

Dissertation
submitted to the
Combined Faculties of the Natural Sciences and
Mathematics
of the Ruperto-Carola-University of Heidelberg,
Germany
for the degree of
Doctor of Natural Sciences

Put forward by
Fabrizio Napolitano

born in Rome

Oral examination on July 29th, 2020

**Search for Dark Matter produced in
association with a hypothetical Dark
Higgs Boson decaying to $W^{\pm} W^{\mp}$ or
 ZZ boson pairs in the fully hadronic
final state at $\sqrt{s} = 13$ TeV using 139
 fb^{-1} of pp collisions recorded with the
ATLAS Detector**

Referees:

Priv.-Doz. Dr. Oleg Brandt

Prof. Dr. André Schöning

Abstract

Searches for Dark Matter, one of the biggest unsolved problems of modern physics, at the LHC often comprise the most exotic signatures. At the same time, these pose experimental challenges, but also opportunities: new techniques are developed, which can find useful application also beyond the scope of their original conception. This thesis presents a search for Dark Matter produced in association with an hypothetical Dark Higgs Boson decaying to VV , $V = W, Z$, in the fully hadronic final state. The reconstruction of the boosted V bosons is performed with a novel technique, the Track-Assisted Reclustered jets, used in a data analysis for the first time. This is shown to achieve a more robust performance and flexibility with respect to standard methods. The search was performed using 139 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13 \text{ TeV}$ collected with the ATLAS detector during Run 2. No significant excess is found in the observed data over the Standard Model; the upper limits on the production cross-section of $E_T^{\text{miss}} + VV$ in the fully hadronic final state for the Dark Higgs scenario are set at 95% confidence level.

Kurzzusammenfassung

Die Suche nach Dunkler Materie, eines der größten ungelösten Probleme der modernen Physik, umfasst am LHC oft die exotischsten Signaturen. Diese stellen gleichzeitig experimentelle Herausforderungen, aber auch Chancen dar: Es werden neue Ansätze entwickelt, die auch über den Rahmen ihrer ursprünglichen Konzeption hinaus nützliche Anwendung finden können. Diese Dissertation präsentiert eine Suche nach Dunkler Materie, die in Verbindung mit einem hypothetischen Dunklen Higgs-Boson produziert wird, das im vollhadronischen Endzustand zu VV , $V = W, Z$, zerfällt. Die Rekon-

struktions der geboosteten V -Bosonen erfolgt mit einem neuartigen Ansatz, den Track-Assisted Reclustered Jet, der zum ersten Mal in einer Datenanalyse verwendet wird. Es hat sich gezeigt, dass damit im Vergleich zu herkömmlichen Methoden eine robustere Leistung und Flexibilität erreicht wird. Die Suche wurde mit 139 fb^{-1} an Proton-Proton-Kollisionsdaten bei $\sqrt{s} = 13 \text{ TeV}$ durchgeführt, die mit dem ATLAS-Detektor während Run 2 aufgezeichnet wurden. In den analysierten Daten wurde kein signifikanter Überschuss gegenüber dem Standardmodell gefunden; die Obergrenzen für den Produktionsquerschnitt von $E_T^{\text{miss}} + VV$ sind im voll-hadronischen Endzustand für das Szenario mit dem Dunklen Higgs-Boson auf 95% Konfidenzniveau festgelegt.

Acknowledgements

I would like to thank my supervisor, Priv.-Doz. Dr. Oleg Brandt, for the incredible support and guidance he gave me over these years, introducing me to the High Energy Physics studied when I was still a master student back in 2014, up to today with the completion of this PhD. Thank you, Oleg. I would like then to thank Prof. Dr. Hans-Christian Schultz-Coulon for always keeping an eye on me, directing and supporting me through the academic system, and for allowing me to spend a full year at CERN. I would like to thank as well Prof. Dr. André Schöning for agreeing to referee yet another thesis of mine. I am grateful to the Max Planck Institut für Kernphysik and the IMPRS-PTFS for admitting me to their prestigious program, and their financial support. I owe gratitude to the Heidelberg HGSFP and the Heidelberg University for the interesting and very formative lectures, and for the opportunity to get teaching experience as tutor. Thank you also to all my ATLAS KIP colleagues for the friendly environment and the help during these years. Thanks to Sebastian and Philipp for proofreading part of this work. Thanks to Prof. Clara Matteuzzi for the interesting discussions over a coffee we always have when passing by in building 13, and for her contagious enthusiasm when exploring new ideas.

I am grateful for the opportunity to explore the forefront of science with the ATLAS detector at the LHC, and for the high standards of research and the support of the other members of the collaboration. I want to thank my colleagues of the TAR/ R -scan effort, and the members of the Mono- $s(VV)$ analysis for this incredible journey at the heart of fundamental physics, in particular Dr. Stanislava Sevova and Philipp Paul Gadow, for the incredible energy and focus to get this analysis done and published. Thanks a lot the JDM subgroup and the JetEtMiss group, in particular to Dr. Jason Veatch, Dr. Samuel Mehan for the help and scientific support. I want to thank all the colleagues of L1Calo who worked with me at CERN and allowed me to conduct the most exciting detector operations I could imagine. Finally I want to thank the INFN and the group lead by Dr. Catalina Curceanu for trusting that I would indeed get a

PhD at some point, and for allowing me to contribute to their scientific effort from day one.

I wish to thank all my dear friends in particular Federico and Stanislav. Not enough words can be spent to thank my family & my extended family with Antonello, Simona, Michele and Lucia. Finally and most importantly I want to thank my wife for always believing (maybe too much!) in me, and making me look at things from a different perspective, and for the incredible support.

Contents

1. Introduction	1
1.1. Author's Contributions	2
I. Theoretical Framework	5
2. The Standard Model of Particle Physics	7
2.1. Fundamental Particles	8
2.2. Mathematical Formulation	10
2.2.1. Electroweak Theory	10
2.2.2. The Higgs Mechanism	12
2.2.3. Quantum Chromodynamics	14
2.3. Shortcomings of the Standard Model	15
3. The puzzle of Dark Matter	19
3.1. Dark Matter Searches	28
3.2. The Dark Higgs Model	30
3.2.1. Theoretical Framework	31
3.2.2. Searches at LHC	32
II. Experimental Techniques	37
4. The ATLAS Detector at the LHC	39
4.1. The LHC	39
4.2. The ATLAS Experiment	43
4.2.1. The ATLAS Coordinates	44
4.2.2. The Inner Detector	45
4.2.3. The Calorimeters	47
4.2.4. The Muon Spectrometers	52
4.2.5. The ATLAS Magnets	53
4.2.6. The Trigger system	54

5. The Track-Assisted Reclustered (TAR) jets	57
5.1. Jets and Reconstruction Techniques	57
5.1.1. Large- R Jets and Substructure	64
5.2. TAR Definition, Performance and Potential	70
5.2.1. Track-Assistance	70
5.2.2. Jet Re-clustering	73
5.2.3. TAR Definition	73
5.2.4. TAR Performance	76
5.3. Jet Calibration and Uncertainties	77
III. Search for Dark Matter produced in association with a Dark Higgs boson decaying to VV in the fully hadronic final state	83
6. Motivation and Overview	85
7. Data, Triggers and Simulated Samples	89
7.1. Data	89
7.2. Trigger	89
7.3. Simulated Samples	90
7.3.1. Background Samples	90
7.3.2. Signal Samples	91
8. Objects, Selections and Simulations	95
8.1. Objects	95
8.1.1. Electrons	95
8.1.2. Muons	96
8.1.3. Taus	96
8.1.4. Track Selection	97
8.1.5. Small- R Jets	98
8.1.6. TAR Jets	98
8.1.7. Missing Transverse Momentum E_T^{miss} and Object based E_T^{miss} Significance \mathcal{S}	99
8.1.8. Variable- R Track Jets	100
8.2. Event Selection	101
8.2.1. Baseline Selection	101
8.2.2. Categories Definition	103

8.3. Data and Monte Carlo Comparison	109
9. Systematic Uncertainties	111
9.1. Experimental Systematic Uncertainties	111
9.2. Theoretical Systematic Uncertainties	115
10. Statistical Framework	121
11. Results	125
11.1. Fit Data to Model	125
11.1.1. Post-fit Distribution	128
11.2. Experimental Search Significance	128
11.3. Exclusion Limits on the Dark Higgs Model	131
12. Conclusions and Outlook	135
IV. Appendices	137
A. Analysis Appendix	139
A.1. Constraints on the Dark Higgs Model	139
A.2. Triggers used	140
A.3. Generic Limits Procedure	140
A.3.1. The $E_T^{\text{miss}} + H(b\bar{b})$ RECAST Example	142
A.3.2. The $E_T^{\text{miss}} + s \rightarrow W^+ W^- (had)$ Sensitivity Estimate	144
A.4. Overlap Removal	146
A.5. TAR+Comb	147
A.6. Full Fit Results	149
A.7. Resolved Category	153
A.8. Event Yield	155
A.8.1. Pre-fit Event Yield Tables	155
A.8.2. Post-fit Event Yield Tables	157
A.9. Expected Significance Breakdown	159
A.10. Exclusion Limits for Different Z' masses	162
A.10.1. Exclusion Limits Tables	163
B. Full Data/MC Comparison	165
B.0.1. Signal Region	165

B.0.2. One μ Control Region	175
B.0.3. Two Leptons Control Region	185
C. Systematics Uncertainties Additional Tables	195
C.0.1. Acceptance Uncertainties	195
D. Pre-fit Distribution	201
E. Introduction to QFT Lagrangians	205
E.0.1. Introduction: QED	205
E.0.2. Electroweak Theory	207
Bibliography	209
List of figures	221
List of tables	235

“Quid ei potest videri magnum in rebus humanis, cui aeternitas omnis, totiusque mundi nota sit magnitudo.”

“What, among the things of this world, can be important to those who know the eternity and the vastness of the universe?”

— Marcus Tullius Cicero – Tusculanae Disputationes

1. Introduction

“If you do not expect the unexpected, you will not find it; for it is hard to be sought out and difficult.”

— Heraklitus, fragment B 18

Particle Physics accounts for the most successful theories and experimental validations among all the quantitative sciences. Most of them required an enormous effort of minds and hands to be achieved, often spanning through the lifetime of single individuals. Perhaps the most striking of them is the Standard Model: built with dedication during several of the last decades, it was always accompanied by the successful operation of the biggest scientific endeavours with increasing complexity and precision. Often the experimental activity paved the way to a more complete and profound understanding of the underlying laws of Nature.

This in fieri development of the Standard Model continues until today, where the predictions of the theory are tested with an even increasingly accuracy, and new hypothesis are suggested and investigated in attempt to shed light on the challenges that the Standard Model faces. One of the most notable examples of these challenges is the puzzle of Dark Matter. Dark Matter is estimated to compose about 85% of the total matter content of the universe and around a quarter of the total energy density. Solid astrophysical evidence through the years corroborate the argument of its physical existence; yet its very nature remains unveiled, and our successful theories fail to address the problem of its origin and its properties.

If Dark Matter interacts with Standard Model particles via the exchange of new hypothetical mediators, it might be produced at the Large Hadron Collider, leaving a characteristic signatures that could be detected by the ATLAS detector. In the first part of this thesis, a theoretical model called the *Dark Higgs* will be introduced, in which the Dark Matter mass is generated through a Higgs mechanism in the Dark Sector. The final state of such characteristic signature is dependent on the mass of the Dark Higgs: above a kinematic threshold it is possible for the Dark Higgs to decay to a pair of vector bosons.

The second part will explain the experimental observables and techniques that have been devised in order to deal with the increasingly challenging final state, in particular in reference to the reconstruction of highly boosted hadronically decaying object. The *Track-Assisted Reclustered* (TAR) jet reconstruction algorithm achieves an optimal performance in such extreme final states, providing excellent background rejection using flexible track-based jet substructure and mass reconstruction. The Track-Assisted Reclustered jets required a dedicated chain of calibrations and uncertainties, which are presented as well.

The third and final part of the thesis is dedicated to the search for Dark Matter produced in association with a Dark Higgs decaying to VV bosons, $V = W/Z$, in the fully hadronic final state using 139 fb^{-1} of data collected during Run 2 at the LHC, referred to as $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ throughout this thesis. The search divides the phase space into different kinematic regions, in order to account for the different regimes of the Dark Higgs as a function of its transverse momentum.

1.1. Author's Contributions

Large collaborations are needed in High Energy Physics due to the utmost technical complexity on every side of the scientific operation and data analysis of the experiments. No results would have been possible without the effort of the work of hundreds (in this case notably even thousands) of physicist and engineers. The author is very grateful to the ATLAS collaboration for being part of this effort. Yet, for clarity, the author lists here his personal contributions to this project, starting with the detector operational activities, over the development of dedicated performance observables, and ending with the search for Dark Matter.

The author has worked in fulfilment of his ATLAS author qualification task on the first level calorimeter trigger (L1Calo). This contribution encompasses: monitoring and performing the calibration of the timing of the system, using collimator beam splash events at start of 2017 data taking, special physics runs with extended readout, and calibration runs. Moreover, the author has contributed to the factual operation of the detector as L1Calo on-call expert and as Run Controller directly in the ATLAS Control Room.

Subsequently, the author has significantly contributed to the commissioning of the TAR jets, which he co-developed already during the time of his master thesis. The author focused on the required calibration chain for the anti- k_t $R=0.2$ jets: he provided the Monte Carlo jet energy scale calibration framework setup, and performed the Global Sequential Calibration in the n_{trk} , w_{trk} and N_{seg} variables. The author provided uncertainties on the anti- k_t $R=0.2$ jets: the flavour uncertainty and out-of-cone uncertainty, which he re-derived for anti- k_t $R=0.4$ and PFlow jets as well. These were used to produce the latest ATLAS recommendations [1].

The author's contributions to the Dark Higgs search are the exploration and validation of the phase space and the kinematic properties of the Dark Higgs model, the estimation of the sensitivity of the ATLAS detector using the Generic Limits procedure from the $E_T^{\text{miss}} + H(b\bar{b})$ [2] and $E_T^{\text{miss}} + V(qq)$ [3] analyses, which was used to request the Monte Carlo production of the signal samples. This was done both for $s \rightarrow W^+W^-$ final state and for the reinterpretation of the $E_T^{\text{miss}} + H(b\bar{b})$ analysis in the $s \rightarrow b\bar{b}$ final state [4]. In addition, the author provided and validated a data reduction scheme that allows the use of the TAR jets for different beyond the Standard Model searches. He implemented, validated and investigated the performance of this novel technique for the first time in the analysis framework, optimizing its parameters in the signal region for the different kinematic regimes and evaluating the sensitivity gains with respect to standard methods. The author developed a new algorithm to maximize the signal significance, combining TAR jets and anti- k_t $R=0.2$ jets, and investigated optimized kinematic observables. Moreover, he derived the theoretical uncertainties associated to the main background of the search, the V +jets process, evaluating the impact of scale, PDF, and alternative hadronisation models. Finally, the author has investigated the compatibility of the ATLAS standard anti- k_t $R=0.4$ jet energy resolution uncertainties for anti- k_t $R=0.2$ jets with a dedicated in-situ direct balance measurement. The author did produce all the plots present in this thesis unless stated otherwise and/or referred to the proper sources.

Part I.

Theoretical Framework

2. The Standard Model of Particle Physics

The Standard Model of particle physics (SM) is a quantum field theory that describes all the properties and interactions of the known particles. It is the product of a number of brilliant ideas, intuitions and experiments; but also false starts and misunderstandings that made them possible. It took its first steps in the 50's, amid the confusion that characterized the field of particle physics in that period. Despite the success of the quantum electrodynamics (QED) in the previous decade, the weak and the strong force and the problem of symmetries had still to be understood from a theoretical standpoint. In 1954 Yang and Mills [5] extended a gauge (or local) theory from the one-dimensional group $U(1)$ of QED to the $SU(2)$ of the isotopic spin conservation. This enabled a way to formalize, via the non-abelian property, the self interaction of the gauge bosons. However, it was soon realized that gauge symmetries forbid gauge boson from being massive. Mass term inserted by hand are in fact non renormalizable. This problem paved the way to the development of the concept of spontaneous symmetry breaking, which was developed by Higgs [6] and independently by Englert and Brout [7] and Guralnik, Hagen, and Kibble [8]. Goldstone demonstrated [9] [10] that for every broken symmetry there must be a massless boson, unless moving from global symmetries to gauge symmetries. In 1961 Glashow [11] found the global group structure $SU(2) \times U(1)$. This group allows the presence of charged massive particles W and Z , and the photon, combining the QED with the weak interaction into the electroweak (EW) force. Starting from this work, Weinberg [12] and Salam [13] incorporated the Higgs mechanism into the EW force: the SM was getting its modern form.

The strong force was still missing, and it had to wait until the 1970's. In 1973 Gross, Wilczek [14] and Polizer [15] discovered the property of asymptotic freedom in non-abelian gauge theories. Following the previous findings of Gell-Mann [16] and Zweig [17] on the quark model, it became clear [18] that the gauge symmetry of the strong interaction was described by the $SU(3)$ group with a massless gauge boson, the gluon. The theory was named quantum chromodynamics or QCD because of the existence of the three colours. The resulting unified symmetry was established then for both electroweak and strong interaction to be $SU(3) \times SU(2) \times U(1)$. The SM showed

its full predicting power in a series of groundbreaking experimental milestones. In 1973 the neutral current, caused by the Z boson exchange, was discovered. The W^\pm and Z bosons itself followed in 1983 at CERN's Super Proton Synchrotron. Finally, the Higgs boson was discovered in 2012 at LHC.

In this Chapter, the SM will be introduced according to its historical development: first exploring the particle content and then the electroweak force, the Higgs mechanism and finally the quantum chromodynamics.

2.1. Fundamental Particles

The particle content of the Standard Model is summarized in this Section. Particles are divided into fermions, obeying the Fermi-Dirac statistic, if they have spin 1/2, or bosons, if they have an integer spin and behave according to the Bose-Einstein statistic.

Fermions Fermions are the building component of the matter, from protons to astrophysical objects. They are classified according to their quantum numbers associated to $SU(3) \times SU(2) \times U(1)$: colour for $SU(3)$ (more formally $SU(3)_C$), weak isospin I for $SU(2)$ ($SU(2)_L$) and weak hypercharge Y for $U(1)$ ($U(1)_Y$). The relation with the electric charge Q is given by $Q = I_3 + Y/2$, where I_3 is the third component of the weak isospin. Quarks are triplets of $SU(3)$, with red, blue and green colours, and leptons are singlets. $SU(2)$ gives rise to left handed doublets, with weak isospin $I = 1/2$, and right handed singlets with $I = 0$. The associated gauge bosons couple to left handed doublets only. Moreover, the fermions are arranged into three identical *families*, with the same quantum numbers and different masses. Higher families, or generations, are associated to higher masses and smaller lifetimes. There are six quarks and six leptons. The three families of leptons contain the electron e , the muon μ and the tau τ . The left handed ones form doublets with their associated neutrinos: ν_e, ν_μ and ν_τ that are electrically neutral (the doublet is more formally indicated as $(\nu_e)_L$ and the singlets as e_R for the first generation of leptons). The quarks, with electric charge of 2/3 are also called *up type* and are they up (u), charm (c) and top (t). The ones with electric charge of -1/3 are called *down type*, and are down (d), strange (s) and bottom (b). They form doublets and singlets under $SU(2)$ (e.g. the $(\begin{smallmatrix} u \\ d \end{smallmatrix})_L$ doublet and the u_R singlet). A schematic table of the fermions is shown on the left of Figure 2.1. The exact value of the masses and upper limits for neutrinos can be found in the PDG [19].

Bosons Five bosons are present in the SM. Four of them are gauge bosons, have a spin 1 and act as carrier of the strong, weak and electromagnetic force. The photon is the mediator of the electromagnetic force and interacts with all the particles having non-zero electric charge. It has no mass, making the electromagnetic interaction long range, and no electric charge. The two W^\pm and Z bosons mediate the weak interactions. They are rather massive, from which the short range of the interaction. The W^\pm have an electric charge of ± 1 and they interact also with one another. The gluon (g) mediates the strong force, is massless and have no electric charge. There are eight gluons following the colour octet, and they interact not only the coloured fermions, but also with themselves. Finally the Higgs boson (H) is a scalar, has spin 0 and no electric charge. It results from the spontaneous electroweak symmetry breaking, as discussed in this Chapter. On the right of Figure 2.1, an overview of the bosons present in the SM is shown.

2.2. Mathematical Formulation

The SM is formulated in the Lagrangian formalism. In this Section, a brief overview of the mathematical formulation of the SM is presented, but for a more complete overview the reader can refer to standard introductory texts [21] [22] [23] or specialized text and lecture notes. A more complete introduction can also be found in the Appendix E.

Within the framework of quantum field theory (QFT), particles are associated to the excitation of fields, in general expressed as a function of time and space, i.e. $\phi(\vec{x}, t)$. The Lagrangian *density*, \mathcal{L} can be expressed as a function of the fields and their derivatives $\partial\phi = \frac{\partial\phi}{\partial x^\mu}$, where with μ the four momentum notation is implied.

2.2.1. Electroweak Theory

The electroweak theory is a quantum field theory that describes both quantum electrodynamics and weak force. It follows the Yang-Mills theories to achieve the description of left-handed interactions that are typical of the latter, relying on the $U(1) \times SU(2)$ gauge group. The spin 1/2 fermion fields are introduced in their left- and right-handed components, $\psi_{L,R} = P_{L,R}\psi = \frac{1\pm\gamma_5}{2}\psi$. The gauge invariance of this theory is ensured by the presence of additional fields. B_μ is the generator of the $U(1)$ group and couples to the weak hypercharge Y . The fields W_μ^a , $a = 1..3$ are instead the generators of the

$SU(2)$ group and couple to the weak isospin, therefore to left-handed components of the fermion fields only. Conveniently they are indicated also with $W_\mu = W_\mu^a \sigma^a / 2$.

The gauge fields restore the invariance of the Lagrangian via the *covariant derivative* \not{D} , which is also the point of contact of the B_μ and W_μ fields with the fermion fields:

$$\mathcal{L} = i\bar{\psi}_R \not{D} \psi_R + i\bar{\psi}_L \not{D} \psi_L - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{2} \text{Tr}(W_{\mu\nu} W^{\mu\nu})$$

with the covariant derivative defined as:

$$\not{D}\psi = \gamma^\mu D_\mu \psi = \gamma^\mu (\partial_\mu + ig_1 B_\mu + ig_2 W_\mu^a \sigma^a / 2) \psi$$

where g_1 and g_2 are the coupling constants that determine the strength of the interaction, for $U(1)$ and $SU(2)$, respectively. The covariant derivative can be expanded for the left and right handed components:

$$\begin{aligned} D_\mu \psi_L &= (\partial_\mu + ig_1 B_\mu + ig_2 W_\mu^a \sigma^a / 2) \psi_L \\ D_\mu \psi_R &= (\partial_\mu + ig_1 B_\mu) \psi_R \end{aligned} \quad (2.1)$$

where the B_μ gauge field can be explicitly seen to interact the same way to ψ_L and ψ_R , while W_μ by construction only to ψ_L .

The physical fields for the W^\pm bosons can be obtained from W_μ^1 and W_μ^2 since:

$$W_\mu = \begin{pmatrix} W_\mu^3 & W_\mu^1 - iW_\mu^2 \\ W_\mu^1 + iW_\mu^2 & -W_\mu^3 \end{pmatrix} = \begin{pmatrix} W_\mu^3 & W_\mu^+ \sqrt{2} \\ W_\mu^- \sqrt{2} & -W_\mu^3 \end{pmatrix} \quad (2.2)$$

where $W_\mu^\pm = 1/\sqrt{2}(W_\mu^1 \mp iW_\mu^2)$. The fields W_μ^3 and B_μ will instead mix with one another, recovering the physical fields for the photon and for the Z boson:

$$\begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} = \begin{pmatrix} \cos(\theta_W) & \sin(\theta_W) \\ -\sin(\theta_W) & \cos(\theta_W) \end{pmatrix} \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} \quad (2.3)$$

Here θ_W is the Weinberg angle that sets the scale of the electroweak mixing, with $\sin(\theta_W) = g_1 / \sqrt{g_1^2 + g_2^2}$. The electromagnetic field A_μ was identified recognizing $g_2 \sin(\theta_W) = e$, and the neutral field Z_μ associated then to the Z boson.

This is now a skeleton of an EW theory that is for the moment massless: mass terms inserted by hand violate the local gauge symmetry of the Lagrangian. Moreover, terms that explicitly break the symmetry are not renormalizable. This contrast can be solved with the Higgs mechanism.

2.2.2. The Higgs Mechanism

The Higgs mechanism introduces a scalar field, doublet under $SU(2)$, with an ad-hoc choice of its potential. In this case, the symmetry of the Lagrangian will not be broken by hand, but by the intrinsic properties of the fields. The complex scalar field reads:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi_0 \end{pmatrix}, \Phi^* = \begin{pmatrix} \phi^- \\ \phi_0 \end{pmatrix}$$

the covariant derivative defined as before in the EW theory is used ($D_\mu \Phi = (\partial_\mu + ig_1 B_\mu + ig_2 W_\mu)\Phi$). The complex scalar field can be seen as a composition of four real scalar fields ϕ_1, ϕ_2, ϕ_3 and ϕ_4 , noting that e.g. $\phi^+ = (\phi_1 + i\phi_2)/\sqrt{2}$ and $\phi_0 = (\phi_3 + i\phi_4)/\sqrt{2}$ and similarly for the conjugates. The Lagrangian for this field (Higgs Lagrangian) can be written as usual for the scalar fields:

$$\mathcal{L}_H = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi) \quad (2.4)$$

where $V(\Phi)$ is the scalar potential, and can be written in its general form as

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 = \mu^2 |\Phi|^2 + \lambda |\Phi|^4 \quad (2.5)$$

It can be noticed that the potential is by construction invariant under $SU(2) \times U(1)$. Moreover, it is clear, expressing $\Phi = (\phi_1, \phi_2, \phi_3, \phi_4)^T$, that the general transformation represented as $R \in O(4)$ leaves this Lagrangian invariant, where $O(4)$ is the orthogonal group. The shape of this potential depends on the choice of the parameters μ and λ . $\lambda > 0$ ensures that the potential is bounded from below. Choosing also $\mu < 0$ (as also shown in Figure 2.2) however, the symmetry of the potential $V(\Phi)$ is broken, since the ground states is not at zero, but at:

$$|\Phi|^2 = \phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2 = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2}$$

where v is the *vacuum expectation value* (VEV) of the Higgs field. This way the system has *spontaneously* chosen one of the minimal configurations and it is not any more symmetric under the gauge symmetry. This process is called *Spontaneous Symmetry Breaking* (SSB).

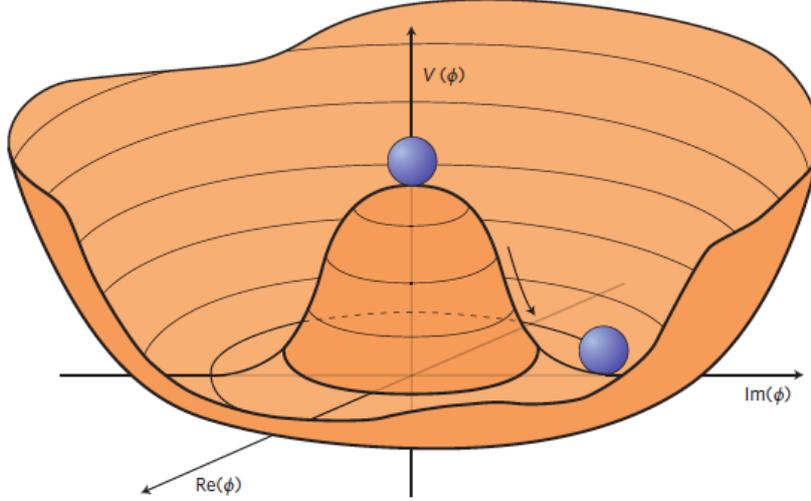


Figure 2.2.: Representation of the Higgs potential, for the parameter choice $\lambda > 0$ and $\mu < 0$. From [24].

Expanding around the minimum, Φ can be expressed as:

$$\Phi = \begin{pmatrix} \phi^+ \\ \frac{1}{\sqrt{2}}[v + H(x) + i\chi(x)] \end{pmatrix} = \frac{1}{\sqrt{2}} \exp \left[\frac{i\sigma_a \theta_a(x)}{v} \right] \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}$$

with H , χ and θ_a real fields (this choice can be done without loss of generality). Since Φ is a complex doublet under $SU(2)$, the phases can be rotated away. This is achieved with $U \in SU(2)$ such that $U = \exp \left[\frac{i\sigma_a \theta_a(x)}{v} \right]$ with $\Phi \rightarrow \Phi' = U\Phi$. The field becomes now:

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

This gauge choice is called *Unitary Gauge* and it is useful at tree level to show the physical fields, but the identical results in the physical observables can be retrieved with all the other gauge choices. The original global symmetry of the Lagrangian $O(4)$ is now broken into $O(3)$. This three missing degrees of freedom were associated with the massless fields θ_a (or ϕ^+ , ϕ^- and χ), which are not any more present in the Lagrangian after fixing the gauge. This is indeed not a chance, but recovers the result

of the Goldstone theorem. These degrees of freedom can be now ‘inherited’ as the transverse component W^\pm and Z boson, which will then acquire mass. They will arise from the kinematic term of the Higgs Lagrangian, $(D_\mu H)^\dagger (D^\mu H)$ (the covariant derivative has constant terms in the gauge fields), which will have terms of the form:

$$(D_\mu H)^\dagger (D^\mu H) \supset \frac{g_2^2 v^2}{4} W_\mu^+ W_\mu^- + \frac{g_2^2 v}{4} H W_\mu^+ W_\mu^- + \frac{1}{8} \frac{g_2^2 v^2}{\cos^2 \theta_W} Z_\mu Z^\mu + \frac{1}{4} \frac{g_2^2 v}{\cos^2 \theta_W} H Z_\mu Z^\mu$$

This shows how the W and Z fields now acquire a mass (the terms $W_\mu^+ W_\mu^-$) with $m_W^2 = g_2^2 v^2 / 4$, $m_Z^2 = g_2^2 v^2 / 4 \cos^2 \theta_W$. Also, the fields couple to the Higgs boson in terms like $H W_\mu^+ W_\mu^-$ and $H Z_\mu Z^\mu$. Expanding the full terms of the covariant derivative using the relations 2.2 and 2.3, it can be verified that the photon is again massless¹. But this means that there is a local symmetry that is not broken, the one associated with the photon field: $U(1)_{em}$.

The fermion masses arise from the Higgs mechanism, when a Yukawa term is added, coupling the fermionic and the Higgs fields. It has the form:

$$\mathcal{L}_{Yukawa} \supset -g_\psi (\bar{\psi}_L \Phi \psi_R) + h.c.$$

with g_ψ is the coupling between the Higgs field and ψ . Repeating the same procedure as for the EW case, the mass terms for the fermions are found to be $m_\psi = g_\psi v \sqrt{2}$. Here the mass matrix of the ψ_R and ψ_L fields is not diagonal after EW symmetry breaking. Diagonalizing to mass eigenstates, the mixing of the quarks in weak interactions is recovered, described by the CKM (Cabibbo–Kobayashi–Maskawa) matrix.

2.2.3. Quantum Chromodynamics

Quantum chromodynamics (QCD) is the final missing piece in this construction of the SM, which describes the strong force and its interaction with quarks and gluons. Similarly to the EW theory, it is also based on a non-abelian structure, but using instead the group $SU(3)$. This group has eight generators, which are associated with the gluons. The modified covariant derivative reads:

$$\not{D} Q_L^i = (\not{\partial} + ig_3 \gamma^\mu G_\mu + ig_2 \not{W} + ig_1 \not{B}) Q_L^i$$

¹Alternatively, it can be seen that the mass matrix has a vanishing determinant, therefore implying the existence of an eigenstate with zero mass.

where the fields associated to $SU(2)$ and $U(1)$ can be recognized, but now a new set of eight spin-1 fields $G_\mu = G_\mu^a \lambda^a / 2$ and its coupling constant to the quark Q was introduced. The kinetic term associated to these fields is, noting the similarities again with the EW theory:

$$\mathcal{L} = -\frac{1}{2} \text{Tr} G_{\mu\nu} G^{\mu\nu}$$

The gluon tensors are defined as $G_{\mu\nu} = \partial_\mu G_\nu - \partial_\nu G_\mu + ig_3 [G_\mu, G_\nu]$, where $[G_\mu, G_\nu]$ can be written in terms of the structure constants of $SU(3)$, f^{abc} , that satisfy the relation $[\lambda^a / 2, \lambda^b / 2] = if^{abc} \lambda^c$.

One important difference with respect to the EW theory, is the behaviour of the coupling g_3 (or more commonly expressed as $\alpha_s = g_3^2 / 4\pi$). In QED, taking into account higher loop diagrams of the processes, the strength of the interaction depends on the momentum transfer q , increasing for decreasing distances or q . This can be interpreted as vacuum polarization, where the vacuum is filled with particle antiparticle pairs, behaving effectively like a dielectric medium and screening the charge. The QCD vacuum, however, is intrinsically different, since it will be filled with self interacting gluons (arising from the non-abelian structure of $SU(3)$ as can be seen in the kinematic term of the Lagrangian). The behaviour for QCD is then the opposite, an anti-screening, where the coupling is large at larger distances (small q) and small at small distances (large q). This leads to peculiar characteristics of the theory: *confinement* and *asymptotic freedom*. The former is caused by the large coupling at small q , which confines the quarks within the nucleon. As results, no free coloured particle is observed. The latter effect takes place at higher momentum transfers: asymptotically, coloured particles behave as free particles. This is particular useful: at high q perturbation theory can be used. The quantitative behaviour of α_s is governed by the so-called renormalization group equation.

2.3. Shortcomings of the Standard Model

The SM is one of the most successful theories ever conceived in physics; however, there are few shortcomings that are quickly described in this Section. The most important of these are: Dark Matter, Dark Energy and neutrinos oscillation.

