

# 1 Search for Dark Matter Produced in Association with a 2 Hypothetical Dark Higgs Boson Decaying to $W^+W^-$ in 3 the $q\bar{q}\ell\nu$ Final State Using $pp$ Collisions Recorded with 4 the ATLAS Detector

Danika MacDonell

July 2, 2020

# **7 1 Introduction**

8 The proposed thesis is a search for dark matter using high energy proton-proton ( $pp$ ) collision  
 9 data recorded with the ATLAS detector at the Large Hadron Collider (LHC). The search targets a  
 10 final state signature of dark matter production in association with the emission of a hypothetical  
 11 higgs boson  $s$  in the dark sector, which subsequently decays to a pair of W bosons. The search  
 12 is motivated by and optimized with a “dark higgs model” [1] shown in figure 2, in which the  $s$  is  
 13 emitted from a hypothetical  $Z'$  gauge boson in the dark sector, which itself mediates the production  
 14 of dark matter particles from the high energy  $pp$  collisions.

15 Any dark matter that may be produced in the  $pp$  collisions is expected not to interact to any  
 16 measurable extent with the ‘normal’ matter constituting the detector. As such, it is assumed that  
 17 the momentum carried by the dark matter would escape the detector undetected. However, the law  
 18 of momentum conservation requires that the momenta of all particles produced by a  $pp$  collision  
 19 in the plane transverse to the beam line sum to zero. This fundamental requirement implies that  
 20 dark matter produced at the LHC will exhibit a signature of high missing transverse momentum  
 21 in the final state. The magnitude of this two dimensional missing transverse momentum vector is  
 22 typically denoted “ $E_T^{\text{miss}}$ ”.

The aim of the proposed thesis work is to apply selections, including a requirement of high  $E_T^{\text{miss}}$ , to data from the ATLAS detector to optimize the sensitivity of the data to the dark higgs signal model. The data will subsequently be compared with standard model (SM) background and signal processes simulated with Monte Carlo (MC) to search for an above-background excess in the data consistent with the signal model.

28 The search is split into two separate analyses:

- 29    1. **The hadronic decay channel:** In this channel, each of the two W bosons in the final state  
30    decay to quarks, which subsequently hadronize in the detector and are measured as jets in  
31    the ATLAS calorimeter.

$$WW \rightarrow q\bar{q} + q\bar{q}$$

Given that the branching fraction for hadronic W decay is 0.68 [2], the fully hadronic WW decay occurs with a branching fraction of 0.46 ( $= 0.68^2$ ). In addition to its substantial branching fraction, this channel has the advantage of being able to fully reconstruct the momenta of both W bosons, and as such the  $s$  mass. The drawback is that there are standard model processes with large production cross section which also produce a final state of multiple jets in the detector, a subset of which pass the other signal selection criteria and represent a sizeable background in the analysis.

2. **The semileptonic decay channel:** One W boson decays to quarks in this channel, and the other decays to a lepton and a neutrino, where the lepton is either an electron or a muon.

$$WW \rightarrow q\bar{q} + \ell\nu$$

Despite its lower branching fraction of 0.29 [2] compared with the fully hadronic decay channel, the semileptonic channel has the advantage that the requirement of having one lepton in the final state reduces the background from standard model processes. However, this comes at the cost of additional  $E_T^{\text{miss}}$  from the neutrino production, which both inhibits reconstruction of the leptonically-decaying W boson and cannot easily be distinguished from the  $E_T^{\text{miss}}$  associated with dark matter production in the signal model.

There is an ongoing analysis searching for the dark higgs signal in the hadronic decay channel, and the proposed thesis will focus on the semileptonic decay channel.

The WW decay can also proceed via a fully leptonic channel, in which both W bosons decay to a lepton and a neutrino. This channel would give a clean signature of two leptons and no jets, but is currently not considered as a viable search channel due to its relatively low branching fraction of 0.046 [2].

Once the analysis is fully developed in the semileptonic decay channel, this channel will be statistically combined with the hadronic search channel. In the event that the combined search does not show an above-background excess, the analysis will place exclusion limits - to within some level of confidence, typically 95% - on the range of  $Z'$  and  $s$  masses in the dark higgs model to which the combined search is sensitive.

## 2 Motivation

The proposed thesis work would contribute to the extensive ongoing dark matter search program at the LHC. The LHC search program is motivated by compelling evidence from observational astronomy for the existence of dark matter constituting 85% [3] of all matter in the universe. Despite clear evidence from observational astronomy for its gravitational interactions with normal matter, dark matter has yet to be detected through the weak, strong or electromagnetic interactions. As such, its composition and non-gravitational interactions - if any - remain largely a mystery. The current most widely accepted and theoretically motivated dark matter candidates take the form of fundamental particles, yet there is no particle in the standard model of particle physics that could represent a viable dark matter candidate [4]. Dark matter is therefore widely assumed to be a “beyond-standard-model” (BSM) particle.

## 69 2.1 Dark Matter Search Methods

70 There are three complementary approaches used to search for particle dark matter: direct and  
 71 indirect detection, and collider searches. Direct detection searches [5, 6] aim to directly detect  
 72 evidence of a recoil induced by elastic scattering between a dark matter particle in the galactic  
 73 halo passing through the detector and a target particle in the detector. Indirect searches [7, 8]  
 74 use observational data to search for evidence of products produced by dark matter annihilation or  
 75 decay in particular regions of the observable universe expected to have a high dark matter density.  
 76 Collider searches [9], of which the proposed thesis work is an example, study the decay products  
 77 from high-energy collisions of subatomic particles to search for an above-background excess of  
 78 events that could be consistent with dark matter having been produced in some of the collisions.

## 79 2.2 Models of Dark Matter Production at Colliders

80 Models of dark matter production in colliders can range in complexity from an effective field the-  
 81 ory (EFT), where the dark matter production mechanism is completely unspecified, to a complete  
 82 model such as supersymmetry [10], which predicts viable dark matter candidates as part of a hy-  
 83 pothesized extension to the standard model designed to address a range of phenomena unexplained  
 84 by the standard model.

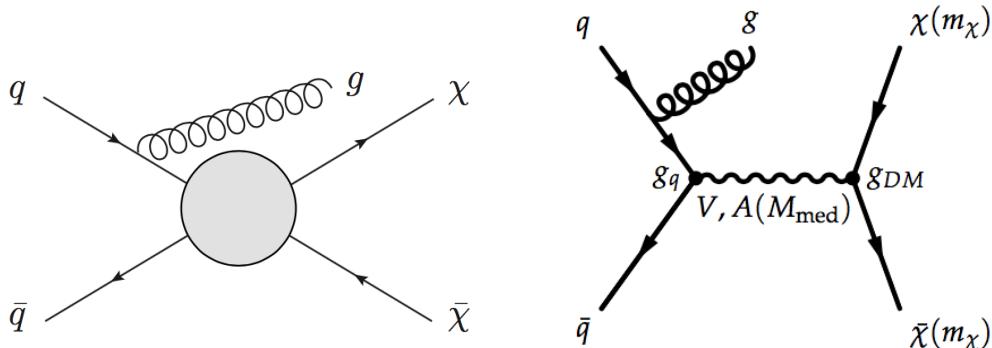


Figure 1: Left: Mono-jet process in the EFT framework (source: [11]). Right: Mono-jet process in a simplified model framework, where the pair production of dark matter occurs via a new vector or axial-vector ( $V, A$ ) mediator of mass  $M_{\text{med}}$ , which couples to quarks and dark matter with coupling constants  $g_q$  and  $g_{\text{DM}}$ , respectively (source: [12])

85 In principle, complete theories of physics beyond the standard model, such as the minimal  
 86 supersymmetric standard model (MSSM) [13] can offer theoretically motivated and experimentally  
 87 accessible models which specify the details of candidate processes by which the colliding partons  
 88 may annihilate to produce dark matter. However, these theories tend to be quite complex, with  
 89 many free parameters - over 100 in the case of MSSM [9] - most of which need to be fixed to  
 90 generate a reasonably testable model. Relying on complete theories alone to guide experimental  
 91 signatures may run the risk of missing important parameter space of new physics for which a  
 92 complete theory has not yet been developed.

93 Simplified models, widely used in recent and ongoing dark matter searches at the LHC, are  
 94 designed to bridge the gap between EFT and complete theories. They provide a ‘first-order’ de-

95 description of theoretically motivated new physics scenarios that could be accessible at collider ener-  
 96 gies. They provide guidance for experimental searches without fully specifying the details of any  
 97 additional new physics at energies above the collider scale that would be needed for a complete  
 98 theory [9]. In terms of dark matter production at the LHC, one or more new mediators associated  
 99 with new physics scenarios may be considered which allow for mixing between standard model  
 100 particles and dark matter. The process by which the mixing occurs is represented with a tree-level  
 101 diagram whose experimental signature would be accessible at LHC energies, such as the diagram  
 102 shown in figure 1 which represents a dark matter benchmark model featured in the 2015 report of  
 103 the ATLAS/CMS Dark Matter Forum [12].

