM particles scales with $\sin \theta$.

1220 Some of these parameters are already constrained by other results. Mixing be-
1221 tween the DH and the SM Higgs in the model would introduce unobservable decay
1222 modes of the SM Higgs to dark sector particles, which would modify the yield of
1223 events observed at the LHC in which a Higgs boson is produced from the SM expec-
1224 tation. Given the agreement between the observed yield and the SM expectation [90,
1225 91], an upper limit of $|\sin \theta| < 0.25$ can be placed (assuming that no other appreciable
1226 unobservable decay channels exist), as discussed in Section 4.2.1 of Ref. [92].

1227 Dijet resonance searches, discussed in Section 2.1.1, also place upper bounds on
1228 the coupling strength g_q of the Z' to SM quarks by searching for a resonant peak in
1229 the dijet spectrum associated with the scenario shown in Figure 2.2b in which a Z' is
1230 produced by high-energy $q\bar{q}$ collisions at the LHC, and decays immediately back to

1231 a pair of quarks. Based on the agreement with SM predictions observed in the most
 1232 recent dijet resonance searches performed by the ATLAS and CMS experiments, [93,
 1233 94, 95, 96] upper bounds on g_q ranging from 0.04 to 0.4, depending on $m_{Z'}$, have been
 1234 established. Figure 2.5 shows upper bounds on g_q from two recent dijet searches, one
 1235 of which (Figure 2.5a) uses “trigger-level jets” [94] to boost sensitivity in the low- $m_{Z'}$
 1236 range.



(a) Limits from ATLAS low-mass dijet resonance search [94]. Black solid line shows observed limits from a previous dijet resonance search [97]. Limits for $m_{Z'} < 700$ GeV are obtained using a subset of ATLAS data collected with non-standard trigger (see Ref. [94] for details). Figure from Ⓜ [94].

(b) Limits from a recent CMS dijet resonance search [95]. The high-mass (low-mass) region is performed with $m_{jj} > 1.25$ TeV ($m_{jj} > 0.49$ TeV). See Section 5.1 of Ref. [95] for details pertaining to the horizontal dotted lines. Figure from Ⓜ [95].

Figure 2.5: Upper bounds on the g_q coupling between the Z' and quarks, as reported by recent dijet resonance searches performed by ATLAS (left) and CMS (right).

1237 2.3 Search for the Dark Higgs Model at the LHC

1238 For certain choices of the model parameters discussed in Section 2.2.1, the DH model
 1239 predicts a unique and measurable signature by which it could be detected at the LHC.
 1240 This LHC signature, of which some of the most contributing Feynman diagrams are
 1241 shown in Figure 2.1, would occur when a q in one of the colliding protons pair-
 1242 annihilates with a \bar{q} in the other proton to produce a Z' , which subsequently decays
 1243 to DM. A DH s is emitted either from an initial-state q or from the Z' mediator, and
 1244 subsequently decays to a pair of SM particles. This produces a final state with the SM
 1245 products from the s decay recoiling against E_T^{miss} in the detector due to the undetected

1246 DM pair. In particular, assuming that the Z' is relatively low in mass compared with
 1247 the centre of mass energy of the $q\bar{q}$ collision, the Z' may be imparted with a large
 1248 momentum (a.k.a. boost). As a result, the diagrams shown in Figures 2.1a and 2.1b
 1249 in which the s is emitted from the Z' can produce highly boosted, collimated SM
 1250 decay products. These boosted SM decay products make the signature distinct from
 1251 typical $E_T^{\text{miss}} + X$ final states in which the SM products ('X') recoiling against the DM
 1252 are assumed to be produced exclusively as initial state radiation, as in Figure 2.1c.

1253 2.3.1 Dark Higgs Decay Channels

1254 As discussed in Section 2.2, since the DH decays to SM particles via mixing with
 1255 the SM Higgs boson, its decay mechanisms - including the branching fractions for its
 1256 decay channels to various SM particles - would be analogous to that of the SM Higgs.
 1257 As a result, the signature of boosted SM decay products recoiling against E_T^{miss} may
 1258 be used to probe various ranges of m_s in the model depending on the particular choice
 1259 of SM products (a.k.a. “decay channel”, or simply “channel”) from the s decay that
 1260 the search targets. Figure 2.6 shows the branching fraction that the SM Higgs - and
 1261 consequently the DH s - would be expected to have to SM particles if its mass were
 1262 allowed to float in the SM. For low m_s , the decay to $b\bar{b}$ is dominant. At $m_s \approx 160$ GeV,
 1263 the decay to WW becomes kinematically accessible¹, and is the dominant s decay
 1264 channel for $m_s \gtrsim 160$ GeV. The decay to ZZ also becomes kinematically accessible
 1265 at $m_s \approx 180$ GeV, though its branching fraction remains sub-dominant compared
 1266 with the WW decay channel. The decay to SM Higgs bosons $s \rightarrow HH$ opens up
 1267 at $m_s \approx 250$ GeV, and for $m_s > 250$ GeV the branching fraction for this HH
 1268 decay channel becomes appreciable, though still smaller than the WW channel. This
 1269 suggests that searches in all four of these s decay channels could complement one
 1270 another to collectively probe the full available m_s range, as summarized in Table 2.1.

1271 2.3.2 Completed and Ongoing Searches for the Dark Higgs Model

1272 The search presented in this thesis is one of several ongoing and completed searches
 1273 for the DH model at the LHC, which target different decay channels of the DH after
 1274 it is emitted in the model’s LHC signature, of which some of the most contributing
 1275 Feynman diagrams are shown in Figure 2.1.

¹A particular decay mode $X \rightarrow YY$ generally becomes kinematically accessible for $m_X > 2m_Y$, because in this regime the on-shell parent particle X has a sufficient rest mass energy to decay to the daughter products YY while satisfying energy conservation.

Table 2.1: Summary of DH decay channels to SM particles that could be targeted to probe various m_s ranges in DH model.

m_s Range	Sensitive DH Decay Channels
$m_s < 160$ GeV	$s \rightarrow b\bar{b}$
160 GeV $< m_s < 180$ GeV	$s \rightarrow WW$
180 GeV $< m_s < 250$ GeV	$s \rightarrow WW, s \rightarrow ZZ$
$m_s > 250$ GeV	$s \rightarrow WW, s \rightarrow ZZ, s \rightarrow HH$

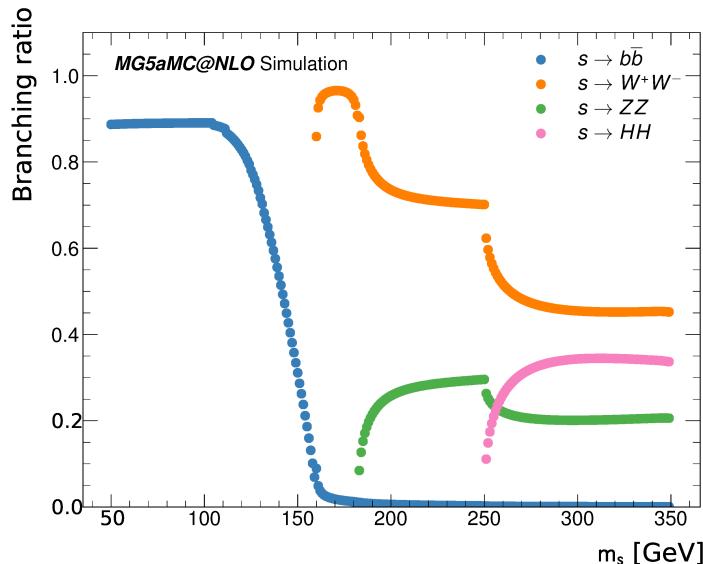


Figure 2.6: m_s dependence of the branching fraction for decays of the DH boson to SM particles. The branching fractions of the $s \rightarrow WW$ and $s \rightarrow ZZ$ decay channels experience a drop-off at $m_s = 2m_H (= 250$ GeV), above which the decay to $s \rightarrow HH$ to SM Higgs becomes kinematically accessible. Note that the branching fraction for $s \rightarrow WW$ decays is only simulated for $m_s \geq 160$ GeV, but remains appreciable ($\gtrsim 0.1$) for $m_s \gtrsim 120$ GeV. Figure from © [98].

1276 **Model Parameters Probed by LHC Searches**

1277 Given the number of parameters in the DH model, computational resource limitations
 1278 make it impractical to scan over all parameters when performing dedicated searches
 1279 for the model at the LHC. Therefore, in the search presented in this thesis - and in all
 1280 other searches for the model performed to date at the LHC - the coupling constants
 1281 and mixing angle θ are fixed as follows:

1282 • $g_q = 0.25$

1283 • $g_\chi = 1$

1284 • $\sin \theta = 0.01$

1285 The choice of $\sin \theta = 0.01$ is well within the range $|\sin \theta| < 0.25$ required by
 1286 measurements of the SM Higgs - see above discussion in Section 2.2.1. The values
 1287 of the coupling constants g_χ and g_q are chosen to be consistent with other searches
 1288 performed at the LHC for dark sector benchmark models [99] involving a Z' mediator,
 1289 which facilitates comparisons with these existing searches. It is worth noting that the
 1290 choice of $g_q = 0.25$ is in fact ruled out in the approximate range $500 \text{ GeV} < m_{Z'} <$
 1291 3000 GeV by the recent dijet searches discussed above in Section 2.2.1. However, the
 1292 choice of $g_q = 0.25$ is maintained in this search because it is considered important to
 1293 maintain consistency with other searches performed with this benchmark choice. In
 1294 particular, it allows for the search results to be easily compared with other searches
 1295 for the DH model, which used the same parameter choices. It is also worth noting
 1296 that the computational procedure required to interpret the analysis with an arbitrary
 1297 BSM physics model has been preserved and automated in the RECAST framework
 1298 [3]. Using this framework, it should be straightforward to re-interpret the search
 1299 results in the future, if needed, for a DH model with alternative choices of the fixed
 1300 parameters.

1301 In the search for the DH model presented in this thesis, the DM mass m_χ is fixed
 1302 to 200 GeV , consistent with the choice used by the other two published searches
 1303 for the DH model [98, 100] performed by the ATLAS collaboration. This choice of
 1304 $m_\chi = 200 \text{ GeV}$ was found in early sensitivity studies undertaken by the search in the
 1305 $s \rightarrow WW(q\bar{q}q\bar{q})$ channel [100] to predict a relatively large measurable yield of events
 1306 in the ATLAS detector from the DH process, which allows the searches to probe a
 1307 large range of the remaining model parameters relative to other m_χ choices. The

1308 recently-published search by the CMS collaboration in the $s \rightarrow WW(\ell\nu\ell\nu)$ channel
 1309 [101] considered four candidate m_χ : 100, 150, 200 and 300 GeV.

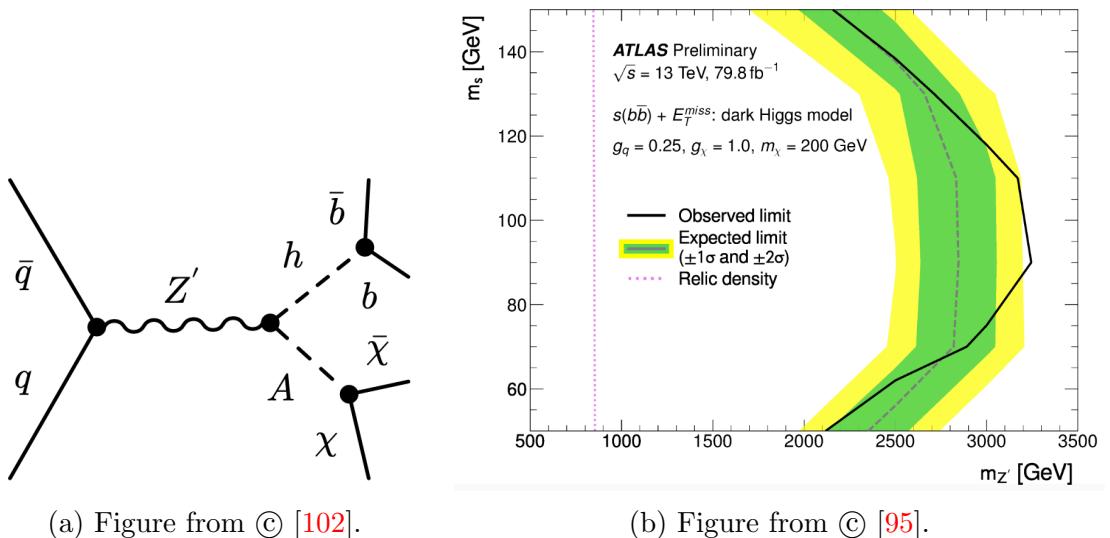
1310 The mediator masses m_s and $m_{Z'}$ are left as floating parameters in the searches.
 1311 The m_s range covered by the search depends on the range to which the s decay
 1312 channel considered in each search is sensitive on the basis of the predicted branching
 1313 ratios in Figure 2.6, as summarized in Table 2.1. In addition, the sensitivity of
 1314 searches for the DH model using this boosted SM+ E_T^{miss} signature drops off sharply for
 1315 $m_s > 2m_\chi$, because in this m_s range the decay mode $s \rightarrow \chi\chi$ becomes kinematically
 1316 accessible, and the radiated s would be expected to decay predominantly via this
 1317 invisible channel, rather than to visible SM particles via mixing with the SM Higgs.

1318 The available $m_{Z'}$ range to which these searches may be sensitive covers the ap-
 1319 proximate range of 500 to 3500 GeV. This range is bound from above by the fall-off
 1320 in cross section for producing such a massive Z' mediator from $q\bar{q}$ collisions at the
 1321 LHC. The production rate of the model begins to fall off quite rapidly for $m_{Z'}$ be-
 1322 low ~ 500 GeV due to the minimum virtual mediator mass required to produce to
 1323 a pair of 200 GeV DM particles (in addition to radiating the DH as shown in the
 1324 contributing Feynman diagrams in Figures 2.1a and 2.1b).

1325 $s \rightarrow bb$ Channel

1326 The $s \rightarrow bb$ decay channel was probed by the search presented in Ref. [98]. The search
 1327 used the RECAST framework [3] to re-interpret an earlier DM search [102], which
 1328 targeted a signature of DM produced in association with a Higgs boson decaying
 1329 to b quarks, in the context of the DH model. This re-interpretation was possible
 1330 because the model that was used to optimize and interpret the original search, for
 1331 which the most contributing Feynman diagram is shown in Figure 2.7a, is very similar
 1332 in structure to the DH model, and predicts the same final state of a boosted $b\bar{b}$ pair
 1333 recoiling against E_T^{miss} in the detector. As shown in Figure 2.7b, this re-interpretation
 1334 was able to place upper limits on the $m_{Z'}$ in the DH model ranging from ~ 2000 GeV
 1335 to 3200 GeV, with the given choices of coupling strengths and $\sin\theta$, for m_s in the
 1336 range $50 \text{ GeV} < m_s < 150 \text{ GeV}$.

1337 There is also a dedicated search in the $s \rightarrow bb$ channel under development within
 1338 the ATLAS collaboration. In addition to optimizing the search strategy to maximize
 1339 sensitivity to the DH model, this search also plans to improve upon the earlier search
 1340 performed in the $s \rightarrow bb$ channel by scanning over additional model parameters such



(a) Figure from © [102].

(b) Figure from © [95].

Figure 2.7: Left: Most important Feynman diagram for the original DM search [102] that was re-interpreted in Ref. [98] to probe the DH model in the $s \rightarrow b\bar{b}$ channel. Figure from © [102]. Right: Exclusion limits on the mediator masses m_s and $m_{Z'}$ in the DH model from the search in the $s \rightarrow b\bar{b}$ channel. Values of m_s and $m_{Z'}$ to the left of the solid black line are excluded by the search. The dashed pink line reproduces the relic density of DM observed in the Universe for the choices of coupling constants $g_\chi = 1$ and $g_q = 0.25$ used in the search. Figure from © [98].

1341 as g_q and m_χ that were fixed in the earlier search, targeting values for these parameters
1342 that would maximize the sensitivity of the search to the DH model while reproducing
1343 the observed relic abundance of DM in the Universe.

1344 **$s \rightarrow WW$ and $s \rightarrow ZZ$ Channels**

1345 The $s \rightarrow WW$ and $s \rightarrow ZZ$ ($s \rightarrow VV$) channels, which collectively dominate the
1346 branching fraction of s decays for $m_s > 160$ GeV, are somewhat more complex in
1347 terms of their final state in the detector compared with the $s \rightarrow bb$ channel. This
1348 is because, whereas the $b\bar{b}$ pair in the latter decay channel produce a characteristic
1349 signature of either two distinct b -tagged jets² or one two-pronged large-radius jet³ in
1350 the calorimeter, the vector bosons in the $s \rightarrow WW$ and $s \rightarrow ZZ$ final states can
1351 decay via several possible channels, which leads to a number of distinct final states in
1352 the detector for these $s \rightarrow VV$ channels. For this reason, a number of different LHC
1353 searches in the $s \rightarrow VV$ channels have been completed or are ongoing, each of which
1354 targets only a subset of the VV decay channels.

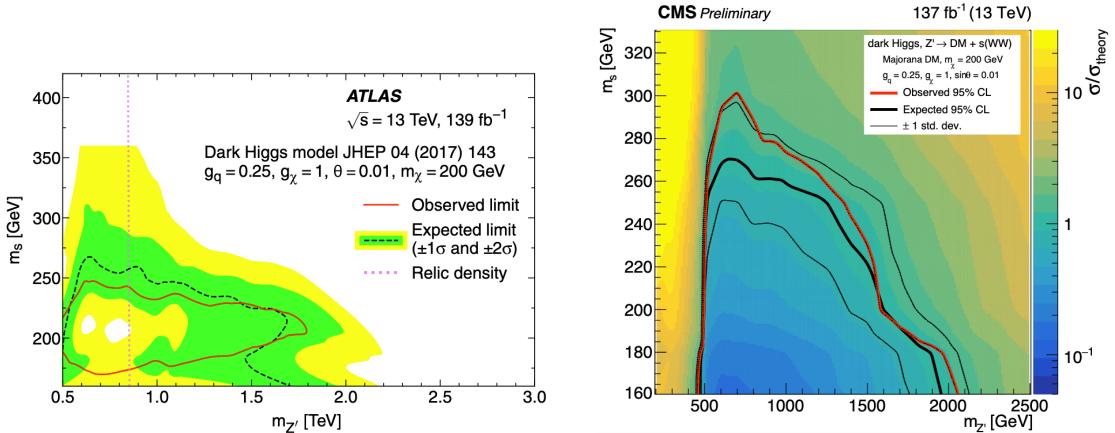
1355 Table 2.2 summarizes the available decay channels for the WW and ZZ final
1356 states, as well as the branching fraction of each channel and information regarding
1357 any completed or ongoing efforts to search for the DH model in each VV decay
1358 channel. Figure 2.8 shows the range of mediator masses ($m_s, m_{Z'}$) that are excluded,
1359 for $m_\chi = 200$ GeV, by completed searches performed by ATLAS in the $s \rightarrow VV(q\bar{q}q\bar{q})$
1360 channel [100], and by CMS in the $s \rightarrow WW(\ell\nu\ell\nu)$ channel [101]. As would be
1361 expected on the basis of the branching fractions in Figure 2.6, these searches exclude
1362 parameter space roughly in the range $m_s > 160$ GeV. The search presented in this
1363 thesis, which covers the $s \rightarrow WW$ decay channel in the semileptonic final state
1364 ($WW \rightarrow q\bar{q}\ell\nu$), complements these existing searches, and extends the excluded range
1365 of m_s and $m_{Z'}$.

²See Section 5.1.2 for a description of the method used to tag jets in the calorimeter as having originated from a b quark.

³see Section 3.4.2 for more details on hadronic jets in the calorimeter

Table 2.2: Summary of decay channels for WW and ZZ pairs and their branching fractions.

Decay channel	Branching fractions [22] (from WW or ZZ pair)	Search effort(s)
$WW \rightarrow q\bar{q}q\bar{q}$	0.45	ATLAS published search: Ref. [100]
$ZZ \rightarrow q\bar{q}q\bar{q}$	0.49	
$WW \rightarrow q\bar{q}\ell\nu$ ($\ell = e$ or μ)	0.29	This thesis
$ZZ \rightarrow q\bar{q}\ell\ell$ ($\ell = e$ or μ)	0.094	Ongoing within ATLAS collaboration
$ZZ \rightarrow q\bar{q}\nu\nu$	0.14	N/A (expected sensitivity too low)
$WW \rightarrow \ell\nu\ell\nu$ ($\ell = e$ or μ)	0.046	CMS published search: Ref. [101]. Effort also ongoing within ATLAS collaboration.
$ZZ \rightarrow LLLL$ ($L = \ell$ or ν)	0.071	N/A (expected sensitivity too low)



(a) Exclusion limits from search in $s \rightarrow VV(q\bar{q}q\bar{q})$ channel. Figure from © [100]. (b) Exclusion limits from search in $s \rightarrow WW(\ell\nu\ell\nu)$ channel. Figure from © [101].

Figure 2.8: Range of $(m_s, m_{Z'})$ in the DH model excluded by searches performed by ATLAS in the $s \rightarrow VV(q\bar{q}q\bar{q})$ channel (left), and by CMS in the $s \rightarrow WW(\ell\nu\ell\nu)$ channel (right), for the following choices of remaining parameters in the model: $g_q = 0.25$, $g_\chi = 1$, $\sin \theta = 0.01$, $m_\chi = 200$ GeV.

1366

Chapter 3

¹³⁶⁷ Introduction to the LHC and the ATLAS Detector

¹³⁶⁸ The Large Hadron Collider (LHC) [52] is a circular proton-proton (pp) collider, which
¹³⁶⁹ resides in a 27 km tunnel near the European Organization for Nuclear Research
¹³⁷⁰ (CERN). Superconducting magnets are used to accelerate counter-rotating bunched
¹³⁷¹ proton beams to near the speed of light, and direct the beams into head-on collisions
¹³⁷² at four interaction points around the ring. The collisions take place at a world-
¹³⁷³ leading centre of mass energy of up to 13 TeV. Each interaction point is surrounded
¹³⁷⁴ by a detector, which measures the energetic debris of particles produced by the high
¹³⁷⁵ energy collisions to perform precision measurements of the SM and search for new
¹³⁷⁶ physics.

¹³⁷⁷ The large 13 TeV centre of mass energy of the collisions makes it possible for the
¹³⁷⁸ colliding proton constituents, known as “partons”, to pair annihilate and subsequently
¹³⁷⁹ produce massive unstable particles such as the Higgs boson, which cannot presently
¹³⁸⁰ be produced by any other experimental means. Experiments at the LHC can study
¹³⁸¹ hypothetical models of physics beyond the SM (“BSM physics”) by searching for
¹³⁸² evidence of the production of the massive particles involved in these models from
¹³⁸³ their subsequent decay to SM particles.

¹³⁸⁴ The LHC collides protons at an approximate rate of 1 billion collisions per second,
¹³⁸⁵ \sim 100 times higher than the proton collision rate at the Tevatron collider [103], which
¹³⁸⁶ operated from 1983-2011 and collided protons and anti-protons ($p\bar{p}$) at a peak centre
¹³⁸⁷ of mass energy of 1.8 TeV. Over several years of data-taking, the high collision rate
¹³⁸⁸ at the LHC has enabled experiments to collect rich data sets. The large data sets can
¹³⁸⁹ be used to probe new phenomena in highly selective final states.

1390 **3.1 The Parton Model**

1391 Before discussing pp collisions at the LHC in detail, it is important to first intro-
 1392 duce the parton model, which describes the substructure of protons involved in the
 1393 collisions.

1394 The proton has an internal structure comprised of constituent quarks, antiquarks
 1395 and gluons - collectively known as “partons” - and their interactions (for a review
 1396 of the parton model and its applications, see [104]). When a proton collides with
 1397 another particle in particle colliders such as the LHC, the probability density $f(x, Q^2)$
 1398 that a particular species of parton, for example a quark with “up” flavour u , will be
 1399 involved in the collision is a function of both the fraction x of the proton’s momentum
 1400 carried by the parton, and the squared momentum scale Q^2 of the collision. Detailed
 1401 parametrized models of the parton distribution function (PDF), such as MSHT20
 1402 [105] have been developed using combined fits to data from deep inelastic scattering
 1403 (DIS) experiments at proton colliders. MSHT20 PDF models at Q^2 of 10 GeV^2 and
 1404 10^4 GeV^2 are shown in Figure 3.1.

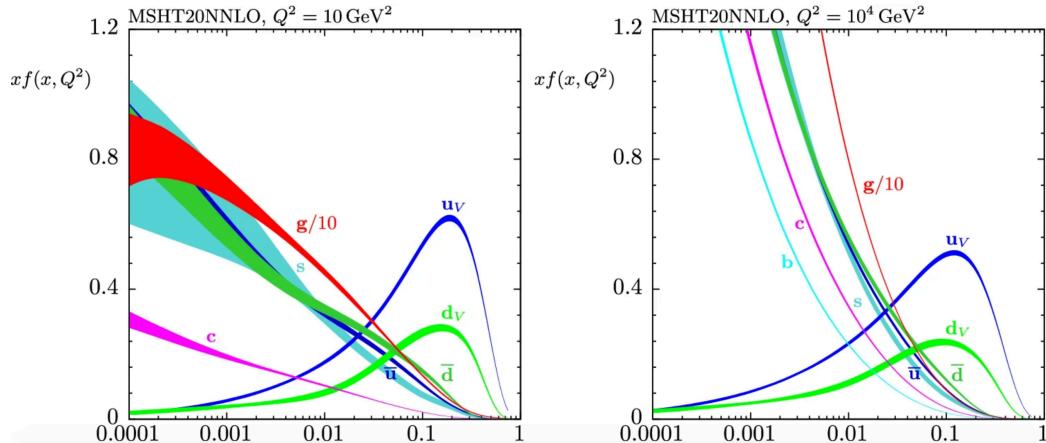


Figure 3.1: Parton distribution functions, with respect to the proton momentum fraction x carried by the parton, modelled with MSHT20 at $Q^2 = 10 \text{ GeV}^2$ and 10^4 GeV^2 . Figure from © [105].

1405 Based on the PDFs shown in Figure 3.1, u and d quarks carry the highest prob-
 1406 ability density for parton momentum fractions above $\sim 10\%$, with the u carrying
 1407 approximately double the probability density of the d . These dominant quarks are
 1408 known as the proton’s “valence” quarks, of which there are two u and one d , and they
 1409 carry the proton’s quantum numbers.

1410 **3.2 Decay Processes from Parton Collisions**

1411 As discussed in Section 1.1.3, each process by which colliding partons may annihilate
 1412 to form observable products proceeds with a certain cross section σ , which is related
 1413 to the production rate $\frac{dN}{dt}$ of the process by the beam luminosity \mathcal{L} :

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

1414 Figure 3.2 shows a summary of cross sections for the production of SM particles -
 1415 or particle combinations (eg. “ Wt ” represents the production of a W boson along with
 1416 a top quark) - from pp collisions at the LHC, as measured by the ATLAS detector.

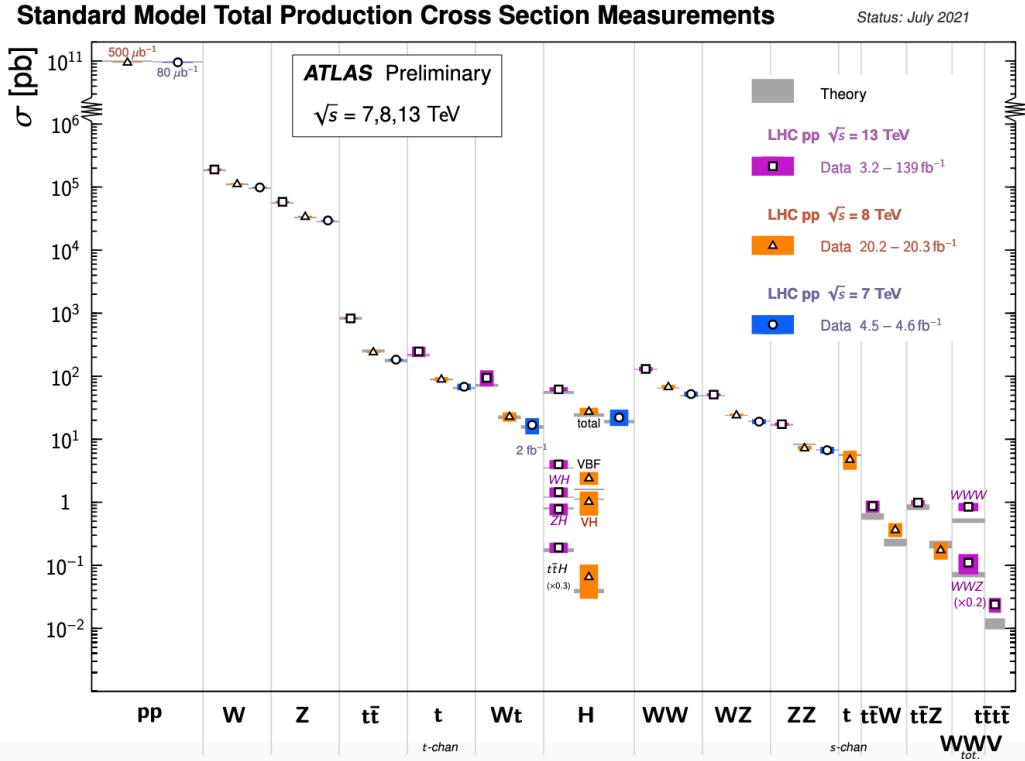


Figure 3.2: Summary of SM cross sections for particle production processes measured by the ATLAS detector. Figure from © [106].

1417 **3.2.1 Branching Fractions and W Boson Decays**

1418 Unstable particles produced by the parton collisions will subsequently decay to less-
 1419 massive particles, typically with multiple possible mechanisms, also known as “chan-
 1420 nels”, by which the decay can occur. Each such channel has an associated “branching

fraction”, which quantifies the relative probability with which the decay will proceed by the given channel. The search presented in this thesis focuses on DM production in association with a pair of oppositely-charged W bosons. Figure 3.3 shows the two W boson decay routes. Due to energy and momentum conservation, W bosons can only decay to a pair of particles whose combined mass is smaller than the W mass. Charge conservation additionally requires that the decay products have a combined charge equal to that of the parent W boson. These two requirements allow the W to decay either “hadronically” to a quark-antiquark pair with one up-type quark/antiquark (U) and one down-type (D), or “leptonically” to a charged lepton (L) and a neutrino (ν).



Figure 3.3: W boson decay mechanisms.

3.3 Detectors at the LHC

The DM search presented in this thesis uses data collected from the ATLAS (A Toroidal LHC ApparatuS) detector [53]. ATLAS is one of four particle detectors at the LHC, which are designed to measure the energetic debris of particles produced by high energy particle collisions to perform precision measurements of the SM and search for new physics using the resulting particle collision data. This section briefly introduces each of the four particle detectors at the LHC, and what each contributes to the LHC physics programme.

The two largest, **ATLAS** (A Toroidal LHC ApparatuS) [53] and **CMS** (Compact Muon Solenoid) [54], are both general-purpose detectors designed to record all SM decay products from the collisions, with the exception of neutrinos, which pass through due to their very low interaction cross sections. Thanks to their near-complete detection of decay products, data from these detectors can be used to study a wide range of physics processes resulting from the collisions, including both measurements of the SM and searches for new physics beyond the SM. By taking a general-purpose approach, rather than specializing in the study of one particular collision or decay

process, these experiments seek to maintain sensitivity to the broadest possible range of particle processes, in the hopes of allowing physicists to detect and measure new physics processes in whatever form they may take. While the physics goals of these two detectors are very similar, they are accomplished using different detector designs and technologies, and as such they are able to produce complementary physics results.

The Large Hadron Collider beauty (LHCb) detector [55] is designed to measure heavy quark (b and c) decays resulting from pp collisions. Precise measurement of heavy quark decays are of particular interest for the study of CP violation in the SM, and in the search for potential sources of CP violation beyond the SM. Rather than providing full coverage of all collision products, the LHCb detector is comprised of a series of sub-detectors that provide “forward angle” coverage to detect particles produced with a large boost along the direction of one of the two proton beams. This forward region is of particular interest for measurements of heavy quark decays, because this is the angular region in which heavy quark pairs are predominantly produced by high-energy collisions.

A Large Ion Collider Experiment (ALICE) [107] is designed to measure the products of heavy-ion collisions produced during special LHC runs in which the proton beams are replaced by Pb beams, which are collided at a centre of mass energy of 5 TeV. The high-energy Pb collisions produce a sufficiently high temperature and density to form an unbound state of quarks and gluons known as “quark-gluon plasma” that would have occurred in the early universe. The study of this exotic state could give novel insights into the theory of quantum chromodynamics¹, including the phenomena of quark colour confinement and chiral-symmetry restoration (for reviews, see Refs. [108, 109]).

3.4 Introduction to the ATLAS detector

The ATLAS detector [53], shown schematically in Figure 3.5, is the largest detector by volume to have been built at any particle collider, with a length of 44 m and a height of 25 m, constituting a total weight of approximately 7,000 tonnes. The tremendous amount of detector material is needed to absorb and measure the highly energetic decay products of the pp collisions with sufficient resolution to enable a detailed reconstruction of the particles involved and their kinematic properties. Such a

¹The theory of quantum chromodynamics (QCD), introduced in Section 1.1.6, describes interactions that proceed via the strong force.

complete and detailed reconstruction of the collisions and subsequent decay processes enable physicists to carry out the impressive range of physics goals shared by the ATLAS and CMS collaborations. These goals include precision measurements of the SM, which profit both from the enormous collision rate, and from the large centre of mass collision energy that enables on-shell production of all known SM particles. The detector is also designed to be sensitive to as wide a range of new physics signatures as possible. Particular emphasis was placed on designing the detector to be sensitive to the anticipated production modes of the Higgs boson, which was jointly discovered by the ATLAS and CMS collaborations in 2012 [6, 7].

The detector provides full 4π coverage around the pp interaction point, with the exception of the beam pipe. It consists of several layers of sub-detectors, each of which is specialized for recording specific kinematic information and particle types.

The ATLAS detector is described spatially using the standard coordinate system of (x, y, z) coordinates and (θ, ϕ) angles shown in Figure 3.5b. The origin of the coordinate system is placed at the nominal interaction point of the colliding proton beams, and the z axis lies along the beam line. The angle of a particle or detector component in the plane transverse to the beam line is given by the angle ϕ , and its angle relative to the beam line is given by θ . The “pseudorapidity” η is a quantity related θ according to:

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] \quad (3.2)$$

Pseudorapidity is often used rather than θ because differences $\Delta\eta$ in pseudorapidity are invariant under Lorentz boosts along the z axis.

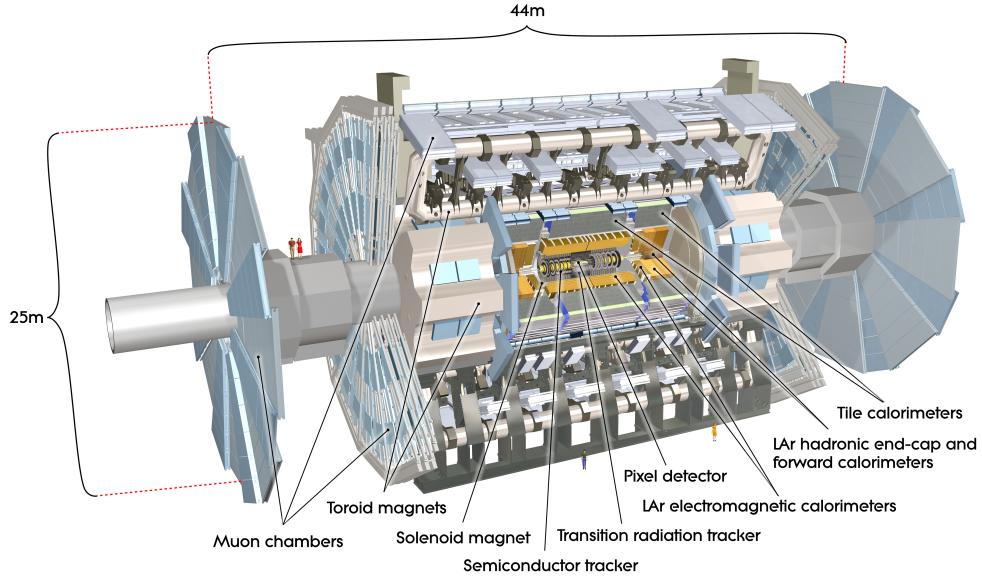


Figure 3.4: Schematic diagram of the ATLAS detector. Figure from © [53]

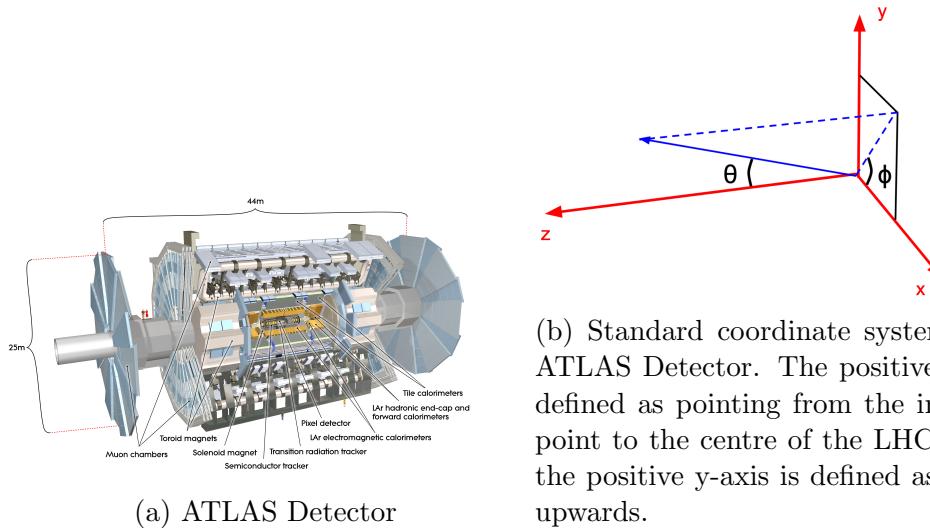


Figure 3.5: Left: Schematic diagram of the ATLAS detector (figure from © [53]). Right: Standard coordinate system used for the ATLAS detector.

1499

The ATLAS sub-detectors are described in some detail in the following sections.

1500 **3.4.1 The Inner Detector**

1501 The inner detector (ID) [110], located nearest the beam pipe, is specialized for charged
 1502 particle tracking. It is immersed in a 2T magnetic field oriented parallel to the beam
 1503 pipe, which bends the trajectories (“tracks”) of electrically charged particles as they
 1504 pass through the field. Three distinct but complementary high-resolution tracking
 1505 technologies are employed along with pattern recognition tools to map the trajectories
 1506 of charged particles passing through the ID. The tracking is accomplished with as little
 1507 material as possible in the ID, such that the particle trajectories can be mapped with
 1508 minimal scattering and energy loss before they reach the calorimeters.

1509 Tracks from the inner detector are reconstructed by assembling clusters of “hits” in
 1510 channels of the ID tracking layers. The reconstructed tracks are a critical component
 1511 of vertex reconstruction, and the degree of bending and direction of the bent tracks at
 1512 the production vertex provide information about the momentum, charge, and identity
 1513 of the charged particles that produced them.

1514 **3.4.2 The Calorimeters**

1515 The calorimeter is designed to measure the energy of all particles that pass through
 1516 it by initiating cascades of secondary particle production in the high-density detector
 1517 material known as “showers”, and fully absorbing the energy of each shower. The only
 1518 particles that cannot be absorbed by the calorimeter are muons and neutrinos, which
 1519 pass through without showering. The calorimeter is divided into two sub-detectors,
 1520 the electromagnetic and hadronic calorimeters. Both are “sampling calorimeters”,
 1521 which means they are comprised of repeated layers of dense absorbing material with
 1522 “sampling” layers in between. The sampling layers track the location of the shower
 1523 and record a small fraction of its energy, to which a calibration factor is applied to
 1524 infer the full shower energy.

1525 **The Electromagnetic Calorimeter**

1526 The electromagnetic (EM) calorimeter forms the inner calorimeter layer, and is de-
 1527 signed to fully absorb and measure the energies of electrons, positrons and photons.
 1528 Energy is primarily deposited in the lead absorbing layers in the form of EM showers
 1529 [111], in which the initial electron or photon interacts via bremsstrahlung with the
 1530 absorbing material to produce a cascade of photon radiation and electron pair pro-
 1531 duction (see, for example, Ref. [112] for a review of the physics of EM and hadronic

1532 showers). The sampling layers are filled with liquid argon (LAr), which absorbs rela-
 1533 tively little energy compared with the lead absorbing layers due to its lower density.
 1534 Ionization is produced when a charged particle passes through the LAr [113] layer,
 1535 which drifts through an electric field generated by a high voltage placed between
 1536 absorber plates and readout electrodes on either end of the LAr layer to produce a
 1537 triangular current pulse [114].

1538 Candidate EM showers are reconstructed into clusters from energy deposits in
 1539 calorimeter cells using a “seed-cluster” algorithm described in Ref. [115]. The seed-
 1540 cluster algorithm works by dividing the $\eta \times \phi$ space of the EM calorimeter into a grid
 1541 of $\eta \times \phi = 0.025 \times 0.025$ solid angle elements called “towers”. For each such element,
 1542 the energy detected in all calorimeter layers that lie within the given patch of solid
 1543 angle is summed to form the energy of the tower. A sliding window algorithm of size
 1544 3×5 towers is used to construct the energy clusters that constitute EM clusters from
 1545 localized energy deposits.

1546 Reconstruction and selection of electron candidates is performed with the use of
 1547 a complex matching algorithm (details of which can be found in Section 5 of Ref.
 1548 [115]) to match candidate reconstructed tracks in the ID with EM clusters in the
 1549 calorimeter based on their proximity in $\eta \times \phi$ space. This matching is complicated
 1550 by the fact that electrons may radiate photons via bremsstrahlung in the ID prior
 1551 to reaching the calorimeters. The radiated photons can subsequently decay into an
 1552 electron-positron pair, which themselves can generate tracks in the ID. As a result,
 1553 it is possible to reconstruct multiple tracks in the ID, all originating from the same
 1554 electron, and match these tracks to EM clusters in the calorimeter. In case several
 1555 tracks in the ID can be matched to the same EM cluster, the track considered to
 1556 be associated with the primary electron is selected by an algorithm that accounts
 1557 for, among other quantities, the number of hits in the ID and the distance in (η, ϕ)
 1558 between the extrapolated track position in the calorimeter and the barycentre of the
 1559 EM cluster.

1560 Reconstructed objects selected as electron candidates are subsequently passed
 1561 through a likelihood-based electron identification algorithm, described in detail in
 1562 Section 6 of Ref. [115], which uses as a discriminant the ratio

$$d_L = \frac{L_S}{L_S + L_B} \quad (3.3)$$

1563 where the likelihood function $L_S(B)$ is a product of signal (background) PDFs for

1564 various quantities related to the reconstructed object such as track conditions, details
 1565 of the track-cluster matching, and reconstructed EM shower widths for the EM cluster
 1566 in various layers of the EM calorimeter. The signal S is “prompt” electrons, which
 1567 originate from the primary interaction, and the background B is a combination of
 1568 jets that mimic prompt electron signatures, electrons from photon conversions in the
 1569 detector and non-prompt electrons originating from hadron decays.

1570 The Hadronic Calorimeter

1571 As discussed in Section 1.1.1, the vast majority of collision events that occur in the
 1572 ATLAS detector ultimately result in the production of quarks and gluons. Due to
 1573 the phenomenon of colour confinement, quarks and gluons cannot exist in isolation,
 1574 and immediately “hadronize” to form colour-neutral combinations of quarks called
 1575 “hadrons”. As these hadrons pass through the detector, they eventually undergo a
 1576 showering process similar in principle to the EM showers described in Section 3.4.2. In
 1577 the case of these “hadronic showers”, the shower is initiated by the strong interaction
 1578 of a hadron with the detector material to produce a cascade of secondary hadrons
 1579 (see, for example, Ref. [112] for a review of the physics of EM and hadronic showers).
 1580 Unlike the EM showers, which proceed exclusively via electromagnetic interactions,
 1581 hadronic showers proceed via both the strong and EM interactions, where the EM
 1582 interactions are primarily induced by electromagnetic decays of neutral pions (π^0)
 1583 [112]. Because hadronic showers involve strong interactions, they are in general much
 1584 more complex in terms of the variety of particles and mechanisms involved in the
 1585 showering, and as a result are in general more variable and less localized compared
 1586 with EM showers.

1587 The hadronic calorimeter surrounds the EM calorimeter, and is designed to fully
 1588 absorb and measure hadronic showers. The primary hadrons that initiate these
 1589 hadronic showers generally pass through the EM calorimeter without showering due
 1590 to their relatively long interaction length [53]. The hadronic calorimeter is comprised
 1591 of a tile calorimeter, which encircles the EM calorimeter barrel, and a LAr calorimeter
 1592 with copper and tungsten absorbers in the end-cap region that encloses the two ends
 1593 of the barrel. The tile calorimeter uses steel as the absorber material and scintillators
 1594 read out by photomultiplier tubes (PMTs) in the sampling layers.

1595 Hadronic showers are reconstructed as “jets” using clusters of energy deposited
 1596 in the hadronic calorimeter cells (for a review of jet reconstruction and calibration

1597 with the ATLAS detector, see [116]). The jets can be reconstructed using a variety
 1598 of different reconstruction algorithms depending on the use case.

1599 Most jet reconstruction algorithms use clusters of topologically connected calorime-
 1600 ter cells known as “topo-clusters” as basic building blocks for jet reconstruction.
 1601 Topo-clusters are designed with the goal of extracting significant signals of energy de-
 1602 position originating from energetic hadrons from the background of detector noise and
 1603 other sources of fluctuation in the calorimeter cells. Candidate clusters are formed
 1604 from “seed cells” in which the deposited energy E is $E > S\sigma_{\text{cell}}$, where σ_{cell} is the
 1605 average noise for the given cell and S is the “seed threshold” significance, set to 4
 1606 by default [117]. Cluster construction proceeds by collecting neighbouring cells with
 1607 energy $E > N\sigma_{\text{cell}}$, where N is the “growth threshold”, set to 2 by default. If a neigh-
 1608 bouring cell passes the $E > N\sigma_{\text{cell}}$ requirement, its neighbours will also be added to
 1609 the cluster if their energy significance exceeds the growth threshold, and this process
 1610 repeats until there are no remaining neighbouring cells with significance above the
 1611 growth threshold. Lastly, one set of neighbouring cells that satisfy $E > P\sigma_{\text{cell}}$ are
 1612 added to the cluster, where p is the “principal cell filter”, set to 0 by default.

1613 Jets are reconstructed from these topo-clusters using the anti- k_t clustering algo-
 1614 rithm described in Ref. [118] in a cone with an angular radius R in (η, ϕ) space, given
 1615 by:

$$R = \sqrt{\eta^2 + \phi^2} \quad (3.4)$$

1616 The jet radius R determines the angular radius within which the anti- k_t algorithm
 1617 will include calorimeter deposits in the vicinity of a topo-cluster or a set of topo-
 1618 clusters and attempt to group the energy deposits into jets. Different choices of
 1619 R can be used in the algorithm depending on the identity and kinematics of the
 1620 shower parent particle(s) that one is interested in reconstructing [116]. Jets produced
 1621 by quarks and gluons that either originate from different parent particles, or whose
 1622 shared parent particle has a relatively low momentum in the lab frame (a.k.a. “low
 1623 boost”), generally have sufficient angular separation that they can be individually
 1624 reconstructed using relatively small radius parameters such as $R = 0.2$ or $R = 0.4$.

1625 Jets that originate from boosted massive particles, such as the hadronically de-
 1626 caying W boson in the search presented in this paper, generally contain two or more
 1627 significant topo-clusters (a.k.a. “prongs”) in close angular proximity, each having
 1628 been induced by the hadronization of a strongly interacting daughter particle pro-

duced by the hadronic decay of the massive parent particle. In this latter case, the angular proximity of these significant topo-clusters can make it challenging to usefully reconstruct them as individual small-radius jets due to the resulting jet overlap. In these cases, it may be more useful to capture all the decay products in a single jet reconstructed with a larger radius parameter such as $R = 0.8$ or $R = 1.0$, such that the resulting large-radius jet fully reconstructs the massive parent particle. Methods such as the TAR algorithm [2] employed in this analysis are subsequently applied to the reconstructed large-radius jet to obtain useful jet sub-structure information, including the likely number of prongs contained within the jet.

3.4.3 The Muon Spectrometer

The muon spectrometer [119] surrounding the calorimeter is specialized for tracking muons and measuring their momentum. It employs the same principle used in the inner detector of applying a strong magnetic field and measuring the resulting bent trajectories of the electrically charged muons passing through to infer their momenta. Details of muon reconstruction and identification using the ATLAS muon spectrometer can be found in Ref. [120].

The magnetic field is generated by rectangular superconducting “toroid magnets” arranged azimuthally in radial planes around the beam axis, which set up a toroidal field concentric to the beam axis. In the region containing the strong field established by the toroid magnets, muon tracks are recorded by three cylindrical layers of muon tracking chambers in the barrel region and three layers of chambers arranged in wheels perpendicular to the beam axis in the end-cap region. Additional layers of fast trigger chambers deliver muon track information to the ATLAS trigger system (see Section 3.4.5) so it can be incorporated into the event readout decision.

Precise measurements of muon track coordinates are provided by monitored drift tubes (MDTs) in the barrel region, which cover a pseudorapidity range of $|\eta| < 2.0$), and by cathode strip chambers (CSCs) in the end-cap region ($2.0 < |\eta| < 2.7$). The CSCs are designed to withstand the relatively high flux of energetic particles that bombard the detector in the end-cap region. These MDTs and CSCs measure track coordinates with nominal resolutions of $60 \mu\text{m}$ and $80 \mu\text{m}$, respectively, in the magnetic bending plane [120].

High precision tracking is needed both to achieve the design performance goal of reconstructing muon transverse momenta with at most 10% resolution for 1 TeV

tracks [53], and to distinguish prompt muons that originate from the primary interaction from the background of non-prompt muons that arise from secondary interactions in the detector. Reconstruction and identification of prompt muons is performed by combining tracking information from the muon spectrometer and the inner detector, along with energy deposition measurements from the calorimeters. Various muon reconstruction strategies are employed, which attempt to match tracks between the muon spectrometer and the inner detector, or to match inner detector tracks with calorimeter deposits [120].

3.4.4 Missing Transverse Momentum

Many signatures of hypothesized processes involving physics beyond the SM - including the DH model considered in this thesis - involve the production of particles that would pass through the ATLAS detector without being detected due to their extremely low interaction cross section with SM particles. Similarly, neutrinos produced in weak interactions will also pass undetected for the same reason.

The law of momentum conservation, which requires the vector sum of momenta of all measured particles produced by a collision to match that of the initial state quarks, can be used to infer the presence of undetected particles in the collision. Because the fraction of proton momentum carried by each of the initial state quarks - described by the PDFs presented in Section 3.1 - is statistical in nature, the momenta in the direction of the pp beam line cannot be known precisely. However, initial state quark momenta in the plane transverse to the beam line are in general negligibly small compared with the collision energy, so it can be expected to a high degree of precision that the final state particle momenta in this transverse plane will sum to zero. This expectation implies that collision events that produce undetected particles with an appreciable momentum in the transverse plane can be expected to exhibit a signature of high missing transverse momentum in the final state.

This two dimensional missing transverse momentum vector for a collision event is typically denoted “ \vec{E}_T^{miss} ”, and its magnitude “ E_T^{miss} ”. A detailed description of E_T^{miss} reconstruction with the ATLAS detector can be found in Section 3 of Ref. [121]. The components of \vec{E}_T^{miss} lie in the transverse (x, y) plane of the detector, and are given by:

$$E_{x(y)}^{\text{miss}} = - \sum_i p_{x(y),i} \quad (3.5)$$

where the sum is over all fully reconstructed electrons, muons, photons, hadronically-decaying tau leptons and jets (a.k.a. “hard objects”), and additionally over all other detector signals that were recorded as part of the event, but were not used as part of the construction of the hard objects (a.k.a the “soft term”).

\vec{E}_T^{miss} and E_T^{miss} are constructed from the components $E_{x(y)}^{\text{miss}}$ as follows:

$$\vec{E}_T^{\text{miss}} = (E_x^{\text{miss}}, E_y^{\text{miss}}) \quad (3.6)$$

$$E_T^{\text{miss}} = |\vec{E}_T^{\text{miss}}| = \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2} \quad (3.7)$$

3.4.5 The Trigger System

The majority of collision events in the ATLAS detector result in “soft quantum chromodynamics” (soft QCD) interactions, which proceed via the strong force and produce only quarks and gluons with a relatively low momentum in the lab frame - and hence are referred to as “soft” - and which do not result in the production of any high-mass ($\mathcal{O}(\text{GeV})$) particles. Figure 3.5 compares the overall cross section of all interactions resulting from pp collisions with the cross sections of processes that produce various massive SM particles. The overall collision cross section is ~ 6 orders of magnitude larger than the highest cross section for producing a massive particle, namely that of the W boson. The soft QCD interactions that dominate the total cross sections are in general not of interest for the ATLAS physics programme. It would be impossible to process and save all collision events produced in the ATLAS detector, as the massive flux of these soft QCD interaction events would very quickly overwhelm the bandwidth of the data acquisition system and fill up the available offline data storage capacity.

The ATLAS trigger [122] is designed to efficiently select a tiny minority of collision events to be read out, processed and saved to offline storage (a.k.a. “recorded”) based on a set of criteria applied to preliminary event information from the sub-detector. The criteria are designed to select for signatures from the sub-detectors that are considered likely to represent processes that will be of interest to the various analyses that study the data either to measure parameters of the SM, or to search for evidence of BSM physics.

The trigger system is divided into a hardware-based “first level trigger” (L1 trigger) and software-based “high-level trigger” (HLT), which collectively select events

at an average rate of \sim 1000 Hz from the total collision rate of 40 MHz [122]. Events must be accepted by both the L1 trigger and the HLT in order to be recorded.

The L1 trigger is based on candidate objects identified within regions of interest (“RoIs”) defined by their η and ϕ ranges. Candidate objects are divided into muons, EM calorimeter clusters, jets in the hadronic calorimeter and taus. In addition, the sums of missing transverse momentum (E_T^{miss}) and total energy are also constructed. Hardware level trigger decisions are designed using combinations of these reconstructed objects and sums. Events that pass the L1 trigger are subsequently processed by the HLT. The L1 trigger is designed to form the “trigger decision” of whether to reject the event or accept it for processing by the HLT within at most 2.5 μs per event [122].

Thanks to the reduction in event rate by the L1 trigger selections, the HLT is able to use software-based algorithms to produce more complex reconstructions of the candidate objects and sums compared with the L1 trigger, and can apply more sophisticated selections on these objects. The HLT trigger decision is typically formed within 300 ms per event [122]. Objects reconstructed and considered in the HLT trigger decision include muons measured by the muon spectrometer, electrons, photons, jets, E_T^{miss} and tau leptons.

The L1 trigger and the HLT each have their own “trigger menu”, which represents the compilation of all sets of selection criteria, or “triggers”, that are considered for each event. Any event that satisfies the criteria of any of the triggers in the trigger menu is kept. Both the L1 trigger and the HLT trigger menus include both triggers applied to individual objects, such as the “ E_T^{miss} trigger” or the “photon trigger”, as well as triggers applied to combinations of different objects.

Importantly for the analysis presented in this thesis, the E_T^{miss} that is reconstructed for each event by the L1 trigger and the HLT is based on calorimeter measurements, which register very little energy deposition from muons, and does not incorporate the momenta of muons detected by the muon spectrometer. As a result, in cases where the majority of E_T^{miss} in the event is recoiling against one or more energetic muons, the E_T^{miss} reconstructed for the triggers will underestimate the actual E_T^{miss} in the event due to the absence of muon momentum information. Because this analysis includes many events with an energetic muon in the final state, it was found that requiring events to have passed the E_T^{miss} trigger as part of the analysis selections removes some events with energetic muons in the final state that would have passed the other event selections. These analysis selections include a stringent lower bound

1757 on the offline-reconstructed E_T^{miss} , which accounts for muon momenta. The solution,
1758 presented in more detail in Section 5.3, was to additionally allow for events that pass
1759 the “single muon” trigger.

1760

Chapter 4

1761

Modelling of Signal and Background Processes

1762 To search for evidence of new physics in the ATLAS data, it is necessary to develop
 1763 an accurate model of the expected yield of both SM events (a.k.a. “background”)
 1764 and hypothesized BSM events (a.k.a. “signal”) in the data, as well as their kinematic
 1765 distributions. The yields and kinematic distributions of events in the data are then
 1766 compared with those in the signal and background models to check for any “above-
 1767 background” significant excesses in the data that could point to the presence of BSM
 1768 physics. If no significant excesses are observed, the search can exclude the signal
 1769 model over the range of parameters for which the model predicts a significant above-
 1770 background excess in the data.

1771 **4.1 Introduction to the Monte Carlo Method**

1772 Like many collider experiments, the ATLAS collaboration uses the “Monte Carlo”
 1773 (MC) method to model the expected yield and kinematic distributions of SM and
 1774 hypothesized BSM events in the data collected by the detector. The MC method is
 1775 a computational algorithm that uses repeated sampling of random variables, where
 1776 each set of randomly sampled variables represents a randomly generated “event”.