Dark Energy and Dark Matter The SM does not have contents and mechanisms able to describe the energy and matter distribution in the universe. From astrophysical observations, it is known that the universe is currently undergoing an expansion at accelerating rates. In order to justify this observation, the concept of Dark Energy was established already couple of decades ago, under the concordance model of cosmology. This unknown form of energy is estimated to comprise to around 68% of the total energy of the universe at the present day. Dark Matter is an older concept with respect to Dark Energy, and started off as a problem in modern astrophysics from the unexplained behaviour of matter in the halos of galaxies. This behaviour could be explained admitting large amounts of matter in the galaxies that was not visible, hence the name. Dark Matter is covered in more details in the next Chapter.

Neutrino Oscillation In the formulation of the SM, the neutrinos are present as massless left-handed spinors. Mass terms, however, would require the presence of right-handed spinors. Massive neutrinos are not easy to incorporate in the SM. If they are Dirac fermions, right-handed neutrinos are needed to acquire mass from the Higgs mechanism. Right-handed neutrinos would hardly be detectable, interacting only with the Higgs field. If they are Majorana fermions, a violation of flavour and lepton number conservation would be introduced in the theory. Neutrinos have been confirmed to oscillate between flavours in the last decades by many experiments using solar, atmospheric, reactor and accelerator neutrinos. This is only possible admitting that these particles indeed carry a mass, albeit small. Mass eigenstates of the neutrinos are those propagating in the vacuum, while their flavour eigenstates are free to oscillate. The oscillation is described by a rotation matrix, which connects the mass to flavour eigenstates. It is similar to the CKM matrix, but it is called Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix. Moreover, while the oscillations of the neutrinos demonstrate that these are in fact massive particles, it is not possible to infer the value of their masses from the oscillation alone, only their differences. In particular, two of them are close in mass ($\Delta m \sim 10^{-5}$ eV) and the other is more distant ($\Delta m \sim 10^{-3}$ eV). The upper limit for the neutrino masses can be measured from e.g. tritium decay, and are around the eV. It is striking how different these scales are with respect to all the other SM particles. The origin of this difference is an open question.

Matter-Antimatter Asymmetry CP violations, which translates to matter-antimatter asymmetries, is present in the formulation of the SM. However, the amount of matter

that is left from the Big Bang, where both should have been created in identical amounts, cannot be justified with the known effects alone. This is the *matter-antimatter* problem: the CP violating processes of the SM are not enough to explain the large amount of matter present nowadays. The QCD sector of the SM does not contain any CP violating phase, although there is no reason a priori to exclude it. This is known as the *strong CP problem*, questioning why the CP violating phase is negligible.

Hierarchy Problem and More The Hierarchy Problem is strictly connected to the concept of *naturalness*, which is the property of a given theory of having parameters that are not too large or too small without an appropriate explanation. The *fine-tuning* of parameters is the condition where these parameters have to be adjusted to very precise values for the theory to fit the observations. The scales of the gravitational and electroweak force can be for example expressed in terms of the Newton's constant G_N (conversely in terms of the Planck mass $M_{pl} = G_N^{-1/2} \sim 10^{19}$ GeV) and the VEV (~ 246 GeV). This hierarchy between the gravitational and electroweak scales, with tens of orders of magnitude difference, shows a naturalness problem. How such a big difference in scales between fundamental parameters arises is a crucial question. Another indication of missing naturalness in the SM is the fine-tuning needed to recover the observed Higgs mass. The radiative corrections to the Higgs mass can in general be written as $m_H^2 = m_{H,bare}^2 + \Delta m_H^2$, where $m_{H,bare}$ is the tree-level bare Higgs mass and Δm_H^2 the radiative correction. Δm_H^2 can be further expressed from the loop diagrams and is proportional to the masses m_i of the particles that run in the loop and an ultraviolet (UV) cut-off Λ : $\Delta m^2 \sim \Lambda^2 m_i^2$, (with a -1 factor if the particle is a fermion). One obvious choice for the UV cut-off is the Planck mass scale, where the gravitational effects become relevant and the SM will breakdown. With this choice, however, it becomes evident that the factors leading to corrections to the Higgs boson mass have to be carefully chosen to cancel each other to counter balance the large Λ , leading to a severe fine-tuning.

There are of course more points that need to be addressed in the Standard Model, both experimentally and theoretical. There are few measurements in tensions with the SM predictions, such as the muon anomalous magnetic dipole moment, the proton radius puzzle and the neutron lifetime. Theoretically, the main challenge the SM faces is the inclusion of gravity.

3. The puzzle of Dark Matter

The idea of the presence of 'Dark' objects in the universe can be traced back to the nineteenth century (for a more complete historical introduction see [25]). Friedrich Bessel inferred the existence of a not yet discovered astrophysical object based on its gravitational influence alone in the system of Sirius and Procyon. The assumption of the presence of a much less luminous companion star could help to explain their relative motion. A few years later, a new planet was predicted based on the anomalies of the motion of Uranus, and another one was expected to exist based on the strange precession of the perihelion of Mercury, but was never found. Even if the first planet was indeed discovered, the second phenomenon had to wait still many years for the advent of General Relativity. Dark clouds were also hypothesized in regions of the sky with absence of stars that looked like structures to the astronomers. Lord Kelvin concluded that the majority of the objects in the galaxy were dark bodies, perhaps extinct stars, using the relation between the size and the dispersion of velocities. However, it is in the last century that the Dark Matter puzzle first took its shape. The astronomer Zwicky noticed among others in 1933 a large scatter of the apparent velocities in eight galaxies of the Coma cluster, which could not be justified by the mass of the visible objects using the virial theorem [26]. Therefore Dark Matter, he concluded, must be present in much larger quantity with respect to ordinary matter. Despite the 1939 measurement performed by Babcock [27] of the rotation curves from the Andromeda Galaxy, the astrophysics community was in a state of confusion. Objections were made about the state of equilibrium of the systems investigated with the virial theorem, which is a prerequisite for its application. It was realized, however, that this interpretation was in tension with the estimated ages of the galaxies and of the universe. More studies were needed to understand the galactic dynamics. In the 60's systematic efforts took place to understand the nature of Dark Matter, which was a concept not yet commonly accepted but also not rejected. Gaseous hydrogen, free and ionized, was estimated to contribute to the gravitational bound at a couple of order of magnitude less than the observed values. In what is now known as the 1970's revolution (around the same time when the Standard Model of particle physics took its modern form), the enhanced sensitivity of the measurements allowed the discovery of more and more galaxies with large amount of mass needed to justify the distribution of velocities. At this time, more exotic hypotheses began to be considered, such as

dwarf stars, and collapsed objects. One by one, they were subsequently ruled out, pointing more and more to a non-baryonic nature of the Dark Matter. The existence of Dark Matter (DM) today is corroborated through observations at different scales: galactic scales are investigated with rotation curves, cluster scales with gravitational lensing and cosmological scales via the Cosmic Microwave Background.

Galactic Scales The presence of the Dark Matter in the universe was first investigated at galactic scales. The velocity of an orbiting object placed at a distance r from the centre of a system scales like $v(r) \propto \sqrt{M(r)/r}$, where $M(r)$ is the mass within the radius r . For stars on the outskirts of the galaxies, $M(r)$ is essentially constant since the vast majority of the stars are found in or close to the galactic centre. Their velocities should then scale like $v(r) \propto \sqrt{1/r}$, contradicting the measurements where approximately a constant behaviour of the velocities can be observed at increasing distances. This can be seen e.g. in Figure 3.1, where the distribution of velocities in NGC 3198¹ as a function of the distance from the galactic centre can be explained with the presence of a Dark Matter halo. Galaxies without significant amounts of Dark Matter are less than a handful [28], making a striking evidence for the existence of Dark Matter at galactic scales in the observable universe.

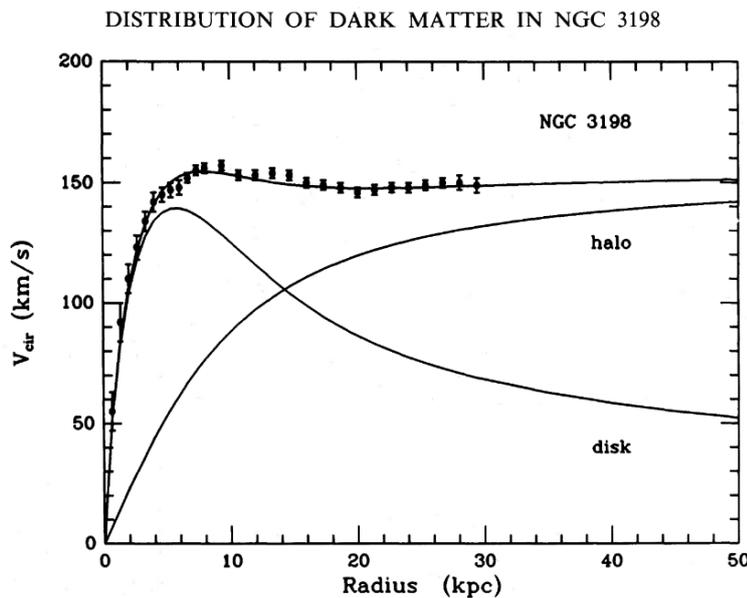


Figure 3.1.: The galactic rotation curve of NGC 3198, from the 1985 measurement [29]. In separate components, the contribution to the observed values from the galactic disk and the Dark Matter halo are shown.

¹Herschel 146 is a barred spiral galaxy in the constellation of Ursa Major.

Cluster Scales Dark Matter at galaxy clusters scales can be probed with gravitational lensing. This effect happens when light rays are deflected by massive objects since, as predicted by General Relativity, massive objects curve the space itself. It can be caused by e.g. stars such as the sun, with which the General Relativity was firstly confirmed in 1919 by Arthur Eddington and Frank Watson Dyson, or galaxies and galaxy clusters. The angle of deflection θ can be shown to be $\theta = 4GM/rc^2$, with G the gravitational constant, M the mass of the object and r the distance from the source. According to the magnitude of the deflection, the gravitational lensing can be divided into strong, weak and micro-lensing. In the strong lensing the deflection is so hard that the image is highly distorted and can be observed in rings (the Einstein Rings) around the massive object. In the weak lensing the image does not form rings, but it is distorted in the perpendicular direction with respect to the massive object. Finally, the micro-lensing usually can be detected only as a decreased magnitude of the luminosity of the source. The mass measurements in lensing, particularly in weak lensing, is a strong indication of the presence of Dark Matter in galaxy clusters. The most significant example is the so-called *bullet cluster*, which is a system of two colliding galaxy clusters placed at a distance of 3.72 billion light years, shown in Figure 3.2. Here the baryonic and Dark Matter component separated during the impact between the two clusters: the baryonic component, shown in red, was slowed down mainly by drag force. The Dark Matter component, which represents the majority of the mass of the clusters, did not experience the drag force, since it only interacts gravitationally. The baryonic component was therefore left behind. The bullet cluster makes difficult for Modified Newtonian Dynamics (MOND) theories, which have found wide application in explaining the galactic rotational curves, to also justify the behaviour at cluster scales. Moreover, it also strongly constrains the self-interaction of the Dark Matter, since self-interaction would have slowed down the DM component by drag force as well.

Cosmological Scales The Cosmic Microwave Background (CMB) is the main tool that allows to explore the cosmological scales. In the history of the universe, after the nuclei started to condense from the quark gluon plasma, space itself was still opaque to electromagnetic radiation due to the scattering of photons off charged particles. When a temperature of ~ 0.1 eV was reached, the kinetic energies of the particles started to be low enough to allow the *recombination* into neutral atoms. The photons were left free to depart from the surface of last scattering. They moved throughout the entire

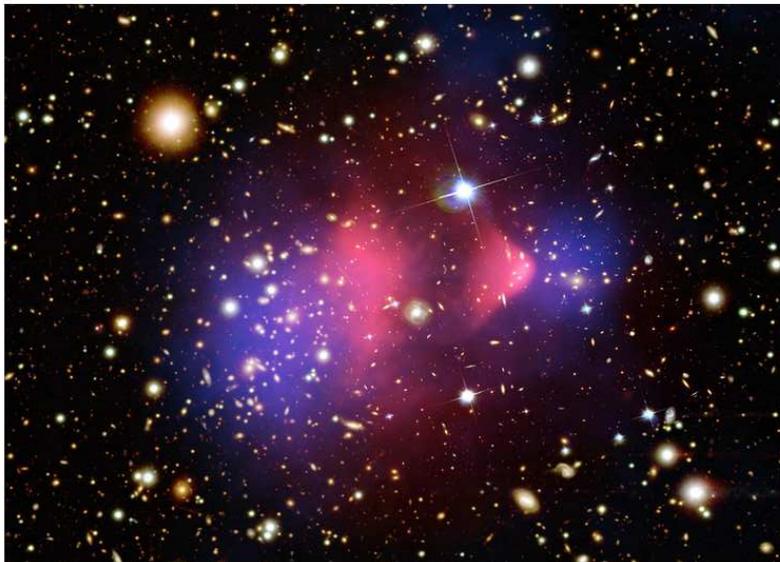


Figure 3.2.: The *bullet cluster* as shown in different superimposed components: the pink clumps in the image are the hot gas distribution in the as measured by the Chandra-X in the X-rays spectrum. The one on the right shows the bullet shape, which passed through the left one. The optical image is taken from the Magellan and Hubble Space Telescope. The blue areas represent the mass of the two systems as measured from micro-lensing. The hot gas was slowed down during the impact by drag force, surpassed by the Dark Matter component [30].

universe as it expanded, with an energy as measured now of $\sim 10^{-4} \text{ eV}^2$, reaching the microwave domain and showing an almost perfect black body distribution peaked at a temperature of 2.725 K. The relic photon background was predicted by Alpher, Herman and Gamow [31] [32] [33] and actually discovered by Penzias and Wilson. They were testing a Dicke type radiometer that they intended to use for radioastronomy and communication with artificial satellites [34], subsequently winning the Nobel prize for their discovery. The first aspect to note is that the CMB mean temperature is so remarkably uniform, that it must have been casually connected at some point. This implies that the universe underwent a fast expansion, the *inflation*. The CMB was subjected to more dedicated analyses, the most recent and important are made by Wilkinson Microwave Anisotropy Probe (WMAP) satellite and Planck Surveyor, which provided a map of the CMB for the entire sky. In Figure 3.3 the Planck temperature power spectrum as measured from the anisotropies (that are at the $10^{-4} - 10^{-5}$ K level) of the CMB. These anisotropies give meaningful information if analyzed in terms of the angular scale, or spherical harmonics. They are given by:

$$\frac{T(\theta, \phi) - T_0}{T_0} = \sum_{l=0}^{\infty} \sum_{m=-l}^l a_{lm} Y_{lm}(\theta\phi)$$

where $Y_{lm}(\theta\phi)$ are the spherical harmonics and a_{lm} the multipole moments. The power spectrum C_l can be defined as:

$$C_l = \langle |a_{lm}|^2 \rangle = \frac{1}{2l+1} \sum_{m=-l}^{+l} |a_{lm}|^2$$

It is also more conveniently expressed by $D_l^{TT} = l(l+1)C_l/2\pi$, where TT indicates temperature-temperature measurement.

As a rule of thumb, the relation between the multiple moment and the observed angle is $\sim \pi(\text{rad})/l$. The monopole is simply the CMB mean temperature. The dipole asymmetry does not carry cosmological value, since it represents the relative motion of the solar system and of the galaxy, and it is therefore subtracted. The peaks showing in the power spectrum in Figure 3.3 are connected to cosmological parameters, such as the baryonic, the Dark Matter content and the geometry of the universe. The peaks can in fact be interpreted according to the behaviour of the photons and baryons (sometimes called *photon-baryon fluid*) before the recombination. As Dark Matter

²Their wavelengths were shifted by $1+z$, with z the redshift due to the expansion.

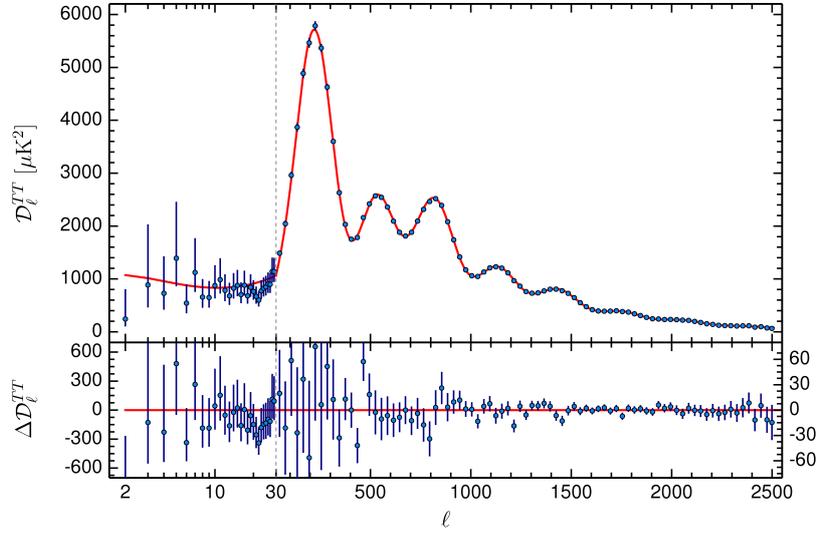


Figure 3.3.: The temperature power spectrum of spherical harmonics as measured by *Planck* 2015 [35].

and baryonic matter started forming gravitational wells, more and more matter got attracted inside. However, at some point, the photons started exerting a pressure on the matter (but not Dark Matter) causing it to expand again. Soon the baryons would fall another time into the gravitational dips, causing the pressure to increase again and so on. This oscillations are referred to as *baryon acoustic oscillations* where the term 'acoustic' is made in analogy to the sound oscillations in fluids. The behaviour of this phenomenon consents to access few important cosmological values. First of all, at large angular scales ($l < 200$), the so-called Sachs-Wolf effect dominates the spectrum. The anisotropies at these scales are in fact mostly influenced by the photons that were redshifted or blueshifted as they gained or lost energy while moving through gravitational potentials. This effect is also present after the photons left the surface of last scattering, moving through structures. Because of the expansion of the universe, the gravitational potential that they face becomes less and less intense as time passes. This is known as the Integrated Sachs-Wolf effect. The first peak has a specific position in the power spectrum, which is connected to the angular scale of the gravitational wells. This scale can be different according to the curvature of the universe. In case of spherical universe, the angles would appear to be larger with respect to a flat universe and smaller in case of a hyperbolic one. The curvature is described with the *density parameter* Ω . The position of the first acoustic peak is at $l \sim 200$, fully compatible with a flat universe, i.e. with Euclidean geometry. The amplitude of the first baryon-

acoustic peak is related to the matter content of the universe. It represents the modes corresponding to the maximum pressure in the gravitational peak, which were frozen at the time of recombination. The more (ordinary) matter, the deeper the potential and therefore the higher the pressure and the peak amplitude. The second peak shows the mode corresponding to the maximum pressure *outside* the wells, which is roughly independent from the depth of the potential. The third peak again corresponds to maximum pressure inside the potential. This time, the amplitude is also sensible to the Dark Matter content. With more DM, the potential gets deeper, but, as pressure increases, it is exerted only to the baryon-photon gas, leaving the DM untouched. As DM is still present and not expelled from the well, gravity starts dominating again, increasing again the quantity of ordinary matter inside the gravitational potential. This in turn increases the pressure again and the amplitude of the third peak that is frozen at time of recombination. The ratio of the first to the third peak in the power spectrum allows to access the Dark Matter content of the universe. Finally, the modes with a wavelength smaller than the thickness of the surface of last scattering, are suppressed (*diffusion damping*) and this happens for $l > 1000$.

From the analysis of the CMB, the Planck collaboration estimated that ordinary matter can account only to 4.9% of the total energy content of the universe, 26.8% the Dark Matter and 68.3% by Dark Energy, which is identified with the cosmological constant in the context of the Λ CDM (Lambda Cold Dark Matter) model.

Dark Matter properties

As just seen, the evidence for the existence of Dark Matter stretches throughout several orders of magnitude in scales: galaxies, clusters of galaxies and cosmological. The amounts of Dark Matter in the universe correspond to a staggering five times of the ordinary matter, strongly favouring the structure formation in the early universe. Being dark, i.e. not reflecting light or interacting electromagnetically it must be composed of neutral particles, or their charge must be small in order to explain the observations. It could interact weakly, but colour neutral. Also its self-interaction is rather limited. Finally, it interacts gravitationally. DM candidates can be classified as *hot*, *warm* and *cold* Dark Matter, according to their kinematic status. Hot DM particles must be relativistic, warm DM was relativistic earlier in the history of the universe, but it cooled down as the universe expanded. Cold Dark Matter, which is the scenario favoured by the observations, is comprised of non-relativistic particles. The mass

of the DM candidate has a strong impact on whether it behaves as hot, warm or cold. Heavier particles would have less kinetic energy than lighter ones, making them non-relativistic. Neutrinos, which couple to matter only weakly, are a good and investigated candidate for DM. However, they are a hot DM candidate, making difficult to explain structure formation. Moreover, the upper bounds on their masses are so low that they would not contribute significantly to the total DM density. New hypothetical particles have therefore been proposed as Dark Matter. *Sterile* neutrinos could be right handed neutrinos, and have significant larger mass than the left handed ones. *Axions* and *Axion-like particles* that were firstly introduced to solve the strong CP problem, and have a tiny coupling with photons. Weakly Interacting Massive Particles (WIMPs) offer interesting properties that make them a good DM candidate.

The quantitative considerations that are derived from the interaction rate of the DM particles can be useful to understand the observed densities in the universe. DM is assumed to be in thermal equilibrium in the early universe, when the process $DM \Leftrightarrow SM \ SM$ could proceed in both directions. At some point, the expansion of the universe was too fast to allow the process, effectively freezing the SM and DM densities. This phenomenon is called *Freeze-out*, and the remaining densities are called *relic densities*. The Boltzmann equations describe the Freeze-out, which is expressed in terms of the number density, n , is given by:

$$\frac{dn}{dt} + 3H_0 n = -\langle\sigma c\rangle(n^2 - n_{eq}^2)$$

where $\langle\sigma c\rangle$ is the averaged DM annihilation cross-section, n_{eq} is the number density at equilibrium (that can be derived as a function of the particle mass) and H_0 is the Hubble constant. Solving this equation it can be seen that, as the universe expands, the densities become so low that the DM annihilation eventually stops. The evolution of the number density can be seen in Figure 3.4. Interestingly, with the mass and the cross-sections that are typical of the electroweak scale, the observed relic abundances in the universe are recovered. This phenomenon is called *WIMP miracle* and makes the WIMPs with masses in the scale GeV-TeV and interactions at the weak scale a good DM candidate.

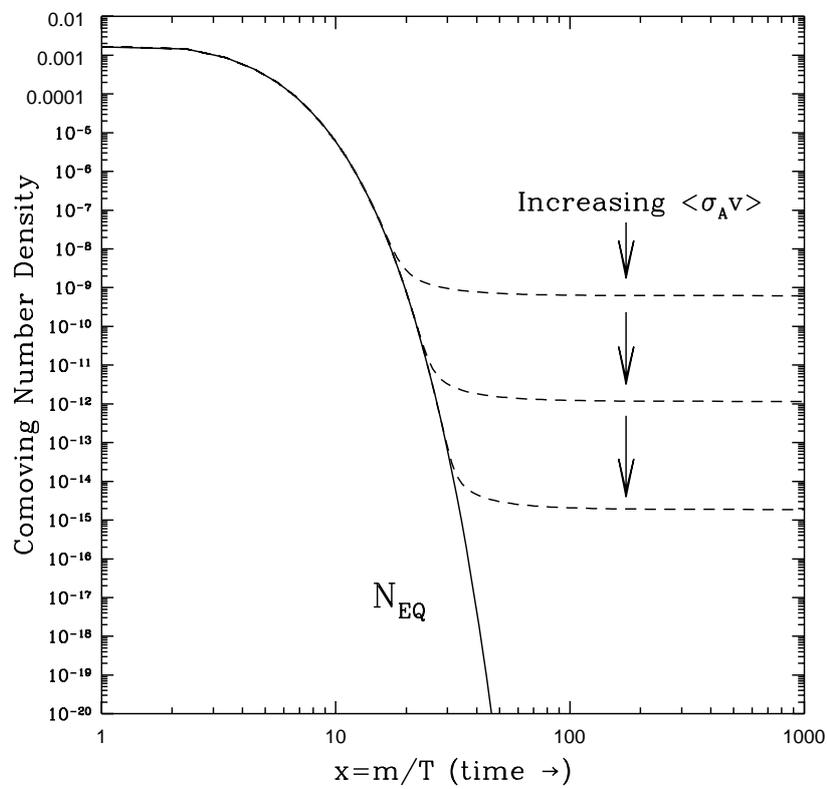


Figure 3.4.: Schematic of the evolution of the number density that undergoes the process of the Freeze-out [36]. As the temperature decreases (or m/T increases) due to expansion of the universe, the density is frozen at a certain value according to the interaction rate. Higher interaction rates make DM annihilation more likely, decreasing its density.

3.1. Dark Matter Searches

Searches for Dark Matter can be grouped in three big areas: *direct*, *indirect* and *production*, depending on the type of interaction that is used for the detection.

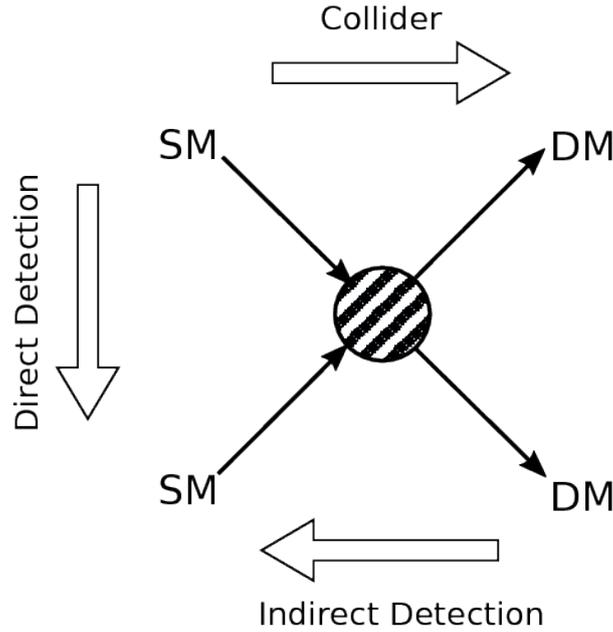


Figure 3.5.: The three main experimental lines to address the experimental search for Dark Matter [37]. From top to bottom, the DM state scatters off SM state that is then detected: this is the direct detection. From right to left, two DM states annihilate into SM states, which are then detected, in what is called the indirect detection. Finally, from left to right, the production at colliders, where two SM states annihilate into DM. What is looked for in this case is missing momentum in the event, as DM states fly away undetected right after production.

Direct Detection The strategy of the direct detection lies in the ability to resolve an impact of a DM particle on a nucleus in a well controlled environment. The scatter off the atomic nucleus will transmit a momentum in the shape of nuclear recoil, which can be expressed as:

$$E_R = \frac{q^2}{2m_N} \sim 50\text{keV} \left(\frac{m_\chi}{100\text{ GeV}} \right)^2 \left(\frac{100\text{ GeV}}{m_N} \right)$$

where q is the momentum transfer, m_N the mass of the nucleus and m_χ the mass of the DM particle. If the DM is present in our solar system, it cannot have a velocity higher than the escape velocity of the Milky Way, $\sim 500\text{-}600\text{ km s}^{-1}$ [38], and the average

velocity would be around 200 km s^{-1} . The designed target nucleus has to be chosen to increase the differential rate and the recoil. For WIMP-like particles, the recoil is of the order of tens of keV when using heavy enough nuclei such as Germanium or Xenon. The different technologies used for the direct detection experiments have different advantages and disadvantages. There are three main effects that can be exploited in direct detection experiments: ionization, scintillation light and phonon propagation in crystals. Detectors at the present time can combine two of these effects.

Scintillator crystals (NaI(Tl) or CsI(Tl)) generate detectable light from the excitation of the atoms, such as in the DAMA/LIBRA apparatus [39]. Germanium detectors operated in cryogenic environment measuring the ionization allow sensitivities to low E_R , meaning almost to sub-GeV domain in WIMP mass, and can measure both charge and heat. As an example the CoGeNT [40], the CDEX-0 [41], the CDMS [42] and CRESST-II [43] detectors. Bolometers use the phonons, both thermal and non-thermal, which are produced in crystals. Thermal phonons are measured using temperature differences. Non thermal phonons are detected using charge signal e.g. in cryogenic Germanium crystals (such as in the EDELWEISS-II [44] detector). Finally, in liquid noble gas detectors, liquid Argon or liquid Xenon is used as scintillation and ionization medium. Photomultipliers can be used to measure the first effect and time projection chambers to extract the second. A notable example is XENON1T [45] with its 3.2 tons of ultra radio-pure liquid Xenon used as active target.

Indirect Detection Reading Figure 3.5 from left to right, it is possible that two DM particles annihilate into SM states, such as photons, anti-particles or neutrinos. These particles could travel through the universe to be finally detected by dedicated experiments. The annihilation is expected to be more likely in regions with higher DM densities, such as galaxy cores or massive objects. This processes would be observed as an enhancement in the number of detected gamma rays, positrons or anti-protons and neutrinos, with respect to the astrophysical expectations. One of the main difficulties is in fact to disentangle the DM candidate events from those of astrophysical origin. Gamma rays are detected both with orbiting detectors, such as Fermi Large Area Telescope [46], or arrays of Cherenkov telescopes such as H.E.S.S. [47]. Cosmic rays can again be detected e.g. with orbiting instruments such as AMS on the international space station [48]. Neutrinos offer the advantage of freely travel through matter after they are produced. As an example of a neutrinos detector e.g. IceCube [49] can be quoted, which uses Antarctic ice as Cherenkov medium.

Production at Colliders At colliders such as the LHC, which is introduced in the next Chapter, the annihilation of SM states could produce DM particles. The DM particles, unlike other SM objects apart from neutrinos, would remain undetected, leaving a characteristic signature of momentum imbalance in the transverse plane with respect to the beam axis. Since it would be impossible to detect a process like $pp \rightarrow \chi\bar{\chi}$ (where χ is the DM particle), as it would not be triggered, searches for Dark Matter at the LHC require the presence of another, detectable SM object. It can be radiated off the initial states, or produced as the consequence of a hypothetical mediator. The process that is addressed in the context of DM searches is therefore $pp \rightarrow \chi\bar{\chi} + X$. X is an additional SM state, such as a quark, and electroweak boson etc. The challenge of this kind of searches is posed by the presence of backgrounds that cannot be reduced, originating from Standard Model processes that are comprised of neutrinos plus the same SM X state that is looked for. The presence of neutrinos makes impossible to distinguish a DM signal from the background, as they both leave momentum imbalance in the transverse plane.

In the DM searches at the LHC, the DM states are often integrated into the frameworks of beyond the Standard Model theories, which can be effective or simplified models, or more complex ones (such as Supersymmetry). A simplified model called the Dark Higgs will be introduced here, which suggests a mechanism for the production of DM that has not been fully investigated yet, through the interaction with an hypothetical Dark Higgs-like particle.

3.2. The Dark Higgs Model

The Dark Higgs model [50] explores the idea of the existence of an Higgs-like particle that is responsible for the generation of the masses in the Dark Sector, such as DM states. This mechanism is similar to the Higgs mechanism taking place in the Standard Model 2.2.2. The Dark Sector can interact with the SM through the Dark Higgs (s) and an additional spin 1 particle, like a new Z' massive gauge boson. The DM searches at the LHC strongly constrain the existence of these additional mediators³, however

³ The existence of the Z' boson is already constrained from different di-jet searches (see Figure A.1 in the Appendix), or searches for the associated production of an hadronically decaying vector boson plus the Dark Higgs boson [51]. The constraints can be stringent on the coupling of the Z' to quarks. However, di-jet limits become less and less stringent if the coupling g_χ has higher values, where the limits are relaxed.

this tension can be released if the DM particles are not the lightest states. The Dark Higgs can be lighter than the DM states; in this case the relic abundance is set by the process $\chi\chi \rightarrow ss$, following the decay of s into SM states. The relic density will depend only on the coupling between the Dark Higgs and DM states, g_χ , while the couplings to the SM states can be small. Conventional direct DM searches would be therefore insensitive, and the Dark Sector would remain secluded [52] even for large g_χ . Also indirect searches would not allow to probe this scenario, since the DM annihilation into SM states would be velocity suppressed, and astrophysical constraints would only apply in case of large mass difference between DM and Dark Higgs.

3.2.1. Theoretical Framework

In this simplified model, the Dark Matter particle χ is considered to be a Majorana fermion⁴. It obtains its mass from the VEV w of a complex Dark Higgs field S and behaves as a singlet under SM gauge group. The Z' boson is associated to a new hypothetical gauge group $U(1)'$, which acquires its mass as well from the spontaneous symmetry breaking of the S field, and generates the Dark Higgs s . If the DM particles have an axial interaction with the new gauge boson, the interaction Lagrangian can be written as:

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

The first term describes the interaction of the DM particles with the Z' , the second term the interaction between DM particles and the Dark Higgs. Finally, the last terms show the Z' to Dark Higgs interaction. Moreover, a vector interaction between the SM quarks (q) and the Z' boson is added in an additional term:

$$\mathcal{L}_\chi = -g_q Z'^\mu \bar{q}\gamma_\mu q$$

There are in total four independent parameters in the Dark Higgs model for the Dark Sector: the masses of the Dark Higgs, Dark Matter and Z' gauge boson, m_s , m_χ and $m_{Z'}$, respectively and the DM coupling g_χ (can be expressed as $g_\chi = g'q_\chi$ where g' is the $U(1)'$ gauge coupling and q_χ the charge of χ). Two more parameters regulate

⁴ The model can be extended to Dirac DM fermion as well. A Majorana fermion corresponds to its own anti-particle.

the contact between the Dark Sector with the SM. They are: the coupling of quarks to Z' , g_q , and the mixing angle θ of the Dark Higgs to SM Higgs boson that allows its decay to SM states.