### 104 **2.3 Dark Higgs Model**

105 The proposed search is motivated by and interpreted with a simplified model described in [1]  
 106 where dark matter is pair produced from a Z boson in the dark sector - known as the  $Z'$  boson - in  
 107 association with the emission of an additional dark sector higgs boson  $s$ . Figure 2 shows one of the  
 108 leading order Feynman diagrams representing this simplified model. The dark sector higgs boson  
 109 subsequently decays to a pair of massive standard model particles through a small mixing with the  
 110 SM higgs boson.

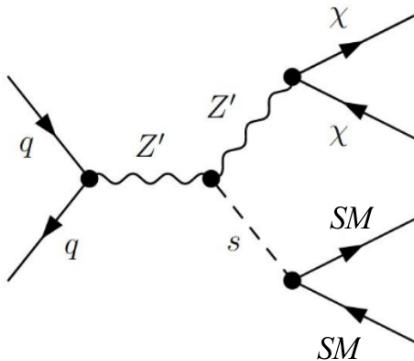


Figure 2: Dark higgs model

111 The generic hypothesis of ‘dark sector’ mediators which exhibit a small mixing with SM parti-  
 112 cles is motivated in part by the theoretical need for creation and annihilation mechanisms between  
 113 DM and SM particles in the well-motivated and popular ‘thermal freeze-out’ hypothesis [14]. In  
 114 the thermal freeze-out hypothesis, DM and SM particles interacted at a sufficient rate in the high  
 115 temperature and matter density of the early universe to remain in thermal equilibrium. Once the  
 116 DM-SM interaction rate dropped below the expansion rate of the universe, the DM rapidly “froze  
 117 out” of thermal equilibrium with the SM particles, at which point the relic abundance of DM was  
 118 fixed to the value observed in the present-day universe.

119 The dark higgs boson in particular is motivated by the need to generate the masses of dark  
 120 matter and any other particles residing in the dark sector [1]. If dark sector particle masses are  
 121 generated via a ‘dark sector higgs mechanism’, this would naturally imply the existence of a dark  
 122 higgs boson.

123        The model presented in [1] considers a scenario in which there is another dark sector mediator  
124 - taken to be the  $Z'$  - with a mass within the range accessible to measurement at the LHC, which  
125 mediates dark matter production from parton collisions. In this case, the  $Z'$  could radiate a dark  
126 higgs boson. Theoretically, the observed relic abundance of dark matter in the universe suggests  
127 that couplings between particles in the dark sector should be quite large, which could result in a  
128 sufficiently high probability of dark higgs emission from the  $Z'$  to produce a measurable signature  
129 in LHC collisions [1].

130        Experimentally, the dark higgs signature would be distinct from generic mono-X signatures  
131 where the SM particle that the dark matter recoils against is produced via initial state radiation,  
132 because as long as the mass of the  $s$  is well below the typical momentum transfer of the collisions,  
133 the  $s$  will tend to be highly boosted in the plane transverse to the beam line. As a result, the  
134 SM decay products will tend to be much higher in transverse momentum and highly collimated  
135 compared with particles produced via initial state radiation.

## 136        2.4 Experimental Background

137        This section discusses a previous experimental search which set exclusion limits on the dark higgs  
138 model described in Section 2.3, in the case where the dark higgs decays to a pair of b quarks.  
139 It goes on to describe the  $m_s$  range that the ongoing and proposed searches in the  $s \rightarrow WW$   
140 decay channel target, and explain how this target is complementary to the  $m_s$  range explored by  
141 the previous search in the  $s \rightarrow bb$  channel.

### 142        2.4.1 Re-interpretation of Mono-H(bb) Search with a Dark Higgs Mediator

143        Last year, the “mono-H(bb)” dark matter search, originally published in 2018 [15] with  $79.8 \text{ fb}^{-1}$   
144 of data taken with the ATLAS detector at a 13 TeV centre of mass energy, was re-interpreted [16]  
145 using the RECAST framework [17] to set exclusion limits on the dark higgs model in the case  
146 where the dark higgs decays to a pair of b-quarks.

147        The mono-H(bb) search selects for a final state of missing transverse energy along with two  
148 jets in the calorimeter, both of which must be tagged as having originated from the hadronization  
149 of b quarks. When the mono-H(bb) analysis was originally published, it had been developed  
150 and interpreted using a simplified model shown in figure 3 left, in which the dark matter is pair  
151 produced from a new pseudoscalar higgs boson, along with the emission of a SM higgs boson,  
152 which subsequently decays to two b quarks.

153        The original analysis was preserved using RECAST [17], a framework designed to preserve  
154 searches for new physics with high energy collision data in such a way that the searches can  
155 be readily re-analyzed with alternative models of new physics which predict the same final state  
156 signature. The RECAST framework was used to re-interpret the mono-H(bb) search using the  
157 dark higgs model shown in figure 3 right, where the dark higgs, whose mass is left as a floating  
158 parameter, decays to a pair of b quarks.

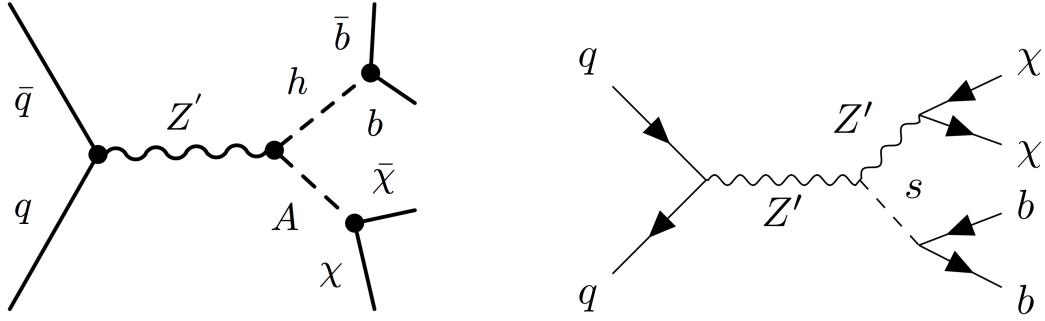


Figure 3: Left: Leading-order Feynman diagram representing the model used to interpret the mono-H(bb) dark matter search. Right: A leading-order Feynman diagram representing the model used to guide and interpret the mono-s(bb) re-interpretation of the mono-H(bb) search. The dark higgs mediator mass  $m_s$  is allowed to float in the dark higgs model.

159      The re-interpreted mono-H(bb) search set limits, shown in figure 4, on the phase space of  $Z'$   
 160     and  $s$  masses excluded by the data. To set these exclusion limits, several free parameters in the  
 161     dark higgs model need to be fixed. Specifically:

- 162     • The dark matter mass  $m_\chi$  is fixed to 200 GeV.
- 163     • The constant  $g_q$  associated with the coupling of the  $Z'$  boson to quarks is fixed to 0.25.
- 164     • The constant  $g_\chi$  associated with the coupling of the  $Z'$  boson to dark matter is fixed to 1.0.
- 165     • The mixing angle  $\theta$  between the dark and SM higgs bosons is fixed to 0.01.

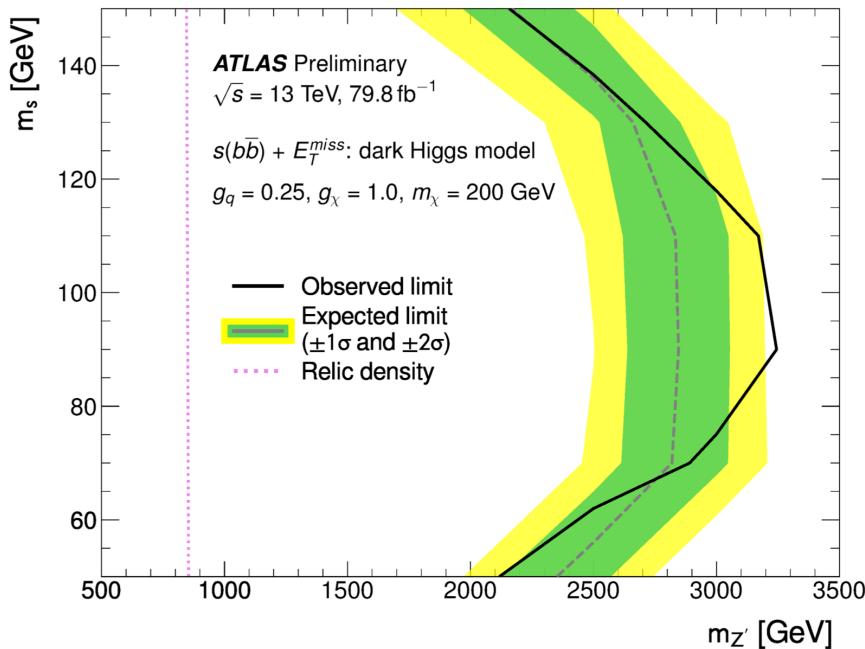


Figure 4: Exclusion limits set on the  $Z'$  and  $s$  mediator masses by the mono-H(bb) search re-interpreted with the dark higgs model. Mass combinations to the left of the solid black curve are excluded at a 95% confidence level. Source: [16]

166 The values of  $g_q$  and  $g_\chi$  were chosen to match the conventional choices used by other LHC DM  
 167 searches, and  $\theta$  matches the choice used in the original paper [1] that introduced the dark higgs  
 168 model.