1777 To simulate the predicted behaviour of any given process using the MC method, a
 1778 parametric model for the process is needed. The parametric model receives as input
 1779 a set of variables associated with a single event (eg. kinematic information describing
 1780 the partons involved in a high-energy parton-parton collision at the LHC, prior to
 1781 the collision), and predicts the values of all variables of interest for the event after
 1782 it has undergone the process being modelled (eg. the energy and momentum of the
 1783 massive particle that would be produced from the parton-parton collision for a given
 1784 particle production process). The MC method proceeds as follows: for each randomly

generated event, the random variables associated with the event are passed into the model to produce a resulting set of output variables. For physical models, one is particularly interested in the set of “observable” output variables, meaning those that can be measured experimentally in the modelled system. Given a set of these so-called “MC simulated events”, the associated set of values for each observable, which were generated by passing the events through the model, represents a random sampling of the underlying probability density distribution for that observable according to the model. The method is useful for complex models with many free parameters, for which it would be unfeasible to develop analytical formulations of the distributions of observables predicted by the model.

The MC simulated events can then be binned into histograms in one or more observables. Assuming that, for each event, the sampling of random variables is performed “independently” - i.e. in a manner such that the sampling of random variables for each event is unaffected by that of any other event - the number N_i of events in each bin i will vary randomly according to Poisson statistics with, on average, a standard deviation of $\sigma_{N_i} = \sqrt{N_i}$. Consequently, as the number of MC simulated events is increased by a factor of α , the relative size $\frac{\sigma_{N_i}}{N_i}$ of fluctuations in each bin will, on average, decrease according to $\frac{1}{\sqrt{\alpha}}$. As a result, as one increases the number of MC simulated events, the shapes of histograms binned in the model’s observables for the simulated events will become an increasingly precise approximation of the underlying probability distributions for these observables according to the model.

4.2 Monte Carlo Simulation of Events in the ATLAS Detector

Signal and SM background models used to perform searches for BSM physics with the ATLAS detector are produced using sophisticated MC simulation of both the passage of the final-state particles through the ATLAS detector and of the physical production mechanisms for the particle collision, production and decay processes involved. For a given process, “truth-level” information for each MC simulated event generated to simulate the process is first obtained from a random proton-proton collision by simulating the physical production mechanism for the process. An example of such a process would be the dominant $W + \text{jets}$ SM background in this DM search, shown in Figure 4.3a.

The set of simulated final state particles, along with their kinematic information, are collectively known as the “truth-level” event. Truth-level events can subsequently

1818 be passed through a highly detailed simulation of the ATLAS detector [123] produced
 1819 using the Geant4 toolkit [124], which models how these events would actually be mea-
 1820 sured by the detector at this so-called “reconstruction-level”. ATLAS requires very
 1821 large MC generated data sets (millions of simulated events per process) to adequately
 1822 model the predicted probability distributions for kinematic observables over their full
 1823 range of interest for the SM measurements and BSM searches that use the ATLAS
 1824 data.

1825 **4.2.1 Use of Alternative MC Generators**

1826 ATLAS uses various MC simulation packages (also known as “generators”) to per-
 1827 form truth-level MC simulation of different physics processes. For many processes,
 1828 particularly SM background processes, independent MC simulations have been per-
 1829 formed using several different packages, and the yields and distributions of events
 1830 predicted by the different packages can be compared to evaluate a systematic uncer-
 1831 tainty associated with the choice of generator used to simulate the process. The
 1832 specific generators used to model the physics processes considered in this search will
 1833 be discussed in Sections 4.3 and 4.4.

1834 **4.2.2 Weighting and Normalization of MC Simulated Processes**

1835 Given a set of MC simulated events for a given process, it is often necessary to apply
 1836 multiplicative “weights” to the simulated events in order to correct their distributions
 1837 and amplitude before comparing with the measured data. The weights are designed
 1838 to modify the amount that a given event contributes to the amplitude of the bin
 1839 to which it is assigned when the MC simulated events are binned into histograms.
 1840 Rather than simply summing the number of simulated events that fall into a given
 1841 bin, the simulated amplitude, or “yield”, of each bin is evaluated instead as the sum
 1842 of event weights w for all events that fall into the bin:

$$\text{simulated yield in bin } k = \sum_{\text{event } i \text{ in bin } k} w_i \quad (4.1)$$

1843 The weights are broadly categorized into “event-level” weights and “scaling fac-
 1844 tors”. Event-level weights may differ between one event and the next, and are designed
 1845 to modify the shapes of simulated yield distributions in one or more observables to
 1846 better represent their expected distributions in the measured data. These shape mod-

ifications may be motivated by a variety of factors, such as to account for data-taking conditions that were not known or incorporated at the time of simulation. Individual sources of event-level weights for processes simulated in the ATLAS detector are discussed in more detail in Section 4.2.2 below.

After applying event-level weights to correct the shapes of the distributions, scaling factors are applied identically to all the events generated for a given process. The scaling factors are designed to scale the total simulated yield of events such that it matches the total number of events that are expected to have been produced by the simulated process in the actual measured data set.

Weighting of MC Simulated Events of Particle Collision Processes

Event-level weights are applied to events generated by the MC method to model particle collision and decay processes in the ATLAS detector. For a given process, the overall event-level weight applied to each event is a product individual weights arising from various sources:

$$\text{event-level weight } i = (\text{generator weight})_i \times (\text{pileup reweighting weight})_i \times \times \prod_j (\text{reconstruction weight } j) \quad (4.2)$$

The “generator weight” is a weight applied by some generators during the generation of truth-level events for various purposes. These purposes may include correcting for the generation of duplicate events at various stages of the calculation and correcting leading-order calculations to achieve the expected distributions that a more precise “next-to-leading-order” calculation would be expected to produce.

The “pileup reweighting weight” is designed to account for the effects of “pileup” (see Ref. [125] for a description of the methods currently used to model pile-up in the ATLAS detector). Due to the oscillatory electric fields used to accelerate protons in the counter-rotating LHC beams, protons in the beams do not form a continuous stream of particles, but are instead concentrated into regularly-spaced “bunches” in the longitudinal beam direction [126]. Superconducting magnets are used to direct proton bunches in the counter-rotating beams into head-on collisions, known as “bunch crossings”, at the centre of the ATLAS detector. Pileup constitutes the soft QCD collision events that take place in the same (or closely-surrounding)

1875 bunch crossings as the “hard interaction” that actually triggered the event readout.
 1876 The nominal procedure of simulating the “hard interactions” that would produce the
 1877 process being modelled does not account for the presence of these pileup interactions
 1878 that would be measured by the detector as part of the readout for the triggered event.

1879 To correctly model the actual pileup conditions during data-taking, the soft QCD
 1880 collision events that constitute these pileup interactions are either simulated or col-
 1881 lected from actual LHC collisions as so-called “zero-bias data”¹ (see Refs. [128] and
 1882 [127], respectively, for recent discussions of these alternative pileup modelling meth-
 1883 ods). Each simulated hard interaction event is then overlaid with a variable number of
 1884 the simulated pileup events, and the hard interaction and pileup events are weighted
 1885 to produce the distribution of pileup events in the data. Since the MC simulated
 1886 datasets were in many cases produced before or during data-taking, it is necessary to
 1887 reweight the MC simulated events using the so-called pileup reweighting weight such
 1888 that the distribution of pileup events accurately reflects the actual pileup distribution
 1889 during the data-taking. To do so, the full ATLAS data-taking period is divided into
 1890 “luminosity blocks”, and the average rate of pileup interactions is measured within
 1891 each such luminosity block. MC simulated events are then associated with specific
 1892 luminosity blocks. The pileup reweighting weights are evaluated for MC simulated
 1893 events within each luminosity block to match the pileup rate in the simulated events
 1894 to the average pileup rate measured during the associated data-taking period.

1895 The “reconstruction weights” collectively refer to weights assigned to apply cor-
 1896 rections to quantities such as data-driven measurements of efficiency or resolution
 1897 associated with the reconstruction of objects such as electrons, muons and jets that
 1898 are produced in the simulated passage of events through the ATLAS detector.

1899 Scaling Distributions for Comparison with Data

1900 In addition to the event-level weights described in Section 4.2.2 above, scaling factors
 1901 are applied to each signal and background process such that the predicted yield of
 1902 events per bin in the MC simulated distributions for each process properly predicts the
 1903 actual yield of events expected in the ATLAS data based on the integrated luminosity
 1904 of the collected data set.

¹Zero-bias data is collected using a dedicated trigger, which fires one LHC turn after a high- p_T L1 trigger fires [127]. This method ensures that the rate at which the zero-bias events are triggered is proportional to the instantaneous luminosity of the collisions. Since no other triggers are applied, the event readout for this zero-bias data is expected to be representative of the pileup background conditions.

1905 **Sum of Weights Normalization**

1906 While the event-level weights can adjust the shapes of distributions of MC simulated events to match those expected in the data, the sum of these weights does not
 1907 in general have any particular physical significance, and depends on the number of
 1908 events that were simulated. Prior to scaling the sum of weights to the predicted yield
 1909 in the data, it is therefore necessary to first normalize the sum of weights to unity by
 1910 dividing by the sum over all event weights for the process:

$$\text{event-level weight (normalized)} i = \frac{\text{event-level weight } i}{\sum_j \text{event-level weight } j} \quad (4.3)$$

1912 where the index j runs over all MC simulated events for the given process.

1913 **Scaling to Expected Data Yield**

1914 As discussed in Section 3.2, the total predicted yield N of events for a given
 1915 process of particle production and decay initiated by a proton-proton collision in the
 1916 LHC is given by the total integrated luminosity:

$$N = \sigma \int_{t_1}^{t_2} \mathcal{L}(t) dt = \sigma \mathcal{L}_{\text{int}} \quad (4.4)$$

1917 where $\int_{t_1}^{t_2} \mathcal{L}(t) dt$ is the integrated beam luminosity over the full data-taking period
 1918 from t_1 to t_2 and σ is the cross section for the process, which quantifies the rate
 1919 at which the proton-proton collisions will produce events via the process for a given
 1920 beam luminosity \mathcal{L} .

1921 Therefore, the final step in weighting the MC simulated events for comparison
 1922 with data is to scale all the normalized weights in Eq. 4.3 by the product of the cross
 1923 section σ for the given process and the integrated ATLAS luminosity \mathcal{L}_{int} such that
 1924 they sum to the total predicted yield N of events for the process:

$$\text{event-level weight (normalized, scaled)} i = [\text{event-level weight (normalized)} i] \times \sigma \times \mathcal{L}_{\text{int}} \quad (4.5)$$

1925 The following calculation confirms that summing all event weights in Eq. 4.5, and
 1926 combining with 4.3 and 4.4 produces the total predicted yield N for the given process:

$$\sum_i \text{event-level weight (normalized, scaled)} i = \frac{\sum_i \text{event-level weight } i}{\sum_j \text{event-level weight } j} \times \sigma \times \mathcal{L}_{\text{int}} \\ = \sigma \mathcal{L}_{\text{int}} = N \quad (4.6)$$

1927 4.2.3 Comparing Data and MC Simulation to Search for New Physics

1928 With the MC simulated data properly weighted and normalized as described in Sec-
 1929 tion 4.2.2 above, event selections are applied to both data and MC simulated events
 1930 based on their final-state observables, such as the identity, momenta and directions
 1931 of final-state particles measured by the detector. The selections, which are discussed
 1932 in more detail in Section 5.2, are designed to define one or more regions of the data,
 1933 known as “signal regions” within which the MC simulation of the signal process pre-
 1934 dicted a relatively large yield of the hypothesized BSM process compared with the MC
 1935 prediction of SM background processes. Within each signal region, the data and MC
 1936 simulated events may be additionally binned in one or more final-state observables,
 1937 and the resulting distributions of ATLAS data are compared with those of the total
 1938 MC simulated SM background yields to check for any significant yield excesses or
 1939 shape differences in ATLAS data that could be indicative of new physics.

1940 4.3 Simulation of the DH Signal Model

1941 The DH signal model presented in Chapter 2 is simulated using a program called
 1942 MadGraph 5 [129]², which generates proton-proton collision events and calculates the
 1943 matrix element at leading order to produce events associated with the Lagrangian for
 1944 a given process. The Lagrangian for the DH signal model is encoded in MadGraph,
 1945 with the coupling constants g_q , g_χ , the mixing angle θ , and the DM mass m_χ fixed to
 1946 the values specified in Section 2.2.1. The DH and Z' masses m_s and $m_{Z'}$ are left as
 1947 floating parameters in the search. Therefore, MC simulated data sets are generated
 1948 over a grid of m_s and $m_{Z'}$. The grid was designed to cover masses for which the
 1949 search is expected to be reasonably sensitive to the model. Figure 4.1 shows the m_s
 1950 and $m_{Z'}$ masses for which MC simulated data sets were generated for the DH signal
 1951 process.

²MADGRAPH5_AMC@NLO 2.7.2 (ATLAS, LCG) [130] is the particular version of MadGraph 5 used to generate the MC simulated signal events used in this search.

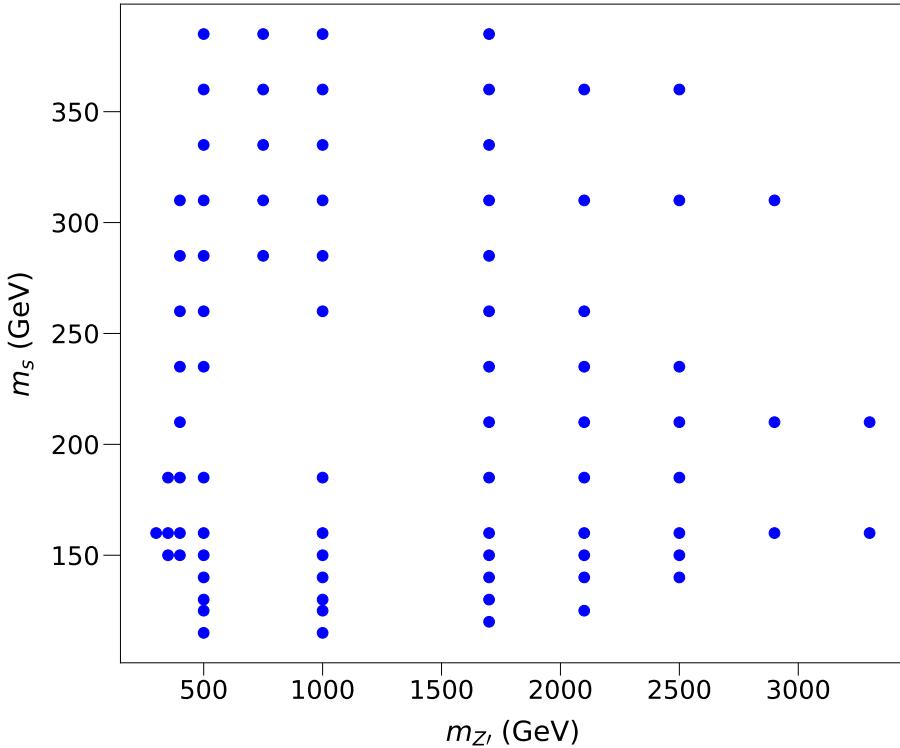


Figure 4.1: Grid of produced signal samples with different choices of $m_{Z'}$, m_s and the other parameters fixed to $g_q = 0.25$, $g_\chi = 1.0$, $m_\chi = 200$ GeV, $\theta = 0.01$.

As discussed in Chapter 2, the signal model considered in this search produces two final-state partons from the $s \rightarrow WW(qq\ell\nu)$ decay. MadGraph also includes production mechanisms in the matrix element calculation for which up to one additional parton is radiated in the final state. These final-state partons initiate cascades of radiation produced by QCD processes, which are modelled using the Pythia8³ [131] program.

4.4 Simulation of SM Background Processes

The selection criteria applied to final-state observables in the ATLAS and MC simulated events are designed to define signal regions, which contain events that exhibit the final-state signature of the DH signal model, namely of a WW pair that decays semileptonically and recoils against missing transverse momentum produced by the DM pair in the final state. However, some SM processes can produce final state

³The PYTHIA 8.230 [131] is the particular version of Pythia8 used to generate the MC simulated signal events used in this search.

observables that are similar enough to that of the signal model as to create an appreciable yield of events in the signal regions. In addition to targeting the signal model, the selections are also optimized to minimize the predicted yield of SM background processes in the signal regions. This section presents the background processes that have a non-negligible yield in the signal regions even with the optimized signal region selections.

Dominant backgrounds to the search are the $W + \text{jets}$, Diboson and $t\bar{t}$ processes described in Section 4.4.1. Figure 4.2 shows the yield breakdowns in the signal regions of all SM background processes considered in the analysis.

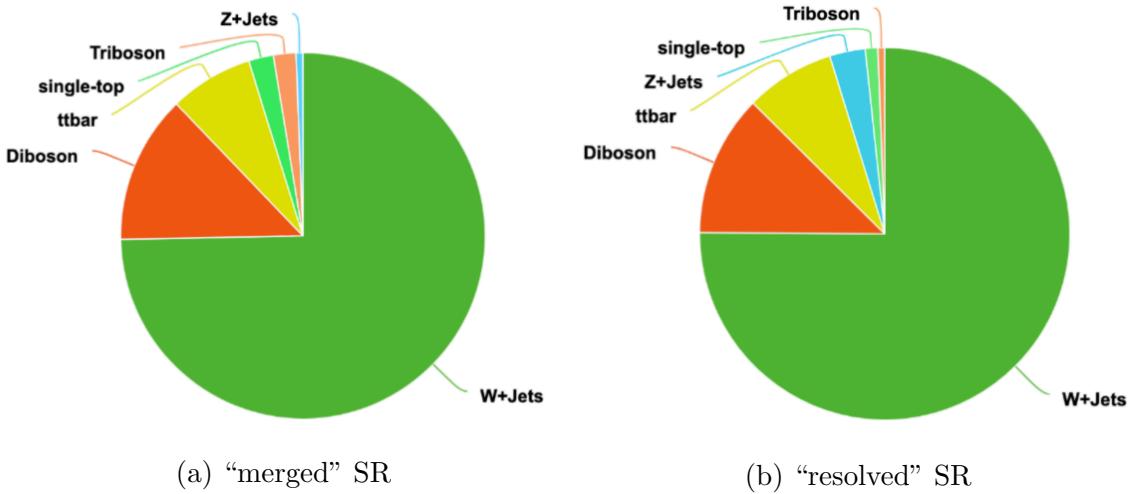


Figure 4.2: Relative contributions of all SM background processes considered in the signal regions.

4.4.1 Dominant Background Processes

W+jets

The dominant SM background in the signal regions comes from the $W + \text{jets}$ process, wherein a leptonically decaying W is produced from the initial parton-parton collision, along with hadronic activity that fakes the hadronically decaying W in the signal model. A leading Feynman diagram for the $W + \text{jets}$ background is shown in Figure 4.3a.

The SHERPA 2.2 [132] MC generator is used to model both the hard $W + \text{jets}$ process, as well as the parton shower initiated by the final-state partons in the process.

Statistical Enhancement in $W + \text{jets}$ Samples for $m_W > 120$ GeV

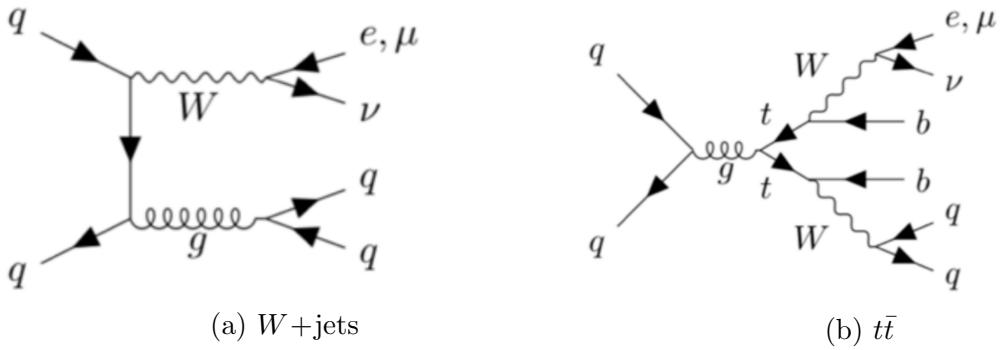


Figure 4.3: Feynman diagrams for $W + \text{jets}$ and $t\bar{t}$ SM background processes.

Due to application of a high $m_T(\ell, E_T^{\text{miss}})$ requirement, described in Section 5.2, to reduce the $W + \text{jets}$ background in the signal region, it was found that the majority of MC simulated events simulated for the $W + \text{jets}$ process that make it into the signal regions are generated with a very high off-shell mass of the leptonically decaying W boson in the process. The default SHERPA 2.2 generator is not optimized to produce large MC statistics in this high- m_W regime. Therefore, in addition to using samples produced by the default SHERPA 2.2 generator, this search also makes use of a recently-developed set of specialized SHERPA 2.2 $W + \text{jets}$ MC generated samples⁴ that are generated with enhanced MC statistics for large off-shell masses ($m_W > 120$ GeV) of the leptonically decaying W .

1993 Diboson

The next-leading SM background after $W + \text{jets}$ comes from the diboson process, in which a pair of vector bosons - WW , ZZ or WZ - are produced from the initial parton-parton collision. The diboson events that make it into the signal region are dominated by the production mechanism in which both bosons decay leptonically ($W \rightarrow \ell\nu$, $Z \rightarrow \nu\nu$ or $Z \rightarrow \ell\ell$) to produce a final-state lepton in addition to missing transverse momentum from the neutrino production, and one or more partons are radiated as part of the diboson production process to produce QCD activity that fakes the hadronically decaying W in the signal model.

The diboson process, as well as the parton showers initiated by partons produced in the process, is modelled using the SHERPA 2.2 MC generator.

⁴The specialized samples generated with enhanced MC statistics for large off-shell m_W are not yet discussed in any published work, but are described in the following internal ATLAS document: [ATL-COM-PHYS-2021-063](#)

2004 **t \bar{t}**

2005 The $t\bar{t}$ process represents the third-leading SM background in the signal regions. A
 2006 leading Feynman diagram for the process is shown in Figure 4.3b. In this process, two
 2007 t quarks are produced from the initial parton-parton collision, both of which decay
 2008 to a b quark and a W boson. The WW pair decays semileptonically, thus faking the
 2009 semileptonically decaying WW pair produced in the signal model. The final-state ν
 2010 from the leptonic W decay produces the missing transverse momentum required in
 2011 the signal region selection. The signal region selection includes a veto on the presence
 2012 of b -tagged quarks in the final state to reduce the yield of $t\bar{t}$ events, but some events
 2013 from the process pass the veto and make it into the signal region due to the limited
 2014 efficiency of the b quark tagging algorithm [133].

2015 The production of $t\bar{t}$ events is modelled using the POWHEG Box v2 [134, 135,
 2016 136, 137] generator, which calculates matrix elements for the process. Parton showers
 2017 initiated by final-state partons produced in the $t\bar{t}$ process are modelled using PYTHIA
 2018 8.230 [131].

2019 **4.4.2 Sub-dominant Background Processes**

2020 **Z+jets**

2021 The Z +jets process is analogous to the W +jets process, but with the leptonically
 2022 decaying W boson replaced by a Z boson, which also decays leptonically. For the
 2023 majority of Z +jets events that are classified into the signal region, the Z boson decays
 2024 to a $\ell\ell$ pair, of which one of the ℓ s is not properly identified as a e or μ during event
 2025 reconstruction.

2026 **Triboson**

2027 The triboson process is similar in structure to the diboson, except that three vector
 2028 bosons rather than two are produced from the initial parton-parton collision. Triboson
 2029 events that pass the signal region selection predominantly exhibit a “ $VVjj$ ” final
 2030 state, in which two of the vector bosons decay leptonically to produce a final-state e
 2031 or μ in addition to missing transverse momentum from ν production, and the third
 2032 vector boson decays hadronically. The triboson process, as well as parton showers
 2033 initiated by final-state partons in the process, is modelled using the SHERPA 2.2 MC
 2034 generator.

2035 **single-top**

2036 The single-top process in the signal region is dominated by “ Wt ” events, in which
2037 a single t quark is produced in association with a W boson from the initial collision
2038 of a quark and a gluon. The t quark subsequently decays to Wb to produce the
2039 signature WW final state of the signal model. As with the $t\bar{t}$ background, the yield
2040 of single-top events in the signal region is reduced by the application of a b -veto in
2041 the event selection for the signal region.

2042 Single-top Wt associated production is modelled using the POWHEG Box [138,
2043 135, 136, 137] v2 generator, which provides matrix elements for the process. Parton
2044 showers initiated by final-state partons produced in the single-top Wt process are
2045 modelled using PYTHIA 8.230 [131].

2046

Chapter 5

2047

Data Reconstruction, Selection and Triggers

2048 This chapter describes the physics objects that are reconstructed on an event-by-event
 2049 basis using collision data from the ATLAS detector, and used in this DM search. It
 2050 also discusses the triggers and event selection cuts that are applied to define the
 2051 subsets of collision data and MC simulated data, also known as “analysis regions”,
 2052 used for the search.

2053 **5.1 Object Definitions**

2054 The goal of the ATLAS detector is to identify particles that are produced by the
 2055 proton-proton collisions that take place in the centre of the detector, and to recon-
 2056 struct their kinematic properties. The particle identification and reconstruction is
 2057 performed using collections of measured signals in the detector sub-systems, which
 2058 are broadly referred to as “physics objects” (See Ref. [139] for a review of physics
 2059 object reconstruction with the ATLAS detector). The physics objects used to recon-
 2060 struct all particles considered in this search are described in the following sections.

2061 **5.1.1 Charged Leptons**

2062 The final state charged lepton produced from the leptonic decay of a W in the DH
 2063 model could with approximately equal probability [22] be an electron e , a muon μ
 2064 or a tau τ . Electrons are stable and as such do not decay before depositing their
 2065 energy in the ATLAS detector. This allows them to be reconstructed directly using
 2066 information from the inner tracker and the EM calorimeter, as discussed in Sections
 2067 3.4.1 and 3.4.2. Muons are unstable and will ultimately decay to a ν_μ and a $e\bar{\nu}_e$ pair
 2068 via a virtual W boson mediator, as shown in Figure 5.1a. However, their mean lifetime
 2069 of $2.2\mu s$, which is the average time after they are produced before they undergo the

decay to $\nu_\mu + e\bar{\nu}_e$, is long enough that muons do make it through the ATLAS detector before they decay, and are reconstructed using information from the inner tracker and the muon spectrometer, as discussed in Sections 3.4.1 and 3.4.2. Only the $W \rightarrow e\bar{\nu}_e$ decay is possible in the 3-body muon decay, because any other particles that a virtual W could otherwise decay to are too massive to be produced from the initial 106 MeV rest mass energy of the muon.

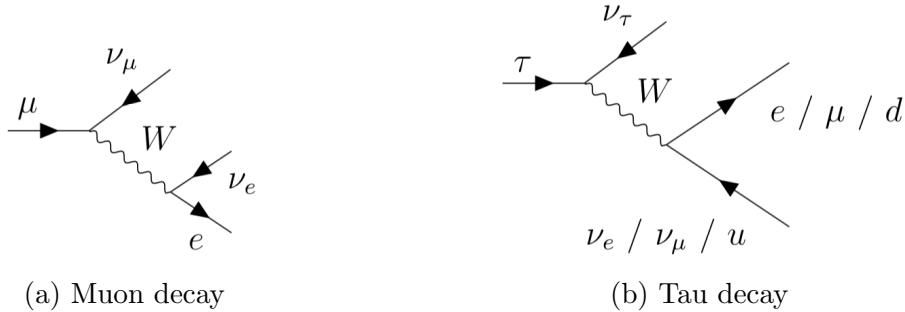


Figure 5.1: Decay mechanisms for muons and taus.

Due to the relatively large tau mass of 1.8 GeV, the virtual W boson in the 3-body tau decay can itself decay either leptонically to $e\bar{\nu}_e$ or $\mu\bar{\nu}_\mu$, or hadronically to $d\bar{u}$, as shown in Figure 5.1b. As a result of the additional decay channels, tau decays proceed with a much shorter mean lifetime of 0.3ps compared with that of the muon. Because of this relatively short lifetime, taus will decay before passing through the ATLAS detector, and as such it is their leptonic or hadronic decay products that are actually measured in the detector [140]. As will be discussed in Section 5.2 below, the selections applied for this analysis require that events have exactly one electron or muon in the final state in order to be considered for the search. As a result, the search is sensitive to $s \rightarrow WW$ decays in which the leptонically decaying W decays to $\tau\nu_\tau$ only in the case where the τ decays leptонically to produce a single energetic electron or muon in the final state. This leptonic τ decay mode occurs with a 35% branching fraction [140].

2089 Electrons

As discussed in Section 3.4.2, electron objects are reconstructed from clusters of energy deposits in the electromagnetic calorimeter that are associated with tracks in the inner detector, and calibrated to the EM scale. Detailed information about electron reconstruction, identification, and calibration can be found in Refs. [141],

[142] and [143]. To accommodate the differing needs of the various studies that make use of electron objects, ATLAS reconstructs these objects at several levels of identification and isolation efficiency, where the various efficiency levels are referred to as “working points”, the names of which are typically variants of *Loose*, *Medium* and *Tight* for reasons that will be discussed in the following paragraphs.

The identification efficiency refers to the probability that an electron passing through the detector will be correctly reconstructed and identified as such. In general, a higher efficiency is achieved by loosening electron identification criteria, which comes at the cost of an increased background acceptance. The increased background acceptance means that reconstructed objects have a higher probability of being incorrectly identified as having originated from an electron.

Electron isolation tackles a slightly different, though related, challenge in comparison with identification. The goal of isolation is to separate the so-called “prompt” electrons that are produced from the primary decay processes of heavy mediators produced in the pp collisions from background processes such as semileptonic quark decays, hadrons misidentified as leptons and photons that convert into e^+e^- pairs before reaching the EM calorimeter. It is generally found that reconstructed objects that originate from prompt electrons can be characterized by a relative absence of (i.e. isolation from) significant activity in a small angular radius R around the object in the space of $\eta \times \phi$. In analogy with the identification efficiency, a high isolation efficiency is achieved by loosening the criteria for defining an object as isolated. As such, loosening isolation criteria will improve the probability that the prompt electrons targeted by the isolation requirement are identified as isolated objects, at the cost of an increase in the rate at which objects that originate from background processes are also identified as isolated.

Two types of electrons are defined for the search based on different sets of criteria. *Baseline* electrons use the *Loose* working point for both identification and isolation. Isolation is measured within a fixed angular radius of $\Delta R = 0.2$ around the reconstructed electron object [142]. The *Loose* identification working point is measured in dedicated studies performed within the ATLAS collaboration to have an efficiency of 93% [142] for identifying prompt electrons with $E_T = 40$ GeV. The *Loose* isolation working point has a total measured efficiency of 98% [142]. Because of their relatively high efficiency, *baseline* electrons are used to veto the presence of additional electrons in the final state.

Signal electrons are designed to reconstruct prompt electrons with high purity.

2129 They are required to satisfy the *Medium* identification criteria, which are measured
 2130 to have an 88% efficiency [142], and *Loose* isolation criteria.

2131 Both types of electrons are required to have $p_T > 7$ GeV and a pseudorapidity in
 2132 the range of $|\eta| < 2.47$.

2133 **Muons**

2134 As described in Section 3.4.3, muons are reconstructed using information from the
 2135 the inner detector and the muon spectrometer. Detailed information about muon
 2136 reconstruction, identification and calibration can be found in Refs. [144] and [145].
 2137 As is the case with electron objects, muon objects are reconstructed at several iden-
 2138 tification and isolation working points, and two definitions for muons are considered
 2139 for this analysis:

2140 *Baseline* muons do not have any isolation requirement, but are required to satisfy
 2141 the *Loose* identification criteria, with a measured efficiency of 98% for $20 \text{ GeV} <$
 2142 $p_{T,\mu} < 100 \text{ GeV}$ [144].

2143 *Signal* muons are designed to have a relatively high purity, and must satisfy the
 2144 *Medium* identification criteria, with a 96% efficiency for $20 \text{ GeV} < p_{T,\mu} < 100 \text{ GeV}$
 2145 [144]. Signal muons are additionally required to pass a set of tight isolation criteria
 2146 referred to as *TightTrackOnly_VarRad* [145]. These tight isolation criteria use infor-
 2147 mation from the inner tracker, and are defined within an angular radius ΔR around
 2148 the reconstructed muon object that depends on the p_T of the muon object.

2149 Both types of muons use a threshold of $p_T > 7$ GeV.

2150 Baseline muons are required to have pseudorapidity in the range of $|\eta| < 2.7$. For
 2151 signal muons, a tighter pseudorapidity range of $|\eta| < 2.5$ is required to ensure that
 2152 the muons are well measured in the inner detector as well as the muon spectrometer.

2153 **5.1.2 Small-radius anti- k_t ($R = 0.4$) jets**

2154 As discussed in detail in Section 3.4.2, quarks and gluons induce showers of energy
 2155 deposits in the calorimeter known as jets. This search uses the “particle flow algo-
 2156 rithm” [146] to reconstruct objects associated with the energy deposits in the hadronic
 2157 calorimeter. The particle flow algorithm matches signals from the inner tracker with
 2158 topologically connected clusters of energy deposits in the calorimeter known as “topo-
 2159 clusters” with the aim of forming objects that represent individual charged particles.
 2160 The energy deposited in the calorimeter by these identified charged particle objects

is removed, leaving behind an ensemble of “particle flow objects” that consist of the remaining calorimeter energy and tracks. The anti- k_t algorithm described in Ref. [118] is then used to reconstruct jets using these particle flow objects. A range of jet radii R ¹ may be chosen within which the anti- k_t algorithm should include particle flow objects for jet reconstruction. The choice of R depends on the kinematics, and on anticipated origins of the quark or gluon that initiated the shower (see discussion in Section 3.4.2 for more details).

As discussed in Chapter 2, the final state signature of the DH model targeted in this search involves a pair of energetic W bosons in the final state, one of which decays leptonically to a $\ell\nu$ pair, and the other hadronically to a pair of quarks. If the boost of the hadronically decaying W is sufficiently low, the angular separation between the two quarks may be large enough that the quarks are most effectively reconstructed as two separate jets, each with a small jet radius R . In the so-called “resolved” regime of the search, the two sets of energy deposits in the calorimeter produced by the $W \rightarrow qq$ decay are so separated that it is not even possible to reconstruct the two quarks within a single large-radius jet. For this search, these so-called “small- R ” jets are reconstructed with a jet radius of $R = 0.4$.

After all small- R jets in the final state are reconstructed and fully calibrated, as described in Ref. [147], only jets with $p_T > 20$ GeV and $|\eta| < 2.5$ are considered for the search. Jet cleaning [148] with the *TightBad* working point is applied to suppress noise in the calorimeter, as well as background jets that are not produced from the primary pp collision. The jet vertex tagger [149] is applied with the *Tight* working point to suppress pileup jets (see Ref. [125] for a discussion of pileup and its simulation in the ATLAS detector) from other pp interactions in the same and neighbouring bunch crossings - see Section 4.2.2 for a more detailed discussion of pileup events.

2187 ***b*-tagging**

When b quarks are produced by pp collisions in the ATLAS detector, they immediately form “ b -hadrons” due to colour confinement. The b -hadrons subsequently decay primarily via the weak force to form lighter hadrons. The so-called secondary decays of b -hadrons occur with a typical lifetime of $\sim 1.5\text{ps}$ [22], and as a result the b -hadrons can travel several millimeters from the primary pp interaction point

¹See Eq. 3.4 for the definition of the angular radius R .

[150] before undergoing secondary decay. The displaced secondary vertex represents a signature of b -hadron decay in the ATLAS detector. It can be reconstructed using precision tracking of charged particles provided by the inner detector, and used to assign a “ b tag” to hadronic jets in the ATLAS calorimeter to identify them as having originated from the decay of a b -hadron. This so-called “ b -jet tagging” is performed with the `DL1r` algorithm [150], which uses a deep learning method for the identification. A fixed working point with a 77% efficiency is used. b -tagged jets are vetoed in the signal region to reduce the background of SM $t\bar{t}$ and single-top processes (see Section 4.4.1 for details).

5.1.3 Resolved W Candidate

As described in Section 5.1.2 above, the pair of quarks produced by the hadronic decay of the W boson in the signal model are reconstructed as two resolved small- R jets in the less-boosted resolved regime. The parent W boson can then be reconstructed from small- R jets induced by its daughter quarks using the combined energy and momentum of the small- R jet pair. Given that small- R jets can also be produced by, for example, initial-state radiation and pileup, it is quite common for there to be more than two small- R jets reconstructed in the final state. These additional jet sources complicate the task of identifying which of the reconstructed small- R jets in a given event should be associated with the $W \rightarrow qq$ in the signal model. For events with more than two small- R jets in the final state, the pair of small- R jets whose combined invariant mass is closest to the W boson mass is assumed to have originated from the W decay, and used to reconstruct the W boson candidate. The algorithm for this jet identification and W boson reconstruction is as follows:

- Construct all possible combinations of two small- R jets (a.k.a. “dijet pairs”) in the final state.
- For each such candidate dijet pair, j_1 and j_2 , sum the four-momenta of the reconstructed jets, $\mathbf{p}_{j_1,j_2} = \mathbf{p}_{j_1} + \mathbf{p}_{j_2}$, and calculate their combined invariant mass:

$$M_{j_1,j_2} = \sqrt{\mathbf{p}_{j_1,j_2} \cdot \mathbf{p}_{j_1,j_2}} \quad (5.1)$$

- Select the dijet pair whose invariant mass is closest to the W boson mass of 80.4 GeV [22] as the small- R jets to be associated with the $W \rightarrow q\bar{q}$ decay.

- 2223 • Reconstruct the hadronically decaying W boson candidate using the dijet pair
 2224 with four-momentum \mathbf{p}_{j_1,j_2} .

2225 Figure 5.2 shows distributions of the reconstructed W boson candidate mass for
 2226 MC simulated events generated for a range of m_s and $m_{Z'}$ after application of the
 2227 baseline event selections presented in Section 5.2, with the additional requirement
 2228 that there be at least two small- R jets in the final state. The distributions are in
 2229 general well centred around the W boson mass.

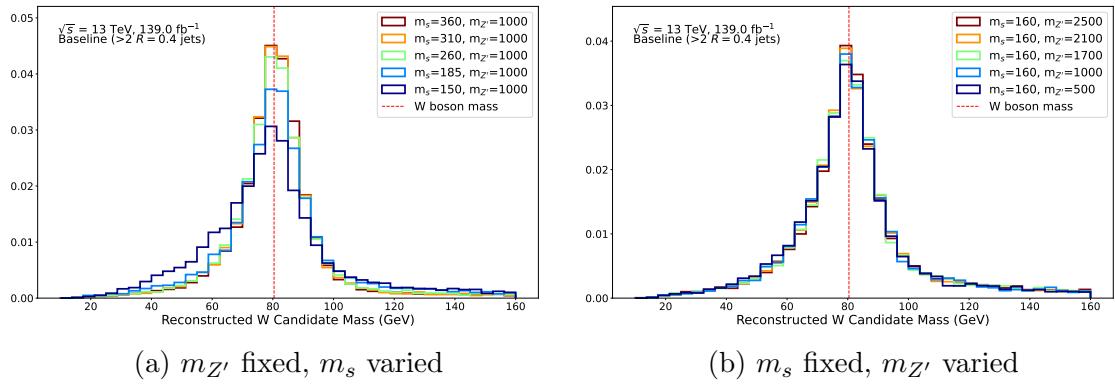


Figure 5.2: Distributions of the reconstructed W candidate mass for MC simulated events produced with the DH signal model over a range of m_s and $m_{Z'}$. All events included in the distributions are required to have passed the baseline event selection described in Section 5.2, and to have at least two small- R jets in the final state. The red dashed vertical line is placed at the W boson mass of 80.4 GeV. Distributions are normalized to unit area.

2230 **5.1.4 Track-Assisted Reclustered Jets (Merged W Candidate)**

2231 If the hadronically decaying W boson is produced with a sufficiently large momentum
 2232 (i.e. boost), the jets produced by the $q\bar{q}$ pair may be sufficiently collimated (i.e.
 2233 “merged”) that they are most effectively reconstructed as a single multi-pronged
 2234 large- R jet, as opposed to the resolved small- R jets used for W reconstruction in the
 2235 resolved regime (see Sections 5.1.2 and 5.1.3 above for details).

2236 Since the signal model predicts that charged particle tracks and energy deposits
 2237 in the detector will have originated primarily from the two quarks produced by the
 2238 $W \rightarrow q\bar{q}$ decay in the signal model, it is important to reconstruct the large-radius
 2239 jet in this so-called merged regime with as much detailed kinematic and substructure
 2240 information as possible. This information is used in the search to help identify whether

the hadronic activity contained within a large- R jet is consistent with having been induced by two energetic quarks originating from a W parent, as predicted by the signal model. Such features include the combined invariant mass $m^{\text{TAR Jet}}$ of all particles associated with the jet, which would be expected to be consistent with the W boson mass within detector resolution. Important substructure information includes variables that aim to quantify the number of distinct “prongs” of localized energy deposition within the jet, which can be correlated to the number of high- p_{T} strongly interacting particles whose energy deposits are included in the jet (two such prongs would be expected for the signal model).

This search uses the track-assisted reclustered (TAR) jet algorithm [2] for large- R jet reconstruction in the merged regime. In this regime, the highest- p_{T} TAR jet reconstructed with a radius parameter of $R = 1.0$ is used to reconstruct the hadronically decaying W boson (W_{had}) in the DH signal model. Whereas most large- R jet reconstruction techniques rely on energy deposits in the calorimeter to reconstruct the jet substructure information, TAR jets are designed to profit from the superior resolution of the inner tracker for improved substructure reconstruction by matching charged particle tracks with energy deposits in the calorimeter.

2258 **TAR Algorithm**

For this search, $R = 0.2$ small- R jets are used to reconstruct energy deposits in the calorimeter, and are input to the TAR algorithm along with tracks measured by the inner detector that satisfy a set of quality criteria summarized in Table 5.1. The TAR algorithm [2] is as follows: the input $R = 0.2$ small- R “subjets” are reclustered using the anti- k_t algorithm with $R = 1.0$ to form large- R jets. A trimming procedure is applied to mitigate the effects of pileup and background QCD processes within the triggered event that do not originate from the hard interaction. The trimming procedure removes any of the input subjets that carry less than a fraction $f_{\text{cut}} = 0.05$ of the total transverse momentum of the large- R jet: $p_{\text{T}}^{\text{subjett}}/p_{\text{T}}^{\text{large-}R \text{ jet}} < 0.05$. The tracks from the inner detector are then matched to the remaining small- R subjets using the ghost association procedure described in Ref. [151], if possible. Any tracks that cannot be matched to subjets using ghost association are instead matched to the nearest subjett, provided that there is a jet within an angular radius $\Delta R = 0.3$ of the track. To account for the energy of the neutral hadronic jet components, which do not leave tracks in the inner detector, the p_{T} of each track is scaled such that the

2274 summed p_T of all tracks matched to a given subjet will evaluate to the energy of the
 2275 subjet as measured by the calorimeter:

$$p_T^{\text{track, new}} = p_T^{\text{track, old}} \times \frac{p_{T,j}^{\text{subjet}}}{\sum_{i \in j} p_{T,i}^{\text{track, old}}} \quad (5.2)$$

2276 where the index i runs over all tracks matched to the subjet j . The rescaled tracks
 2277 and remaining subjets are again reclustered using the anti- k_t algorithm to form the
 2278 final large- R TAR jet.

2279 While the kinematic properties of the TAR jets are calculated from the constituent
 2280 small- R jets, the jet substructure and mass m^{TAR} are calculated from the constituent
 2281 tracks. Figure 5.5 shows a visual summary of the basic TAR algorithm.

2282 TAR-lepton Disentanglement

2283 The analysis applies the TAR algorithm to $R = 0.2$ small- R jets and tracks that have
 2284 undergone a “TAR-lepton disentanglement” preselection to remove tracks associated
 2285 with any reconstructed baseline electrons or muons, as well as any $R = 0.2$ jets that
 2286 overlapped with the baseline electron tracks. This preselection is helpful given the
 2287 final state targeted in the search, because the charged lepton produced by the leptonic
 2288 $W \rightarrow \ell\nu$ decay often falls within the $R = 1.0$ cone of the TAR jet, as illustrated in
 2289 Figure 5.3. This TAR-lepton overlap disrupts the jet reconstruction, particularly due
 2290 to the additional jet energy induced by calorimetric clusters created in the large- R
 2291 jet by the overlapping electron.

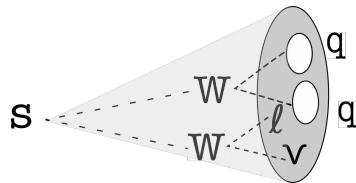


Figure 5.3: Illustration of the final state scenario in which the charged lepton produced by the leptonic $W \rightarrow \ell\nu$ decay overlaps with the large- R TAR jet reconstructed from the hadronic $W \rightarrow qq$ decay.

2292 Figure 5.4 shows a comparison of the distributions of reconstructed TAR jet mass
 2293 $m^{\text{TAR Jet}}$ either without or with the TAR-lepton disentanglement preselection applied,
 2294 for MC simulated events produced with the DH signal model at several representative
 2295 m_s and $m_{Z'}$, in which the reconstructed electron in the final state overlaps with

the highest- p_T reconstructed TAR jet. For all the signal points, the TAR-lepton disentanglement preselection is found to substantially improve the ability of the TAR algorithm to reconstruct TAR jets with $m^{\text{TAR Jet}}$ near the W boson mass, as would be expected for the signal model.

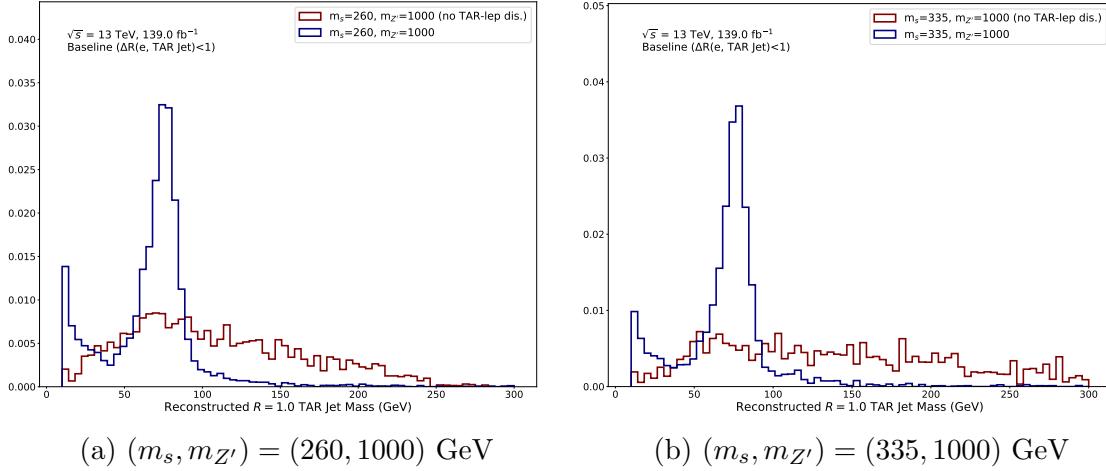


Figure 5.4: Distributions of $m^{\text{TAR Jet}}$ for the leading- p_T TAR jet in MC simulated events generated for the DH signal model process with semileptonic WW decay at several representative m_s and $m_{Z'}$, with and without application of the lepton disentanglement preselection. Events included in the distributions are required to have one signal electron and at least one reconstructed TAR jet, both within an angular radius of $\Delta R = 1.0$. Distributions are normalized to unit area.

2300 Summary of the TAR Procedure

2301 The following steps summarize the algorithm used to construct the TAR jets used in
2302 this search (steps with a * are included to disentangle leptons):

- 2303 • Tracks and calibrated anti- k_t $R = 0.2$ jets are chosen as input to the algorithm.
- 2304 • Tracks associated with a baseline muon or electron are removed from the input
2305 collection (*).
- 2306 • $R = 0.2$ jets overlapping with a baseline electron ($\Delta R < 0.2$) are removed from
2307 the input collection (*).
- 2308 • The remaining $R = 0.2$ subjets are reclustered using the anti- k_t algorithm into
2309 $R = 1.0$ jets, and trimmed using the p_T fraction $f_{\text{cut}} = 0.05$.

- 2310 • Input tracks are matched to $R = 0.2$ subjets that remain after trimming, if
 2311 possible, using ghost association.
- 2312 • Tracks that remain unassociated are matched to the nearest anti- k_t $R = 0.2$ jet
 2313 within $\Delta R < 0.3$.
- 2314 • The p_T of each track is rescaled using the p_T of the jet to which it is matched
 2315 using Eq. 5.2. This rescaling accounts for the missing neutral momentum,
 2316 which is measured at calorimeter level but is not present at tracker level.
- 2317 • Finally, jet substructure variables and m^{TAR} are calculated using the rescaled
 2318 matched tracks.

2319 The parameters of the TAR algorithm used are summarized in Table 5.1.

2320

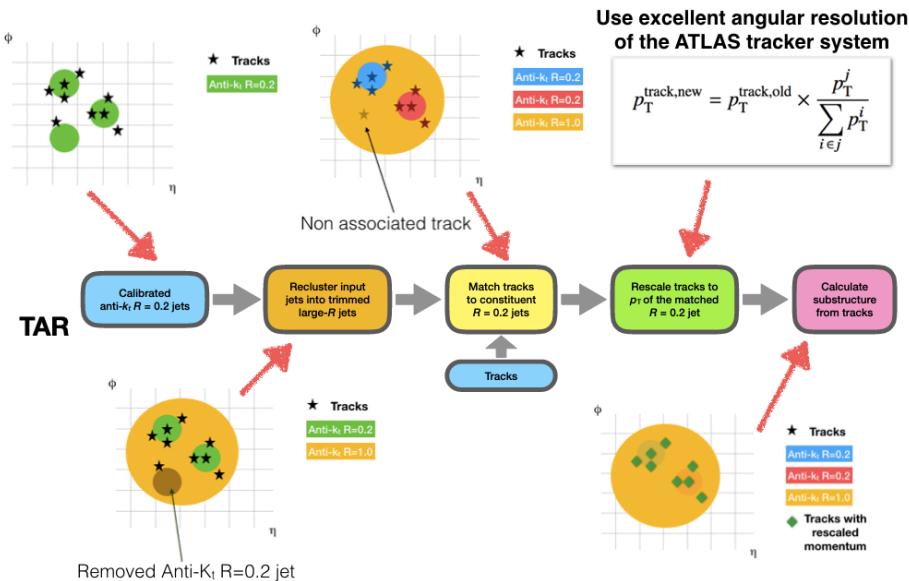


Figure 5.5: TAR jet reconstruction algorithm depicted without lepton disentanglement. Figure adapted from © [2]

2321 **5.1.5 $E_{\text{T}}^{\text{miss}}$**

2322 The missing transverse momentum $E_{\text{T}}^{\text{miss}}$, introduced in Section 3.4.4, quantifies the
 2323 imbalance of momentum in the plane transverse to the beam line. In the event that
 2324 all particles produced in a pp collision are fully detected, conservation of momentum
 2325 implies that the transverse momenta of all objects produced by the event should sum

Table 5.1: TAR jet reconstruction parameters.

	Loose quality
Track selection	$p_T > 0.5 \text{ GeV}$ $ \eta < 2.5$
Tracks removed if associated to	electrons, muons
	$R = 0.2$ anti- k_t jets
Input jet selection	$p_T > 20 \text{ GeV}$ $ \eta < 2.5$
Reclustering radius	$R = 1.0$
TAR jet p_T	$p_T^{\text{TAR}} > 100 \text{ GeV}$
Trimming radius	$R = 0.2$
Trimming p_T fraction	$f_{\text{cut}} = 0.05$
Track-to-jet association	$\Delta R(\text{jet}, \text{track}) < 0.3$
jet-electron overlap removal	$\Delta R(\text{jet}, \text{electron}) < 0.2$

2326 to zero within detector resolution. As a result, large E_T^{miss} in an event is indicative of
 2327 the production of undetected energetic particles. The semileptonic $s \rightarrow WW(qq\ell\nu)$
 2328 decay channel of the LHC signature for the DH model probed in this DM search (see
 2329 Chapter 2 for details) predicts large (i.e. above-detector-resolution) E_T^{miss} in the final
 2330 state. The large E_T^{miss} would be due both to the DM pair produced from the decay of
 2331 the hypothetical Z' , and to the neutrino produced by the leptonic decay of one of the
 2332 W bosons in the final state. Both the DM pair and the neutrino would be expected
 2333 to pass through the detector without any appreciable interactions due to their very
 2334 low interaction cross sections with SM particles, and hence constitute undetected (i.e.
 2335 missing) momentum in the event.

2336 The E_T^{miss} is calculated using fully calibrated and reconstructed physics objects
 2337 (for details, see Ref. [152]). For this search, baseline electrons and muons (see Sec-
 2338 tion 5.1.1) and $R = 0.4$ jets (see Section 5.1.2) are used to construct the E_T^{miss} . A
 2339 soft term is additionally included, which uses tracks that are not associated with any
 2340 of these reconstructed objects.

2341 This search also makes use of the object-based E_T^{miss} significance \mathcal{S} [153], which is
 2342 designed to be positively correlated with the likelihood that the measured E_T^{miss} was
 2343 actually produced by undetected particles in the event, rather than by fluctuations
 2344 arising from the limited detector resolution. The E_T^{miss} significance is calculated on an
 2345 event-by-event basis using the uncertainties associated with the reconstructed objects
 2346 involved in the E_T^{miss} calculation for the given event.

2347 5.1.6 Overlap Removal

2348 To avoid double-counting any physics objects in an event, a priority-based overlap
 2349 removal (OR) strategy is employed, which eliminates any overlap between the physics
 2350 objects. This is accomplished by removing all but the highest-priority object from
 2351 any region in which objects overlap. The strategy presented in this section resolves
 2352 any overlap between electrons, muons and $R = 0.4$ small- R jets. The overlap removal
 2353 between leptons and TAR jets is described in Section 5.1.4. No overlap removal
 2354 between $R = 0.4$ jets and TAR jets is applied, as they are not used in the same
 2355 selection (see Section 5.2 for details). Table 5.2 summarizes the criteria under which
 2356 overlap is removed between a given pair of objects.

2357 Overlap removal is performed for baseline objects, and only the remaining objects
 2358 are considered as candidate signal objects. Note that the calorimeter-tagged (CT)

2359 muons listed in Table 5.2, details of which can be found in Section 4 of Ref. [120], are
 2360 identified and reconstructed using only inner detector tracks and calorimeter energy
 2361 deposits consistent with a minimum-ionizing particle, and do not have any associ-
 2362 ated hits identified in the muon spectrometer. Due to the absence of any associated
 2363 signal in the muon spectrometer, these CT muons are given a relatively low priority
 2364 in the OR procedure compared with non-CT muons, which do activate the muon
 2365 spectrometer.

Table 5.2: Object priorities and overlap removal criteria for each pair of physics objects considered in the OR procedure. Object pairs and removal criteria are listed in the sequence by which they are considered for OR, with the top row considered first.

Removed Object	Retained Object	Criteria for OR
Electron (lower p_T)	Electron (higher p_T)	shared inner detector track
Muon	Electron	shared ID track, and muon is CT
Electron	Muon	shared ID track, and muon is not CT
anti- k_t 4 Jet	Electron	Angular separation $\Delta R < 0.2$
Electron	anti- k_t 4 Jet	$\Delta R < \min(0.4, 0.04 + 10 \text{ GeV}/p_T(e))$
anti- k_t 4 Jet	Muon	fewer than 3 tracks in jet, and (muon is ghost-associated to jet, or $\Delta R < 0.2$)
Muon	anti- k_t 4 Jet	$\Delta R < \min(0.4, 0.04 + 10 \text{ GeV}/p_T(\mu))$

2366 5.1.7 Dark Higgs Candidate Mass

2367 In principle, the four-momentum of the Dark Higgs boson s in the DH signal model
 2368 is simply the sum of the four-momenta of the WW pair that it decays to:

$$\mathbf{p}_s = \mathbf{p}_{W_{\text{had}}} + \mathbf{p}_{W_{\text{lep}}} \quad (5.3)$$

2369 where W_{had} (W_{lep}) denotes the hadronically (leptonically) decaying W boson. The
 2370 W_{had} four-momentum is reconstructed in the resolved regime using the pair of anti- k_t
 2371 4 jets whose invariant mass is closest to the on-shell W mass of 80.4 GeV (see Section
 2372 5.1.3), or in the merged regime as the four-momentum of the highest- p_T $R = 1.0$ TAR
 2373 jet (see Section 5.1.4). The four-momentum of the W_{lep} is the sum of four momenta
 2374 of its lepton and neutrino daughters:

$$\mathbf{p}_{W_{\text{lep}}} = \mathbf{p}_\ell + \mathbf{p}_\nu \quad (5.4)$$

If the neutrino were the only anticipated source of “true E_T^{miss} ” (i.e. E_T^{miss} arising from undetected particles rather than limited detector resolution) in the final state, the final state E_T^{miss} could be unambiguously assigned to the neutrino, i.e. $(p_{x,\nu}, p_{y,\nu}) = (E_x^{\text{miss}}, E_y^{\text{miss}})$, at which point the only missing information would be the z-component $p_{\nu,z}$ of the neutrino momentum. However, since the DH signal model additionally predicts E_T^{miss} originating from the DM pair in the final state, there is some ambiguity involved with assessing how much of the measured E_T^{miss} is accounted for by the neutrino vs. the DM pair.

An approximate solution is obtained by determining the minimum m_s that would be required in order for the s decay to have produced a lepton and W_{had} with the observed momenta, subject to the constraint that the invariant mass of the reconstructed $m_{W_{\text{lep}}}$ be equal to the on-shell W mass of 80.4 GeV. Although this minimum m_s may not necessarily evaluate to the actual modelled m_s , by providing an absolute lower bound on the possible range of m_s that could produce the observed final state, it is expected to at least be positively correlated with the actual modelled m_s .