3.2.2. Searches at LHC

Since the massive gauge boson Z' can radiate off a Dark Higgs (relic density constraints on g_χ have large enough values to allow this process with non negligible probability) and then decay to DM, LHC searches offer an unique opportunity to explore such model. The Feynman diagram of the process that can be targeted is shown in Figure 3.6, where the DM states recoil against the Dark Higgs, generating large amounts of momentum imbalance ($\mathbf{p}_T^{\text{miss}}$). This is usually the case if the mass of the DM particle and the mass Dark Higgs are smaller with respect to the mass of the Z' , leading to a substantial *Lorentz boost* of the Dark Higgs.

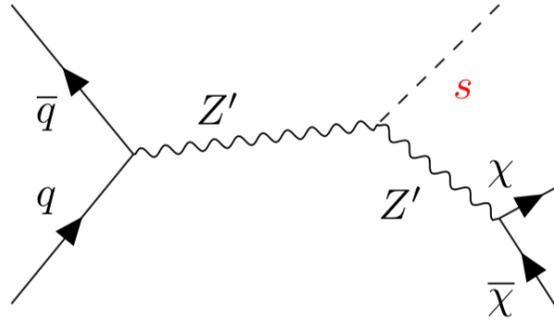


Figure 3.6.: Feynman diagram of the Dark Higgs-Strahlung. s is the Dark Higgs, Z' and χ are the gauge boson and the Dark Matter particle, respectively.

Since the Dark Higgs mixes with the SM Higgs boson, its branching ratio (BR) can vary according to its mass, as shown in Figure 3.7. As it can be seen, the BR to b-quarks dominates in the low mass range, and above 160 GeV rapidly decreases as soon as new decay channels are kinematically allowed, such as the W^+W^- , ZZ .

If m_s is close to the SM Higgs mass, searches for the Higgs boson associated with $\mathbf{p}_T^{\text{miss}}$ can be used to constrain the Dark Higgs model. One example is a search for DM produced in association with a SM Higgs boson decay to $b\bar{b}$, which used 79.8fb^{-1} of $\sqrt{13}$ TeV of proton-proton collisions with the ATLAS detector [2]. This search was subsequently re-interpreted in terms of the Dark Higgs model using the RECAST

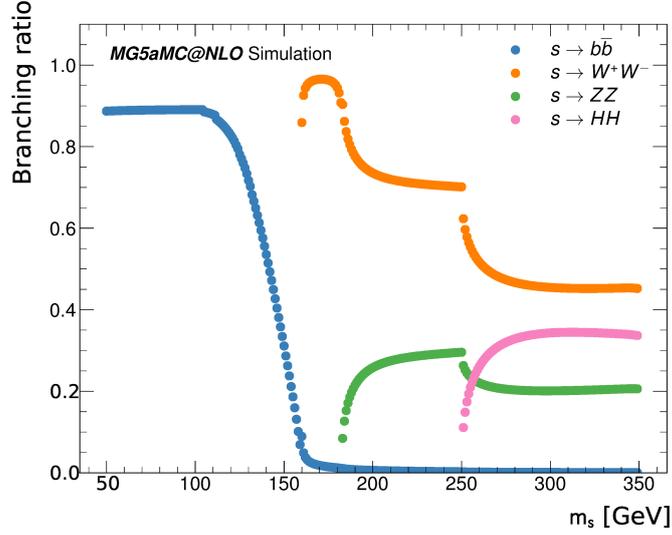


Figure 3.7.: Branching ratio of the Dark Higgs into SM states: $b\bar{b}$, W^+W^- , ZZ and HH as calculated from MADGRAPH5_aMC@NLO 2.6.6 at leading order [53], from [4].

framework [4]. The signature in this final state is also called $E_T^{\text{miss}} + s \rightarrow b\bar{b}$, because of the presence of a resonant b -quark system and momentum imbalance. In Figure 3.8, the observed 95% CL_s ⁵ exclusion contours in the parameter space $(m_{Z'}, m_s)$ are shown. The calculated relic density in the Dark Higgs scenario that matches the one observed by Planck [35] is also shown, and therefore excluded. As the mass of the Dark Higgs approaches the value of around 160 GeV, the production of a pair of W boson is not any more suppressed. This leads to a reduction of the $b\bar{b}$ decay mode and hence a degradation of the sensitivity of this search.

If, however, m_s is higher than 160 GeV, the BR to $b\bar{b}$ is so low that searches for this final state become insensitive to the Dark Higgs model. Specialized searches have to be carried out in the resonant W^+W^- or in general VV final state. The two most contributing Feynman diagrams are shown in Figure 3.9. Both processes produce a highly energetic Dark Higgs, resonantly decaying to a W boson pair. According to the energy of the pair, the bosons will be produced with small angular separation between the two in case of high boosts, or with a large one in case of smaller boosts.

As it will be discussed in the following Chapters, the search for Dark Matter produced in association with a Dark Higgs boson decaying to a W (or $V = W, Z$) pair can be carried out according to the decay channels of the bosons.

⁵ CL_s is introduced in Chapter 10.

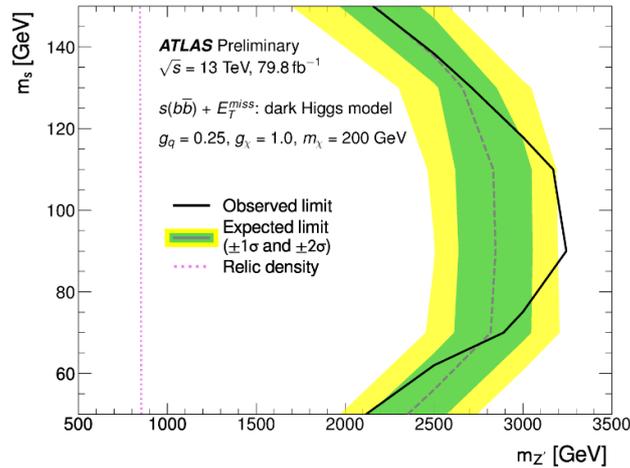


Figure 3.8.: The exclusion contour for the Dark Higgs model with benchmark parameters $m_\chi = 200$ GeV, $g_q = 0.25$ and $g_\chi = 1.0$, as a function of $m_{Z'}$ (x-axis) and m_s (y-axis). The solid line represents the observed limit. The dashed line represents the expected limit; green and yellow band show the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty on the expected limit respectively. The pink dotted line indicates the parameter points, for which the observed relic density is reproduced [4].

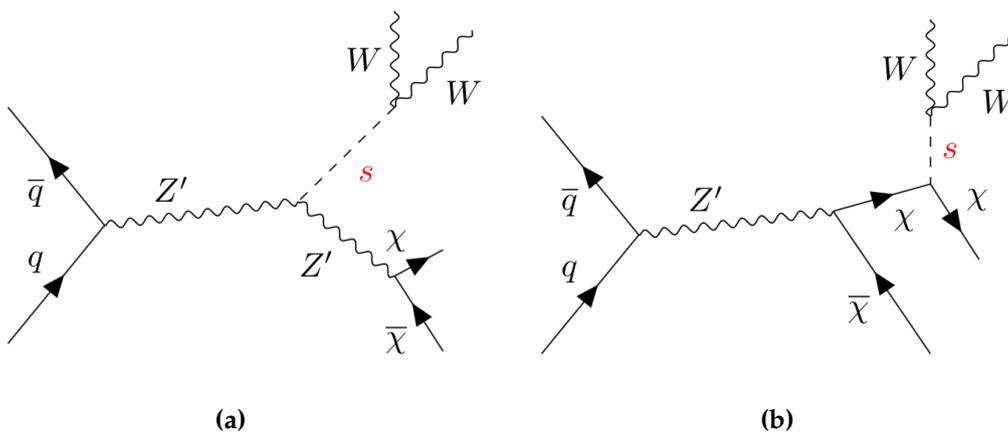


Figure 3.9.: On the left, the most contributing Feynman diagram for the W^+W^- final state: the Dark Higgs from the Z' that subsequently decays to Dark Matter. On the right, the second most contributing diagram: the Dark Higgs is radiated off a Dark Matter particle.

The *fully hadronic* search targets the hadronic decay of both the W^+ and W^- into quarks and hence into jets. The main advantage is given by the higher branching ratio of this final state, which is around 46%, benefiting then from an higher total cross-sections. However, the hadronic channels typically suffer from higher background, since SM processes with much higher cross-sections can produce the same final states. This can be addressed with specialized techniques in the reconstruction of the hadronically decaying W bosons, especially if these objects are energetic, which is uncommon to happen in background processes.

The *semi-leptonic* search targets the case in which one of the W bosons decays to electron (muon) and electron (muon) neutrino, and the other one hadronically. Since this final state still has a sizeable branching ratio (around 30%) and a lepton is present, the amount of background processes can be reduced. However, the invariant mass of the system has to be taken special care of, because of the presence of the neutrino.

The *fully leptonic* channel has in scope the leptonic decay of the W pair into neutrinos and leptons. In this case the total BR is around 4%, drastically reducing the cross-section. In the case of the searches for Dark Matter, the presence of the neutrinos could partially restore the momentum imbalance, making it more difficult to disentangle this final state from background processes. However, the presence of the two leptons makes a clean final state, easy to reconstruct and with lower background levels.

This thesis will focus on the fully hadronic search, since it offers the higher cross-sections for the signal, and allows to experiment with new tools that pertain the reconstruction of boosted hadronically decaying object as explained in the following Chapters.

Part II.

Experimental Techniques

4. The ATLAS Detector at the LHC

This Chapter introduces the experimental apparatus of the LHC accelerator complex and the ATLAS (A Toroidal LHC ApparatuS) detector, which, together with the CMS¹ detector, is one of the two general purpose detectors at the LHC.

4.1. The LHC

The Large Hadron Collider (LHC) [54] is placed along the border between the Swiss canton of the Republic of Geneva and the Arrondissement de Gex, France, at the Conseil Européenne pour la Recherche Nucléaire (CERN), the European Organization for Nuclear Research.

It is the world's largest and most powerful particle accelerator, placed inside a 27 km long tunnel, excavated for the Large Electron-Positron Collider (LEP), and designed to accelerate protons to a center-of-mass energy of 14 TeV with a peak instantaneous luminosity $\mathcal{L} = 10^{34} \text{ cm}^{-2} \text{ s}^{-12}$, reaching regions of the phase-space that were unexplored before, allowing precision Standard Model physics and searches for new physics.

The luminosity is defined as:

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

in the case of Gaussian-shaped identical beams. The N_b is the number of particles per *bunch*, circa 10^{11} protons on average in Run 2. The bunches are packets of protons organized in *trains*; there are, by design, $n_b = 2808$ of them. Each bunch arrives at the *interaction point* or IP, the point where the two beams collide, after 25 ns from the previous one, and with a revolution frequency of $f_{rev} = 11.2455$ kHz. Other important factors are the relativistic gamma factor, γ_r , the normalized transverse beam emittance, ϵ_n , which quantifies how the protons are spread in the position-momentum space. At a low emittance protons are close to each other, with similar momenta. The amplitude

¹Compact Muon Solenoid.

²The designed luminosity has already been surpassed by more than a factor of two.

function at the interaction point, β^* , determines how much the beam is squeezed at the IP. Finally the F is a geometrical function that corrects for the angle between the beams.

The luminosity is one of the most important parameters for an accelerator: the rate at which interactions are produced, $\frac{dN}{dt}$ is given by the \mathcal{L} times the cross-section σ . The higher the luminosity at a given time, the higher the number of events produced.

LHC achieves beam control through the 1232 dipole and 392 quadrupole superconducting NiTi magnets, which are operated at 1.9 K with a field of 8.33 T. The 16 radiofrequency cavities can be modulated to a frequency of 400 MHz, with a maximum voltage of 2 megavolts, accelerating the beams up to 13 TeV.

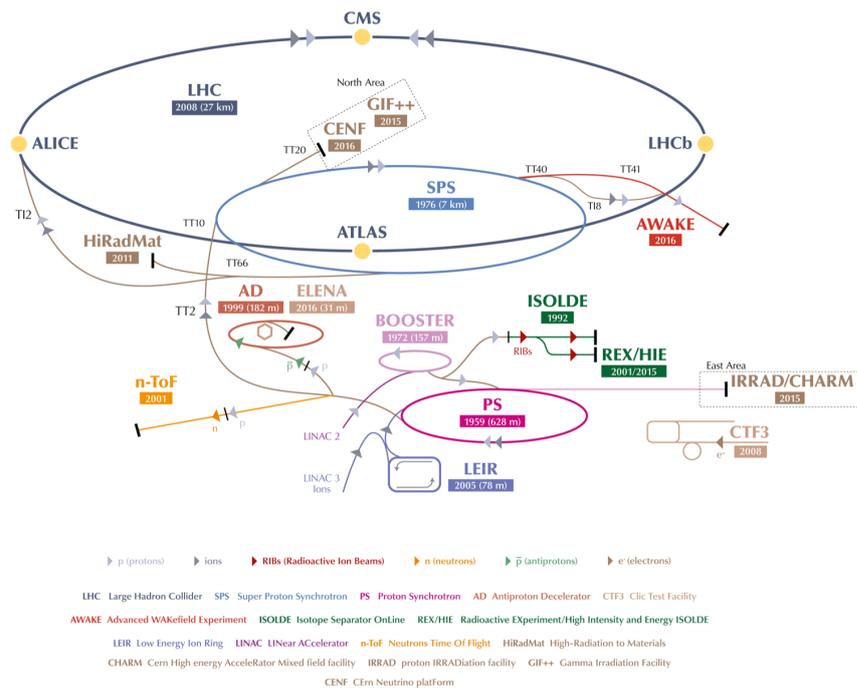


Figure 4.1.: Schematic of CERN's accelerator complex, with the experiments served. Taken from [55].

The full accelerator chain is depicted in Figure 4.1. The proton accelerators serving the LHC are: the LINAC 2 (LINAC 3 for heavy ions) provides initially the protons with and energy $E = 50$ MeV. The proton source is a bottle of hydrogen at one end of the linear accelerator. The hydrogen is passed through electric fields that get rid of the atomic electrons, leaving only protons to enter the accelerator. They are further

accelerated with the proton Synchrotron Booster at $E = 1.4$ GeV, to avoid the limitation on the number of protons that could enter the next machine, the Proton Synchrotron (PS). The PS has a circumference of 628 m, and accelerated first protons already during autumn 1959, at the time the world's highest energy particle accelerator. Now it injects protons with an $E = 25$ GeV to the next step, the Super Proton Synchrotron (SPS). It has almost 7 km in circumference, and accelerates the proton beams at $E = 450$ GeV. It is operational since 1976, becoming famous for the discovery of the W and Z bosons with the detectors UA1 and UA2. The beams are then injected further from the SPS to the LHC, or redirected to the NA61/SHINE and NA62 experiments or the COMPASS experiment. In the LHC there is a further acceleration of protons up to a total centre-of-mass energy of 13 TeV with maximum design of 14 TeV. The beams can be maintained and collided into each other providing luminosity to the experiments for long periods of time, with about 38 hours being the record time for a single *fill*, which identifies the period of time between the completion of the injections to the *beam dump*. The beam dump is an automatically or scheduled way to remove the beam from the accelerator in a safe manner. This happens for example in case of critical losses or in case the beam becomes unstable, or if the beam is providing low luminosity (luminosity goes down exponentially with time) where it becomes more convenient to dump the beam and refill it. This is meant to provide protection to the accelerator systems and to the experiments. It is realized using the so-called *abort gap*, which is a designed gap between proton bunches in the beam train. This gives 3 microseconds time for the *deflecting extraction kicker* to deflect it into a 7 m long segmented carbon cylinder absorber, water cooled and heavily shielded, which is housed in a dedicated tunnel segment.

The LHC is not a perfect ring, but a sequence of eight straight sections and eight arc sections. It lies between 45 and 170 m below the surface and has a slight inclination of about one degree towards the *Lac Léman*. Two transfer tunnels connect the main ring to the SPS for the beam injections. Figure 4.2 shows the schematics of the underground architecture. There are eight access to the tunnel, at the correspondence of the centre of the eight *octants* in which the accelerator tunnel is divided. These are named clockwise from the southernmost. Point 1 (or P1) is the location of the ATLAS detector. Point 2 and Point 8 host the ALICE³ and LHCb⁴ detectors; they previously hosted L3 and

³A Large Ion Collider Experiment.

⁴Large Hadron Collider beauty.

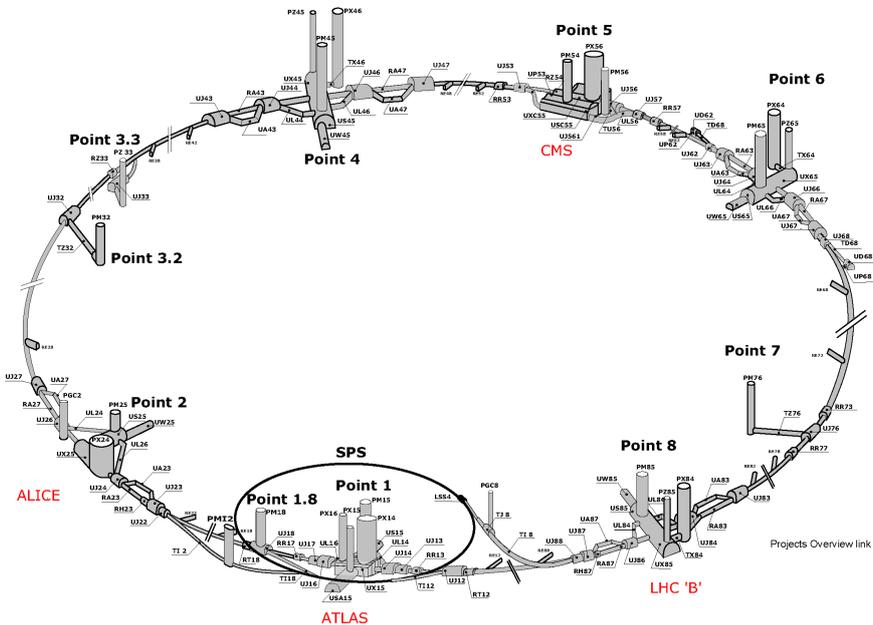


Figure 4.2.: The LHC underground layout [56].

Delphi detectors from LEP (Point 4 and 6 were devoted to the Aleph and OPAL detectors). Point 6 now hosts the beam dump system. CMS is located at Point 5.

ATLAS and CMS are two general purpose detectors and they address a similar physics program. ALICE is devoted to the study of the strong interactions that are present at densities where a phase of matter called Quark-Gluon Plasma forms. This is the state of matter that formed in the early universe, and a key component to the understanding of quantum chromodynamics. LHCb is a single arm forward spectrometer, which is specialized to the study of bottom quarks (mostly produced in the forward region). Its physics case involve the study of CP violating processes with B and K mesons and general aspects of flavour physics. These are the "big four" experiments, in terms of dimensions and participation; however, few much smaller experiments can be found at the LHC: TOTEM⁵, LHCf⁶ and MoEDAL⁷. The first two investigate proton physics at small angles, and are located few hundred meters along the beam line from the CMS and ATLAS detectors. MoEDAL investigates the possible existence of the magnetic monopoles.

⁵Total cross-section, Elastic scattering and diffraction dissociation Measurement at the LHC.

⁶Large Hadron Collider forward.

⁷Monopole and Exotics Detector at the LHC.

4.2. The ATLAS Experiment

The ATLAS detector [57], is a multi-purpose detector located at Point 1 of the LHC ring. It is cylindrically shaped, covering almost the entire solid angle around the IP. ATLAS is composed of several sub-detectors that are for the great part arranged in cylindrical layers. The different sub-detectors technologies address the variety of the physical program that the detector was planned to be sensitive to. The first layer after the IP, only 3.3 cm away from the beam axis, is the *inner detector*. It is able to measure the momentum and charge of charged particles via the measurement and identification their tracks; therefore it is also called *tracker system*. A superconducting solenoid is placed around the tracker system, which provides the bending power through a magnetic field of about 2 T for the spectrometry of particles in the inner detector. The *calorimeters* measure the energy and direction of particle, yielding as well an identification method for electromagnetic and hadronic objects. Finally, *muon spectrometers* can be found as the outermost sub-detector. They provide identification and momentum measurement for the muons, whose trajectory is bended by the superconducting *toroid magnets* that generate around 4 T of magnetic field. The trigger system, comprised of an ultra fast, hardware based section, and of a computer farm, is the last step of the ATLAS detector. The ATLAS sub-detectors are arranged with two different geometries: the *barrel* is chosen for the very central pseudorapidities (depending on the sub-detector), and consists of cylindrical arrangements around the beam pipe. The *end-cap* geometry instead focuses on the forward pseudorapidities, where the sub-detectors are placed in disks centred on the beam pipe. The end-cap can be distinguished in A-side, for positive z-axis in the direction of the airport, and C-side for negative in z-axis pointing to the Jura mountains. All the aforementioned sub-detectors are detailed in the next sub-sections. An overview of the detector can be seen in Figure 4.3.

At increasing luminosities, the chance that more than one proton collision happens in the same bunch crossing also increases. At the typical LHC values in Run 2, up to 70 interactions per bunch crossing can be achieved. Apart from the interesting hard scattering ones, most of them produce low energies collisions, which have to be disentangled from the high energy ones. All the additional interactions are referred to as *pile-up*. The ATLAS detector has been designed (and is being upgraded) also to cope with high levels of occupancy and pile-up.

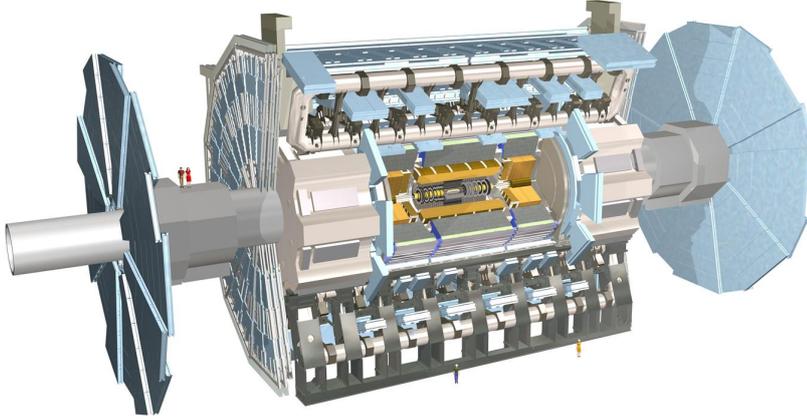


Figure 4.3.: Overview of the ATLAS detector, with its sub-components indicated on the picture. The detector is 25 m high and 44 m long, weighting approximately 7000 tonnes. From [58].

4.2.1. The ATLAS Coordinates

In ATLAS, the nominal IP is defined as the origin of the coordinate system, right-handed with the z-axis on the beam direction, towards the airport. The y-axis points upwards, slightly tilted to account for the general tilt of the LHC ring; the x-axis points towards the centre of the LHC itself. The azimuthal angle ϕ is measured around the z-axis in the x-y plane, relatively to the x-axis. The polar angle θ is measured from the z-axis in the x-z plane. Since the polar angle is not a Lorentz invariant, the *pseudorapidity* is instead used, defined as $\eta = -\ln[\tan(\theta/2)]$, with the rapidity given by $y = \frac{1}{2} \ln \frac{E+p_z}{E-p_z}$. E is the energy of the particle and p_z its longitudinal momentum. The pseudorapidity and the rapidity are the same in the limit of massless particles. Transverse quantities, such as the missing transverse momentum E_T^{miss} or the p_T are projected into the x-y plane so that e.g. $p_T = \sqrt{p_x^2 + p_y^2}$ and similarly for E_T^{miss} . The distance between objects is conveniently defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, a Lorentz invariant in the massless limit.

4.2.2. The Inner Detector

The ATLAS inner detector (ID) is designed to provide a measurement of charged particles trajectory, momentum, identification as well as primary and secondary vertex reconstruction. The phase-space covered is $p_T > 0.5$ GeV and $|\eta| < 2.5$. The vertexes reconstruction capability is a key requirement for all the experiments at the LHC, given the instantaneous luminosity that can be achieved. The ID is immersed in a 2 T magnetic field generated by the solenoid magnet, having dimensions of circa 6.2 m and 2.1 m, as can be see in Figure 4.4. The magnetic field in the ID is a key component of the track measurement. As charged particles travel through the detector in a curved trajectory, the curvature can be used to determine the charge of the particles and their momenta. If the energy increases, however, the track will increasingly look like a straight line, resulting in a degraded momentum resolution. The ID is expected to provide a measurement of the transverse momenta with a precision of $\sigma(p_T)/p_T = 0.05\% \times p_T[\text{GeV}] \oplus 1\%$. To fulfil the requirement needed to cope with the dense environment, both in terms of reconstruction and radiation, the highest granularity possible close to the interaction point is required. The ID is therefore divided into three different and interdependent sub-detectors, each relying on a complementary technology to address this challenge. The closest to the interaction point is the *Pixel* layer, using discrete space-points of silicon pixels. At higher radii the *Silicon Microstrip* (SCT) layer also provides space-points. The outermost of the ID layer is the *Transition Radiation Tracker* (TRT) that, on the other hand, is based on gaseous straw tubes, and can provide continuous hits at a cheaper cost than silicon-based tracker.

Pixel and SCT layers

Pixel and SCT were design to withstand the high radiation dose that occur in the vicinity of the interaction point. This harsh environment poses in fact limitations on the sensors, electronics, mechanical structure and services. The Pixels and SCT required state of the art technologies to meet the radiation hardness requirement, and to cope with the high occupancy while providing the high resolution position measurements possible. Pixel sensors are 250 μm thick detectors, using oxygenated n-type wafers with readout pixels on the n^+ -implanted side of the detector. This choice allows a good charge collection efficiency, and the oxygenated material has a demonstrated history of increased radiation tolerance. 1744 pixel sensor, each of

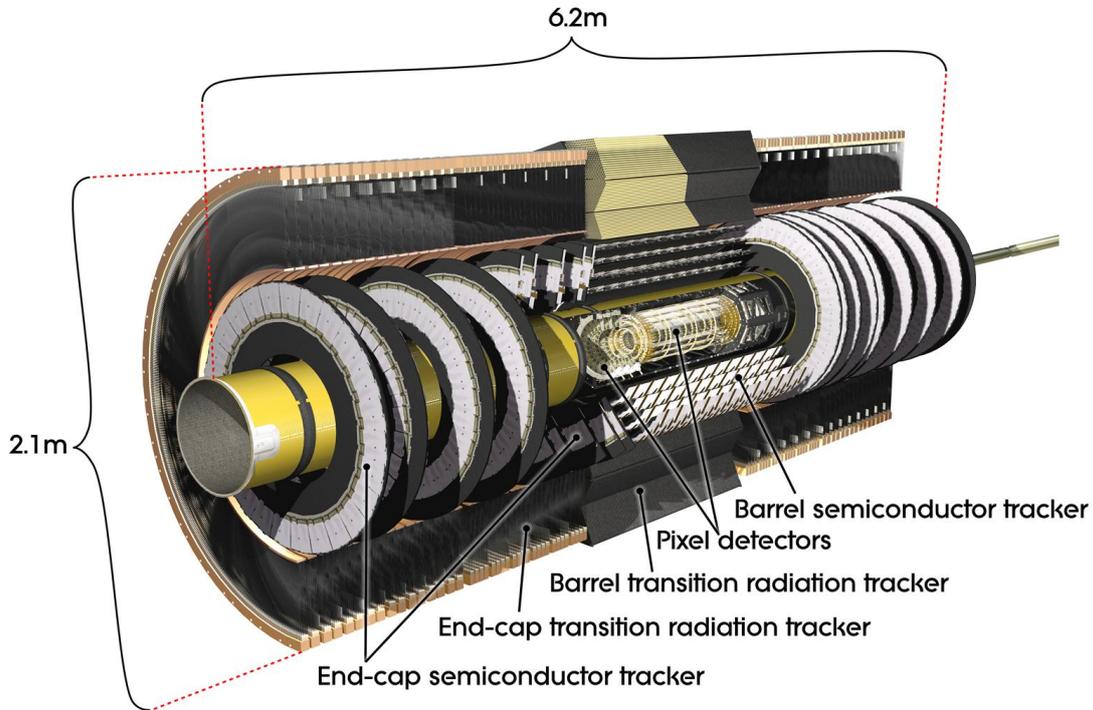


Figure 4.4.: Computer generated image of the ATLAS inner detector [59].

dimensions $19 \times 63 \text{ mm}^2$ and containing 47232 individual pixels, are arranged in three layers in the barrel, with radii 50.5 mm, 88.5 mm, and 122.5 mm and three end-cap disks, distant 495 mm, 580 mm, and 650 mm from the interaction point. A fourth layer in the barrel, called Insertable B-Layer (IBL) was installed with a radius of 25.7 mm in the ID during the long shut-down between 2013 and 2015 to fulfil the requirements imposed from the increased luminosity of Run 2. The voltages applied to operate the pixels range from a starting value of 150 V up to 600 V to allow the charge collection even after radiation damages, and at a temperature of around -10°C also to mitigate leakage currents and annealing. The spacial resolution of the pixel sensors has been measured in a test beam to be $4.7 \mu\text{m}$ before irradiation, and $6.0 \mu\text{m}$ after, at incident angles of around 10° .

The SCT is also based on silicon sensors, but for reasons of costs and reliability, uses instead a single-sided p-in-n technology with coupled readout strips with a thickness is around $285 \mu\text{m}$. There are 15912 sensors each containing 768 active strips of 12 cm length. The SCT modules are 4088 with about 4 (2) sensors for the barrel (end-cap), arranged in four layers in the barrel with radii of 299 mm, 371 mm, 443 mm and 5541 mm and nine disk layers in the end-cap, at distances of 853 mm, 934 mm, 1091 mm,

1299 mm, 1399 mm, 1771 mm, 2115 mm, 2505 mm, and 2720 mm on the z-axis. Their resolution is about 17 μm on the later plane and 580 μm on the longitudinal plane, as measured in the test beam.

TRT layers

The Transition Radiation Tracker relies on a different technology, based on Polyimide drift straw tubes of 4 mm diameter. Each straw is a small gas proportional chamber, with an anode of gold-plated tungsten wire in the centre, and cathode of a specially manufactured walls. The cathodes and anodes are operated at 1.5 kV potential difference, achieving a gain of 2.4×10^4 for the gas mixture used (70% Xe, 27% CO₂ and 3% O₂, at around 5 to 10 mbar over-pressure). This setup provide a relatively inexpensive and radiation resistant sub-detector, yet supplying a large number of hits (around 36 for high p_T tracks) within a pseudorapidity of $|\eta| < 2.0$. The barrel region is occupied, in radii from 55 to 108 cm, by 50000 straws; the end-caps, from 64 to 103 cm, by 320000 radial straws organized in 14 wheels. The straw tubes are embedded into a matrix of polypropylene fibres designed and optimized to stimulate the transition radiation, which is emitted by relativistic charged particles as they traverse a material boundary. The TRT allows particle identification, in particular for electrons and pions. The typical resolutions are about 130 μm on the later plane.

4.2.3. The Calorimeters

The ATLAS calorimeters are designed to measure the energy and direction of electromagnetic or hadronic interacting particles by stopping them completely (or at least significantly) in their volumes. The incoming particles face alternate layers of a high Z material, also called *passive*, which stimulates the interactions and produces a cascade of secondary particles, the *shower*. The *active* material is used measure the ionization induced by the secondary particles. This type of calorimeters are called *sampling calorimeters*, since the energy is sampled only in the active layer, and not in the passive one. The basic unit of the calorimeters is the so-called calorimeter *cell*, where the energy is measured in its active part. The geometric size of the cell is critical, as it impacts the angular and longitudinal resolution. ATLAS has optimized the design of the calorimeter system, providing two different sub-detectors. The *electromagnetic calorimeter* (or EM calorimeter) is optimized for the measurement of electromagnet-

ically interacting particles, such as electrons or photons. The shower is induced by pair production or bremsstrahlung, and its length is characterized by X_0 , the *radiation length*. A particle travelling an X_0 into a material has $1/e$ of its original energy left. The *hadronic calorimeter* (or HAD calorimeter) instead, focuses on hadronically interacting particles, such as pions or protons. The basic distance is the *interaction length* λ_0 , which is the mean distance travelled by those particles before undergoing inelastic nuclear interaction. λ_0 is typically much larger than X_0 , up to 30 times more. Therefore the hadronic calorimeter requires a much larger depth than the electromagnetic one. The entire calorimeter system layout can be seen in Figure 4.5. The usage of the calorimeters has two main advantages. Firstly, unlike the tracker system, the calorimeters provide measurements for both charged and neutral particles produced in the hard interactions. Secondly, the precision of the measurements increases with increasing energy (with more energy, more secondary particles are produced). This also is the opposite of the tracker system, where the precision degrades with increasing particle's energy, as seen above. The general behaviour for calorimeters precision is $\frac{\sigma(E)}{E} = \frac{\alpha}{\sqrt{E[\text{GeV}]}} \oplus \frac{\beta}{E} \oplus \gamma$, where α and β and γ are some constants that depend on the calorimeter system. α represents the sampling term, β is the noise term and γ a constant term.

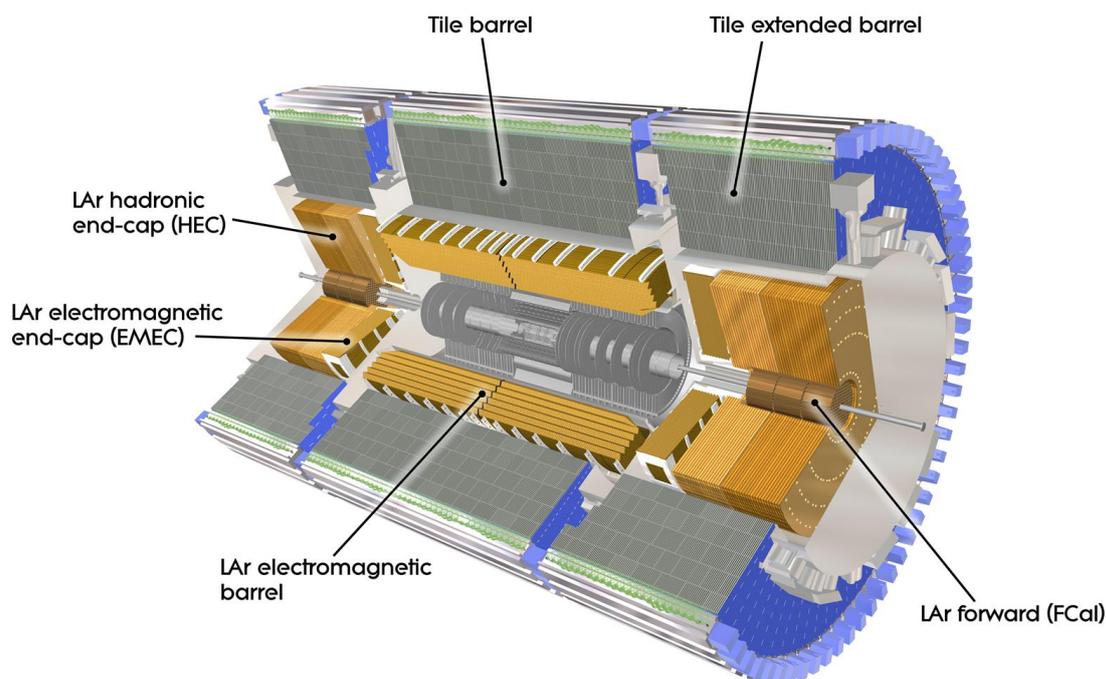


Figure 4.5.: Computer generated image of the ATLAS calorimeters [60].