#### 169 2.4.2 Dark Higgs Decay Modes

170 The re-interpreted mono-H(bb) search probed the dark higgs model in the case where the  $s$  emitted  
 171 from the  $Z'$  mediator decays to a pair of b quarks. However, predictions of the dark higgs branching  
 172 ratio presented in the re-interpretation paper [16] and shown in figure 5 indicate that the  $s \rightarrow bb$   
 173 decay mode is only sensitive to a range of dark higgs masses up to  $\sim 150$  GeV, above which the  
 174 WW decay mode dominates in sensitivity.

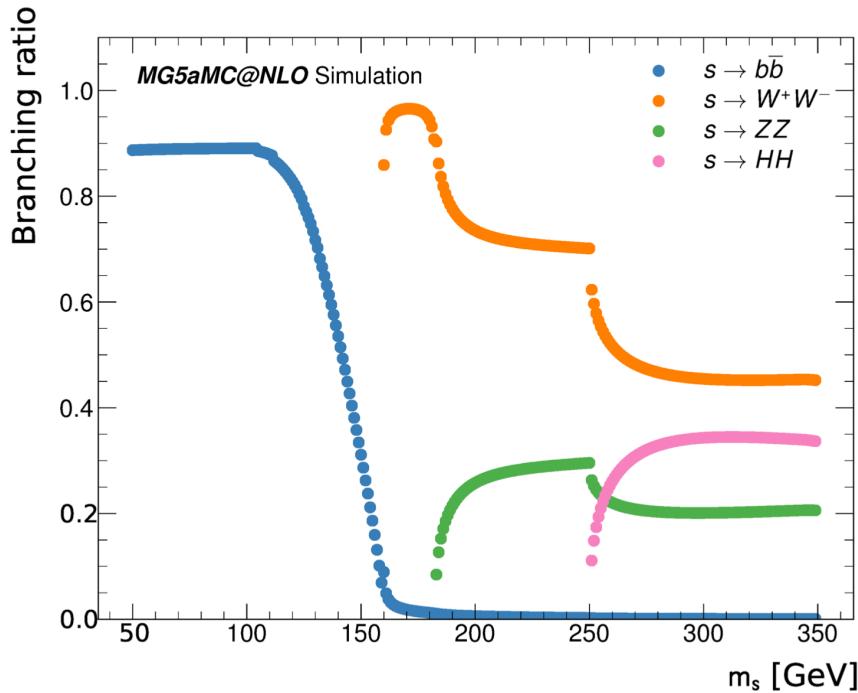


Figure 5: Predicted dark higgs decay branching ratios as a function of dark higgs mass. Source: [16]

175 There is no a-priori reason to expect the dark higgs mass  $m_s$  to necessarily lie below 150 GeV.  
 176 Therefore, it's important to probe the model in the higher  $m_s$  range by analyzing the  $s \rightarrow WW$   
 177 decay mode. The proposed thesis work will analyze the latter decay mode using the full  $139 \text{ fb}^{-1}$   
 178 ATLAS run 2 dataset, focusing on the semileptonic final state  $WW \rightarrow q\bar{q}\ell\nu$ .

### 179 3 Introduction to the LHC and the ATLAS detector

180 The Large Hadron Collider (LHC) [18] is a circular proton-proton collider which resides in a 27 km  
181 tunnel near the European Organization for Nuclear Research (CERN). Superconducting magnets  
182 are used to accelerate counter-rotating bunched proton beams to high energy, and direct the beams  
183 into head-on collisions at four interaction points around the ring. The collisions take place at a  
184 world-leading centre of mass energy of up to 13 TeV.

185 Each interaction point is surrounded by a detector, which measures the energetic debris of  
186 particles produced by the high energy collisions to perform precision measurements of the standard  
187 model and search for new physics. ATLAS (A Toroidal LHC ApparatuS) [19] is one of two multi-  
188 purpose detectors at the LHC, designed to record and study a wide range of physics processes  
189 resulting from the collisions.

190 The ATLAS detector, shown schematically in figure 6, provides full  $4\pi$  coverage around the  
191 interaction point, with the exception of the beam pipe. It consists of several layers of sub-detectors,  
192 each of which is specialized for recording certain kinematic information and particle types. The  
193 sub-detectors are described in some detail below.

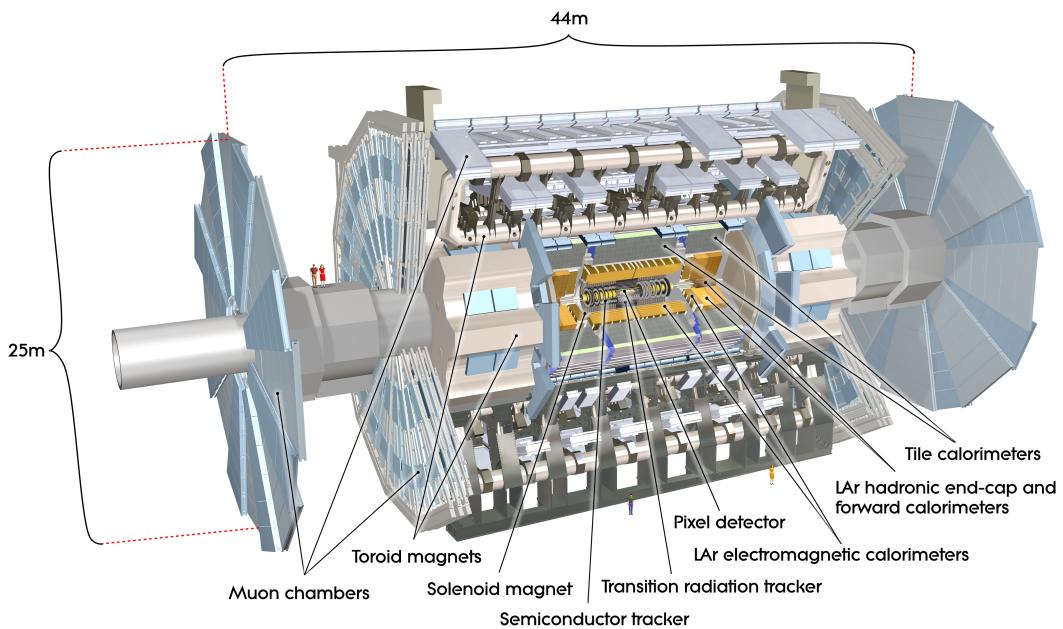


Figure 6: Schematic diagram of the ATLAS detector. Source: [19]

#### 194 3.1 The Inner Detector

195 The inner detector (ID), located nearest the the beam pipe, is specialized for charged particle  
196 tracking. It is immersed in a 2T magnetic field oriented parallel to the beam pipe, which bends  
197 the trajectories of electrically charged particles as they pass through the field. Three distinct but  
198 complementary tracking technologies are employed along with pattern recognition tools to map the  
199 trajectories of charged particles passing through the ID. These reconstructed tracks are a critical

200 component of vertex reconstruction, and the degree of bending and direction of the bent tracks  
201 at the production vertex provide information about the momentum, charge, and identity of the  
202 charged particles that produced them.

## 203 **3.2 The Calorimeter**

204 The calorimeter is designed to measure the energy of all particles which pass through it by initiating  
205 “showers” and fully absorbing their energy. The only particles which cannot be absorbed by the  
206 calorimeter are muons and neutrinos, which pass through without showering. The calorimeter is  
207 divided into two sub-detectors, the electromagnetic and hadronic calorimeters. Both are “sampling  
208 calorimeters”, which means they are comprised of repeated layers of dense absorbing material with  
209 “sampling” layers in between. The sampling layers track the location of the shower and record a  
210 small fraction of its energy, to which a calibration factor can be applied to infer the full shower  
211 energy.

### 212 **3.2.1 Electromagnetic Calorimeter**

213 The electromagnetic (EM) calorimeter forms the inner calorimeter layer, and is designed to fully  
214 absorb and measure the energies of electrons and photons. Energy is deposited in the iron-steel  
215 absorbing layers in the form of EM showers [20], in which the initial electron or photon interacts  
216 with the absorbing material to produce a cascade of photon radiation and electron pair production.  
217 The sampling layers are filled with liquid argon (LAr), with sensors to measure the ionization  
218 produced by charged particles passing through the LAr [21].

### 219 **3.2.2 Hadronic Calorimeter**

220 The hadronic calorimeter surrounds the EM calorimeter, and is designed to fully absorb and mea-  
221 sure hadronic showers, also known as “jets”, initiated by hadrons, which can make it through the  
222 EM calorimeter due to their relatively long interaction length [19]. Unlike the EM showers de-  
223 scribed above which proceed exclusively via electromagnetic interactions, hadronic showers pro-  
224 ceed via both the strong and EM interactions, and as a result are in general more variable and less  
225 localized.