To simplify the math involved in determining the minimum m_s , the coordinate system is rotated without loss of generality such that the lepton is strictly traveling along the z axis, and the hadronically decaying W boson W_{had} is in the xz plane, as shown in Figure 5.6.

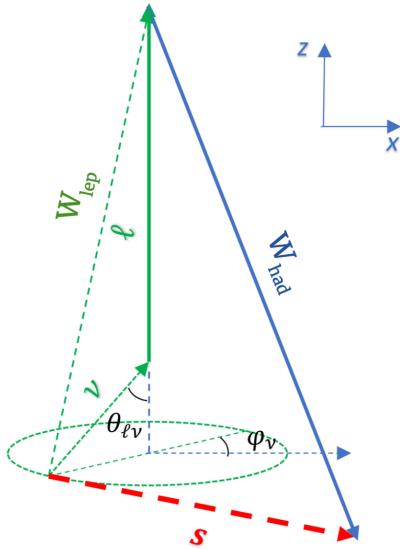


Figure 5.6: Coordinate system used to evaluate the minimum DH mass m_s that is kinematically required to produce the observed final state, subject to the constraint that $m_{W_{\text{lep}}} = 80.4$ GeV.

2394 In this rotated coordinate system, the four-momenta \mathbf{p}_ν , \mathbf{p}_ℓ of the neutrino, lepton
2395 and hadronically decaying W boson, respectively, are given by:

$$\mathbf{p}_\nu = E_\nu(1, \sin \theta_{\ell\nu} \cos \phi_\nu, \sin \theta_{\ell\nu} \sin \phi_\nu, \cos \theta_{\ell\nu}) \quad (5.5)$$

$$\mathbf{p}_\ell = E_\ell(1, 0, 0, 1) \quad (5.6)$$

2396 and

$$\mathbf{p}_{W_{\text{had}}} = (E_{W_{\text{had}}}, p_{W_{\text{had}},x}, 0, p_{W_{\text{had}},z}) \quad (5.7)$$

2397 where $\theta_{\ell\nu}$ is the angular separation between the lepton and the neutrino, and ϕ_ν is the
2398 angle of the neutrino relative to the x axis in the xy plane. The m_s is then obtained
2399 by squaring the four-momenta in Eq. 5.3:

$$\begin{aligned} m_s^2 &= (\mathbf{p}_{W_{\text{had}}} + \mathbf{p}_{W_{\text{lep}}})^2 = (\mathbf{p}_{W_{\text{had}}} + \mathbf{p}_\ell + \mathbf{p}_\nu)^2 \\ &= (E_{W_{\text{had}}} + E_\ell + E_\nu)^2 - (p_{W_{\text{had}},x} + E_\nu \sin \theta_{\ell\nu} \cos \phi_\nu)^2 - (E_\nu \sin \theta_{\ell\nu} \sin \phi_\nu)^2 - (E_\ell + p_{W_{\text{had}},z} + E_\nu \cos \theta_{\ell\nu})^2 \end{aligned} \quad (5.8)$$

2400 It can be shown by taking derivatives of Eq. 5.8 that the minimum m_s occurs when
2401 $\phi_\nu = 0$ (i.e. when the neutrino is in the same plane as the $\mathbf{p}_{W_{\text{had}}}$).

2402 Setting $\phi_\nu = 0$ in Eq. 5.8 and using the Pythagorean identity $\sin \theta = \sqrt{1 - \cos^2 \theta}$:

$$\begin{aligned} m_s^2 &= (E_{W_{\text{had}}} + E_\ell + E_\nu)^2 - \left(p_{W_{\text{had}},x} + E_\nu \sqrt{1 - \cos^2 \theta_{\ell\nu}} \right)^2 \\ &\quad - (E_\ell + p_{W_{\text{had}},z} + E_\nu \cos \theta_{\ell\nu})^2 \end{aligned} \quad (5.9)$$

2403 This leaves an equation for m_s with two unknowns: the energy E_ν of the neutrino,
2404 and the cosine $\cos \theta_{\ell\nu}$ of the angle between the lepton and the neutrino. The neutrino
2405 energy is determined as a function of $\cos \theta_{\ell\nu}$, by imposing the constraint that the mass
2406 $m_{W_{\text{lep}}}$ of the leptonically decaying W boson be set to the on-shell W boson mass of
2407 $m_W = 80.4$ GeV:

$$m_{W_{\text{lep}}}^2 = m_W^2 = (p_\ell + p_\nu)^2 = 2p_\ell p_\nu = 2E_\ell E_\nu(1 - \cos \theta_{\ell\nu}) \quad (5.10)$$

2408 Solving for E_ν :

$$E_\nu = \frac{m_W^2}{2E_\ell(1 - \cos \theta_{\ell\nu})} \quad (5.11)$$

2409 With this independent determination of E_ν , the minimum m_s in Eq. 5.9 is eval-
2410 uated numerically by scanning over $\cos \theta_{\ell\nu} \in [-1, 1]$ and identifying the value of
2411 $\cos \theta_{\ell\nu}$ that minimizes m_s (excluding $\cos \theta_{\ell\nu} = 1$ to avoid a singularity in Eq. 5.11).

2412 Figure 5.7 shows distributions of this minimized “ $\min(m_s)$ ” for MC simulated
2413 events produced with the DH signal model over a range of m_s (left column) or $m_{Z'}$
2414 (right column). The distributions are more sharply peaked for lower m_s , and for
2415 higher m_s the location of peak in $\min(m_s)$ becomes increasingly shifted to the left
2416 of (i.e. below) the actual modelled m_s . The minimal variation between the different
2417 modelled values of $m_{Z'}$ presented in distributions in the right-hand column of Figure
2418 5.7, which scan over a range of $m_{Z'}$ for the same m_s , offers an encouraging indication
2419 that the value of the $\min(m_s)$ variable is primarily a function of the m_s parameter in
2420 the model that it is designed to approximate.

2421 Despite the shifted location of the peaks at higher m_s , the presence of distinct
2422 peaks in the approximate vicinity of the modelled m_s imply that the $\min(m_s)$ variable
2423 can be a valuable tool to aid in discriminating events in the data that could be
2424 consistent with the DH signal process from the SM background processes. For this
2425 reason, events in the signal regions are binned in m_s when searching for evidence of
2426 the signal model in the ATLAS collision data. The binning in $\min(m_s)$ is presented
2427 in detail in Section 7.2.

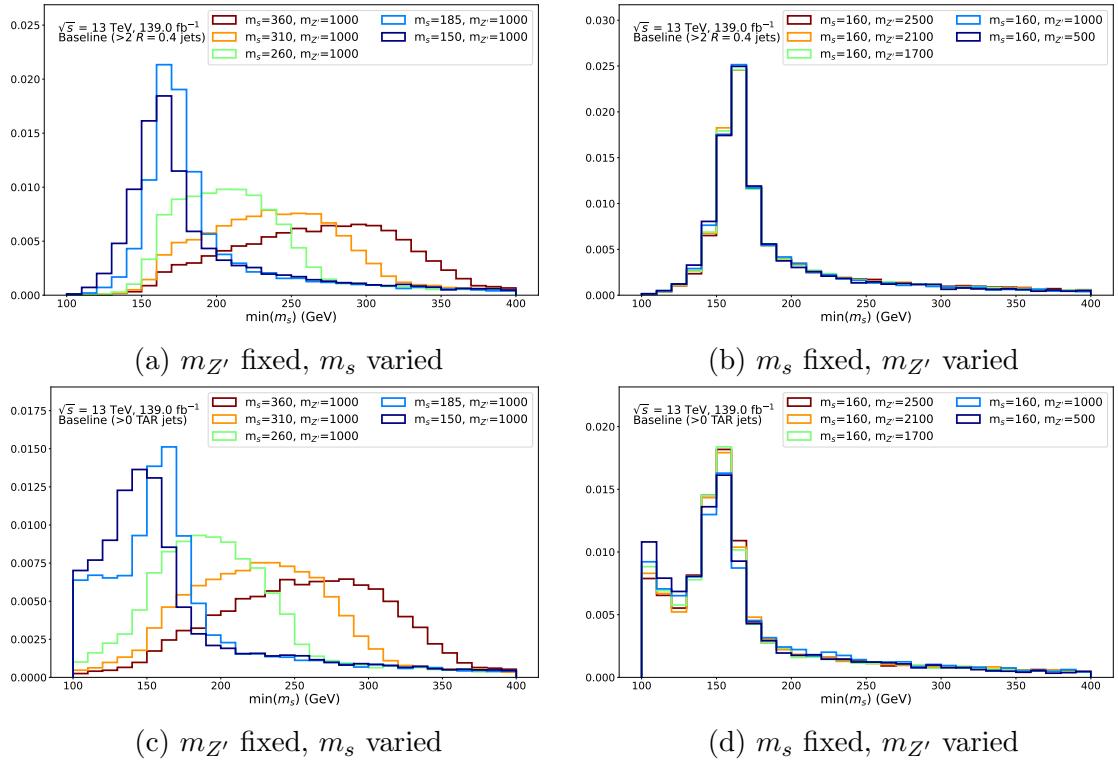


Figure 5.7: Distributions of the DH candidate mass $\min(m_s)$, reconstructed using the minimization strategy presented in Section 5.1.7, for MC simulated events produced for the DH signal model over a range of m_s and $m_{Z'}$. Distributions are normalized to unit area. Events included in all distributions are required to pass the baseline selection requirements presented in Section 5.2. For the distributions shown in the top (bottom) row, the W_{had} is reconstructed as a resolved candidate (TAR jet) using the strategy presented in Section 5.1.3 (5.1.4), and events are required to have $\geq 2 R = 0.4$ small- R jets ($\geq 1 R = 1.0$ TAR jets) reconstructed in the final state.

2428 5.1.8 Transverse Mass

2429 The transverse mass $m_T(E_{\text{T}}^{\text{miss}}, \ell)$ between the lepton and $E_{\text{T}}^{\text{miss}}$ is considered in this
 2430 search because it is sensitive to the presence of additional $E_{\text{T}}^{\text{miss}}$ beyond that arising
 2431 from the neutrino in the leptonic decay of the W_{lep} . It is computed for events measured
 2432 in the ATLAS detector as:

$$m_T(\ell, E_{\text{T}}^{\text{miss}}) = \sqrt{2p_{T,\ell}E_{\text{T}}^{\text{miss}}(1 - \cos\theta_{\ell,E_{\text{T}}^{\text{miss}}})} \quad (5.12)$$

2433 which is derived from the more general transverse mass definition found in Section
 2434 48.6.1 of Ref. [22].

$$\begin{aligned}
m_{T, \text{full}}^2(\ell, E_T^{\text{miss}}) &= (p_{T,\ell} + p_{T,E_T^{\text{miss}}})^2 \\
&= m_\ell^2 + m_{E_T^{\text{miss}}}^2 + 2E_{T,\ell}, E_{T,E_T^{\text{miss}}} (1 - \cos \theta_{\ell,E_T^{\text{miss}}}) \quad (5.13)
\end{aligned}$$

under the assumptions that the masses associated with the lepton and E_T^{miss} are negligibly small compared with their momenta. The assumption of negligible lepton mass is in general justified given the high energy of the pp collisions at the LHC. The assumption of negligible mass associated with E_T^{miss} is justified if the true E_T^{miss} arises only from the neutrino in the leptonic W_{lep} decay, as it would in the leading SM backgrounds. In the signal model, however, there is additional mass associated with the E_T^{miss} , which arises from the decay of the massive Z' mediator to an invisible DM pair. The result, shown in figure 5.8 after applying the baseline selection, is that the bulk of the SM background has $m_T(\ell, E_T^{\text{miss}})$ below the W mass peak, but the signal distribution tends to be peaked closer to ~ 250 GeV.

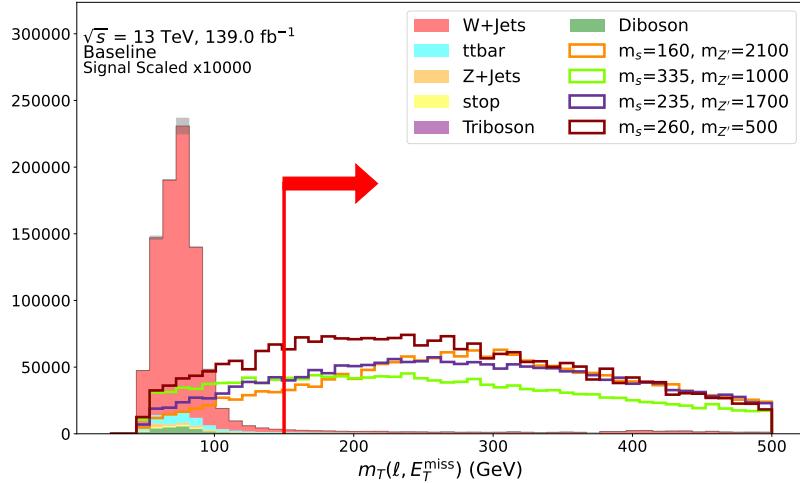


Figure 5.8: Transverse mass distribution for SM background and several signal points with baseline selections, excluding the lower bound on $m_T(\ell, E_T^{\text{miss}})$. The red line and arrow indicate the placement of the baseline selection on $m_T(\ell, E_T^{\text{miss}})$.

5.2 Event Selections

Selections are applied to the ATLAS collision data and MC simulated events with the aim of defining subsets, or “regions” of data that are enriched in a particular process of interest for the search. The regions are designed and optimized using MC simulated

2449 data (see Chapter 4 for a detailed discussion of MC simulation and its application to
 2450 simulating data produced by the ATLAS detector). The use of MC simulated data
 2451 makes it possible to quantify the relative contribution to the expected yield of events
 2452 in the region arising from each physics process predicted in the data.

2453 Selections that define the “signal regions” (SRs) are optimized to produce an
 2454 enriched yield of MC simulated events produced using the DH signal model (referred
 2455 to as “signal events”), with a minimal yield of simulated events generated to model SM
 2456 background processes (referred to as “background events”). A discrepancy between
 2457 the ATLAS collision data and the predicted yield of SM backgrounds in the signal
 2458 regions would indicate the presence of a BSM physics process with a production
 2459 mechanism at the LHC consistent with that of the signal model. “Control regions”
 2460 (CRs) are optimized to have an enriched yield of MC simulated events modelled for
 2461 one particular SM background process. $W + \text{jets}$ and $t\bar{t}$ CRs are defined for this DM
 2462 search to obtain data-driven constraints on the total yields of these SM background
 2463 processes in the signal region.

2464 5.2.1 Kinematic Categories

2465 Within each of the SRs and CRs, the analysis selection is divided into two kinematic
 2466 regimes, referred to as “categories”. The “merged” category is designed to target
 2467 the merged regime discussed in Section 5.1.4 in which the hadronic decay products
 2468 are sufficiently boosted as to be reconstructed as a single $R = 1.0$ TAR jet. The
 2469 leading p_T TAR jet is then used to reconstruct the candidate hadronically decaying
 2470 W boson W_{had} in the signal model. The “resolved” category targets the lower- p_T
 2471 regime in which the hadronic decay products have a sufficient angular separation
 2472 that they cannot be reclustered into a TAR jet, and are instead reconstructed as
 2473 resolved small- R anti- k_t $R = 0.4$ jets. As described in Section 5.1.3, the W_{had} is
 2474 reconstructed in the resolved category using the two small- R jets whose combined
 2475 invariant mass is nearest to $m_W = 80.4$ GeV.

2476 Figure 5.9 illustrates the two kinematic categories.

2477 Within each of the signal and control regions, the selection is divided into the
 2478 merged and resolved kinematic categories presented in Section 5.2.1. Events are
 2479 classified into the merged category if there is at least one reconstructed $R = 1.0$
 2480 TAR jet in the final state ($N(\text{TAR Jets}) > 0$), and the resolved category if there
 2481 are at least two reconstructed $R = 0.4$ small- R jets ($N(\text{Jets}) > 1$). Selections are

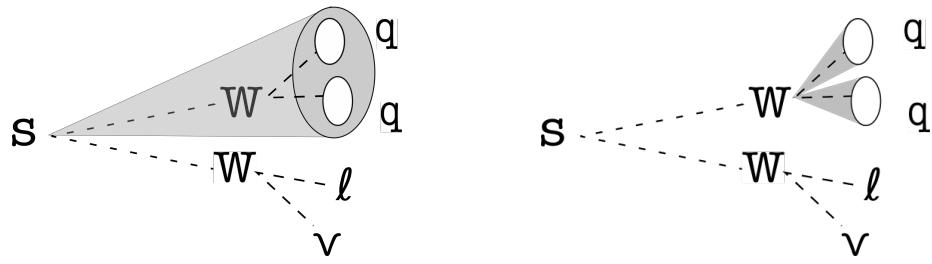


Figure 5.9: A graphical representation of the two kinematic categories used in the search based on the characteristics of the final state. Left: merged category, in which the jets produced by the $w \rightarrow q\bar{q}$ are sufficiently boosted as to be reconstructed within a single large- R TAR jet. Right: resolved category, in which the two quarks are reconstructed separately as two small- R anti- k_t $R = 0.4$ jets.

further refined and optimized separately within each category. It is worth noting that, as is, the requirements for events to be classified into the merged or resolved categories are not mutually exclusive - i.e. there are events that have least one $R = 1.0$ TAR jet **and** at least two small- R jets. To avoid double-counting the same events in the search, orthogonality between the merged and resolved categories is therefore enforced in all of the signal and control regions by explicitly requiring events that pass the requirements to be classified into the resolved category in any region to have additionally failed the merged category selection in all regions.

5.2.2 Baseline selection

The baseline preselection is common to all analysis regions and categories used in the search, and is used to roughly define the region of interest for the search prior to any detailed optimization of selection requirements or sub-division into separate analysis regions. The baseline selection is defined as follows:

- (1 signal muon or 1 signal electron) and no additional baseline muons or electrons
- (E_T^{miss} trigger passed) OR ((single muon trigger passed) AND (signal muon matched to muon trigger))
- $E_T^{\text{miss}} > 200 \text{ GeV}$
- $\mathcal{S} > 5$
- $m_T(\ell, E_T^{\text{miss}}) > 150 \text{ GeV}$
- $N(\text{TAR Jets}) > 0 \text{ or } N(\text{Jets}) > 1$

See Section 5.1.1 for the definitions of baseline and signal muons and electrons, and Section 5.3 for details of the E_T^{miss} and single muon triggers as well as the single muon trigger matching requirement. The veto on baseline muons or electrons in addition to requirement of a single signal lepton is designed to improve the purity of the single lepton final state predicted by the signal model by removing events in which an additional lepton was produced from the hard scatter and reconstructed as a baseline lepton, yet failed the criteria to be identified as a signal lepton.

2510 **5.2.3 Signal Region Definition**

2511 In addition to the baseline selection, a veto on jets identified as having been induced
 2512 by a b quark (a.k.a. a b -jet veto) is applied in the SR to reduce the yield of SM $t\bar{t}$
 2513 and single-top processes discussed in Section 4.4. See Section 5.1.2 for details of the
 2514 b -jet tagging algorithm.

2515 Within each category, the baseline selection is further refined by optimizing the
 2516 exact placements of upper or lower bounds, also referred to as “cuts”, on variables
 2517 for which the distributions of MC simulated events generated according to the signal
 2518 model differs appreciably from MC simulated distributions of the SM background
 2519 processes.

2520 Broadly, the cut placements are optimized to maximize the predicted yield of MC
 2521 simulated events that model the DH signal process relative to the predicted yield
 2522 of events that model the SM background processes. However, this basic benchmark
 2523 fails to account for the fact that, as discussed in Section 4.1, the relative “statistical
 2524 uncertainty” associated with the limited number of MC simulated events used to
 2525 calculate yield predictions will increase as the selections are tightened (i.e. as lower
 2526 bounds are increased or upper bounds are reduced) due to the resulting reduction
 2527 in the number of MC simulated events that pass the selections. As the relative
 2528 statistical uncertainty of the predicted yields increases, the ability of the search to
 2529 confidently identify an excess of ATLAS collision events above the predicted yield of
 2530 SM background processes - which would be indicative of additional events produced
 2531 by a BSM process - is reduced. Furthermore, as the predicted yield of events is
 2532 reduced, the relative statistical uncertainty associated with the number of observed
 2533 events, derived from the Poisson distribution, will also increase and similarly impact
 2534 the sensitivity of the search. Therefore, while it is often valuable to tighten certain
 2535 selections in order to increase the relative predicted yield of the signal process, it is also
 2536 important to avoid over-tightening them to the point that the statistical uncertainty
 2537 of the predicted and observed event yields begins to reduce the sensitivity of the
 2538 search.

2539 A metric known as the “Asimov discovery significance” [154] Z is used as a means
 2540 of quantifying the sensitivity of a signal region given the predicted yields s and b
 2541 of the signal and background processes, respectively, while also accounting for the
 2542 statistical uncertainty σ_b associated with the limited number of MC simulated events
 2543 used to predict b :

$$Z(s, b, \sigma_b) = \left[2(s+b) \left(\ln \left[\frac{(s+b)(b+\sigma_b^2)}{b^2 + (s+b)\sigma_b^2} \right] - \frac{b^2}{\sigma_b^2} \ln \left[1 + \frac{\sigma_b^2 s}{b(b+\sigma_b^2)} \right] \right) \right]^{\frac{1}{2}} \quad (5.14)$$

2544 **Optimization Strategy for Signal Region Definition**

2545 After applying the baseline selections and b -jet veto, the placement of upper and lower
 2546 bounds is optimized for the selection variables listed in Table 5.3. The SR was kept
 2547 “blinded” during optimization, which means that only the MC simulated signal and
 2548 background events were considered, and not the ATLAS collision data. Blinding is
 2549 done in order to avoid biasing the cut choices on the basis of any trends or fluctuations
 2550 that may be present in the distributions of ATLAS data in the SR.

2551 The choice of selection variables was made based on a visual assessment of the
 2552 impact of cuts on all the variables considered as candidate selection variables on the
 2553 Asimov discovery significance Z . Visualization of Z with respect to the cut placement
 2554 on each variable is done using so-called “N-1” plots, which are shown for the finalized
 2555 selections in the merged and resolved SRs, respectively, in Figures 5.10 and 5.11. The
 2556 N-1 plots are produced for a given variable v and a given set of candidate selections
 2557 on all other variables as follows:

- 2558 • Place the candidate selections on all variables except for v .
- 2559 • In the upper panel, plot the distributions of the background and signal processes,
 2560 binned in v .
- 2561 • In the lower panel, plot the distributions of the Asimov discovery significance
 2562 Z for each signal process plotted in the upper panel.
 - 2563 – If comparing the placement of an **upper bound** v_u on the variable v :
 2564 calculate the Asimov significance $Z(v_u)$ for all events with $v < v_u$:

$$\begin{aligned} s_u &= \sum_{i \in \{\text{signal}, v(i) < v_u\}} w(i) \\ b_u &= \sum_{i \in \{\text{SM backgrounds}, v(i) < v_u\}} w(i) \\ \sigma_{b_u} &= \sqrt{\sum_{i \in \{\text{SM backgrounds}, v(i) < v_u\}} [w(i)]^2} \end{aligned} \quad (5.15)$$

2565 where each event i is implicitly required to have passed all the other candidate
 2566 selections for the given process (signal or SM backgrounds) in addition
 2567 to $(v < v_u)$, and $w(i)$ is the event weight associated with event i (see dis-
 2568 cussion of event weights in Section 4.2.2). The statistical variance $(\sigma_b)^2$
 2569 is evaluated in Eq. 5.16 as the sum of squared weights for events in the
 2570 background process that pass all other candidate selections in addition to
 2571 $(v < v_u)$. Inserting s_u , b_u and σ_{b_u} into Eq. 5.14:

$$Z(v_u) = Z(s_u, b_u, \sigma_{b_u}) \quad (5.16)$$

- 2572 – Conversely, if comparing the placement of a **lower bound** v_d : calculate the
 2573 Asimov significance $Z(v_d)$ for all events with $v > v_d$ by replacing “ $v < v_u$ ”
 2574 in Eq. 5.16 with “ $v > v_d$ ”.

2575 The optimization was performed first in the merged category of the SR, after
 2576 placing an additional requirement of at least one $R = 1.0$ TAR jet. Once the selections
 2577 defining the merged SR were finalized, optimization was subsequently performed in
 2578 the resolved category, with the $N(\text{TAR Jets}) > 0$ requirement replaced by $N(\text{Jets}) >$
 2579 1 in addition to a veto on any events that pass the finalized merged SR selections.

2580 Since the optimal placement of selections was found to vary to some extent for MC
 2581 simulated signal data sets with different m_s and $m_{Z'}$, the cut placements were initially
 2582 optimized with the aim of maximizing the average Asimov discovery significance for
 2583 MC simulated data sets at the following four mass points, which cover most of the
 2584 m_s range considered in the search: $(m_s, m_{Z'}) = \{(210, 2100), (285, 1700), (310, 500),$
 2585 $(335, 1000)\}$ GeV. The $m_{Z'}$ values of these four mass points were chosen because
 2586 they are near the edge of the so-called “exclusion range”, which represents the range
 2587 of m_s and $m_{Z'}$ within which the search was expected to be sensitive to the presence
 2588 (or absence) of events produced by the DH signal model in the ATLAS data, at a
 2589 95% confidence level on the basis of sensitivity projections obtained using the method
 2590 presented in Section 7.3.3 with the Asimov data set. It is particularly desirable to
 2591 optimize the cut placements for mass points near the edge of the exclusion range in
 2592 order to extend this range as much as possible.

2593 An iterative approach was used to optimize cut placements at the four signal
 2594 points, which combined:

- 2595 • repeated grid searches, which scanned over 1,000,000 candidate multi-dimensional

2596 combinations of cut placements on some or all of the optimized variables to
 2597 identify combinations that maximized Z , and

- 2598 • visual analysis of N-1 plots such as those shown in Figures 5.10 and 5.11 to
 2599 visually validate the optimal placements found by the grid searches.

2600 It is worth noting that, inspecting Eq. 5.14, the Asimov discovery significance does
 2601 not account for the statistical uncertainty σ_s arising from limited MC simulated events
 2602 in the signal sample. Therefore, in order to ensure that there were sufficient events in
 2603 the signal samples, the grid search included an option to avoid cut combinations that
 2604 reduced the predicted signal yield below some acceptable minimum set by the user.
 2605 After some testing, it was found that setting a minimum acceptable predicted yield of
 2606 15 for the signal point $(m_s, m_{Z'}) = (210, 2100)$ GeV was adequate to ensure that the
 2607 signal samples of interest for cut optimization had a sufficient number of events as to
 2608 prevent their statistical uncertainty from becoming appreciable compared with other
 2609 sources of uncertainty. The signal point $(m_s, m_{Z'}) = (210, 2100)$ GeV was chosen
 2610 to define the minimum yield because, at the time of optimization, it was among the
 2611 mass points with the lowest predicted yield over the set of $m_{Z'}$ and m_s masses for
 2612 which MC simulated datasets were produced for the search (also referred to as the
 2613 “signal grid”).

2614 After an optimized set of selections were determined for the four signal points
 2615 considered in the iterative optimization procedure described above, these selections
 2616 were validated and further refined by examining the projected sensitivity of the search
 2617 over the signal grid. Some minor refinements were made to cut placements with the
 2618 aim of maximizing the exclusion range. See Section 7.3.3 for a description of how
 2619 the sensitivity is quantified for each MC simulated signal data set, and visualized
 2620 over the full signal grid. For expediency, only the statistical uncertainties associated
 2621 with the MC simulation of signal and background processes were considered when
 2622 evaluating the Asimov discovery significance and the sensitivity projections used for
 2623 optimization, and the systematic uncertainties presented in Chapter 6 were neglected.
 2624 The choice to neglect systematic uncertainties was justified by the fact that, following
 2625 initial efforts to include the dominant theoretical sources of systematic uncertainty in
 2626 evaluation of the Asimov discovery significance, their inclusion was found to have a
 2627 negligible impact on the evaluation of optimal cut placements.

2628 Figures 5.10 and 5.11 show N-1 plots with the finalized selections, as well as the
 2629 placements of cuts, for the $m^{\text{TAR Jet}}$ and $\Delta R(\text{TAR Jet}, \ell)$ ($m(W_{\text{Cand}})$ and $\Delta R(W_{\text{Cand}}, \ell)$)

Table 5.3: List of selection variables, with descriptions, for which the placements of cuts were optimized when designing the merged and resolved signal regions. The third and fourth columns indicate whether the variable is used in the merged category, the resolved category, or both.

Variable	Description	Mgd	Res
E_T^{miss}	A lower bound is placed on E_T^{miss} to select for the production of undetected energetic particles.	✓	✓
\mathcal{S}	A lower bound is placed on \mathcal{S} to select for a high likelihood that the measured E_T^{miss} arises from undetected particles rather than limited detector resolution.	✓	✓
$m_T(\ell, E_T^{\text{miss}})$	A lower bound is placed on $m_T(\ell, E_T^{\text{miss}})$ to select for a high likelihood of there being sources of E_T^{miss} in the final state in addition to the neutrino produced by a leptonic $W \rightarrow \ell\nu$ decay (See Section 5.1.8 for details).	✓	✓
$m^{\text{TAR Jet}}$	Reconstructed mass of the highest- p_T $R = 1.0$ TAR jet. A window cut around the W boson mass of 80.4 GeV is placed on $m^{\text{TAR Jet}}$ to select for events in which the highest- p_T $R = 1.0$ TAR jet actually reconstructs the hadronic decay of a boosted W boson, rather than other potential sources of strongly interacting particles.	✓	✗
$D_2^{\beta=1}(\text{TAR Jet})$	Energy correlation function of the highest- p_T $R = 1.0$ TAR jet. Used to discriminate large- R jets with a two-pronged substructure from those with a single-pronged substructure using the angular separation and transverse momenta of combinations of the jet constituents [155, 156]. An upper bound is placed on $D_2^{\beta=1}(\text{TAR Jet})$, because the $D_2^{\beta=1}(\text{TAR Jet})$ distribution for two-pronged signal events is found to be peaked at lower $D_2^{\beta=1}(\text{TAR Jet})$ compared with the SM background events.	✓	✗
$\Delta R(\text{TAR Jet}, \ell)$	Angular separation in $\eta \times \phi$ space between the TAR jet and lepton. An upper bound is placed on $\Delta R(\text{TAR Jet}, \ell)$ to select for the expected signal topology in which the TAR jet and the $\ell\nu$ form a collimated system, having originated from the decay of the boosted s .	✓	✗
$p_T(W_{\text{Cand}})$	p_T of the reconstructed W candidate in the resolved regime. A lower bound is placed on the $p_T(W_{\text{Cand}})$ to select for the signal topology in which the W is produced with a large momentum from the decay of a boosted s .	✗	✓
$m(W_{\text{Cand}})$	Mass of the reconstructed W candidate in the resolved regime. A window cut is placed around the W boson mass of 80.4 GeV.	✗	✓
$\Delta R(W_{\text{Cand}}, \ell)$	Angular separation in $\eta \times \phi$ space between the reconstructed W candidate and lepton in the resolved regime. An upper bound is placed on $\Delta R(W_{\text{Cand}}, \ell)$ to select for the expected signal topology in which the W candidate and the $\ell\nu$ form a collimated system, having originated from the decay of the boosted s .	✗	✓

variables in the merged (resolved) SR. N-1 plots for the other selection variables listed in Table 5.3 can be found in Appendix A. The finalized selections are summarized in Table 5.4 for the merged and resolved SRs, respectively. The lower bound on E_T^{miss} in the merged SR was ultimately kept at the same value of 200 GeV applied in the baseline selection, because it was found that with the other optimized selections applied there were very few signal and background events with $E_T^{\text{miss}} < 250$ GeV, and explicitly increasing the lower bound was not found to produce any appreciable improvements in sensitivity.

Table 5.4: Optimized selection criteria for the signal region in the merged and resolved categories.

Merged Selection	Resolved Selection
Passes baseline selection	
$N(\text{TAR Jets}) > 0$	Fails merged selection
$m_T(\ell, E_T^{\text{miss}}) > 220$ GeV	$N(\text{Jets}) > 1$
$E_T^{\text{miss}} > 200$ GeV	$m_T(\ell, E_T^{\text{miss}}) > 200$ GeV
$\mathcal{S} > 16$	$E_T^{\text{miss}} > 250$ GeV
$68 \text{ GeV} < m^{\text{TAR Jet}} < 89 \text{ GeV}$	$\mathcal{S} > 16$
$\Delta R(\text{TAR Jet}, \ell) < 1.2$	$68 \text{ GeV} < m(W_{\text{Cand}}) < 89 \text{ GeV}$
$D_2^{\beta=1}(\text{TAR Jet}) < 1.1$	$\Delta R(W_{\text{Cand}}, \ell) < 1.4$
	$p_T(W_{\text{Cand}}) > 150$ GeV

Tables 5.5 and 5.6 show the predicted yields of MC simulated events for the three dominant SM backgrounds after successive application of selections in the merged and resolved SRs respectively, beginning with the baseline selection. Tables 5.5 and 5.6 show the same information for the DH signal model at three sample mass points in $(m_s, m_{Z'})$.

5.2.4 Control regions

There are in general numerous uncertain theoretical parameters involved in modelling signal and SM background processes produced by pp collisions at the LHC, and it is necessary to fix their values when generating the MC simulated events and weights used to predict the yields of these processes in the ATLAS collision data. These include, for example, parameters associated with the parton distribution function used to model the protons involved in the high-energy collisions. A recent review of progress in the determination of parton distribution functions can be found in Ref.

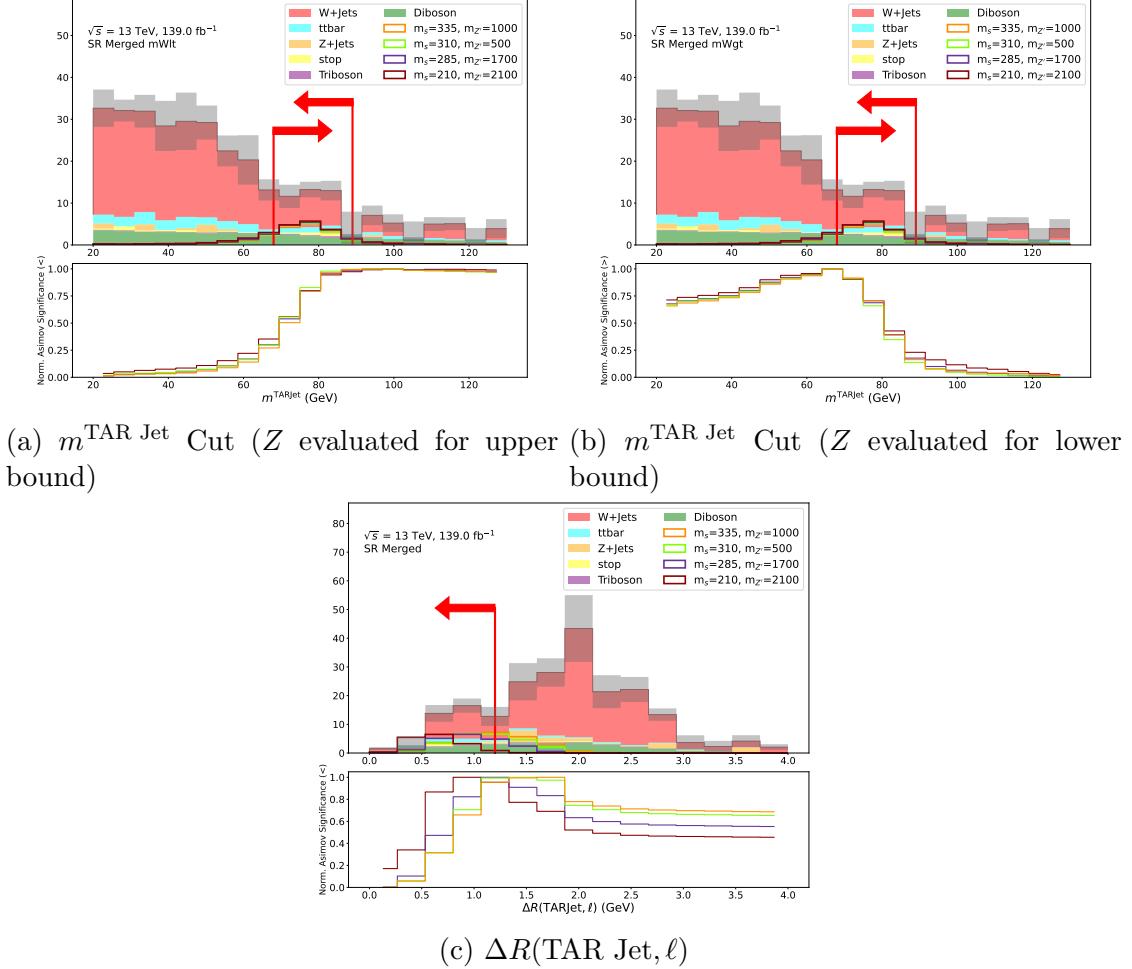


Figure 5.10: N-1 Distributions for the $m^{\text{TAR Jet}}$ and $\Delta R(\text{TAR Jet}, \ell)$ variables used in the merged signal region definition. Grey bands show statistical uncertainty on background estimate. The lower panel shows the cumulative Asimov significance normalized to unit peak, where the direction ($>$ or $<$) specified in the y label indicates whether the significance is being summed from above ($>$) or from below ($<$). Red vertical line and arrow show placement and direction of selection on the given variable in this region.

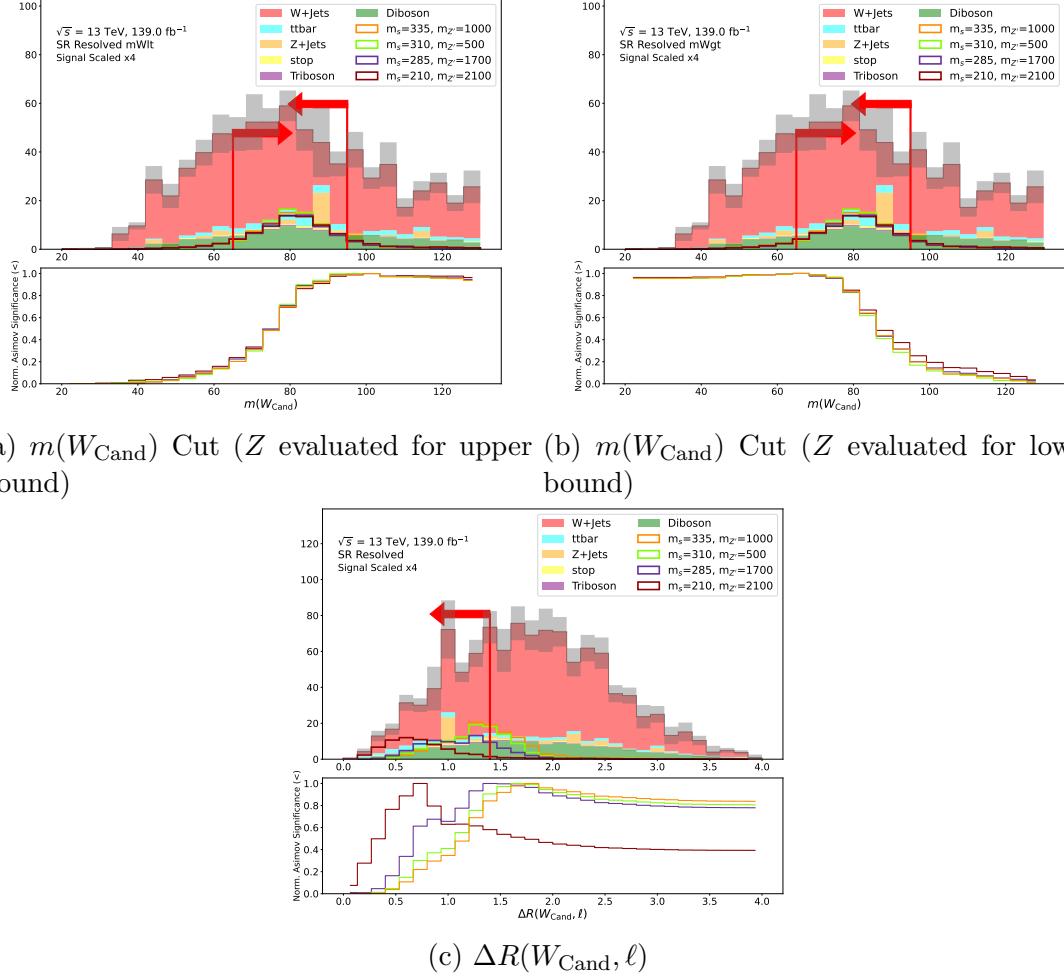


Figure 5.11: N-1 Distributions for the $m(W_{\text{Cand}})$ and $\Delta R(W_{\text{Cand}}, \ell)$ variables used in the resolved signal region definition. Grey bands show statistical uncertainty on the background estimate. The lower panel shows the cumulative Asimov significance normalized to unit peak, where the direction ($>$ or $<$) specified in the y label indicates whether the significance is being summed from above ($>$) or from below ($<$). Red vertical line and arrow show placement and direction of selection on the given variable in this region.

Table 5.5: Cutflow yields after application of preselection cuts for the dominant SM backgrounds in the merged SR. Percent passage is reported relative to the number of weighted events after application of the baseline selection.

Cut	W+jets	$t\bar{t}$	Diboson
None	11352445.72	101404110.82	1270378.37
baseline selection	93740.14 (100.00%)	51263.07 (100.00%)	8469.92 (100.00%)
$N(\text{TAR Jets}) \geq 1$	36957.80 (39.43%)	44993.63 (87.77%)	6887.00 (81.31%)
b -jet veto	34077.64 (36.35%)	5082.68 (9.91%)	6199.93 (73.20%)
$m_T(\ell, E_T^{\text{miss}}) > 220 \text{ GeV}$	21927.98 (23.39%)	2994.69 (5.84%)	3450.61 (40.74%)
$m^{\text{TAR Jet}} \in (68, 89) \text{ GeV}$	942.31 (1.01%)	165.13 (0.32%)	206.63 (2.44%)
$\mathcal{S} > 16$	437.27 (0.47%)	21.52 (0.04%)	85.00 (1.00%)
$\Delta R(\text{TAR Jet}, \ell) < 1.2$	65.79 (0.07%)	10.30 (0.02%)	21.46 (0.25%)
$D_2^{\beta=1}(\text{TAR Jet}) < 1.1$	24.29 (0.03%)	4.29 (0.01%)	7.61 (0.09%)

[2651] [157], and recommendations for the evaluation of uncertainties associated with parton
[2652] distribution function parameters in the context of studies performed with LHC data
[2653] can be found in Ref. [158].

[2654] In some cases, the range of plausible choices for an uncertain parameter, or a set
[2655] of related uncertain parameters, can produce an associated uncertain range of yield
[2656] predictions that is non-negligible compared with the statistical yield uncertainties.
[2657] The resulting uncertainties in predicted yields in the SR can be reduced by constraining
[2658] the predicted yields using the ATLAS collision data in a kinematically similar
[2659] region.

[2660] $W + \text{jets}$ and $t\bar{t}$ control regions (CRs), with enriched yields of these respective SM
[2661] background processes, are defined in this search with the aim of providing data-driven
[2662] constraints on the total yield of the $W + \text{jets}$ and $t\bar{t}$ backgrounds in the SRs within each
[2663] of the merged and resolved kinematic categories. As discussed in detail in Section 7.3,
[2664] the data-driven constraints are obtained within each category by comparing the total
[2665] yield of MC simulated SM background events to the observed yield of data in the
[2666] CRs for the given category, and scaling the total predicted yield of the $W + \text{jets}$ and $t\bar{t}$
[2667] backgrounds in both the CRs and the SRs by multiplicative ‘‘normalization factors’’,
[2668] $\mu_{W+\text{jets}, \text{category}}$ and $\mu_{t\bar{t}, \text{category}}$, such that the total predicted yield of SM background
[2669] processes within each control region is in close agreement with the observed yield of

Table 5.6: Cutflow yields after application of preselection cuts for the dominant SM backgrounds in the resolved SR. Percent passage is reported relative to the number of weighted events after application of the baseline selection.

Cut	W+jets	t <bar>t</bar>	Diboson
None	11352445.72	101404110.82	1270378.37
baseline selection	93740.14 (100.00%)	51263.07 (100.00%)	8469.92 (100.00%)
Fails merged selection	93552.83 (99.80%)	51178.42 (99.83%)	8435.36 (99.59%)
$N(\text{Jets}) \geq 2$	41566.23 (44.34%)	49150.32 (95.88%)	6624.88 (78.22%)
b -jet veto	38106.85 (40.65%)	5696.57 (11.11%)	5888.39 (69.52%)
$m_T(\ell, E_T^{\text{miss}}) > 200$ GeV	30794.06 (32.85%)	4053.26 (7.91%)	4045.52 (47.76%)
$E_T^{\text{miss}} > 250$ GeV	14965.69 (15.97%)	1639.75 (3.20%)	2366.03 (27.93%)
$m(W_{\text{Cand}}) \in (65, 95)$ GeV	5214.91 (5.56%)	816.93 (1.59%)	941.15 (11.11%)
$\mathcal{S} > 16$	3041.82 (3.24%)	166.35 (0.32%)	429.93 (5.08%)
$\Delta R(W_{\text{Cand}}, \ell) < 1.4$	771.85 (0.82%)	58.37 (0.11%)	114.24 (1.35%)
$p_T(W_{\text{Cand}}) > 150$ GeV	242.64 (0.26%)	21.19 (0.04%)	51.12 (0.60%)

2670 ATLAS collision data.

2671 The selections for the $W+\text{jets}$ and $t\bar{t}$ CRs were designed with the following aims in
2672 order to provide effective and reliable constraints on the total yield of the respective
2673 SM background processes in the SR:

- 2674 • To obtain a **high purity** of the background process of interest in the CR, as
2675 evaluated by its MC simulated yield in the CR relative to other MC simulated
2676 background processes. It is important that the predicted event yield in the CR
2677 be dominated by the background process of interest, to ensure that the ratio of
2678 total MC simulated yield from all SM background processes, $\sum_{\text{process } i} (N_{\text{MC}})_{i, \text{CR}}$
2679 to the observed yield of ATLAS collision events $(N_{\text{data}})_{\text{observed, CR}}$ in the CR
2680 represents a reasonable approximation of the equivalent ratio for the background
2681 process $p_{\text{constraint}}$ of interest:

$$\frac{\sum_{\text{process } i} (N_{\text{MC}})_{i, \text{CR}}}{(N_{\text{data}})_{\text{total observed, CR}}} \approx \frac{(N_{\text{MC}})_{p_{\text{constraint}}, \text{CR}}}{(N_{\text{data}})_{p_{\text{constraint}}, \text{CR}}} \quad (5.17)$$

2682 where the index i runs over all simulated SM background processes.

Table 5.7: Cutflow yields after application of preselection cuts for three sample signal points in the merged SR. Signal points are labelled in column headers as $(m_{Z'}, m_s)$ (units of GeV). Percent passage is reported relative to the number of weighted events after application of the baseline selection.

Cut	(1000, 360)	(1700, 335)	(2100, 210)
None	798.61	396.57	537.03
baseline selection	168.35 (100.00%)	111.66 (100.00%)	143.64 (100.00%)
$N(\text{TAR Jets}) \geq 1$	147.88 (87.84%)	98.12 (87.87%)	124.98 (87.01%)
b -jet veto	131.89 (78.34%)	87.51 (78.37%)	111.32 (77.50%)
$m_T(\ell, E_T^{\text{miss}}) > 220 \text{ GeV}$	102.83 (61.08%)	70.82 (63.42%)	92.86 (64.65%)
$m^{\text{TAR Jet}} \in (68, 89) \text{ GeV}$	45.00 (26.73%)	30.85 (27.63%)	31.88 (22.19%)
$\mathcal{S} > 16$	33.84 (20.10%)	23.97 (21.47%)	24.34 (16.95%)
$\Delta R(\text{TAR Jet}, \ell) < 1.2$	15.80 (9.39%)	13.33 (11.94%)	23.20 (16.15%)
$D_2^{\beta=1}(\text{TAR Jet}) < 1.1$	11.10 (6.59%)	9.50 (8.51%)	16.19 (11.27%)

- The CR should contain a **relatively large number of MC simulated events** (and consequently a lower relative statistical uncertainty) for the background process of interest compared with the SR. This requirement is designed to ensure that the normalization factors $\mu_{W+\text{jets,category}}$ and $\mu_{t\bar{t},\text{category}}$ will be constrained predominantly in the overall fit of MC simulated yields to the observed data (see Section 7.3 for details on the fitting strategy) by the comparison with data in the CRs, and also to minimize the statistical uncertainties of the fitted normalization factors associated with the limited number of MC simulated events in the CRs.
- To ensure **orthogonality with the SR**. In order for the CR to provide an unbiased data-driven constraint on the amplitude of a simulated SM background used in the signal region, it must be defined in such a way that no events are shared between these two regions.
- To obtain a **negligibly small contamination of MC simulated signal events**. Negligible signal contamination is needed to ensure that the constraints on normalization factors are obtained purely from known physics processes in the data.

Table 5.8: Cutflow yields after application of preselection cuts for three sample signal points in the resolved SR. Signal points are labelled in column headers as $(m_{Z'}, m_s)$ (units of GeV). Percent passage is reported relative to the number of weighted events after application of the baseline selection.

Cut	(1000, 360)	(1700, 335)	(2100, 210)
None	798.61	396.57	537.03
baseline selection	168.35 (100.00%)	111.66 (100.00%)	143.64 (100.00%)
Fails merged selection	152.86 (90.80%)	100.58 (90.08%)	127.22 (88.57%)
$N(\text{Jets}) \geq 2$	138.08 (82.02%)	90.56 (81.10%)	110.52 (76.94%)
b -jet veto	121.13 (71.95%)	79.35 (71.06%)	96.14 (66.93%)
$m_T(\ell, E_T^{\text{miss}}) > 200 \text{ GeV}$	101.11 (60.06%)	68.49 (61.34%)	84.25 (58.65%)
$E_T^{\text{miss}} > 250 \text{ GeV}$	78.78 (46.80%)	55.43 (49.64%)	67.87 (47.25%)
$m(W_{\text{Cand}}) \in (65, 95) \text{ GeV}$	55.87 (33.19%)	39.09 (35.01%)	42.40 (29.52%)
$\mathcal{S} > 16$	40.37 (23.98%)	28.77 (25.77%)	29.22 (20.34%)
$\Delta R(W_{\text{Cand}}, \ell) < 1.4$	16.84 (10.00%)	14.51 (12.99%)	23.77 (16.55%)
$p_T(W_{\text{Cand}}) > 150 \text{ GeV}$	13.35 (7.93%)	11.35 (10.16%)	15.87 (11.05%)

- 2700 • Events in the CR should occupy a **kinematic phase space that is as similar**
 2701 **as possible to the SR**. This is important to ensure that the ratio of the
 2702 MC simulated yield of SM background events to the observed yield of ATLAS
 2703 collision data in the CR is representative of the equivalent ratio in the SR.

2704 Table 5.9 summarizes the CRs used in the analysis and the cut changes relative
 2705 to the SR that are used to define each CR. The CRs are discussed in more detail in
 2706 Section 5.2.4.

2707 W+jets Control Region Definition

2708 A W+jets CR (a.k.a. CRW) is defined to obtain a high purity and yield of the W+jets
 2709 background process by reversing the cut on $\Delta R(W, \ell)$. Orthogonality with the SR is
 2710 ensured by the reversal of the $\Delta R(W, \ell)$ cut. The lower bound on E_T^{miss} significance
 2711 is also reduced to 12 in the merged category of the CR in order to reduce the relative
 2712 statistical uncertainty by boosting the number of MC simulated events in the region.
 2713 In addition, after some study the cut on $\Delta R(W, \ell)$ in the merged CRW was tightened

Table 5.9: Summary of control regions.

Control region	Modified selections relative to the SR
Merged W+jets CR	$(\Delta R(\text{TAR Jet}, \ell) < 1.2) \rightarrow (\Delta R(\text{TAR Jet}, \ell) > 1.8)$ $(\mathcal{S} > 16) \rightarrow (\mathcal{S} > 12)$
Resolved W+jets CR	$(\Delta R(W_{\text{Cand}}, \ell) < 1.4) \rightarrow (\Delta R(W_{\text{Cand}}, \ell) > 1.4)$
Merged $t\bar{t}$ CR	$(N(\text{b-tagged jets}) < 1) \rightarrow N(\text{b-tagged jets}) \geq 2$ $(\mathcal{S} > 16) \rightarrow (\mathcal{S} > 12)$
Resolved $t\bar{t}$ CR	$(N(\text{b-tagged jets}) < 1) \rightarrow N(\text{b-tagged jets}) \geq 2$

from a simple reversal - ($\Delta R(\text{TAR Jet}, \ell) < 1.2$) \rightarrow ($\Delta R(\text{TAR Jet}, \ell) > 1.2$) - to a slightly increased lower bound of ($\Delta R(\text{TAR Jet}, \ell) > 1.8$) in order to reduce the predicted yield of signal processes in the CR to an acceptable level.

The motivation for reversing the ΔR selection is that, as shown schematically in Figure 5.12, a ΔR reversal would largely reverse the directions of the lepton and neutrino in the $W + \text{jets}$ process described in Section 4.4.1, without modifying the kinematic details of the hadronic activity that fakes a hadronically-decaying W candidate for this process in the SR. As a result, the modelling of $W + \text{jets}$ events in this ΔR -reversed region would be expected to be similar to $W + \text{jets}$ events in the SR, with the exception of the reversed lepton and neutrino directions in the $W \rightarrow \ell\nu$ decay.

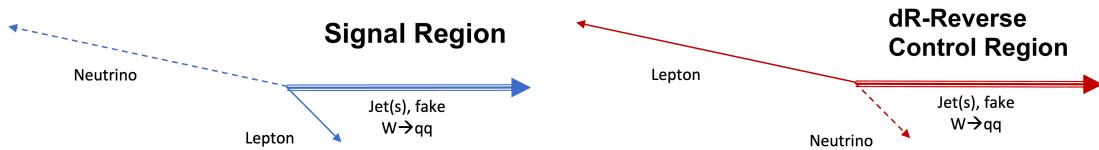


Figure 5.12: Comparison of typical event topologies between the SR and the $W + \text{jets}$ CR, showing that the typical directions of the lepton and neutrino relative to hadronic activity faking the hadronically-decaying W boson are simply reversed.

Kinematic distributions of interest are compared between the $W + \text{jets}$ CRs and the SRs in Figures B.1 and B.2 in Appendix B.1.

A comparison of the E_T^{miss} distributions between the merged SR and CRW in Figures B.1i and B.1j indicates some bias towards lower E_T^{miss} for events that pass the merged CRW selection. This difference is attributable to two factors. First, because the \mathcal{S} is in general expected to be positively correlated with E_T^{miss} , the loosened \mathcal{S} cut of $\mathcal{S} > 12$ in the merged CRW compared with $\mathcal{S} > 16$ in the SR would be expected to allow a larger proportion of low- E_T^{miss} events into the region. The impact of the loosened \mathcal{S} selection in the merged CRW is shown in Figure 5.13 by comparing the E_T^{miss} distribution between the merged SR and the merged CRW, either with or without the loosened \mathcal{S} selection in the merged CRW. Tightening the \mathcal{S} lower bound is seen to remove most of the low- E_T^{miss} bias in the merged CRW. The small remaining bias towards low E_T^{miss} in the merged CRW even after tightening the \mathcal{S} selection is attributed to the expected reversal of the neutrino direction in the ΔR -reversed topology of $W + \text{jets}$ background in the CRW, such that it becomes approximately aligned with the high-pt TAR jet rather than recoiling against both the TAR jet and the lepton. The neutrino p_T required to conserve momentum in the ΔR -reversed

topology will on average be somewhat smaller due to the approximately opposing momenta of the lepton and the TAR jet. Since the neutrino constitutes the only source of true E_T^{miss} in the $W+\text{jets}$ background, this implies that the E_T^{miss} distribution would be expected to become somewhat biased to lower values in the CRW topology, as observed.

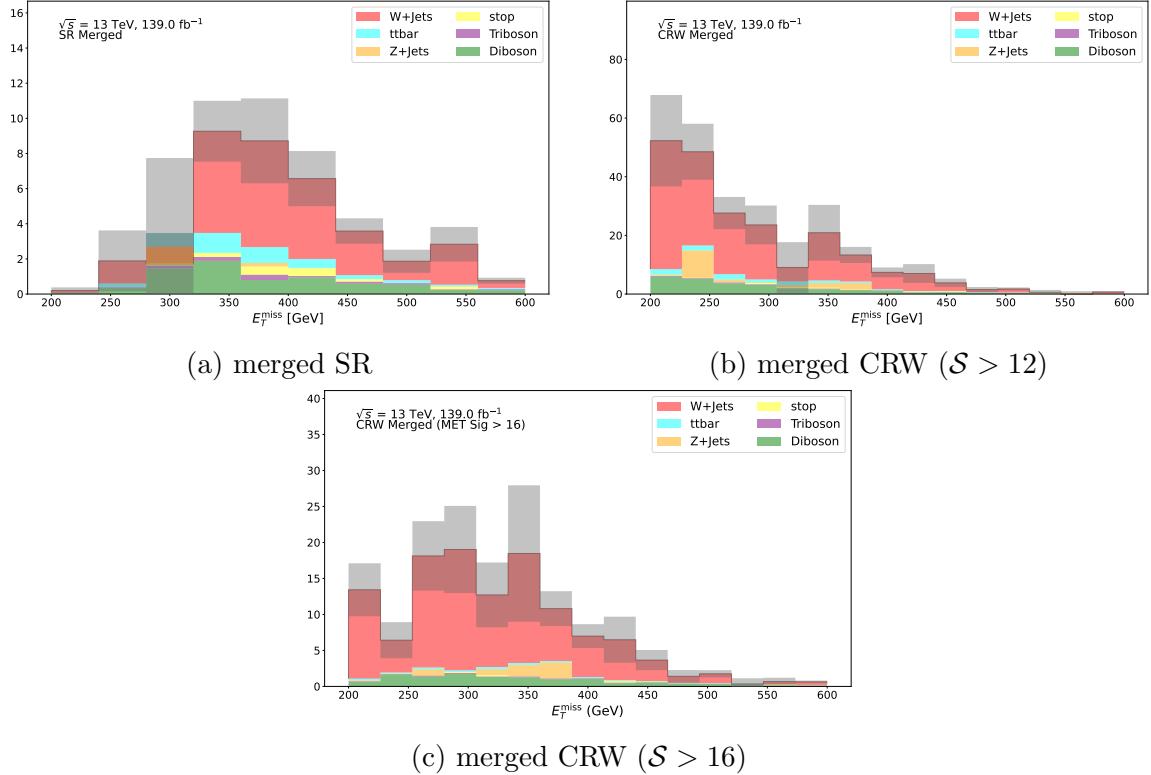


Figure 5.13: Comparison of N-1 distributions between the merged SR (top left) and the merged CRW, with the lower bound on \mathcal{S} in the merged CRW either kept at its nominal value of 12 (upper right), or tightened (bottom) to match the $\mathcal{S} > 16$ cut applied in the merged SR.