Electromagnetic calorimeter

The electromagnetic calorimeter covers up to $|\eta| < 4.9$ and with full ϕ coverage. It is divided in a barrel (EMB) for a precision measurement within $0 < |\eta| < 2.5$, obtained by finer granularity in the first layers, and even a pre-sampler at $0 < |\eta| < 1.8$. In the forward region, the electromagnetic end-cap (EMEC) covers $2.5 < |\eta| < 3.2$. All these calorimeters are housed in different cryostats. The precision electromagnetic calorimeters use lead as passive material and liquid argon (LAr) as active, with accordion-shaped absorbers and electrodes. LAr was chosen for its stability and radiation-hardness; it is maintained at a temperature between 88.5 K and 88.6 K. The lead makes an optimal passive material both because of its affordability and its enhanced shower containment, within $X_{0,Pb} \simeq 0.56\text{cm}$. The accordion shapes assures a full ϕ coverage and a fast signal extraction at the electrodes. The accordion sheets have a thickness of 1.53 mm (1.13 mm) for $|\eta| < 0.8$ ($|\eta| > 0.8$) in the barrel and in the end-cap 1.7 mm (2.2 mm) for $|\eta| < 2.5$ ($|\eta| > 2.5$). The sheets are thicker at higher pseudorapidities to limit the decrease of the sampling fraction. The electrodes of the readout electronics are arranged in three conductive copper layers insulated with polyamide sheets. The two external electrode are kept at a potential of 2000 V, and the signal is read out from the middle electrode. In the end-caps the voltage varies, depending on η , between 1000 and 2500 V. As the secondary particles from the shower hit the LAr, they ionize the medium. The electron from the ionization are collected by the electrodes within a characteristic time of 450 ns. Since the LAr is uniform in the calorimeters, the signal at the electrodes results in a clean triangular shape from the electron collection. Furthermore, a signal shaper is applied. The EMB has three layers in depth with different granularities, as can be seen in Figure 4.6. The first layer has a fine $\Delta\eta = 0.0031$ spacing that allows discriminating photons from neutral pions decay. The second and the third layer have a total radiation length of about $20 X_0$, and their purpose is to collect as much energy from the shower as possible. The second layer has a granularity of 0.025 and 0.0245 in $\Delta\eta$ and $\Delta\phi$. The EMEC has a two wheel geometry of radii 330 mm to 2098 mm in the A- and C-side, covering from 1.375 to 3.2 in $|\eta|$. The transition between the EMEC and the EMB is referred to as *crack region* of the calorimeter, and contains several radiation lengths of materials, which mainly serve the electronic of the inner detectors. The total thickness of the EM calorimeter is about 22 to 24 radiation lengths. The energy resolution of the EMB and of the EMEC is about $\frac{\sigma(E)}{E} = \frac{10.1\%}{\sqrt{E[\text{GeV}]}} \oplus 0.17\%$ for the former and $\frac{\sigma(E)}{E} = \frac{12.1\%}{\sqrt{E[\text{GeV}]}} \oplus 0.4\%$ for the latter [61].

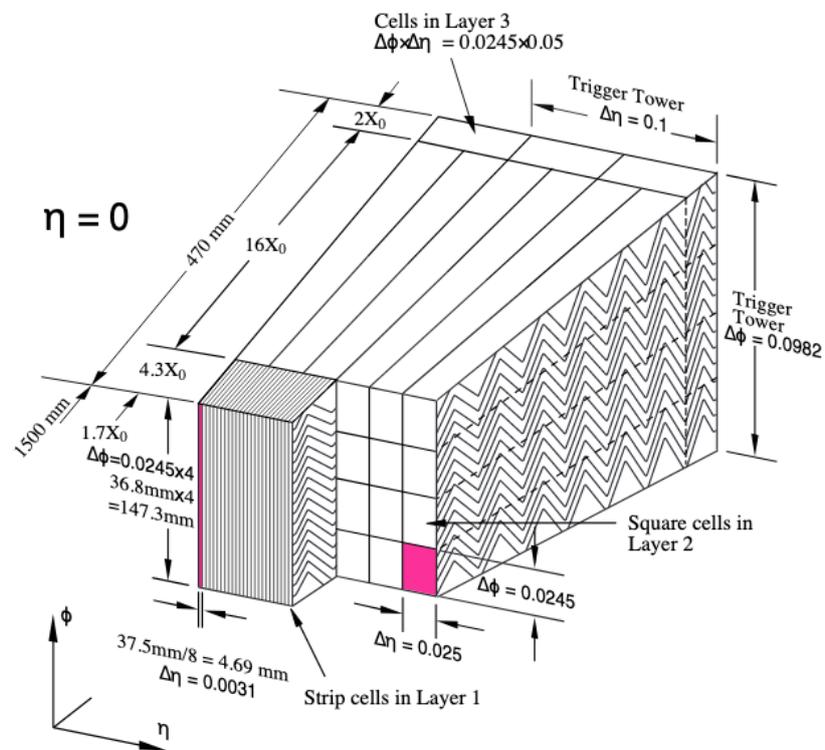


Figure 4.6.: Structure of the EMB calorimeter with the granularities in η - ϕ dimensions shown [57].

Hadronic calorimeter

As strongly interacting particles enter the hadronic calorimeter, they initiate showers that are different by the ones initiated by electron or photons, being dominated by hadronic processes. Hadronic showers are typically much wider and longer. Moreover, part of the energy of the shower is absorbed in the nuclear binding energy, and also carried away by invisible particles such as neutrinos. This leads to non linearities as its response is different to electromagnetic and hadronic component. This type of calorimeter is called *non compensating*. The hadronic calorimeter is just outside, in radial direction, of the EM calorimeters, designed to address the measurement of energy and direction of particles subject to the strong force. Similarly to the EM calorimeter, it provides a sampling measurement with active and passive material. The hadronic barrel is also called *tile* calorimeter. It uses plastic scintillating tiles as active medium and steel as passive, chosen to have a moderate interaction length $\lambda_I=16.8$ cm and a good trade off between material cost and its massive extension. The measurement principle is based on photon readout with photomultiplier tubes from the scintillating fibres. The tile calorimeter fully envelopes the EMB and EMEC. It is subdivided in three parts: one barrel ($|\eta| < 1.0$) and two extended barrels ($1.0 < |\eta| < 1.7$). A tile module is illustrated in Figure 4.7. The hadronic end-cap (HEC) uses instead LAr and copper, due to the higher radiation levels associated to the forward regions, which are closer to the beam. It is placed at $1.7 < |\eta| < 3.2$. The total number of interaction lengths is about 7 to 16 λ_I , assuring the containment of the hadronic showers. The typical granularity of the hadronic calorimeter is 0.1×0.1 ($\Delta\eta \times \Delta\phi$) in the central region ($|\eta| < 2.5$) and 0.2×0.2 for the more forward ($|\eta| > 2.5$). It can achieve a performance of $\frac{\sigma(E)}{E} = \frac{52.7\%}{\sqrt{E[\text{GeV}]} \oplus 5.7\%$ for tile and $\frac{\sigma(E)}{E} = \frac{70.6\%}{\sqrt{E[\text{GeV}]} \oplus 5.8\%$ for HEC.

Forward calorimeter

The Forward Calorimeter (FCal) covers the most forward region in pseudorapidity $3.1 < |\eta| < 4.9$. It is dovetailed between the EMEC and the beam pipe, providing a coverage of the transition regions between the sub-detectors. Divided into three layers, one for EM calorimetry the other two for hadronic calorimetry, it is exposed to intense radiation, requiring radiation-hard components at expense of the measurement performance. Copper is therefore chosen as passive material in the EM layer, being

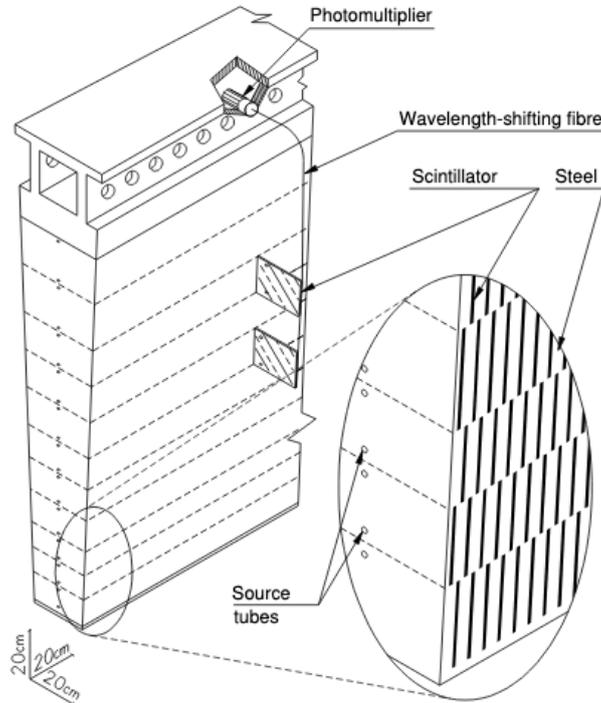


Figure 4.7.: Tiles, fibres and photomultipliers composing the optical readout of a tile calorimeter module [57].

able to disperse the heat better than lead, and requiring smaller cooling infrastructure. LAr was preferred as active material, and a special geometry was designed in the gaps to avoid ion accumulation. In the last two layers, tungsten was chosen to allow better shower containment. FCal achieves a performance of $\frac{\sigma(E)}{E} = \frac{92.2\%}{\sqrt{E[\text{GeV}]} } \oplus 7.5\%$.

4.2.4. The Muon Spectrometers

The muon spectrometers, or muon systems (MS), is the most external from the IP of all the sub-detectors. They are designed to measure the direction and momentum of minimum ionizing particles (MIP), such as muons, which escape the upstream sub-detectors. The size of calorimeter, as seen above, assures that few non-MIPs from the showers reach the spectrometers (this is effect is called *punch-through*). The muon spectrometers are immersed into a strong magnetic field (4 T on the superconductors) generated by the toroid magnets. The field allows to reconstruct the curvature (in the η direction since the toroid field is oriented in the transverse plane) of the outgoing muons and therefore their charge and momenta (the bending power is approximately

1 to 5.5 Tm). Four sub-detectors compose the MS: two in the barrel and two in the end-caps. The redundancy in each η region is chosen to provide two kind of measurements. One is quick, integrated with the trigger system and used for triggering, bunch crossing identification and cosmic ray veto using its exceptional time resolution of 1.5 to 4 ns. For such a purpose the Resistive Plate Chambers (RPC) are used in the barrel ($|\eta| < 1.05$) and the Thin Gap Chambers (TGC) in the end-caps ($1.05 < |\eta| < 2.7$). Instead, the Muon Drift Tubes (MDT) and the Cathode Strip Chambers (CSC) are specialized on excellent measurement of the track parameters, with a resolution of about is 35 μm per chamber for the former, and 40 μm in the longitudinal and 5 mm in the transverse plane. The MS therefore is able to provide precision measurements, achieving a resolution of $\frac{\sigma_{p_T}}{p_T} = 10\%$ for $p_T = 1$ GeV muons.

4.2.5. The ATLAS Magnets

The magnet system consists of one solenoid and three toroids. The solenoid is aligned with the beam axis and provides 2 T magnetic field for the inner detector while minimizing the material budget in front of the calorimeters. The toroids are placed two at the end-caps and one at barrel, and generate the magnetic field of about 0.5 to 1 T needed by the muon spectrometers. This system is unique given its size and field strength, covering approximately 12000 m^3 of volume in the experimental cavern. The solenoid operates at 7730 kA with a layout carefully optimized to contribute only with 0.66 radiation lengths. In order to achieve this, the EM calorimeter and the solenoid share the same vacuum vessel. Only a 2 mm thick aluminium panel, that acts as heat shield, is placed between the solenoid and the inner wall of the cryostat. The material choice of the superconducting single-layer coil was also chosen to allow it to be as thin as possible. Al-stabilized NbTi conductor is in fact used, and it can fit into 12 mm thick Al support cylinders. The total energy stored in the coil is 40 MJ, and it weights approximately 5.4 tonnes. The steel of the hadronic calorimeter and its structure act as field returner. Operationally it can be charged and discharged in 30 minutes, and in case of quench (the condition where the magnet loses its superconducting properties), the cold mass temperature rises from 4.5 K to 120 K. It can be cooled down again in about a day. The toroids use a pure Al-stabilized Nb/Ti/Cu conductor, operated at 4.6 K and arranged in toroidal (hence the name) shapes. The operational current is about 21 kA, storing about 1.1 GJ of energy. In case of quench, it needs a much longer time of 50 hours to return to the operational temperature.

The field is carefully monitored since its precise knowledge is essential in the momentum measurement of tracks in the inner detector and in the muon spectrometers. Since the magnetic field is quite uniform inside the solenoid, 4 NMR probes are placed around the ID, and can measure the field strength with an accuracy of about 0.01 mT. The magnetic field generated by the toroid, on the contrary, is highly non-uniform, resulting in strong gradients. It is therefore monitored by around 1730 Hall cards, which can stand high field gradients.

4.2.6. The Trigger system

A trigger and data acquisition system is a key component of all high energy detectors and especially those at the LHC. The magnitude of the inelastic proton-proton cross-section at the LHC energies is about a hundred mb, by far the dominant process observed. At the design luminosity, about tens of inelastic events can be expected at each bunch crossing, every 25 ns. By contrast, W or Z bosons production cross-sections lie at around 6 orders of magnitude below, and processes that are typically addressed by the ATLAS physics program, like $t\bar{t}$ or Higgs boson production are even rarer. Since the recording of the data is limited by the rate at which this can be technically done (and also by the available storage capability), the ATLAS detector makes use of a two-step trigger system. The aim is to reject those processes, such as low energetic QCD events from the inelastic scattering, which are overwhelmingly produced (at a rate of \sim GHz) and of scarce interest. The size of a single recorded event is roughly around a MB, and only a recording rate of about 1 KHz with the current technological capabilities can be afforded. The ATLAS trigger system is a multi-step trigger, based on a first layer (L1) [62], using a ultrafast, custom hardware-based trigger, capable of rejecting the majority of the events, bringing the rate down from 40 MHz to \sim 75 kHz. To achieve this performance, the first layer relies on the measurement of the calorimeter system and the muon spectrometers. The L1 trigger system is therefore composed of the *L1Calo*, responsible for the calorimeters, and the *L1Muon* ([63] [64] and references therein), responsible to handle the trigger decision based on the muon sub-detectors. The recently commissioned L1Topo [65] combines the information of the two, being able to compute basic observable such the di-jet or di-muon invariant mass e.g. accessing the angular separation of the pair. The second and last layer of the trigger system is the *High Level Trigger* (HLT) ([66] and references therein), which is instead software based, employing thousands of CPUs on a dedicated computer

cluster. Typical variables that are used to perform the trigger decision are: transverse momenta of electron, muons, taus, photons, jets and missing transverse energy.

Triggers for common objects such as jets and leptons are typically *prescaled*. This is because, especially for low energy objects, the production rates are so high that triggering all of them would quickly saturate the available bandwidth. Therefore, only a fraction of them is registered by the trigger system. In case of *un-prescaled* triggers, the target object is required an high enough energy, and its rate is not overwhelming. In this case, all these objects activate the trigger system.

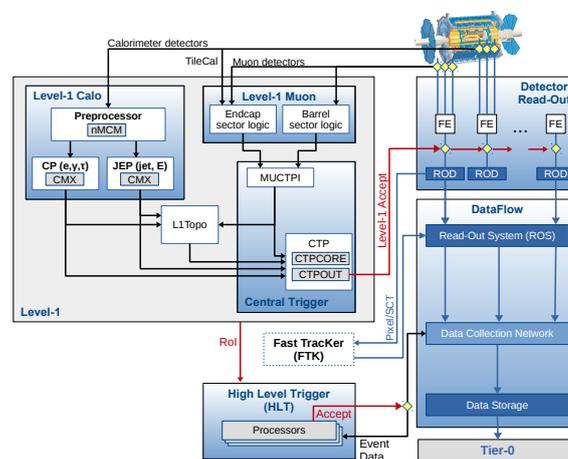


Figure 4.8.: The ATLAS Trigger and Data Acquisition (TDAQ) system in Run 2 [67]. The Fast Tracker was decommissioned and will not be part of the TDAQ in future.

The complete flow of the ATLAS *Trigger and Data Acquisition (TDAQ)* in place for Run 2 is shown in Figure 4.8. The L1 decision is formed by the *Central Trigger Processor (CTP)*, which receives inputs from the L1Calo and the L1Muon and other subsystem. It also enforces the dead-time on the L1 decision in order to prevent data overflow. Events triggered by the L1 are buffered in the *Read-Out System (ROS)* and processed by the HLT. If they are accepted by the HLT, they are transported to local storage first at the experiment site, and subsequently at CERN's Tier-0 facilities for offline reconstruction.

L1Calo The L1Calo system is based on analogue signals from the ATLAS calorimeters. Instead of using all of the calorimeter cells, the L1Calo uses a coarser granularity

by an analogue summation of signals from all the cells in a $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ ($\Delta\eta \times \Delta\phi = 0.4 \times 0.4$) in the central (forward $|\eta| > 3.4$) region. The calorimeter is divided in 7168 *Trigger Towers* (TT), split between the EM and hadronic calorimeter. The analogue signals are digitized at a standard sampling rate of 40 MHz (can be raised to 80 MHz for small instantaneous luminosities, e.g. at the end of the fill). A critical task, known as *timing*, is to make sure that all the inputs are properly timed in together. If the system is not well timed, in the best case scenario trigger inefficiencies could be introduced, in the worst case mis-triggering. The timing calibration is not simple, as several effects determine the arrival time of the signal, such as cable lengths and hardware maintenance replacements. Timing has to be determined with calibration runs and with data. The author has contributed to the calibration of the ATLAS L1Calo timing during his qualification task [68] [69]. The digitized signal is then converted to GeV by the PreProcessor Module (PPM), and is passed to the Cluster Processor (CP) and the Jet-Energy sum Processor (JEP). CP and JEP can run simple algorithms for photon, electron, tau and jet identification, as well as global variables such total and missing transverse energy. These algorithms form *Region of Interest* (RoI), which group interesting TT and transmits them to the CTP.

5. The Track-Assisted Reclustered (TAR) jets

In this Chapter, the definition of a novel observable for the reconstruction of highly energetic hadronically decaying objects will be discussed. Hadronic objects are the vast majority of those produced in proton-proton collisions at the LHC. As outlined in the previous Chapter 4.2.3, the ATLAS detector addresses the reconstruction of hadronic objects with a multi-facet approach based on dedicated sub-detectors, such as the calorimeters. The next Sections will firstly introduce the concept of *jets* in high energy physics, and provide a brief description of how they are treated in regard of their calibration and the derivation of their uncertainties. If the hadronically decaying object is energetic enough, it can be reconstructed with a single *large-radius* (Large- R) jet. Finally, the performance and versatility of Large- R jets can be greatly improved using a technique called *Track-Assistance* [70], which can be applied constructing the Large- R jet. This procedure is called *Track-Assisted Reclustering* (TAR) and will be detailed in this chapter as well. TAR jets are demonstrated to achieve a better performance with respect to standard techniques both in resolving with high accuracy the mass of the decaying object, but also in the background rejection power. This is realized exploiting the structure of the jet. The jet *substructure* is therefore an important tool that can be used with this scope in mind, and it is also briefly introduced here.

The complex final state that is offered by the Dark Higgs scenario, comprising a pair of energetic W or Z bosons, is a demonstration of the utility of this method, as shown in the search presented in the last Part of this thesis.

5.1. Jets and Reconstruction Techniques

Unlike other objects that are studied with the ATLAS detector, quarks and gluons present in the final state cannot be individually reconstructed. Being dominated by QCD interactions, they undergo a series of processes resulting in a number of secondary particles, typically mesons such as pions, but also kaons, protons and neutrons. Each of them carries only a fraction of the momentum of the initial quark/gluon. The high particle multiplicity arising from these stochastic processes and taking place

in relatively collimated area, poses per-se a reconstruction challenge. This "spray" of particles is referred to as *jet*, which is algorithmically defined to select the energy deposits in the calorimeter corresponding as much as possible to the energy of initial quark/gluon. The jet is then regarded as a mean to "integrate out" the QCD processes that the final state quarks or gluons experience, since they are not relevant to the physics taking place in the hard proton-proton scattering.

In a LHC proton-proton event, the *hard scattering* between two partons is characterized by large momentum transfers, and it can be treated perturbatively. It is described by the Matrix Element (ME) at a given order based on Feynman diagrams. *Initial State Radiation* (ISR) is the radiation composed of quarks/gluons emitted by incoming parton before the hard scattering. *Final State Radiation* (FSR) instead stems from the outgoing parton, after the hard scattering. The *Underlying Event* (UE) refers to all the other processes, apart from the hard scattering, which take place in the proton-proton interaction. Being the proton a multi-partonic object, the low momentum transfer interactions that take place with the other partons also participate to the event.

Final state partons undergo then the so-called *Parton Showering* [71], which is a cascade of QCD radiation typically consisting of soft quarks and gluons. This happens at the scale of $1/E \ll 1 \text{ fm}$, and the emitted radiation is mostly collinear with respect to the final state parton, $dN/d\Theta^2 \sim 1/\Theta^2$ [72]. Finally, *Hadronisation* takes place. Since the quark/gluons now present in the shower start entering a low momentum transfer regime, effects of colour confinement occur. Coloured partons begin to recombine and form mesons and baryons. QCD becomes non-perturbative and its description has to rely on a series of phenomenological models. The most important are the Lund string and the cluster model. In the Lund string, the interactions is modelled with a string-like potential, $V \sim kr$, with r the distance between quarks/gluons. When the string reaches a critical value, it splits, or fragments. In the cluster model instead, the parton cascade is arranged in clusters of colour-neutral, small invariant mass sets of quarks/gluons. The colour in this case is pre-confined. Mesons and baryons are then free to travel through the ATLAS detector. The electromagnetically charged ones will leave tracks in the tracker system, but both charged and neutral component will finally interact with the calorimeters.

Since quarks/gluons are the most common objects found in high energy proton-proton collisions, they are present also in the up to 70 interactions of an LHC event. This pile-up partons (*in-time* if coming from the same bunch crossing, *out-of-time* if in a neighbouring one) will also undergo the PS and hadronisation, forming particles that

interact with the calorimeters. Because of the high multiplicities, pile-up particles can lead to inefficiencies in resolving the primary interaction¹.

As the energy is deposited in the calorimeter cells (the cells structure is detailed in 4.2.3), these are grouped together into bigger three-dimensional objects called topological clusters or *topoclusters* [73]. The topoclusters are formed to minimize the noise and pile-up contribution, rejecting cells above a certain threshold in significance $\zeta_{cell} = E_{cell}/\sigma_{noise}$, and retaining those above it. σ_{noise} contains both effects from electronic noise and pile-up. The clustering starts with cells with the highest significance and goes on with the lower ones. In order to take care of the case when two particles hit the calorimeters in close proximity, the topocluster is split if there are more energy deposits $E_{cell}^{EM} > 500$ MeV. This is because the topoclusters are then defined to have a total energy that is the sum of the energies of its components, a direction that is a weighted average of the directions of the components, but are assumed massless. This is a crucial point in fact, which relates the angular resolution of the calorimeter with the efficiency in the mass reconstruction. In extreme kinematic regimes such as those in Lorentz boosted topologies, decay products can be so close together that eventually they are reconstructed into a single massless object, endangering the precision of the measurement. To overcome this challenge, calorimeter-based measurements can be integrated with track-based ones, which inherently have an higher angular resolution, as shown in the next Section 5.2.

Because of the non-compensating nature of the ATLAS calorimeters, the different behaviour at the electromagnetic and hadronic level has to be taken into account. Moreover, signal losses present due to the clustering and due to the passive materials, can also be taken care of. This very first calibration, which restores the same response of the two different sub-detectors and addresses the other sources of inefficiencies, is called *Local Cluster Weighting* (LCW) calibration. Topoclusters corrected with the LCW are called *LCTopoClusters*. If they are not corrected with the LCW, but are used for the reconstruction at the electromagnetic scale only, are referred to as *EMTopoClusters*. LCTopoClusters and EMTopoClusters have been used during Run 1 and 2 as the building blocks of the jets in ATLAS.

¹The primary interaction is the one containing the hard scattering event.

The Clustering Algorithms

As shown, the topoclusters are used to build the jets, which are algorithmically defined [74]. During the earlier HEP studies on jet production, the necessity of rigorous definitions for jets, both on the theoretical and experimental side emerged. Since the spray of particles in the jet appear to be well contained within a narrow cone, the first jet definitions used this kind of geometrical objects². However, this naive approach showed its limitation and was disfavoured later at hadron colliders, because of ambiguities in how to correctly define angular and energy scale [74] even if the stablest cone was looked for iteratively. The basic requirements were set e.g. in [76]. Apart from the requests to be as simple as possible, both in theory and computationally, a jet clustering algorithm should be *collinear* and *infrared* safe.

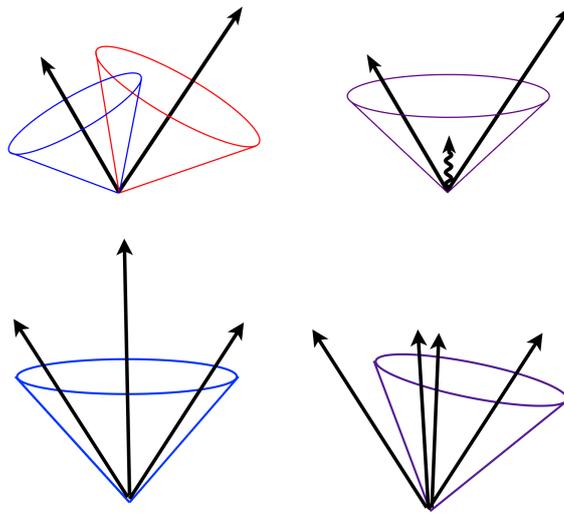


Figure 5.1.: Illustration of the infrared and collinear safety that a jet clustering algorithm should have. The jets are shown as cones and their components as arrows. In the top left diagram, two different jets were identified. In case of non infrared safety, the algorithm merges the two jets if a soft component is added between the two (top right diagram). On the bottom left, a jet is identified from three components. In case of non collinear safety of the jet algorithm, a collinear splitting of the component in the middle causes the jet axis to shift and skip the left most component (bottom right).

As it can be seen in Figure 5.1 and caption, an infrared safe jet algorithm does not change its jet if, for example, a soft component is added in the area of the jet. A

²George Stermann and Steven Weinberg defined the first jet algorithm for e^+e^- colliders back in 1977 [75].

collinear safe one does not change if a component is split into a pair of collinear constituents.

The most common jet algorithms in use nowadays at the LHC experiments are the sequential recombination algorithms [77]. These are a class of algorithms parametrized by the power of the energy scale in the distance measure. The three algorithms in this class are the k_t , *Cambridge/Aachen* (C/A) and anti- k_t . They are all defined via the distance measure between the entities³ that are being clustered together, i and j :

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

To determine when to stop the clustering, the distance between the input i and the beam B is used:

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

where Δ_{ij} is the distance in pseudorapidity and azimuth between i and j , and k_t is the transverse momentum.

The algorithms start by identifying the smallest of distance d_{ij} . If it is smaller than d_{iB} the inputs i and j are combined together, and if it is instead bigger than d_{iB} , then the entity i is called jet and removed from the list of inputs.

R is a parameter denoting the *radius* of the jet and p can be used to distinguish the algorithms of this class:

- $p=1$ defines the k_t algorithm, which clusters first the low p_T entities at larger distance.
- $p=0$ defines the C/A algorithm, which considers only the geometrical scales for the clustering.
- $p=-1$ defines the anti- k_t algorithm, which starts with high p_T entities at smaller distance first.

The *active area* is an area in the $\eta - \phi$ plane, defined as follows. Adding so-called *ghosts* in the plane, which are massless and with zero p_T (such as not to alter the clustering procedure with physical objects), the active area is the area where the ghosts are being clustered with the jet. It can be seen in Figure 5.2 as the coloured areas.

³Entities such as topoclusters or everything being clustered.

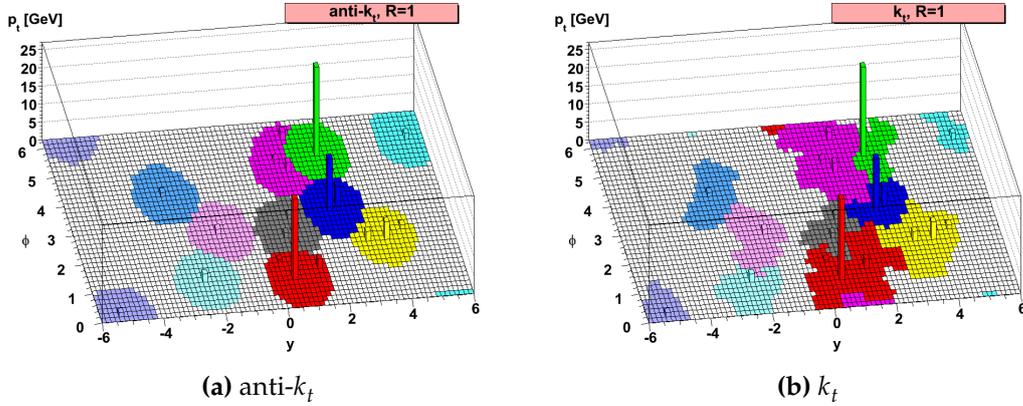


Figure 5.2.: The anti- k_t (left) and k_t (right) jet clustering algorithms, shown in the rapidity azimuth plane with p_T on the z-axis for the same event. The anti- k_t exhibits circular shapes of its active area. In case of overlap between a hard and a softer jet, the harder one will extend its border into the softer one as much as R . If two neighbouring jets are at the same p_T scale, none of them will prevail and the border will be in-between. The k_t algorithm clusters with preference to softer components, and results in asymmetrical shapes. Hard jets could be clustered with a smaller area with respect to softer ones. From [77].

Ghost-association is one of the methods of associating objects to jets. The objects to be associated (can be electrons, tracks etc.) have their masses and transverse momenta set to zero, retaining only their positions. If they end up in the clustering of the jet, they are associated to it. The ghost-association is a well defined way to associate objects and has no ambiguities. The ΔR association looks instead for the $\eta - \phi$ distance of the object with respect to the jet; if it is smaller than the radius of the jet, it is associated to it. This procedure is quick and easy to use, safe in general for anti- k_t jets, where the shapes are circular. In case of k_t jets, ghost-association is a preferable alternative.

Now that the jet is defined, it needs to be calibrated, which is done both using simulated event and data, and it needs to have uncertainties associated to it, such that it can be used in physical analyses. In the next paragraphs the most common jet definitions used in ATLAS will be detailed.

Jets and Radius: Small- R , Large- R and R -scan jets

During Run 1 and 2, more jets definitions have been used; for clarity they are briefly explained here.

Small-R jets are the most typical ATLAS jets. They are anti- k_t based jets using topoclusters at the electromagnetic scale. The clustering radius is $R = 0.4$, therefore not suited for studies of Lorentz boosted objects, but rather to resolve individual quarks/gluons. They are called ANTIKT4EMTOPO.

Large-R jets are the main tool that allows the reconstruction of highly energetic hadronically decaying objects. They are based on the anti- k_t algorithm with larger radius, $R = 1.0$, and constructed on LCW topoclusters. They are named ANTIKT10LCTOPOTRIMMEDPTFRAC5SMALLR20. The *TrimmedPtFrac5SmallR20* refers to the grooming technique that is applied to them and its configuration. It will be explained in the following Section 5.1.1.

R-scan jets are named after the effort of providing a scan in the radius of anti- k_t jets. The radii that have been considered are $R = 0.2$ and $R = 0.6$, since $R = 0.4$ is already provided. They are based on LCW clusters and they are named ANTIKT2LCTOPO and ANTIKT6LCTOPO. The $R = 0.6$ jets can help understanding the theory to data differences if compared to standard anti- k_t $R=0.4$ jets; furthermore they can constrain the modelling of the hadronisation processes and underlying event. The anti- k_t $R=0.2$ jets can find wide applications in the field of re-clustering of jets: they can be used as input in the construction of larger jets. The intrinsic advantage is that more complex objects can be built out of them, such as Large- R jets of arbitrary radius, which can be helpful in analyses addressing different kinematic regimes. One example of this is the TAR jet that can be constructed with anti- k_t $R=0.2$ jets (but also anti- k_t $R=0.4$ jets), and is discussed in the following.

The Next Generation: PFlow and TCC *Particle Flow* (PFlow) [78] and *Track-Calorimeter* (TCC) [79] are alternative jet definitions that started being used in ATLAS in the later Run 2, which are named for completeness. Their shared philosophy is the necessity to use the tracks as a mean to improve the jet reconstruction efficiency. In the PFlow algorithm, the charged hadron's tracks are matched to EM topoclusters; these topoclusters are removed from the clustering. Topoclusters that are not matched to tracks are assumed to be generated by neutral hadrons. This way the superior angular resolution of the tracker system is used for the reconstruction of the charged particles, and the neutral ones are still measured with the calorimeters. The tracks and remaining topoclusters are then clustered together using anti- k_t algorithm of radius $R = 0.4$. Apart from the enhanced resolution, the use of tracks in PFlow jets

achieves improvements in the pile-up stability, making them the preferred alternative to standard Small- R jets at the end of Run 2.