226 The hadronic calorimeter is comprised of a tile calorimeter which encircles the EM calorimeter  
227 barrel, and a LAr calorimeter with copper and tungsten absorbers in the end-cap region which  
228 encloses the two ends of the barrel. The tile calorimeter uses steel as the absorber material and  
229 scintillators read out by photomultiplier tubes (PMTs) in the sampling layers.

## 230 **3.3 The Muon Spectrometer**

231 The muon spectrometer [19] surrounding the calorimeter is specialized for tracking muons and  
232 measuring their momentum. It employs the same principle used in the inner detector of applying  
233 a strong magnetic field and measuring the resulting bent trajectories of the electrically charged  
234 muons passing through to infer their momentum.

235 The magnetic field is generated by rectangular superconducting “toroid magnets” arranged  
236 azimuthally in radial planes around the beam axis, which set up a toroidal field concentric to the

237 beam axis. In the region containing the strong field established by the toroid magnets, muon  
238 tracks are recorded by three cylindrical layers of muon tracking chambers in the barrel region and  
239 three layers of chambers arranged in wheels perpendicular to the beam axis in the end-cap region.  
240 Additional layers of fast trigger chambers deliver muon track information to the ATLAS trigger  
241 system so it can be incorporated into the event readout decision.

## 242 **4 Monte Carlo Simulation of Signal and Background**

243 To search for evidence of new physics in the ATLAS data, it is necessary to model the number of  
244 SM “background” events that are expected in the data, as well as their distributions in the many  
245 observables, such as  $E_T^{\text{miss}}$ , that can be computed from the raw data. Only the SM processes which  
246 are expected to represent statistically significant backgrounds in the signal region are considered in  
247 the analysis. The data can then be compared with the SM background model to search for evidence  
248 of an excess of events above the SM background which may be consistent with a hypothesized  
249 signal model.

250 Signal and SM background models are generated by sophisticated Monte Carlo simulations,  
251 both of the physical production mechanisms and the passage of the final-state particles through the  
252 ATLAS detector.

253 For a given process, for example the Z+jets background shown in figure 10 left, “truth-level” in-  
254 formation for each MC event is first obtained from a random proton-proton collision by simulating  
255 the physical production mechanism for the process. The simulated final state products, along with  
256 their kinematic information, are collectively known as the “truth-level” event. Truth-level events  
257 can subsequently be passed through a simulation of the ATLAS detector to statistically model how  
258 they would actually be measured by the detector at “reconstruction-level”.

259 Given sufficient MC statistics, the kinematic distributions of reconstruction-level events, scaled  
260 by the integrated LHC beam luminosity and production cross section for the process, provide a  
261 statistical model of how the given process would appear in the ATLAS data.

## 262 **5 Ongoing Search for the Dark Higgs Model in the Hadronic 263 Decay Channel**

264 This section briefly summarizes the “mono-s(WW)” search for the dark higgs model in the fully-  
265 hadronic final state channel, which is currently undergoing approval within the ATLAS collabora-  
266 tion.

267 MC simulated samples were generated for the dark higgs model over the grid of dark higgs and  
268 Z’ boson masses shown in figure 7.

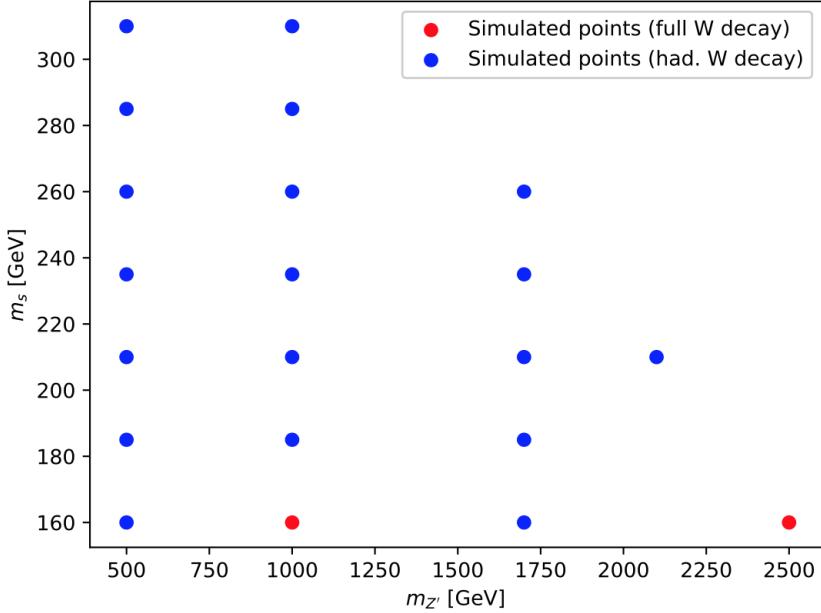


Figure 7: Grid of signal models generated using MC for the hadronic mono-s(WW) analysis at various dark higgs and  $Z'$  boson masses. Red points are simulated with all three possible WW decay channels ( $WW \rightarrow qqqq$ ,  $WW \rightarrow qql\nu$  and  $WW \rightarrow \ell\nu l\nu$ ) and blue points are simulated exclusively with fully hadronic WW decays ( $WW \rightarrow qqqq$ ) to save computing time.

## 269 5.1 Signal Region Selection

### 270 5.1.1 Pre-selection

271 Selection cuts are applied to the data and MC samples in the “signal region” (SR) with the aim  
 272 of maximizing the sensitivity of the data to the dark higgs signal of interest in the fully hadronic  
 273 decay channel. To avoid biasing the selection based on features in the data, the SR is ‘blinded’,  
 274 which means that the cuts are developed and optimized using the signal and background MC alone.

275 Prior to cut optimization, loose pre-selection cuts are applied to broadly capture the region of  
 276 interest for the search. The pre-selection cuts used by the hadronic analysis are listed in the first  
 277 column of Table 1, which is taken from the support note for the hadronic analysis.

278 The first three cuts and the final cut listed in Table 1 - 0 baseline lepton ( $\ell$ ),  $E_T^{\text{miss}} > 200$  GeV,  
 279  $\tau$  veto and  $N(\text{small-}R \text{ jets}) \geq 2$  - perform the basic selection for a fully hadronic final state with  
 280 high  $E_T^{\text{miss}}$ .

281 A veto is placed on variable radius (VR) track jets tagged as having been produced from b  
 282 quarks (i.e. “b-tagged”) is designed to reduce the “ $t\bar{t}$ bar” background of standard model top quark  
 283 pair production, as top quarks decay almost exclusively to b quarks. A lower bound of 15 is placed  
 284 object based  $E_T^{\text{miss}}$  significance  $\mathcal{S}$  with the aim of removing events with high  $E_T^{\text{miss}}$  originating from  
 285 poor detector energy resolution. Lastly, to a lower bound of 0.35 is placed on the angular separation  
 286 between the leading 3 small-radius jets and the  $E_T^{\text{miss}}$  to select for the desired topology of dark matter  
 287 recoiling against jets.

Table 1: The baseline selection which is required in all the signal and control regions.

$0\ell$ SR	$1\mu$ control region	$2\ell$ control region
0 baseline $\ell$ $E_T^{\text{miss}} > 200 \text{ GeV}$ $\tau$ veto	0 baseline e, 1 baseline+signal $\mu$ $E_T^{\text{miss,no}\mu} > 200 \text{ GeV}$ $\tau$ veto	2 baseline+signal $ee$ / opposite sign $\mu^+\mu^-$ $p_T^{\ell\ell} > 200 \text{ GeV}$ $\tau$ veto
0 $b$ -tagged VR track jets in event VR track jet $\Delta R$ overlap veto	0 $b$ -tagged VR track jets in event VR track jet $\Delta R$ overlap veto	0 $b$ -tagged VR track jets in event VR track jet $\Delta R$ overlap veto
$\Delta\phi_{\text{jets}_{1,2,3}, E_T^{\text{miss}}} > 0.35$ $\mathcal{S} > 15$ $N(\text{small-}R \text{ jets}) \geq 2$	$\Delta\phi_{\text{jets}_{1,2,3}, E_T^{\text{miss}}} > 0.35$ $\mathcal{S} > 15$ $N(\text{small-}R \text{ jets}) \geq 2$	$\mathcal{S} < 15$ $N(\text{small-}R \text{ jets}) \geq 2$

### 5.1.2 Kinematic Categories

In addition to the pre-selection cuts described in Section 4.1.1, further selections are optimized separately in the three kinematic categories illustrated in figure 8, known as “merged”, “intermediate” and “resolved”.