A comparison of Figures B.1i and B.1j shows that, though present, this bias in the E_T^{miss} distribution is much more subtle in the resolved SR. This is attributed to the fact that the lower bound on \mathcal{S} is not loosened compared with the SR in the resolved category, in addition to the relatively low p_T of the hadronic activity that the neutrino is approximately aligned with in the CRW.

Based on the comparable shapes of the other kinematic distributions in the merged and resolved categories, it is concluded that the kinematics of events in the $W+\text{jets}$ CRs are sufficiently similar to those in the SR of the corresponding category that the

2754 constraints on normalization of the W +jets background evaluated in the W +jets CRs
 2755 can be reasonably applied in the SR.

2756 Table 5.10 compares the yield and relative composition of the W +jets background
 2757 in the SRs with the W +jets CRs. Figures 5.14 and 5.15 compare the signal yield
 2758 and signal/background between the SR and the W +jets CR in the merged and
 2759 resolved categories, respectively. The W +jets background constitutes over 75% of
 2760 the predicted yield of SM background processes in both CRs, making it the dominant
 2761 SM background process, with signal contamination below 2.5% for all signal points.
 2762 Thanks to the boost in the number of MC simulated events in the CRs afforded by the
 2763 reversal of the ΔR cut, as well as the reduced lower bound on S in the merged CR, the
 2764 relative statistical uncertainties of the predicted W +jets yields are reduced by factors
 2765 of 2 and 1.5 in the CRW compared with the SR in the merged and resolved categories,
 2766 respectively. This ensures that the normalization of the W +jets background process
 2767 can be predominantly established in the W +jets CR with a reasonably low statistical
 2768 uncertainty.

2769 The benchmark used to evaluate the maximum acceptable level of signal con-
 2770 tamination in a given CR is as follows: if the predicted yield of the signal process
 2771 is negligibly small (i.e. less than \sim half as large) compared with the statistical un-
 2772 certainty associated with the predicted yield of all SM background processes in the
 2773 CR, then the level of signal contamination is considered acceptably small, since the
 2774 presence or absence of the signal process would not be detectable given the statistical
 2775 uncertainty of the total predicted yield. In the merged W +jets CR, the predicted
 2776 signal yield is at most 5.1 (from Figure 5.14a, which is negligibly small compared with
 2777 the total background yield uncertainty of 24.5 (from the third row in Table 5.10). The
 2778 signal yields in the resolved CRW are similarly small compared with the statistical
 2779 uncertainty of the total background yield.

2780 $t\bar{t}$ Control Region Definition

2781 Given that b -tagged jets are vetoed in the SR to reduce the yield of events produced
 2782 by the $t\bar{t}$ process, a reversal of this veto presents a straightforward opportunity to
 2783 define an orthogonal CR enriched with $t\bar{t}$ events. Early studies found that while a
 2784 simple reversal of the b -jet veto in the SR selection to instead require at least one
 2785 b -tagged jet was sufficient to obtain a $t\bar{t}$ -enriched region with a substantially larger
 2786 sample of simulated $t\bar{t}$ events compared with the SR, the contamination of signal

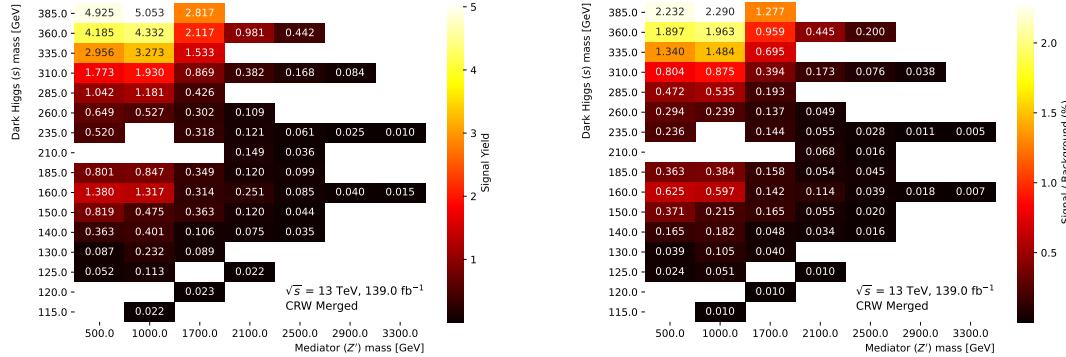


Figure 5.14: Predicted yields of MC simulated events (left), and ratio of predicted signal / SM background yields (right) for all signal points in the merged $W + \text{jets}$ control region.

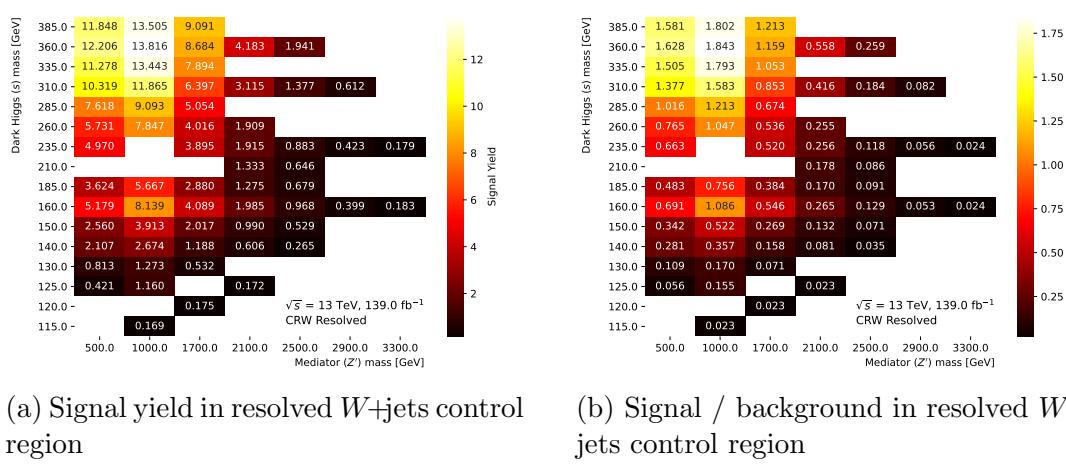


Figure 5.15: Predicted yields of MC simulated events (left), and ratio of predicted signal / SM background yields (right) for all signal points in the resolved $W + \text{jets}$ control region.

Table 5.10: Comparison of the $W + \text{jets}$ background yield, total background yield, and the composition of the $W + \text{jets}$ background relative to the total background. Uncertainties on the relative composition are obtained from the sum of squared event weights.

Region	W+jets Yield	Total Background Yield	Relative W+jets Composition
Merged SR	24.3 ± 7.2	40.1 ± 7.3	$(60.6 \pm 21.1)\%$
Resolved SR	243.5 ± 19.6	339.9 ± 24.0	$(71.6 \pm 7.7)\%$
Merged CRW	164.0 ± 23.6	220.6 ± 24.5	$(74.3 \pm 13.5)\%$
Resolved CRW	598.1 ± 30.4	749.5 ± 31.0	$(79.8 \pm 5.2)\%$

2787 events in the region was found to be too high for a control region.

2788 Further tightening the b -jet veto reversal to require at least two b -tagged jets
 2789 was found to reduce the signal contamination in this $t\bar{t}$ CR to a reasonable level on
 2790 the basis of the benchmark requirement of negligible signal yield compared with the
 2791 statistical uncertainty of the total background yield discussed above in the context
 2792 of the $W + \text{jets}$ CR. However, due to the associated reduction in the number of MC
 2793 simulated $t\bar{t}$ events that pass this tightened requirement, it was deemed necessary to
 2794 reduce the lower bound on \mathcal{S} in the merged $t\bar{t}$ CR to $\mathcal{S} > 12$, as is done in the merged
 2795 $W + \text{jets}$ CR, in order to increase the number of simulated events admitted for the $t\bar{t}$
 2796 process.

2797 Table 5.11 summarizes the modifications made to the SR selections to define the
 2798 $t\bar{t}$ CR in both the merged and resolved categories.

2799 Table 5.12 compares the yield and relative composition of the $t\bar{t}$ background in
 2800 the SRs with those in the $t\bar{t}$ CRs. Figures 5.16 and 5.17 compare the signal yield
 2801 and signal/background between the SR and the $t\bar{t}$ CR in the merged and resolved
 2802 categories, respectively. As expected, the reversal of the $t\bar{t}$ veto produces a region
 2803 that is highly enriched in $t\bar{t}$ events, which constitute 90% of the predicted yield in
 2804 the $t\bar{t}$ CR. The merged and resolved $t\bar{t}$ CRs admit comparable yields of $\sim 65 - 70$ $t\bar{t}$
 2805 events, which in both categories constitutes a several-fold increase in the predicted
 2806 yield compared with the SR. Comparing the predicted signal yields in the merged and
 2807 resolved CRTT, shown in Figures 5.16a and 5.17a respectively, with the statistical
 2808 uncertainties associated with the total yield of events in these CRTTs from Table 5.11,
 2809 the predicted signal yields are in all cases well below the statistical uncertainty of the
 2810 background yield, and thus constitute a reasonably low level of signal contamination
 2811 in the CRTT.

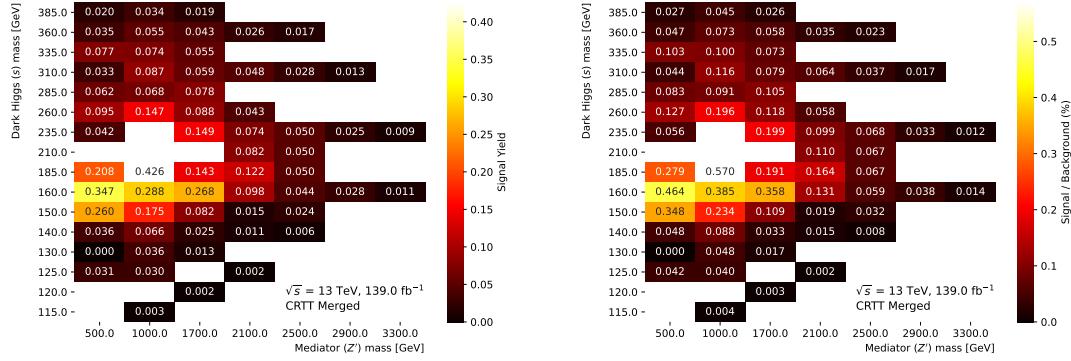


Figure 5.16: Predicted yields of MC simulated events (left), and ratio of predicted signal / SM background yields (right) for all signal points in the merged $t\bar{t}$ control region.

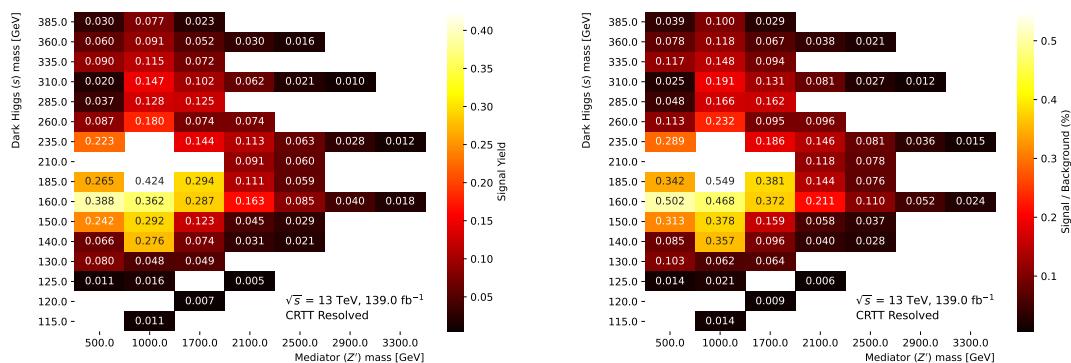


Figure 5.17: Predicted yields of MC simulated events (left), and ratio of predicted signal / SM background yields (right) for all signal points in the resolved $t\bar{t}$ control region.

Table 5.11: Summary of differences in selections on $N(b\text{-jet})$ and \mathcal{S} between signal regions and $t\bar{t}$ control regions.

Region	b -tagged Jet Veto/Requirement	\mathcal{S} Cut
Merged SR	$N(b\text{-jet}) = 0$	$\mathcal{S} > 16$
Resolved SR	$N(b\text{-jet}) = 0$	$\mathcal{S} > 16$
Merged CRTT	$N(b\text{-jet}) > 1$	$\mathcal{S} > 12$
Resolved CRTT	$N(b\text{-jet}) > 1$	$\mathcal{S} > 16$

Table 5.12: Comparison of the $t\bar{t}$ background yield, total background yield, and the composition of the $t\bar{t}$ background relative to the total background. Uncertainties on the relative composition are obtained from the sum of squared event weights.

Region	$t\bar{t}$ Yield	Total Background Yield	$t\bar{t}$ Relative Composition
Merged SR	4.3 ± 0.3	40.1 ± 7.3	$(10.7 \pm 2.1)\%$
Resolved SR	21.2 ± 0.7	339.9 ± 24.0	$(6.2 \pm 0.5)\%$
Merged CRTT	70.5 ± 1.7	74.8 ± 5.0	$(94.2 \pm 6.7)\%$
Resolved CRTT	64.2 ± 1.2	73.3 ± 4.9	$(83.1 \pm 3.4)\%$

2812 Kinematic distributions of interest are compared between the $t\bar{t}$ CRs and the SRs
 2813 in Figures B.3 and B.4 in Appendix B.2. A similar bias towards lower E_T^{miss} that was
 2814 discussed in the context of the merged $W+\text{jets}$ CR is also seen in the merged CRTT,
 2815 and is similarly attributable to the loosened lower bound on \mathcal{S} in this region. The
 2816 level of shape agreement in the other distributions is considered to be close enough
 2817 to conclude that events in the $t\bar{t}$ CR occupy a sufficiently similar region of kinematic
 2818 phase space as events in the SR to justify applying data-driven $t\bar{t}$ normalization
 2819 constraints obtained in this CR to the SR.

2820 5.2.5 Background Yields

2821 Tables 5.13 and 5.14 show the overall background yields in the merged and resolved
 2822 analysis regions, respectively, after application of all the analysis selections described
 2823 in this section.

Table 5.13: Background yields after application of all analysis cuts in the merged analysis regions. Uncertainty on the yields is statistical.

Background	Merged SR	Merged CRW	Merged CRTT
W+jets	24.2 ± 7.2	164.0 ± 23.6	-1.3 ± 4.6
t<bar>t</bar>	4.3 ± 0.3	9.9 ± 0.6	70.4 ± 1.6
Diboson	7.6 ± 0.3	26.8 ± 1.5	0.2 ± 0.0
Triboson	1.2 ± 0.1	1.9 ± 0.1	0.0 ± 0.0
Z+jets	1.1 ± 0.9	15.2 ± 6.3	0.0 ± 0.0
single t	1.6 ± 0.5	2.7 ± 0.6	5.4 ± 0.9
Total	40.1 ± 7.3	220.6 ± 24.5	74.8 ± 5.0

Table 5.14: Background yields after application of all analysis cuts in the resolved analysis regions. Uncertainty on the yields is statistical.

Background	Resolved SR	Resolved CRW	Resolved CRTT
W+jets	243.5 ± 19.6	598.1 ± 30.4	1.5 ± 2.2
t<bar>t</bar>	21.2 ± 0.7	20.0 ± 0.7	64.2 ± 1.2
Diboson	51.1 ± 1.1	105.5 ± 1.5	0.6 ± 0.0
Triboson	3.0 ± 0.2	5.6 ± 0.3	0.0 ± 0.0
Z+jets	17.4 ± 13.7	17.1 ± 6.0	0.0 ± 0.0
single t	3.7 ± 0.8	3.2 ± 0.7	11.0 ± 1.2
Total	339.9 ± 24.0	749.5 ± 31.0	77.3 ± 2.8

2824 5.3 Triggers

2825 As discussed in Section 3.4.5, the ATLAS trigger system only saves collision events
 2826 that pass both the hardware-based level-1 (L1) trigger and the software-based high-
 2827 level trigger (HLT). The L1 trigger and the HLT are each comprised of numerous sets
 2828 of selection criteria, which are also referred to as triggers. Any collision event that
 2829 satisfies at least one of the triggers that comprise the L1 trigger is processed by the
 2830 HLT. Likewise, if the event satisfies any of the triggers that comprise the HLT, it will
 2831 be saved for later analysis.

2832 The search presented in this thesis is interested in events that produce a single

2833 energetic lepton due to the $s \rightarrow WW(q\bar{q}\ell\nu)$ decay, in addition to high E_T^{miss} due
 2834 both to the undetected boosted DM in the final state, and to the undetected ν
 2835 from the $W \rightarrow \ell\nu$ decay. It is important to determine the efficiency with which the
 2836 ATLAS trigger system accepts events in the region of phase space defined by the event
 2837 selections described in Section 5.2 above. This efficiency quantifies the probability
 2838 that an event that the triggers are designed to accept successfully passes the trigger
 2839 criteria and gets accepted. If the trigger efficiency is < 100% in any area of the phase
 2840 space considered in the analysis, it is in general necessary to apply scale factors to any
 2841 MC simulated events that fall into this phase space to account for the fact that some
 2842 of these events would have been rejected by the trigger during actual data-taking. It
 2843 is also then necessary to evaluate and propagate uncertainties associated with these
 2844 scale factors.

2845 To simplify the trigger efficiency analysis and determine whether any scale factors
 2846 may be needed, it is helpful to identify a minimal list of triggers that all events
 2847 considered in the analysis would be expected to pass. One of the event selection
 2848 criteria for the analysis, presented in Section 5.2, requires all events to have $E_T^{\text{miss}} >$
 2849 200 GeV. Since the ATLAS E_T^{miss} trigger, described in Refs. [159] and [160], is
 2850 designed to efficiently select events with $E_T^{\text{miss}} > 150$ GeV, it is reasonable to expect
 2851 events that pass the event selection criteria to have also passed the E_T^{miss} trigger
 2852 with a high efficiency. The specific E_T^{miss} triggers in the ATLAS trigger menu that
 2853 are considered in this study are chosen following ATLAS recommendations, and vary
 2854 between different data collection periods defined by ATLAS. The full list of E_T^{miss}
 2855 triggers used, along with the associated data collection period for each, is listed in
 2856 Table 5.15.

2857 The ATLAS trigger system also includes single-muon and single-electron triggers,
 2858 which are designed to pass events in which a single muon (electron) is reconstructed in
 2859 the final state and satisfies some minimum p_T requirement. Since all events considered
 2860 in the search are required to have a single lepton in the final state, events which pass
 2861 the event selection would also be expected to pass these charged lepton triggers with
 2862 high efficiency. The specific single muon and electron triggers considered in this
 2863 study, along with the ATLAS data-taking period(s) in which they were applied and
 2864 the minimum lepton p_T requirement associated with each trigger, are listed in Tables
 2865 5.16 and 5.17, respectively.

2866 The lepton triggers are known to be < 100% efficient, but the resulting scale
 2867 factors and associated systematic uncertainties are in general well calibrated by ded-

Table 5.15: Summary of E_T^{miss} triggers from the ATLAS trigger menu used for the search, along with the associated data collection period for each trigger.

Period	MET Trigger
2015	HLT_XE70_MHT
2016 (A-D3)	HLT_XE90_MHT_L1XE50
2016 (D4-F1)	HLT_XE100_MHT_L1XE50
2016 (F2-)	HLT_XE110_MHT_L1XE50
2017 (B-D5)	HLT_XE110_PUFIT_L1XE55
2017 (D6-K)	HLT_XE110_PUFIT_L1XE50
2018 (B-C5)	HLT_XE110_PUFIT_XE70_L1XE50
2018 (C5-)	HLT_XE110_PUFIT_XE65_L1XE50

icated measurements performed within the ATLAS collaboration. As a result, the charged lepton triggers are useful as a means of independently quantifying the efficiency of the E_T^{miss} trigger, as will be shown in a moment. However, if the E_T^{miss} trigger can be shown to pass the events considered in this search with 100% efficiency then there is no need to apply scaling factors or evaluate related uncertainties.

The efficiency of the E_T^{miss} trigger for a set of event selection criteria that define a given region “X” is defined equivalently for ATLAS data (“data”) and MC simulated events (“MC”). Events considered for the calculation of trigger efficiency are also required to have passed the single lepton trigger (defined as the logical OR of the single muon trigger and the single electron trigger), to independently ensure that all data events considered passed a trigger that is relevant to the final state of interest.

The trigger efficiency is given by:

$$\text{eff}_{E_T^{\text{miss}}, \text{ region X}} = \frac{\sum_i w_i \text{ passing } (E_T^{\text{miss}} \text{ triggers}) \& (\text{single lepton triggers}) \& (\text{selection cuts for region X})}{\sum_i w_i \text{ passing } (\text{single lepton triggers}) \text{ AND } (\text{selection cuts for region X})} \quad (5.18)$$

where w_i is the total event weight for event i ($w_i = 1$ in the case of data). See Section 4.2.2 for a detailed discussion of weights that are assigned to the MC simulated events. Correction scale factors, dependent on the p_T and η of the final-state lepton in each event, are included in the MC event weights in Eq. 5.18 to account for the < 100% trigger efficiency of the single lepton triggers.

Table 5.16: Summary of single muon triggers from the ATLAS trigger menu used for the search, along with the associated data collection period for each trigger. The minimum muon p_T threshold of each trigger is also listed.

Periods	Single Muon Trigger	Muon p_T threshold
2015	HLT_MU20_ILOOSE_L1MU15	20 GeV
2016 (A, B-D3, D4-E, F-G2, G3-I3, I4-), 2017 (B-), 2018	HLT_MU50	50 GeV
2016	HLT_MU24_ILOOSE	24 GeV
2015, 2016 (A)	HLT_MU40	40 GeV
2016 (B-D3, D4-E)	HLT_MU24_IVARMEDIUM	24 GeV
2016 (D4-E, F-G2, G3-I3, I4-), 2017 (B-), 2018	HLT_MU26_IVARMEDIUM	26 GeV

2885 Figure 5.18 compares the E_T^{miss} trigger efficiency defined in Eq. 5.18 for MC
 2886 simulated events and ATLAS data for the region defined with the baseline selections,
 2887 with the following modifications:

- 2888 • A range of lower bounds on the E_T^{miss} are considered, from ~ 100 GeV to
 2889 ~ 500 GeV.
- 2890 • The single charged lepton is required to be an electron (a.k.a. the “electron
 2891 channel”) in Figure 5.18a and a muon (a.k.a. the “muon channel”) in Figure
 2892 5.18b.

2893 Comparing the trigger efficiencies in the electron channel (Figure 5.18a) and the
 2894 muon channel (Figures 5.18b and 5.18c), the efficiency in the electron channel con-
 2895 verges to 100% for $E_T^{\text{miss}} > 200$ GeV, but in the muon channel it instead converges to
 2896 $\sim 97\%$ for $E_T^{\text{miss}} > 200$ GeV. After some investigation, the inefficiency in the muon
 2897 channel was found to be due to events with large E_T^{miss} arising from high- p_T muons.
 2898 This is because high- p_T muons leave very little energy in the ATLAS calorimeter
 2899 and are detected instead by the muon spectrometer. As a result, such high- p_T muon
 2900 events can be missed by the E_T^{miss} triggers, which don’t use information from the muon
 2901 spectrometer (see Section 3 of Ref. [160] for details of the construction of E_T^{miss} for the

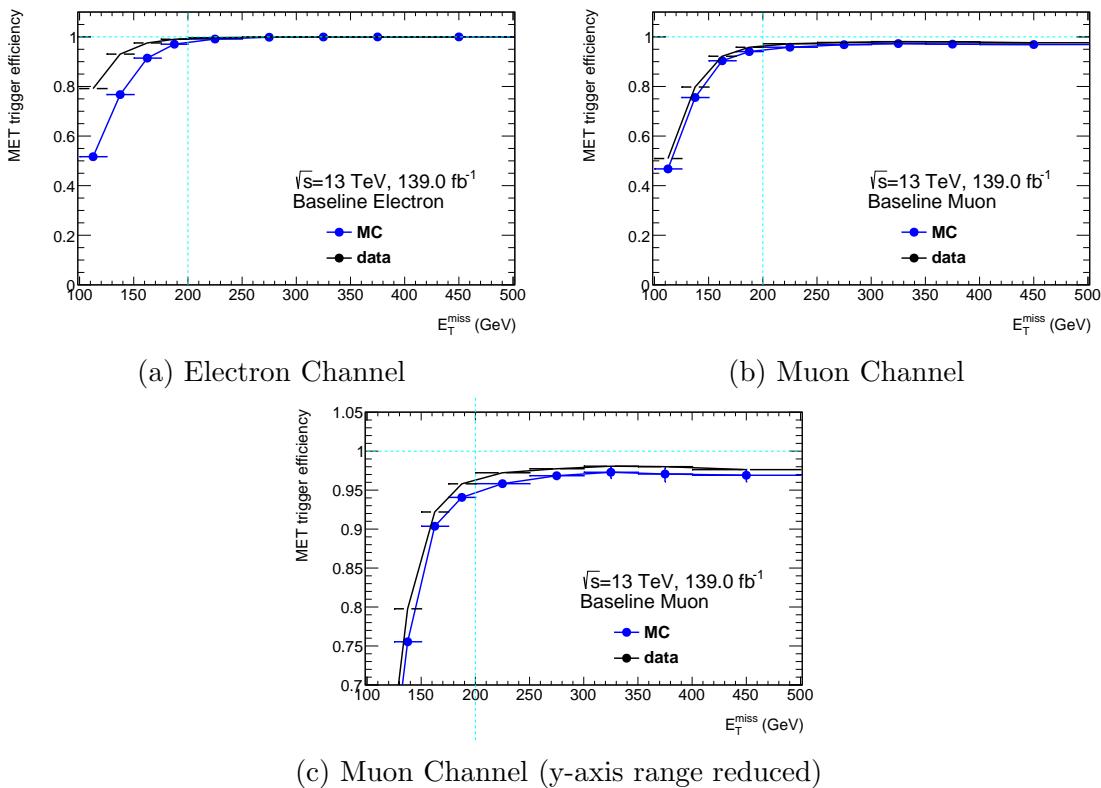


Figure 5.18: Comparison of the E_T^{miss} trigger efficiency defined in Eq. 5.18, as a function of the E_T^{miss} lower bound in the event selection, between MC simulated events and ATLAS data in a region defined by the baseline selection. The event selection is separated into electron (top left) and muon (top right and bottom center) channels.

Table 5.17: Summary of single electron triggers from the ATLAS trigger menu used for the study presented in Section 5.3, along with the associated data collection period for each trigger. The minimum electron p_T threshold of each trigger is also listed.

Periods	Single Muon Trigger	Electron p_T threshold
2015	HLT_E24_LHMEDIUM_L1EM20VH	24 GeV
2015	HLT_E60_LHMEDIUM	60 GeV
2015	HLT_E120_LHLOOSE	120 GeV
2016 (A, B-D3)	HLT_E24_LHTIGHT_NOD0_IVARLOOSE	24 GeV
2016 (A, B-D3, D4-F, G-), 2017 (B-), 2018	HLT_E60_LHMEDIUM_NOD0	60 GeV
2016 (A, B-D3, D4-F, G-)	HLT_E60_MEDIUM	60 GeV
2016 (A, B-D3, D4-F, G-), 2017 (B-), 2018	HLT_E300_ETCUT	300 GeV
2016 (A, B-D3, D4-F, G-), 2017 (B-), 2018	HLT_E140_LHLOOSE_NOD0	140 GeV
2016 (D4-F, G-), 2017 (B-) 2018	HLT_E26_LHTIGHT_NOD0_IVARLOOSE	26 GeV

2902 ATLAS E_T^{miss} trigger). This can be seen by plotting the E_T^{miss} trigger efficiency using
 2903 a calculation of E_T^{miss} in the event selection that ignores the muon p_T (a.k.a. “muon
 2904 invisible”) in Figure 5.19, and observing that in this case the efficiency converges to
 2905 100% for E_T^{miss} (muon invisible) > 200 GeV.

2906 It was found that high- E_T^{miss} events in the muon channel that fail the E_T^{miss} trigger
 2907 do, however, pass the muon trigger with high efficiency. For this reason, the efficiency
 2908 of a logical OR of the E_T^{miss} and single muon triggers is studied in the muon channel.
 2909 The (E_T^{miss} OR single muon) trigger efficiency is calculated as follows for MC simulated
 2910 events in a given region “X”:

$$\text{eff}_{E_T^{\text{miss}} \text{ OR single muon, MC, region X}} = \frac{\sum_i w_i \text{ passing } (E_T^{\text{miss}} \text{ OR single muon triggers}) \text{ AND } (\text{in region X})}{\sum_i w_i \text{ in region X}} \quad (5.19)$$

2911 where the event weight w_i in the numerator includes the scale factors to correct for

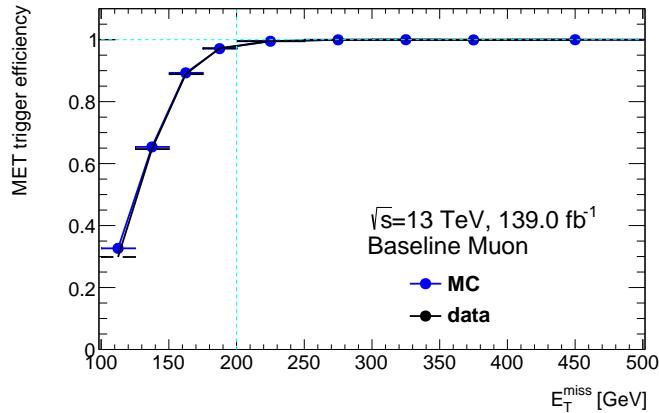


Figure 5.19: E_T^{miss} trigger efficiency, as a function of the E_T^{miss} lower bound with the baseline event selection applied in the muon channel, with muons treated as invisible in the calculation of E_T^{miss} .

the known < 100% trigger efficiency of the single muon trigger. Note that since Eq. 5.19 is evaluated only for MC simulated events, there is no need for an independent trigger in the numerator and denominator such as the single lepton trigger included in Eq. 5.18. As shown in Figure 5.20, the E_T^{miss} OR single muon trigger is found to be effectively 100% efficient, given the application of appropriate scale factors for the single muon trigger, in the muon channel with the baseline event selection for all lower bounds on the E_T^{miss} down to ~ 100 GeV.

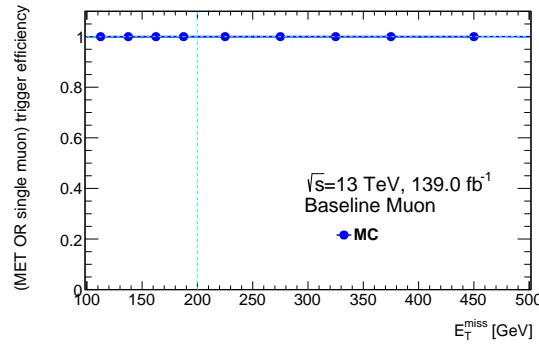


Figure 5.20: Efficiency of the E_T^{miss} OR single muon trigger, as a function of the E_T^{miss} lower bound, for the baseline event selection in the muon channel.

Based on the analysis presented in this section, it is concluded that, if all events considered in the analysis are explicitly required to have passed the (E_T^{miss} OR single muon) trigger, the trigger efficiency is known to be 100% for all events, except for the small subset of events with a high- p_T muon that pass the muon trigger but fail the E_T^{miss} trigger. For these events, scaling factors are included in the event weight to correct for the known < 100% efficiency of the single muon trigger. An additional “trigger-matching” requirement is applied for events that fail the E_T^{miss} trigger but pass the single muon trigger. This requirement ensures that the final state muon object that activated the single muon trigger can be identified as the same muon object (i.e. as having originated from the same muon) that was used to reconstruct the signal muon used in the search.

2930

Chapter 6

2931

Systematic Uncertainties

2932 In order to properly assess the significance of any discrepancies between predicted
 2933 yields of signal and SM background and of the observed collision data, it is important
 2934 to assign uncertainties to all sources that may limit the precision and accuracy of yield
 2935 predictions. In addition to a limited precision arising from statistical uncertainty¹,
 2936 there may be inaccuracies in the values of various parameters input to the simulation
 2937 due to a limited precision with which their values are known. These inaccuracies
 2938 can systematically shift predicted production rates and kinematic properties of the
 2939 simulated processes. Systematic uncertainties aim to quantify the uncertainty of
 2940 predicted yields of signal and SM background processes within each region and bin²
 2941 used in the search that could result from each source of potential inaccuracy in the
 2942 modelling.

2943 Systematic uncertainties (or simply “systematics”) are broadly classified into two
 2944 categories according to their origin: theoretical and experimental. Theoretical sys-
 2945 tematics account for inaccuracies that could result from a limited knowledge of pa-
 2946 rameters involved in modelling the production and decay mechanisms resulting from
 2947 pp collision events at the LHC. Experimental systematics are evaluated to account
 2948 for the limited accuracy and precision involved with the highly detailed model used
 2949 to simulate the operation of the ATLAS detector. In addition to the limited preci-
 2950 sion of the measured LHC beam luminosity, uncertainties arise from a wide range of
 2951 sources involved with simulating the detection and reconstruction of collision events.
 2952 These include limitations associated with the Geant4 model [124] used to simulate
 2953 the passage of particles through the detector, with modelling the data acquisition

¹See Chapter 4.1 for a discussion of the origin of statistical uncertainty in MC simulations.

²See Section 7.2 for details of the binning in $\min(m_s)$ performed in the SR

and readout systems employed by each sub-detector³, and with simulating the reconstruction and identification of physics objects⁴ using the simulated readouts from all the sub-detectors. Additional uncertainties arise from time-dependent modifications to the simulation, such as the reweighting of simulated events to account for time-varying pileup⁵ conditions in the detector.

Experimental systematics are generally constrained in the process of calibrating particular physics objects reconstructed in the ATLAS detector using the ATLAS collision data, and as a result are generally considered quite well-defined and robust. In contrast, the choice of uncertainty to assign to theoretical parameters can be much less clear, as constraints from experimental results may be sparse, or in some cases such as appropriate choice of renormalization scale in perturbative QCD calculations, essentially non-existent (see Section 9.1 of Ref. [22] for a review of running coupling and the renormalization scale).

Experimental and theoretical systematic uncertainties are evaluated in all analysis regions and bins for the SM background processes. Due to the negligible yield of the DH signal process in the CRs, systematic uncertainties are only evaluated for the signal process in the SRs.

6.1 Experimental Systematics

Experimental systematics are evaluated for all physics objects considered in the search, and for the LHC beam luminosity.

The integrated luminosity \mathcal{L}_{int} recorded by the ATLAS detector for the full data set considered in this search is known with a precision of 1.7% [161]. Since, from Eq. 1.2, the total number of recorded collision events scales linearly with the integrated beam luminosity, propagating this $\pm 1.7\%$ systematic to the yields results in coherent 1.7% up and down shifts of the predicted yield for each process across all analysis regions and bins.

For a given systematic uncertainty on a parameter k used in the reconstruction of physics objects, the general procedure for propagating the systematic uncertainty to the predicted yield N_p of a process p in a bin j is as follows:

³See Section 3.4 for a description of the ATLAS sub-detectors.

⁴See Section 5.1 for a presentation of the reconstructed physics objects used in this search.

⁵See Section 4.2.2 for a discussion of the pileup reweighting weight used to account for time-varying pileup conditions in the detector.

- Repeat the reconstruction with k shifted up by one standard deviation: k up = $k + \sigma_k$.
- Evaluate the predicted yield in the bin $N_{p, k \text{ up}, \text{bin } j}$ with updated simulation. The “up” systematic yield uncertainty is:

$$\text{syst}(\text{exp, } p, k \text{ up, bin } j) = N_{p, k \text{ up}, j} - N_{p, \text{nom}, j} \quad (6.1)$$

where $N_{p, \text{nom}, j}$ is the nominal yield.

- Repeat the above process with k shifted down by one standard deviation to evaluate the “down” systematic uncertainty.

Due to statistical uncertainties arising from the limited number of MC events available to evaluate shifted yields, asymmetries between the up and down yield shifts in the above procedure can occur simply due to statistical fluctuations in the number of events that fall into each bin, rather than from actual asymmetries in the underlying distribution being sampled. For this reason, the propagated yield systematics are symmetrized in each bin as follows:

$$\text{syst}(\text{exp, } p, k \text{ symm, bin } j) = \pm \left(\frac{N_{p, k, \text{up}, j} - N_{p, k, \text{down}, j}}{2} \right) \quad (6.2)$$

Table 6.1 summarizes the sources of experimental systematics considered for all physics objects in the search.

Figure 6.1 shows the envelope of all symmetrized shifts in the total predicted yield of all SM backgrounds in the SRs due to the sources of experimental systematics considered for each type of physics object. Figure 6.2 shows the same set of envelopes for a sample signal point at $(m_s, m_{Z'}) = (210, 2100)$ GeV. Comparing the yield shifts associated with the reconstruction of the various physics objects, the experimental systematics are dominated by sources of uncertainty associated with the reconstruction of $R = 0.4$ jets and $R = 0.2$ TAR subjets in the merged SR. The $R = 0.2$ TAR subjets uncertainties produce relatively small yield shifts in the resolved SR because TAR jets are not used in this region. Shifts in predicted yield are also notably smaller for the signal model compared with the SM background processes.

Tables 6.2 and 6.3 show a sample breakdown of systematic uncertainties associated with the total predicted yield of SM background processes in the SRs and CRs,

Table 6.1: Sources of experimental systematic uncertainty considered for all physics objects used in the search.

Physics Object	Systematic Uncertainties Considered
Electrons	<ul style="list-style-type: none"> – Reconstruction, ID and isolation efficiency – Energy scale and resolution
Muons	<ul style="list-style-type: none"> – Reconstruction, ID and isolation efficiency – Trigger efficiency – Energy scale and resolution – Track-to-vertex association
$R = 0.4$ Jets	<ul style="list-style-type: none"> – Energy scale and resolution – Quark flavour composition – Jet-vertex association – Pileup – b-tagging efficiency
TAR jets	<ul style="list-style-type: none"> – $R = 0.2$ subjets input to TAR algo (consider same sources as for $R = 0.4$ jets) – Uncertainties associated with tracks input to TAR algo
E_T^{miss}	<ul style="list-style-type: none"> – Soft terms used in E_T^{miss} calculation (note that uncertainties associated with other objects used in the E_T^{miss} calculation are propagated to E_T^{miss}, so are not classified as E_T^{miss} systematics)

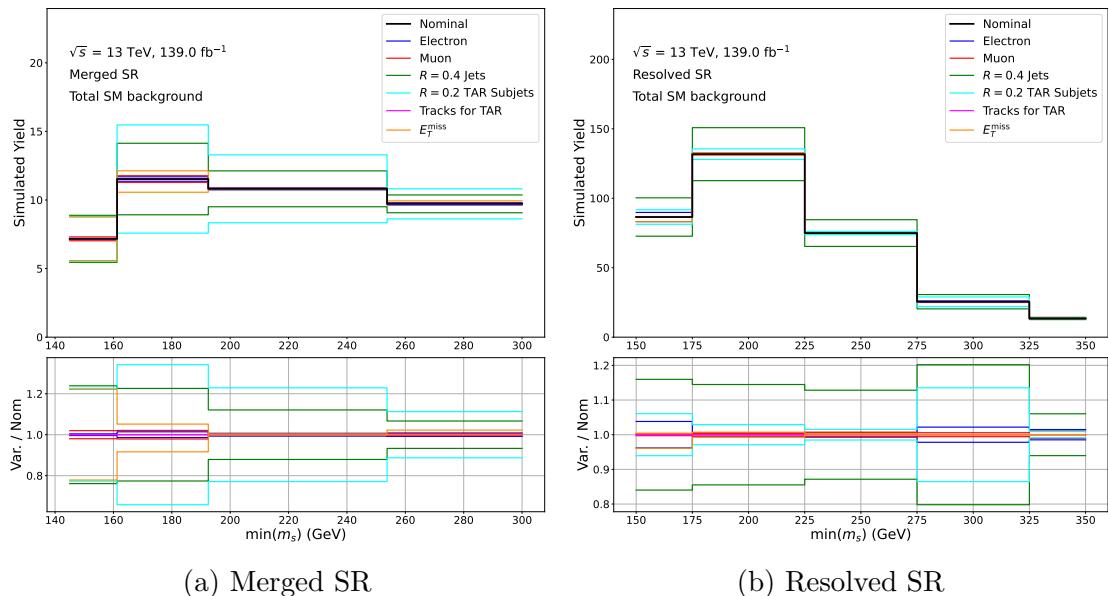


Figure 6.1: Envelope of shifts in the total predicted yield of SM background processes in the merged (left) and resolved (right) SRs due to experimental systematics associated with each physics object considered in the search. The predicted yield is binned in $\min(m_s)$ using the binning strategy employed in the fit used to search for evidence of the DH signal model in the data (see Chapter 7 for details). Bottom panel shows ratio of the shifts relative to the nominal yield.

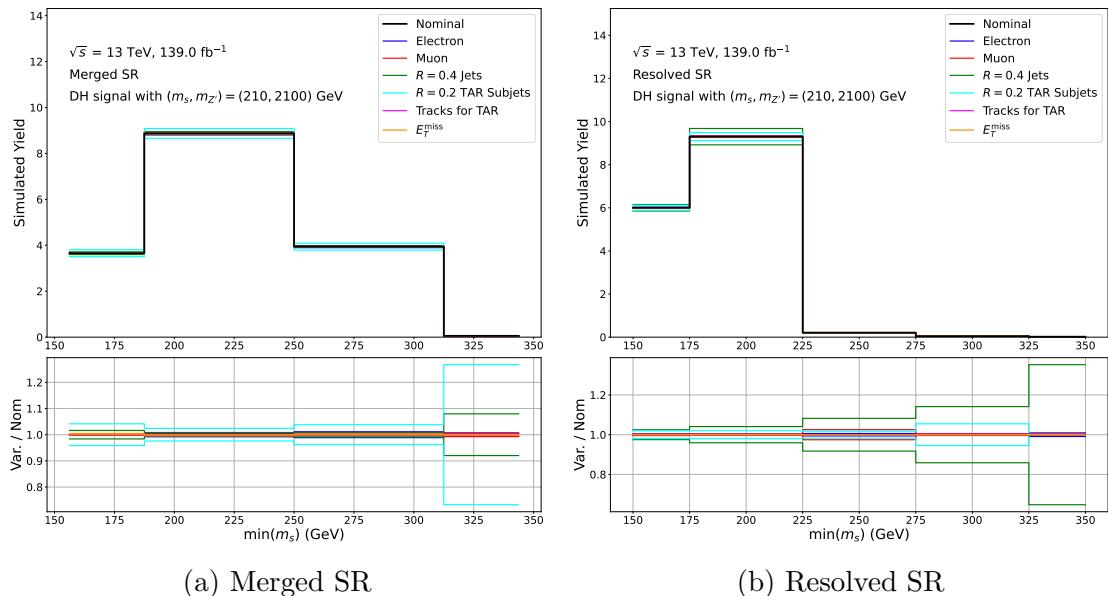


Figure 6.2: Envelope of shifts in the predicted yield of the DH signal process at $(m_s, m_{Z'}) = (210, 2100)$ in the merged (left) and resolved (right) SRs due to experimental systematics associated with each physics object considered in the search. The predicted yield is binned in $\min(m_s)$ using the binning strategy employed in the fit used to search for evidence of the DH signal model in the data (see Chapter 7 for details). Bottom panel shows ratio of the shifts relative to the nominal yield.

3010 respectively, which arise from shifting uncertain parameters associated with the re-
 3011 constructed jet energy resolution (JER) of $R = 0.4$ jets by $\pm\sigma$. The $R = 0.4$ JER
 3012 systematics are shown because they produce relatively large shifts in predicted yields
 3013 compared with other experimental sources.

Table 6.2: Symmetrized uncertainties in the total predicted yields of all SM background processes in the merged and resolved SRs due to varying parameters associated with jet energy resolution (JER) by $\pm 1\sigma$. Uncertainties are reported as percent shift relative to nominal yield.

Name of JER parameter	Mgd SR	Res SR
JET_JER_DataVsMC_MC16	$\pm 2.92\%$	$\pm 0.85\%$
JET_JER_EffectiveNP_10	$\pm 0.29\%$	$\pm 0.11\%$
JET_JER_EffectiveNP_11	$\pm 0.62\%$	$\pm 1.41\%$
JET_JER_EffectiveNP_12restTerm	$\pm 0.78\%$	$\pm 1.30\%$
JET_JER_EffectiveNP_1	$\pm 1.80\%$	$\pm 0.10\%$
JET_JER_EffectiveNP_2	$\pm 0.25\%$	$\pm 2.48\%$
JET_JER_EffectiveNP_3	$\pm 0.46\%$	$\pm 2.33\%$
JET_JER_EffectiveNP_4	$\pm 0.29\%$	$\pm 0.55\%$
JET_JER_EffectiveNP_5	$\pm 1.94\%$	$\pm 2.62\%$
JET_JER_EffectiveNP_6	$\pm 1.05\%$	$\pm 1.64\%$
JET_JER_EffectiveNP_7	$\pm 0.42\%$	$\pm 1.46\%$
JET_JER_EffectiveNP_8	$\pm 0.73\%$	$\pm 0.01\%$
JET_JER_EffectiveNP_9	$\pm 0.86\%$	$\pm 0.49\%$

3014 6.2 Theoretical Systematics

3015 The sources of theoretical uncertainty considered in this search, as well as the methods
 3016 used to evaluate the resulting uncertainties in predicted yields for the various processes
 3017 relevant to the search, are discussed in the following sections. Some sources are only
 3018 evaluated for a subset of the processes considered in the search.

3019 Table 6.4 summarizes the theoretical uncertainties that are evaluated for each

Table 6.3: Symmetrized up and down uncertainties in total predicted yield of all SM background processes in the merged and resolved CRs due to varying parameters associated with jet energy resolution (JER) by $\pm 1\sigma$. Uncertainties are reported as percent shift relative to nominal yield.

Name of JER parameter	Mgd CRW	Res CRW	Mgd CRTT	Res CRTT
JET_JER_DataVsMC_MC16	$\pm 0.27\%$	$\pm 0.24\%$	$\pm 12.07\%$	$\pm 15.00\%$
JET_JER_EffectiveNP_10	$\pm 0.17\%$	$\pm 0.32\%$	$\pm 12.07\%$	$\pm 22.03\%$
JET_JER_EffectiveNP_11	$\pm 0.17\%$	$\pm 0.65\%$	$\pm 12.07\%$	$\pm 18.64\%$
JET_JER_EffectiveNP_12restTerm	$\pm 0.42\%$	$\pm 1.02\%$	$\pm 12.07\%$	$\pm 2.92\%$
JET_JER_EffectiveNP_1	$\pm 0.63\%$	$\pm 0.21\%$	$\pm 12.07\%$	$\pm 4.31\%$
JET_JER_EffectiveNP_2	$\pm 1.03\%$	$\pm 0.25\%$	$\pm 12.07\%$	$\pm 6.03\%$
JET_JER_EffectiveNP_3	$\pm 0.44\%$	$\pm 0.14\%$	$\pm 12.07\%$	$\pm 23.77\%$
JET_JER_EffectiveNP_4	$\pm 0.92\%$	$\pm 1.24\%$	$\pm 0.00\%$	$\pm 39.84\%$
JET_JER_EffectiveNP_5	$\pm 0.01\%$	$\pm 0.06\%$	$\pm 12.07\%$	$\pm 20.10\%$
JET_JER_EffectiveNP_6	$\pm 0.54\%$	$\pm 1.24\%$	$\pm 12.07\%$	$\pm 6.32\%$
JET_JER_EffectiveNP_7	$\pm 0.21\%$	$\pm 0.30\%$	$\pm 0.00\%$	$\pm 5.43\%$
JET_JER_EffectiveNP_8	$\pm 0.05\%$	$\pm 1.04\%$	$\pm 0.00\%$	$\pm 10.51\%$
JET_JER_EffectiveNP_9	$\pm 0.86\%$	$\pm 0.61\%$	$\pm 12.07\%$	$\pm 13.51\%$

3020 process⁶.

3021 **6.2.1 Use of Acceptance for Evaluating Theoretical Systematics of the**
 3022 **DH Signal Model**

3023 For the SM background processes, the systematic uncertainties of predicted yields
 3024 due to theoretical sources are in general evaluated directly using yield shifts induced
 3025 by varying the uncertain parameter. Following ATLAS guidelines, theoretical uncer-
 3026 tainties associated with the modelling of the DH signal process are instead evaluated
 3027 in a given bin by calculating the relative shift in the so-called “acceptance”. The
 3028 acceptance is defined for a given process p as the ratio of yield in the given bin with
 3029 all analysis selection applied, relative to the inclusive yield of the process with no
 3030 selections applied:

⁶Due the relatively low yield and small number of MC simulated events admitted into the analysis regions for the $Z + \text{jets}$ process and the similarity in the production mechanisms for the $W + \text{jets}$ and $Z + \text{jets}$ processes, theoretical uncertainties are not explicitly calculated for the $Z + \text{jets}$ process. Instead, all relative theoretical uncertainties evaluated in each bin j for the $W + \text{jets}$ process $\text{syst}(\text{theo unc, } W + \text{jets, } j)/N_{W + \text{jets, nominal, } j}$ are also assigned to the $Z + \text{jets}$ process.

Table 6.4: Summary of theoretical uncertainties evaluated for each process considered in the search.

Source	W+jets	Diboson	Triboson	Z+jets	t <bar>t</bar>	Single Top	DH Signal
PDF	✓	✓	✓	✓	✓	✓	✓
α_s (PDF)	✓	✓	✓	✓	✗	✗	✗
α_s (ISR)	✗	✗	✗	✗	✓	✓	✗
μ_R and μ_F	✓	✓	✓	✓	✓	✓	✓
ME	✗	✗	✗	✗	✓	✓	✗
PS	✓	✓	✓	✓	✓	✓	✓
Wt/t <bar>t Int.</bar>	✗	✗	✗	✗	✗	✓	✗

$$\text{acc}(p, \text{ bin } j) = \frac{N(p, j)}{N(p, \text{inclusive})} \quad (6.3)$$

3031 The uncertainty in the predicted yield in each bin is then evaluated by multiplying
 3032 the relative shift in acceptance by the nominal yield in the bin:

$$\text{syst}(p, \text{ up, bin } j) = N_{p, \text{ nom, } j} \left(\frac{\text{acc}(p, \text{ up, } j) - \text{acc}(p, \text{ nom, } j)}{\text{acc}(p, \text{ nom, } j)} \right) \quad (6.4)$$

3033 Unless a given uncertainty applies exclusively to the signal model, the methods
 3034 used to evaluate the uncertainties in the yield associated with each theoretical source
 3035 are for the sake of brevity presented in their direct application to the yield, as would
 3036 be used for background processes. In general, the presented methods would be ap-
 3037 plied rather to the acceptance when considering the DH signal process, and the yield
 3038 uncertainty would subsequently be derived from the relative acceptance uncertainty
 3039 using Eq. 6.4.

3040 6.2.2 Modelling the Parton Distribution Function

3041 Uncertainties associated with the parton distribution function⁷ (PDF) used to model
 3042 the substructure of protons involved in LHC collisions can affect the predicted cross
 3043 sections of production and decay processes at the LHC. Following ATLAS guidelines
 3044 and recommendations from the *PDF4LHC* [158] working group, generator weights⁸
 3045 of events simulated for all processes are re-evaluated for 100 replicas of the nominal
 3046 PDF. Each replica is obtained by randomly re-sampling all uncertain inputs to the
 3047 PDF model and re-evaluating the PDF with the new inputs [162]. The yield in each
 3048 region and bin is re-evaluated with each set of generator weights, and the uncertainty
 3049 of the yield in each bin for a given process is estimated as the standard deviation over
 3050 all yield variations [158].

3051 6.2.3 Strong Coupling Constant α_s

3052 The nominal PDF is produced with the strong coupling constant α_s set to its currently
 3053 accepted value of $\alpha_s(m_Z^2) = 0.1180 \pm 0.0015$ [158] at a momentum scale m_Z .

⁷See Section 3.1 for an introduction to the parton model and parton distribution functions.

⁸See discussion of generator weights in Section 4.2.2

3054 **α_s in PDF Modelling**

3055 For the $W + \text{jets}$, diboson and triboson processes generated using SHERPA 2.2, alter-
 3056 native generator weights are produced using PDFs re-evaluated with α_s varied up or
 3057 down by its ± 0.0015 uncertainty. Following the *PDF4LHC* prescription, the yield
 3058 N_p for a given process p is re-evaluated in each bin j with the alternative generator
 3059 weights, and the uncertainty is evaluated as:

$$\text{syst}(\alpha_s, p, \text{bin } j) = \pm \left(\frac{N_{p, \alpha_s+1\sigma, j} - N_{p, \alpha_s-1\sigma, j}}{2} \right) \quad (6.5)$$

3060 and added in quadrature with the PDF uncertainties discussed in Section 6.2.2.

3061 **Strong Coupling Constant in the Modelling of Initial State Radiation (ISR)**

3062 For the single-top and $t\bar{t}$ processes generated using the POWHEG Box matrix element
 3063 generator interfaced with the PYTHIA 8 parton shower generator, the effect of varying
 3064 the strong α_s used to model initial state radiation (ISR) from $\alpha_{s, \text{ISR}} = 0.1$ to $\alpha_{s, \text{ISR}} =$
 3065 0.15 is evaluated using the up and down Var3c A14 tune variation [163], and the
 3066 associated yield uncertainty in each bin is evaluated as half the resulting difference
 3067 of yields, as in Eq. 6.5.

3068 **6.2.4 Renormalization and Factorization Scales (μ_R and μ_F) in QCD**

3069 In the framework of perturbative QCD, the strong coupling constant is expressed as a
 3070 function of the “renormalization scale” μ_R , which is unphysical in the sense that the
 3071 values of physical observables should be independent of μ_R . However, the choice of
 3072 μ_R can impact the calculated values of observables in perturbative QCD calculations
 3073 due to missing higher-order terms in truncated expansions. See Section 9.1 of Ref.
 3074 [22] for a review of running coupling and the renormalization scale. A similar effect
 3075 is seen with the “factorization scale” μ_F used in calculations of the proton PDF (see,
 3076 for example, Section 9.2 of Ref. [22]), which effectively quantifies the resolution with
 3077 which the proton is probed in a collision. As with μ_R , the choice of μ_F has no physical
 3078 meaning, but affects the calculated values of observables due to missing higher-order
 3079 terms in truncated QCD expansions.

3080 To account for the unphysical impact of the choice of μ_R and μ_F on simulated
 3081 observables, generator weights are re-evaluated for each process with μ_R and μ_F varied
 3082 by factors of either $\frac{1}{2}$ or 2 in pairwise combinations (i.e. $(\mu_R, \mu_F) \rightarrow \{(\frac{1}{2}\mu_R, \mu_F), (\mu_R, 2\mu_F)\}$,

3083 etc.) from their values used for the nominal event generation. Following ATLAS
 3084 guidelines, the systematic uncertainty in the yield in each bin is evaluated as the
 3085 envelope of yield variations over all pairwise combinations of μ_R and μ_F variations,
 3086 excluding the extreme off-diagonals $\{(\frac{1}{2}\mu_R, 2\mu_F), (2\mu_R, \frac{1}{2}\mu_F)\}$:

$$\text{syst}(\text{scale, up, } p, \text{ bin } j) = \max \left[0, \max_{k \in \{\mu_R, \mu_F \text{ var'ns}\}} \{N_{p, k, j} - N_{p, \text{nom}, j}\} \right] \quad (6.6)$$

3087 where “max” → “min” for the “down” variation.

3088 6.2.5 Matrix Element (ME) Generator Comparison

3089 For the $t\bar{t}$ and single-top processes, for which separate generators are used to calculate
 3090 the ME (POWHEG BOX) for the hard scatter process and model the subsequent parton
 3091 shower (PYTHIA 8), an uncertainty associated with the choice of using POWHEG Box
 3092 to calculate the ME is obtained by comparing the nominal yield predictions with those
 3093 obtained using an alternate MadGraph5_aMCNLO [130] generator. The associated
 3094 uncertainty in the predicted yield for a given process p ($t\bar{t}$ or single-top) in bin j is
 3095 evaluated using the difference in yield obtained with the alternate ME calculator:

$$\text{syst}(\text{ME, } p, \text{ bin } j) = \pm \left(N_{p, \text{ME=MadGraph5_aMCNLO}, j} - N_{p, \text{nom}, j} \right) \quad (6.7)$$

3096 6.2.6 Parton Showering (PS)

3097 Uncertainties associated with the modelling of the PS are evaluated for all processes.
 3098 The approach used to evaluate these uncertainties for a given process differs depending
 3099 on which package, or set of packages, was used to generate the nominal set of MC
 3100 simulated events for the process.