The TCC jet algorithm targets the reconstruction of hadronically decaying Lorentz boosted objects. Differently from PFlow, which removes topoclusters to avoid double counting, the TCC algorithm uses the tracks to assess the jet structure. Therefore it exploits the different behaviour of the tracker and calorimeter system, the former has a superior angular resolution but a transverse momentum resolution that degrades with increasing p_T . The latter has instead a good energy resolution at increasing energies. This is achieved using the tracks to establish the position of the clusters when possible, but the calorimeter for the energy measurement. The final clustering uses the same receipt as in Large- R jets, with $R = 1.0$ and identical grooming settings.

Finally the PFlow and the TCC algorithm can be combined together in what is called *Unified Flow Object* (UFO), which is being explored as of the end of Run 2.

5.1.1. Large- R Jets and Substructure

The ATLAS physics program addresses a wide spectrum of kinematic regimes, requiring the development of tools to better enhance the precision and the sensitivity of the measurements. In the context of Dark Matter and SUSY searches, the final states are expected to have an hard recoil against invisible particles. If these final states consist of hadronically decaying objects, the possibility of reconstructing the decay products individually starts being limited by the angular separation between them. As a rule of thumb, a decaying object of mass m and transverse momentum p_T , has a ΔR between the decay products as:

$$\Delta R \sim \frac{2m}{p_T}$$

Jets originating from individual quarks are approximated to be massless⁴. When these jets approach each other at a distance smaller than the clustering radius (for Small- R this is $R = 0.4$), the final object can be mistakenly reconstructed with the mass of a quark/gluon in the extreme case. This issue can be appreciated in Figure 5.3. On the right, the angular separation between quarks decaying from a highly energetic W boson is shown to be below $\Delta R = 0.5$ for transverse momenta bigger than 300 GeV.

⁴The typical p_T is much higher than the masses of the light quarks.

This strongly limits the ability to reconstruct the quarks from the W decay individually. On the left, the angular separation between W bosons and b quarks in top decays. The high p_T kinematic regime in hadronic final states inherently offers an higher sensitivity than the low p_T regime, since the QCD background is exponentially falling with rising p_T . Therefore, the importance to develop methods and techniques to overcome this limitation in the reconstruction strategy is crucial. In order to fully exploit the possibilities offered by the highest than ever integrated luminosities and energies available at the LHC, Large- R jets are employed in ATLAS.

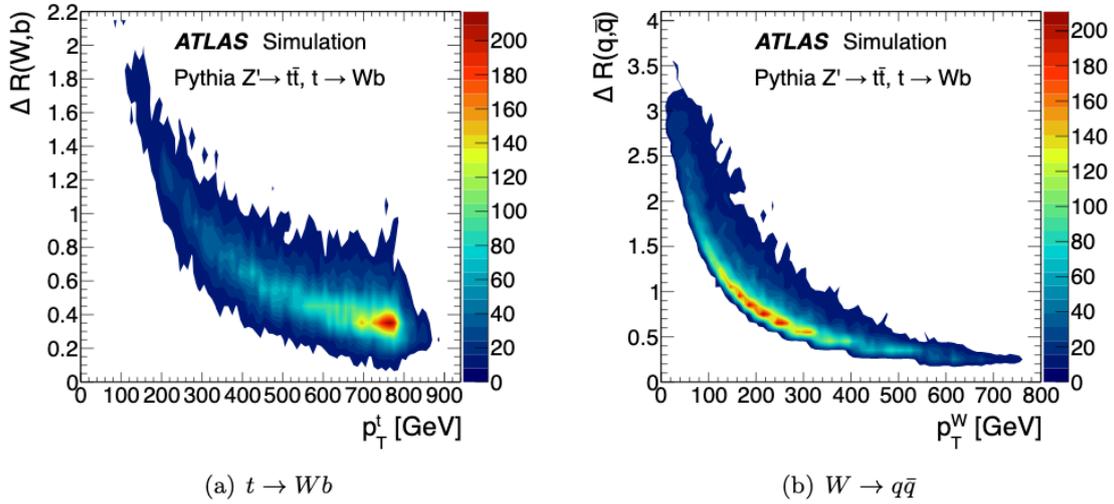


Figure 5.3.: Both pictures show the angular separation between their decay products as a function of the p_T of the decaying particle, in a simulated sample with energetic di-tops [80]. On the left (a) the ΔR between the W boson and the b quark from the top quark decay; the hyperbolic behaviour follows the $\sim 1/p_T$ law. While a Small- R jet is not sufficient to contain both W and b , a Large- R with $R = 1.0$ assures the containment for $p_T^{top} > 300$ GeV. On the right (b), the ΔR between the quarks from the W bosons decay. If $p_T^W > 300$ GeV, the ability to resolve the quarks individually with Small- R jets starts degrading quickly, but on the other hand a Large- R jet can contain both of them.

A Large- R jet of radius $R = 1.0$ has more than six times a bigger area than a Small- R jet, meaning that a jet so large is much more likely to contain soft radiation from the high pile-up environment of an LHC event in Run 2. This additional radiation could critically spoil the mass measurement, which is based on the addition of the 4-momenta of the topoclusters constituting the Large- R jet. A relatively low p_T constituent, but placed at larger distance could significantly rise the value of the measurement. Moreover, a larger area is more likely to acquire an higher number of topoclusters, building up the value of the mass also in the case of pure pile-up jets.

In order to mitigate this issue, *substructure* techniques can be applied to the Large- R jet. The decay of an object such as a W boson can be distinguished from pile-up jets: the former shows a structure with two hard components compatible with a two-body decay $W \rightarrow q\bar{q}'$. The latter coming from soft QCD radiation typically shows instead only a single core with diffuse spread.

Grooming Techniques

The algorithms dealing with the removal of the soft QCD radiation from the Large- R jets are called *grooming* algorithms. They are designed to retain the original hard structure from the decay of energetic objects, while discarding the components coming from pile-up and underlying event. Enhancements in the resolution of the mass measurement can in this way be achieved. Additionally, a jet that has been groomed can be easier distinguished from a background one by means of substructure variables, which are briefly described here.

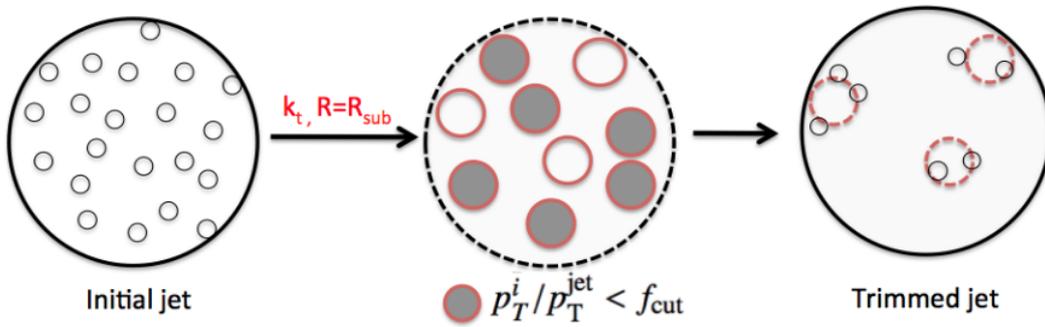


Figure 5.4.: The flow of the trimming algorithm [80].

There are three main grooming techniques used in ATLAS: *Mass-drop filtering*, *trimming* and *pruning*. Their common goal is to find and remove energy deposits that are at a much lower scale and further away from the hard components. The trimming algorithm has found a wider application. It is a two step algorithm with two parameters shown in Figure 5.4. First, the original jet's constituents are clustered again with the k_t algorithm, but with a much smaller radius R_{sub} , into *subjets* i . In order to remove the soft components, the p_T of each of the subjets is compared to the p_T of the original jet. If the ratio the two is smaller than a threshold parameter, $p_T^i / p_T^{jet} < f_{cut}$ the subjet is discarded, otherwise it is kept. The two parameter, R_{sub} and f_{cut} are set in ATLAS to 0.2 for the clustering radius and 0.05 (5%) for the threshold, respectively.

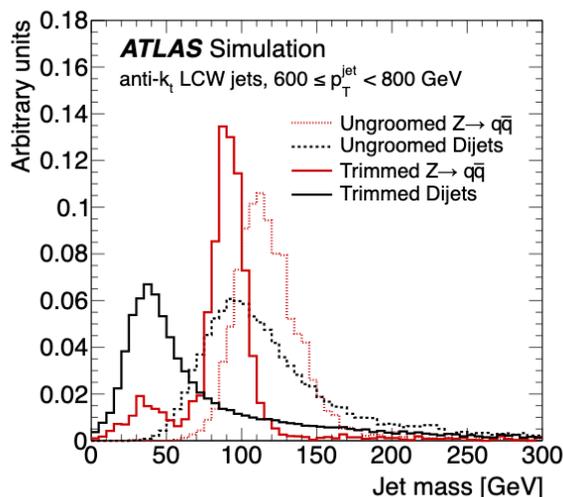


Figure 5.5.: The mass distribution of Large- R jets in high p_T regimes ($600 \text{ GeV} < p_T < 800 \text{ GeV}$) before trimming (dashed) and after trimming (continuous line), from boosted Z (red) and QCD di-jet background (black) [80]. The trimmed jets from boosted Z lose a small fraction of their mass, but they see an increase in the resolution. Trimming removes the majority of the energy from the QCD jet, achieving a strong suppression of the background at the higher end of the mass spectrum.

The advantage of the use of trimming in Large- R jets can be seen e.g. in Figure 5.5. The hard structures that are present in the jets and originated from boosted objects are mostly untouched. Instead, the majority of the components of a QCD di-jet⁵ background are instead removed, bringing down the total mass of the jet. In this way a simple mass cut can get rid of a large fraction of the background jets. The mass is not the only discriminant that can be used: other observables can be investigated.

Substructure Observables

QCD jets that survive the grooming, retaining masses of the order of the Lorentz boosted objects looked for, pose an obstacle in achieving high sensitivities. In fact, cross-sections of the QCD processes at LHC are overwhelming with respect to the ones from the signals and can introduce large amount of background contamination even for large masses. A further rejection can be achieved exploiting even more the peculiarity of two and three-prong decays. This can be qualitatively seen in the η - ϕ space in Figure 5.6. The constituents of a QCD jet (left), a W jet (centre) and a top jet

⁵*Di-jet* means 'a pair of QCD hard jets recoiling against each other'. *Multijet* more generally refers to a number of jets greater than two.

(right) are shown. Boosted objects such as the W and the top quark tend to present harder structures that are well separated. QCD jets present mainly collinear radiation that is close to the core of the jet.

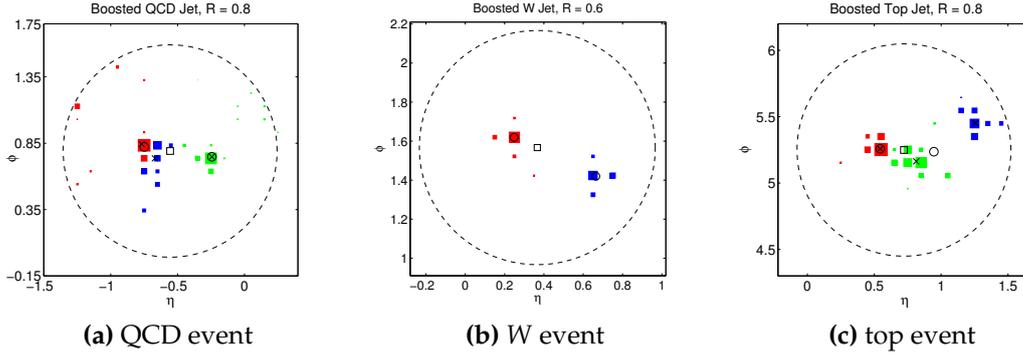


Figure 5.6.: $\eta - \phi$ map of Large- R jets constituents originated from QCD (a), W bosons (b) and Top quarks decays (c) [81]. Constituents of the same subjet are indicated by the same colour, the open square indicates the total jet direction, the open circles indicate the two subjet directions, and the crosses indicate the three-subject directions.

This difference in the angular and transverse momenta distributions can be exploited by means of the so-called *substructure observables*. These observables are built in a way to put emphasis on the structure of the jet, which is different in boosted objects with respect to QCD events. There are many observables that have been defined and are used in HEP at ATLAS and CMS. For a more complete overview the reader can refer to [82] [83] that also contain details on state-of-the-art machine learning based algorithms. For the purposes of this thesis, two classes of observables are shown: the N -Subjettiness and *Energy Correlation Functions*. They offer high signal tagging capabilities and background rejection at a reasonable computational complexity. The N -Subjettiness, τ_N is calculated via [81]:

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min(\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k})$$

where k runs over the constituents⁶ in a given Large- R jet, $p_{T,k}$ is their transverse momenta and $\Delta R_{i,k}$ the usual angular distance between the constituents and the k_t clustered subjets. The normalization factor d_0 is taken as:

⁶Constituents can be topoclusters, tracks etc.

$$d_0 = \sum_k p_{T,k} R_0$$

where R_0 is the clustering radius of the Large- R jet. Smaller values of the N-Subjettiness indicate a vicinity of the constituents to their respective subjets, indicating a compatibility with a number of N subjets. Larger values favour instead a compatibility with $N + 1$ subjets. It is convenient to use ratios of N-Subjettiness observables for the most common scenarios of two to four prong decays, such as W , top quarks or pairs of Lorentz boosted W bosons. They are indicated as $\tau_{21} = \tau_2/\tau_1$, $\tau_{32} = \tau_3/\tau_2$ and $\tau_{42} = \tau_4/\tau_2$ (sometimes also $\tau_{43} = \tau_4/\tau_3$).

Energy Correlation Functions (ECF) are defined [84] without the need of using k_t clustered subjets, evaluating directly the constituents. The generalized ECF functions are written as:

$$\text{ECF}(N, \beta) = \sum_{i_1 < i_2 < \dots < i_n} \left(\prod_{a=1}^N p_{T_{i_a}} \right) \left(\prod_{b=1}^{N-1} \prod_{c=b+1}^N \Delta R_{i_b i_c} \right)^\beta$$

where the sum runs over all the constituents of the jet, and the relative importance of the angular distances with respect to the transverse momenta is modified by the parameter β . The angular distance is calculated with all the pairwise combinations of the constituents. More explicitly in the case of interest of two and three prong decays:

$$\text{ECF}(2, \beta) = \sum_{i < j} p_{T_i} p_{T_j} (\Delta R_{i,j})^\beta$$

$$\text{ECF}(3, \beta) = \sum_{i < j < k} p_{T_i} p_{T_j} p_{T_k} (\Delta R_{i,j} \Delta R_{i,k} \Delta R_{k,j})^\beta$$

and $\text{ECF}(1, \beta)$ is just the p_T normalization⁷. For harder constituents compatible with two or three prongs structures, the ECF tends to have higher values with respect to jets with soft, diffuse radiation such in the case of QCD jets. Similarly as in the case of the N-Subjettiness, it is convenient to define normalized ratios of the ECF, to achieve an even higher separation power of these observables. Two ratios are used the most, C_2 and D_2 . They are define as:

⁷Sometimes the already normalized version of the ECF are used: $e_2^\beta = 1/p_T^2 \text{ECF}(2, \beta)$, $e_3^\beta = 1/p_T^3 \text{ECF}(3, \beta)$ etc.

$$C_2 = \frac{ECF(3, \beta = 1)ECF(1, \beta = 1)}{ECF(2, \beta = 1)^2}$$

$$D_2 = \frac{ECF(3, \beta = 1)ECF(1, \beta = 1)^3}{ECF(2, \beta = 1)^3}$$

The choice $\beta = 1$ is common in the ATLAS collaboration, but different values can be used.

Substructure and mass variables are an extremely useful tool to probe hadronic final states originated from energetic, boosted objects. They can be used with simple cuts to select the substructure of the jet, and the mass of the object of interest. Additionally, more complex taggers can be built with these variables, based e.g. on machine learning techniques such as Deep Neural Networks.

5.2. TAR Definition, Performance and Potential

The angular resolution of the calorimeter is limited by the geometrical size of the calorimeters cells; however its energy resolution increases with increasing momentum of the incoming particles. In final states with highly energetic hadronically decaying objects, the limitation in the angular resolution starts degrading the performance of the reconstruction of these objects. As seen, other methods have been studied and developed, such as the PFlow and TCC, to enhance the reconstruction capabilities of Large- R jets using the associated tracks. The angular resolution of the tracker system is higher than that of the calorimeters, and can be used to overcome this limitation.

5.2.1. Track-Assistance

One simple technique that can improve the jet reconstruction using both tracker and calorimeter measurement is the *Track-Assistance*, which was firstly proposed in the context of the top tagger HPTTOPTAGGER [70]. The Track-Assistance procedure makes use of the tracks to ‘help’ or ‘assist’ the measurement performed by the calorimeters. In a typical proton-proton collision, about 65% of the energy is carried by charged hadrons, while about 25% by photons from π_0 decays and only around 10% by neutral hadrons (mostly neutrons and K_L^0) ([70] and references therein). The prospect of

reconstruction based on tracks alone is therefore not viable or competitive with respect to calorimeter-based measurement. This is due to the absence of tracks associated to photons and neutral hadrons. The procedure for correcting (or rescaling) calorimeter-based observables was outlined using hadronic calorimeters to correct electromagnetic-only measurements [85] [86] [87], employing ratios of quantities measuring the charged component (tracks) and charged plus neutral component (calorimeter) such as:

$$\alpha = \frac{E_{jet}}{E_{track}}$$

The ratio α is a proxy to access the charged to neutral ratio.

One of the most important calorimeter observables is the jet mass. In Large- R jets, J , the calorimeter mass (m^{calo}) is measured adding the massless 4-vectors of the constituent topoclusters i as:

$$m^{calo} = \sqrt{\left(\sum_{i \in J} E_i\right)^2 - \left(\sum_{i \in J} \vec{p}_i\right)^2}$$

As noted before, in case of highly energetic hadronic decays, the collimation of the particles can be comparable with the calorimeter granularity. The *Track-Assisted Mass* [85] (m^{TA}) applies the ratios α used in the Track-Assistance to correct a measurement performed by the tracker alone, m^{track} , compensating for the missing neutral component. It is defined as:

$$m^{TA} = \frac{p_T^{calo}}{p_T^{track}} \times m^{track}$$

where again the superscript *calo* or *track* indicates a measurement from calorimeter or tracker, respectively, and the p_T ratio can be identified with the ratios α .

Within ATLAS, the Track-Assisted mass is demonstrated to achieve better performance than the calorimeter mass at high p_T of the decaying W/Z (also called W/Z jet). This is expected since at low p_T , the decay products are still well separated, and the calorimeter only measurement can achieve good performances. To obtain a better resolution for the mass of the Large- R jets, the m^{calo} and m^{TA} are combined into the combined mass m^{comb} [85]. The combination can be chosen to minimize the resolution of the combined mass response. m^{comb} is used in ATLAS as standard mass definition

for Large- R jets. A comparison of the performance for the three observables is shown in Figure 5.7.

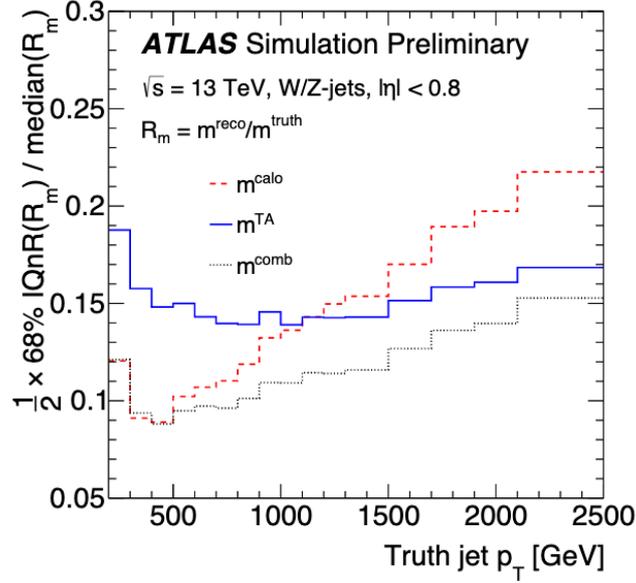


Figure 5.7.: The performance of the reconstruction of W/Z jets using the m^{TA} , m^{calo} and their combination m^{comb} (the standard in ATLAS) is compared. On x-axis, the p_T of the MC truth jet and on y-axis the interquartile range (IQnR) divided by twice the median of the response distribution is shown. The IQnR is a robust way to access the resolution. It can be seen that the Track-Assisted mass achieves better resolutions at higher p_T 's and that the combination m^{comb} achieves an even superior performance. From [85].

Another similar way to perform the Track-Assistance procedure, but with the benefit of an higher freedom in the calculation of Track-Assisted observables, is to apply it on the p_T of the individual tracks. A track can in fact be corrected for the overall missing neutral component of the jet as:

$$p_T^{track,new} = p_T^{track,old} \times \frac{p_T^{calo}}{\sum_{i \in J} p_{T,i}^{track,old}} \quad (5.1)$$

where the index i runs over all the tracks associated to the jet and *old*, *new* refers to the track before and after the Track-Assistance procedure [88]. It should be noted here that this prescription does not affect the η - ϕ position of the tracks, but only their transverse momenta. Again, the fraction represents the α of above. These tracks can be used, e.g. summing their four momenta, to calculate the Track-Assisted jet mass or substructure observables.

One important characteristic of the missed neutral component is that its fractional contribution fluctuates separately in the parton showering of each of the quarks from the hadronic decays of the objects. With m^{TA} (therefore with m^{comb}) these fluctuations are ignored and an average is used for the entire Large- R jet. If the Track-Assisted α to correct for the neutral component for each of the individual parton-showered quarks could be produced, higher performances can be achieved. This possibility was firstly explored with the *Track-Assisted-Subjet* (TAS) mass [85], which brought forward the idea of a correction for the individual subjets of the Large- R jet. The TAR algorithm is a generalization of the TAS procedure, which gains independence from the underlying definition of the Large- R jet and, at the same time, has the freedom of choosing the Large- R jet radius that is optimal to the final state. This is achieved via the jet re-clustering, which is the reclustering of smaller jets into bigger ones.

5.2.2. Jet Re-clustering

Jet re-clustering (or reclustering) makes use of smaller jets to build Large- R jets [89]. The smaller jets can be Small- R jets or R -scan jets $R = 0.2$, and they are clustered together (with clustering algorithms such as anti- k_t or k_t) with a radius R bigger than their own radii. The main advantage of reclustered jets is that they can address scenarios where the standard radius $R = 1.0$ is not optimal. Moreover, no additional calibration is needed as they inherit that of their smaller constituents. Their performance is found to be similar to that of the Large- R jets. As shown in the next sub-sections, the jet reclustering allows the application of an improved version of the Track-Assistance for Large- R jets. The neutral component is corrected for in each individual small jets (Small- R jets or R -scan jets $R = 0.2$), which the reclustered jet is built of.

5.2.3. TAR Definition

The Track-Assisted Reclustered jet algorithm (TAR) [88] makes use of the Track-Assistance procedure to the tracks associated to a Large- R jet reclustered from R -scan $R = 0.2$ or Small- R jets.

In the following Section, the small jet used is the R -scan $R = 0.2$ jets, since this is the TAR original configuration explored in [88]. The configuration used in the analysis described in Part III uses instead Small- R jets.

The correction for the missed neutral component is performed for each of the individual smaller jets and it is applied to each track individually, according to the second receipt illustrated (Equation 5.1). It can be rewritten as:

$$p_T^{track,new} = p_T^{track,old} \times \frac{p_{T,j}^{R=0.2jet}}{\sum_{i \in j} p_{T,i}^{track,old}} \quad (5.2)$$

where this time j is the $R = 0.2$ jet. The calibrated smaller $R = 0.2$ jets with $p_T > 20$ GeV and $|\eta| < 2.5$ are clustered with anti- k_t and then trimmed using $f_{cut} = 0.05$, which improves the resilience to pile-up and underlying event radiation. Tracks are matched to jets that survive the trimming. They are required to fulfil quality criteria and be well-reconstructed⁸, and to be associated to the primary vertex⁹. Only tracks with $p_T > 500$ MeV and $|\eta| < 2.5$ are used. Unmatched tracks are still associated with the jets if they fall within a $\Delta R < 0.3$ of the jet axis in case of $R = 0.2$ jets ($\Delta R < 0.5$ for Small- R jets). After the track-to-jet matching procedure, the p_T of the tracks is rescaled according to the formula 5.2, therefore depending on the jets they are associated to.

The Large- R jet is built from the smaller jets with rescaled associated tracks and is referred to as TAR jet. The rescaled tracks are then used to calculate the mass m^{TAR} and all the other substructure observables. The TAR algorithm is also depicted in Figure 5.8.

With respect to standard methods, this procedure allows an extended freedom in the choice of the Large- R jet reclustering radius or grooming procedure, and provides a way to use the superior track-based substructure. Because of the reclustering step, no additional calibration is needed for the Large- R jet, as the constituent jets are already calibrated. The uncertainties are propagated *bottom-up*, from the tracks and smaller jets to the final observables. Using the tracks to access the substructure of the Large- R jets, the TAR procedure can also be utilized with samples produced with ATLAS Fast II (AFII) simulation [90] [91], which makes use of parametrized showers development in the calorimeters, but a full track simulation. This offers the opportunity to generate

⁸At least seven hits in the pixel and SCT detectors, not more than one module allowed to be shared between multiple tracks in the pixel or SCT detector. Less than three holes are requested per track, and not more than one in the pixel detector.

⁹The primary vertex is selected as the vertex with the highest scalar p_T sum of tracks associated with it using transverse and longitudinal impact parameter requirements.

around ten times more events with respect to standard full simulation (FullSim), with the same computing resources, greatly improving the available MC statistics.

5.2.4. TAR Performance

The TAR is observed to enhance the mass resolution in regions of the phase-space that are relevant to the identification of hadronically decaying objects such as W/Z , Higgs bosons and top quarks. This can be seen e.g. in Figure 5.9, where the mass and D_2 resolutions are compared against the standard m^{comb} and cluster based D_2 . The mass in particular is shown to improve the resolution of hadronically decaying W boson in the regime $p_T < 1$ TeV. Similar conclusions can be drawn for the reconstruction of other heavy objects, such as top quarks and Higgs bosons (also in the W^*W final state).

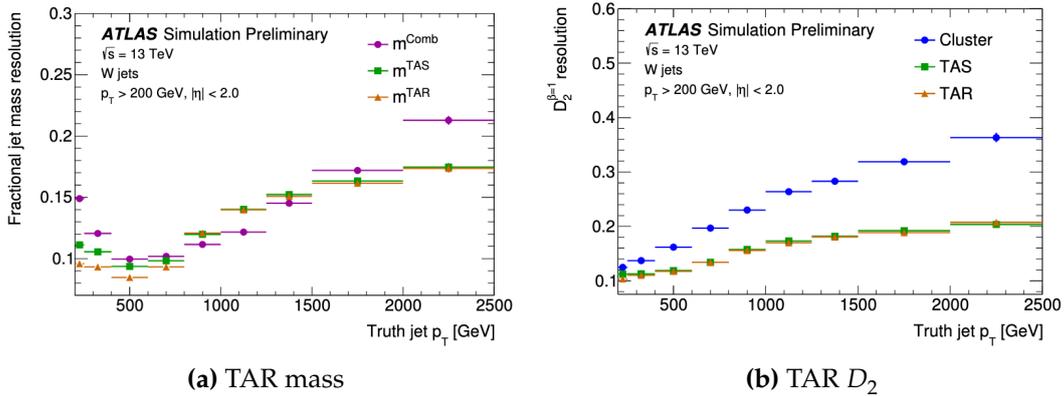


Figure 5.9.: The resolution of the mass observable for W jets (left) and for the $D_2^{\beta=1}$ substructure observable (right) as a function of the MC truth p_T . The TAR algorithm is shown in beige. The m^{comb} is shown on the left in magenta and an alternative (and older) definition of TAR, TAS (Track-Assisted Subjet) is shown in green. On the right, the substructure observable D_2 for TAR and TAS is compared against calorimeter-based substructure only, since there exists no combined version as for the mass. The Figure is based on the author's work in [88].

Simple cut-based taggers can be constructed with mass and substructure variables. These can be considered as a proxy for arbitrarily complex ones such as Deep Neural Networks (DNN) or Boosted Decision Trees (BDT). DNN or BDTs constructed on TAR variables are expected to maintain the improvements with respect to DNN or BDT based on cluster-based observables and standard techniques. The simple taggers show improvements in the identification of Lorentz boosted objects in all the scenarios considered. The *Receiver Operating Characteristic* (ROC) curve is typically used to show

the efficiency of the selection of the signal and the rejection of the background. An example of a ROC in the case relevant to the Dark Higgs model decaying to a pair of W bosons is shown for the case $H \rightarrow W^*W$ with the SM Higgs boson in Figure 5.10. The performance of TAR exhibits an enhancement of a factor of two or more (at fixed signal efficiency) in the background rejection with respect to cluster-based observables.

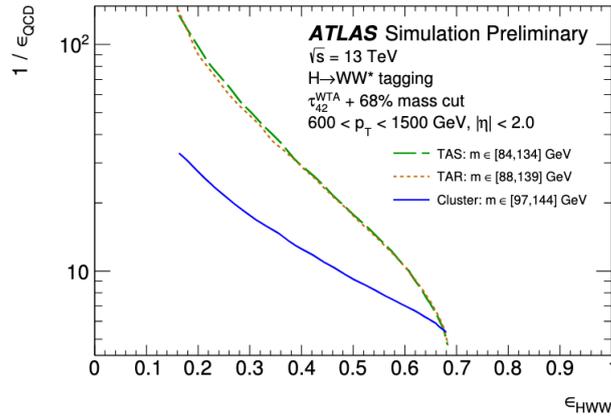


Figure 5.10.: ROC curves comparing the separation power of a 68% efficient mass cut plus a substructure observable for the identification of four-prong decays, τ_{42} , against QCD jets for $250 \text{ GeV} < p_T^{jet} < 600 \text{ GeV}$. On the x-axis the efficiency of the identification of the pair of W from Higgs decay, and on the y-axis the efficiency of the background rejection. Again TAR (TAS) in beige (green) outperforms standard techniques, in blue. The Figure is based on the author’s work in [88].

The full list of figures and tables for the performance of TAR in all the signal scenarios can be found in Reference [88].

5.3. Jet Calibration and Uncertainties

The TAR jet technique requires the usage of smaller jets with a full set of calibrations and systematic uncertainties. If Small- R jets are used as input, their calibration and systematic uncertainties are fully available and centrally provided to the ATLAS analyses. However, for R -scan $R = 0.2$ jets, these are not yet available and have to be derived separately. Since jets of smaller radii can better probe the individual fluctuations of the charge to neutral ratios of the quarks/gluons, R -scan $R = 0.2$ jets provide the best choice, therefore justifying the effort of producing the calibrations and the uncertainties. The TAR jets as used in the analysis illustrated in Part III, are constructed with Small- R jets as a way to explore this novel approach, yet for

future analyses looking for challenging final states, R -scan $R = 0.2$ is expected to be used instead. Calibration and systematic uncertainties are then a crucial aspect of the development of the TAR jets, in this Section an overview of what are the steps in ATLAS for deriving these important ingredients is presented, for both the Small- R and R -scan $R = 0.2$ jets.

Jet Energy Scale

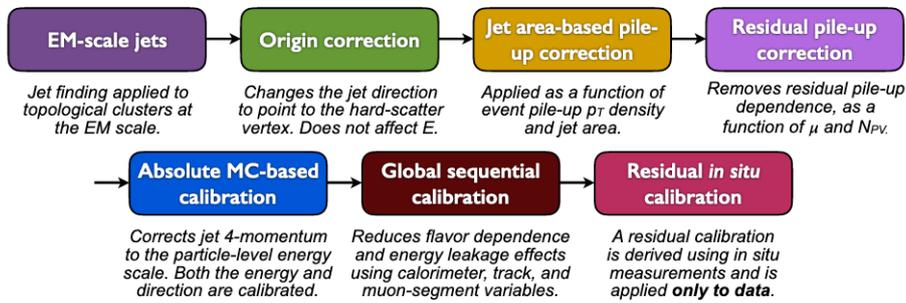


Figure 5.11.: The calibration sequence for jets at the EM scale [92].

The jet energy scale (JES) calibration consists of a multiple-step process that restores the energy scale to that of truth MC jets reconstructed at particle level. In Figure 5.11, the ATLAS calibration sequence applied to Small- R EM jets [92] is shown. After they are built (first step), the jet direction is corrected to point to the primary vertex (second step, Origin Correction). In order to reduce the impact of the pile-up on the jet, a correction is applied as a function of the pile-up p_T density and the jet area. Another step is taken to account for the number of primary vertexes and the average number of visible interactions per bunch crossing μ (third and fourth step, Jet area-based pile-up correction and Residual pile-up correction). The energy scale of the jet is derived as a function of the pseudorapidity and the transverse momentum (fifth step, Absolute MC-based calibration). The so-called Global Sequential Calibration, makes use of variables that access the dependence of the jet response to flavour composition, and leakage effects (sixth step). Up to this stage, all the corrections were derived in Monte Carlo simulations. An example of the Absolute MC-based calibration and the Global Sequential Calibration is shown in Figure 5.12, the former as a function of η , and the latter as a function of the number of tracks associated to the jet. Finally, an in-situ measurement in data is performed to determine the absolute energy scale of the jet by balancing it against a well measured object (e.g. Z boson decaying to leptons, or a

γ , or another jet). From this measurement, the last correction factor (Residual in-situ calibration) is determined and applied to data only.