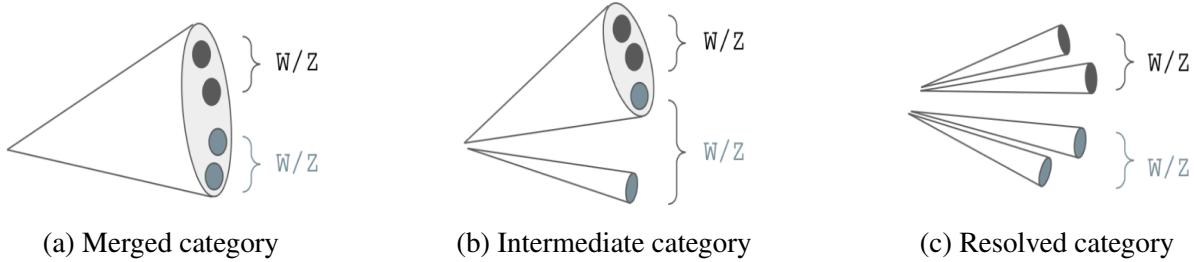


Figure 8: Illustration of the kinematic categories used for event selection optimization in the analysis of the mono-s(WW) hadronic channel. Source: Adapted from a figure shown by P. Gadow in an ATLAS internal approval presentation.

The “merged” category represents regime in which the hadronic final state is sufficiently boosted in the transverse plane that the quark hadronizations identified as having originated from the hadronic WW decays cannot be individually resolved, and are instead reconstructed as a single large-radius jet. In the “resolved” category, the boost is sufficiently small that all four quark hadronizations identified as having originated from W decays can be resolved into individual jets. In between these two extremes is the “intermediate” category, which is characterized by two or three of the hadronized quarks being reconstructed together as a large-R jet and the others being resolved into individual jets.

## 5.2 Control regions

Control regions (CRs) are designed to provide data-driven normalization constraints for the dominant Z+jets and W+jets standard model background processes, shown in figure 10. The CR is chosen by selecting a region of phase space that is kinematically similar to the SR, and which contains a high purity of the background process to be normalized. The CR must be orthogonal to the SR to avoid unblinding any part of the SR. To minimize the possibility of any signal events in

306 the data affecting the normalization of the standard model background process, it is also desirable  
 307 to select a CR in which negligible contamination of the MC signal events is observed. Data and  
 308 MC distributions are compared in the intermediate kinematic category of the two CRs in figure 9.

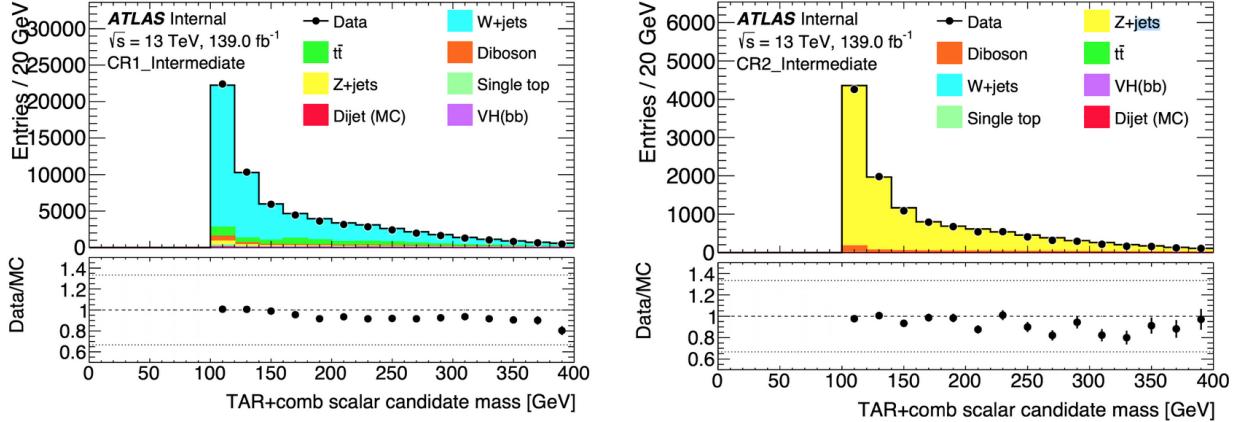


Figure 9: Data-MC comparisons of reconstructed dark higgs candidate mass distributions in the intermediate kinematic category of the 1-lepton (left) and 2-lepton (right) CRs

### 309 5.2.1 1 Lepton CR

310 The 1 lepton CR is identical to the SR, except that the lepton veto is replaced by requirement of  
 311 one signal muon. This CR is designed to obtain a purified sample of the SM W+jets background  
 312 in which the  $E_T^{\text{miss}}$  arises from a neutrino produced in the leptonic decay of a W boson, and thereby  
 313 provide a constraint on the normalization of this background process. In the fully hadronic signal  
 314 region, the lepton in this process is not properly reconstructed, and escapes detection.

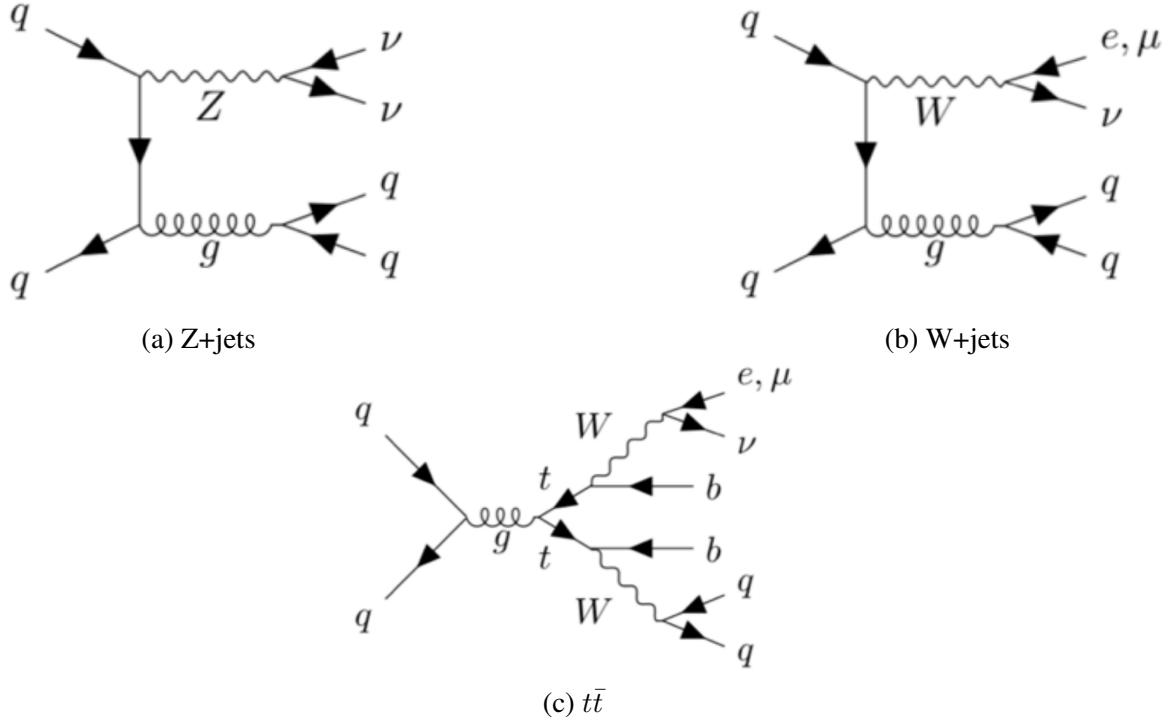


Figure 10: Dominant SM background processes for the mono-s(WW) search. The Z+jets background is only dominant in the hadronic channel, as it is significantly reduced by the 1-lepton requirement in the semileptonic channel.

### 315 5.2.2 2 Lepton CR

316 The 2 lepton CR is also largely identical to the SR. The primary difference is the replacement of  
 317 the lepton veto by a requirement of two signal muons or electrons. This CR is designed to obtain  
 318 a purified sample of the Z+jets background produced by the leptonic decay of a Z boson. This is  
 319 used to obtain a normalization for the Z+jets background where the Z boson decays hadronically  
 320 in the signal region. In addition to the two-lepton requirement, the MET significance is reversed  
 321 and the  $\Delta\phi_{\text{jets}_{1,2,3} E_T^{\text{miss}}}$  cut is removed to boost the contribution from QCD background.

### 322 5.3 Systematic Uncertainties

323 In addition to the statistical uncertainty arising from the limited number of events present in the  
 324 data and simulated in the Monte Carlo, systematic uncertainties arising from both theoretical and  
 325 experimental sources which can affect the normalization and shape of observables used in the  
 326 analysis. These systematic uncertainties are accounted for in the analysis by evaluating the effects  
 327 of variations in the Monte Carlo generated samples arising from parameters in the MC production  
 328 being varied from their nominal values by their respective systematic uncertainty.

329 **5.3.1 Experimental systematics**

330 Experimental systematic uncertainties arising from various aspects of modelling the ATLAS de-  
331 tector and LHC machine are provided by the relevant ATLAS performance groups. These include  
332 the limited precision of the LHC beam luminosity, uncertainties associated with lepton and muon  
333 reconstruction and identification, and jet energy resolution. Uncertainties associated with the re-  
334 construction of the missing transverse momentum are also considered.