3101 Alternate PS Generator

3102 For the $t\bar{t}$, single-top and DH signal processes for which separate packages are used to
 3103 calculate the ME and to model the subsequent PS, an uncertainty associated with the
 3104 choice of PS generator is evaluated by comparing the predicted yields obtained with
 3105 the nominal PYTHIA 8 PS generator with an alternate HERWIG 7 [164, 165] generator.
 3106 In analogy with the evaluation of ME uncertainty described in Section 6.2.5 above,
 3107 the associated uncertainty in yield is evaluated for the $t\bar{t}$ and single-top processes as:

$$\text{syst}(\text{PS}, p, \text{bin } j) = \pm \left(N_{p, \text{PS=HERWIG 7, } j} - N_{p, \text{nom, } j} \right) \quad (6.8)$$

3108 For the DH signal model, MC simulated samples are generated with the alternate
3109 HERWIG 7 PS generator at three sample points in m_s and $m_{Z'}$:

- 3110 • $(m_s, m_{Z'}) = (160, 1000)$ GeV
- 3111 • $(m_s, m_{Z'}) = (160, 2100)$ GeV
- 3112 • $(m_s, m_{Z'}) = (310, 1000)$ GeV

3113 For each choice of PS generator - nominal or alternate - all three sets of samples
3114 generated at the mass points listed above are statistically combined, with the cross
3115 section excluded from the event weights⁹. The cross section is excluded when com-
3116 bining the samples to prevent the signal points at $m_s = 160$ GeV with larger cross
3117 sections from dominating the shape of the combined sample. Using the combined
3118 DH sample, the relative uncertainty in the predicted acceptance associated with the
3119 alternate PS generator is evaluated as:

$$(\text{rel. syst})(\text{PS, comb. DH, bin } j) = \pm \left(\frac{\text{acc}_{\text{comb. DH, PS=HERWIG 7, } j} - \text{acc}_{\text{comb. DH, nom, } j}}{\text{acc}_{\text{comb. DH, nom, } j}} \right) \quad (6.9)$$

3120

3121 This relative uncertainty is then applied to the DH signal yields at all m_s and $m_{Z'}$:

$$\text{syst}_{(\text{PS, } (m_s, m_{Z'}), \text{ bin } j)} = \pm \left[(\text{rel. syst})(\text{PS, combined DH, } j) \right] \times \left[N_{(m_s, m_{Z'}), \text{ nom, } j} \right] \quad (6.10)$$

3122

3123 CKKW and QSF PS Systematics

3124 For the $W + \text{jets}$, diboson and triboson processes, the SHERPA 2.2 generator is used
3125 to model both the ME calculation and the PS. Since it is not in general possible
3126 to disentangle the ME and PS components of the SHERPA 2.2 event simulation, the
3127 uncertainty associated with the PS shower generation is instead evaluated by varying

⁹see Section 4.2.2 for a detailed discussion of event weights

3128 the following theoretical parameters in the parton showering model implemented in
 3129 SHERPA 2.2:

- 3130 • The “merging scale” Q_{cut} is used to merge the matrix element associated with
 3131 partonic emissions with the ensuing parton shower evolution in the CKKW
 3132 scheme [166, 132]. This merging scale is set to 20 GeV for the nominal event
 3133 generation. Following ATLAS guidelines, the uncertainty associated with the
 3134 choice of Q_{cut} is evaluated by producing alternate MC simulated samples with
 3135 Q_{cut} varied to either $Q_{\text{cut, down}} = 15$ GeV or $Q_{\text{cut, up}} = 30$ GeV.
- 3136 • The “resummation scale” μ_{QSF} is used in modelling parton shower evolution (see
 3137 Section 9.2.3.3 of Ref. [22] for a review of resummation in QCD calculations).
 3138 Following ATLAS guidelines, the uncertainty associated with the choice of μ_{QSF}
 3139 is evaluated producing alternate samples with μ_{QSF} varied by factors of either
 3140 $\frac{1}{2}$ or 2.

3141 *Evaluation Strategy for the W+jets Process*

3142 For the $W+\text{jets}$ process, event weights associated with up and down variations of
 3143 the Q_{cut} and μ_{QSF} scales are evaluated by the method described in Ref. [167] using
 3144 truth-level¹⁰ MC simulated samples that were generated with each scale variation.
 3145 These event weights are used to re-evaluate the predicted yield of $W+\text{jets}$ events
 3146 in all regions and bins for each scale variation, and the associated uncertainty is
 3147 calculated as half the difference in yields between up and down variations:

$$\text{syst}_{(\text{CKKW (QSF)}, W+\text{jets}, \text{bin } j)} = \pm \left(\frac{N_{W+\text{jets}, Q_{\text{cut up}} (\mu_{\text{QSF, up}}), j} - N_{W+\text{jets}, Q_{\text{cut down}} (\mu_{\text{QSF, down}}), j}}{2} \right) \quad (6.11)$$

3148 *Evaluation Strategy for Diboson and Triboson Processes*

3149 For the diboson and triboson processes, predicted yields are evaluated for each
 3150 truth-level sample of MC simulated diboson events generated with variations of the
 3151 Q_{cut} and μ_{QSF} scales, and for the nominal truth-level sample, in a single “inclusive”
 3152 region defined by the following selection on truth-level observables:

¹⁰See Section 4.2 for information on the distinction between truth-level and reconstruction-level simulation

- 3153 • (1 signal muon) or (1 signal electron)
- 3154 • (1 or more $R = 1.0$ large- R jet¹¹) or (2 or more $R = 0.4$ jets)
- 3155 • $E_T^{\text{miss}} > 150 \text{ GeV}$
- 3156 • $m_T(\ell, E_T^{\text{miss}}) > 150 \text{ GeV}$

3157 The inclusive region is designed to contain all of the analysis regions used in the
 3158 search, with the lower bounds on E_T^{miss} and $m_T(\ell, E_T^{\text{miss}})$ placed as high as possible in
 3159 order to reflect the kinematics of events in the analysis regions, while still admitting
 3160 enough simulated events to ensure that statistical yield fluctuations are sub-dominant
 3161 in comparison to the yield shifts resulting from varying Q_{cut} and μ_{QSF} . Within this
 3162 inclusive region, the relative uncertainty in yield is evaluated for each set of scale
 3163 variations as:

$$(\text{rel. syst})_{(\text{CKKW (QSF), diboson})} = \pm \left(\frac{N_{\text{diboson (tr.)}, Q_{\text{cut up}} (\mu_{\text{QSF, up}}), \text{incl.}} - N_{\text{diboson (tr.)}, Q_{\text{cut down}} (\mu_{\text{QSF, down}}), \text{incl.}}}{2 \times N_{\text{diboson (tr.)}, \text{nom, incl.}}} \right) \quad (6.12)$$

3164
 3165 The relative uncertainty $(\text{rel. syst})_{(\text{CKKW (QSF), diboson})}$ is applied to the yield in each
 3166 analysis region and bin in the reconstruction level diboson and triboson¹² samples:

$$\text{syst}_{(\text{CKKW (QSF), diboson (triboson), bin } j)} = \pm \left[(\text{rel. syst})_{(\text{CKKW (QSF), diboson})} \right] \times \left[N_{\text{diboson (triboson), nom, } j} \right] \quad (6.13)$$

- 3167
- 3168 *Combination of CKKW and QSF Systematics*

3169 For each process, the CKKW and QSF systematic uncertainties are combined in
 3170 quadrature to obtain an overall estimate of the total uncertainty associated with the
 3171 parton shower modelling.

¹¹The TAR jets used in the search are by default built at reconstruction level. At truth level, the TAR jets are approximated by $R = 1.0$ jets constructed using the anti- k_t algorithm [118].

¹²The relative CKKW and QSF uncertainty evaluated for the diboson process is also applied to the triboson process because there are currently no MC simulated samples available for the triboson process with Q_{cut} and μ_{QSF} varied.

3172 6.2.7 Interference Between single-top Wt and $t\bar{t}$ Processes at NLO

3173 Interference between the mechanisms for single-top (Wt) and $t\bar{t}$ production (details
 3174 of which can be found in Ref. [168]) is accounted for when simulating the single-
 3175 top process to obtain an accurate prediction of the combined production rate for
 3176 the two processes. Two theoretical schemes, diagram removal (DR) and diagram
 3177 subtraction (DS) have been developed to handle this interference. The DR scheme is
 3178 used to generate the nominal MC simulated samples for the single-top process, and
 3179 an uncertainty associated with the choice of interference handling scheme is evaluated
 3180 by generating alternate samples for the single-top process using the DS scheme. The
 3181 associated uncertainty in the predicted yield of the single-top process is evaluated in
 3182 each bin j as:

$$\text{syst}(Wt/t\bar{t}, \text{single-top, bin } j) = \pm \left(N_{\text{single-top, DS, } j} - N_{\text{single-top, nom, } j} \right) \quad (6.14)$$

3183 6.2.8 Shifts in Predicted Yields due to Theoretical Systematic Uncer- 3184 tainties

3185 Figure 6.3 shows yield variations associated with all sources of theoretical uncertainty
 3186 presented in this section in the merged SR for sample processes.

3187 Tables 6.5 and 6.6 summarize the uncertainties associated with the total predicted
 3188 yield from all theoretical sources for each background process in the SRs and CRs,
 3189 respectively.

3190 6.3 Relative Impact of Statistical and Systematic Uncertainties in Anal- 3191 ysis Regions

3192 Table 6.9 compares the statistical uncertainty associated with the total MC simulated
 3193 yield of all background processes in each analysis region with the combined statis-
 3194 tical+systematic uncertainty. The combination of systematic uncertainties from all
 3195 sources is done within the HistFitter framework used to perform the statistical anal-
 3196 ysis for the search (see Chapter 7 for details). Table 6.10 shows the same comparison
 3197 for the DH signal model at a sample mass point. As discussed in Section 7.1.2, shifts
 3198 in the predicted yield of a given process induced by varying a given systematic source
 3199 are treated as correlated between all regions and bins in the fit. Yield shifts asso-
 3200 ciated with experimental sources are additionally treated as correlated between all

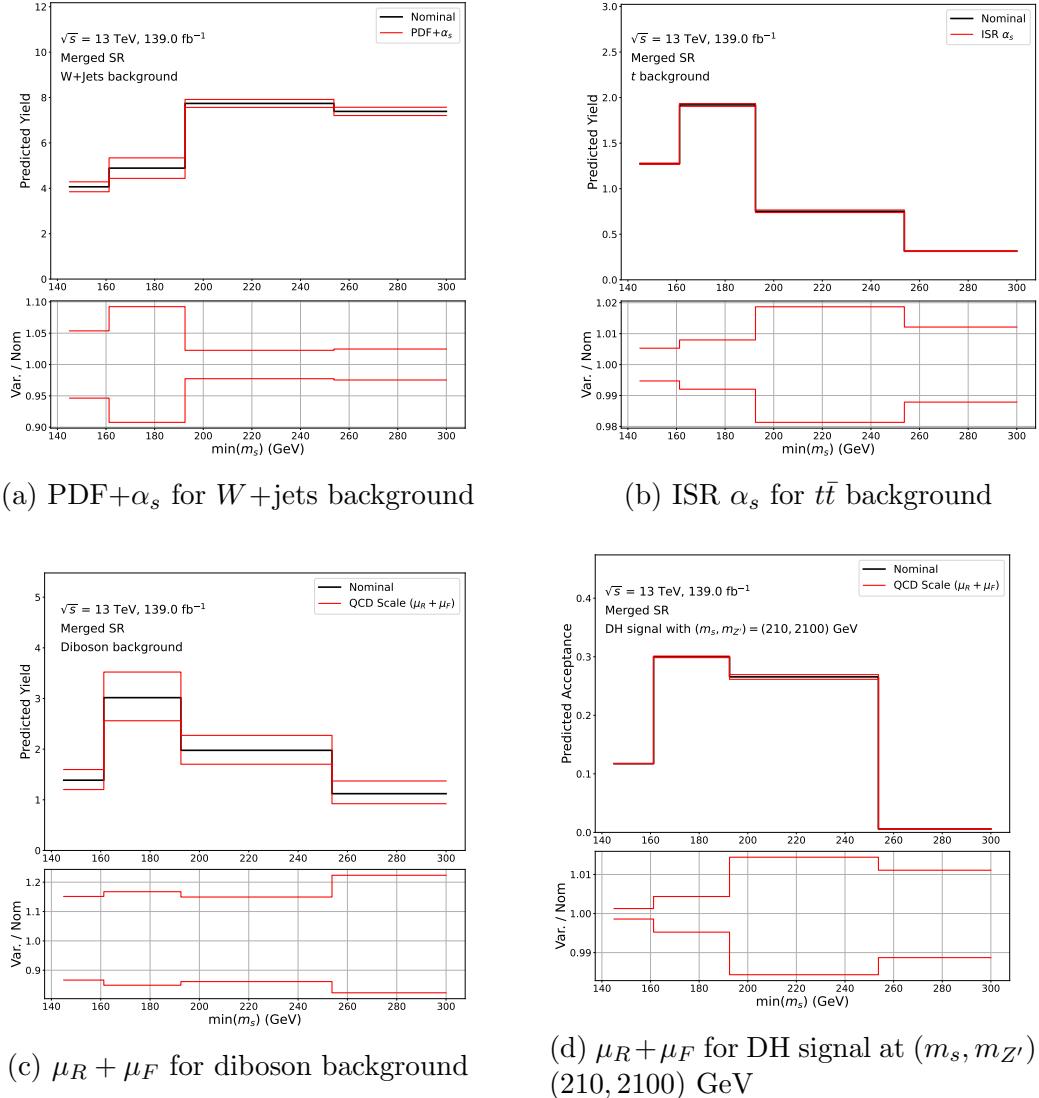
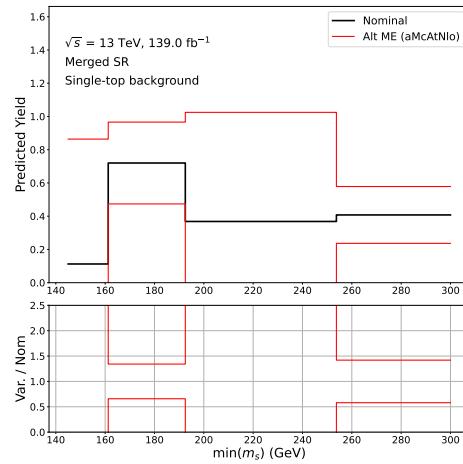


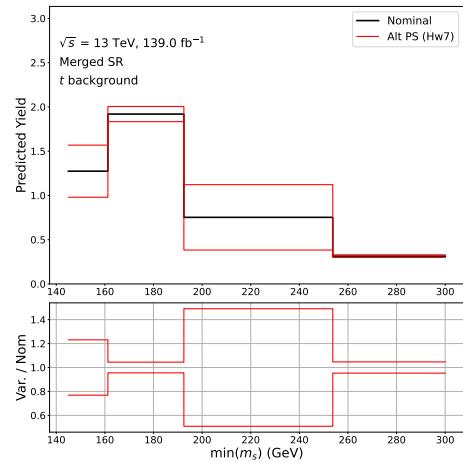
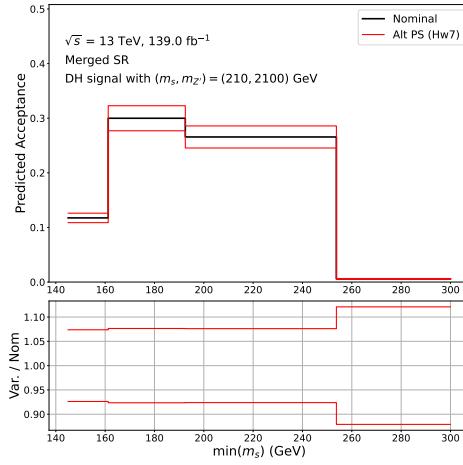
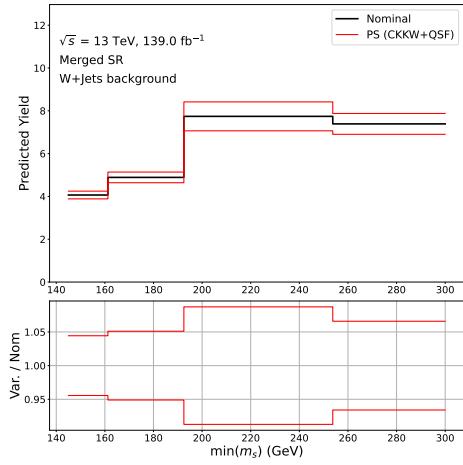
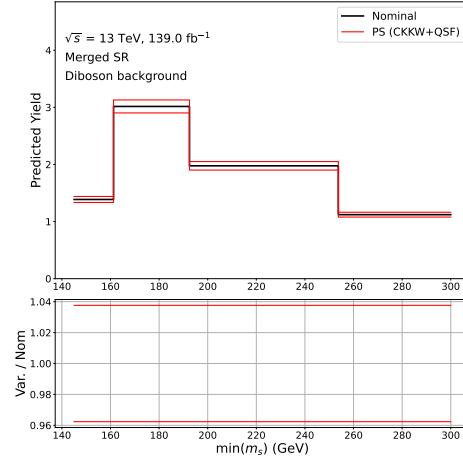
Figure 6.3: Shifts in the predicted yield of various processes considered in the search in the merged SRs due to theoretical systematic uncertainties. The predicted yield is binned in $\min(m_s)$ using the binning strategy employed in the fit used to search for evidence of the DH signal model in the data (see Chapter 7 for details). Bottom panel shows ratio of the shifts relative to the nominal yield.

3201 processes.

3202 For the SM background, the combined systematic uncertainty is comparable in
 3203 size to the statistical uncertainty in the merged SR, but all other analysis regions
 3204 the combined systematic uncertainty is dominant. For the DH signal process, the
 3205 statistical uncertainty dominates in the merged SR, but is comparable to the combined



(e) ME for single-top background

(f) PS for $t\bar{t}$ background(g) PS for DH signal at $(m_s, m_{Z'}) = (210, 2100) \text{ GeV}$ (h) PS (CKKW+QSF) for $W + \text{jets}$ background

(i) PS (CKKW+QSF) for diboson background

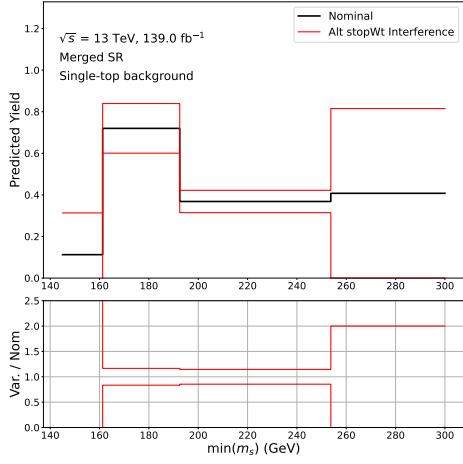
(j) $Wt/t\bar{t}$ Interference for single-top background

Figure 6.3: Shifts in the predicted yield due to theoretical systematic uncertainties (continued)

3206 systematic uncertainties in the resolved SR.

Table 6.5: Theoretical systematic uncertainties associated with predicted yields of dominant backgrounds in the SRs, reported both in terms of absolute yield, and (in parentheses) as percent relative to nominal yield.

Bkg and Syst.	Mgd SR	Res SR
W+Jets $\mu_R+\mu_F$	$+4.85 \text{ (+20.15\%)}$ -2.65 (-10.99\%)	$+33.92 \text{ (+14.33\%)}$ $-42.77 \text{ (-18.07\%)}$
W+Jets PDF+ α_s	$\pm 1.03 \text{ (\pm 4.27\%)}$	$\pm 3.18 \text{ (\pm 1.34\%)}$
W+Jets PS	$\pm 1.25 \text{ (\pm 6.22\%)}$	$\pm 16.31 \text{ (\pm 7.54\%)}$
diboson $\mu_R+\mu_F$	$+1.26 \text{ (+16.82\%)}$ -1.11 (-14.85\%)	$+7.62 \text{ (+15.22\%)}$ -6.61 (-13.19\%)
diboson PDF+ α_s	$\pm 2.50 \text{ (\pm 33.35\%)}$	$\pm 16.30 \text{ (\pm 32.55\%)}$
diboson PS	$\pm 0.28 \text{ (\pm 3.77\%)}$	$\pm 1.89 \text{ (\pm 3.77\%)}$
$t\bar{t}$ PS	$\pm 0.76 \text{ (\pm 17.94\%)}$	$\pm 1.82 \text{ (\pm 8.64\%)}$
$t\bar{t}$ ME	$\pm 2.15 \text{ (\pm 50.36\%)}$	$\pm 4.91 \text{ (\pm 23.26\%)}$
$t\bar{t}$ PDF	$\pm 0.03 \text{ (\pm 0.80\%)}$	$\pm 0.09 \text{ (\pm 0.43\%)}$
$t\bar{t}$ $\mu_R+\mu_F$ (ME)	$+0.16 \text{ (+3.74\%)}$ -0.21 (-4.91\%)	$+0.59 \text{ (+2.79\%)}$ -0.86 (-4.08\%)
$t\bar{t}$ ISR	$\pm 0.04 \text{ (\pm 0.94\%)}$	$\pm 0.24 \text{ (\pm 1.13\%)}$
$t\bar{t}$ $\mu_R+\mu_F$ (FSR)	$\pm 0.60 \text{ (\pm 14.12\%)}$	$\pm 0.73 \text{ (\pm 3.44\%)}$
stop PS	$\pm 1.04 \text{ (\pm 64.74\%)}$	$\pm 3.00 \text{ (\pm 80.10\%)}$
stop ME	$\pm 1.82 \text{ (\pm 113.45\%)}$	$\pm 3.07 \text{ (\pm 82.10\%)}$
stop $Wt/t\bar{t}$	$\pm 0.78 \text{ (\pm 48.58\%)}$	$\pm 2.33 \text{ (\pm 62.20\%)}$
stop PDF	$\pm 0.04 \text{ (\pm 2.58\%)}$	$\pm 0.09 \text{ (\pm 2.54\%)}$
stop $\mu_R+\mu_F$ (ME)	$+0.15 \text{ (+9.24\%)}$ -0.11 (-7.02\%)	$+0.39 \text{ (+10.45\%)}$ -0.33 (-8.78\%)
stop ISR	$\pm 0.15 \text{ (\pm 9.04\%)}$	$\pm 0.48 \text{ (\pm 12.87\%)}$
stop $\mu_R+\mu_F$ (FSR)	$\pm 1.19 \text{ (\pm 74.13\%)}$	$\pm 1.59 \text{ (\pm 42.47\%)}$
triboson $\mu_R+\mu_F$	$+0.04 \text{ (+3.58\%)}$ -0.04 (-3.11\%)	$+0.12 \text{ (+4.39\%)}$ -0.11 (-3.83\%)
triboson PDF+ α_s	$\pm 0.02 \text{ (\pm 1.78\%)}$	$\pm 0.04 \text{ (\pm 1.37\%)}$
triboson PS	$\pm 0.07 \text{ (\pm 6.22\%)}$	$\pm 0.21 \text{ (\pm 7.54\%)}$
Z+Jets $\mu_R+\mu_F$	$+0.32 \text{ (+20.15\%)}$ -0.18 (-10.99\%)	$+2.49 \text{ (+14.33\%)}$ -3.14 (-18.07\%)
Z+Jets PDF+ α_s	$\pm 0.07 \text{ (\pm 4.27\%)}$	$\pm 0.23 \text{ (\pm 1.34\%)}$
Z+Jets PS	$\pm 0.07 \text{ (\pm 6.22\%)}$	$\pm 1.30 \text{ (\pm 7.54\%)}$

Table 6.6: Theoretical systematic uncertainties associated with predicted yields of dominant backgrounds in the CRs, reported both in terms of absolute yield, and (in parentheses) as percent relative to nominal yield.

Bkg and Syst.	Mgd CRW	Res CRW	Mgd CRTT	Res CRTT
W+Jets $\mu_R+\mu_F$	$+35.90 \quad (+21.89\%)$ $-23.55 \quad (-14.36\%)$	$+42.02 \quad (+7.03\%)$ $-60.43 \quad (-10.10\%)$	$+3.50 \quad (-278.89\%)$ $-0.87 \quad (+68.98\%)$	$+4.83 \quad (+319.96\%)$ $-0.57 \quad (-37.98\%)$
W+Jets PDF+ α_s	$\pm 3.38 \quad (\pm 2.06\%)$	$\pm 13.10 \quad (\pm 2.19\%)$	$\pm 0.68 \quad (\pm -54.00\%)$	$\pm 0.30 \quad (\pm 19.81\%)$
W+Jets PS	$\pm 12.17 \quad (\pm 8.02\%)$	$\pm 35.66 \quad (\pm 6.86\%)$	$\pm 0.19 \quad (\pm 8.44\%)$	$\pm 0.18 \quad (\pm 10.77\%)$
diboson $\mu_R+\mu_F$	$+1.09 \quad (+4.06\%)$ $-4.20 \quad (-15.66\%)$	$+18.37 \quad (+17.40\%)$ $-15.45 \quad (-14.64\%)$	$+0.03 \quad (+20.96\%)$ $-0.02 \quad (-15.84\%)$	$+0.13 \quad (+22.80\%)$ $-0.10 \quad (-16.85\%)$
diboson PDF+ α_s	$\pm 8.62 \quad (\pm 32.18\%)$	$\pm 33.34 \quad (\pm 31.59\%)$	$\pm 0.05 \quad (\pm 32.59\%)$	$\pm 0.17 \quad (\pm 29.12\%)$
diboson PS	$\pm 1.00 \quad (\pm 3.77\%)$	$\pm 3.99 \quad (\pm 3.77\%)$	$\pm 0.00 \quad (\pm 3.77\%)$	$\pm 0.02 \quad (\pm 3.77\%)$
$t\bar{t}$ PS	$\pm 0.02 \quad (\pm 0.20\%)$	$\pm 0.20 \quad (\pm 0.98\%)$	$\pm 3.14 \quad (\pm 4.46\%)$	$\pm 4.93 \quad (\pm 7.68\%)$
$t\bar{t}$ ME	$\pm 2.74 \quad (\pm 27.66\%)$	$\pm 0.13 \quad (\pm 0.64\%)$	$\pm 1.39 \quad (\pm 1.97\%)$	$\pm 9.00 \quad (\pm 14.02\%)$
$t\bar{t}$ PDF	$\pm 0.03 \quad (\pm 0.33\%)$	$\pm 0.06 \quad (\pm 0.30\%)$	$\pm 0.29 \quad (\pm 0.41\%)$	$\pm 0.14 \quad (\pm 0.22\%)$
$t\bar{t}$ $\mu_R+\mu_F$ (ME)	$+0.28 \quad (+2.83\%)$ $-0.21 \quad (-2.10\%)$	$+1.27 \quad (+6.35\%)$ $-0.87 \quad (-4.38\%)$	$+2.42 \quad (+3.44\%)$ $-3.76 \quad (-5.34\%)$	$+0.98 \quad (+1.52\%)$ $-1.57 \quad (-2.44\%)$
$t\bar{t}$ ISR	$\pm 0.03 \quad (\pm 0.28\%)$	$\pm 0.11 \quad (\pm 0.53\%)$	$\pm 0.89 \quad (\pm 1.26\%)$	$\pm 0.04 \quad (\pm 0.07\%)$
$t\bar{t}$ $\mu_R+\mu_F$ (FSR)	$\pm 0.95 \quad (\pm 9.60\%)$	$\pm 0.46 \quad (\pm 2.31\%)$	$\pm 5.97 \quad (\pm 8.47\%)$	$\pm 2.89 \quad (\pm 4.49\%)$
stop PS	$\pm 2.10 \quad (\pm 77.12\%)$	$\pm 1.18 \quad (\pm 36.86\%)$	$\pm 1.97 \quad (\pm 36.44\%)$	$\pm 4.57 \quad (\pm 41.68\%)$
stop ME	$\pm 0.21 \quad (\pm 7.65\%)$	$\pm 2.03 \quad (\pm 63.64\%)$	$\pm 1.80 \quad (\pm 33.28\%)$	$\pm 2.05 \quad (\pm 18.72\%)$
stop $Wt/t\bar{t}$	$\pm 3.05 \quad (\pm 112.13\%)$	$\pm 1.09 \quad (\pm 34.30\%)$	$\pm 3.85 \quad (\pm 71.16\%)$	$\pm 7.90 \quad (\pm 71.98\%)$
stop PDF	$\pm 0.05 \quad (\pm 2.02\%)$	$\pm 0.22 \quad (\pm 6.90\%)$	$\pm 0.08 \quad (\pm 1.44\%)$	$\pm 0.25 \quad (\pm 2.31\%)$
stop $\mu_R+\mu_F$ (ME)	$+0.25 \quad (+9.08\%)$ $-0.18 \quad (-6.72\%)$	$+0.48 \quad (+14.99\%)$ $-0.34 \quad (-10.61\%)$	$+0.37 \quad (+6.89\%)$ $-0.27 \quad (-4.98\%)$	$+1.58 \quad (+14.44\%)$ $-1.08 \quad (-9.87\%)$
stop ISR	$\pm 0.02 \quad (\pm 0.75\%)$	$\pm 0.28 \quad (\pm 8.65\%)$	$\pm 0.17 \quad (\pm 3.09\%)$	$\pm 0.43 \quad (\pm 3.91\%)$
stop $\mu_R+\mu_F$ (FSR)	$\pm 0.24 \quad (\pm 8.87\%)$	$\pm 0.24 \quad (\pm 7.67\%)$	$\pm 0.05 \quad (\pm 0.99\%)$	$\pm 1.57 \quad (\pm 14.32\%)$
triboson $\mu_R+\mu_F$	$+0.11 \quad (+5.86\%)$ $-0.10 \quad (-4.99\%)$	$+0.34 \quad (+6.05\%)$ $-0.29 \quad (-5.20\%)$	$+0.00 \quad (+6.16\%)$ $-0.00 \quad (-5.27\%)$	$+0.00 \quad (+10.12\%)$ $-0.00 \quad (-8.51\%)$
triboson PDF+ α_s	$\pm 0.02 \quad (\pm 1.12\%)$	$\pm 0.06 \quad (\pm 1.12\%)$	$\pm 0.00 \quad (\pm 1.76\%)$	$\pm 0.00 \quad (\pm 3.98\%)$
triboson PS	$\pm 0.07 \quad (\pm 3.77\%)$	$\pm 0.21 \quad (\pm 3.77\%)$	$\pm 0.00 \quad (\pm 3.77\%)$	$\pm 0.00 \quad (\pm 3.77\%)$
Z+Jets $\mu_R+\mu_F$	$+3.34 \quad (+21.89\%)$ $-2.19 \quad (-14.36\%)$	$+1.21 \quad (+7.03\%)$ $-1.73 \quad (-10.10\%)$	$+0.00 \quad (-278.89\%)$ $-0.00 \quad (+68.98\%)$	$+0.00 \quad (+319.96\%)$ $-0.00 \quad (-37.98\%)$
Z+Jets PDF+ α_s	$\pm 0.31 \quad (\pm 2.06\%)$	$\pm 0.38 \quad (\pm 2.19\%)$	$\pm 0.00 \quad (\pm -54.00\%)$	$\pm 0.00 \quad (\pm 19.81\%)$
Z+Jets PS	$\pm 0.95 \quad (\pm 8.02\%)$	$\pm 1.24 \quad (\pm 6.86\%)$	$\pm 0.00 \quad (\pm 8.44\%)$	$\pm 0.00 \quad (\pm 10.77\%)$

Table 6.7: Theoretical systematic uncertainties associated with predicted yields of the DH signal model at several m_s and $m_{Z'}$ in the merged SR, reported both in terms of absolute yield, and (in parentheses) as percent relative to nominal yield.

Signal Point ($m_{Z'}$, m_s)	PDF	$\mu_R + \mu_F$	PS
(500, 260) GeV	$\pm 0.71\%$	$+1.76\%$ -1.70%	$\pm 4.74\%$
(1000, 335) GeV	$\pm 1.58\%$	$+1.64\%$ -1.48%	$\pm 4.74\%$
(1700, 235) GeV	$\pm 1.29\%$	$+0.98\%$ -1.04%	$\pm 4.74\%$
(2100, 210) GeV	$\pm 1.05\%$	$+1.13\%$ -1.23%	$\pm 4.74\%$

Table 6.8: Theoretical systematic uncertainties associated with predicted yields of the DH signal model at several m_s and $m_{Z'}$ in the resolved SR, reported both in terms of absolute yield, and (in parentheses) as percent relative to nominal yield.

Signal Point ($m_{Z'}$, m_s)	PDF	$\mu_R + \mu_F$	PS
(500, 260) GeV	$\pm 0.73\%$	$+1.35\%$ -1.21%	$\pm 8.23\%$
(1000, 335) GeV	$\pm 1.30\%$	$+1.40\%$ -1.27%	$\pm 8.23\%$
(1700, 335) GeV	$\pm 1.33\%$	$+0.49\%$ -0.52%	$\pm 8.23\%$
(2100, 210) GeV	$\pm 0.67\%$	$\pm 0.37\%$	$\pm 8.23\%$

Table 6.9: Comparison of the statistical uncertainties associated with the total predicted yield of SM background events in each analysis region with the total combined statistical+systematic uncertainties.

Analysis Region	Predicted SM Bkg Yield	Stat Unc	Stat+Sys Unc
Merged SR	39.74	± 7.29 (18%)	± 11.05 (28%)
Resolved SR	331.73	± 23.92 (7%)	± 64.68 (19%)
Merged CRW	220.62	± 24.48 (11%)	± 45.12 (20%)
Resolved CRW	749.53	± 31.03 (4%)	± 110.68 (15%)
Merged CRTT	77.27	± 2.28 (3%)	± 11.32 (15%)
Resolved CRTT	77.28	± 2.75 (4%)	± 16.57 (21%)

Table 6.10: Comparison of the statistical uncertainties associated with the predicted yield of the DH signal process at $(m_s, m_{Z'}) = (210, 2100)$ GeV in each analysis region with the total combined statistical+systematic uncertainties.

Analysis Region	Predicted SM Bkg Yield	Stat Unc	Stat+Sys Unc
Merged SR	16.19	± 2.89 (18%)	± 3.00 (18%)
Resolved SR	15.86	± 1.49 (9%)	± 2.11 (13%)

3207

Chapter 7

3208

Statistical Framework

3209 This chapter presents the statistical framework used to search for evidence of new
 3210 physics in the ATLAS collision data and, if no such evidence is found, establish the
 3211 range of parameters over which the DH model is excluded by the search. Evidence
 3212 of new physics would take the form of a statistically significant discrepancy between
 3213 distributions of the observed collision data in the SRs compared with distributions of
 3214 the predicted yield of SM background processes.

3215 Owing to the observation of m_s -dependent peaks in the $\min(m_s)$ distribution
 3216 of the DH signal model in the SRs¹, the sensitivity of the search was found to be
 3217 dramatically boosted by binning all data - MC simulated and ATLAS collision events
 3218 - in the SRs into several bins of $\min(m_s)$ prior to performing the statistical analysis.
 3219 Details of the binning strategy are presented in Section 7.2. Within each analysis
 3220 region and bin, the statistical analysis is performed by fitting the predicted yields
 3221 of background and signal processes to the ATLAS collision data with the aim of
 3222 maximizing the likelihood function presented in the following section.

3223 The computational construction and analysis of the likelihood function is per-
 3224 formed within the HistFitter [1] statistical analysis framework.

3225 **7.1 Likelihood Function**

3226 The likelihood function that lies at the heart of the statistical framework is a product
 3227 of Poisson distributions of event counts in all regions and bins:

¹See Section 5.1.7 for details.

$$\begin{aligned}
L(\mathbf{n}, \boldsymbol{\theta}^0 | \mu_{\text{Sig}}, \mathbf{s}, \boldsymbol{\mu}, \mathbf{b}, \boldsymbol{\theta}) &= P_{\text{SR}} \times P_{\text{CRs}} \times C_{\text{syst}} \\
&= \prod_{i \in \text{SR bins}} P(n_{S,i} | \lambda_{S,i}(\mu_{\text{Sig}}, \boldsymbol{\mu}_i, s_i, \mathbf{b}_i, \boldsymbol{\theta})) \times \\
&\quad \prod_{j \in \text{CRs}} P(n_j | \lambda_j(\mu_{\text{Sig}}, \boldsymbol{\mu}_j, s_j, \mathbf{b}_j, \boldsymbol{\theta})) \times \\
&\quad C_{\text{syst}}(\boldsymbol{\theta}^0, \boldsymbol{\theta})
\end{aligned} \tag{7.1}$$

3228 where:

- 3229 • $\mathbf{n} \in n_{S,i}, n_j$ is the set of all observed event counts in each region and bin.
- 3230 • $n_{S,i}$ and n_j are the number of observed events in the i^{th} bin of the SR and in
3231 the j^{th} CR, respectively.
- 3232 • The signal strength μ_{Sig} is an overall factor that scales the predicted yield of
3233 collision events produced by the DH signal model in all regions and bins.
- 3234 • $\boldsymbol{\mu}_{i(j)}$ represents the normalization factors in each bin i or region j , which scale
3235 the $W + \text{jets}$ and $t\bar{t}$ backgrounds in all regions and bins of the fit. Separate
3236 normalization factors are used in the merged and resolved categories. The
3237 normalization factors are treated as unconstrained nuisance parameters (NPs)
3238 in the fit. They are initialized as 1, with an uncertainty of ± 1 .
- 3239 • $s_{i(j)}$ is the yield of the signal process in each bin i or region j predicted by MC
3240 simulation. Likewise, $\mathbf{b}_{i(j)}$ is the set of yield predictions for all SM background
3241 processes in each bin i or region j .
- 3242 • $\boldsymbol{\theta}$ represents the set of NPs that parametrize all uncertainties associated with
3243 the MC simulated yields. For each uncertainty source k , the corresponding NP
3244 θ_k continuously interpolates between the nominal value of the predicted yield
3245 and the up/down shifts associated with varying the given uncertainty source by
3246 $\pm 1\sigma$, as discussed in Section 4.4 of Ref. [1]. For systematic sources, the NPs
3247 are normalized in the HistFitter framework such that $\theta_k = 0$ for the nominal
3248 yield and $\theta_k = \pm 1$ corresponds to $\pm 1\sigma$ variations of the uncertainty source.
3249 NPs associated with statistical uncertainty of the predicted yields are instead
3250 normalized such that $\theta = 1$ represents the nominal yield in the given bin and
3251 $\theta_k = 1 \pm 1\sigma_{\text{stat}}/N$ represents $\pm 1\sigma$ yield shifts, where N and σ_{stat} are the expected
3252 yield in the bin and its statistical uncertainty, respectively.

- 3253 • $C_{\text{syst}}(\boldsymbol{\theta}^0, \boldsymbol{\theta})$ is a composite function of Gaussian priors, which is used to constrain
 3254 the floating NPs $\boldsymbol{\theta}$ in the fit based on their central values $\boldsymbol{\theta}^0$ and uncertainties
 3255 κ :

$$C_{\text{syst}}(\boldsymbol{\theta}^0, \boldsymbol{\theta}) = \prod_{k \in S} \frac{1}{\kappa_k \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{\theta_k^0 - \theta_k}{\kappa})^2} \quad (7.2)$$

3256 where S is the full set of uncertainties considered in the fit.

- 3257 • The Poisson expectation functions $\lambda_{S,i}(\mu_{\text{Sig}}, \boldsymbol{\mu}_i, s_i, \mathbf{b}_i, \boldsymbol{\theta})$ and $\lambda_j(\mu_{\text{Sig}}, \boldsymbol{\mu}_j, s_j, \mathbf{b}_j, \boldsymbol{\theta})$,
 3258 represent the total expected yield in each SR bin i and CR bin j , respectively.
 3259 The Poisson expectation function is discussed in further detail in the following
 3260 section.

3261 7.1.1 Poisson Expectation Function

3262 Within a given SR bin or CR k , the Poisson expectation function λ_k in Eq. 7.1 is
 3263 given by:

$$\lambda_k(\mu_{\text{Sig}}, \boldsymbol{\mu}_k, s_k, \mathbf{b}_k, \boldsymbol{\theta}) = \mu_{\text{Sig}} s_k \eta_{k,\text{Sig}}(\boldsymbol{\theta}) + \sum_{p \in \{\text{SM bkg processes}\}} \mu_{k,p} b_{k,p} \eta_{k,p}(\boldsymbol{\theta}) \quad (7.3)$$

3264 where s_k ($b_{k,p}$) is the yield of the signal process (SM background process p) predicted
 3265 by MC simulation, as discussed in Section 4.2. $\eta_{i,s}(\boldsymbol{\theta})$ is a scaling factor, nominally
 3266 set to 1, which parametrizes the variations in expected yield induced by varying the
 3267 NPs $\boldsymbol{\theta}$ in the fit:

$$\eta_{k,p}(\boldsymbol{\theta}) = 1 + \sum_{s \in \text{all systematics}} I(\theta_s(k, p), N_{\text{down}}(k, p), N_{\text{up}}(k, p)) \quad (7.4)$$

3268 where $I(\theta_p, \sigma_{\text{down}}(i, s), \sigma_{\text{up}}(i, s))$ is a continuous function of $\theta_s(k, p)$ that interpolates²
 3269 the relative shift in predicted yields between the up $N_{\text{up}}(k, p)$ and down $N_{\text{down}}(k, p)$
 3270 extrema, associated with varying the given uncertainty source by $\pm 1\sigma$.

²HistFitter employs a 6th-order polynomial to interpolate between $\sigma_{\text{down}}(i, s)$ and $\sigma_{\text{up}}(k, p)$ using a linear extrapolation beyond these extrema. See Section 4.1 of Ref. [169] for details.

3271 **7.1.2 Nuisance Parameter Nomenclature and Correlations**

3272 Many of the NPs that appear in the likelihood function are correlated between bins,
 3273 categories, and/or processes. For example, if a given NP is correlated between all
 3274 bins in the fit, the value of the NP will be the same within all bins: $\theta_s(k, p) \rightarrow \theta_s(p)$.
 3275 The names given to NPs in the HistFitter framework reflect the presence of any such
 3276 correlations by omitting the correlated information in the NP name. Taking the above
 3277 example of an NP that is correlated between all bins, this correlation will be reflected
 3278 by the omission of any bin number specification (eg. `_bin1`).

3279 The background normalization factors μ and uncertainty-related NPs θ are broken
 3280 down as follows:

$$\mu \in \{\mu_{W+\text{jets, merged}}, \mu_{W+\text{jets, resolved}}, \mu_{t\bar{t}, \text{merged}}, \mu_{t\bar{t}, \text{resolved}}\} \quad (7.5)$$

3281 and

$$\theta \in \{\theta_{\text{statistical}}, \theta_{\text{sys, experimental}}, \theta_{\text{sys, theory}}\} \quad (7.6)$$

3282 Table 7.1 summarizes the correlations present for each type of NP, and the scheme
 3283 used by HistFitter to assign names to each.

3284 Table C.1 (C.2) summarizes the scheme used to name NPs associated with the
 3285 experimental (theoretical) uncertainty sources presented in Chapter 6.

3286 **7.2 Binning Strategy**

3287 As shown in Figure 7.1, the $\min(m_s)$ variable presented in Section 5.1.7 has a m_s -
 3288 dependent peaked distribution, which for most m_s values is distinct in shape from the
 3289 distribution of SM background processes in this variable. This shape discrimination
 3290 between the signal and background processes is exploited when performing the likeli-
 3291 hood fit by binning the MC simulated events and collision data into several $\min(m_s)$
 3292 bins.

3293 The exact placement of bin edges in $\min(m_s)$ used when performing the fit was
 3294 optimized with the aim of maximizing the projected exclusion of the DH signal model
 3295 for a background-only hypothesis using the limit-setting strategy presented in Section
 3296 7.3.3, while maintaining a total predicted yield of > 5 SM background events in each
 3297 bin to justify the use of the asymptotic approximation during limit-setting (see Section
 3298 7.3.3 for details).

Table 7.1: Summary of naming scheme and correlation information for nuisance parameters in the likelihood function. See Tables C.1 and C.2 for details of the names assigned to NPs associated with individual sources of systematic uncertainty.

NP Type	Description	Correlation Info	Naming Scheme
$-\mu_{W+\text{jets, mgd}}$ $-\mu_{W+\text{jets, res}}$ $-\mu_{t\bar{t}, \text{mgd}}$ $-\mu_{t\bar{t}, \text{res}}$	Normalization factor that scales the $W + \text{jets}$ ($t\bar{t}$) background in all analysis regions and bins of the merged (resolved) category.	Correlated between all analysis regions and bins for the $W + \text{jets}$ ($t\bar{t}$) background in the merged (resolved) category. Each normalization factor scales one process ($t\bar{t}$ or $W + \text{jets}$) - i.e. no correlation between processes.	$-\mu_{W\text{jets_mgd}}$ $-\mu_{W\text{jets_res}}$ $-\mu_{t\bar{t}\text{bar_mgd}}$ $-\mu_{t\bar{t}\text{bar_res}}$
$\theta_{\text{statistical}}$	NPs that parametrize the statistical uncertainty associated with the predicted yield in each region and bin from MC simulation.	Correlated between all processes. Not correlated between regions or bins.	$\gamma_{\{\text{region}\}_{\{\text{category}\}}_{\{\text{bin type}\}_{\text{bin index}}}}$
		bin type = $\begin{cases} \text{"mS_bin"} & \text{in SR bins} \\ \text{"cuts"} & \text{in CRs} \end{cases}$	
$\theta_{\text{sys, experimental}}$	NPs that parametrize systematic uncertainties from experimental sources.	Correlated between all processes, regions and bins.	$\alpha_{\{\text{uncertainty source}\}}$
$\theta_{\text{sys, theory}}$	NPs that parametrize systematic uncertainties for each process from theoretical sources.	Correlated between all regions and bins. Not correlated between processes.	$\alpha_{\text{process}_{\{\text{uncertainty source}\}}}$

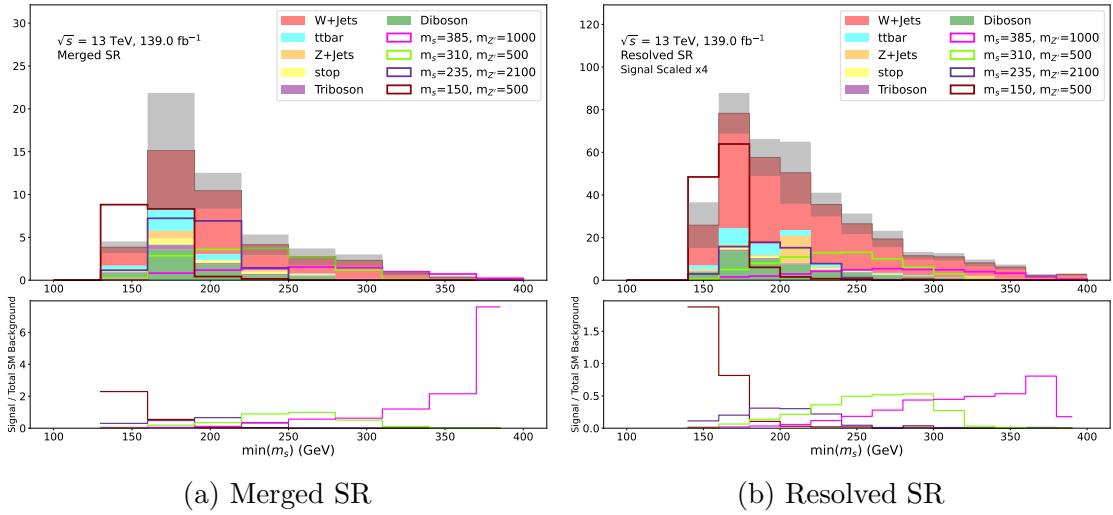


Figure 7.1: Predicted yields of SM background processes (stacked filled) and the DH signal model (unfilled lines) at several mass points in the merged (left) and resolved (right) SRs, binned in $\min(m_s)$. The lower panel shows the ratio of yields predicted for the DH signal process over the sum of MC background processes. Grey bands show the statistical uncertainty of the background estimate.

3299 Distributions of $\min(m_s)$ are shown with the optimized binning in Figure 7.2.
 3300 In the merged SR, binning the data into four $\min(m_s)$ bins with non-equidistant
 3301 bin edges was found to provide maximal sensitivity while maintaining > 5 predicted
 3302 SM background events per bin. In the resolved SR, binning into five $\min(m_s)$ bins
 3303 was found to be optimal, and equidistant bin edges are used because no appreciable
 3304 sensitivity improvements were found with non-equidistant edges.

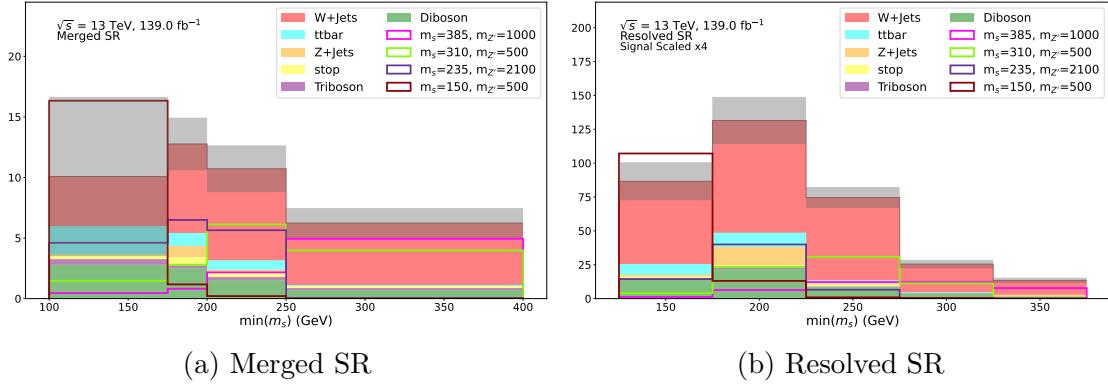


Figure 7.2: Predicted yield of SM background processes (stacked filled) and the DH signal model (unfilled lines) at several mass points in the merged (left) and resolved (right) SRs, binned in $\min(m_s)$ with the optimized bin edges presented in Table 7.2. Grey bands show the statistical uncertainty of the background estimate.

3305 The CRs are left unbinned in the fit to provide constraints on the overall yield of
 3306 the $W + \text{jets}$ and $t\bar{t}$ processes in each kinematic category. Table 7.2 summarizes the
 3307 binning strategy in all analysis regions.

Table 7.2: Binning used in the analysis regions.

Region	Binning in Merged Category	Binning in Resolved Category
SR	4 non-equidistant bins in m_s bin edges: [125, 165, 190, 225, 375] GeV	5 equidistant bins in m_s bin edges: [125, 175, 225, 275, 325, 375] GeV
$W + \text{jets}$ CR	none	none
$t\bar{t}$ CR	none	none

3308 7.3 Fit Setup

3309 With the MC simulated yields and ATLAS collision data binned into CRs and
 3310 $\min(m_s)$ bins within the SRs, as detailed in Section 7.2 above, the statistical analysis

3311 is performed by fitting the simulated yields to the collision data in all regions and
 3312 bins. The fits are performed within the HistFitter framework by varying the signal
 3313 strength parameter μ_{Sig} and all floating NPs $\{\boldsymbol{\mu}, \boldsymbol{\theta}\}$ in the likelihood function pre-
 3314 sented in Eq. 7.1, until the set of parameter values that maximizes the likelihood is
 3315 determined.

3316 7.3.1 Pruning of Systematics

3317 Prior to performing a fit, NPs associated with systematic uncertainties are pruned
 3318 within the HistFitter framework at a user-defined threshold. The pruning is done as
 3319 follows: for a given source of systematic uncertainty, HistFitter compares the relative
 3320 up and down yield shifts due to a given NP within each analysis region and bin. If
 3321 the relative shift in both directions is below the user-defined threshold in all regions
 3322 and bins, the given NP is ‘pruned’, which means that it is not included in fit. Pruning
 3323 helps to reduce fitting time, as well as minimizing numerical instabilities in the fit.

3324 For this search, a 1% pruning threshold is applied. Dedicated studies were done
 3325 to ensure that pruning systematics at a 1% threshold has no appreciable impact on
 3326 the overall fit behaviour or search results.

3327 7.3.2 Background-only Fit and Signal Region Extrapolation

3328 To initially check for an excess of data in the SRs, a “background-only” fit of the
 3329 SM background yield prediction to the observed collision data is performed in the
 3330 CRs to constrain the normalization factors (eg. $\mu_{W+\text{jets, mgd}}$) for the $W+\text{jets}$ and
 3331 $t\bar{t}$ backgrounds within each category, along with all other NPs $\boldsymbol{\theta}$ that affect yields
 3332 expectations in the CRs. The likelihood function presented in Eq. 7.1 is modified
 3333 such that μ_{Sig} is fixed to 0 (i.e. “background-only”), and the SR component P_{SR} is
 3334 fixed to 1 (i.e. “CR-only”).

3335 The values of normalization factors and other NPs that are constrained in the
 3336 background-only fit to maximize this modified likelihood function are subsequently
 3337 extrapolated to the SR to incorporate these constraints into the predicted yields of
 3338 background processes in the SR. The updated distributions of the total predicted SM
 3339 background yield in the SR are compared with distributions of the observed collision
 3340 data to check for the presence of any discrepancies that, if statistically significant
 3341 considering all statistical and systematic uncertainties, would be indicative of new
 3342 physics in the data.

3343 The extrapolation to the SR is performed by building the total expected yield of
 3344 SM background events in each SR bin using the Poisson expectation formula in Eq.
 3345 7.3 with the constrained NPs, and with μ_{Sig} still set to 0. Any NPs that are specific
 3346 to the SR and thus not constrained in the background-only fit are maintained at their
 3347 nominal pre-fit values when constructing the extrapolated Poisson expectations. Fur-
 3348 ther details of the extrapolation method and its implementation within the HistFitter
 3349 framework can be found in Section 5.2 of Ref. [1].

3350 Evaluation of Systematic Uncertainties by Means of the Transfer Factor

3351 After the values of all background normalization factors $\boldsymbol{\mu}$ are constrained in the
 3352 background-only fit, the total yield $((b_{\text{fit}})_{p, \text{mgd CRW}} + (b_{\text{fit}})_{p, \text{mgd CRTT}})$ of a process
 3353 p (either $W + \text{jets}$ or $t\bar{t}$) in the merged CRs whose normalization factor $\mu_{p, \text{mgd}}$ is
 3354 constrained in the fit can be expressed using the formalism of Eq. 7.3 (neglecting for
 3355 now the $\boldsymbol{\theta}$ constraints) as:

$$(b_{\text{fit}})_{p, \text{mgd CRW}} + (b_{\text{fit}})_{p, \text{mgd CRTT}} = \mu_{p, \text{mgd}} \left[b_{p, \text{mgd CRW}} + b_{p, \text{mgd CRTT}} \right] \quad (7.7)$$

3356
 3357 where an equivalent formula would apply in the resolved CRs with “mgd” → “res”.

3358 Extrapolating the constrained normalization factor $\mu_{p, \text{mgd}}$ to any region or bin j
 3359 in the merged category scales the predicted yield of the process p in a bin j of the
 3360 merged SR as follows:

$$\begin{aligned} (b_{\text{fit}})_{p, j} &= \mu_{p, \text{mgd}} \times b_{p, j} \\ &= \left[(b_{\text{fit}})_{p, \text{mgd CRW}} + (b_{\text{fit}})_{p, \text{mgd CRTT}} \right] \times \left[\frac{b_{p, j}}{b_{p, \text{mgd CRW}} + b_{p, \text{mgd CRTT}}} \right] \end{aligned} \quad (7.8)$$

3361 The ratio of raw MC yields in brackets, which extrapolates the fitted yield $((b_{\text{fit}})_{p, \text{mgd CRW}} +$
 3362 $(b_{\text{fit}})_{p, \text{mgd CRTT}})$ of process p in the merged CRs to the estimated yield $(b_{\text{fit}})_{p, j}$ in a
 3363 given region or bin j , is referred to as the “transfer factor” TF .

3364 When considering how the systematic uncertainties, which are correlated between
 3365 regions and bins, can be expected to impact the predicted yields $(b_{\text{fit}})_{p, j}$, the fitted
 3366 yield of the process $p \in \{W + \text{jets}, t\bar{t}\}$ in the CRs will be predominantly set by
 3367 the observed yield of collision data, and consequently should be effectively unshifted
 3368 when systematic uncertainty sources are varied. As a result, the primary impact of

3369 systematic uncertainty sources on the predicted yields $(b_{\text{fit}})_{p, \text{SR bin } j}$ in the SR will
 3370 take the form of correlated shifts of the predicted yields on the top and bottom of
 3371 the TF ratio. For this reason, relative systematic uncertainties associated with the
 3372 post-fit predicted yields in all regions are evaluated for the $W+\text{jets}$ and $t\bar{t}$ backgrounds
 3373 using systematic shifts in the TF rather than in the pre-fit yields. In a given SR bin
 3374 or CR j in the merged category:

$$\begin{aligned} \text{rel. syst}(\text{TF, sys up, } p, j \text{ (mgd)}) &= \frac{TF(\text{sys up, } p, j \text{ (mgd)})}{TF_{\text{nom}}} \\ &= \left[\frac{b_{p, \text{sys up, } j \text{ (mgd)}}}{b_{p, j \text{ (mgd)}}} \right] \times \\ &\quad \left[\frac{b_{p, \text{mgd CRW}} + b_{p, \text{mgd CRTT}}}{b_{p, \text{sys up, mgd CRW}} + b_{p, \text{sys up, mgd CRTT}}} \right] - 1 \end{aligned} \quad (7.9)$$

3375 where an analogous equation holds in the resolved category with "mgd" → "res".