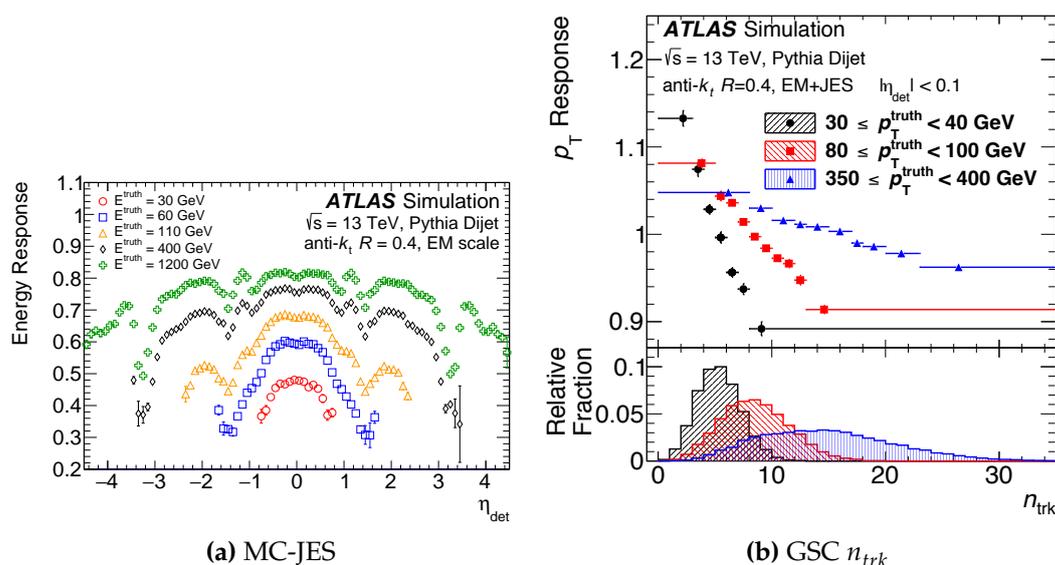


Figure 5.12.: An example of the absolute MC-based calibration on the left, showing the jet energy response as a function of the jet η for different values of the jet p_T in different markers, before the correction is applied. On the right, a stage of the Global Sequential Calibration using the number of tracks n_{trk} (x-axis) and the response (y-axis), for three different p_T bins in different markers, before the correction is applied. The bottom plot shows the relative fraction per p_T bin [92].

The final in-situ calibration is necessary to ensure the comparability of data to simulated samples. It is performed exploiting a series of processes: first the di-jet η -intercalibration, where a jet in a well instrumented region of the calorimeter ($|\eta| < 0.8$) is used as reference against jets in the more forward regions. The $Z + \text{jets}$, $\gamma + \text{jets}$ and multijet balance is used to cover the phase-space available in transverse momentum (from around 20 GeV to around 2 TeV).

Jet Energy Scale Uncertainty

The JES uncertainty is computed from the around 80 different sources of systematic uncertainty that affect the energy scale, propagating from the individual calibration effects. The majority of them are related to the in-situ measurements of the $Z/\gamma + \text{jets}$ and multijet balance. The remaining are derived from pile-up uncertainties, physics mis-modelling, differences in the jet response depending on the jet's flavour (light quarks, gluons and b-quark initiated jets) and statistical uncertainty. A relevant source

of uncertainty is the out-of-cone (OOC), which quantifies the energy loss outside the cone of the jet and therefore affects the in-situ measurements.

In Figure 5.13, the overview of the propagation of the uncertainties from e.g. γ +jets in-situ balance to the final JES uncertainty that embodies all the different effects (right) is depicted.

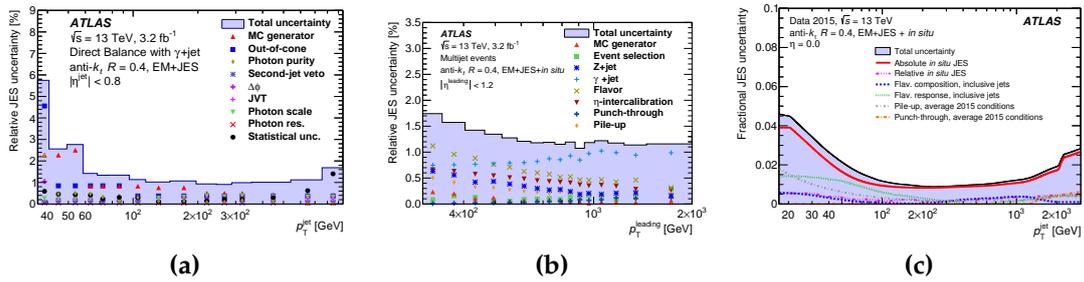


Figure 5.13.: Overview of the composition of the JES systematic uncertainty in early Run 2, as a function of p_T . On the left, the total JES uncertainty as calculated from the γ +jets balance, where the main components are the MC generator, which accounts for potential mis-modelling, and out-of-cone effects. In the central plot it is shown the overview of the collective impact of the in-situ (γ +jets measurement depicted with cyan pluses, which dominates at higher p_T 's), pile-up and flavour uncertainties, which dominates at the low end of the spectrum. Finally, on the right, the plot showing the combined effect of the in-situ uncertainties, together with the flavour composition and response, pile-up uncertainties and punch-through [92].

R-scan Jets For the jet of radius $R=0.2$ or $R=0.6$ and built on LCTopoClusters (the R -scan jets), the derivation of the calibration is done separately for the MC correction, which in great part follows the procedure applied to Small- R jets. The in-situ calibration uses Small- R jets as reference, and benefits a partial inheritance of the systematic uncertainties calculated for them. However, additional components (e.g. the out-of-cone or the flavour response) had to be specifically re-derived. The author contributed to the R -scan effort in the MC jet energy scale, Global Sequential Calibration, the flavour uncertainty and the out-of-cone uncertainty.

Jet Energy Resolution and Uncertainties

The jet energy resolution (JER) [93] refers to the overall fluctuation that a measurement of the jet energy at fixed energy experiences. The JER, expressed as σ_{p_T}/p_T can be parametrized in terms of a noise factor (N), which represents the effect of pile-up and

electronic noise of the calorimeters, a stochastic term (S), which expresses the sampling properties of the ATLAS calorimeters, and constant term (C), which accounts for the passive materials present within the detector. The transverse momentum dependence is expressed via the formula:

$$\frac{\sigma_{p_T}}{p_T} = \frac{N}{p_T} \times \frac{S}{\sqrt{p_T}} \times C$$

The JER is evaluated in a in-situ measurement as well. The method used is a balance of a well measured *reference* object, e.g. a jet in a well instrumented central region such as $0.2 < |\eta| < 0.7$ against another *probe* jet (this is called *di-jet direct balance* measurement). For this procedure, the asymmetry $A = \frac{p_T^{ref} - p_T^{probe}}{p_T^{average}}$ is compared in data and simulations at particle and reconstruction level. The JER extracted from A is validated against the Monte Carlo responses, and their difference, the *non-closure*, is used as a systematic uncertainty. Additionally, the data/MC difference in the JER is taken care of applying a smearing to MC jets or with an ulterior uncertainty.

At low p_T , where the JER extraction from A breaks down due to the dominant contribution of the noise term contribution, additional constraints are needed, and are evaluated with a separate measurement, the so-called *random cones* method. The energies contained in a pair random and not overlapping cones are summed (so to get p_T^{cone1} and p_T^{cone2}) in random events with filled bunches (*zero-bias*). In this way the noise can be extracted from the width of the momentum difference $p_T^{cone1} - p_T^{cone2}$, taking into account any non Gaussian behaviour.

The most important components of the JER systematic uncertainty are shown in Figure 5.14. The major sources can be seen to originate from the random cone method for the constraints on the noise term and from the systematics of the di-jet direct balance measurement. The former are more relevant at low p_T , and the latter, dominated by the non-closure uncertainty, are relevant from medium to high p_T .

R-scan $R = 0.2$ Jets In R -scan $R = 0.2$ jets the JER is expected to be similar to that of the Small- R jets: in fact typically the core of $R = 0.1$ of the Small- R jets retains the vast majority of the jet's energy and therefore is shared with the R -scan $R = 0.2$ jets. However, differences in the behaviour as a function of η of the JER associated uncertainties require special care, and in general the derivation of the systematic components has to be repeated. The author has performed the evaluation of the JER

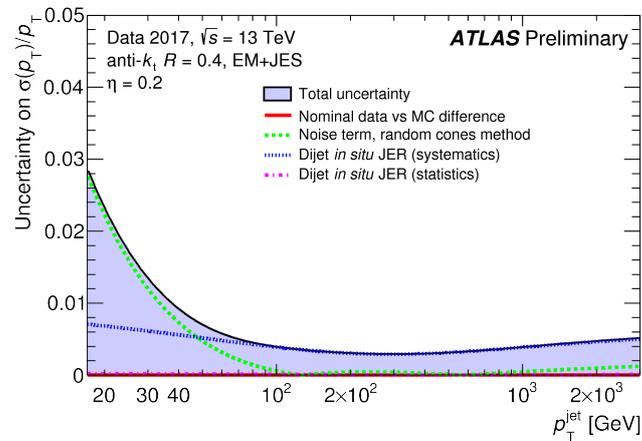


Figure 5.14.: The breakdown of the systematic uncertainties components on the fractional jet energy resolution as a function of p_T [94]. The black line indicates the total uncertainty. The green line shows the impact of the noise term component, which is dominant at the lower end of the spectrum. The blue component is derived from the in-situ di-jet direct balance method.

for R -scan $R = 0.2$ jets with the asymmetry method in a dedicated in-situ direct balance measurement. He extracted the non-closure systematic uncertainty, and compared the total JER and systematic uncertainty between R -scan and Small- R jets [95].

Part III.

**Search for Dark Matter produced in
association with a Dark Higgs boson
decaying to VV in the fully hadronic
final state**

6. Motivation and Overview

The last Part of this thesis describes a search for Dark Matter produced in association with an hypothetical Dark Higgs decaying to VV bosons in the fully hadronic final state using 139 fb^{-1} of data collected during Run 2 at the LHC. It is also called $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ analysis for brevity. The Dark Higgs model has already been introduced in Section 3.2. It provides a mechanism for the generation of the masses in the Dark Sector, via a scalar particle that is lighter than Dark Matter. In this case, the strong constraints placed by the LHC on the existence of new mediators can be relaxed. The point of contact with the SM is a new gauge boson Z' , which couples to quarks. The branching ratio of the Dark Higgs to other SM states is similar to the branching ratio of the SM Higgs as a function of its mass. This is a consequence of the mixing between the two bosons in this scenario. At low masses ($m_s < 160 \text{ GeV}$) the main decay channel is $s \rightarrow b\bar{b}$. However, as soon as kinematically allowed ($m_s > 160 \text{ GeV}$), the W^+W^- production becomes dominant, as shown already in Figure 3.7. The reconstruction of the W^+W^- system has the advantage of being able to exploit its resonance at the value of the Dark Higgs mass, allowing the search to focus on excesses in the invariant mass spectrum. The $E_T^{\text{miss}} + s \rightarrow ZZ(\text{had})$ final state is experimentally indistinguishable from the $E_T^{\text{miss}} + s \rightarrow W^+W^-(\text{had})$, therefore the search for the Dark Higgs is performed in the combined $s \rightarrow VV$ decay channel. The vector boson pair can be fully reconstructed if it decays hadronically, benefiting from a large branching fraction of around 45% for W^+W^- . Another advantage is the clear event topology, where the missing transverse momentum originates only from the Dark Matter particles and it is not altered by the presence of neutrinos, as in semi-leptonic or fully leptonic channels. The final state comprised of a pair of resonant vector bosons decaying hadronically and large quantities of momentum imbalance is at the time of writing a new signature, unexplored with the ATLAS detector. The Dark Higgs model can be considered as well a 'signature generator', a way to optimize the search in the $E_T^{\text{miss}} +$ resonant $VV(\text{had})$ state and investigate this unknown territory, regardless of the underlying theoretical model. It is in fact of utmost relevance to experimentally scrutinize these processes before proceeding with Run 3. This is also described by the motto 'turn every stone': nothing must be left unattempted searching for Dark Matter and Beyond the Standard Model physics at the LHC.

The identification and reconstruction of the pair of hadronically decaying vector bosons is addressed in the analysis with the TAR jet technique (defined in Section 5.2). TAR jets are employed in the most sensitive regions of the phase-space. Moreover, TAR based jet substructure observables are used to fully exploit the shape and peculiarities of this hadronic signature with strong Lorentz boosts. As the Dark Matter particles escape the detector undetected in the signal hypothesis, they generate large quantities of missing transverse momentum E_T^{miss} (introduced more precisely later in the Chapter 8). The E_T^{miss} is correlated with the Lorentz boost that the Dark Higgs s experiences. The boost originates from the recoil with the Dark Matter particles; this topologies are also called *back-to-back*. As already seen in the previous Part of this thesis, the collimation of the decay products as a function the mass of the decaying particle and its transverse momentum is approximately given by the relation $\Delta R \sim 2m/p_T$.

This dependence creates different kinematic scenarios that are taken advantage of in this analysis, to fully exploit the physics potential of the ATLAS detector. The events are divided in categories according to the Lorentz boost of the candidate Dark Higgs. The categories are called *merged* and *intermediate*, and they are briefly described here and in more details in the next Chapters. All events are assigned to the categories based on a *priority selection*: first, events are checked for the merged category, and then for the intermediate, making sure that the highest possible sensitivity is achieved.

The merged category addresses the case where the transverse momentum of the Dark Higgs is so high that the spread of the VV pair is below the typical radius of Large- R jets, $\Delta R = 1$. This category represents the one with the highest sensitivity, since it is associated with large quantities of momentum imbalance that is unlikely to happen in background processes. The pair of V bosons is confined into a small area in the $\eta - \phi$ space, challenging the angular resolution of the calorimeters and of the standard reconstruction techniques. In order to achieve an even better signal identification and background rejection in this kinematic regime, TAR jets are employed. The power of TAR-based jet substructure assures a further distinction between background-like and signal-like jets. The former typically present an energetic core surrounded by soft radiation. The latter instead present four energetic components, roughly at the same momentum scale. The N-Subjettiness observables τ_{42} and τ_{43} are a well understood tool to exploit the number of hard prongs to increase the sensitivity. As explained in the previous Part, these observables are constructed upon the tracks with rescaled momenta associated with the TAR jet.

The intermediate category targets instead the case with moderate transverse momentum of the Dark Higgs. Here, the V with the highest momentum will be energetic enough to still be contained within a Large- R jet. The quarks from the hadronic decay of the sub-leading V in p_T will be instead separated enough to acquire larger distances from each other and with a Large- R jet from the leading V in p_T . To address this complexity, a simple algorithm was studied, exploiting the individual reconstruction of the leading V with a TAR jet, required to be consistent with the mass of the boson. Additionally, Small- R jets are selected within a certain distance from the TAR jet, and added to the mass of the TAR jet if the invariant mass of the pair is compatible with mass of the second boson.

The analysis is divided into a signal region (SR) and control regions (CRs) by vetoing or requiring leptons. There is one SR where leptons are vetoed (0LSR) and two control regions with one or two leptons (1LCR and 2LCR). These two CRs are used to assess and constrain the impact of the two main background processes that affect the analysis: W +jets and Z +jets. They can enter the SR in case of leptonic decay of the W , where one lepton is misidentified (and the neutrino undetected), and in the case of the Z , when it decays invisibly to neutrinos. Other background processes in order of relevance are: the $t\bar{t}$, single-top, di-boson process and multijet production.

The statistical evaluation of the analysis is done with a profile likelihood fit, performed with events selected as described below. The Dark Higgs boson would manifest itself as an increase of events in the mass spectrum at m_s . The signal strength is extracted through a fit to the invariant mass distribution of the candidate Dark Higgs in the signal region. The relative yield of the background processes is simultaneously extracted via the control regions. In order to maintain the simplicity of the statistical treatment, the mass spectrum in the control regions is not considered, but only its relative normalization.

7. Data, Triggers and Simulated Samples

This Chapter describes the details of the dataset used, the related triggers and the simulated samples. All the samples are processed with the XAMPP framework, which uses DxAOD samples in the version EXOT27 in the Athena software version 21.2.

7.1. Data

The analysis uses the entire dataset collected in Run 2. The runs are required to be taken while the detector was operational in all its subcomponents and the beam conditions provided by the LHC were optimal. This type of runs are called *good runs*. The runs are checked for defects both online during data taking, and offline during a first reprocessing. The discarded runs are well below 5% of the total. The individual runs are selected based on the good-run lists (GRLs):

- **data 2015:** `data15_13TeV.periodAllYear_DetStatus-v89-pro21-02_Unknown_PHYS_StandardGRL_All_Good_25ns.xml`
- **data 2016:** `data16_13TeV.periodAllYear_DetStatus-v89-pro21-01_DQDefects-00-02-04_PHYS_StandardGRL-
_All_Good_25ns.xml`
- **data 2017:** `data17_13TeV.periodAllYear_DetStatus-v99-pro22-01_Unknown_PHYS_StandardGRL-
_All_Good_25ns-Triggerno17e33prim.xml`
- **data 2018:** `data18_13TeV.periodAllYear_DetStatus-v102-pro22-04_Unknown_PHYS_StandardGRL-
_All_Good_25ns-Triggerno17e33prim.xml`

The integrated luminosities amount to 3.2 fb^{-1} for data 2015, 33.0 fb^{-1} for data 2016, 44.3 fb^{-1} for data 2017, and 58.4 fb^{-1} for data 2018, considering only the GRLs. The total integrated luminosity is 139 fb^{-1} . The proton bunch gap used in the LHC for the entire Run 2 is 25 ns.

7.2. Trigger

Since the final state targeted in this analysis is expected to contain large amount of momentum imbalance, the lowest un-prescaled¹ E_T^{miss} triggers items are used. For the 0LSR and the 1LCR this is HLT_XE70_MHT for data 2015 and HLT_XE90_MHT_L1XE50

¹See Section 4.2.6.

Table 7.1.: Summary of the simulated background processes. The V +jets cross-section refers to the NNLO normalization. For the Z +jets process this is calculated for $66 \text{ GeV} < m_{ll} < 116 \text{ GeV}$. The alternative W/Z +jets sample are used to evaluate the theoretical modelling uncertainties. The $t\bar{t}$ cross-section is normalized at NNLO+NNLL, using a top quark mass of $m_t=172.5 \text{ GeV}$. The di-boson process is treated differently in the leptonic, semi-leptonic and loop-induced, in the accuracy of the additional partons. Similarly, single-top is generated and treated differently in tW associated production, single top s- and t-channel production.

Process	Generator	PDFs	σ_{norm}	comment
W+jets	SHERPA 2.2 NLO	NNPDFNNLO	$20100 \pm 1000 \text{ pb}$	nominal
W+jets	MadGraph5+Pythia8 NLO	NNPDFNNLO and NNPDF23LO	$20100 \pm 1000 \text{ pb}$	alternative
Z+jets	SHERPA 2.2 NLO	NNPDFNNLO	$1906 \pm 95 \text{ pb}$	nominal
Z+jets	MadGraph5+Pythia8 NLO	NNPDFNNLO and NNPDF23LO	$1906 \pm 95 \text{ pb}$	alternative
$t\bar{t}$	Powheg+Pythia8 NLO	NNPDF3.0NLO	$831 \pm 76 \text{ fb}$	nominal
Diboson	SHERPA 2.2 NLO	NNPDFNNLO	81.3 pb	nominal
Single-top	Powheg+Pythia8 NLO	NNPDF3.0NLO	214 pb	nominal
Multijet	Pythia8	NNPDF23LO	77 mb	nominal

for data 2016, 2017 and 2018. For the 2LCR instead, the events are selected by requesting an additional lowest un-prescaled single lepton trigger. All triggers can be found on the Table A.1 in the Appendix. The E_T^{miss} triggers have been explicitly checked to be fully efficient, and well described in Monte Carlo simulation. Because of the large value of E_T^{miss} that is used as baseline requirement of the analysis, no additional correction was applied.

7.3. Simulated Samples

Simulated samples are used to describe and understand the background processes, and to optimize the sensitivity to the model's final state with the expected Dark Higgs signal. They are all produced using a full GEANT4 simulation [96] of the ATLAS detector.

7.3.1. Background Samples

The main background processes are the $W/Z + \text{jets}$, $t\bar{t}$, single-top, di-boson and multijet production, which are estimated with MC samples only and constrained with dedicated control regions. The full list can be found on Table 7.1.

The multijet background has been additionally evaluated in a data-driven way, to ensure a complete understanding of this process. In fact, given the large cross-sections, multijet events in which e.g. one of the jets is mis-measured, can give rise to significant amount of momentum imbalance. For this reason, scale factors are derived in multijet-enriched regions orthogonal to the signal region, and cross checked in validation regions. The scale factors account for the multijet events present in data and not in Monte Carlo simulation².

Finally, the effect of pile-up (in-time and out-of-time) is obtained overlaying minimum-bias³ events generated with Pythia8 on the hard-scattering events. The minimum-bias events are re-weighted as a function of the average interaction per bunch crossing μ , in order to match the pile-up conditions of Run 2.

7.3.2. Signal Samples

The signal samples are generated with the parameters choice as described in Section 3.2, and which will be always the same in this Part when referring to the signal:

- $g_q = 0.25$
- $g_\chi = 1$
- $\theta = 0.01$
- $m_\chi = 200 \text{ GeV}$

The values of the couplings chosen reflect the common choice of the benchmark values that are commonly employed in beyond the SM Dark Matter searches at the LHC. The choice of $m_\chi = 200 \text{ GeV}$ is motivated by the cross-section (that decreases with increasing m_χ), which is needed to probe the parameter space. The last two parameter values are: the mass of the Dark Higgs, m_s and the mass of the Z' boson, $m_{Z'}$. Given the finite ATLAS computational resources, the optimal signal production grid for the values $(m_{Z'}, m_s)$ was determined evaluating the sensitivity of the ATLAS detector to the model.

This procedure makes use of the *generic limits* (GL) that are calculated in the context of ATLAS analyses with similar target final state, i.e. large amounts of momentum

²Scale factor is defined as $SF^{multijet} = (N_{data} - N_{MC}^{non-multijet}) / N_{MC}^{multijet}$.

³Events with low E_T^{miss} .

imbalance (generated by Dark Matter) and jets. The $E_T^{\text{miss}} + H(b\bar{b})$ analysis [97] targets DM produced in association to a SM Higgs boson decaying to $b\bar{b}$. The $E_T^{\text{miss}} + V(\text{had})$ analysis [98] targets a similar final state, DM produced in association with an hadronically decaying W/Z boson. Even if the final state is different with respect to $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$, the sensitivity is expected to be smaller than $E_T^{\text{miss}} + H(b\bar{b})$, but higher than $E_T^{\text{miss}} + V(\text{had})$. The GL are provided in bins i of E_T^{miss} , with the sensitivity calculated for the two analyses as:

$$S_i^{H(b\bar{b})} = \frac{\sigma_i(pp \rightarrow s + E_T^{\text{miss}})_{\text{Dark Higgs}} \cdot (\mathcal{A} \cdot \varepsilon)_i}{\sigma_i(pp \rightarrow H \rightarrow b\bar{b} + E_T^{\text{miss}})_{\text{obs}}}$$

$$S_i^{V(\text{had})} = \frac{\sigma_i(pp \rightarrow s + E_T^{\text{miss}})_{\text{Dark Higgs}} \cdot (\mathcal{A} \cdot \varepsilon)_i}{\sigma_i(pp \rightarrow V(q\bar{q}) + E_T^{\text{miss}})_{\text{obs}}}$$

where $\sigma_i(pp \rightarrow s + E_T^{\text{miss}})_{\text{Dark Higgs}}$ is the partonic cross-section of the Dark Higgs signal per E_T^{miss} category. $\sigma_i(pp \rightarrow H \rightarrow b\bar{b} + E_T^{\text{miss}})_{\text{obs}}$ and $\sigma_i(pp \rightarrow V(q\bar{q}) + E_T^{\text{miss}})_{\text{obs}}$ are the observed upper 95% CL_s limits on the production cross-section of the $H \rightarrow b\bar{b} + E_T^{\text{miss}}$ and $V(q\bar{q}) + E_T^{\text{miss}}$ processes, respectively.

For the sensitivity estimate, measure $S = \sum_i S_i^4$ is defined as the sum of the sensitivities in each E_T^{miss} category i . A point in the parameter space is expected to be excluded if $S > 1$.

The results of the GL estimate and the consequent choice of the signal grid generation are shown in Figure 7.1. On the left, the S as a function of the two mass parameters and on the right the signal grid can be seen. The bulk of the sensitivity of this analysis is located in the area of Z' masses between about 500 and 1000 GeV and Dark Higgs masses between about 160 GeV and 250 GeV. The sensitivity degrades as the cross-section decreases in both positive directions. A detailed description of the GL estimate can be found in Appendix A.3.

⁴Not to be confused with s the Dark Higgs.

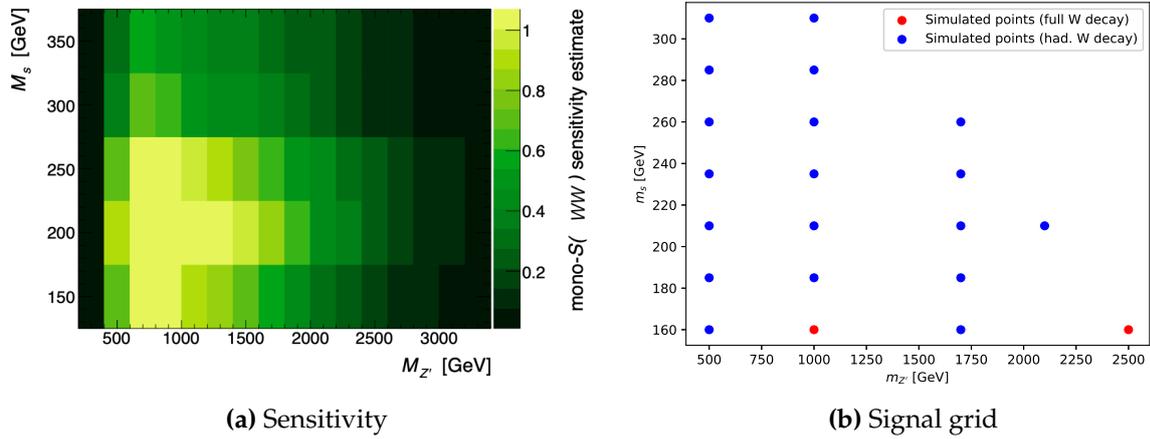


Figure 7.1.: The expected sensitivity S , as calculated with the GL procedure from the $E_T^{\text{miss}} + H(b\bar{b})$ [97] and $E_T^{\text{miss}} + V(\text{had})$ [98] analyses (left), is depicted on the $(m_{Z'}, m_s)$ plane on the x-y axis and the S on the z-axis. The points with $S > 1$ are expected to be excluded in the parameter space. The grid of produced signal samples in the plane $(m_{Z'}, m_s)$, chosen on the basis of the GL estimate (right). Blue points are generated with hadronic-only decays of the W boson, which is the target of this analysis. Red points are generated instead with the full decay, for preliminary studies of the semi-leptonic analysis.

8. Objects, Selections and Simulations

In this Chapter, the relevant objects and the reconstruction methods used in the signal and control regions of the analysis are briefly summarized.

8.1. Objects

The objects used in the analysis are listed here: they are leptons such as electron/positrons, muons, taus, and various definition of jets. To avoid double counting of the measurements in the calorimeters and the tracker system, an overlap removal is performed as explained in the Appendix A.4.

8.1.1. Electrons

Electrons are measured with the tracker and the calorimeter systems. To assure the quality of the measurements, electron reconstruction takes place only in the regions covered by the tracker system $|\eta| < 2.47$. First, a *seed-cluster* of sufficient energy is identified using the energy deposits in the calorimeters [99]. Subsequently, likelihood identification [100] is performed to associate the electrons track candidates to the seed-clusters, also accounting for effects such as bremsstrahlung and photon conversion in the detector material. In order to be able to reject the electron candidates actually originating from hadronic jets or converted photons, a dedicated identification procedure is applied. Properties such as the calorimeter shower shape and profile (different between hadronically interacting particle and EM ones), track and cluster spatial matching, and track quality criteria based on the number of hits in the tracker subsystems are exploited. Finally, a ΔR isolation is required around the calorimeter clusters and the associated tracks. The electron energy is then calibrated [101]. The identification criteria that are listed can be made more relaxed or stringent based on the purity of the selection and the background rejection desired. These working points are called *Loose*, *Medium* and *Tight*.

Electrons in the Analysis Electrons are divided into *baseline* and *signal* electrons in order to address the different needs of the analysis regions. Baseline electrons make use of Loose identification and isolation requirements and are required to have a $p_T > 7$ GeV, allowing a higher reconstruction efficiency for vetoes that are placed in the 0LSR and the 1LCR. In the di-electron control region, signal electrons are used instead, to gain advantage from the higher purity, using a Loose identification and isolation, but requiring higher transverse momentum threshold $p_T > 25$ GeV.

8.1.2. Muons

Muon candidates are reconstructed exploiting both the tracker system and the muon spectrometers with of a combined fit [102], where the hits in the inner detector, the energy losses in the calorimeters and the hits in the spectrometers are considered. The pseudo-rapidity range is extended to $|\eta| < 2.7$, using the muons segments from the spectrometers only. In the range $|\eta| < 2.5$, tracks in the inner detector are required to satisfy the quality criteria. Finally in the $|\eta| < 0.1$ region muon selection algorithms are used, requiring only the presence of segments in the spectrometers, and energy losses in the calorimeter compatible with minimum ionizing particles.

Muons in the Analysis Similarly as in the electrons case, muons are divided into high efficiency baseline muons used for the vetoes in the signal region, and the signal muons used in the control regions. Baseline muons make use of Loose identification and isolation requirements, and are selected to have $p_T > 7$ GeV in the $|\eta| < 2.5$. Signal muons make use of Medium selection and Tight isolation and are selection to have $p_T > 25$ GeV in the extended pseudo-rapidity range $|\eta| < 2.7$.

8.1.3. Taus

Tau lepton candidates are seeded by jets, with an identified tau track secondary vertex having a large momentum fraction [103]. It is required to have $p_T > 20$ GeV and $|\eta| < 2.5$, and at least three charged tracks. The identification of taus is further enhanced making use of boosted decision trees discriminants using tracking and calorimeter information.

Taus in the Analysis Tau vetoes are applied in all analysis regions; the Loose working point is used, and the crack region of the calorimeters is excluded.

8.1.4. Track Selection

Tracks are reconstructed using hits in the inner detector with an iterative algorithm seeded on a combined measurement of the silicon detector. A Kalman filter with a stringent ambiguity solver [104] is used for this purpose (for a detailed description and performance see e.g. [105]). The number of hits in the tracker and the relative impact parameter are used with Loose working point for track quality and primary vertex compatibility:

- number of silicon hits ≥ 7 ,
- number of shared modules ≤ 1 ,
- number of silicon holes ≤ 2 ,
- number of pixel holes ≤ 1 .

where a 'shared module' is a hit in a pixel or SCT sub-detector that is shared by more than one track, and a hole is the lack of a hit along the trajectory of a track. Moreover, the track must be part of the primary vertex fit or, when not associated with other vertexes, the absolute value of $z_0^{BL} \times \sin\theta$ must be less than 3 mm^1 . The primary vertex associated with the hard-scattering is selected in the event among all the vertexes, having the highest scalar quadratic sum of the p_T of the associated tracks, and satisfying the transverse and longitudinal impact parameter requirements.

Tracks in the Analysis Only tracks fulfilling the quality criteria with $p_T > 0.5 \text{ GeV}$ and $|\eta| < 2.5$ are used, employed to construct the TAR jets. Tracks in fact can account for the in-jet charged component, and their measurement directly benefits from the superior angular resolution of the tracker system, as discussed in the previous Part.

¹ z_0^{BL} is the difference between the longitudinal position of the primary vertex and the longitudinal position of the track along the beam line at the point where d_0^{BL} (the transverse impact parameter calculated with respect to the measured beam line position) is measured.

8.1.5. Small- R Jets

Small- R jets have been described in details in Section 5.1. They are reconstructed using the anti- k_t algorithm with radius $R = 0.4$, from EMTopoClusters and fully calibrated². Small- R jets are required to have $p_T > 20$ GeV and $|\eta| < 4.5$; they are divided into *central* and *forward* jets. Central jets fall within the acceptance of the tracker system $|\eta| < 2.5$, which can help selecting those originating from the primary vertex, using the *Jet Vertex Tagger* (JVT) [106]. The JVT is a potent tool that has been optimized to allow discerning jets from the hard-scattering and those from pile-up (in time and out-of-time) using tracking measurements. The forward jets reach the maximum extension of the calorimeters at $|\eta| < 4.5$, and the transverse momentum required is at least 30 GeV.

Small- R Jets in the Analysis Small- R jets are used in the intermediate category of the analysis in order to reconstruct the Dark Higgs candidate. Moreover, they are employed in all analysis regions to calculate the missing transverse momentum E_T^{miss} , and to compute event variables.

$R=0.2$ Jets and Optimization of the Analysis R -scan $R = 0.2$ jets or simply $R = 0.2$ jets have also been described in details previously 5.1. They are reconstructed using the anti- k_t algorithm with radius of $R = 0.2$, using LCTopoClusters, and fully calibrated, and are required to have a $p_T > 20$ GeV and fall within $|\eta| < 2.5$. The $R=0.2$ jets are a potential input to TAR jets that could be used in the two most sensitive categories of the analysis, the merged and the intermediate, substituting the Small- R jets. Having a smaller radius with respect to Small- R jets, they can better access the charged to neutral fluctuations. The use of $R=0.2$ jets in the context of a re-optimization of this analysis as well as other beyond the Standard Model searches is envisioned, expecting sizeable sensitivity gains.

8.1.6. TAR Jets

Two types of Large- R jets are considered in the analysis, the standard and the TAR jets. The standard Large- R jets are used as a benchmark to assess the performance gains of the TAR jets, which are the only definition of Large- R jets used in the analysis.

²See Section 5.3.

TAR jets have been thoroughly described in the previous Part. They are built from tracks and calorimeter deposits. First, the Small- R jets are clustered together with an anti- k_t algorithm of radius $R = 0.8$. The radius parameter can be fully optimized, and $R = 0.8$ has been found to be the optimal for the final state considered. Subsequently, the reclustered jet is trimmed with identical settings as in standard Large- R jets. Tracks are matched to their respective Small- R jets, and their momenta are rescaled according to the track-assisted procedure.