335 **5.3.2 Theoretical Modelling Uncertainties**

336 Theoretical modelling uncertainties associated with the generation of Monte Carlo samples for  
337 both signal and the dominant Z+jets and W+jets SM background processes are also evaluated and  
338 incorporated in the analysis. Variations arising from the modelling of the parton density functions  
339 (PDFs) [22] of the colliding protons are considered, along with the renormalization and factoriza-  
340 tion scales [23] used to model QCD processes.

341 For the W+jets and Z+jets background processes, which are nominally modelled with the  
342 SHERPA generator [24] the variation arising from generating the events with an alternative event  
343 generator called PYTHIA 8 [25] are also considered. The alternative generator uses a different  
344 algorithm to model parton showering, with all other settings in the alternative generator kept the  
345 same.

346 **5.4 Sensitivity Estimates**

347 To estimate the analysis sensitivity for the hadronic channel, the data is binned in  $E_T^{\text{miss}}$  and dark  
348 higgs candidate mass. Mock binned data is generated for each signal point to be statistically  
349 consistent with the simulated SM background, and is subsequently fit to the SM background and  
350 signal model for events passing the event selection criteria. The W+jets and Z+jets background  
351 shapes are obtained in the fit from Monte Carlo, and their normalization factors from the 1-lepton  
352 and 2-lepton CRs, respectively.

353 The fit is used to obtain a CLs value [26] for each signal point, which can also be expressed  
354 as a “significance” by treating the CLs as a p-value and computing the corresponding standard  
355 deviation from the null hypothesis. Signal points with CLs value below 0.05, or significance above  
356 1.96, could be excluded for the hypothesized production cross section with 95% confidence if the  
357 data is statistically consistent with the background.

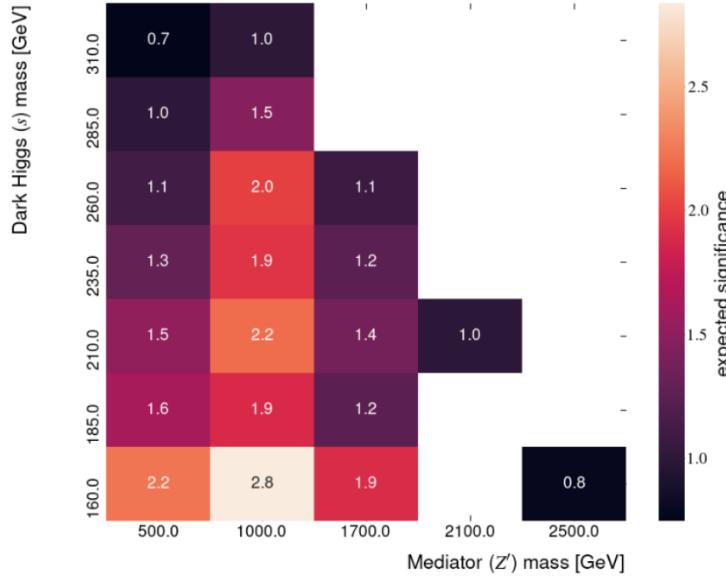


Figure 11: Hadronic channel with all analysis regions and systematic uncertainties

Figure 12: Sensitivity estimate for hadronic channel, expressed in terms of significance. The fit includes all analysis regions and systematic uncertainties.

## 358 6 Signal Sample Generation

### 359 6.1 Parameter Choices

360 MC signal samples for the dark higgs model with  $s \rightarrow WW$  decay are generated over a range of  
 361  $m(s)$  and  $m(Z')$  for which the analysis is expected to be sensitive to the model. There are several  
 362 other free parameters in the dark higgs model which need to be fixed to perform the scan.

- 363 • Coupling  $g_q$  of the  $Z'$  boson to quarks: 1
- 364 • Coupling  $g_\chi$  of the  $Z'$  boson to dark matter: 0.25
- 365 • Dark matter mass  $m_\chi = 200$  GeV
- 366 • Mixing angle  $\theta$  between the SM and dark sector higgs bosons: 0.01

367 The values of  $g_q$ ,  $g_\chi$  and  $m_\chi$ , listed above, are chosen to match the choices used for other LHC  
 368 searches for ease of comparison, and the value of  $\theta$  is chosen at the suggestion of the authors of  
 369 the phenomenology paper [1] which introduced the model.

### 370 6.2 Semileptonic Signal Grid

371 A signal grid has been generated for the mono- $s$ (WW) signal with semileptonic WW decay,  
 372 presently at the same points as were generated with hadronic decay for the hadronic analysis de-  
 373 scribed above. There are additionally two signal points generated with all possible WW decay  
 374 modes (i.e. ‘inclusive’).

375 Work is ongoing to re-generate the grid with a more recent release of the ATLAS simulation  
 376 software, and extend it as shown in Fig. 13 to allow for a more complete coverage of the phase  
 377 space region to which the analysis could be sensitive.

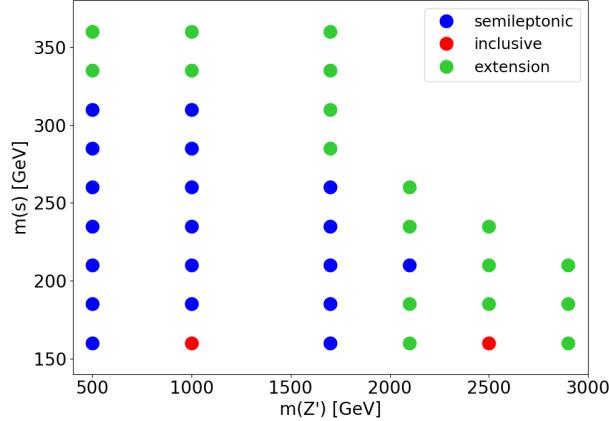


Figure 13: Grid of signal points with respect to dark higgs (s) and  $Z'$  boson masses

## 378 7 Event Selection for Semileptonic Channel

379 This section describes the current event selection used to optimize the sensitivity of the analysis to  
 380 the dark higgs signal with semileptonic WW decay.

### 381 7.1 Preselection

382 As was done for the hadronic analysis, several pre-selection cuts are applied to broadly define the  
 383 channel prior to further cut optimization. The current pre-selection cuts are listed below.

- 384 • Event passes  $E_T^{\text{miss}}$ trigger
- 385 • 1 signal muon or 1 signal electron
- 386 •  $E_T^{\text{miss}} > 150 \text{ GeV}$

387 In addition to the current preselection cuts, the following two cuts are common to both cate-  
 388 gories, and may become listed as pre-selection in the future:

- 389 • Veto on b-tagged jets
- 390 • Lepton- $E_T^{\text{miss}}$ transverse mass  $m_T(\ell, E_T^{\text{miss}}) > 150 \text{ GeV}$

391 The transverse mass  $m_T(E_T^{\text{miss}}, \ell)$  between the lepton and  $E_T^{\text{miss}}$ is considered in the semileptonic  
 392 channel, because it is sensitive to additional  $E_T^{\text{miss}}$ on top of that arising from the neutrino in the  
 393 semileptonic W decay. It is computed in the ATLAS data and MC as:

$$394 m_T(\ell, E_T^{\text{miss}}) = \sqrt{2p_{T,\ell} E_T^{\text{miss}} (1 - \cos \theta_{\ell,E_T^{\text{miss}}})} \quad (1)$$

394 which comes from the full transverse mass definition [27]

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

395 under the assumptions that the masses of the lepton and  $E_T^{\text{miss}}$  are negligibly small compared with  
 396 their momenta. The assumption of negligible lepton mass is in general justified given the energy  
 397 of LHC collisions. The assumption of negligible  $E_T^{\text{miss}}$  mass is justified if the true  $E_T^{\text{miss}}$  arises only  
 398 from the neutrino in the semileptonic W, as it would in the leading SM backgrounds, but not in  
 399 the signal model which has additional  $E_T^{\text{miss}}$  mass arising from dark matter production. The result,  
 400 shown in figure 14, is that the bulk of the SM background has  $m_T(\ell, E_T^{\text{miss}})$  below the W mass  
 401 peak, but the signal distribution tends to be peaked closer to  $\sim 250$  GeV.