3376 If yield shifts associated with a given systematic source are positively correlated
 3377 between the region or bin j and the CRs and comparable in size relative to the
 3378 nominal yield, then the relative systematic uncertainty of the transfer factor can be
 3379 significantly reduced compared with that of the raw yields due to cancellation between
 3380 the two terms in square brackets in Eq. 7.9.

3381 Table 7.3 compares the total relative systematic uncertainty associated with the
 3382 predicted yield of the $W+\text{jets}$ process in the SRs (bins combined) and $W+\text{jets}$
 3383 CRs, evaluated using either the systematic yield uncertainties alone (first term in
 3384 square brackets in Eq. 7.9) or using the TF . The $t\bar{t}$ CRs are excluded from the
 3385 comparison due to the relatively small predicted yield of $W+\text{jets}$ events in these CRs.
 3386 The combination of systematic uncertainties from all sources is performed within the
 3387 HistFitter framework. The relative systematic uncertainties are reduced in all regions
 3388 when evaluated using the TF rather than the yield. In the $W+\text{jets}$ CR (CRW), much
 3389 of the reduction comes from the fact that most of the $W+\text{jets}$ events in the combined
 3390 CRW+CRTT are contained within the CRW, so systematic shifts in the combined
 3391 CRW+CRTT yield are guaranteed to be highly correlated with those in the CRW.

Table 7.3: Comparison of total relative systematic uncertainty associated with the predicted yield of the $W+jets$ process in the SRs and $W+jets$ CRs (CRW), evaluated using either the predicted yield (second column) or the TF (third column). The relative statistical uncertainty is shown in the first column for comparison.

Analysis Region	Stat	Total Sys (yield)	Total Sys (TF)
Merged SR	$\pm 14\%$	$\pm 31\%$	$\pm 27\%$
Resolved SR	$\pm 7\%$	$\pm 24\%$	$\pm 14\%$
Merged CRW	$\pm 11\%$	$\pm 24\%$	$\pm 12\%$
Resolved CRW	$\pm 4\%$	$\pm 16\%$	$\pm 4\%$

3392 7.3.3 Exclusion Hypothesis Test and Limit Setting

3393 In the event that no significant discrepancy is seen in the comparison of SM back-
 3394 ground yield predictions to the ATLAS collision data in the SR following the background-
 3395 only fit and extrapolation of constraints to the SR, an exclusion hypothesis test of
 3396 the DH signal model is performed at each simulated m_s and $m_{Z'}$ to assess the range,
 3397 or “limits”, of model parameters that can be excluded by the search. This procedure
 3398 is referred to as “limit setting”.

3399 The exclusion hypothesis test is performed using “signal+background fits”, which
 3400 incorporate all regions and bins in the likelihood function presented in Eq. 7.1, and
 3401 allow for a nonzero signal strength μ_{Sig} . The ultimate product of the hypothesis test
 3402 at each m_s and $m_{Z'}$ of the DH model is a “ CL_s value”. A CL_s value below 0.05 is
 3403 considered to exclude the signal+background hypothesis at a 95% confidence level
 3404 (CL).

3405 Typical approaches to evaluating a CL associated with an alternative hypothesis
 3406 directly consider the level of agreement between the observed data and the expecta-
 3407 tion of the alternative hypothesis. In the context of a search for new phenomena in
 3408 experimental particle physics, this amounts to evaluating the confidence level CL_{s+b}
 3409 associated with the signal+background hypothesis. The CL_{s+b} approach encounters
 3410 issues when one considers a regime in which the data is expected to be predominantly
 3411 comprised of background events ($s \ll b$). In this regime, CL_{s+b} becomes increas-
 3412 ingly similar to the confidence level CL_b associated with the null background-only
 3413 hypothesis; this effect tends to diminish the ability of the search to exclude the sig-
 3414 nal+background hypothesis at a fixed reference CL (eg. 95%). To address this issue,

3415 the CL_s method considers instead the CL_s ratio:

$$\text{CL}_s = \frac{\text{CL}_{s+b}}{\text{CL}_b} \quad (7.10)$$

3416 which compares the confidence with which the observed data can be explained by
3417 the alternative signal+background hypothesis relative to the null background-only
3418 hypothesis. The CL_s is generally found to provide more powerful exclusion than the
3419 CL_{s+b} in the high-background regime.

3420 CL_s Evaluation Strategy

3421 For a given value of the hypothesized signal strength μ_{Sig} , a “profiled log-likelihood
3422 ratio” $q_{\mu_{\text{Sig}}}$ is constructed as follows:

$$q_{\mu_{\text{Sig}}} = -2 \log \left(\frac{L(\mu_{\text{Sig}}, \hat{\boldsymbol{\theta}})}{L(\hat{\mu}_{\text{Sig}}, \hat{\boldsymbol{\theta}})} \right) \quad (7.11)$$

3423 where $\hat{\mu}_{\text{Sig}}$ and $\hat{\boldsymbol{\theta}}$ are the values of μ_{Sig} and of the NPs $\boldsymbol{\theta}$ that maximize the likelihood
3424 function with all parameters left floating (i.e. profiled over) in the signal+background
3425 fit. $\hat{\boldsymbol{\theta}}$ is the set of floating NP values that maximize the likelihood function for a given
3426 fixed μ_{Sig} in the fit. Note that the background normalization parameters $\boldsymbol{\mu}$ are folded
3427 into the set of NPs $\boldsymbol{\theta}$ in Eq. 7.11, and in all equations that follow in this section.

3428 Two p-values are evaluated to quantify the CL_{s+b} (CL_b) by integrating the dis-
3429 tribution $f(q_{\mu_{\text{Sig}}} | \mu_{\text{Sig}}, \boldsymbol{\theta})$ above (below) the value $q_{\mu_{\text{Sig}}, \text{obs}}$ evaluated for $\mu_{\text{Sig}} = 1$
3430 ($\mu_{\text{Sig}} = 0$) with the observed ATLAS collision data and the central NP values in-
3431 put to the fit:

$$\text{CL}_{s+b} = p_{\mu_{\text{Sig}}=1} = \int_{q_1, \text{obs}}^{\infty} f(q_1 | \mu_{\text{Sig}} = 1, \boldsymbol{\theta}) dq_1 \quad (7.12)$$

$$\text{CL}_b = p_{\mu_{\text{Sig}}=0} = \int_{q_0, \text{obs}}^{\infty} f(q_0 | \mu_{\text{Sig}} = 0, \boldsymbol{\theta}) dq_0 \quad (7.13)$$

3432 The distribution $f(q_{\mu_{\text{Sig}}} | \mu_{\text{Sig}}, \boldsymbol{\theta})$ can be obtained by throwing pseudo-experiments
3433 that randomize the number of observed events and the central values of the NPs $\boldsymbol{\theta}$,
3434 and re-calculating $q_{\mu_{\text{Sig}}}$ for each pseudo-experiment. However, for sufficiently high
3435 statistics in the analysis regions - typically taken to be at least $\mathcal{O}(5)$ expected SM
3436 background events per region or bin - $f(q_{\mu_{\text{Sig}}} | \mu_{\text{Sig}}, \boldsymbol{\theta})$ is known to follow a χ^2 distri-
3437 bution according to Wilks’ theorem [170]. In this “asymptotic regime”, asymptotic

3438 formulae [171] are used to evaluate the p-value, thus avoiding the need to generate
 3439 pseudo-experiments. The CL_s value is then evaluated using Eq. 7.10 to test the
 3440 exclusion of the DH signal hypothesis at the given m_s and $m_{Z'}$. In addition to the
 3441 “observed” CL_s value $\text{CL}_{s, \text{obs}}$ obtained with the observed ATLAS collision data, an
 3442 additional “expected” CL_s value $\text{CL}_{s, \text{exp}}$ is evaluated by replacing the observed col-
 3443 lision data with the background-only “Asimov data set”, which is simply set to be
 3444 equal to the predicted yield of SM background processes in each region and bin.

3445 Visualization and Limit Setting

3446 Having obtained a $\text{CL}_{s, \text{obs}}$ and $\text{CL}_{s, \text{exp}}$ for each m_s and $m_{Z'}$ by the above procedure,
 3447 interpolation is performed with respect to both sets of CL_s values within the $(m_s,$
 3448 $m_{Z'})$ plane to obtain an expected (observed) “exclusion contour”, which corresponds
 3449 with $\text{CL}_{s, \text{exp}} = 0.05$ ($\text{CL}_{s, \text{obs}} = 0.05$). The $\text{CL}_{s, \text{exp}} = 0.05$ contour is evaluated with
 3450 a $\pm 1\sigma$ uncertainty band, and a fine-grained colour-coded distribution of the $\text{CL}_{s, \text{exp}}$
 3451 is plotted within the modelled range of m_s and $m_{Z'}$. The interpolation is performed
 3452 internally by HistFitter using a radial basis interpolation function³.

³Interpolation is performed using the `interpolate.Rbf()` class function in python’s `scipy` module. This work uses the ‘linear’ (r) interpolation function rather than the default ‘multiquadratic’, because linear interpolation was found to provide a smoother interpolation of the signal grid.

3453

Chapter 8

3454

Results

3455 This chapter presents the results of applying the statistical analysis presented in
 3456 Chapter 7 to the MC simulated and observed yields of ATLAS collision events in all
 3457 analysis regions and bins¹ to search for evidence of new physics in the SRs. No signif-
 3458 icant discrepancy is found between the predicted yields of SM background processes
 3459 and the collision data in the SRs, so the exclusion hypothesis test described in Section
 3460 7.3.3 is used to determine the parameter space of the DH model that is excluded by
 3461 the search.

3462 **8.1 Background-only Fit**

3463 This section presents the results of performing the “background-only” fit of the pre-
 3464 dicted yields of SM background processes obtained from MC simulation to the ob-
 3465 served yields of ATLAS collision events in the CRs, as described in Section 7.3.2, with
 3466 the goal of obtaining data-driven constraints for the normalizations of the $W + \text{jets}$
 3467 and $t\bar{t}$ backgrounds within each kinematic category.

3468 **8.1.1 Pre- and Post-Fit Yields of MC Simulated Background Events**

3469 Figure 8.1 compares the total predicted yield of all SM background processes consid-
 3470 ered in the CRs with the yield of observed events in the ATLAS collision data, both
 3471 before and after the background-only fit in the CRs. The observed yield of ATLAS
 3472 collision events is consistent with the pre-fit predicted yield of SM background events
 3473 within the combined statistical and systematic uncertainties in all CRs, with the ex-
 3474 ception of the merged CRTT. In the merged CRTT, a slight overprediction of the SM

¹See Section 5.2 for details of the selections applied to define the regions used for the search, and Section 7.2 for details of the strategy used to bin data in the SRs.

3475 background yield is observed.

3476 The uncertainty of the total expected yield of SM background processes is reduced
 3477 in all CRs after the fit, as the post-fit uncertainty is predominantly set by the Poisson
 3478 uncertainty associated with the observed event counts.

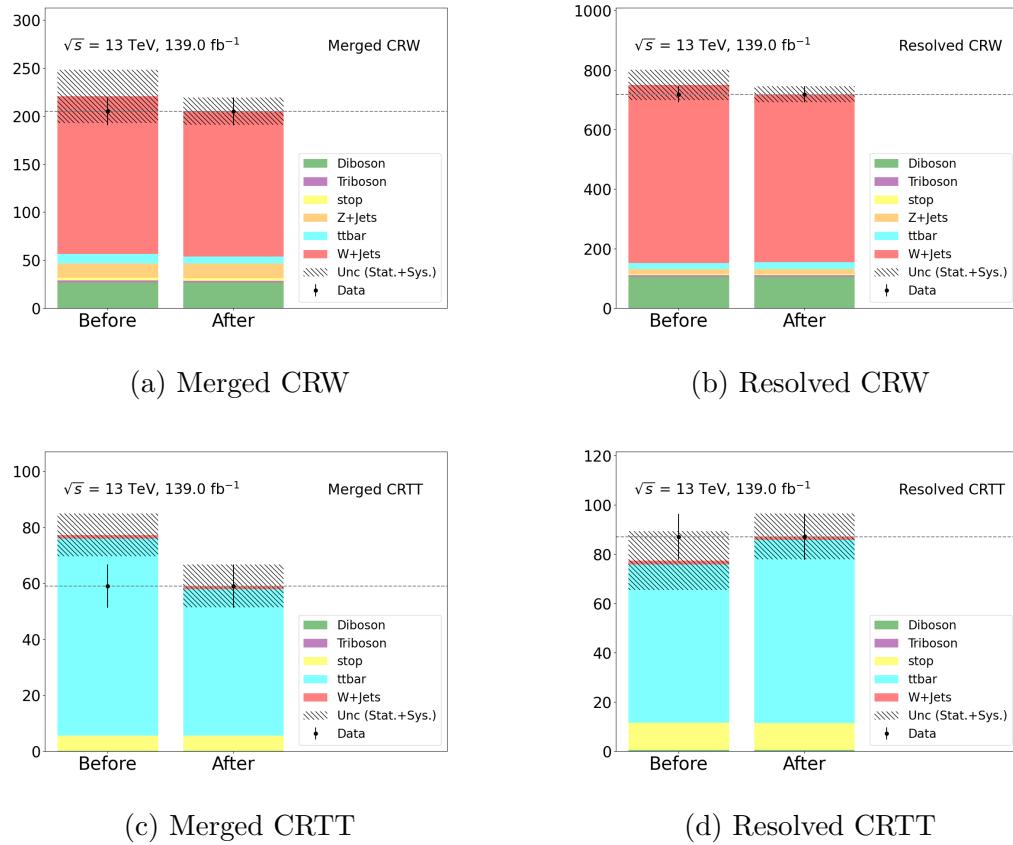


Figure 8.1: Comparison between the predicted yields of SM background processes and observed yields of ATLAS collision data in the CRs, before and after the background-only fit. Hatched band shows the combined statistical and systematic uncertainty of the total yield prediction of SM background processes. Black error bars represent the Poisson uncertainty associated with the observed event count in each CR.

3479 8.1.2 Nuisance Parameter Pulls and Correlations

3480 Figure 8.2 summarizes the post-fit shifts (a.k.a. “pulls”) of the values of all NPs
 3481 included in the background-only fit, as well as their post-fit uncertainties. Both the
 3482 values and uncertainties of the μ normalization factors, which scale the $W+jets$ and

3483 $t\bar{t}$ backgrounds separately in the merged and resolved categories, are constrained by
 3484 the fit to obtain the post-fit agreement seen in Figure 8.1 between the total predicted
 3485 yield of SM background processes and the observed yields in the ATLAS collision
 3486 data. The values of all other NPs, which parametrize sources of statistical and sys-
 3487 tematic uncertainty, receive negligible pulls in the fit, since agreement with data can
 3488 be obtained in all CRs just by varying the $W + \text{jets}$ and $t\bar{t}$ normalization factors.
 3489 The uncertainties of the γ_-^* parameters² that parametrize the statistical uncertainty
 3490 associated with the MC simulated yield predictions in each CR are reduced by the fit
 3491 to data, which results in the reduction seen in Figure 8.1 of the uncertainty associated
 3492 with the post-fit expected yield of SM background events.

3493 Figure 8.3 shows the Pearson correlation coefficient r between NPs in the fit. There
 3494 is some appreciable correlation ($|r| \gtrsim 0.2$) between separate background normalization
 3495 factors μ_-^* , and between normalization factors and a few of the γ_-^* NPs - for example,
 3496 $r = -0.7$ between $\mu_{W\text{jets_mgd}}$ and $\gamma_{\text{stat_CRW_Merged_cuts_bin_0}}$. However, there
 3497 is in general very little cross correlation ($r < 0.01$) between the γ_-^* and α_-^* NPs,
 3498 which collectively parametrize all uncertainty sources in the fit due to the negligible
 3499 shifts induced on these parameters by the background-only fit (see Figure 8.2).

3500 8.2 Comparison of SM Background Expectation and Data in the Signal 3501 Region

3502 After performing the background-only fit in the CRs, the constraints on the back-
 3503 ground normalization factors $\boldsymbol{\mu}$ and the other NPs $\boldsymbol{\theta}$ summarized in Figure 8.2 are
 3504 extrapolated to the SR following the procedure described in Section 7.3.2. Figure
 3505 8.4 compares the predicted yields of SM background processes in the SR - binned
 3506 in $\min(m_s)$ using the binning strategy presented in Section 7.2 - before and after
 3507 the background-only fit and extrapolation procedure. Table 8.1 summarizes the total
 3508 predicted and observed yields in the merged SRs, combined over all bins within each
 3509 SR.

3510 Both before and after the background-only fit extrapolation, Figures 8.4a and
 3511 8.4b, respectively, show an excess of observed ATLAS collision events compared with
 3512 the predicted yield of SM background processes in the first three bins of the merged
 3513 SR. However, the difference is within uncertainty in all three bins, and a comparison
 3514 of the combined yields reported in the first row of Table 8.1 reveals that the overall

²See Table 7.1 for a description of the scheme used to name NPs.

Table 8.1: Comparison of the total observed yields of events in the SRs with the predicted yield of SM background processes, either before (pre-fit) or after (post-fit) extrapolation of constraints from the background-only fit in the CRs.

Region	Observed Yield	SM Bkg Prediction (Pre-fit)	SM Bkg Prediction (Post-fit)
Merged SR	51	39.7 ± 11	36.8 ± 10
Resolved SR	285	331.7 ± 65	321.2 ± 40

yield of ATLAS collision data in the merged SR a factor of 1.3 larger than the total post-fit uncertainty associated with the expected yield of SM background processes (a.k.a. a 1.3σ excess). If the distribution of measured discrepancies between the predicted yield of SM background processes and the observed yield of collision data is assumed to be reasonably approximated as Gaussian, the observed 1.3σ discrepancy corresponds to a two-sided p-value of 0.18 (see, for example, Section 11 of [172] for an introduction to the p-value and its use in hypothesis testing). If a p-value below 0.05 is taken to represent a statistically significant deviation from the null hypothesis, the observed p-value of 0.18 implies that the observed yield of ATLAS collision events in the merged SR is statistically compatible with the null hypothesis that all events observed in the SR were produced by SM background processes.

Figures 8.4c and 8.4d show an over-prediction of the SM background yield compared with the observed number of ATLAS collision events in all bins of the resolved SR, both before and after NP constraints from the background-only fit are extrapolated to the SRs. Comparing the observed yield of collision events from Table 8.1 in the resolved SR with the post-fit predicted yield of SM background events reveals a -0.8σ discrepancy. The resulting two-sided Gaussian p-value of 0.42 indicates that, as in the merged SR, the observed yields are statistically compatible with the null background-only hypothesis.

8.3 Exclusion of the Dark Higgs Signal Model

Given the absence of any statistically significant discrepancy between the observed yields of ATLAS collision events in the SRs and the predicted yields of events from SM background processes, the exclusion hypothesis test presented in Section 7.3.3 is used to determine the range of m_s and $m_{Z'}$ for which the DH signal model can be confidently excluded by the search for the fixed choices of the DM mass m_χ , mixing

3540 angle $\sin \theta$ and coupling choices g_q and g_χ in the DH model used for the search³:

3541 • $m_\chi = 200$ GeV

3542 • $g_q = 0.25$

3543 • $g_\chi = 1$

3544 • $\sin \theta = 0.01$

3545 8.3.1 Signal+Background Fit

3546 Figures 8.5, 8.6 and 8.7 compare the observed yield of ATLAS collision events in the
 3547 SRs, binned in $\min(m_s)$, with the predicted yields for the SM background and DH
 3548 signal processes at the three sample combinations of m_s and $m_{Z'}$ in the signal model.
 3549 The sample m_s and $m_{Z'}$ combinations, listed below, are chosen to cover the range of
 3550 production cross sections for the DH model considered in the search:

3551 • $(m_s, m_{Z'}) = (160, 1000)$ GeV (cross section: 93.3 fb^{-1})

3552 • $(m_s, m_{Z'}) = (210, 2100)$ GeV (cross section: 3.9 fb^{-1})

3553 • $(m_s, m_{Z'}) = (310, 2900)$ GeV (cross section: 0.31 fb^{-1})

3554 As shown in Figure 8.5, the DH signal model at $(m_s, m_{Z'}) = (160, 1000)$ GeV
 3555 with a cross section of 93.3 fb^{-1} has a sufficiently large production rate that the
 3556 signal+background fit constrains the signal strength parameter to $\mu_{\text{Sig}} = 0.09 \pm 0.08$.
 3557 Therefore, the signal+background hypothesis ($\mu_{\text{Sig}} = 1$) is excluded with a high degree
 3558 of confidence.

3559 Figure 8.6 shows the results of the signal+background fit for $(m_s, m_{Z'}) = (210, 2100)$ GeV
 3560 in the signal model, with an intermediate cross section of 3.9 fb^{-1} . The predicted
 3561 yield of the signal process is comparable in the first three bins of the merged SR
 3562 to the discrepancy between the observed yield of data and the predicted yield of SM
 3563 background processes, which results in a fitted signal strength close to 1 ($\mu_{\text{Sig}} = 0.88$).
 3564 However, since the predicted yield of the signal and background processes are still
 3565 consistent with the background-only prediction, there is a large (nearly 100%) un-
 3566 certainty associated with the fitted signal strength. As a result, the fit can neither
 3567 support nor confidently exclude the signal+background hypothesis at this mass point.

³See Sections 2.2.1 and 2.3.2 for a presentation and discussion of the choices for the m_χ , $\sin \theta$ and coupling parameters in the DH model.

3568 Further reduced exclusion power is observed for the fit shown in Figure 8.7d, which
 3569 considers the signal model at the mass combination $(m_s, m_{Z'}) = (310, 2900)$ GeV with
 3570 the lowest cross section of the three combinations considered. The signal yields are
 3571 so small compared with the uncertainty of the total yield prediction that the fitted
 3572 signal strength could range from -3.82 to 5.06 within its post-fit uncertainty. As a
 3573 result, the search can neither meaningfully support nor exclude the DH signal model
 3574 at this mass point.

3575 Figure 8.8 summarizes the post-fit value and uncertainty of the signal strength
 3576 parameter μ_{Sig} for the signal+background fit performed with the DH signal model at
 3577 all m_s and $m_{Z'}$ considered in the search. The fitted value and uncertainty vary de-
 3578 pending on the production cross section and the shape of the signal distribution with
 3579 respect to $\min(m_s)$. The size of the uncertainty for a given m_s and $m_{Z'}$ combination
 3580 reflects the exclusion power of the search at the given mass combination, where a
 3581 larger uncertainty generally implies less exclusion power. In general, the fitted values
 3582 of μ_{Sig} are consistent with 0 within 1.5σ , in agreement with the null background-only
 3583 hypothesis.

3584 Nuisance Parameter Pulls and Correlations

3585 Figure 8.9 summarizes the pulls and uncertainties of all NPs included in the sig-
 3586 nal+background fit using the DH signal model at the sample mass point $(m_s, m_{Z'}) =$
 3587 $(210, 2100)$ GeV. In contrast with the background-only fit, the pulls of some γ_-^* and
 3588 α_-^* NPs, which constrain the statistical and systematic uncertainties associated with
 3589 MC simulated yields, can be appreciably shifted (i.e. pulled) relative to their pre-fit
 3590 values. This is due to differences in the shapes of $\min(m_s)$ distributions in the SRs
 3591 between the observed ATLAS collision data and the predicted yields of the SM back-
 3592 ground and signal processes, which cannot be corrected by varying the μ_-^* normal-
 3593 ization parameters alone. Particularly large pulls are seen for the NPs $\alpha_{\text{-JET_JER}}^*$
 3594 and $\alpha_{\text{-JET_R02_JER}}^*$, which parametrize the systematic jet energy resolution (JER)
 3595 uncertainties associated with the $R = 0.4$ ⁴ and $R = 0.2$ ⁵ jets, respectively (see Tables
 3596 C.1 and C.2 for descriptions of individual NPs used to parametrize systematic uncer-
 3597 tainties). As shown in Figures 6.1 and 6.2, the JER systematic uncertainties induce

⁴See Section 5.1.2 for a description of the $R = 0.4$ jets, and Section 5.1.3 for a description of the method used to reconstruct the W boson candidate in the resolved category.

⁵See Section 5.1.4 for a description of the algorithm used to construct TAR jets using input $R = 0.2$ jets.

3598 the largest yield shifts in the SRs relative to other systematic sources. Therefore,
 3599 it is reasonable to expect that these NPs would in general receive relatively strong
 3600 pulls in the signal+background fit to help correct for the observed differences between
 3601 expected and observed event yields in the SRs.

3602 Figure 8.10 shows the Pearson correlation coefficient r between NPs in the sig-
 3603 nal+background fit for the sample mass point $(m_s, m_{Z'}) = (210, 2100)$ GeV in the
 3604 DH signal model. As with the background-only fit, there is appreciable correlation
 3605 ($|r| \gtrsim 0.2$) between separate background normalization factors μ_-^* , and between
 3606 normalization factors and several of the γ_-^* NPs that parametrize the statistical
 3607 uncertainties associated with MC simulated yield predictions. In contrast to the
 3608 background-only fit, some of the α_-^* parameters also have non-negligible correlations
 3609 ($|r|$ up to ~ 0.3) with one another and with γ_-^* and μ_-^* NPs. The introduction
 3610 of correlations involving α_-^* NPs in the signal+background fit compared with the
 3611 background-only fit is attributed to the non-negligible post-fit shifts of these NPs
 3612 observed in Figure 8.9 for the signal+background fit.

3613 Ranking of Systematic Uncertainties

3614 Figure 8.11 shows the pre- and post-fit values and impacts of NPs in the signal+background
 3615 fit on the fitted signal strength μ_{Sig} . The 30 leading NPs are ranked from top to bot-
 3616 tom in order of the size of their impact. The pre- and post-fit impact of a given NP
 3617 θ is measured as follows:

3618 • **Pre-fit impact:** Shift the value of θ to the upper bound $\theta_0 + \Delta\theta$ of its pre-fit
 3619 uncertainty. Perform the signal+background fit with θ fixed to this upper value,
 3620 and all other NPs left floating. Repeat with the value of θ fixed to the lower
 3621 bound $\theta_0 - \Delta\theta$ of its pre-fit uncertainty. The resulting shifts $\Delta_{\text{up/down}}\widehat{\mu}_{\text{Sig}}$ of
 3622 the best-fit μ_{Sig} are shown as unfilled white boxes with black borders in Figure
 3623 8.11.

3624 • **Post-fit impact:** As above, but with the value of θ shifted instead to the upper
 3625 and lower bounds of its post-fit uncertainty. The resulting shifts $\Delta_{\text{up/down}}\widehat{\mu}_{\text{Sig}}$
 3626 are shown as filled boxes, where the colour of the box indicates whether μ_{Sig} is
 3627 correlated (blue) or anti-correlated (green) with the value of the NP.

3628 The highest-ranked NPs are associated with systematic JER uncertainties, which
 3629 also receive some of the largest pulls in the fit (see Figure 8.9 and related discussion

3630 in Section 8.3.1). Following in rank are some of the statistical γ_-^* NPs.

3631 8.3.2 Hypothesis Testing and Model Exclusion

3632 Hypothesis testing is performed following the method presented in Section 7.3.3 to
 3633 determine the range of m_s and $m_{Z'}$ that can be confidently excluded by the search.
 3634 The colour map in Figure 8.12 shows the expected CL_s value evaluated over the
 3635 range of m_s and $m_{Z'}$ considered in the search. Interpolated contours are drawn at
 3636 $CL_s = 0.05$ for the expected (grey dashed) and observed (solid red) values, and the
 3637 contained areas represent the expected and observed range of m_s and $m_{Z'}$ that are
 3638 excluded by the search for the assumed DM mass and coupling choices.

3639 Figure 8.13 summarizes the excluded range of m_s and $m_{Z'}$ in the DH model -
 3640 for the fixed choices of the DM mass, DH mixing angle and couplings presented at
 3641 the beginning of this section - by this and all other searches presented in Section
 3642 2.3.2 that place constraints on the model by targeting various decay modes of the
 3643 DH boson s . This search (blue) extends the excluded range for on-shell WW decays
 3644 ($m_s > 160$ GeV). By additionally probing off-shell WW decays in the range $m_s < 160$,
 3645 this search largely closes the pre-existing gap in coverage (150 GeV $< m_s < 160$ GeV)
 3646 between searches in the $s \rightarrow WW$ and $s \rightarrow bb$ DH decay channels.

3647 8.3.3 Dependence of Sensitivity on Signal Strength

3648 Neglecting slight changes to the width of the Z' mediator, the production rate of the
 3649 DH model would be expected to scale in proportion to g_q^2 and g_χ^2 [12] if the values
 3650 of the g_q and g_χ coupling strength parameters, respectively, are varied, since both
 3651 correspond to a single annihilation (g_q) or decay (g_χ) vertex in the model (see Figure
 3652 2.1). As noted in Section 2.2.1, upper bounds are placed on the coupling strength g_q
 3653 between the Z' mediator and quarks in the DH model by dijet searches, which range
 3654 from approximately 0.04 to 0.2 depending on the value of $m_{Z'}$, and on the methods
 3655 the assumptions involved in placing the constraints. However, the actual value of g_χ
 3656 in the model is unknown, and is not currently constrained by dijet or other searches.

3657 Given the dependence of the production rate on the choices of g_q and g_χ , a test
 3658 was done following recommendations in Ref. [173] to evaluate the impact of reducing
 3659 the production rate of the DH process on the sensitivity of the search. Implicit in
 3660 this test is a simplifying assumption that the modified values of g_q and g_χ associated
 3661 with reducing μ will result in the same distributions of kinematic variables as the

benchmark choices, such that the predicted yields of the signal model in all regions and bins of the search scale linearly with μ . Figure D.1 compares the range of m_s and $m_{Z'}$ excluded by the search with the value of a “signal strength” parameter μ , which coherently scales the production rate of the DH signal process at all m_s and $m_{Z'}$ considered in the search. The search is able to exclude phase space in the m_s vs. $m_{Z'}$ plane for μ down to 0.3, with the excluded range successively diminished as μ is reduced.

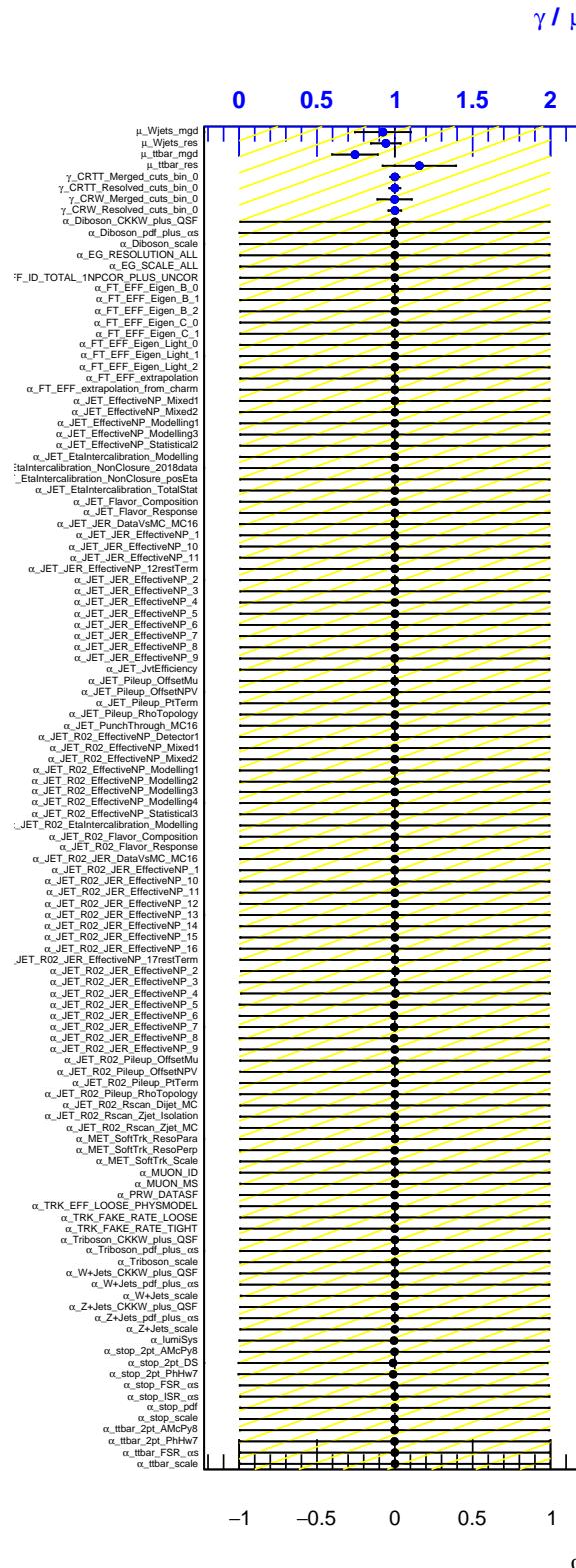


Figure 8.2: Post-fit values and uncertainties of all NPs in the background-only fit. See Tables 7.1, C.1 and C.2 for details of the scheme used to name the NPs. Yellow hatched band shows the pre-fit uncertainty for each NP, and black horizontal error bars show the post-fit uncertainty.

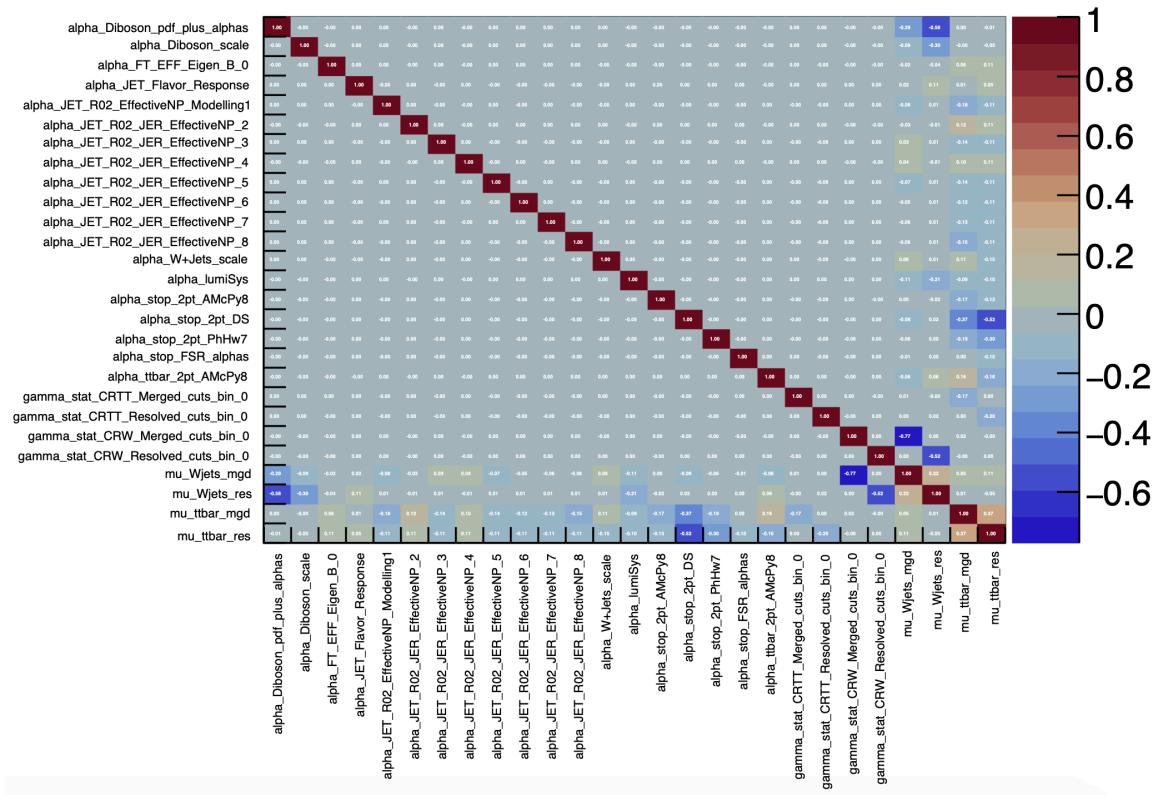


Figure 8.3: Correlation matrix for all NPs considered in the background-only fit for which at least one coefficient of cross-correlation with another NP is larger than 0.1. See Tables 7.1, C.1 and C.2 for details of the scheme used to name the NPs.

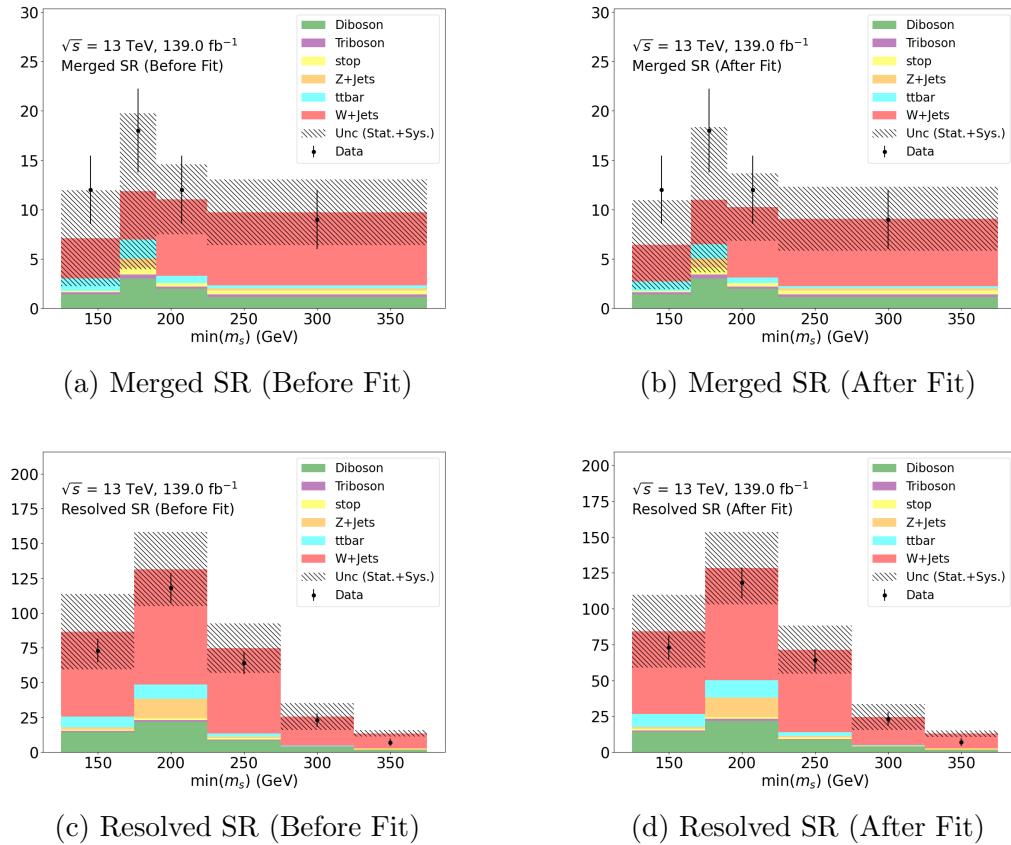


Figure 8.4: Comparison between predicted yields of SM background processes and observed yields in the SRs, before (left) and after (right) the background-only fit and extrapolation to the SR. Yields are binned in $\min(m_s)$ using the binning strategy presented in Section 7.2.

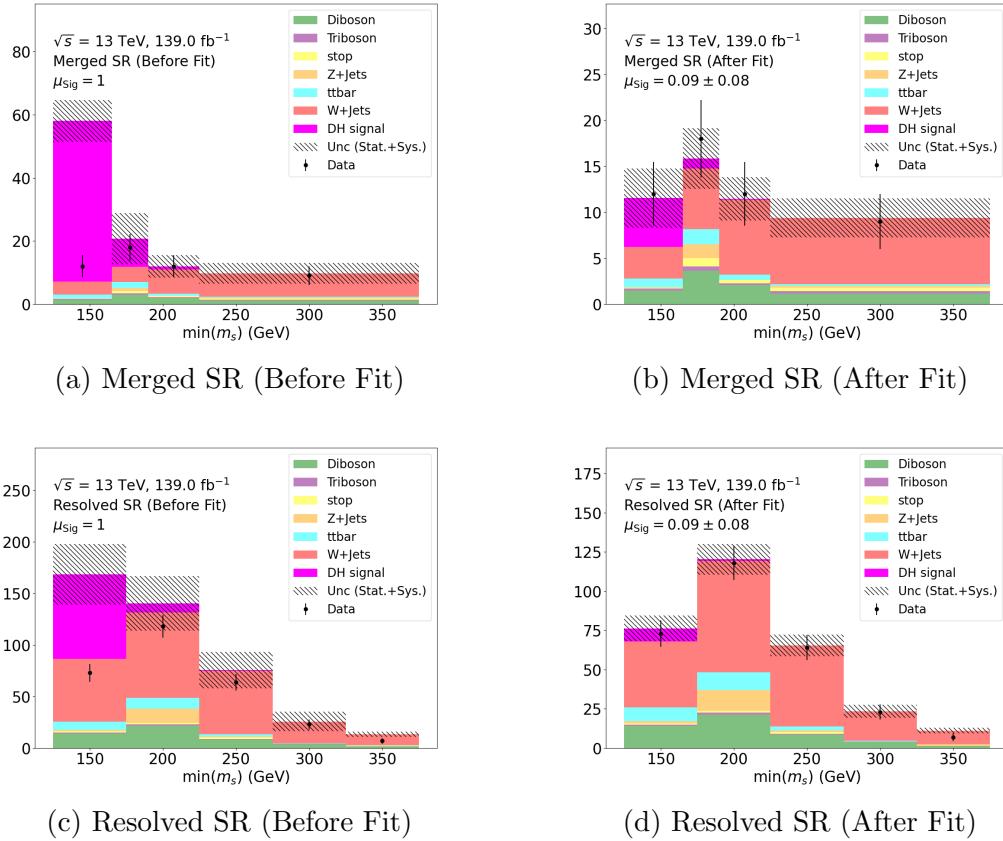


Figure 8.5: Comparison between predicted yields of SM background processes, the DH signal process at $(m_s, m_{Z'}) = (160, 1000)$ GeV and observed yields in the SRs. Yields are shown before (left) and after (right) the signal+background fit. The pre- and post-fit values of the signal strength μ_{Sig} are also reported. Yields are binned in $\min(m_s)$ using the binning strategy presented in Section 7.2.

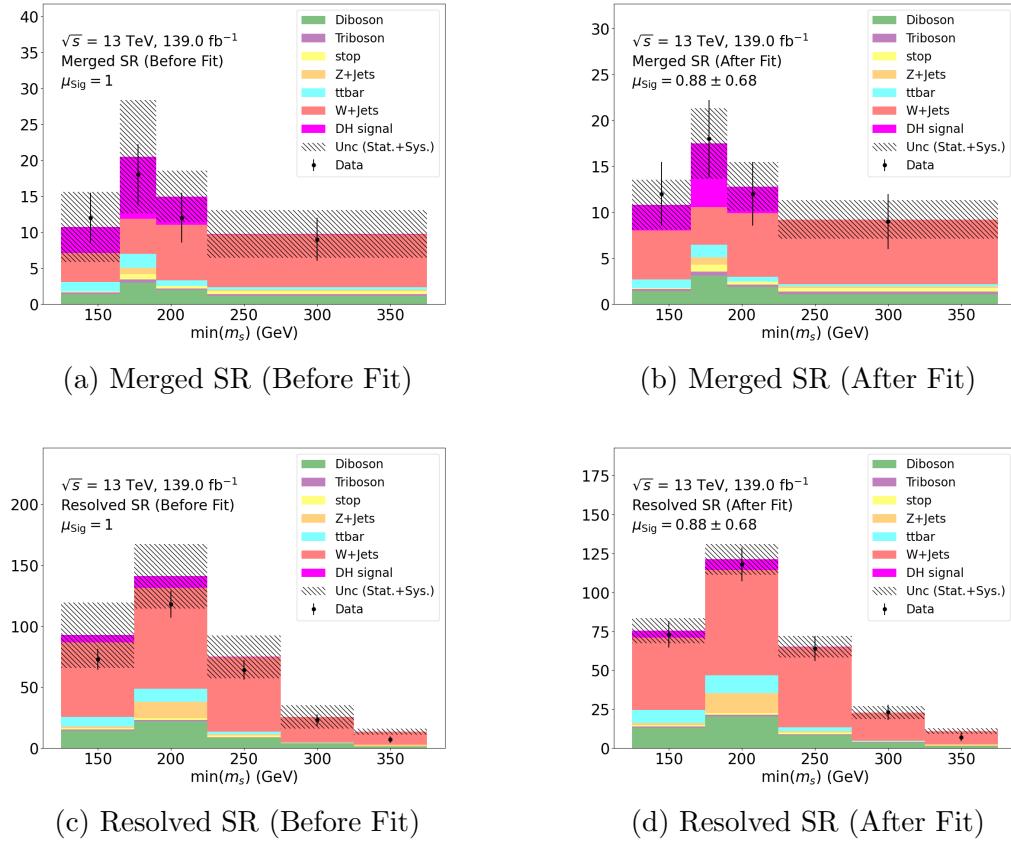


Figure 8.6: Comparison between predicted yields of SM background processes, the DH signal process at $(m_s, m_{Z'}) = (210, 2100)$ GeV and observed yields in the SRs. Yields are shown before (left) and after (right) the signal+background fit. The pre- and post-fit values of the signal strength μ_{Sig} are also reported. Yields are binned in $\min(m_s)$ using the binning strategy presented in Section 7.2.

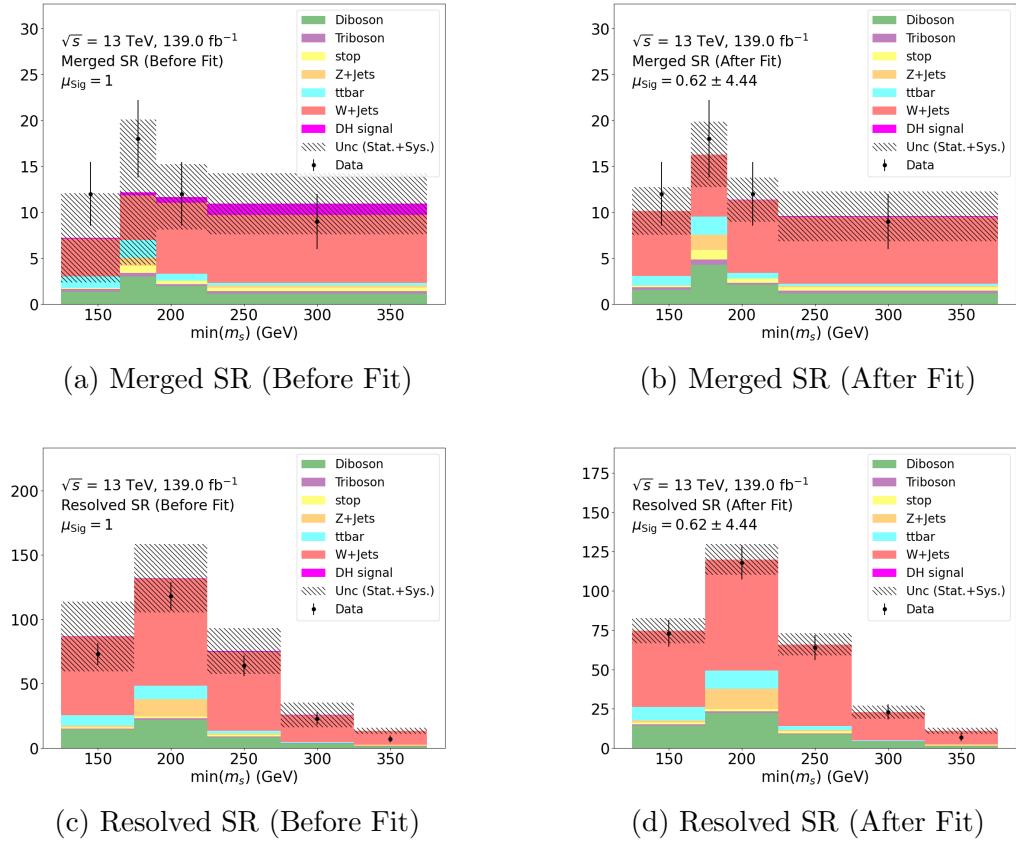


Figure 8.7: Comparison between predicted yields of SM background processes, the DH signal process at $(m_s, m_{Z'}) = (310, 2900)$ GeV and observed yields in the SRs. Yields are shown before (left) and after (right) the signal+background fit. The pre- and post-fit values of the signal strength μ_{Sig} are also reported. Yields are binned in $\min(m_s)$ using the binning strategy presented in Section 7.2.

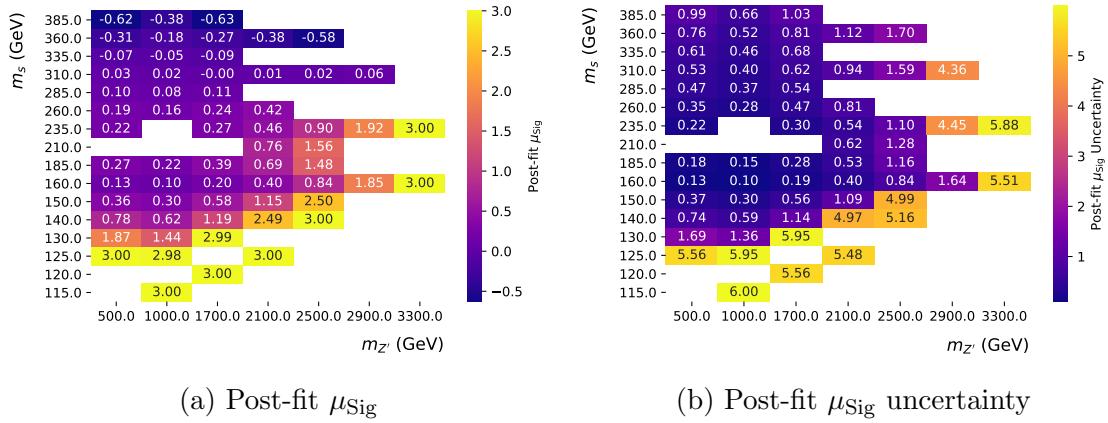


Figure 8.8: Post-fit value (left) and uncertainty (right) of the signal strength parameter μ_{Sig} in the signal+background fit (μ_{Sig} left floating) for each m_s and $m_{Z'}$ in the DH signal model considered in the search.

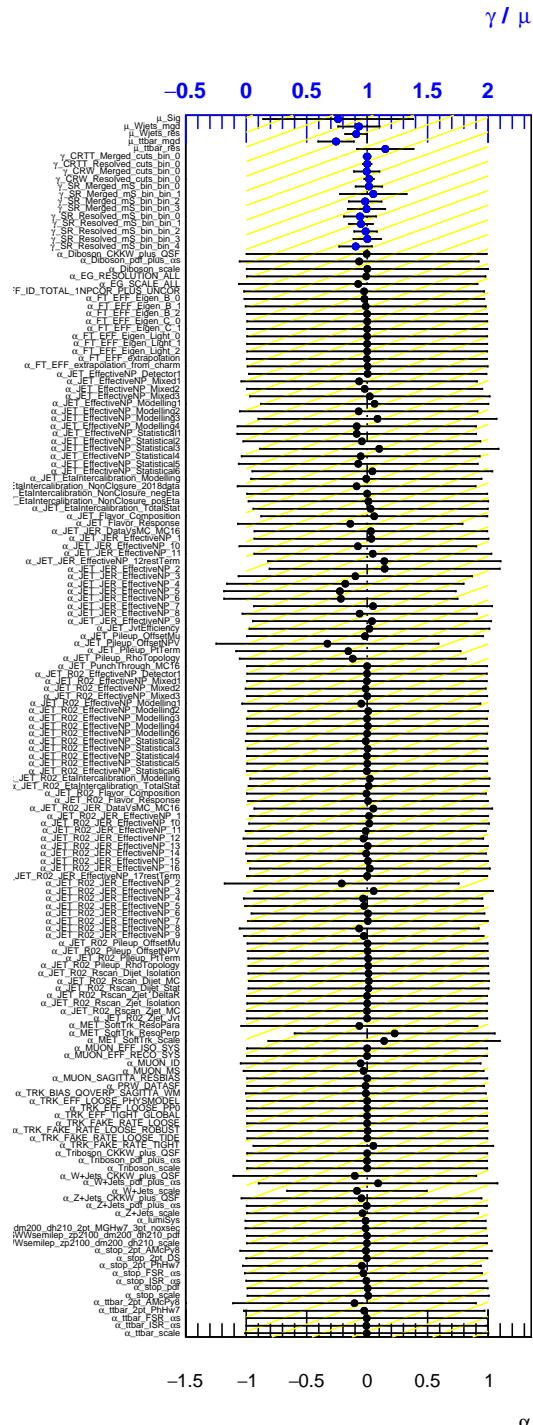


Figure 8.9: Post-fit values of all NPs in the signal+background fit using the DH signal model at $(m_s, m_{Z'}) = (210, 2100)$ GeV. See Tables 7.1, C.1 and C.2 for details of the scheme used to name the NPs. Yellow hatched band shows the pre-fit uncertainty for each NP, and black horizontal error bars show the post-fit uncertainty.

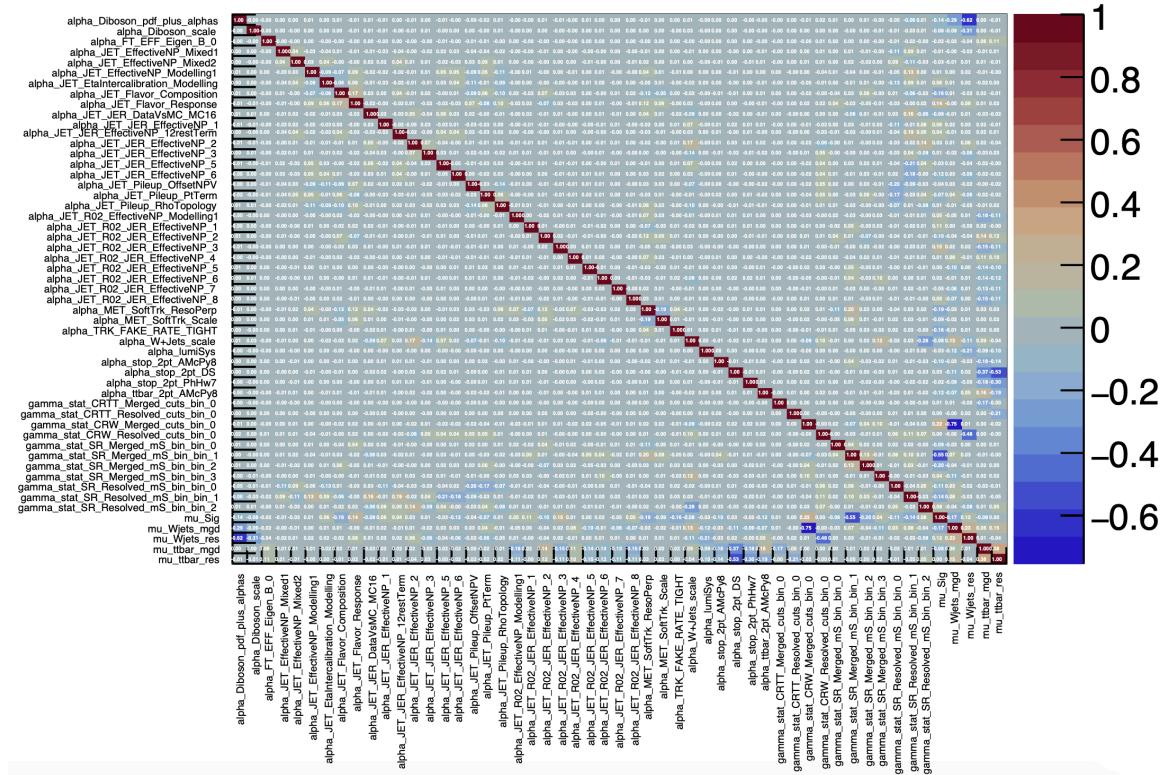


Figure 8.10: Correlation matrix for all NPs considered in the signal+background fit at a sample signal point of $(m_s, m_{Z'}) = (210, 2100)$ GeV, for which at least one coefficient of cross-correlation with another NP is larger than 0.1. See Tables 7.1, C.1 and C.2 for details of the scheme used to name the NPs.