TAR Jets in the Analysis The TAR jet reconstruction technique is exploratorily used for the first time in the ATLAS collaboration, paving the way to its further employment in other searches. TAR jets are required to have a p_T of at least 200 GeV and also fall well within the tracker acceptance $|\eta| < 2.0$. They are used in the most sensitive categories of the analysis, the merged and the intermediate. Benefiting from the excellent track-based jet substructure, few observables are calculated from TAR jets: the τ_{42} , τ_{43} and D_2 . They offer a separation power on the number of prongs in the jet, as will be detailed later.

Standard Large- R Jets Standard Large- R jets are reconstructed from LCTopoClusters using the anti- k_t algorithm, with a radius parameter $R=1.0$. Following the pile-up mitigation strategy outlined in the previous Part, trimming is applied to constituents clustered with k_t algorithm with a radius R_{sub} of 0.2. Only clusters with relative p_T fraction above 5% are kept, while the others are discarded. The mass is calculated as the combination of calorimeter and track-assisted mass.

Standard Large- R Jets in the Analysis In order to compare against TAR jets, standard Large- R jets are required to have a p_T of at least 200 GeV and fall well within the tracker acceptance $|\eta| < 2.0$. The cluster-based substructure observables τ_{42} and D_2 are used as well to compare against TAR jets.

8.1.7. Missing Transverse Momentum E_T^{miss} and Object based E_T^{miss} Significance \mathcal{S}

The missing transverse momentum (E_T^{miss}) is an object of critical importance in searches for Dark Matter. It is used to measure the quantity of momentum that escapes the

ATLAS detector undetected, such as in the case of neutrinos or Dark Matter particles. It is evaluated using the transverse momentum of visible objects that can be reconstructed and identified [107]. These are: electrons and muons selected using the baseline selection, fully calibrated Small- R jets that are described above, and all the tracks as reconstructed in the inner detector. For the reconstruction of the E_T^{miss} , no photons and taus are used.

Object based E_T^{miss} significance \mathcal{S} is a variable based on E_T^{miss} , which relates the E_T^{miss} to the resolutions of the individual jets, electrons, and muons in a event. It is optimized to remove the dependence with respect to pile-up effects and soft terms, as well as detector mis-measurements, resolution and inefficiencies [108].

E_T^{miss} and \mathcal{S} in the Analysis In the analysis, targeting final states with large momentum imbalance, a $E_T^{\text{miss}} > 200$ GeV baseline requirement is used, exploiting the exponential decrease of QCD background as a function of E_T^{miss} to enhance the search sensitivity. \mathcal{S} is used to further reduce the multijet background contribution in the 0LSR and in the 1LCR, requiring $\mathcal{S} > 15$. In the di-lepton control region, 2LCR, object based E_T^{miss} significance is instead used to increase the purity of the selection, by requiring $\mathcal{S} < 15$.

8.1.8. Variable- R Track Jets

Variable- R (VR) track jets are a powerful tool, allowing for b -tagging in Lorentz boosted topologies. Tracks are selected with Loose requirement as listed above (Section 8.1.4), and clustered together using the anti- k_t algorithm. The clustering radius is dependent on the transverse momentum via the relation $R(p_T) = 30\text{GeV}/p_T$ [109], allowing substantial increase in the b -tagging efficiency. The radius can move between a minimum, $R = 0.02$ and a maximum, $R = 0.4$ as optimized in detailed studies [110] and calibrated in $t\bar{t}$ events [111] [112].

Variable- R Track Jets in the Analysis VR track jets are use to veto events containing bottom quarks in the final state such as in the case of $t\bar{t}$ events. They are required to have a $p_T > 0.5$ GeV, $|\eta| < 2.5$ and the vertex has to satisfy $z_0 \sin \theta < 3$ mm. To enhance the b -tagging efficiency, a multivariate algorithm based on boosted decision trees (**MV2c10**) is applied on the VR track jet, exploiting the kinematic variables and the

probability of reconstructing secondary vertexes for different flavour hypothesis [113]. The working point chosen has a 77% tagging efficiency.

8.2. Event Selection

In this Section, the event selection of the analysis is described. First, the baseline selection will be introduced, which applies to all the signal and control regions and to the different categories. Then, the so-called *priority-ranked* event selection and the merged and intermediate categories will be defined. The priority-ranked event selection makes sure that each event is checked first in the most sensitive merged category, and then the intermediate.

8.2.1. Baseline Selection

The first requirement of the baseline selection divides the analysis into the three regions: the 0LSR in case of absence of baseline leptons, the 1LCR in case of a signal muon and no baseline electrons, and finally the 2LCR in case of two signal electrons or opposite sign signal muons. A veto is additionally placed on the presence of taus in all the regions. To address the large amounts of momentum imbalance arising from Dark Matter in the target final state, the missing transverse momentum is required to have $E_T^{\text{miss}} > 200$ GeV in the 0LSR. Furthermore, to ensure the compatibility of the kinematics regimes between the analysis regions, the 1LCR and 2LCR utilizes a modified definition of E_T^{miss} . In the 1LCR, the transverse momentum of the muon with the highest momentum is not considered in the computation of E_T^{miss} , which is then referred to as $E_T^{\text{miss,no}\mu}$. This quantity must be at least 200 GeV. Finally, in the 2LCR, the transverse momentum of the di-lepton system, $p_T^{\ell\ell}$, is required to also have at least 200 GeV. The choice of 200 GeV is motivated by sensitivity arguments, but also efficiency of the E_T^{miss} triggers. In order to reduce the contribution of the $t\bar{t}$ background process into the target phase-space, a veto on b -tagged events using VR track jets is applied, as described in the previous Section. The multijet background can be a dangerous source of artificially high E_T^{miss} in case of energy mis-measurements, where e.g. a jet recoiling against another could be measured with an energy much higher or much lower than it should have. In this case, the E_T^{miss} is pointing in the direction of one of these jets. To drastically reduce this background source, a cut on

the azimuthal distance between the leading three Small- R jets and the E_T^{miss} is applied: $\Delta\phi(E_T^{\text{miss}}, \text{jets}_{1,2,3}) > 0.35$. In the 2LCR this cut is not applied since the requirement of the two leptons is sufficient to greatly reduce the multijet contribution. Moreover, an object based E_T^{miss} significance $\mathcal{S} > 15$ cut is applied in the 0LSR and the 1LCR. To eliminate possible signal contamination in the 2LCR, this cut is inverted ($\mathcal{S} < 15$). Finally, a minimum of two Small- R jets is required in all the regions. A summary of the baseline requirements is summarized on Table 8.1.

Table 8.1.: The baseline selection that is required in all the signal and control regions.

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

Signal distribution at baseline selection

The E_T^{miss} and TAR jet mass distribution for $s \rightarrow W^+W^-$ signal events with different m_s is shown in Figure 8.1. As it can be noted, heavier masses are more difficult to reconstruct with a single Large- R jet object, and the mass distribution exhibits a significant tail. The signal cross-section also plays an important role, and can be appreciated through the magnification factor shown in the top right labels. On the left the sizeable E_T^{miss} extending beyond 500 GeV can be seen.

Event Cleaning The events considered in the analysis have to satisfy basic event cleaning criteria, in order to protect against defects during data taking. First, events are required to be part of the good-run lists as defined in Section 7.1. Additional vetoes are then applied to reject corrupted events based on:

- Noise burst in the LAr calorimeter or with data corruption,
- Corrupted events due to Tile errors,

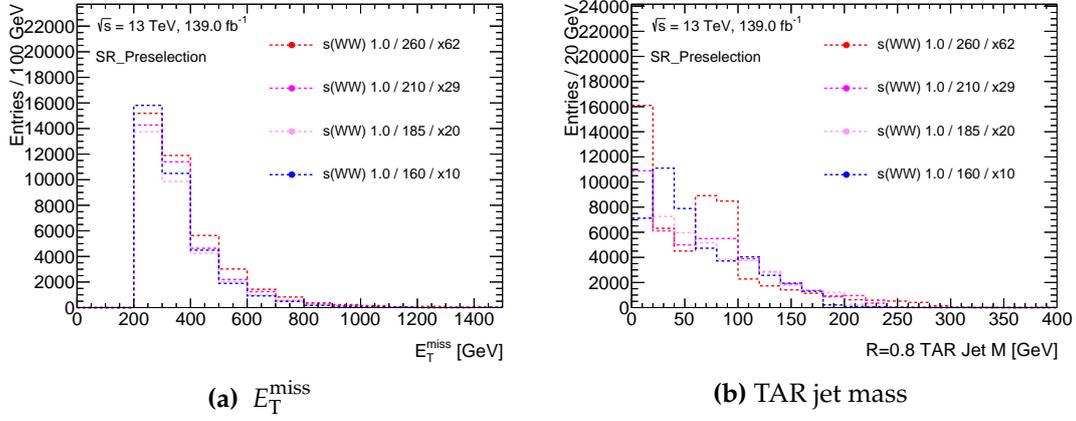


Figure 8.1.: The E_T^{miss} (left) and TAR jet mass (right) for selected signal W^+W^- final state. Different mass hypotheses are shown, magnified as displayed in the label to be visible in the same y-axis range.

- SCT recovery procedure for single event upsets,
- Incomplete events.

Moreover, all the events are required to have a reconstructed primary vertex with at least two tracks as defined above, to make sure that the event contains an hard scattering interaction.

8.2.2. Categories Definition

As mentioned earlier, each of the analysis regions is divided into two categories, the merged and the intermediate category, as a mean to address the different kinematic regimes that are present in the final state as a function of the Lorentz boost of the Dark Higgs boson. This allows an increased sensitivity of the search, addressing the reconstruction challenge with different techniques made available within the ATLAS collaboration. All the categories make use of the objects as defined in the previous Chapter.

The Merged Category

In the merged category, the Dark Higgs, s , is so energetic, that all its decay products are confined within a relatively small radius, which depends also on the m_s hypothesis³. If the decay products of $s \rightarrow VV$ are collimated together, and reconstructed within a single Large- R jet, then the mass of the entire system is calculated from the mass of the single jet. However, if the radius needed to reconstruct this object is too large, the efficiency of the reconstruction is endangered by the high levels of pile-up at the LHC, which are collected via the large area. This creates a sweet spot roughly around $R = 1.0$, which is in fact used for standard Large- R jets, where the pile-up and soft radiation contained within the jet is not prohibitive, and the decaying object is well contained. The Large- R jet used to address this reconstruction challenge is the TAR jet; the optimal choice for the radius is $R = 0.8$. In Figure 8.2 the ratio of the significances for TAR⁴ jets over standard Large- R jets. It can be seen that the TAR jets show throughout the entire mass spectrum substantial gains over standard large- R jets.

To enhance the sensitivity of the merged category, TAR jets are required to satisfy the following cuts:

- $p_T^{\text{TAR}} > 300 \text{ GeV}$
- $0 < \tau_{42}^{\text{TAR}} < 0.3$
- $0 < \tau_{43}^{\text{TAR}} < 0.6$
- $100 \text{ GeV} < m_{\text{TAR}} < 400 \text{ GeV}$

The high p_T requirement ensures to avoid the region dominated by the QCD background, where soft radiation can build up higher masses in the large area of the jet. The substructure requirements are based on N-Subjettiness cuts, with τ_{42} distinguishing four-prongs (signal-like) versus two-prongs (background-like), and similarly τ_{43} distinguishing four-prongs versus three-prongs. Finally, the mass of the TAR jet must be in the region of interest of the parametric space of the signal: between 100 GeV and 400 GeV. The lower bound is given by the rapid decrease of the $s \rightarrow VV$ branching fraction and the higher bound by the decrease of the total cross-section.

³Again recall the rule of thumb $\Delta R = 2m/p_T$.

⁴TAR jets in Figure 8.2 are constructed with R -scan $R = 0.2$ jets. TAR jets are constructed in this analysis with Small- R jets instead.

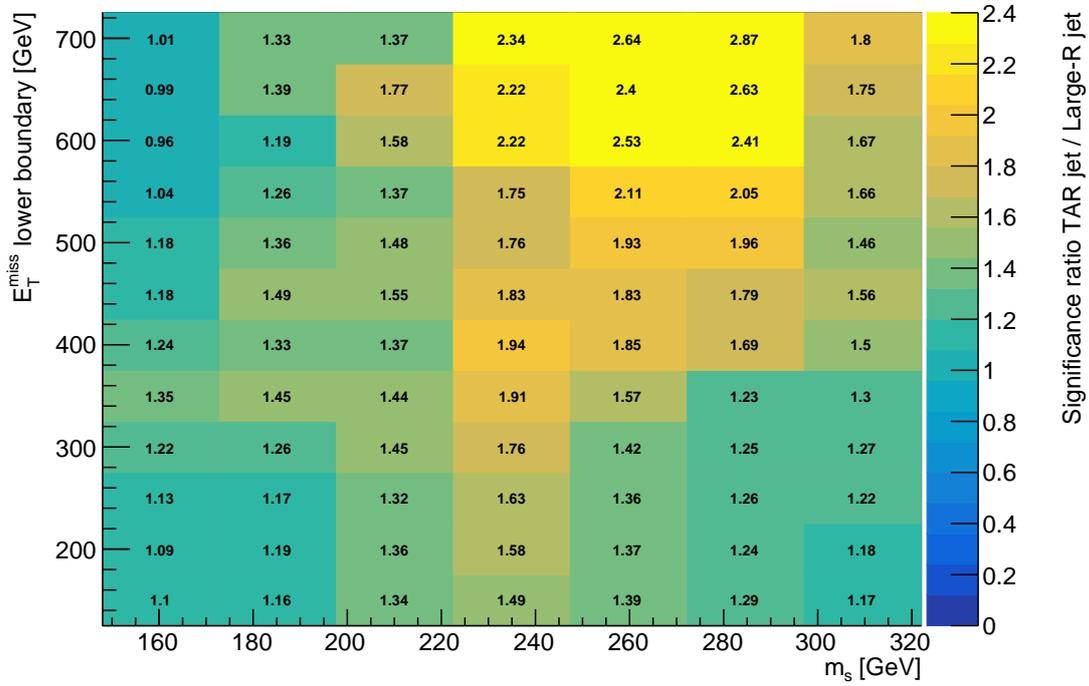


Figure 8.2.: Ratio of significances between TAR jets ($R = 0.8$ constructed on R -scan jets) and standard large- R jet in the merged category for $m_{Z'} = 1$ TeV. The mass of the Dark Higgs is shown on the x-axis, and on y-axis the E_T^{miss} lower boundary. The significance is computed after the requirement that the candidate mass is within 50 GeV window centred around the Dark Higgs mass, and after the $\tau_{42}^{\text{TAR}} < 0.3$ requirement. Significance is defined as $S = \sqrt{2((s + b) \log(1 + s/b) - s)}$.

The Intermediate Category

In the intermediate category, the VV system still retains a big amount of the Lorentz boost associated with the merged category. Since this boost will be, in general, unevenly distributed among the vector bosons, one of them would be leading in p_T , and the other one sub-leading. This creates an interesting kinematic regime, where the decay products of the leading V in p_T are collimated within some radius that can be reconstructed with a single Large- R jet. The sub-leading V in p_T needs instead to be reconstructed via the identification of single jets from the V hadronic decay with Small- R jets.

This particular behaviour of the VV system at moderate transverse momenta is addressed using a dedicated algorithm, the *TAR+Comb*, constructed to improve the sensitivity of the category. The objective is to exploit collimated topologies of the vector bosons as much as possible, in order to avoid the combinatorial ambiguities that arising when Small- R jets are used.

The algorithm proceeds as explained below (Algorithm 1 and with more visualisation in Appendix A.8). The events are separated according to the presence of a TAR jet with given properties. If the TAR jet has a mass and substructure compatible with a single V boson, it is interpreted as fully containing the leading V in p_T . Therefore the sub-leading V boson in p_T will have its decay products further away, at larger angular distances, and not contained within the TAR jet considered or other ones. This V boson can be reconstructed with a pair of Small- R jets. They will be not much far away in the $\eta - \phi$ region, and their invariant mass will be compatible with that of the boson. Therefore, all the Small- R jet pairs around the TAR jets are ordered by the distance of their masses with respect to the m_W ⁵. The pair with the mass closer to m_W is assumed to be the sub-leading p_T V , and its four-momentum is added to the TAR jet.

If the event contains a TAR jet with mass above that of the W/Z boson, however, this can be explained if the leading V in p_T is fully contained within the TAR jet. Additionally, also a jet from the hadronic decay of the sub-leading V in p_T falls within the TAR radius. The missing jet from the sub-leading p_T boson is therefore looked

⁵As written above this is done because of the higher $s \rightarrow W^+W^-$ cross-section and the experimental mass resolution of the vector boson.

for, and the pairs compatible with the W/Z mass are constructed. The closest one is identified and the corresponding jet is added to the four-momentum of the TAR jet.

Algorithm 1: TAR+Comb algorithm.

Result: TAR+Comb returns the invariant mass of the Dark Higgs

if One TAR jet compatible with a W/Z boson is found: $60 \text{ GeV} < m_{TAR} < 100 \text{ GeV}$ and $D_{2,TAR} < 1.5$

then

The sub-leading p_T W/Z boson is looked for using a pair of Small- R jets that is compatible with the W/Z mass, and not far away from the TAR jet (the s still has a significant boost):

1. All Small- R jets within a radius (R') of 2.5 around the TAR jet are considered
2. Small- R jets overlapping with TAR jets are discarded
3. All possible pairs of remaining Small- R jets are formed
4. The pairs are ordered with respect to the difference with the W/Z mass
5. The pair closest in mass to m_W is designated as the W/Z candidate and is required to be within the mass window $60 \text{ GeV} < m_{jj} < 100 \text{ GeV}$
6. The Dark Higgs mass is then computed as the combined 4-momentum of the Small- R jet pair and the 4-momentum of the TAR jet

else

if One TAR jet compatible with a W/Z boson plus a prong from the other boson, $m_{TAR} > 100 \text{ GeV}$

then

In this case, there is only one prong that is not captured within the area of the TAR jet. To identify the most probable, all possible combinations are formed, which are compatible with the W/Z mass:

1. All Small- R jets within a radius (R') of 2.5 are considered
2. Small- R jets are discarded if they have less than 5% of the p_T of the TAR jet
3. All possible unordered pairs of remaining Small- R jets are formed
4. The jet pairs are ordered with respect to the difference with the W/Z mass
5. The jet pair closest in mass to m_W falling within the window $60 \text{ GeV} < m_{jj} < 100 \text{ GeV}$ and overlapping with the TAR jet is designated as a W/Z candidate and not considered in the next steps.
6. The remaining jet pairs must contain **exactly one Small- R jet** each, overlapping with the TAR jet
7. The jet pair that satisfies the above, closest in mass to m_W falling within the mass window (60 - 100 GeV) is designated as the other W/Z candidate
8. The Dark Higgs mass is then computed as the 4-momentum of the TAR jet plus the Small- R jet of the latter W/Z candidate that is partially overlapping with the TAR jet.

else

| Reject the event in the intermediate category

end

end

Table 8.2.: A summary of the different reconstruction algorithms used in each of the prioritized signal categories, explained in detail in this Section. The final kinematic selection requirements of the reconstructed Dark Higgs candidate are also listed.

merged	intermediate
$\geq 1 R = 0.8$ TAR jet	= 1 Dark Higgs candidate
$p_{\text{TAR}} > 300$ GeV	reconstructed with TAR+Comb
$0. < \tau_{42} < 0.3$	$100 \text{ GeV} < m_{\text{TAR+Comb}} < 400 \text{ GeV}$
$0. < \tau_{43} < 0.6$	
$100 \text{ GeV} < m_{\text{TAR}} < 400 \text{ GeV}$	

Summary of the Categories definitions

The requirements made on the above categories are summarized together in the Table 8.2.

Priority-Ranking and $E_{\text{T}}^{\text{miss}}/E_{\text{T}}^{\text{miss,no}\mu}/p_{\text{T}}^{\ell\ell}$ binning

In order to achieve an higher sensitivity, a priority-ranking is applied. Since the signal to background ratio is higher in the merged category with respect to the intermediate one, events are first checked for the merged category. If one or more requirements are not met, the event is successively checked for intermediate category. This procedure avoids discarding events, which can contribute to the overall sensitivity in other categories, ensuring the disjointness of these categories. It is applied in the signal region as well as in the control regions.

As shown in the $E_{\text{T}}^{\text{miss}} + V(\text{had})$ and $E_{\text{T}}^{\text{miss}} + H(b\bar{b})$ analyses, the statistical treatment greatly benefits using $E_{\text{T}}^{\text{miss}}$ bins. This is because the SM background is exponentially falling with $E_{\text{T}}^{\text{miss}}$. A $E_{\text{T}}^{\text{miss}}$ binning is therefore applied to all the categories in the SR. As discussed above, in order to retain a similar kinematics as in the SR, $E_{\text{T}}^{\text{miss,no}\mu}$ is used in the 1LCR for as the binning variable, and $p_{\text{T}}^{\ell\ell}$ in the 2LCR. The bins are summarized in Table 8.3.

The merged category has not been included in the first bin 200-300 GeV as the signal yield for this range does not justify the increase complexity of the analysis. This bin was instead used for a data-driven multijet estimate.

Table 8.3.: The $E_T^{\text{miss}} / E_T^{\text{miss,no}\mu} / p_T^{\ell\ell}$ bins for the different regions 0LSR / 1LCR / 2LCR for the merged and for the intermediate categories are shown. The ranking is merged (M) \rightarrow intermediate (I).

	200-300 GeV	300-500 GeV	more than 500 GeV
0LSR (E_T^{miss})	I	M,I	M,I
1LCR ($E_T^{\text{miss,no}\mu}$)	I	M,I	M,I
2LCR ($p_T^{\ell\ell}$)	I	M,I	M,I

8.3. Data and Monte Carlo Comparison

The data/MC comparison is a critical step to assess the validity of the background description. In order to proceed further with the statistical evaluation, the 1LCR and 2LCR event yields are compared against their MC simulated predictions. The total yield of the background processes in the 0LSR is in fact obtained from MC samples, and their normalization is estimated from the 1LCR and 2LCR in a combined fit. In order to avoid any potential bias from the analysers, searches for new physics are required to proceed with a *blinded* analysis. The data in the SR is compared against simulated samples, only after the CRs and the statistical treatment is well under control. Only as the analysis strategy is evaluated as sound and complete, the analysers may proceed with *unblinding*. The comparison of data and MC yield is presented for the 1LCR, 2LCR and for the 0LSR, in Appendix B. The 0LSR is obtained after the unblinding step.

The distributions shown display the statistical error on data only and before the fit, meaning that no additional scale factors obtained via the statistical treatment are applied. No uncertainty on the background sample is shown. The pre-fit and post-fit distribution with the complete set of uncertainties is shown in Chapter 11.

In Figure 8.3, the pre-fit data/MC comparison for the candidate Dark Higgs mass in all the analysis regions and for all the categories, inclusively in E_T^{miss} , is shown. The mass is the most important observable considered in the analysis, as the only shaped variable taken into account in the statistical treatment. Good agreement of the MC with the observed data is found within the uncertainties. The pre-fit distribution can additionally be seen in Appendix D.

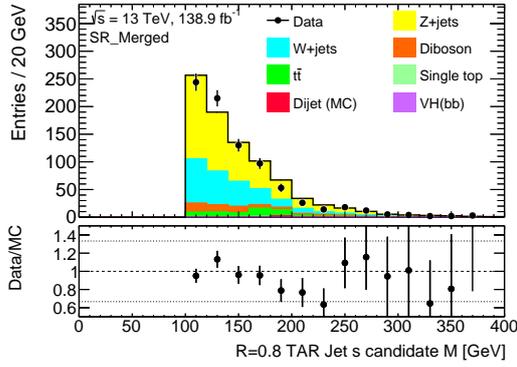
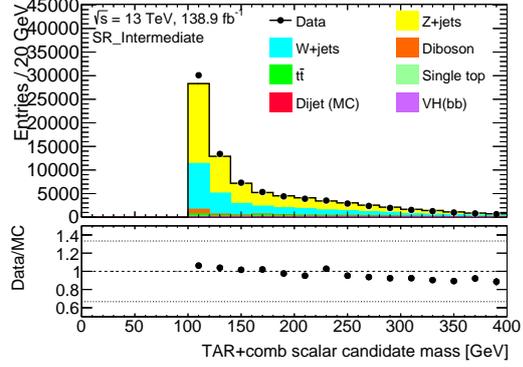
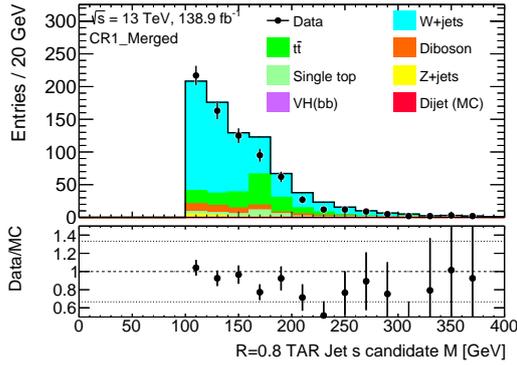
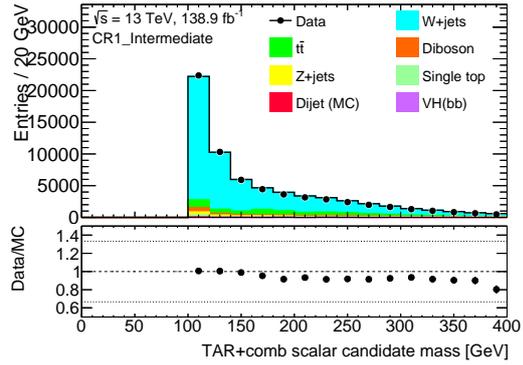
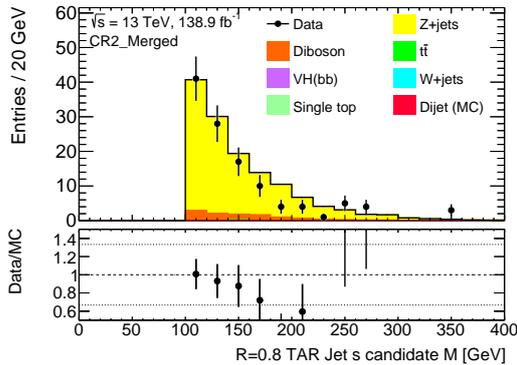
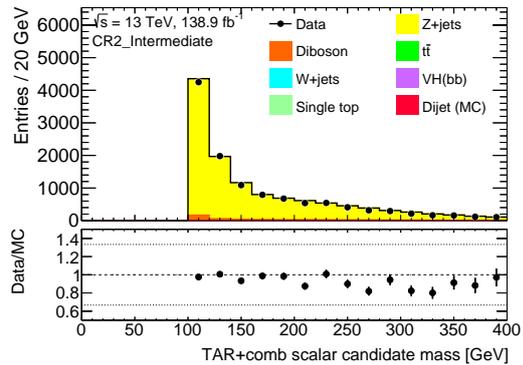
(a) Leading TAR jet m , 0LSR(b) TAR+comb Dark Higgs candidate m , 0LSR(c) Leading TAR jet m , 1LCR(d) TAR+comb Dark Higgs candidate m , 1LCR(e) Leading TAR jet m , 2LCR(f) TAR+comb Dark Higgs candidate m , 2LCR

Figure 8.3.: Data and MC distributions in the inclusive E_T^{miss} selection for the Dark Higgs candidate mass as reconstructed in the merged and in the intermediate category, for the 0LSR, 1LCR and 2LCR.

9. Systematic Uncertainties

The systematic uncertainties are considered in the analysis, as they could impact both the events yield and the shape of signal and background distributions. For example, the W +jet normalization in the control region, or the shape of the invariant mass distributions of the signal could be affected. Since the statistical treatment of the analysis relies these measured quantities, a rigorous and complete treatment is needed. Two general sources of systematic uncertainties can be identified: experimental and theoretical systematic uncertainties. The former refer to all those effects that have repercussions on object reconstruction, identification and calibration, as well as measured quantities in the events, and they are quantified as uncertainties. As an example, the jet energy scale (JES) systematic uncertainty reflects the uncertainty on the determination of the jet energy, and it is evaluated in dedicated measurements. Theoretical systematic uncertainties instead reflect our knowledge (or lack thereof) on the quantities that are predicted theoretically with the Monte Carlo simulations, such as the modelling of the Z +jet background. The systematic uncertainties determined using limited MC samples are *smoothed* to avoid instabilities in the fit. In the statistical treatment described on the next Chapter 10, systematic uncertainties are called Nuisance Parameters. The full list of the systematic uncertainties is provided in Table 9.1.

9.1. Experimental Systematic Uncertainties

The Experimental systematic uncertainties are briefly described here, and they depend both on the ATLAS detector and the LHC accelerator. Many of these uncertainties have been centrally made available by the collaboration.

Luminosity Luminosity is a critical factor in all the measurements at colliders. Its central value and uncertainty directly impacts e.g. the Monte Carlo simulations, which are scaled to match its value. The luminosity for all data taking in the years 2015, 2016, 2017 and 2018 has a precision of 2.1%, 2.2%, 2.4% and 2% respectively. The combined luminosity uncertainty corresponds to 1.7%.

Trigger Trigger uncertainties are expected to have a limited impact on the analysis, since the requirement of $E_T^{\text{miss}} > 200$ GeV places it in the full efficiency region. However, for completeness, the uncertainty on the modelling of the E_T^{miss} triggers in MC scale factor¹ is included as statistical component of the fit.

Electron Reconstruction The electron reconstruction uncertainty is fully provided by the collaboration, by the *E/gamma* group. The uncertainty on the electron reconstruction and identification, the scale factors, and the energy scale and resolution are taken into account. The event yield is calculated after applying the 1σ variation around the smearing and efficiency scale factors.

Muon Reconstruction Muon reconstruction uncertainties take into account the muon efficiency factor, isolation, and momentum scale and resolution. These are provided by the muon combined performance group.

Jets The relevant jet energy scale (JES) and jet energy resolution (JER) uncertainties are provided by the *JetEtMiss* group, and they have been already discussed in the previous Part.

TAR Jets The uncertainties of the TAR jets are derived bottom-up from its constituents: Small- R jets and tracks. Additionally, uncertainties on hadronisation model are studied using different generators in $t\bar{t}$ simulated samples, and were found have a negligible impact. They are therefore neglected and not considered among the systematic uncertainties of the analysis.

Missing Transverse Momentum Similarly, E_T^{miss} is an observable derived from multiple objects. Therefore its uncertainty stems from the different objects used in its construction. Only the uncertainties on the soft term are provided separately.

¹Briefly explained here 7.3.1.

Table 9.1.: Qualitative summary of the experimental systematic uncertainties that are considered in this analysis. The name of the systematic uncertainties and their description is shown. The smoothing is used to mitigate the impact that limited MC statistics has on the fit.

Systematic uncertainty	Short description	smoothing
Event		
Luminosity	uncertainty on the total integrated luminosity	-
PRW_DATASF	pile-up-reweighting uncertainty	yes
Electrons		
EL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR	trigger efficiency uncertainty	no
EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	reconstruction efficiency uncertainty	no
EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	ID efficiency uncertainty	no
EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	isolation efficiency uncertainty	no
EG_SCALE_ALL	energy scale uncertainty	yes
EG_SCALE_AF2	energy scale uncertainty for AFII simulation	yes
EG_RESOLUTION_ALL	energy resolution uncertainty	yes
Muons		
MUON_EFF_TrigSystUncertainty	trigger efficiency uncertainties	no
MUON_EFF_TrigStatUncertainty		
MUON_EFF_RECO_STAT	reconstruction and ID efficiency uncertainty for $p_T > 15$ GeV	no
MUON_EFF_RECO_SYS		
MUON_EFF_RECO_STAT_LOWPT	reconstruction and ID efficiency uncertainty for $p_T < 15$ GeV	yes
MUON_EFF_RECO_SYS_LOWPT		
MUON_EFF_ISO_STAT	isolation efficiency uncertainty	no
MUON_EFF_ISO_SYS		
MUON_EFF_TTVA_STAT	track-to-vertex association efficiency uncertainty	no
MUON_EFF_TTVA_SYS		
MUON_SCALE	energy scale uncertainty	yes
MUON_ID	energy resolution uncertainty from inner detector	yes
MUON_MS	energy resolution uncertainty from muon system	yes
MUON_SAGITTA_RESBIAS	uncertainty in the momentum scale (charge-dependent)	yes
MUON_SAGITTA_RHO	uncertainty in the momentum scale (charge-dependent)	yes
Taus		
TAUS_TRUEHADTAU_SME_TES_DETECTOR	energy scale uncertainty (detector effects)	yes
TAUS_TRUEHADTAU_SME_TES_INSITU	energy scale uncertainty (in-situ correction)	yes

TAUS_TRUEHADTAU_SME_TES_MODEL	tau-related uncertainty (MC modelling)	yes
Small-R jets		
JET_EffectiveNP_Detector	JES uncertainty: detector effects (2 components)	yes
JET_EffectiveNP_Mixed	JES uncertainty: mixed effects (3 components)	yes
JET_EffectiveNP_Modelling	JES uncertainty: modelling effects (4 components)	yes
JET_EtaIntercalibration_Modelling	uncertainties in scale calibration of forward / central jets	yes
JET_EtaIntercalibration_NonClosure_2018data	uncertainties in scale calibration of forward / central jets	yes
JET_EtaIntercalibration_NonClosure_highE	uncertainties in scale calibration of forward / central jets	yes
JET_EtaIntercalibration_NonClosure_negEta	uncertainties in scale calibration of forward / central jets	yes
JET_EtaIntercalibration_NonClosure_posEta	uncertainties in scale calibration of forward / central jets	yes
JET_EtaIntercalibration_TotalStat	uncertainties in scale calibration of forward / central jets	yes
JET_BJES_Response	flavour-related uncertainty	yes
JET_Flavour_Composition	flavour-related uncertainty	yes
JET_Flavour_Response	flavour-related uncertainty	yes
JET_JER_EffectiveNP	jet energy resolution uncertainty split into 7 parameters	yes (symm.)
JET_JER_DataVsMC_MC16	jet energy resolution uncertainty	yes (symm.)
JET_JvtEfficiency	jet-vertex-tagger efficiency uncertainty	yes
JET_Pileup_OffsetMu	Pileup uncertainty	yes
JET_Pileup_OffsetNPV	Pileup uncertainty	yes
JET_Pileup_PtTerm	Pileup uncertainty	yes
JET_Pileup_RhoTopology	Pileup uncertainty	yes
JET_PunchThrough_MC16	punch through uncertainty	yes
JET_SingleParticle_HighPt	absolute in-situ propagation of single-particle, test beam uncertainties	yes
Tracks (used in TAR jets)		
TRK_BIAS_D0_WM	d_0 residual alignment tracking uncertainties	yes (symm.)
TRK_BIAS_Z0_WM	z_0 residual alignment uncertainties	yes (symm.)
TRK_BIAS_QOVERP_SAGITTA_WM	p_T residual alignment tracking uncertainties	yes (symm.)
TRK_EFF_LOOSE_GLOBAL	tracking efficiency (loose working point) uncertainty	yes (symm.)
TRK_EFF_LOOSE_IBL	tracking efficiency (loose working point) uncertainty	yes (symm.)
TRK_EFF_LOOSE_PHYSMODEL	tracking efficiency (loose working point) uncertainty	yes (symm.)
TRK_EFF_LOOSE_PP0	tracking efficiency (loose working point) uncertainty	yes (symm.)
TRK_EFF_LOOSE_TIDE	tracking in dense environments efficiency (loose working point) uncertainty	yes (symm.)
TRK_FAKE_RATE_LOOSE	tracking uncertainties on fake rate	yes (symm.)
TRK_FAKE_RATE_LOOSE_ROBUST	tracking uncertainties on fake rate	yes (symm.)
TRK_FAKE_RATE_LOOSE_TIDE	tracking uncertainties on fake rate in dense environments	yes (symm.)
TRK_RES_D0_DEAD	tracking uncertainties associated with IP d_0 resolution	yes (symm.)