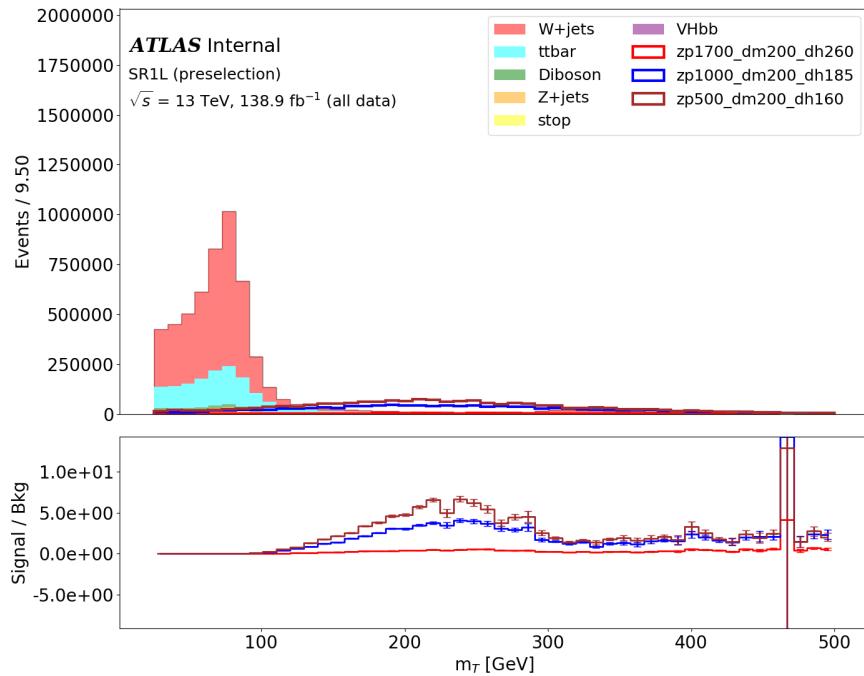


Figure 14: Transverse mass distribution for SM background and several signal points with pre-selection cuts

## 402 7.2 Kinematic Categories

403 Further cuts are optimized within the two kinematic categories illustrated in figure 16. As for the  
 404 hadronic analysis, the kinematic category definitions aim to characterize the degree to which the  
 405 leading hadronic jets in the event are boosted. The “merged” category aims to capture the highly  
 406 boosted regime, in which the final state is sufficiently boosted in the transverse plane that the  
 407 leading hadronic showers are reconstructed as a single large-radius jet. The “resolved” category  
 408 captures the regime in which all hadron showers have a low enough boost as to be individually  
 409 reconstructed as small-radius jets.

410     The intermediate category used in the hadronic analysis, characterized by a combination of  
 411     large- and small-radius jets, is not considered for the semileptonic channel because the semilep-  
 412     tonic WW decay in the signal model produces only two final-state quarks.

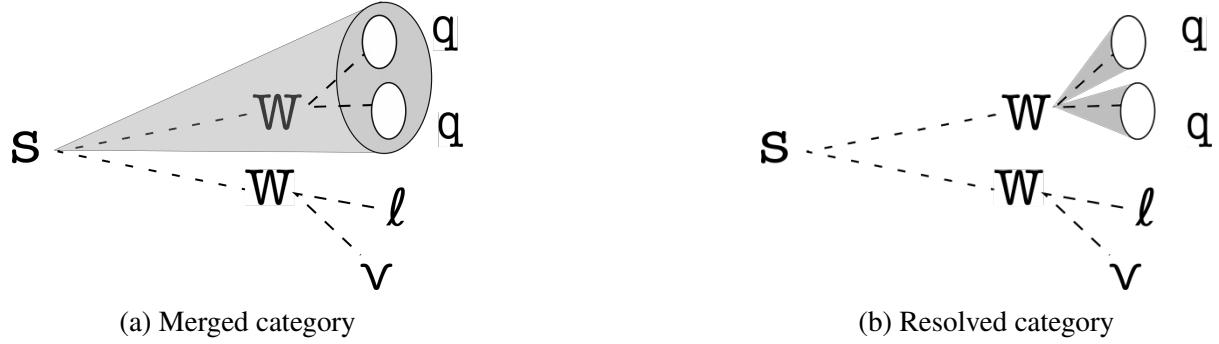


Figure 15: Kinematic categories for the semileptonic WW decay channel

413     The current merged and resolved category cut definitions, listed and discussed in tables 2 and  
 414     3, respectively, have been developed and optimized by two other members of the analysis team.  
 415     The  $E_T^{\text{miss}}$  distributions for each category are shown in figure ???. There is ongoing work within the  
 416     analysis team to investigate the benefit of tuning the selection further by optimizing the selections  
 417     within the  $WW \rightarrow q\bar{q}\nu\bar{\nu}$  and  $WW \rightarrow q\bar{q}\mu\nu$  decay channels separately.

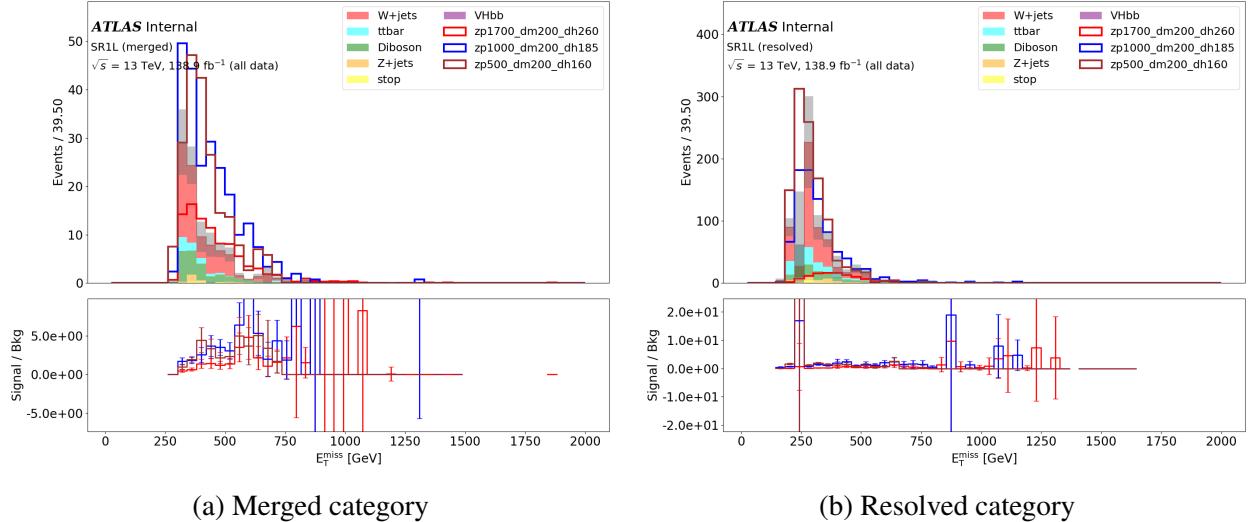


Figure 16:  $E_T^{\text{miss}}$  distributions in the SR in the two kinematic categories developed for the semilep-  
 tonic WW decay channel. Signals shown with amplitudes multiplied by a factor of 5.

Table 2: Current data selection cuts in the **merged** category of the semileptonic decay channel

Selection	Discussion
At least 1 large-radius jet	-
$E_T^{\text{miss}} > 300 \text{ GeV}$	-
Reconstructed mass of large-radius jet $\in [60 \text{ GeV}, 100 \text{ GeV}]$	Increases the likelihood that selected large-R jet originates from a W boson
$E_T^{\text{miss}}$ significance $> 15$	$E_T^{\text{miss}}$ significance gives a measure of the likelihood that the observed $E_T^{\text{miss}}$ arises from true $E_T^{\text{miss}}$ due to invisible particles, rather than from fluctuations associated with limited detector resolution.
$\Delta R(\ell, \text{large-R jets}) < 1.6$	$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ is the angular distance between the lepton and the large-radius jet, where $\Delta\phi$ is the angular separation in the transverse plane and pseudorapidity $\eta$ is a quantity related to the angle $\theta$ of a particle relative to the beam axis: $\eta = -\ln \left[ \tan \left( \frac{\theta}{2} \right) \right]$ . Pseudorapidity is used rather than $\theta$ because differences $\Delta\eta$ in pseudorapidity are invariant under lorentz boosts. An upper bound is placed on the angular distance to select for the boosted topology of interest, in which all decay products of the dark higgs boson are boosted in approximately the same direction.
$D_2$ of leading large-R jet $< 1.4$	The $D_2$ variable [28] uses energy correlation functions to give a measure of the likelihood that a jet contains only two spatially separated energy sources. This allows the semileptonic signal channel - in which only two dominant energy sources (i.e. “prongs”) would be expected due to the $W \rightarrow qq$ decay - to be separated from the QCD background which may contain many such prongs.

Table 3: Current data selection cuts in the **resolved** category of the semileptonic decay channel

Selection	Discussion
At least 2 small-radius jets	-
Reconstructed mass of the $W$ candidate $\in [60 \text{ GeV}, 100 \text{ GeV}]$	The $W$ candidate is selected by considering all combinations of small-radius jets, and identifying the pair whose reconstructed mass is closest to the $W$ boson mass of $\sim 80 \text{ GeV}$ .
$E_T^{\text{miss}}$ significance $> 13$	-
$m_T(\ell, E_T^{\text{miss}}) > 200 \text{ GeV}$	-
$W$ candidate $p_T > 100 \text{ GeV}$	The lower bound on $W$ candidate $p_T$ increases the likelihood that the two jets used to reconstruct the $W$ boson actually originated from the decay of a boosted parent particle.
$\Delta R(W, \ell) < 1$	An upper bound on the angular distance between the $W$ boson candidate and the lepton selects for the topology of interest in which one of the boosted $W$ bosons decays leptonically, such that in the lab frame, the lepton is expected to be measured in approximately the same direction as the reconstructed $W$ .