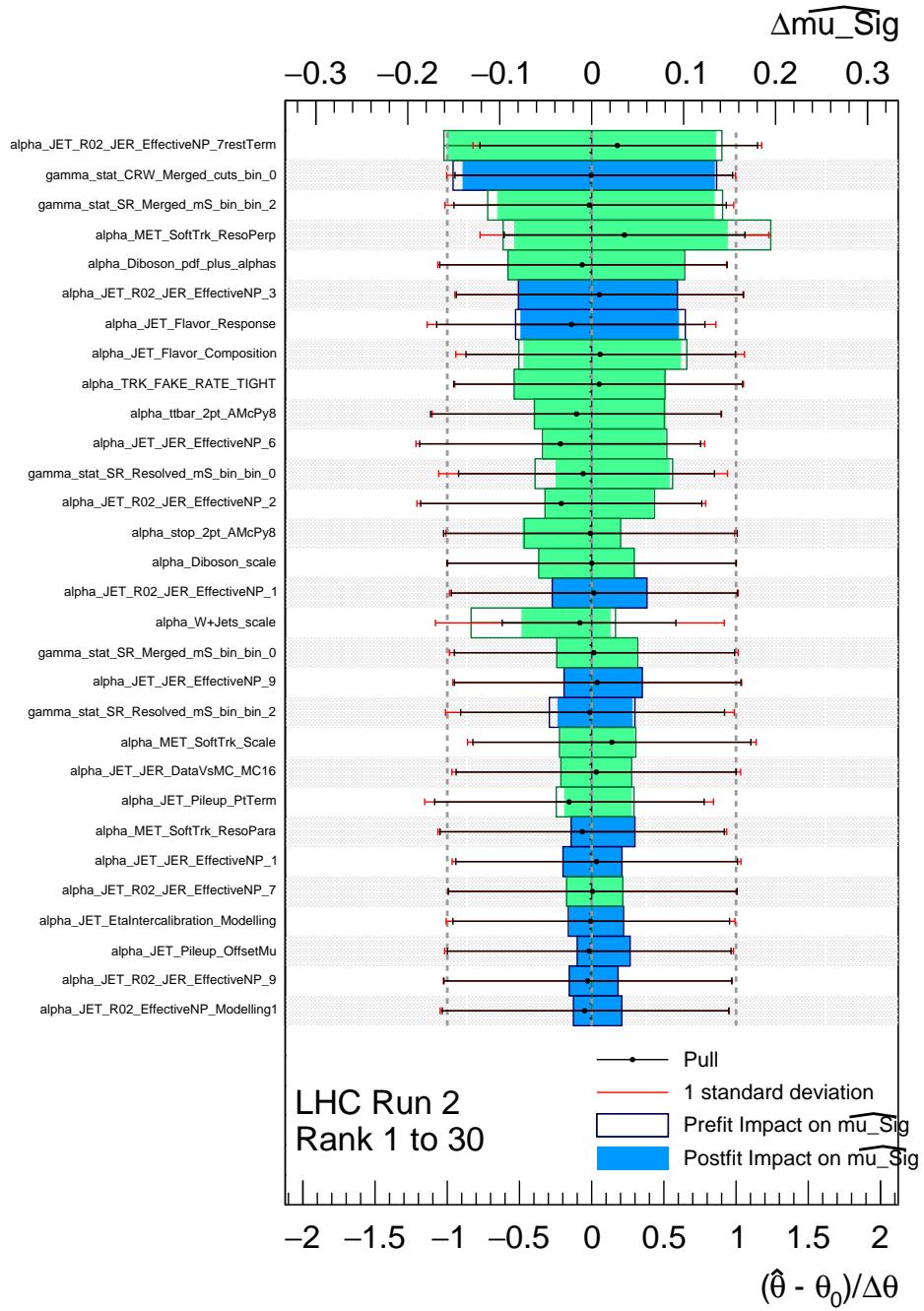


Figure 8.11: Leading 30 pre-and post-fit impacts on μ_{Sig} for NPs associated with experimental and theoretical uncertainties in the signal+background fit at the sample signal point $(m_s, m_{Z'})=(210, 2100)$ GeV. NPs are ranked from top to bottom in order of the size of their impact on μ_{Sig} . Blue (green) colouring of post-fit impacts indicates positive (negative) correlation with the signal strength. Post-fit values of the NPs (a.k.a. “pulls”) are also shown as black circles, where red (black) error bars show the size of the pre-fit (post-fit) uncertainty.

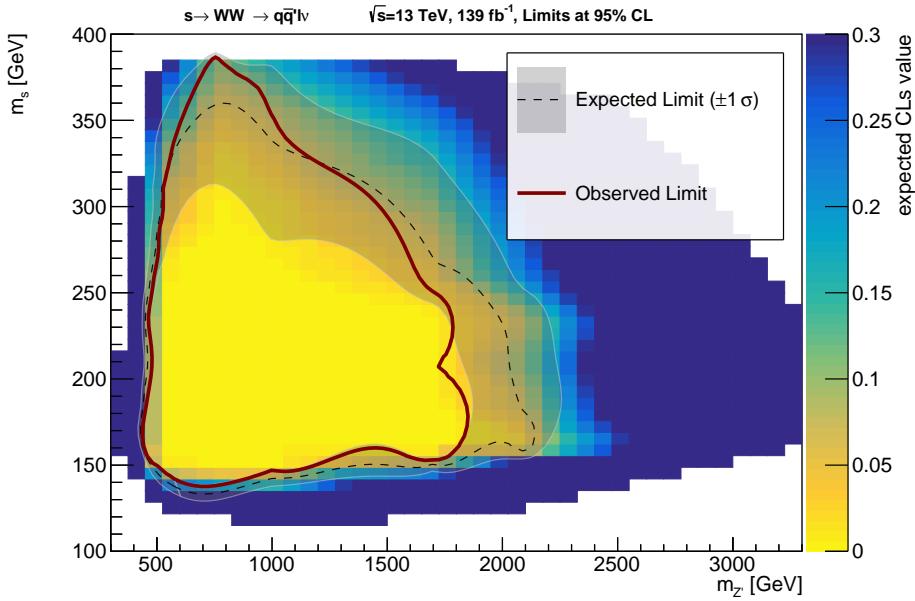


Figure 8.12: Expected (grey dashed with $\pm 1\sigma$ uncertainty band) and observed (solid red) range of m_s and $m_{Z'}$ in the DH model excluded by this search. All m_s and $m_{Z'}$ contained within the solid red line are excluded by the search for the choice of $m_\chi = 200 \text{ GeV}$, $\sin \theta = 0.01$, $g_\chi = 1.0$ and $g_q = 0.25$.

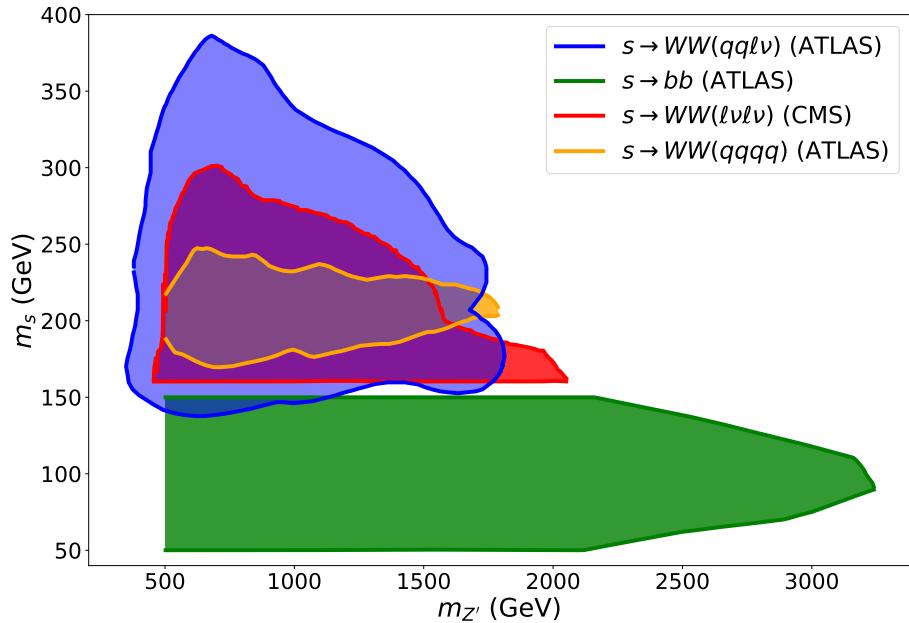


Figure 8.13: Summary of m_s and $m_{Z'}$ parameters in the DH model excluded by all searches for the model by ATLAS and CMS. All values of m_s and $m_{Z'}$ contained within a coloured area are excluded. The range excluded by the search presented in this thesis is shown in blue.

3669

Chapter 9

3670

Conclusion

3671 The study presented in this thesis is part of a worldwide programme to search for
 3672 dark matter using particle physics detectors, and focusses in particular on dark matter
 3673 production at the LHC. Given the potential for yet-unconceived mechanisms by which
 3674 the hypothetical interactions between dark matter and the Standard Model could
 3675 occur, the search programme at the LHC emphasizes a comprehensive coverage of the
 3676 possible final states that could result in the detector from dark matter production in
 3677 the high-energy pp collisions, with searches guided by and interpreted using simplified
 3678 models for the dark matter production mechanisms. While there could be many
 3679 possible dark matter production mechanisms that would predict a signature in the
 3680 $E_T^{\text{miss}} + WW$ final state studied in this thesis, the construction and interpretation
 3681 of the search are guided by the Dark Higgs model [40]. No statistically significant
 3682 deviation was found between distributions of ATLAS collision events in the signal
 3683 regions and Standard Model predictions. The search places exclusion limits on the
 3684 Dark Higgs model for masses of the Dark Higgs mediator in the approximate range
 3685 of 150 GeV to 350 GeV. As shown in the summary plot in Figure 8.13, the parameter
 3686 space of the Dark Higgs model excluded by this search extends the reach of existing
 3687 searches for the model performed by the ATLAS and CMS collaborations [100, 101,
 3688 98], which targeted different final states.

3689 The semileptonic $WW(qq\ell\nu)$ final state studied by this search presented a number
 3690 of opportunities compared with alternative WW decay modes to develop targeted
 3691 data selections and analysis strategies that enhance the sensitivity of the search in
 3692 this final state. The requirement of a single energetic lepton in the final state allows
 3693 for a significant reduction of SM background processes relative to the fully hadronic
 3694 channel, and the $W+\text{jets}$ process that dominates the Standard Model background in
 3695 this semileptonic final state is massively reduced by the application of a lower bound

3696 on the transverse mass between the final-state lepton and the E_T^{miss} . In addition, the
 3697 distinct decay modes of the two W bosons enable a detailed reconstruction of the
 3698 hadronically decaying W . This reconstruction is facilitated in the boosted merged
 3699 regime by a modification to the basic TAR algorithm [2] used to reconstruct hadronic
 3700 activity in the final state within one or more large-radius jets, where the modification
 3701 additionally disentangles the final-state lepton from the hadronic activity. Targeted
 3702 selections involving the reconstructed hadronically decaying W further reduce the
 3703 background of SM processes in the search. Although the final-state neutrino prevents
 3704 a full reconstruction of the Dark Higgs boson, the $\min(m_s)$ strategy allows for an
 3705 approximate reconstruction, which provides valuable shape discrimination between
 3706 the DH signal model and SM background processes in the signal regions.

3707 While the searches that target the WW final state were optimized to probe the
 3708 Dark Higgs model, it is important to acknowledge that appreciable constraints on the
 3709 Dark Higgs model were also obtained in the bb final state [98] by re-interpreting an
 3710 existing dark matter search [102] that targeted the same final state, but which was
 3711 optimized to probe a different model. The impressive sensitivity of the re-interpreted
 3712 search in the bb final state to the Dark Higgs model highlights the value of ensuring
 3713 that searches in this $E_T^{\text{miss}} + WW$ final state can also be re-interpreted in the future
 3714 to constrain any alternative models that may predict a signature in the same final
 3715 state. This search has been preserved for future re-interpretation using the RECAST
 3716 framework [3] developed within the ATLAS collaboration. More generally, given the
 3717 vast multitude of mechanisms by which dark matter could be produced at the LHC,
 3718 and the tremendous amount of human effort and computing resources involved in
 3719 developing searches to probe new final states, it will be important moving forward
 3720 to ensure that all new searches for dark matter can be efficiently re-interpreted to
 3721 constrain alternative models, thus maximizing the potential impact of each search.

3722 Despite longstanding evidence from observational astronomy for the abundance
 3723 of dark matter in the universe, its composition remains one of the open mysteries of
 3724 modern physics. Each time that a new model is tested or new parameter space is
 3725 probed, a collective step is taken towards cracking the mystery of what makes up the
 3726 most abundant form of matter in the universe.

3727

Appendix A

3728

Kinematic Distributions (N-1) in Signal Regions

3729 Figures [A.1](#) and [A.2](#) show N-1 plots with the finalized selections for the variables
3730 listed in Table [5.3](#) that are used to define the analysis regions that are not shown in
3731 Figures [5.10](#) and [A.2](#), respectively, in Section [5.2.3](#). The placements of cuts is also
3732 shown using red arrows.

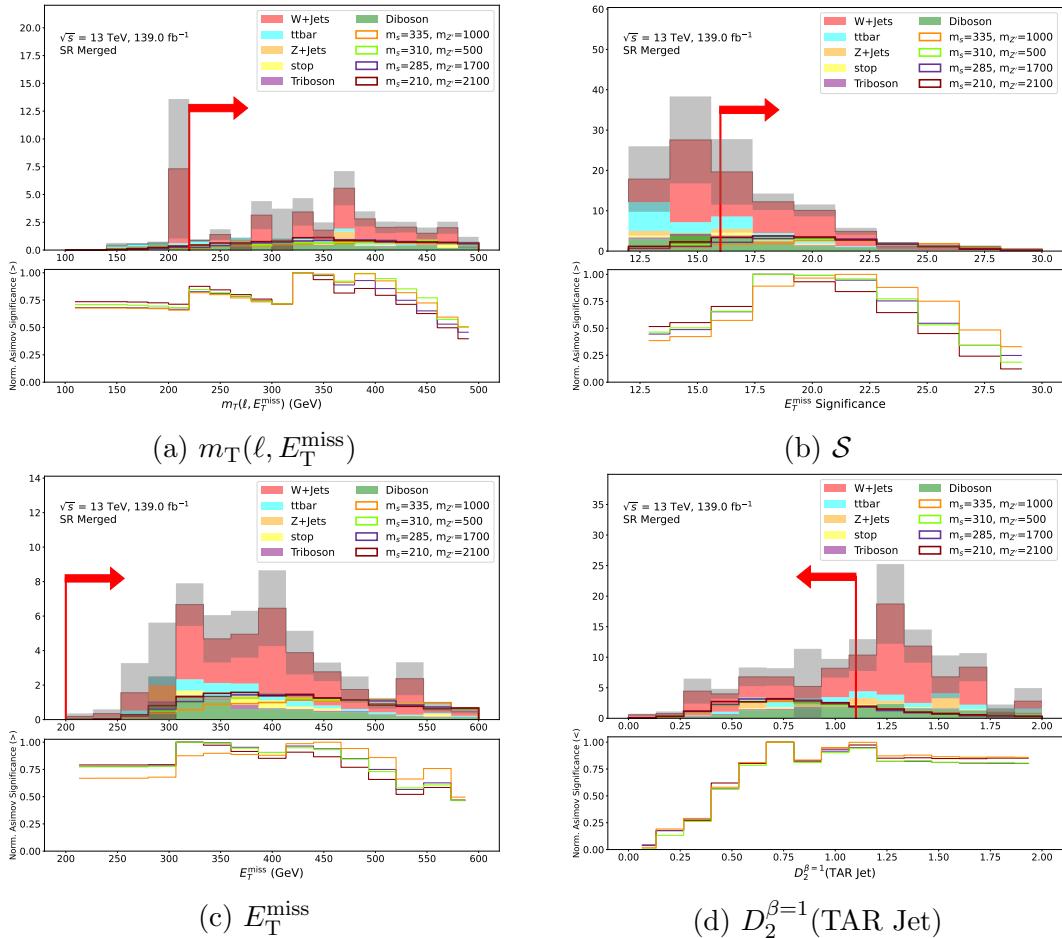


Figure A.1: N-1 Distributions for variables used in the merged signal region definition. Grey bands show statistical uncertainty on background estimate. The lower panel shows the cumulative Asimov significance normalized to unit peak, where the direction ($>$ or $<$) specified in the y label indicates whether the significance is being summed from above ($>$) or from below ($<$). Red vertical line and arrow show placement and direction of selection on the given variable in this region.

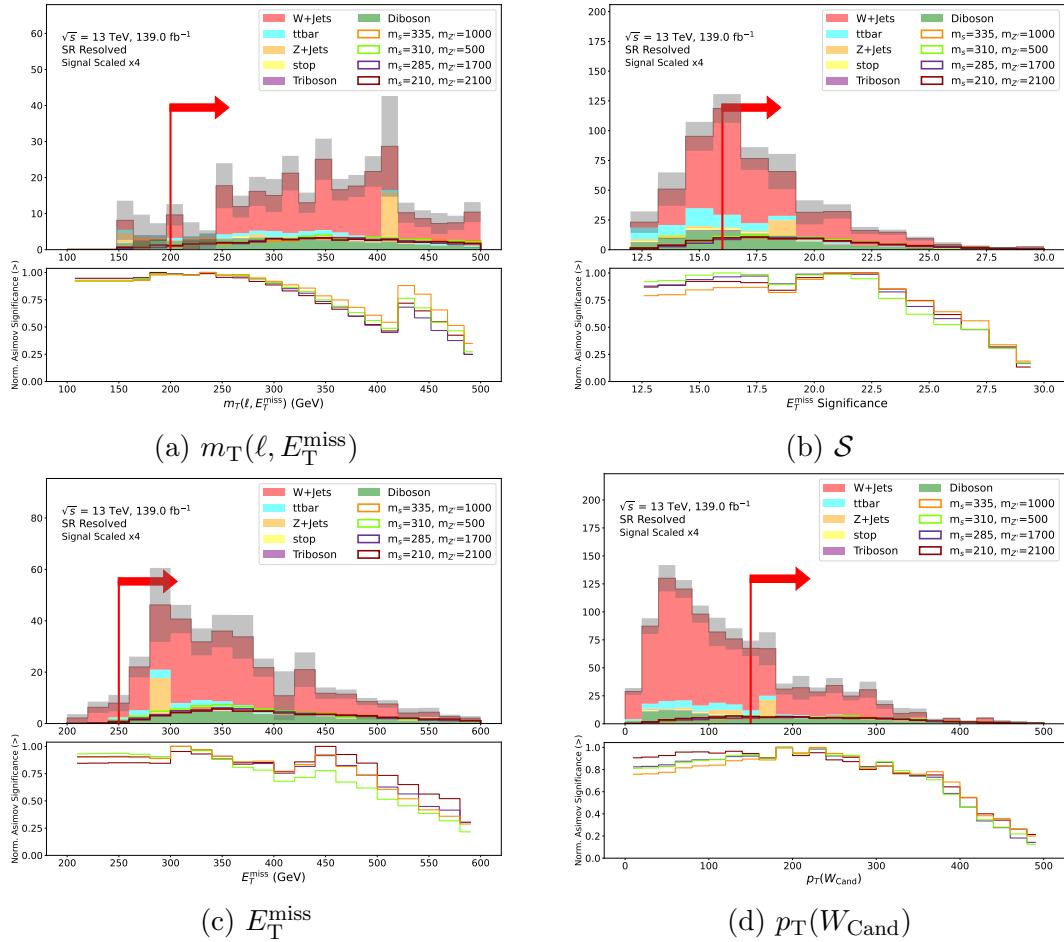


Figure A.2: N-1 Distributions for variables used in the resolved signal region definition. Grey bands show statistical uncertainty on the background estimate. The lower panel shows the cumulative Asimov significance normalized to unit peak, where the direction ($>$ or $<$) specified in the y label indicates whether the significance is being summed from above ($>$) or from below ($<$). Red vertical line and arrow show placement and direction of selection on the given variable in this region.

3733

Appendix B

3734

Kinematic Distributions in Signal and Control Regions

3735

3736 This appendix documents comparisons of the distributions of kinematic variables
 3737 of interest for the search between the signal region (SR) and each control region
 3738 (CR), considering the merged and resolved categories separately. The aim of these
 3739 comparisons is to validate that, within each of the merged and resolved categories,
 3740 the kinematic properties of events in the CRs are sufficiently similar to those in
 3741 the corresponding SR that the data-driven normalization factors for the $W + \text{jets}$
 3742 and $t\bar{t}$ background processes, which are evaluated primarily by comparison with the
 3743 observed yield of ATLAS collision data in the CRs, can be reasonably applied to scale
 3744 the predicted yields of these processes in the SR as well.

3745 **B.1 Signal region vs. $W + \text{jets}$ control region**

3746 Figure B.1 compares distributions of kinematic variables of interest for the analysis
 3747 between the merged SR and the $W + \text{jets}$ CR. Figure B.2 presents the same comparison
 3748 between the resolved SR and the $W + \text{jets}$ CR.

3749 **B.2 Signal region vs. $t\bar{t}$ control region**

3750 Figure B.3 compares distributions of kinematic variables of interest for the analysis
 3751 between the merged SR and the $t\bar{t}$ CR. Figure B.4 presents the same comparisons
 3752 between the resolved SR and the $t\bar{t}$ CR.

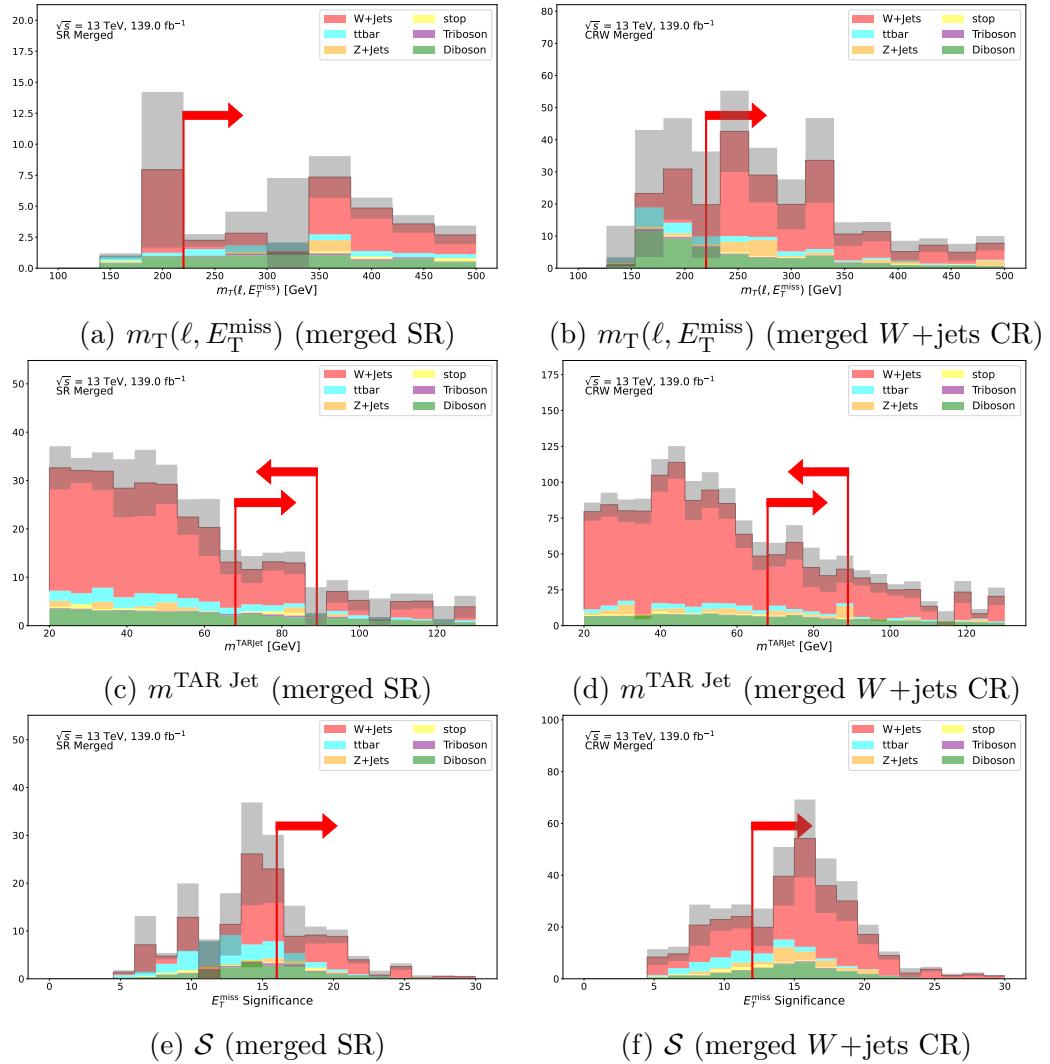


Figure B.1: Comparison of N-1 distributions for kinematic variables of interest between the SR and the $W+\text{jets}$ CR in the merged category. Grey bands show statistical uncertainty on the background estimate.

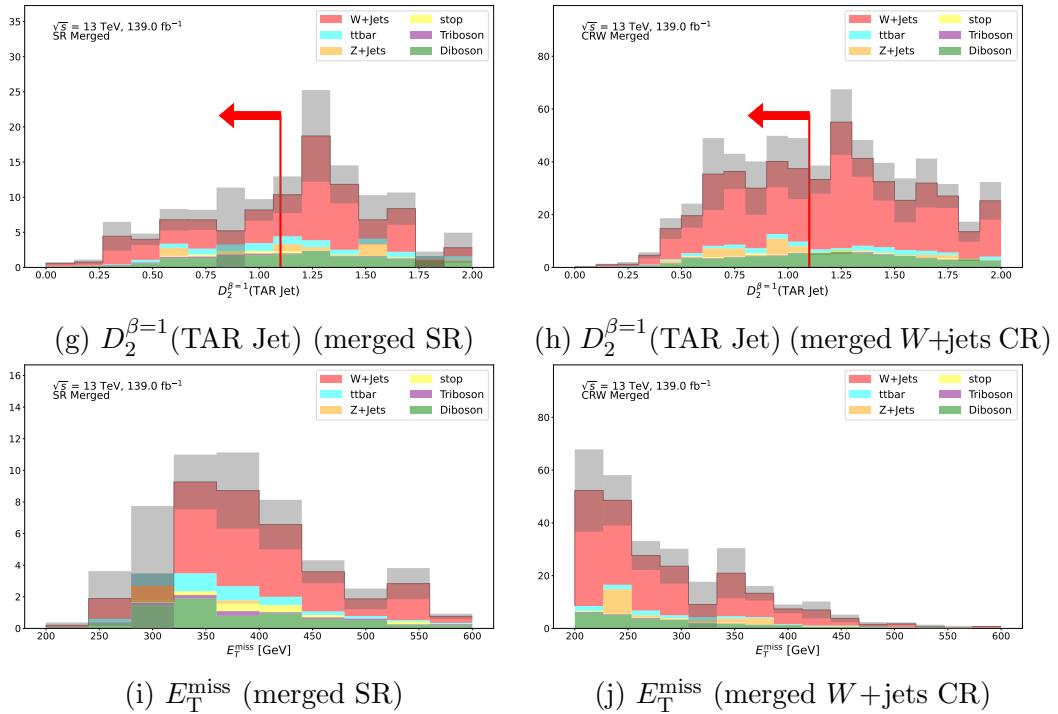


Figure B.1: Comparison of N-1 distributions for kinematic variables of interest between the SR and the $W+\text{jets}$ CR in the merged category (continued).

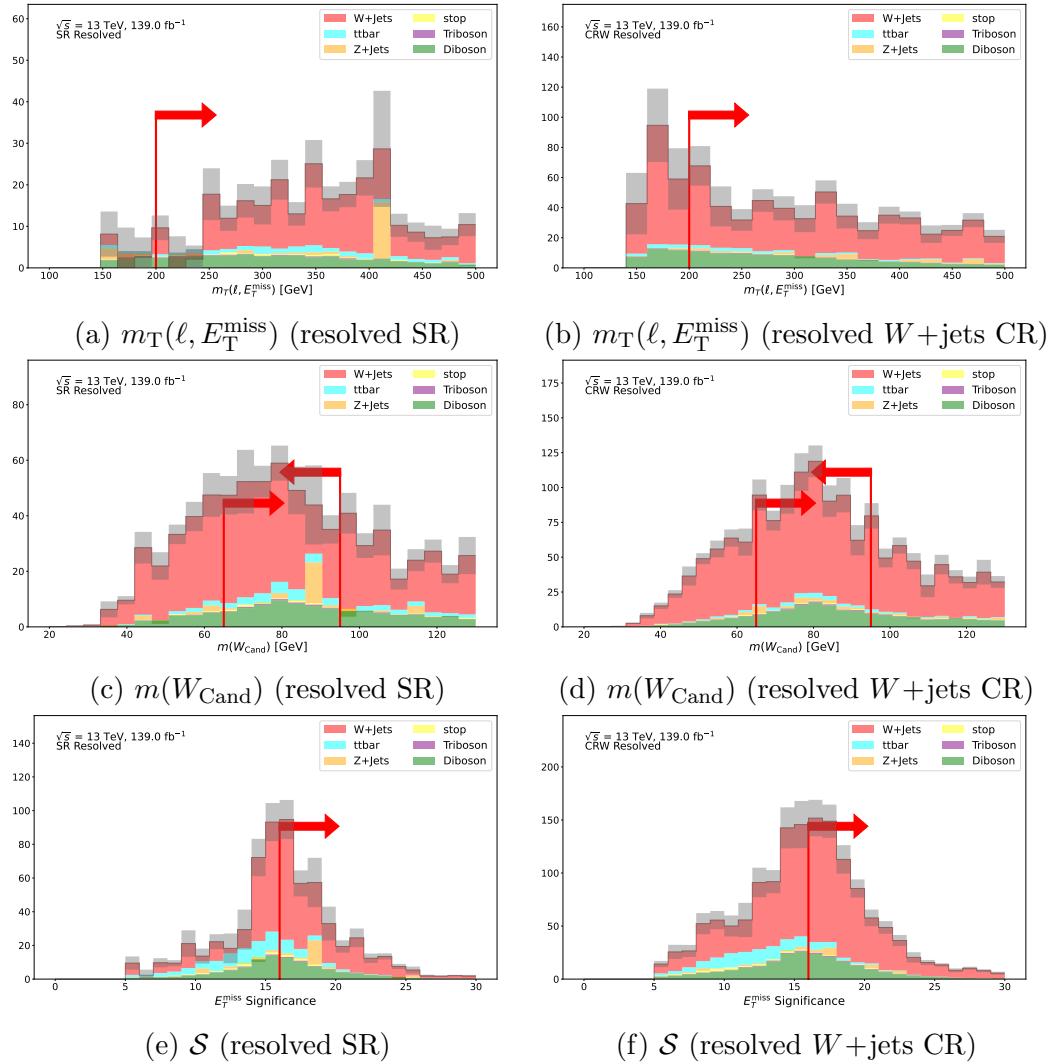


Figure B.2: Comparison of N-1 distributions for kinematic variables of interest between the SR and the $W+\text{jets}$ CR in the resolved category. Grey bands show statistical uncertainty on the background estimate.

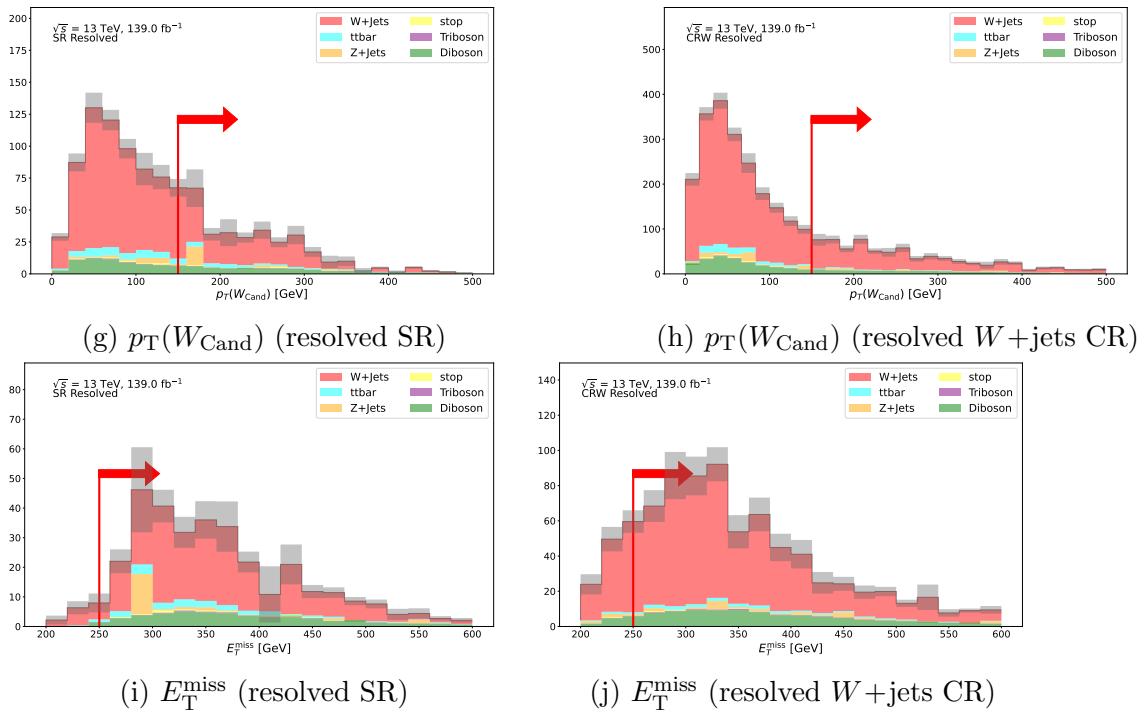


Figure B.2: Comparison of N-1 distributions for kinematic variables of interest between the SR and the $W+\text{jets}$ CR in the resolved category (continued).

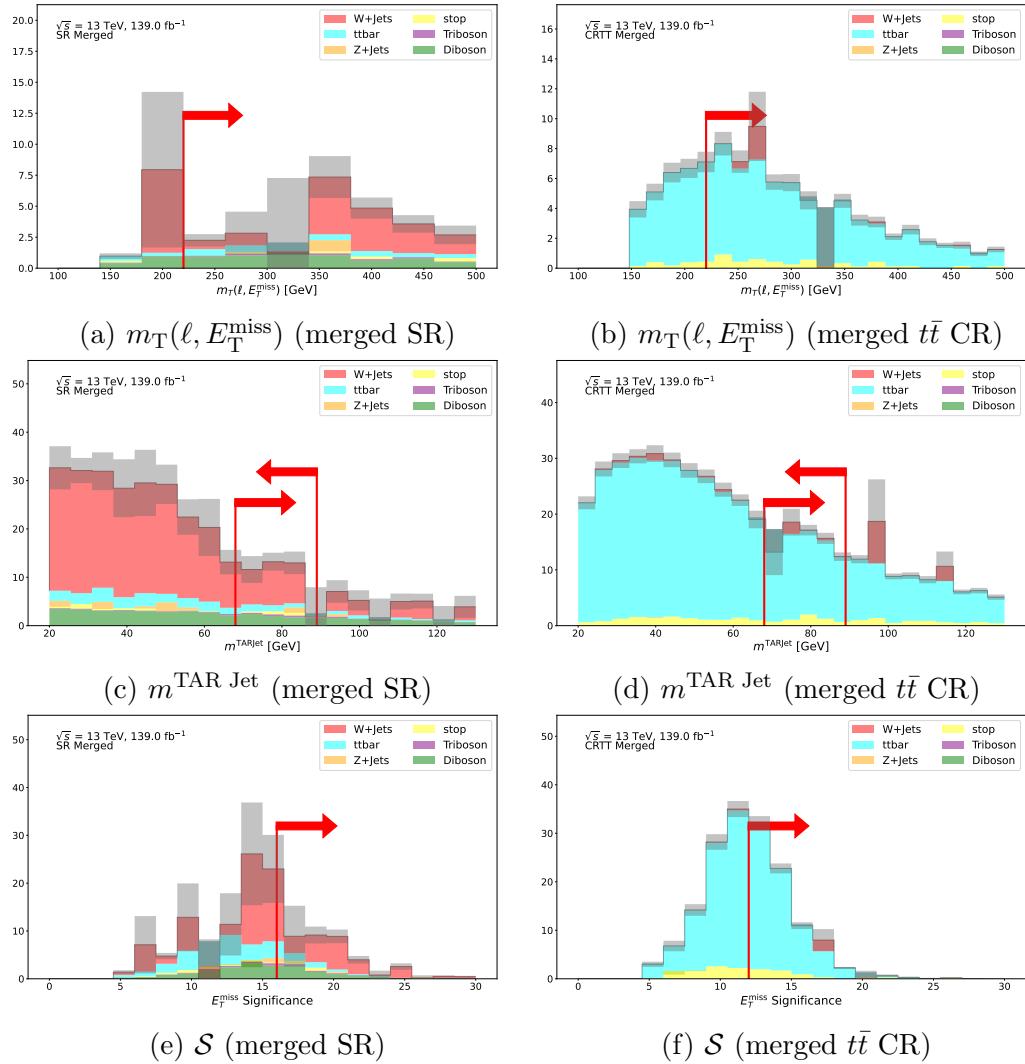


Figure B.3: Comparison of N-1 distributions for kinematic variables of interest between the SR and the $t\bar{t}$ CR in the merged category. Grey bands show statistical uncertainty on the background estimate.

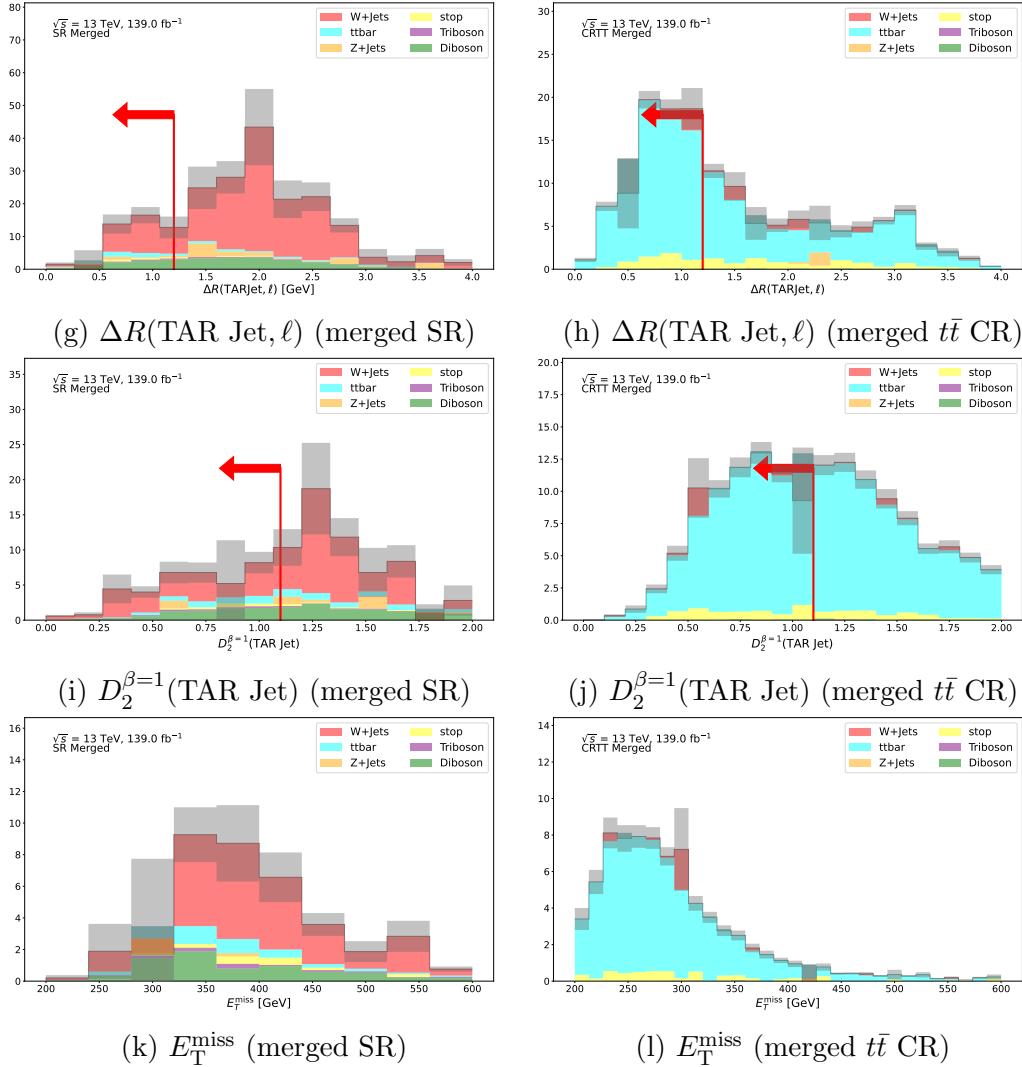


Figure B.3: Comparison of N-1 distributions for kinematic variables of interest between the SR and the $t\bar{t}$ CR in the merged category (continued).

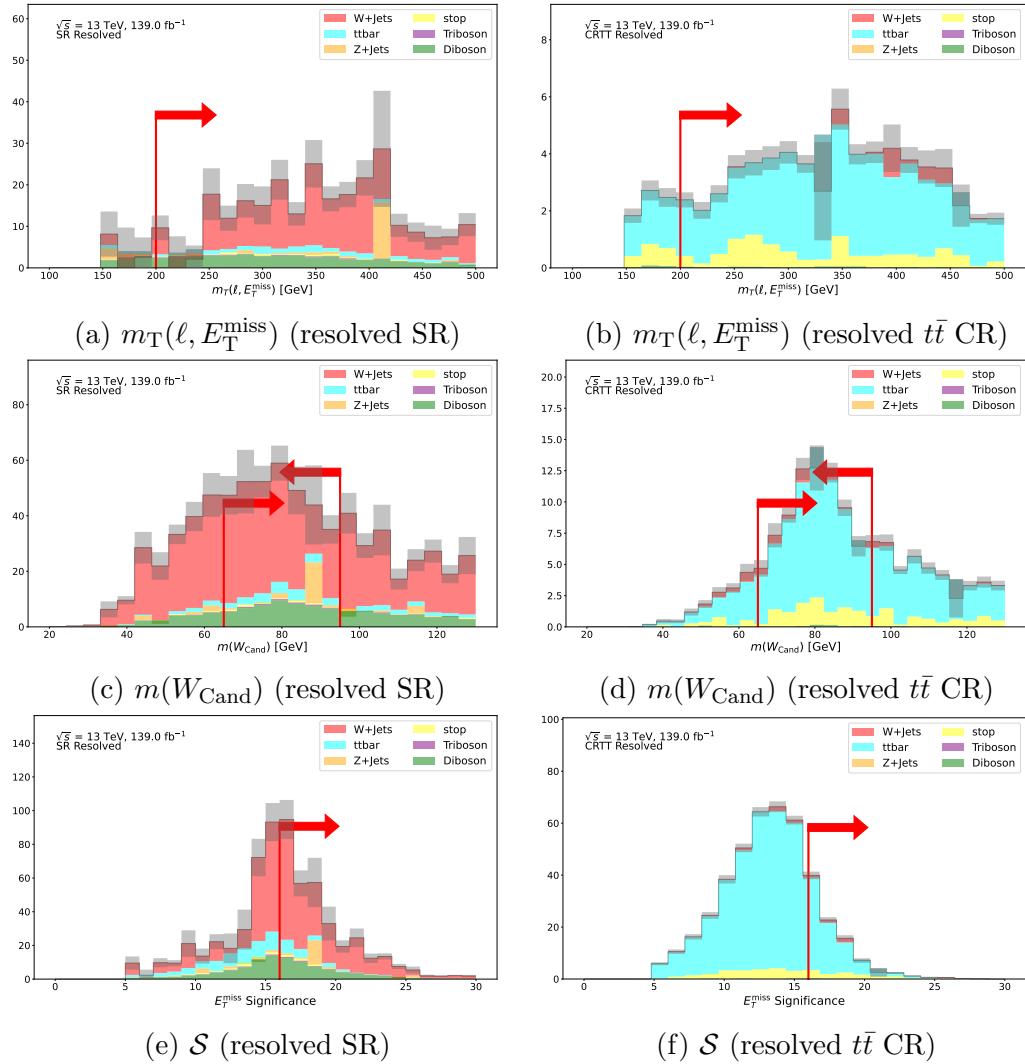


Figure B.4: Comparison of N-1 distributions for kinematic variables of interest between the SR and the $t\bar{t}$ CR in the merged category. Grey bands show statistical uncertainty on the background estimate.

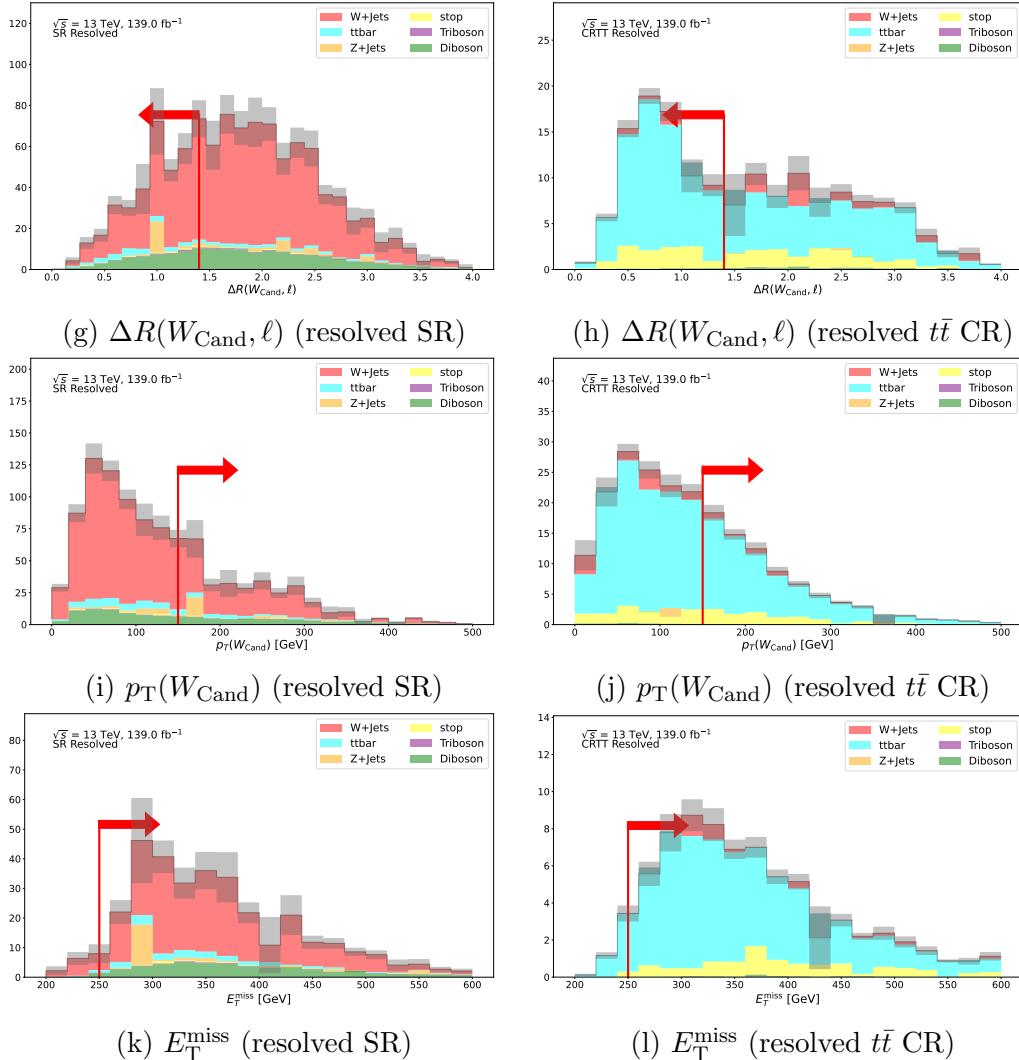


Figure B.4: Comparison of N-1 distributions for kinematic variables of interest between the SR and the $t\bar{t}$ CR in the merged category (continued)

3753

Appendix C

3754

Nuisance Parameter Descriptions for Systematics

3755 Tables C.1 and C.2 provide qualitative descriptions of the NPs used to parametrize
 3756 systematic uncertainties in the likelihood function presented in Section 7.1. Details
 3757 of the systematic uncertainty sources can be found in Chapter 6.

Table C.1: Qualitative description of NPs associated with sources of experimental systematic uncertainty considered in the search.

Nuisance Parameter	Short Description
Event	
α_{lumiSys}	uncertainty on the total integrated luminosity
$\alpha_{\text{PRW_DATASF}}$	pileup reweighting uncertainty
Electrons	
$\alpha_{\text{EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR}}$	reconstruction efficiency uncertainty
$\alpha_{\text{EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR}}$	ID efficiency uncertainty
$\alpha_{\text{EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR}}$	isolation efficiency uncertainty
$\alpha_{\text{EG_SCALE_ALL}}$	energy scale uncertainty
$\alpha_{\text{EG_RESOLUTION_ALL}}$	energy resolution uncertainty
Muons	
$\alpha_{\text{MUON_EFF_TrigSystUncertainty}}$	trigger efficiency uncertainties
$\alpha_{\text{MUON_EFF_TrigStatUncertainty}}$	
$\alpha_{\text{MUON_EFF_RECO_STAT}}$	reconstruction and ID efficiency uncertainty for $p_T > 15$ GeV
$\alpha_{\text{MUON_EFF_RECO_SYS}}$	
$\alpha_{\text{MUON_EFF_RECO_STAT_LOWPT}}$	reconstruction and ID efficiency uncertainty for $p_T < 15$ GeV
$\alpha_{\text{MUON_EFF_RECO_SYS_LOWPT}}$	
$\alpha_{\text{MUON_EFF_ISO_STAT}}$	isolation efficiency uncertainty
$\alpha_{\text{MUON_EFF_ISO_SYS}}$	
$\alpha_{\text{MUON_EFF_TTVA_STAT}}$	track-to-vertex association efficiency uncertainty
$\alpha_{\text{MUON_EFF_TTVA_SYS}}$	
$\alpha_{\text{MUON_SCALE}}$	energy scale uncertainty
$\alpha_{\text{MUON_ID}}$	energy resolution uncertainty from inner detector
$\alpha_{\text{MUON_MS}}$	energy resolution uncertainty from muon system
$\alpha_{\text{MUON_SAGITTA_RESBIAS}}$	uncertainty in the momentum scale

$\alpha_{\text{MUON_SAGITTA_RHO}}$	uncertainty in the momentum scale
anti-k_t $R = 0.4$ jets	
$\alpha_{\text{JET_EffectiveNP_Detector}}$	JES uncertainty: detector effects (2 components)
$\alpha_{\text{JET_EffectiveNP_Mixed}}$	JES uncertainty: mixed effects (3 components)
$\alpha_{\text{JET_EffectiveNP_Modelling}}$	JES uncertainty: modelling effects (5 components)
$\alpha_{\text{JET_EffectiveNP_Statistical}}$	JES uncertainty: statistical uncertainty (6 components)
$\alpha_{\text{JET_EtaIntercalibration_Modelling}}$	
$\alpha_{\text{JET_EtaIntercalibration_NonClosure_2018data}}$	
$\alpha_{\text{JET_EtaIntercalibration_NonClosure_highE}}$	uncertainties in scale calibration of forward / central jets
$\alpha_{\text{JET_EtaIntercalibration_NonClosure_negEta}}$	
$\alpha_{\text{JET_EtaIntercalibration_NonClosure_posEta}}$	
$\alpha_{\text{JET_EtaIntercalibration_TotalStat}}$	
$\alpha_{\text{JET_BJES_Response}}$	
$\alpha_{\text{JET_Flavor_Composition}}$	flavour-related uncertainties
$\alpha_{\text{JET_Flavor_Response}}$	
$\alpha_{\text{JET_JER_EffectiveNP}}$	jet energy resolution uncertainty (12 components)
$\alpha_{\text{JET_JER_DataVsMC_MC16}}$	jet energy resolution uncertainty (data vs. MC)
$\alpha_{\text{JET_JvtEfficiency}}$	jet-vertex-tagger efficiency uncertainty
$\alpha_{\text{JET_Pileup_OffsetMu}}$	
$\alpha_{\text{JET_Pileup_OffsetNPV}}$	Pileup uncertainties
$\alpha_{\text{JET_Pileup_PtTerm}}$	
$\alpha_{\text{JET_Pileup_RhoTopology}}$	
$\alpha_{\text{JET_PunchThrough_MC16}}$	punch through uncertainty
Tagging efficiency (using anti-k_t $R = 0.4$ jets)	
$\alpha_{\text{FT_EFF_EIGEN_B}}$	b -tagging efficiency uncs (medium eigenvector decomp. of flavour tagging uncs)
$\alpha_{\text{FT_EFF_EIGEN_C}}$	3 components for b -jets, 4 for c -jets and 5 for light jets
$\alpha_{\text{FT_EFF_EIGEN_L}}$	
$\alpha_{\text{FT_EFF_EIGEN_extrapolation}}$	b -tagging efficiency uncertainty on the extrapolation on high p_T -jets
$\alpha_{\text{FT_EFF_EIGEN_extrapolation_from_charm}}$	b -tagging efficiency uncertainty on τ -jets
Rscan $R = 0.2$ jets used for TAR jet construction	
$\alpha_{\text{JET_R02_EffectiveNP_Detector}}$	JES uncertainty: detector effects (2 components)
$\alpha_{\text{JET_R02_EffectiveNP_Mixed}}$	JES uncertainty: mixed effects (3 components)
$\alpha_{\text{JET_R02_EffectiveNP_Modelling}}$	JES uncertainty: modelling effects (5 components)
$\alpha_{\text{JET_R02_EffectiveNP_Statistical}}$	JES uncertainty: statistical uncertainty (6 components)
$\alpha_{\text{JET_R02_EtaIntercalibration_Modelling}}$	
$\alpha_{\text{JET_R02_EtaIntercalibration_NonClosure_highE}}$	
$\alpha_{\text{JET_R02_EtaIntercalibration_NonClosure_negEta}}$	uncertainties in scale calibration of forward / central jets
$\alpha_{\text{JET_R02_EtaIntercalibration_NonClosure_posEta}}$	
$\alpha_{\text{JET_R02_EtaIntercalibration_TotalStat}}$	
$\alpha_{\text{JET_R02_BJES_Response}}$	
$\alpha_{\text{JET_R02_Flavor_Composition}}$	flavour-related uncertainty
$\alpha_{\text{JET_R02_Flavor_Response}}$	
$\alpha_{\text{JET_R02_JER_EffectiveNP}}$	jet energy resolution uncertainty (17 components)
$\alpha_{\text{JET_R02_JER_DataVsMC_MC16}}$	jet energy resolution uncertainty (data vs. MC)
$\alpha_{\text{JET_R02_Pileup_OffsetMu}}$	
$\alpha_{\text{JET_R02_Pileup_OffsetNPV}}$	Pileup uncertainty
$\alpha_{\text{JET_R02_Pileup_PtTerm}}$	
$\alpha_{\text{JET_R02_Pileup_RhoTopology}}$	

$\alpha_{\text{JET_R02_PunchThrough_MC16}}$	punch through uncertainty
$\alpha_{\text{JET_R02_SingleParticle_HighPt}}$	High pT term (2012 version)
$\alpha_{\text{JET_R02_Rscan_Dijet_DeltaR}}$	Dijet Direct matching, delta R between Rscan and Ref jets
$\alpha_{\text{JET_R02_Rscan_Dijet_Isolation}}$	Dijet Rscan isolation requirement
$\alpha_{\text{JET_R02_Rscan_Dijet_Jvt}}$	Dijet Rscan LAr JVT
$\alpha_{\text{JET_R02_Rscan_Dijet_MC}}$	Dijet Rscan MC generator difference
$\alpha_{\text{JET_R02_Rscan_Dijet_Stat}}$	Dijet Rscan comb. of stat. components
$\alpha_{\text{JET_R02_Rscan_NonClosure}}$	Non closure observed at low p_T in the Z+jets calibration method
$\alpha_{\text{JET_R02_Rscan_Zjet_DeltaR}}$	Zjet Direct matching, ΔR between Rscan and Ref jets
$\alpha_{\text{JET_R02_Rscan_Zjet_Isolation}}$	Zjet Rscan isolation requirement
$\alpha_{\text{JET_R02_Rscan_Zjet_MC}}$	Zjet Rscan MC generator difference
$\alpha_{\text{JET_R02_Rscan_Zjet_stat}}$	Zjet Rscan comb. of stat. components
$\alpha_{\text{JET_R02_Zjet_Jvt}}$	Zjet Rscan LAr JVT
Tracks used for TAR jet construction	
$\alpha_{\text{TRK_BIAS_D0_WM}}$	d_0 residual alignment tracking uncertainties
$\alpha_{\text{TRK_BIAS_Z0_WM}}$	z_0 residual alignment uncertainties
$\alpha_{\text{TRK_BIAS_QOVERP_SAGITTA_WM}}$	p_T residual alignment tracking uncertainties
$\alpha_{\text{TRK_EFF_LOOSE_GLOBAL}}$	tracking efficiency (loose working point) uncertainty
$\alpha_{\text{TRK_EFF_LOOSE_IBL}}$	tracking efficiency (loose working point) uncertainty
$\alpha_{\text{TRK_EFF_LOOSE_PHYSMODEL}}$	tracking efficiency (loose working point) uncertainty
$\alpha_{\text{TRK_EFF_LOOSE_PP0}}$	tracking efficiency (loose working point) uncertainty
$\alpha_{\text{TRK_EFF_LOOSE_TIDE}}$	tracking in dense environments efficiency (loose working point) uncertainty
$\alpha_{\text{TRK_FAKE_RATE_LOOSE}}$	tracking uncertainties on fake rate
$\alpha_{\text{TRK_FAKE_RATE_LOOSE_ROBUST}}$	tracking uncertainties on fake rate
$\alpha_{\text{TRK_FAKE_RATE_LOOSE_TIDE}}$	tracking uncertainties on fake rate in dense environments
$\alpha_{\text{TRK_RES_D0_DEAD}}$	tracking uncertainties associated with IP d_0 resolution
$\alpha_{\text{TRK_RES_D0_MEAS}}$	tracking uncertainties associated with IP d_0 resolution
$\alpha_{\text{TRK_RES_Z0_DEAD}}$	tracking uncertainties associated with IP z_0 resolution
$\alpha_{\text{TRK_RES_Z0_MEAS}}$	tracking uncertainties associated with IP z_0 resolution
E_T^{miss}	
$\alpha_{\text{MET_SoftTrk_ResoPerp}}$	track-based soft term related to transversal resolution uncertainty
$\alpha_{\text{MET_SoftTrk_ResoPara}}$	track-based soft term related to longitudinal resolution uncertainty
$\alpha_{\text{MET_SoftTrk_Scale}}$	track-based soft term related to longitudinal scale uncertainty
$\alpha_{\text{MET_JetTrk_Scale}}$	track MET scale uncertainty due to tracks in jets

Table C.2: Qualitative description of NPs associated with sources of theoretical systematic uncertainty considered in the search.

Nuisance Parameter	Short Description
$\alpha_{\text{W+jets_scale}}$	QCD scale ($\mu_R + \mu_F$) for $W + \text{jets}$ process
$\alpha_{\text{W+jets_pdf_plus_alphas}}$	combined PDF + α_s for $W + \text{jets}$ process
$\alpha_{\text{W+jets_CKKW_plus_QSF}}$	combined CKKW + QSF PS for $W + \text{jets}$ process
$\alpha_{\text{Z+jets_scale}}$	uncertainty of QCD scale ($\mu_R + \mu_F$) for $Z + \text{jets}$ process
$\alpha_{\text{Z+jets_pdf_plus_alphas}}$	combined PDF + α_s for $Z + \text{jets}$ process

$\alpha_Z + \text{jets_CKKW_plus_QSF}$	combined CKKW + QSF PS for $Z + \text{jets}$ process
$\alpha_{\text{Diboson_pdf_plus_alphas}}$	combined PDF + α_s for diboson process
$\alpha_{\text{Diboson_scale}}$	QCD scale ($\mu_R + \mu_F$) for diboson process
$\alpha_{\text{Diboson_CKKW_plus_QSF}}$	combined CKKW + QSF PS for diboson process
$\alpha_{\text{Triboson_pdf_plus_alphas}}$	combined PDF + α_s for diboson process
$\alpha_{\text{Triboson_scale}}$	QCD scale ($\mu_R + \mu_F$) for diboson process
$\alpha_{\text{Triboson_CKKW_plus_QSF}}$	combined CKKW + QSF PS for diboson process
$\alpha_{\text{ttbar_scale}}$	QCD scale ($\mu_R + \mu_F$) for $t\bar{t}$ process
$\alpha_{\text{ttbar_pdf}}$	PDF for $t\bar{t}$ process
$\alpha_{\text{ttbar_2pt_PhHw7}}$	comparison with alternate PS generator for $t\bar{t}$ process
$\alpha_{\text{ttbar_2pt_AMcPy8}}$	comparison with alternate ME generator for $t\bar{t}$ process
$\alpha_{\text{ttbar_ISR_alphas}}$	ISR α_s for $t\bar{t}$ process
$\alpha_{\text{ttbar_FSR_alphas}}$	FSR α_s ($\mu_R + \mu_F$) for $t\bar{t}$ process
$\alpha_{\text{stop_scale}}$	QCD scale ($\mu_R + \mu_F$) for single top process
$\alpha_{\text{stop_pdf}}$	PDF for single top process
$\alpha_{\text{stop_2pt_PhHw7}}$	comparison with alternate PS generator for single top process
$\alpha_{\text{stop_2pt_AMcPy8}}$	comparison with alternate ME generator for single top process
$\alpha_{\text{stop_ISR_alphas}}$	ISR α_s for single top process
$\alpha_{\text{stop_FSR_alphas}}$	FSR α_s ($\mu_R + \mu_F$) for single top process
$\alpha_{\text{stop_2pt_DS}}$	$t\bar{t}/Wt$ interference at NLO for single top process
$\alpha_{\text{monoSww_*_zp*_dm200.dh*_scale}}$	QCD scale ($\mu_R + \mu_F$) for DH signal process
$\alpha_{\text{monoSww_*_zp*_dm200.dh*_pdf}}$	PDF for DH signal process
$\alpha_{\text{monoSww_*_zp*_dm200.dh*_2pt_MGHw7_3pt_noxsec}}$	comparison with alternate PS generator for DH signal process

3758

Appendix D

3759

Dependence of Sensitivity on Signal Strength

3760 To assess the impact of varying the production rate of the DH signal model on the
3761 sensitivity of the search, Figure D.1 compares the range of m_s and $m_{Z'}$ excluded by
3762 the search with the value of the signal strength parameter μ , which coherently scales
3763 the production rate of the DH signal process at all m_s and $m_{Z'}$ considered in the
3764 search.

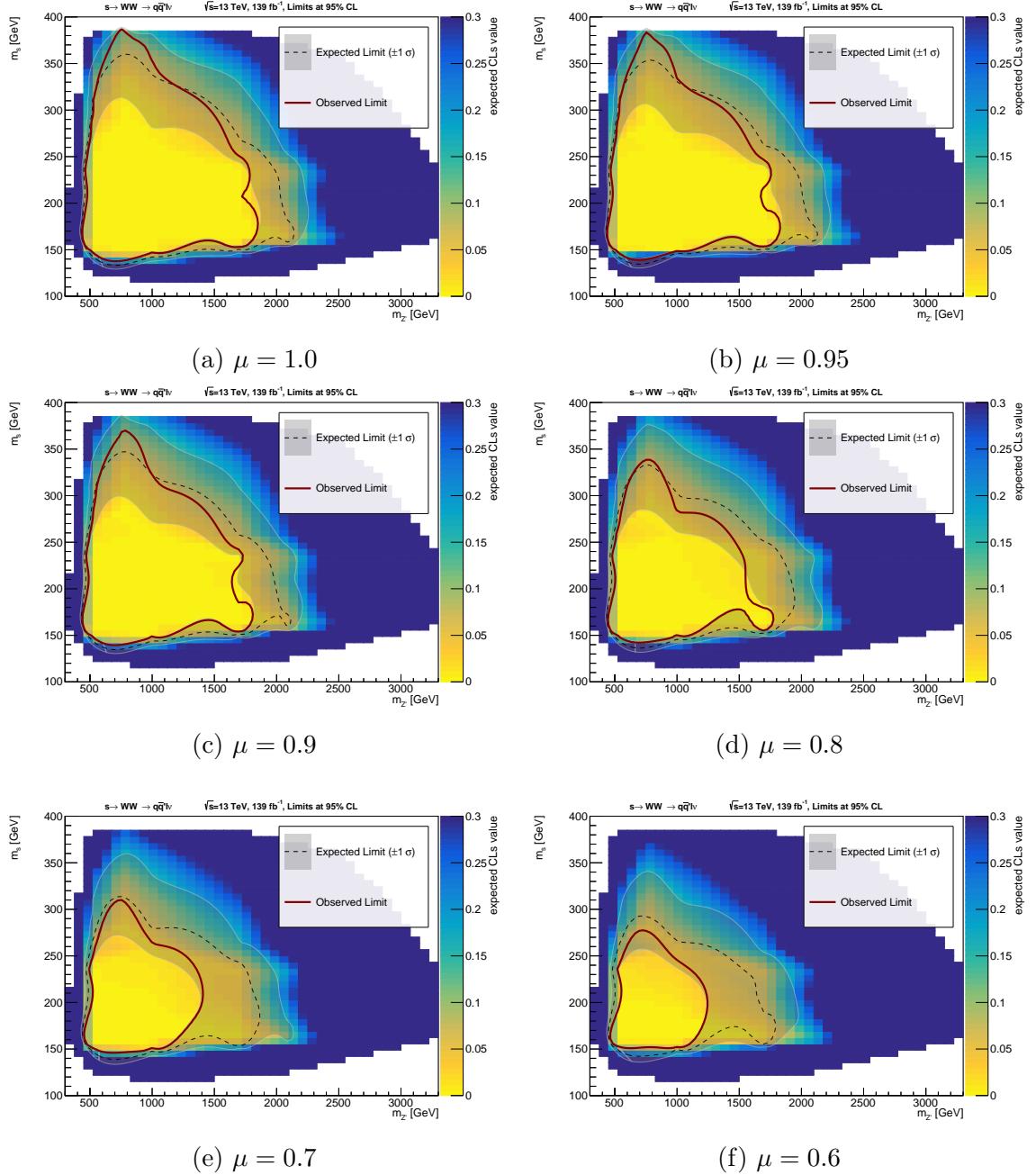


Figure D.1: Range of m_s and $m_{Z'}$ in the DH signal model excluded by the search for various choices of the signal strength μ which coherently scales the production rate of the signal model at all m_s and $m_{Z'}$. Note that the variation with μ assumes changes to coupling combinations yield the same kinematic distributions as the benchmark choices.

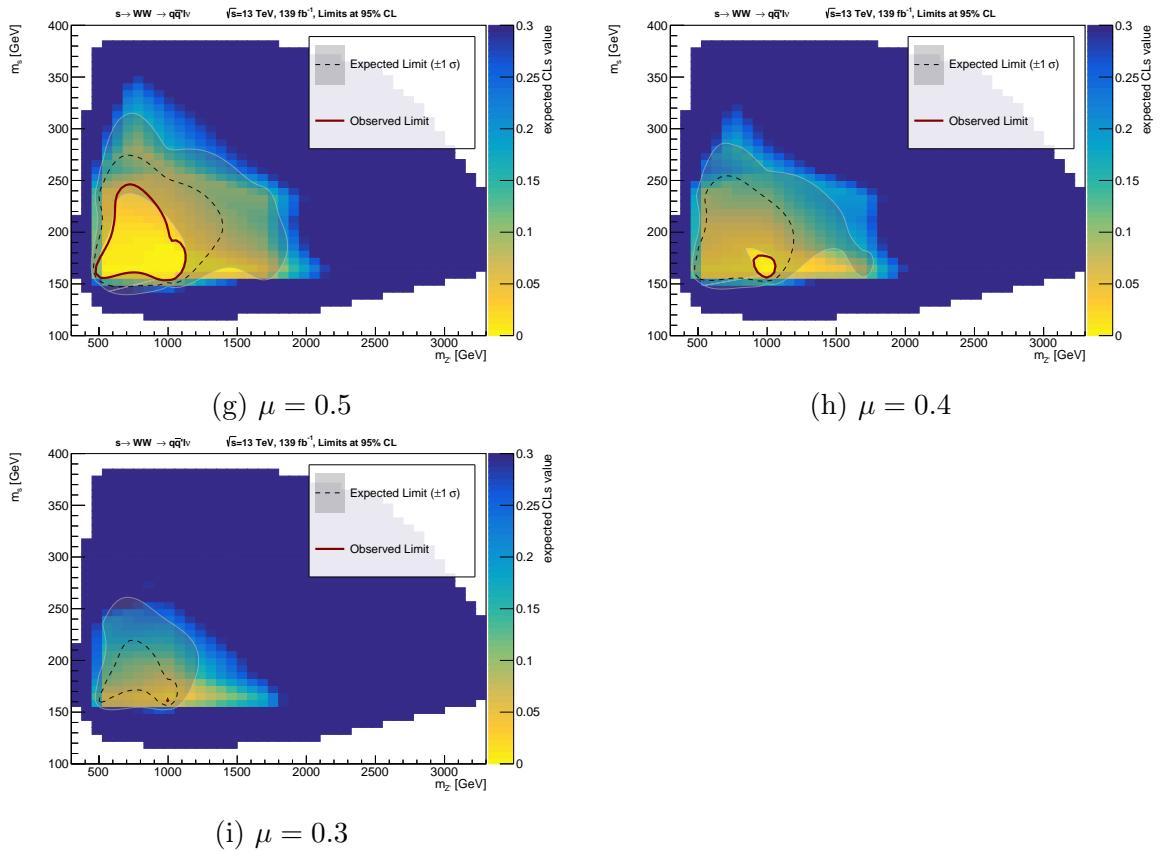


Figure D.1: Range of m_s and $m_{Z'}$ in the DH signal model excluded by the search for various choices of the signal strength μ (continued).

3765

Bibliography

- 3766 [1] M. Baak et al., *HistFitter software framework for statistical data analysis*,
 3767 *Eur. Phys. J. C* **75** (2015) (cit. on pp. v, 141, 142, 149).
- 3768 [2] ATLAS Collaboration, *Track assisted techniques for jet substructure*,
 3769 ATL-PHYS-PUB-2018-012, 2018,
 3770 URL: <https://cds.cern.ch/record/2630864> (cit. on pp. v, 51, 75, 78, 175).
- 3771 [3] K. Cranmer et al., *RECAST - extending the impact of existing analyses*,
 3772 *JHEP* **04** (2011) 038, arXiv: [1010.2506 \[hep-ex\]](https://arxiv.org/abs/1010.2506)
 3773 (cit. on pp. vi, 35, 36, 175).
- 3774 [4] T. Šimko et al., *Scalable Declarative HEP Analysis Workflows for*
 3775 *Containerised Compute Clouds*, *Frontiers in Big Data* **04** (2021)
 3776 (cit. on p. vi).
- 3777 [5] ATLAS Collaboration, *Search for new phenomena in events with an*
 3778 *energetic jet and missing transverse momentum in pp collisions at $\sqrt{s} = 13$*
 3779 *TeV with the ATLAS detector*, *Phys. Rev. D* **103** (11 2021) 112006,
 3780 arXiv: [2102.10874 \[hep-ex\]](https://arxiv.org/abs/2102.10874) (cit. on pp. vi, 22).
- 3781 [6] ATLAS Collaboration, *Observation of a new particle in the search for the*
 3782 *Standard Model Higgs boson with the ATLAS detector at the LHC*,
 3783 *Phys. Lett. B* **716** (1 2012) 1–29, arXiv: [1207.7214 \[hep-ex\]](https://arxiv.org/abs/1207.7214)
 3784 (cit. on pp. 1, 45).
- 3785 [7] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with*
 3786 *the CMS experiment at the LHC*, *Phys. Lett. B* **716** (1 2012) 30–61,
 3787 arXiv: [1207.7235 \[hep-ex\]](https://arxiv.org/abs/1207.7235) (cit. on pp. 1, 45).
- 3788 [8] A. Einstein, *Relativity: The Special and General Theory*,
 3789 New York: Holt, 1921 (cit. on pp. 1, 2).

- 3790 [9] Super-Kamiokande Collaboration,
 3791 *Evidence for Oscillation of Atmospheric Neutrinos*,
 3792 Phys. Rev. Lett. **81** (1998) 1562–1567, arXiv: hep-ex/9807003 [hep-ex]
 3793 (cit. on p. 1).
- 3794 [10] L. Canetti et al., *Matter and antimatter in the universe*,
 3795 New Journal of Physics **14** (2012) 095012, arXiv: 1204.4186 [hep-ph]
 3796 (cit. on p. 1).
- 3797 [11] MissMJ (Wikipedia user), *Standard Model of Elementary Particles*,
 3798 [Online; accessed March 9, 2021], 2021,
 3799 URL: https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg (cit. on p. 2).
- 3800 [12] D. Griffiths, *Introduction to Elementary Particles*,
 3801 Weinheim: Wiley-VCH Verlag GmbH & Co., 2008
 3802 (cit. on pp. 3, 6–9, 14, 161).
- 3803 [13] F. Englert et al., *Broken Symmetry and the Mass of Gauge Vector Mesons*,
 3804 Phys. Rev. Lett. **13** (9 1964) 321–323 (cit. on pp. 4, 9, 23).
- 3805 [14] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*,
 3806 Phys. Rev. Lett. **13** (16 1964) 508–509 (cit. on pp. 4, 9, 23).
- 3807 [15] G. S. Guralnik et al., *Global Conservation Laws and Massless Particles*,
 3808 Phys. Rev. Lett. **13** (20 1964) 585–587 (cit. on pp. 4, 9, 23).
- 3809 [16] G. Breit et al., *Capture of Slow Neutrons*, Phys. Rev. **49** (7 1936) 519–531
 3810 (cit. on p. 5).
- 3811 [17] M. Vogel,
 3812 *Concepts in particle physics: a concise introduction to the standard model*,
 3813 Contemporary Physics **59** (2018) 327–328 (cit. on p. 7).
- 3814 [18] G. Costa et al., *Symmetries and group theory in particle physics: An
 3815 introduction to space-time and internal symmetries*, vol. 823,
 3816 Springer Science & Business Media, 2012 (cit. on p. 8).
- 3817 [19] W. Greiner et al., *Quantum Chromodynamics*, Berlin: Springer-Verlag, 2007
 3818 (cit. on p. 8).
- 3819 [20] A. Pich, *The Standard Model of Electroweak Interactions*, 2012,
 3820 arXiv: 1201.0537 [hep-ph] (cit. on p. 9).

- 3822 [21] C. N. Yang et al.,
 3823 *Conservation of Isotopic Spin and Isotopic Gauge Invariance*,
 3824 *Phys. Rev.* **96** (1 1954) 191–195 (cit. on p. 9).
- 3825 [22] P. Zyla et al., *Review of Particle Physics*, *PTEP* **2020** (8 2020)
 3826 (cit. on pp. 9, 12, 13, 15, 17, 30, 39, 68, 72, 73, 85, 119, 128, 131).
- 3827 [23] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (21 1967) 1264–1266
 3828 (cit. on p. 10).
- 3829 [24] V. C. Rubin et al., *Extended rotation curves of high-luminosity spiral*
 3830 *galaxies. IV. Systematic dynamical properties, Sa → Sc.*,
 3831 *Astrophys. J.* **225** (1978) L107–L111 (cit. on p. 11).
- 3832 [25] L. Bergström,
 3833 *Non-baryonic dark matter: observational evidence and detection methods*,
 3834 *Rep. Prog. Phys.* **63** (2000) 793–841, arXiv: [hep-ph/0002126](#) [hep-ph]
 3835 (cit. on p. 11).
- 3836 [26] M. Persic et al., *Rotation Curves of 967 Spiral Galaxies*,
 3837 *Astrophys. J.* **99** (1995) 501, arXiv: [astro-ph/9502091](#) [astro-ph]
 3838 (cit. on pp. 11, 12).
- 3839 [27] M. Milgrom, *A modification of the Newtonian dynamics as a possible*
 3840 *alternative to the hidden mass hypothesis.*, *Astrophys. J.* **270** (1983) 365–370
 3841 (cit. on p. 12).
- 3842 [28] W. H. Tucker et al., *A Search for “Failed Clusters” of Galaxies*,
 3843 *Astrophys. J.* **444** (1995) 532 (cit. on p. 12).
- 3844 [29] D. Landry et al., *Chandra measurements of a complete sample of X-ray*
 3845 *luminous galaxy clusters: the gas mass fraction*,
 3846 *Mon. Not. R. Astron. Soc.* **433** (2013) 2790–2811,
 3847 arXiv: [1211.4626](#) [astro-ph.CO] (cit. on p. 12).
- 3848 [30] J.-P. Uzan, *The big-bang theory: construction, evolution and status*, 2016,
 3849 arXiv: [1606.06112](#) [astro-ph.CO] (cit. on p. 12).
- 3850 [31] A. A. Penzias et al.,
 3851 *A Measurement of Excess Antenna Temperature at 4080 Mc/s.*,
 3852 *Astrophys. J.* **142** (1965) 419–421 (cit. on p. 13).

- 3853 [32] Planck Collaboration, *Planck 2018 results*, *Astron. Astrophys.* **641** (2020),
 3854 arXiv: [1807.06209 \[astro-ph.CO\]](#) (cit. on pp. 13, 14).
- 3855 [33] B. Famaey et al., *Modified Newtonian Dynamics (MOND): Observational*
 3856 *Phenomenology and Relativistic Extensions*, *Living Rev. Relativ.* **15** (2012)
 3857 (cit. on p. 13).
- 3858 [34] C. Skordis et al., *New Relativistic Theory for Modified Newtonian Dynamics*,
 3859 *Phys. Rev. Lett.* **127** (16 2021) 161302, arXiv: [2007.00082 \[astro-ph.CO\]](#)
 3860 (cit. on p. 13).
- 3861 [35] S. D. M. White et al., *Clustering in a neutrino-dominated universe*,
 3862 *Astrophys. J.* **274** (1983) L1–L5 (cit. on p. 14).
- 3863 [36] H. Baer et al., *Dark matter production in the early Universe: Beyond the*
 3864 *thermal WIMP paradigm*, *Physics Reports* **555** (2015) 1–60,
 3865 arXiv: [1407.0017 \[hep-ph\]](#) (cit. on p. 15).
- 3866 [37] L. Roszkowski et al., *WIMP dark matter candidates and searches – current*
 3867 *status and future prospects*, *Rep. Prog. Phys.* **81** (2018) 066201,
 3868 arXiv: [1707.06277 \[hep-ph\]](#) (cit. on p. 15).
- 3869 [38] M. Pospelov et al., *Secluded WIMP dark matter*,
 3870 *Phys. Lett. B* **662** (1 2008) 53–61, arXiv: [0711.4866 \[hep-ph\]](#)
 3871 (cit. on p. 15).
- 3872 [39] M. Duerr et al., *How to save the WIMP: global analysis of a dark matter*
 3873 *model with two s-channel mediators*, *JHEP* **09** (2016) 042,
 3874 arXiv: [1606.07609 \[hep-ph\]](#) (cit. on pp. 15, 26, 28–30).
- 3875 [40] M. Duerr et al., *Hunting the dark Higgs*, *JHEP* **04** (2017) 143,
 3876 arXiv: [1701.08780 \[hep-ph\]](#) (cit. on pp. 15, 22, 23, 29, 30, 174).
- 3877 [41] J. Billard et al.,
 3878 *Direct Detection of Dark Matter – APPEC Committee Report*, 2021,
 3879 arXiv: [2104.07634 \[hep-ex\]](#) (cit. on pp. 15–17).
- 3880 [42] M. Cirelli, *Indirect searches for dark matter*, *Pramana* **79** (2012) 1021–1043,
 3881 arXiv: [1912.12739 \[hep-ph\]](#) (cit. on pp. 15, 17).
- 3882 [43] A. Boveia et al., *Dark Matter Searches at Colliders*,
 3883 *Annu. Rev. Nucl. Part. Sci.* **68** (2018) 429–459 (cit. on pp. 16, 17).

- 3884 [44] I. Lawson et al.,
 3885 *The SNOLAB Deep Underground Research Facility and Its Science Program*,
 3886 *Nuclear Physics News* **23** (2013) 5–9 (cit. on p. 16).
- 3887 [45] D. Baxter et al., *Recommended conventions for reporting results from direct*
 3888 *dark matter searches*, *Eur. Phys. J. C* **81** (2021),
 3889 arXiv: [2105.00599 \[hep-ex\]](#) (cit. on p. 16).
- 3890 [46] M. Ibe et al., *Migdal effect in dark matter direct detection experiments*,
 3891 *JHEP* **03** (2018) 003, arXiv: [1707.07258 \[hep-ph\]](#) (cit. on p. 17).
- 3892 [47] J. Billard et al.,
 3893 *Assessing the discovery potential of directional detection of dark matter*,
 3894 *Phys. Rev. D* **85** (2012), arXiv: [1110.6079 \[astro-ph.CO\]](#) (cit. on p. 17).
- 3895 [48] J. Billard et al., *Implication of neutrino backgrounds on the reach of next*
 3896 *generation dark matter direct detection experiments*, *Phys. Rev. D* **89** (2014),
 3897 arXiv: [1307.5458 \[hep-ph\]](#) (cit. on p. 17).
- 3898 [49] J. Conrad, *Indirect Detection of WIMP Dark Matter: a compact review*,
 3899 (2014), arXiv: [1411.1925 \[hep-ph\]](#) (cit. on p. 17).
- 3900 [50] N. ArkaniHamed et al.,
 3901 *The hierarchy problem and new dimensions at a millimeter*,
 3902 *Phys. Lett. B* **429** (1998) 263, arXiv: [hep-ph/9803315 \[hep-ph\]](#)
 3903 (cit. on p. 18).
- 3904 [51] P. Brax et al., *LHC signatures of scalar dark energy*, *Phys. Rev. D* **94** (2016),
 3905 arXiv: [1604.04299 \[hep-ph\]](#) (cit. on p. 18).
- 3906 [52] L. Evans et al., *LHC Machine*, *JINST* **03** (2008) S08001 (cit. on pp. 19, 40).
- 3907 [53] ATLAS Collaboration,
 3908 *The ATLAS Experiment at the CERN Large Hadron Collider*,
 3909 *JINST* **03** (2008) S08003 (cit. on pp. 19, 43, 44, 46, 49, 52).
- 3910 [54] CMS Collaboration, *The CMS experiment at the CERN LHC*,
 3911 *JINST* **03** (2008) S08004 (cit. on pp. 19, 43).
- 3912 [55] LHCb Collaboration, *The LHCb Detector at the LHC*,
 3913 *JINST* **03** (2008) S08005 (cit. on pp. 19, 44).
- 3914 [56] N. Trevisani, *Collider Searches for Dark Matter (ATLAS + CMS)*,
 3915 *Universe* **04** (2018) 131 (cit. on p. 19).

- 3916 [57] T. Mombacher, *Dark matter searches at LHCb*, 2021,
 3917 arXiv: [2111.00306 \[hep-ex\]](#) (cit. on p. 19).
- 3918 [58] BABAR collaboration, *The BABAR detector*,
 3919 *Nucl. Instrum. Methods Phys. Res. A* **479** (2002) 1–116,
 3920 arXiv: [hep-ex/0105044 \[hep-ex\]](#) (cit. on p. 19).
- 3921 [59] Belle Collaboration, *The Belle detector*, *Nucl. Instrum. Methods Phys. Res.*
 3922 *A* **479** (2002) 117, Detectors for Asymmetric B-factories (cit. on p. 19).
- 3923 [60] I. Adachi et al., *Detectors for extreme luminosity: Belle II*,
 3924 *Nucl. Instrum. Meth. A* **907** (2018) 46 (cit. on p. 19).
- 3925 [61] K. Akai et al., *SuperKEKB collider*,
 3926 *Nucl. Instrum. Methods Phys. Res. A* **907** (2018) 188–199,
 3927 arXiv: [1809.01958 \[physics.acc-ph\]](#) (cit. on p. 19).
- 3928 [62] M. Campajola, *Dark Sector first results at Belle II*,
 3929 *Phys. Scripta* **96** (2021) 084005 (cit. on p. 19).
- 3930 [63] H. Ray, *The MiniBooNE Experiment: An Overview*, 2007,
 3931 arXiv: [hep-ex/0701040 \[hep-ex\]](#) (cit. on p. 20).
- 3932 [64] A. Habig, *The NOvA Experiment*,
 3933 *Nucl. Phys. B Proc. Suppl.* **229-232** (2012) 460, ed. by G. S. Tzanakos
 3934 (cit. on p. 20).
- 3935 [65] MiniBooNE collaboration, *Dark matter search in nucleon, pion, and electron*
 3936 *channels from a proton beam dump with MiniBooNE*,
 3937 *Phys. Rev. D* **98** (2018), arXiv: [1807.06137 \[hep-ex\]](#) (cit. on p. 20).
- 3938 [66] P. deNiverville et al.,
 3939 *Hunting sub-GeV dark matter with the NOvA near detector*,
 3940 *Phys. Rev. D* **99** (2019), arXiv: [1807.06501 \[hep-ph\]](#) (cit. on p. 20).
- 3941 [67] NA64 Collaboration,
 3942 *Dark Matter Search in Missing Energy Events with NA64*,
 3943 *Phys. Rev. Lett.* **123** (2019), arXiv: [1906.00176 \[hep-ex\]](#) (cit. on p. 20).
- 3944 [68] G. Jungman et al., *Supersymmetric dark matter*, *Phys. Rept.* **267** (1996) 195,
 3945 arXiv: [hep-ph/9506380](#) (cit. on p. 20).

- 3946 [69] O. Buchmueller et al.,
 3947 *Beyond Effective Field Theory for Dark Matter Searches at the LHC*,
 3948 JHEP **01** (2014) 025, arXiv: [1308.6799 \[hep-ph\]](#) (cit. on p. 21).
- 3949 [70] D. Abercrombie et al., *Dark Matter benchmark models for early LHC Run-2*
 3950 *Searches: Report of the ATLAS/CMS Dark Matter Forum*,
 3951 Physics of the Dark Universe **27** (2020) 100371,
 3952 arXiv: [1507.00966 \[hep-ex\]](#) (cit. on p. 21).
- 3953 [71] C. Csaki, *The Minimal Supersymmetric Standard Model (MSSM)*,
 3954 Mod. Phys. Lett. A **11** (1996) 599, arXiv: [hep-ph/9606414 \[hep-ex\]](#)
 3955 (cit. on p. 20).
- 3956 [72] CMS collaboration,
 3957 *Search for new particles in events with energetic jets and large missing*
 3958 *transverse momentum in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$* ,
 3959 JHEP **11** (2021) 153, arXiv: [2107.13021 \[hep-ex\]](#) (cit. on p. 22).
- 3960 [73] CMS Collaboration, *Search for dark matter produced in association with*
 3961 *heavy-flavor quark pairs in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$* ,
 3962 Eur. Phys. J. C **77** (2017), arXiv: [1706.02581 \[hep-ex\]](#) (cit. on p. 22).
- 3963 [74] ATLAS Collaboration, *Search for dark matter produced in association with*
 3964 *bottom or top quarks in $\sqrt{s} = 13 \text{ TeV}$ pp collisions with the ATLAS detector*,
 3965 Eur. Phys. J. C **78** (2018), arXiv: [1710.11412 \[hep-ex\]](#) (cit. on p. 22).
- 3966 [75] ATLAS Collaboration,
 3967 *Search for associated production of a Z boson with an invisibly decaying Higgs*
 3968 *boson or dark matter candidates at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector*,
 3969 (2021), arXiv: [2111.08372 \[hep-ex\]](#) (cit. on p. 22).
- 3970 [76] CMS Collaboration, *Search for dark matter produced in association with a*
 3971 *leptonically decaying Z boson in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$* ,
 3972 Eur. Phys. J. C **81** (2021), arXiv: [2008.04735 \[hep-ex\]](#) (cit. on p. 22).
- 3973 [77] CMS Collaboration, *Search for dark matter particles produced in association*
 3974 *with a Higgs boson in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$* ,
 3975 JHEP **03** (2020) 025, arXiv: [1908.01713 \[hep-ex\]](#) (cit. on p. 22).

- 3976 [78] ATLAS Collaboration, *Search for Dark Matter produced in association with a*
 3977 *Standard Model Higgs boson decaying to b-quarks using the full Run 2*
 3978 *collision data with the ATLAS detector*, (2021), arXiv: [2108.13391 \[hep-ex\]](#)
 3979 (cit. on p. 22).
- 3980 [79] ATLAS Collaboration, *Search for dark matter in events with missing*
 3981 *transverse momentum and a Higgs boson decaying into two photons in pp*
 3982 *collisions at $\sqrt{s} = 13\text{TeV}$ with the ATLAS detector*, **JHEP** **10** (2021) 013,
 3983 arXiv: [2104.13240 \[hep-ex\]](#) (cit. on p. 22).
- 3984 [80] A. Alves et al., *Dark matter complementarity and the Z' portal*,
 3985 *Phys. Rev. D* **92** (2015), arXiv: [1501.03490v2 \[hep-ph\]](#) (cit. on p. 25).
- 3986 [81] M. Fairbairn et al., *Constraints on Z' models from LHC dijet searches and*
 3987 *implications for dark matter*, **JHEP** **09** (2016) 018,
 3988 arXiv: [1605.07940v2 \[hep-ph\]](#) (cit. on pp. 25–27).
- 3989 [82] M. Chala et al., *Constraining dark sectors with monojets and dijets*,
 3990 *JHEP* **07** (2015) 089, arXiv: [1503.05916 \[hep-ph\]](#) (cit. on pp. 25–27).
- 3991 [83] M. T. Frandsen et al., *LHC and Tevatron bounds on the dark matter direct*
 3992 *detection cross-section for vector mediators*, **JHEP** **07** (2012) 123,
 3993 arXiv: [1204.3839v2 \[hep-ph\]](#) (cit. on p. 29).
- 3994 [84] P. J. Fitzpatrick et al.,
 3995 *New Pathways to the Relic Abundance of Vector-Portal Dark Matter*, 2020,
 3996 arXiv: [2011.01240 \[hep-ph\]](#) (cit. on p. 29).
- 3997 [85] M. Blennow et al., *Neutrino portals to dark matter*,
 3998 *Eur. Phys. J. C* **79** (2019), arXiv: [1903.00006 \[hep-ph\]](#) (cit. on p. 29).
- 3999 [86] G. Arcadi et al., *Dark Matter through the Higgs portal*,
 4000 *Physics Reports* **842** (2020) 1–180 (cit. on p. 29).
- 4001 [87] Y. Nomura et al., *Dark matter through the axion portal*,
 4002 *Phys. Rev. D* **79** (2009), arXiv: [0810.5397 \[hep-ex\]](#) (cit. on p. 29).
- 4003 [88] A. Pais, *Remark on Baryon Conservation*,
 4004 *Phys. Rev. D* **08** (6 1973) 1844–1846 (cit. on p. 30).

- 4005 [89] J. Berger et al.,
 4006 *Cosmological constraints on decoupled dark photons and dark Higgs,*
 4007 *J. Cosmol. Astropart. Phys.* **11** (2016), arXiv: [1605.07195 \[hep-ex\]](https://arxiv.org/abs/1605.07195)
 4008 (cit. on p. 30).
- 4009 [90] *A combination of measurements of Higgs boson production and decay using*
 4010 *up to 139 fb⁻¹ of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with*
 4011 *the ATLAS experiment*, tech. rep. ATLAS-CONF-2020-027, CERN, 2020,
 4012 URL: <https://cds.cern.ch/record/2725733> (cit. on p. 31).
- 4013 [91] *Combined Higgs boson production and decay measurements with up to 137*
 4014 *fb⁻¹ of proton-proton collision data at $\sqrt{s} = 13$ TeV,*
 4015 tech. rep. CMS-PAS-HIG-19-005, CERN, 2020,
 4016 URL: <https://cds.cern.ch/record/2706103> (cit. on p. 31).
- 4017 [92] S. Argyropoulos et al.,
 4018 *Collider Searches for Dark Matter through the Higgs Lens,*
 4019 *Symmetry* **13** (2021) 2406, arXiv: [2109.13597 \[hep-ex\]](https://arxiv.org/abs/2109.13597) (cit. on p. 31).
- 4020 [93] CMS Collaboration, *Search for Narrow Resonances in the b-Tagged Dijet*
 4021 *Mass Spectrum in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV,*
 4022 *Phys. Rev. Lett.* **120** (2018), arXiv: [1802.06149 \[hep-ex\]](https://arxiv.org/abs/1802.06149) (cit. on p. 32).
- 4023 [94] ATLAS Collaboration,
 4024 *Search for Low-Mass Dijet Resonances Using Trigger-Level Jets with the*
 4025 *ATLAS Detector in pp Collisions at $\sqrt{s} = 13$ TeV,*
 4026 *Phys. Rev. Lett.* **121** (2018), arXiv: [1804.03496 \[hep-ex\]](https://arxiv.org/abs/1804.03496) (cit. on p. 32).
- 4027 [95] CMS Collaboration, *Search for narrow and broad dijet resonances in*
 4028 *proton-proton collisions at $\sqrt{s} = 13$ TeV and constraints on dark matter*
 4029 *mediators and other new particles*, *JHEP* **08** (2018) 130,
 4030 arXiv: [1806.00843 \[hep-ex\]](https://arxiv.org/abs/1806.00843) (cit. on pp. 32, 37).
- 4031 [96] ATLAS Collaboration, *Search for low-mass resonances decaying into two jets*
 4032 *and produced in association with a photon using pp collisions at $\sqrt{s} = 13$*
 4033 *TeV with the ATLAS detector*, *Phys. Lett. B* **795** (2019) 56–75,
 4034 arXiv: [1901.10917 \[hep-ex\]](https://arxiv.org/abs/1901.10917) (cit. on p. 32).

- 4035 [97] ATLAS Collaboration,
 4036 *Search for new phenomena in dijet events using 37 fb⁻¹ of pp collision data*
 4037 *collected at $\sqrt{s} = 13$ TeV with the ATLAS detector,*
 4038 Phys. Rev. D **96** (5 2017) 052004, arXiv: [1703.09127 \[hep-ex\]](https://arxiv.org/abs/1703.09127)
 4039 (cit. on p. 32).
- 4040 [98] *RECAST framework reinterpretation of an ATLAS Dark Matter Search*
 4041 *constraining a model of a dark Higgs boson decaying to two b-quarks,*
 4042 tech. rep. ATL-PHYS-PUB-2019-032, CERN, 2019,
 4043 URL: <https://cds.cern.ch/record/2686290> (cit. on pp. 34–37, 174, 175).
- 4044 [99] D. Abercrombie et al., *Dark Matter benchmark models for early LHC Run-2*
 4045 *Searches: Report of the ATLAS/CMS Dark Matter Forum,*
 4046 Physics of the Dark Universe **27** (2020) 100371,
 4047 arXiv: [1507.00966 \[hep-ex\]](https://arxiv.org/abs/1507.00966) (cit. on p. 35).
- 4048 [100] ATLAS Collaboration, *Search for Dark Matter Produced in Association with*
 4049 *a Dark Higgs Boson Decaying into $W^\pm W^\mp$ or ZZ in Fully Hadronic Final*
 4050 *States from $\sqrt{s} = 13$ TeV pp Collisions Recorded with the ATLAS Detector,*
 4051 Phys. Rev. Lett. **126** (12 2021) 121802, arXiv: [2010.06548 \[hep-ex\]](https://arxiv.org/abs/2010.06548)
 4052 (cit. on pp. 35, 38, 39, 174).
- 4053 [101] *Search for dark matter particles produced in association with a dark Higgs*
 4054 *boson decaying into W^+W^- in proton-proton collisions at $\sqrt{s} = 13$ TeV with*
 4055 *the CMS detector*, tech. rep. CMS-PAS-EXO-20-013, CERN, 2021,
 4056 URL: <https://cds.cern.ch/record/2776774> (cit. on pp. 36, 38, 39, 174).
- 4057 [102] *Search for Dark Matter Produced in Association with a Higgs Boson decaying*
 4058 *to $b\bar{b}$ at $\sqrt{s} = 13$ TeV with the ATLAS Detector using 79.8 fb⁻¹ of*
 4059 *proton-proton collision data*, tech. rep. ATLAS-CONF-2018-039,
 4060 CERN, 2018, URL: <https://cds.cern.ch/record/2632344>
 4061 (cit. on pp. 36, 37, 175).
- 4062 [103] H. T. Edwards,
 4063 *The Tevatron Energy Doubler: A Superconducting Accelerator*,
 4064 Annu. Rev. Nucl. Part. Sci. **35** (1985) 605–660 (cit. on p. 40).
- 4065 [104] T.-M. Yan et al., *The parton model and its applications*,
 4066 Int. J. Mod. Phys. A **29** (2014) 1430071, arXiv: [1409.0051 \[hep-ph\]](https://arxiv.org/abs/1409.0051)
 4067 (cit. on p. 41).

- 4068 [105] S. Bailey et al., *Parton distributions from LHC, HERA, Tevatron and fixed*
 4069 *target data: MSHT20 PDFs*, *Eur. Phys. J. C* **81** (2021),
 4070 arXiv: [2012.04684 \[hep-ph\]](https://arxiv.org/abs/2012.04684) (cit. on p. 41).
- 4071 [106] *Standard Model Summary Plots June 2021*,
 4072 tech. rep. ATL-PHYS-PUB-2021-032, CERN, 2021,
 4073 URL: <http://cds.cern.ch/record/2777014> (cit. on p. 42).
- 4074 [107] *Performance of the ALICE experiment at the CERN LHC*,
 4075 *Int. J. Mod. Phys. A* **29** (2014) 1430044, arXiv: [1402.4476 \[nucl-ex\]](https://arxiv.org/abs/1402.4476)
 4076 (cit. on p. 44).
- 4077 [108] R Alkofer et al., *Quark confinement: the hard problem of hadron physics*,
 4078 *J. Phys. G* **34** (2007), arXiv: [hep-ph/0610365 \[hep-ph\]](https://arxiv.org/abs/hep-ph/0610365) (cit. on p. 44).
- 4079 [109] P. Braun-Munzinger et al.,
 4080 *Colloquium: Phase diagram of strongly interacting matter*,
 4081 *Rev. Mod. Phys.* **81** (3 2009) 1031–1050, arXiv: [0801.4256 \[hep-ph\]](https://arxiv.org/abs/0801.4256)
 4082 (cit. on p. 44).
- 4083 [110] *ATLAS inner detector: Technical Design Report, 1*,
 4084 Technical design report. ATLAS ATLAS-TDR-4, CERN-LHCC-97-016,
 4085 Geneva: CERN, 1997, URL: <https://cds.cern.ch/record/331063>
 4086 (cit. on p. 47).
- 4087 [111] L. D. Landau et al., *The cascade theory of electronic showers*,
 4088 *Proc. R. Soc. Lond. A* **166** (1938) 213–228 (cit. on p. 47).
- 4089 [112] C. Leroy et al.,
 4090 *Physics of cascading shower generation and propagation in matter: Principles*
 4091 *of high-energy, ultrahigh-energy and compensating calorimetry*,
 4092 *Rep. Prog. Phys.* **63** (2000) 505–606 (cit. on pp. 47, 49).
- 4093 [113] H. Zhang et al.,
 4094 *The ATLAS Liquid Argon Calorimeter: Overview and Performance*,
 4095 *J. Phys. Conf. Ser.* **293** (2011) 012044 (cit. on p. 48).
- 4096 [114] V Grassi,
 4097 *The Atlas Liquid Argon Calorimeter at the CERN Large Hadron Collider: General Performance and Latest Developments of the High Voltage System*,
 4098 tech. rep. ATL-LARG-PROC-2013-013, CERN, 2013,
 4099 URL: <https://cds.cern.ch/record/1628748> (cit. on p. 48).

- 4101 [115] ATLAS Collaboration,
 4102 *Electron reconstruction and identification in the ATLAS experiment using the*
 4103 *2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13 \text{ TeV}$,*
 4104 *Eur. Phys. J. C* **79** (2019), arXiv: [1902.04655 \[physics.ins-det\]](https://arxiv.org/abs/1902.04655)
 4105 (cit. on p. 48).
- 4106 [116] S. Schramm, *ATLAS Jet Reconstruction, Calibration, and Tagging of*
 4107 *Lorentz-boosted Objects*, tech. rep. ATL-PHYS-PROC-2017-236,
 4108 CERN, 2017, URL: <https://cds.cern.ch/record/2291608> (cit. on p. 50).
- 4109 [117] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters*
 4110 *and its performance in LHC Run 1*, *Eur. Phys. J. C* **77** (2017),
 4111 arXiv: [1603.02934 \[hep-ex\]](https://arxiv.org/abs/1603.02934) (cit. on p. 50).
- 4112 [118] M. Cacciari et al., *The anti- k_t jet clustering algorithm*, *JHEP* **04** (2008) 063,
 4113 arXiv: [0802.1189 \[hep-ph\]](https://arxiv.org/abs/0802.1189) (cit. on pp. 50, 72, 132).
- 4114 [119] *ATLAS muon spectrometer: Technical Design Report*,
 4115 Technical design report. ATLAS ATLAS-TDR-10, CERN-LHCC-97-022,
 4116 Geneva: CERN, 1997, URL: <https://cds.cern.ch/record/331068>
 4117 (cit. on p. 51).
- 4118 [120] ATLAS Collaboration, *Muon reconstruction and identification efficiency in*
 4119 *ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13 \text{ TeV}$* ,
 4120 *Eur. Phys. J. C* **81** (2021), arXiv: [2012.00578 \[hep-ex\]](https://arxiv.org/abs/2012.00578)
 4121 (cit. on pp. 51, 52, 81).
- 4122 [121] ATLAS Collaboration,
 4123 *Performance of missing transverse momentum reconstruction with the*
 4124 *ATLAS detector using proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$* ,
 4125 *Eur. Phys. J. C* **78** (2018) (cit. on p. 52).
- 4126 [122] M. zur Nedden,
 4127 *The Run-2 ATLAS Trigger System: Design, Performance and Plan*,
 4128 tech. rep. ATL-DAQ-PROC-2016-039, CERN, 2016,
 4129 URL: <https://cds.cern.ch/record/2238679> (cit. on pp. 53, 54).
- 4130 [123] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*,
 4131 *Eur. Phys. J. C* **70** (2010) 823–874, arXiv: [1005.4568 \[physics.ins-det\]](https://arxiv.org/abs/1005.4568)
 4132 (cit. on p. 58).

- 4133 [124] S. Agostinelli et al., *Geant4—a simulation toolkit*,
 4134 *Nucl. Instrum. Methods Phys. Res. A* **506** (2003) 250–303
 4135 (cit. on pp. 58, 118).
- 4136 [125] Z. Marshall et al., *Simulation of Pile-up in the ATLAS Experiment*,
 4137 *J. Phys. Conf. Ser.* **513** (2014) 022024 (cit. on pp. 59, 72).
- 4138 [126] H. Wiedemann, *Particle Accelerator Physics*, 4th ed., Springer, 2015,
 4139 URL: <https://doi.org/10.1007/978-3-319-18317-6> (cit. on p. 59).
- 4140 [127] A. Haas, *ATLAS Simulation using Real Data: Embedding and Overlay*,
 4141 *J. Phys. Conf. Ser.* **898** (2017) 042004 (cit. on p. 60).
- 4142 [128] Novak, Tadej, *New techniques for pile-up simulation in ATLAS*,
 4143 *EPJ Web Conf.* **214** (2019) 02044 (cit. on p. 60).
- 4144 [129] J. Alwall et al., *MadGraph 5: going beyond*, *JHEP* **06** (2011) 128,
 4145 arXiv: [1106.0522 \[hep-ph\]](https://arxiv.org/abs/1106.0522) (cit. on p. 62).
- 4146 [130] J. Alwall et al.,
 4147 *The automated computation of tree-level and next-to-leading order differential*
 4148 *cross sections, and their matching to parton shower simulations*,
 4149 *JHEP* **07** (2014) 079, arXiv: [1405.0301 \[hep-ph\]](https://arxiv.org/abs/1405.0301) (cit. on pp. 62, 129).
- 4150 [131] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*,
 4151 *Comput. Phys. Commun.* **191** (2015) 159, arXiv: [1410.3012 \[hep-ph\]](https://arxiv.org/abs/1410.3012)
 4152 (cit. on pp. 63, 66, 67).
- 4153 [132] T. Gleisberg et al., *Event generation with SHERPA 1.1*,
 4154 *JHEP* **02** (2009) 007, arXiv: [0811.4622 \[hep-ph\]](https://arxiv.org/abs/0811.4622) (cit. on pp. 64, 131).
- 4155 [133] C. Varni, *Tracking and flavour-tagging performance at ATLAS*,
 4156 tech. rep. ATL-PHYS-PROC-2020-085, CERN, 2020,
 4157 URL: <https://cds.cern.ch/record/2742644> (cit. on p. 66).
- 4158 [134] S. Frixione et al., *A positive-weight next-to-leading-order Monte Carlo for*
 4159 *heavy flavour hadroproduction*, *JHEP* **09** (2007) 126,
 4160 arXiv: [0707.3088 \[hep-ph\]](https://arxiv.org/abs/0707.3088) (cit. on p. 66).
- 4161 [135] P. Nason,
 4162 *A new method for combining NLO QCD with shower Monte Carlo algorithms*,
 4163 *JHEP* **11** (2004) 040, arXiv: [hep-ph/0409146](https://arxiv.org/abs/hep-ph/0409146) (cit. on pp. 66, 67).

- 4164 [136] S. Frixione et al., *Matching NLO QCD computations with parton shower*
 4165 *simulations: the POWHEG method*, JHEP **11** (2007) 070,
 4166 arXiv: [0709.2092 \[hep-ph\]](https://arxiv.org/abs/0709.2092) (cit. on pp. 66, 67).
- 4167 [137] S. Alioli et al., *A general framework for implementing NLO calculations in*
 4168 *shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043,
 4169 arXiv: [1002.2581 \[hep-ph\]](https://arxiv.org/abs/1002.2581) (cit. on pp. 66, 67).
- 4170 [138] E. Re, *Single-top Wt-channel production matched with parton showers using*
 4171 *the POWHEG method*, Eur. Phys. J. C **71** (2011) 1547,
 4172 arXiv: [1009.2450 \[hep-ph\]](https://arxiv.org/abs/1009.2450) (cit. on p. 67).
- 4173 [139] S Xella, *Physics objects reconstruction in the ATLAS experiment*,
 4174 tech. rep. ATL-GEN-PROC-2013-001, CERN, 2013,
 4175 URL: <https://cds.cern.ch/record/1519113> (cit. on p. 68).
- 4176 [140] *Measurement of the tau lepton reconstruction and identification performance*
 4177 *in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV*,
 4178 tech. rep. ATLAS-CONF-2017-029, CERN, 2017,
 4179 URL: <https://cds.cern.ch/record/2261772> (cit. on p. 69).
- 4180 [141] ATLAS Collaboration, *Electron and photon reconstruction and performance*
 4181 *in ATLAS using a dynamical, topological cell clustering-based approach*,
 4182 ATL-PHYS-PUB-2017-022, 2017,
 4183 URL: <https://cds.cern.ch/record/2298955> (cit. on p. 69).
- 4184 [142] ATLAS Collaboration,
 4185 *Electron reconstruction and identification in the ATLAS experiment using the*
 4186 *2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV*,
 4187 Eur. Phys. J. C **79** (2019) 639, arXiv: [1902.04655 \[hep-ex\]](https://arxiv.org/abs/1902.04655)
 4188 (cit. on pp. 70, 71).
- 4189 [143] ATLAS Collaboration, *Electron and photon energy calibration with the*
 4190 *ATLAS detector using 2015–2016 LHC proton–proton collision data*,
 4191 JINST **14** (2019) P03017, arXiv: [1812.03848 \[hep-ex\]](https://arxiv.org/abs/1812.03848) (cit. on p. 70).
- 4192 [144] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS*
 4193 *detector in proton–proton collision data at $\sqrt{s} = 13$ TeV*,
 4194 Eur. Phys. J. C **76** (2016) 292, arXiv: [1603.05598 \[hep-ex\]](https://arxiv.org/abs/1603.05598) (cit. on p. 71).

- 4195 [145] S. Rettie, *Muon identification and performance in the ATLAS experiment*,
 4196 tech. rep. ATL-PHYS-PROC-2018-052, CERN, 2018,
 4197 URL: <https://cds.cern.ch/record/2626330> (cit. on p. 71).
- 4198 [146] ATLAS Collaboration, *Jet reconstruction and performance using particle flow*
 4199 *with the ATLAS Detector*, *Eur. Phys. J. C* **77** (2017) 466,
 4200 arXiv: [1703.10485 \[hep-ex\]](https://arxiv.org/abs/1703.10485) (cit. on p. 71).
- 4201 [147] ATLAS Collaboration,
 4202 *Monte Carlo Calibration and Combination of In-situ Measurements of Jet*
 4203 *Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS*,
 4204 ATLAS-CONF-2015-037, 2015,
 4205 URL: <https://cds.cern.ch/record/2044941> (cit. on p. 72).
- 4206 [148] ATLAS Collaboration, *Selection of jets produced in 13 TeV proton–proton*
 4207 *collisions with the ATLAS detector*, ATLAS-CONF-2015-029, 2015,
 4208 URL: <https://cds.cern.ch/record/2037702> (cit. on p. 72).
- 4209 [149] ATLAS Collaboration,
 4210 *Tagging and suppression of pileup jets with the ATLAS detector*,
 4211 ATLAS-CONF-2014-018, 2014,
 4212 URL: <https://cds.cern.ch/record/1700870> (cit. on p. 72).
- 4213 [150] ATLAS Collaboration, *Calibration of light-flavour b-jet mistagging rates*
 4214 *using ATLAS proton–proton collision data at $\sqrt{s} = 13 \text{ TeV}$* ,
 4215 ATLAS-CONF-2018-006, 2018,
 4216 URL: <https://cds.cern.ch/record/2314418> (cit. on p. 73).
- 4217 [151] M. Cacciari et al., *The catchment area of jets*, *JHEP* **04** (2008) 005
 4218 (cit. on p. 75).
- 4219 [152] ATLAS Collaboration,
 4220 *Performance of missing transverse momentum reconstruction with the*
 4221 *ATLAS detector using proton–proton collisions at $\sqrt{s} = 13 \text{ TeV}$* ,
 4222 *Eur. Phys. J. C* **78** (2018) 903, arXiv: [1802.08168 \[hep-ex\]](https://arxiv.org/abs/1802.08168) (cit. on p. 80).
- 4223 [153] ATLAS Collaboration, *Object-based missing transverse momentum*
 4224 *significance in the ATLAS Detector*, ATLAS-CONF-2018-038, 2018,
 4225 URL: <https://cds.cern.ch/record/2630948> (cit. on p. 80).

- 4226 [154] W. Buttinger et al., *Formulae for Estimating Significance*,
 4227 tech. rep. ATL-COM-GEN-2018-026, CERN, 2018,
 4228 URL: <https://cds.cern.ch/record/2643488> (cit. on p. 90).
- 4229 [155] A. J. Larkoski et al., *Energy Correlation Functions for Jet Substructure*,
 4230 JHEP **06** (2013) 108, arXiv: [1305.0007 \[hep-ph\]](https://arxiv.org/abs/1305.0007) (cit. on p. 94).
- 4231 [156] A. J. Larkoski et al., *Analytic boosted boson discrimination*,
 4232 JHEP **05** (2016) 117, arXiv: [1507.03018 \[hep-ph\]](https://arxiv.org/abs/1507.03018) (cit. on p. 94).
- 4233 [157] S. Forte et al.,
 4234 *Progress in the Determination of the Partonic Structure of the Proton*,
 4235 Annu. Rev. Nucl. Part. Sci. **63** (2013) 291–328, arXiv: [1301.6754 \[hep-ph\]](https://arxiv.org/abs/1301.6754)
 4236 (cit. on p. 98).
- 4237 [158] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*,
 4238 J. Phys. G **43** (2016) 023001, arXiv: [1510.03865 \[hep-ph\]](https://arxiv.org/abs/1510.03865)
 4239 (cit. on pp. 98, 127).
- 4240 [159] ATLAS Collaboration, *Performance of the missing transverse momentum*
 4241 *triggers for the ATLAS detector during Run-2 data taking*,
 4242 JHEP **08** (2020) 080, arXiv: [2005.09554 \[hep-ex\]](https://arxiv.org/abs/2005.09554) (cit. on p. 111).
- 4243 [160] M. Ronzani, *ATLAS Transverse Missing Energy Trigger Performance*,
 4244 tech. rep. ATL-DAQ-PROC-2019-028, CERN, 2019,
 4245 URL: <https://cds.cern.ch/record/2693457> (cit. on pp. 111, 113).
- 4246 [161] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$*
 4247 *TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, 2019,
 4248 URL: <https://cds.cern.ch/record/2677054> (cit. on p. 119).
- 4249 [162] R. D. Ball et al.,
 4250 *A determination of parton distributions with faithful uncertainty estimation*,
 4251 Nuclear Physics B **809** (2009) 1–63, arXiv: [0808.1231 \[hep-ph\]](https://arxiv.org/abs/0808.1231)
 4252 (cit. on p. 127).
- 4253 [163] *ATLAS Pythia 8 tunes to 7 TeV data*, tech. rep. ATL-PHYS-PUB-2014-021,
 4254 CERN, 2014, URL: <https://cds.cern.ch/record/1966419> (cit. on p. 128).
- 4255 [164] M. Bähr et al., *Herwig++ physics and manual*,
 4256 Eur. Phys. J. C **58** (2008) 639, arXiv: [0803.0883 \[hep-ph\]](https://arxiv.org/abs/0803.0883) (cit. on p. 129).

- 4257 [165] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*,
 4258 Eur. Phys. J. C **76** (2016) 196, arXiv: 1512.01178 [hep-ph] (cit. on p. 129).
- 4259 [166] S. Höche et al., *QCD matrix elements and truncated showers*,
 4260 JHEP **05** (2009) 053, arXiv: 0903.1219 [hep-ph] (cit. on p. 131).
- 4261 [167] J. K. Anders et al.,
 4262 *V+Jets theoretical uncertainties estimation via a parameterisation method*,
 4263 tech. rep. ATL-COM-PHYS-2016-044, CERN, 2016,
 4264 URL: <https://cds.cern.ch/record/2125718> (cit. on p. 131).
- 4265 [168] C. D. White et al., *Isolating Wt production at the LHC*,
 4266 JHEP **11** (2009) 074, arXiv: 0908.0631 [hep-ph] (cit. on p. 133).
- 4267 [169] K. Cranmer et al., *HistFactory: A tool for creating statistical models for use*
 4268 *with RooFit and RooStats*, tech. rep. CERN-OPEN-2012-016,
 4269 New York U., 2012, URL: <https://cds.cern.ch/record/1456844>
 4270 (cit. on p. 143).
- 4271 [170] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for*
 4272 *Testing Composite Hypotheses*, Ann. Math. Stat. **09** (1938) 60–62
 4273 (cit. on p. 152).
- 4274 [171] G. Cowan et al.,
 4275 *Asymptotic formulae for likelihood-based tests of new physics*,
 4276 Eur. Phys. J. C **71** (2011), arXiv: 1007.1727 [physics.data-an]
 4277 (cit. on p. 153).
- 4278 [172] D. M. Lane, *Online Statistics Education: A Multimedia Course of Study*,
 4279 [Online; accessed March 11, 2021], Rice University,
 4280 URL: <http://onlinestatbook.com/> (cit. on p. 157).
- 4281 [173] *Recommendations on presenting LHC searches for missing transverse energy*
 4282 *signals using simplified s-channel models of dark matter*,
 4283 Physics of the Dark Universe **27** (2020) 100365,
 4284 arXiv: 1603.04156 [hep-ex] (cit. on p. 161).