418 **7.3 W+jets Control Region**

419 The W+jets control region is defined with the aim of providing a data-driven estimate of the nor-  
 420 malization of the W+jets background, illustrated in figure 10b in the signal region. The exact  
 421 definition of this control region is still being studied and optimized, but at its simplest, it would  
 422 simply reverse the transverse mass cut:

$$m_T(\ell, E_T^{\text{miss}}) > 150 \text{ GeV} \rightarrow m_T(\ell, E_T^{\text{miss}}) \leq 150 \text{ GeV}$$

423 This reversal makes the W+jets CR orthogonal to the SR, and aims to boost the statistics and  
 424 purity of the W+jets background without drastically changing the kinematics of the region. Table  
 425 4 shows that the  $m_T(\ell, E_T^{\text{miss}})$  cut reversal improves the W+jets purity and vastly reduces the signal  
 426 contamination, but only improves the W+jets statistics in the resolved category.

Table 4: Comparison of W+jets statistics and purity, as well as signal contamination, for SR and preliminary W+jets CR. Signal contamination is averaged over all signal points, and the uncertainty is the standard deviation among all points.

Region	W+jets Statistics	W+jets Purity	Mean Signal Contamination
Merged SR	59.1	65.2%	$31\% \pm 11\%$
Resolved SR	337	56.8%	$13\% \pm 12\%$
Merged W+jets CR	53.8	78.5%	$0.49\% \pm 0.30\%$
Resolved W+jets CR	704	83.4%	$0.082\% \pm 0.080\%$

427 Additional cut reversals are being considered to further boost the statistics and purity of the  
 428 W+jets background in this CR, but careful study is needed to determine whether the region re-  
 429 mains sufficiently representative of the SR with these additional cut reversals to provide a reliable  
 430 normalization factor to be applied to the SR.

431 **7.4  $t\bar{t}$  Control Region**

432 An additional CR is being considered, but has yet to be studied, to constrain the second-to-most  
 433 dominant  $t\bar{t}$  background. At its simplest, this CR would simply reverse the b-jet veto, which is  
 434 primarily intended to reduce the  $t\bar{t}$  background in the SR:

$$N(\text{b-tagged jets}) < 1 \rightarrow N(\text{b-tagged jets}) \geq 1$$

435 **8 Sensitivity Studies for the Semileptonic channel**

436 Preliminary sensitivity studies have been performed for the semileptonic decay channel using the  
437 HistFitter statistical analysis framework [29].

438 **8.1 Background on limit setting**

439 **8.2 HistFitter setup**

- 440     • Fit variable
- 441     • Binning
- 442     • Use of temporary systematics proxy

443 **8.3 Preliminary sensitivity limits**

- 444     • Resolved
- 445     • Merged
- 446     • Combined

447 **9 Remaining Work and Outlook**

- 448     • Study variables to help address  $p_\nu$  vs.  $p_{\chi\bar{\chi}}$  MET ambiguity
- 449     • Finalize SR and CR definitions
- 450     • Experimental and modeling systematics
- 451     • Finalize sensitivity estimates with final systematics
- 452     • Statistical combination with hadronic channel for combined limits

453 **References**

- 454 [1] Michael Duerr, Alexander Grohsjean, Felix Kahlhoefer, Bjoern Penning, Kai Schmidt-  
455 Hoberg, and Christian Schwanenberger. Hunting the dark higgs. *Journal of High Energy  
456 Physics*, 2017(4), Apr 2017.
- 457 [2] M. Tanabashi et al. (Particle Data Group), Phys. Rev. D 98, 030001 (2018) and 2019 update.
- 458 [3] N. et al. Aghanim. Planck 2018 results. VI. Cosmological parameters. 7 2018.
- 459 [4] J. L. Feng. Dark matter candidates from particle physics and methods of detection. *Annual  
460 Review of Astronomy and Astrophysics*, 1(48):495–545, 2010.
- 461 [5] Marc Schumann. Direct detection of WIMP dark matter: concepts and status. *Journal of  
462 Physics G: Nuclear and Particle Physics*, 46(10):103003, aug 2019.
- 463 [6] Teresa Marrodan Undagoitia and Ludwig Rauch. Dark matter direct-detection experiments.  
464 *J. Phys. G*, 43(1):013001, 2016.
- 465 [7] Marco Cirelli. Indirect searches for dark matter. *Pramana*, 79(5):1021–1043, Oct 2012.
- 466 [8] Jan Conrad and Olaf Reimer. Indirect dark matter searches in gamma- and cosmic rays.  
467 *Nature Physics*, 13:224–231, 03 2017.
- 468 [9] Antonio Boveia and Caterina Doglioni. Dark matter searches at colliders. *Annual Review of  
469 Nuclear and Particle Science*, 68(1):429–459, 2018.
- 470 [10] Gerard Jungman, Marc Kamionkowski, and Kim Griest. Supersymmetric dark matter. *Phys.  
471 Rept.*, 267:195–373, 1996.
- 472 [11] O. Buchmueller, Matthew J. Dolan, and Christopher McCabe. Beyond Effective Field Theory  
473 for Dark Matter Searches at the LHC. *JHEP*, 01:025, 2014.
- 474 [12] Daniel Abercrombie, Nural Akchurin, Ece Akilli, Juan Alcaraz Maestre, Brandon Allen, Bar-  
475 bara Alvarez Gonzalez, Jeremy Andrea, Alexandre Arbey, Georges Azuelos, Patrizia Azzi,  
476 and et al. Dark matter benchmark models for early lhc run-2 searches: Report of the atlas/cms  
477 dark matter forum. *Physics of the Dark Universe*, 27:100371, Jan 2020.
- 478 [13] Csaba Csaki. The minimal supersymmetric standard model (mssm). *Modern Physics Letters  
479 A*, 11(08):599–613, Mar 1996.
- 480 [14] Stefano Profumo, Leonardo Giani, and Oliver F. Piattella. An introduction to particle dark  
481 matter, 2019.
- 482 [15] Search for Dark Matter Produced in Association with a Higgs Boson decaying to  $b\bar{b}$  at  $\sqrt{s} =$   
483 13 TeV with the ATLAS Detector using  $79.8 \text{ fb}^{-1}$  of proton-proton collision data. Technical  
484 Report ATLAS-CONF-2018-039, CERN, Geneva, Jul 2018.

- 485 [16] RECAST framework reinterpretation of an ATLAS Dark Matter Search constraining a model  
 486 of a dark Higgs boson decaying to two  $b$ -quarks. Technical Report ATL-PHYS-PUB-2019-  
 487 032, CERN, Geneva, Aug 2019.
- 488 [17] Kyle Cranmer and Itay Yavin. Recast - extending the impact of existing analyses. *Journal of*  
 489 *High Energy Physics*, 2011(4), Apr 2011.
- 490 [18] Lyndon Evans and Philip Bryant. LHC machine. *Journal of Instrumentation*, 3(08):S08001–  
 491 S08001, aug 2008.
- 492 [19] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003,  
 493 2008.
- 494 [20] Lev Davydovitch Landau, G. Rumer, and Piotr Leonidovich Kapitza. The cascade theory of  
 495 electronic showers. *Proceedings of the Royal Society of London. Series A. Mathematical and*  
 496 *Physical Sciences*, 166(925):213–228, 1938.
- 497 [21] Huaqiao Zhang and. The ATLAS liquid argon calorimeter: Overview and performance.  
 498 *Journal of Physics: Conference Series*, 293:012044, apr 2011.
- 499 [22] M. Tanabashi et al. (Particle Data Group), Phys. Rev. D 98, 030001 (2018) and 2019 update.  
 500 18. *Structure Functions*.
- 501 [23] M. Tanabashi et al. (Particle Data Group), Phys. Rev. D 98, 030001 (2018) and 2019 update.  
 502 9. *Quantum Chromodynamics*.
- 503 [24] T Gleisberg, S Hoche, F Krauss, M Schonherr, S Schumann, F Siegert, and J Winter. Event  
 504 generation with SHERPA 1.1. *Journal of High Energy Physics*, 2009(02):007–007, feb 2009.
- 505 [25] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. A brief introduction to pythia 8.1.  
 506 *Computer Physics Communications*, 178(11):852?867, Jun 2008.
- 507 [26] P.A. Zyla et al. (Particle Data Group), to be published in Prog. Theor. Exp. Phys. 2020,  
 508 083C01 (2020). 40. *Statistics*.
- 509 [27] P.A. Zyla et al. (Particle Data Group), to be published in Prog. Theor. Exp. Phys. 2020,  
 510 083C01 (2020). 47. *Kinematics*.
- 511 [28] Simone Marzani, Gregory Soyez, and Michael Spannowsky. Looking inside jets. *Lecture*  
 512 *Notes in Physics*, 2019.
- 513 [29] M. Baak, G. J. Besjes, D. Côté, A. Koutsman, J. Lorenz, and D. Short. HistFitter software  
 514 framework for statistical data analysis. *Eur. Phys. J. C*, 75:153, 2015.