TRK_RES_D0_MEAS	tracking uncertainties associated with IP d_0 resolution	yes (symm.)
TRK_RES_Z0_DEAD	tracking uncertainties associated with IP z_0 resolution	yes (symm.)
TRK_RES_Z0_MEAS	tracking uncertainties associated with IP z_0 resolution	yes (symm.)
Tagging efficiency (using VR track jets)		
FT_EFF_EIGEN_B	b -tagging efficiency uncertainties ("BTAG_MEDIUM):	no
FT_EFF_EIGEN_C	3 components for b -jets, 4 for c -jets and 5 for light jets	no
FT_EFF_EIGEN_L		no
FT_EFF_EIGEN_extrapolation	b -tagging efficiency uncertainty on the extrapolation on high p_T -jets	no
FT_EFF_EIGEN_extrapolation_from_charm	b -tagging efficiency uncertainty on τ -jets	no
E_T^{miss}-Trigger and E_T^{miss}-Terms		
METTrigStat	trigger efficiency uncertainty	yes
METTrigSyst		yes
MET_SoftTrk_ResoPerp	track-based soft term related to transversal resolution uncertainty	yes (symm.)
MET_SoftTrk_ResoPara	track-based soft term related to longitudinal resolution uncertainty	yes (symm.)
MET_SoftTrk_ScaleUp	track-based soft term related to longitudinal scale uncertainty	yes (symm.)
MET_SoftTrk_ScaleDown	track-based soft term related to longitudinal scale uncertainty	yes (symm.)
MET_JetTrk_ScaleUp	track MET scale uncertainty due to tracks in jets	yes (symm.)
MET_JetTrk_ScaleDown	track MET scale uncertainty due to tracks in jets	yes (symm.)

9.2. Theoretical Systematic Uncertainties

The modelling uncertainty of the signal and main background process, W/Z +jets have been evaluated. Modelling of the $t\bar{t}$ process is not taken into account because of the small yield and the relatively small impact on the total uncertainty. The uncertainty is considered for the other background processes only on the total normalization. Parton distribution functions (PDF), renormalization and factorization scale used in the computation of the hard scattering process, as well as generator, underlying event and parton shower uncertainties are taken into account for the W/Z +jets background. They affect the shape and the normalization of the distributions, as well as the relative acceptance between analysis regions. The uncertainties are evaluated using a dedicated procedure that makes use of alternative samples, with different generation, parton shower and underlying event treatment. Different eigen-variations of the PDF and a different scheme for the scales is also used for this purpose.

A simplified version of the analysis is used at Monte Carlo truth level, since the interaction of the final states with the ATLAS detector factors out of this uncertainties. Acceptance effects are calculated in this simplified analysis, while shape effects using

Table 9.2.: Summary of the evaluation strategy for the theoretical uncertainties. The strategy is shown for W/Z +jets, other background processes and for the signal. The shape uncertainties are derived at reconstruction level, the acceptance with the simplified analysis using truth level Monte Carlo.

Process	Normalization	Shape (reconstructed quantities)			Acceptance (truth quantities)		
	scale and PDF	scale	PDF	generator	scale	PDF	generator
W/Z +jets	Floating	Variation scheme	Eigen-variations	Sherpa vs MG5	Variation scheme	Eigen-variations	Sherpa vs MG5
Other SM	Fixed	-	-	-	-	-	-
Signal	Fixed	Variation scheme	Eigen-variations	-	Variation scheme	Eigen-variations	-

reconstructed quantities. The structure of the analysis is preserved: three regions, 0LSR, 1LCR and 2LCR with two categories each, the merged and the intermediate. Objects are defined as close as possible to the reconstructed ones, but at truth level. This leads to some discrepancies, which however do not affect the size of the phase-space probed. TAR jets are substituted with reclustered truth jets of $R = 0.8$ using $R = 0.4$ truth jets as input, with trimming parameter as in the standard object (5%). This is done in order to avoid the ill-defined treatment of tracks at truth level.

A summary of the strategy for the evaluation of the theoretical systematic uncertainties is given in Table 9.2. The complete list of theoretical systematic uncertainties considered can be found in Table 9.3.

Uncertainties on the signal

A subset of theoretical uncertainties is considered on the signal: the variation in the nominal renormalization and factorization scales (μ_R and μ_F) and different PDF sets. These variations are evaluated on their impact on the invariant mass distribution of the Dark Higgs candidate.

The uncertainty on the scales is calculated with:

$$\text{scale uncertainty} = \max_i \left(1 - \frac{\sigma_i}{\sigma_0} \right)$$

where σ_i is the acceptance of the i -th variation and σ_0 is the nominal acceptance, calculated as:

$$\sigma_i = \sum_{bins} (\text{histogram with variation } i)$$

The variations are generated with respect to nominal (μ_R, μ_F) are: $(\mu_R, 2 \times \mu_F)$, $(\mu_R, 1/2 \times \mu_F)$, $(2 \times \mu_R, \mu_F)$, $(1/2 \times \mu_R, \mu_F)$, $(2 \times \mu_R, 2 \times \mu_F)$ and $(1/2 \times \mu_R, 1/2 \times \mu_F)$.

The uncertainties on the PDF set are calculated taking the standard deviation over the alternative sets:

$$\text{PDF uncertainty} = \sqrt{\frac{\sum_{i=1}^{N_{\text{variations}}} \sigma_i^2 - \left(\sum_{i=1}^{N_{\text{variations}}} \sigma_i\right)^2}{N_{\text{variations}} - 1}}$$

Uncertainty on the background

The main background processes of the analysis, W/Z +jets, require a dedicated scrutiny in order to assess and cover possible mis-modelling of the MC samples. Similarly as the signal, scales and PDF set variations are calculated. Moreover, a *2-point* variation systematic uncertainty is computed, encompassing the difference of the nominal MC sample with alternative ones. Acceptance effects are calculated as well, for the variations considered, between categories and signal regions.

The nominal W/Z +jets MC samples summarized on Table 7.1 are generated with MadGraph5 interfaced with Pythia8 (MGPy8) for parton shower and underlying event description. The QCD treatment is done at LO, in contrast with the alternative Sherpa setup of two partons, which is done at NLO. The NNPDF2.3 LO PDFs are used. In order to factor out mis-modelling of the p_T of the vector boson, the alternative samples are reweighted in order to achieve the same E_T^{miss} distribution with respect to the nominal sample. This effectively decouples the p_T mis-modelling from the effects that impact the invariant mass distribution of the Dark Higgs candidate.

The shape systematic uncertainties for the background are calculated as the envelope of the fluctuations with respect to the nominal samples.

Acceptance Acceptance uncertainties are calculated by means of ratios (R) between the different regions or categories. They are introduced to take into account the different acceptance in the analysis regions that is observed in nominal and alternative MC samples. These additional uncertainties allow the fit to control the potential mis-modelling of the relative acceptances. The acceptance ratio (ACC_R) is defined as:

$$\text{ACC}_R(i^{\text{th}} \text{ variation})_{\text{region}} = \frac{N(i^{\text{th}} \text{ variation, region})}{N(\text{nominal, region})}$$

where N denotes the acceptance in one region, and *region*, the signal region or category, e.g $N(i^{\text{th}} \text{ variation, merged OLSR})$ is the acceptance of the merged OLSR where the i -th variation (such as scale, PDF or two point) is applied. The systematic uncertainty is the envelope of the ACC_R fluctuations inclusively in E_T^{miss} . Tables with the acceptances uncertainties are provided in the Appendix C.

Table 9.3.: Qualitative summary of the theoretical systematic uncertainties that are considered in this analysis. The name of the systematic uncertainties and their description is shown. The smoothing is used to mitigate the impact that limited MC statistics has on the fit.

Systematic uncertainty	Short description	smoothing
Uncertainties on the normalization		
norm_zjets	floating normalization for Z+jets	-
norm_wjets	floating normalization for W+jets	-
ttbarNorm	combined $t\bar{t}$ cross-section uncertainty due to scale and PDF $(0.05)^2$	-
stopsNorm	combined top quark (s-channel) cross-section uncertainty due to scale and PDF $(0.037)^3$	-
stoptNorm	combined top quark (t-channel) cross-section uncertainty due to scale and PDF (0.039)	-
stopWtNorm	combined top quark ($W t$) cross-section uncertainty due to scale and PDF (0.054)	-
WWNorm	W W cross-section uncertainty (0.25)	-
WZNorm	W Z cross-section uncertainty (0.26)	-
ZZNorm	Z Z cross-section uncertainty (0.20)	-
HiggsNorm	VHbb cross-section uncertainty (0.22)	-
Uncertainties on the acceptance		
(decorrelated in merged, intermediate)		
AccZjets2Pt	generator comparison uncertainty on acceptance in categories	-
AccZjetsPDF	PDF uncertainty on acceptance in categories	-
AccZjetsScale	scale uncertainty on acceptance in categories	-
AccWjets2Pt	generator comparison uncertainty on acceptance in categories	-
AccWjetsPDF	PDF uncertainty on acceptance in categories	-
AccWjetsScale	scale uncertainty on acceptance in categories	-
Uncertainties on the shape		
(decorrelated in merged, intermediate)		
zjets_SCALE	scale shape uncertainty (envelope of 7-point variation)	yes (symm.)
zjets_NNPDF30_nlo_as_0118AlphaS	combined PDF (standard deviation method) + α_s shape uncertainty	yes
zjets_gen	generator comparison shape uncertainty	yes (symm.)
wjets_SCALE	scale shape uncertainty (envelope of 7-point variation)	yes (symm.)
wjets_NNPDF30_nlo_as_0118AlphaS	combined PDF (standard deviation method) + α_s shape uncertainty	yes
wjets_gen	generator comparison shape uncertainty	yes (symm.)

10. Statistical Framework

This Chapter introduces core concepts that are needed for the presentation and the interpretation of the result of this analysis. For a more detailed and precise introduction, the reader can refer to standard textbooks e.g. [114] [115] [116].

In the previous pages the definition of the signal and control regions was shown, together with all the techniques optimized to suppress the background yield as much as possible, and at the same time to maximize the signal. Since measured data cannot be labelled as 'signal' or 'background', the question to answer at this stage is whether it can be concluded from the data that, yes, the data really can be explained only if the signal is present, or no, the background alone can reproduce the data. This is known in statistics as (frequentist) hypothesis testing: the desire is to know which of the two hypothesis, the background only H_b (also called *null hypothesis*) or the background plus signal H_{s+b} hypothesis (*alternative hypothesis*) is more consistent with data. Similarly to other searches for new physics at the high energy frontier, the statistical formalism that the $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ analysis utilizes is the binned profile likelihood.

Staying as general as possible, suppose the outcome of an experiment, x , to be represented in a histogram with N bins. The *hypothesis* is a statement on the outcome of the experiment. The experiment is performed via physical measurements, which are affected by systematics uncertainties as described in the previous Chapter, such as the jet energy resolution (JER) and so on. The entity of this uncertainties is evaluated in ancillary measurements like in-situ di-jet balance as in the case of the JER. All these uncertainties are encoded in the statistical treatment as Nuisance Parameters (NP), indicated as θ . Finally, the Parameter of Interest (POI) is a convenient multiplicative factor that parametrizes the hypotheses, in our case the *signal strength* μ . The signal strength refers to the presence of the signal or not: $\sum_{i=1}^N x_i = n^{\text{obs}} = \mu \times s + b$. Here n^{obs} is the number of events observed; s and b are the signal and background yields, respectively. This translates into two extremes: for $\mu=1$ the alternative hypothesis H_{s+b} is obtained, and for $\mu=0$ the null hypothesis H_b . The conditional probability of observing the data given the signal strength and the NP, known as *likelihood*, can be written as:

$$L(x|H_{s+b}) = L(x|\mu, \theta)$$

The likelihood can be written as the product of few factors. $L_{measurement}(x|\mu, \theta)$ represents the per-bin probability of observing the number of events, following therefore a Poisson distribution: $L_{measurement}(x|\mu, \theta) \propto \text{Pois}(x|\mu, \theta)$ ¹. The ancillary measurements that constrain the NPs are written as $C(\theta)$ and do not depend on data here. They are expressed as a Gaussian, Poisson or Log-Norm distribution.

With the likelihood now defined, still remains the question of how to decide between the null hypothesis $H_b, \mu = 0$, and the alternative hypothesis $H_{s+b}, \mu = 1$. The Neyman-Pearson lemma assures that the likelihood ratio $\frac{L(x|H_b)}{L(x|H_{s+b})}$ is the most potent discriminator when testing the null hypothesis versus the alternative hypothesis. The profile likelihood ratio (PLR), $\Lambda(\mu)$, is therefore defined as:

$$\Lambda(\mu) = \frac{L(x|\mu, \hat{\theta}(\mu))}{L(x|\hat{\mu}, \hat{\theta}(\mu))}$$

where data was let free to choose the optimal values for μ and θ ; these are in fact estimated (profiling) as the ones maximizing the likelihood. $\hat{\theta}(\mu)$ is the maximum likelihood estimator for the NPs, θ , for fixed μ (conditional), while $\hat{\mu}$ and $\hat{\theta}(\mu)$ are the maximum likelihood estimators for both μ and θ (unconditional). The PLR ranges between zero and one for any given value of μ , and values closer to one represent higher compatibility between the data and the μ -parametrized hypothesis.

With the PLR at hand, the *test statistic* can be defined, which will be used to perform the hypothesis test²:

$$t_\mu = -2\ln(\Lambda(\mu))$$

The probability density function of t_μ under the null and alternative hypothesis, $g(t|H_b)$ and $g(t|H_{s+b})$ allows establishing a further quantitative criterion for accepting or rejecting the hypothesis. Generally for the HEP applications, the probability densities for t_μ are hard to compute, therefore an asymptotic analytic form is used instead. According to a theorem from Wilks [117] and Wald [118], if the signal strength is Gaussian distributed, then the test statistic follows the χ^2 distribution in the large sample limit. In case of low number of events, the asymptotic approximation is not valid any more, especially at larger values for t_μ [119], in which case Monte Carlo simulations have to be used. In the *Asimov dataset*, all observed quantities are set to

¹Or more explicitly $L_{measurement}(x|\mu, \theta) \propto \prod_{i=1}^N \frac{(\mu s_i(\theta) + b_i(\theta))^{x_i}}{x_i!} e^{-(\mu s_i(\theta) + b_i(\theta))}$.

²Having care to set $t_0 = 0$ if $\hat{\mu} < 0$ and $t_\mu = 0$ if $\hat{\mu} > \mu$. This is because in the first case the discrepancy points to the presence of systematic errors, and in the second case is irrelevant to limit setting.

their expected values neglecting their statistical uncertainties, which are obtained via the asymptotic formulae. It is used for both discovery and exclusion. The probability α (β) of rejecting the null (alternative) hypothesis H_b (H_{s+b}) given that H_b (H_{s+b}) is true is called error of first (second) kind and depends on the observed test statistic t_μ^{obs} . α and β are given by:

$$\alpha = \int_{t_\mu^{cut}}^{\infty} g(t_\mu|H_b)dt, \quad \beta = \int_{-\infty}^{t_\mu^{cut}} g(t_\mu|H_{s+b})dt$$

For a good test statistic, α (background efficiency) should be as small as possible and $1-\beta$ (signal efficiency) as large as possible. In this case t_μ^{cut} is a predefined value.

The *p-value* quantifies the disagreement between the results of the experiment and the hypothesis taking in consideration the observed level of the test statistic. This is indicated p_b and p_{s+b} for the null and alternative hypothesis respectively:

$$p_b = \int_{t_\mu^{cut}}^{\infty} g(t_\mu|H_b)dt, \quad p_{s+b} = \int_{-\infty}^{t_\mu^{cut}} g(t_\mu|H_{s+b})dt$$

Equivalently, the p-values can be related to the number of standard deviations using a one-sided Gaussian tail until the p-value with the variable Z or significance: $Z = \Phi(1 - p)$, with Φ being the inverse of the cumulative Gaussian distribution. Historically, rejection of the null hypothesis in the HEP community is established by the significance: $Z = 5$ (five sigma) is referred to as discovery. This corresponds to a very small p-value of $\sim 3 \times 10^{-7}$. Instead, the null hypothesis is accepted if the p-value is smaller than 0.05 (known as 95% *confidence level* or CL), which corresponds roughly to $Z = 1.6$.

Using only p_{s+b} (CL_{s+b} exclusion), limits on the model may in fact be established even if the experiment is not sensitive, since the distributions $g(t_\mu|H_b)$ and $g(t_\mu|H_{s+b})$ significantly overlap. Instead, limits are set using CL_s :

$$CL_s = \frac{p_{s+b}}{1 - p_b}$$

A background-like fluctuation would result in large p_b and a false exclusion is prevented. Consistently, the alternative hypothesis is rejected at 95% confidence level when $CL_s < 0.05$.

Fit Model The fit of the statistical model to the data is performed using the events as outlined in the previous Chapters. The events in the 0LSR are binned in E_T^{miss} for all categories, (200-300] GeV, (300-500] GeV and (500 GeV - 13 TeV) with the first bin considered only for the intermediate category. For the 1LCR and 2LCR the $E_T^{\text{miss},\text{no}\mu}$ and $p_T^{\ell\ell}$ is used instead. The summary for the binning can be found on Table 8.3. Furthermore, the distributions are binned, only in the 0LSR, in the mass of the Dark Higgs candidate m_s . This is done between (100-400) GeV, in steps of 20 GeV (merged, and intermediate $E_T^{\text{miss}} > 500$ GeV) and steps of 10 GeV (intermediate $E_T^{\text{miss}} < 500$ GeV). The signal is extracted using a shape fit to the m_s spectrum. The V +jets background is constrained using the control regions. For this reason, and since the background is smoothly falling, the shape of the mass distribution is not taken into account, but only its normalization. In this way the introduction of artificial constraints on experimental uncertainties is avoided. In addition to the signal strength μ , the model contains also the free floating background normalization for the W +jets and Z +jets. In this way the CRs and SR are used to constrain simultaneously the background level in the analysis. Finally, the Nuisance Parameters derived from ancillary measurements are used to constrain the experimental and systematic uncertainties that were listed in the previous Chapter. The statistical uncertainty on the Monte Carlo simulation is taken into account with a separate NP.

11. Results

The results presented in this Chapter refer to the search for a pair of hadronically decaying vector bosons plus Dark Matter using 139 fb^{-1} of data taken with the ATLAS detector during Run 2. The analysis has been outlined in the previous pages, with the objects and selection strategy used. The TAR jet is a novel technique introduced in an ATLAS analysis for the first time, and is employed in the two most sensitive categories, the merged and the intermediate. These results are evaluated using a binned profiled likelihood approach as explained in Chapter 10. First, they are statistically evaluated in the background only hypothesis, using the test statistics t_0 . Then, exclusion limits are set under the signal plus background hypothesis using the CL_s method.

11.1. Fit Data to Model

The binned profile likelihood fit was investigated in order to make sure that the statistical evaluation is sound. One of the most important cross-check on the fit is done on the NPs. The value of the NP that maximizes the likelihood, θ_{fit}^1 , can in general be different from the prior value assigned from the ancillary measurements or the theoretical predictions, θ_0 . The relative difference with respect to the prior uncertainty ($\Delta\theta_0$), is called *pull*: $pull = (\theta_{fit} - \theta_0) / \Delta\theta_0$. The statistical model should not show NPs significantly different from their prior values, or in other words, the pulls should be small. Moreover, the uncertainties on the NPs after the fit, $\Delta\theta_{fit}$, can be compared to the one assigned by the prior and is called *constrain*. If the $\Delta\theta_{fit} < \Delta\theta_0$, this would imply that the fit is able in principle to measure the NP parameter better than the ancillary measurement or theoretical prediction, which should be treated with care. The pulled and constrained NPs are thoroughly checked in order to make sure that the fit setup is solid.

As it can be seen in Figure 11.1a, the NP pulls and constraints are shown, with an Asimov dataset is used instead of observed data for the conditional $\mu = 0$ fit. The pulls are centred around zero as it is expected by construction of the Asimov data. In Figure 11.1b, observed data from the full Run 2 is used for the $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ analysis. As it can be appreciated, most pulls are centred around zero. The value of

¹It is $\hat{\theta}$ or $\hat{\theta}$ in case of conditional or unconditional fit.

mass hypothesis. The fit is repeated as many times as the NPs considered, excluding the NP under exam and evaluating its impact in the final signal yield uncertainty.

The results can be seen in Figure 11.2, for $(m_{Z'} = 1 \text{ TeV}, m_s = 235 \text{ GeV})$ on the left and for $(m_{Z'} = 1 \text{ TeV}, m_s = 310 \text{ GeV})$ on the right, for the 20 most important systematics. The NPs that have the highest impact on the analysis are the theoretical modelling uncertainties on the shape of the main background process, the Z +jets. Other important systematics are the jet energy resolution on the Small- R jets, the b-tagging efficiency and the flavour composition. The tables with all the systematic uncertainties were shown in Chapter 9.

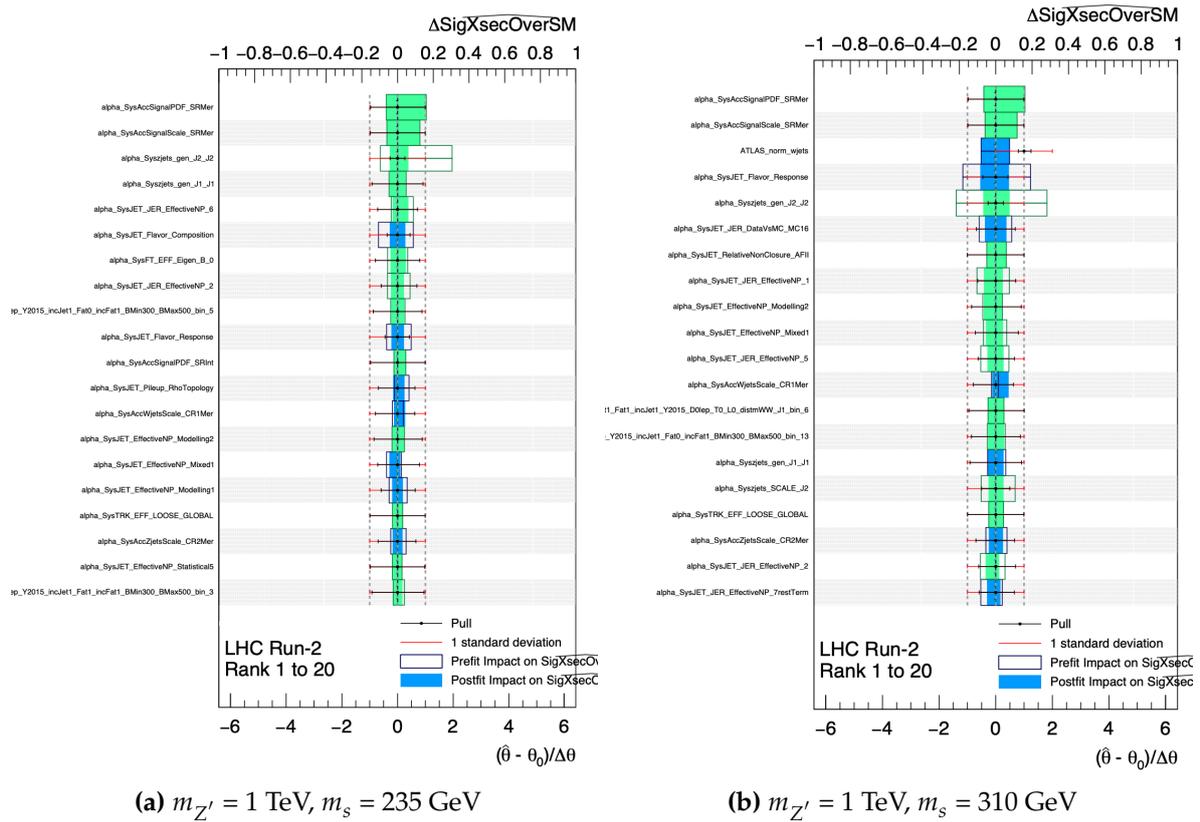


Figure 11.2.: NPs ranking plot based on the Asimov data set generated with $\mu = 1$ for different signal mass hypothesis (in caption). The filled boxes represent the $\sigma_{\hat{\mu}}/\hat{\mu}$ on pre-fit and post fit (sale on top x-axis). The pull of each NP is also overlaid (scale on bottom x-axis). The ranking is ordered based on the post-fit impact on the fitted value of μ , and are displayed from top (most important) downwards.

11.1.1. Post-fit Distribution

The post-fit distributions are shown on Figure 11.3 for the 0LSR, on Figure 11.4 for the 1LCR and on Figure 11.5 for the 2LCR, in all the categories considered.

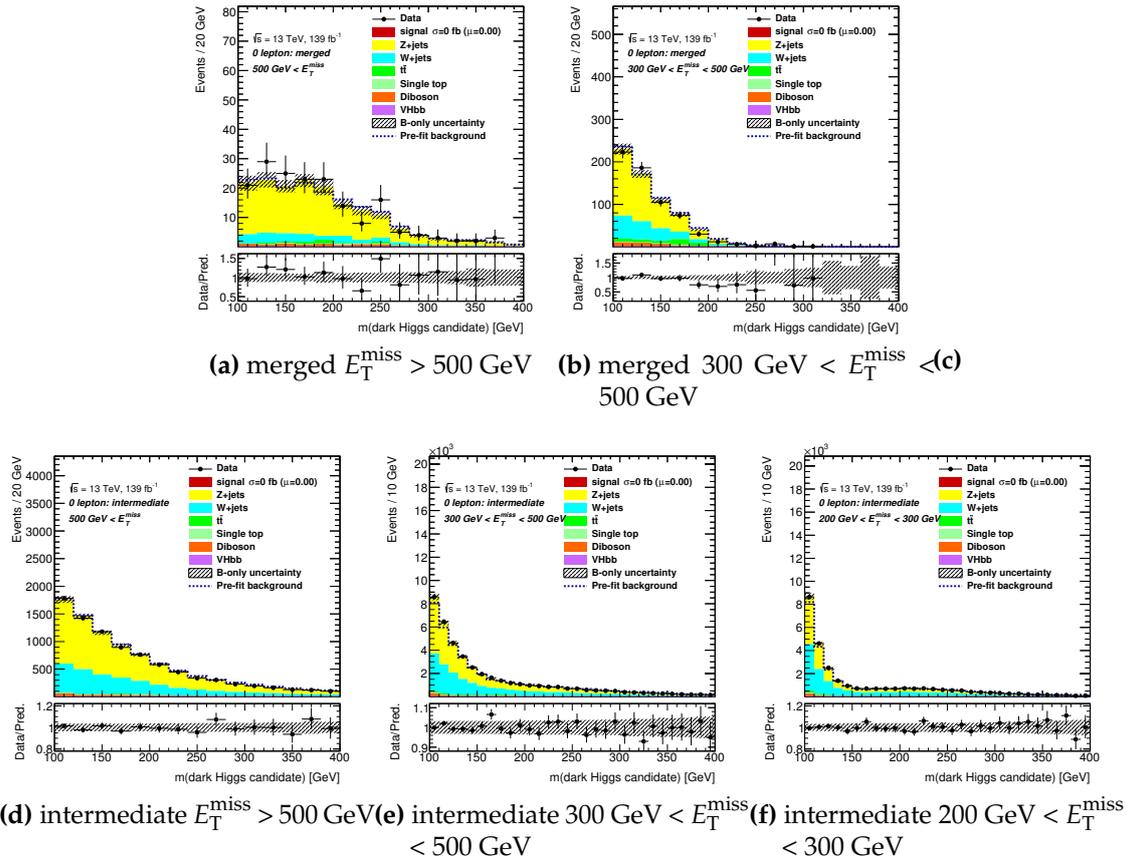


Figure 11.3.: MC distributions for the background model and overlaid data in the 0 lepton signal region for theory prediction (post-fit). The bottom panel shows the relative uncertainty on the background prediction. The hatched band indicates the pre-fit uncertainty on the total background prediction.

The detailed event yield tables for all the background processes and for data are shown in Appendix A.8, both pre-fit A.8.1 and post-fit A.8.2.

11.2. Experimental Search Significance

The *expected significance* is a way to quantify the experimental sensitivity of the search using the Asimov dataset. The expected significance for the $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$

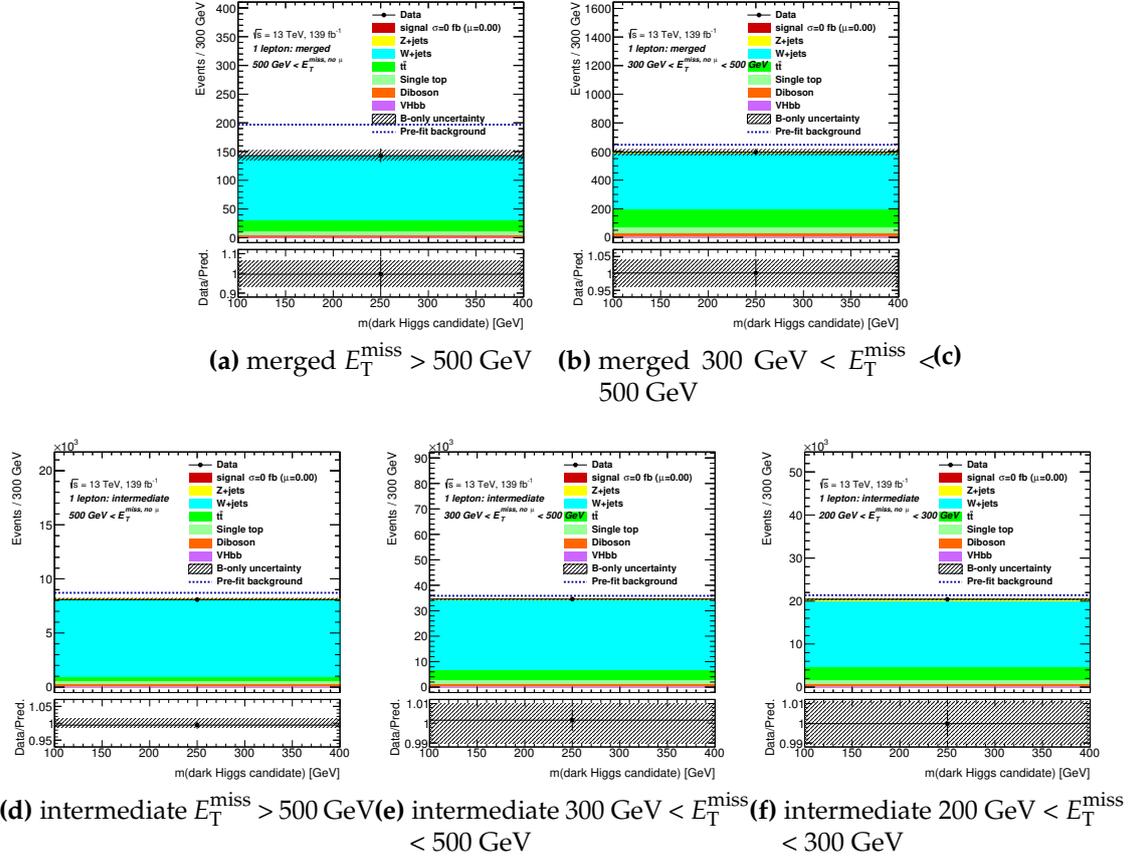


Figure 11.4.: Distributions showing the background model after the conditional $\mu = 0$ fit and overlaid data in the 1 lepton control region (post-fit). The hatched band indicates the post-fit uncertainty on the total background prediction. The dotted-line indicates the theory prediction (pre-fit) for the total background.

analysis targeting the Dark Higgs model using 139 fb^{-1} of data is shown on the left of Figure 11.6. As it can be seen, the ATLAS detector is sensitive to this exotic scenario in a considerable area of the parameter space, from around 0.5 TeV to around 1.7 TeV in the mass of the massive gauge boson, and from 160 GeV to around 260 GeV in the mass of the Dark Higgs boson. The expected sensitivity is higher in the scenario with $m_{Z'} = 1 \text{ TeV}$, which has the benefits of a moderate cross-section, and favours Lorentz boosted topologies with reduced backgrounds. For smaller $m_{Z'}$, the Dark Higgs is not energetic enough, and events migrate to the less sensitive intermediate category. The much reduced cross-section for higher $m_{Z'}$ affects the sensitivity more than the benefits of the shift of signal events to the sensitive merged category.

On the right of Figure 11.6, the observed significance for $he E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ search is presented. In all the parameter space, this is below 1.3.

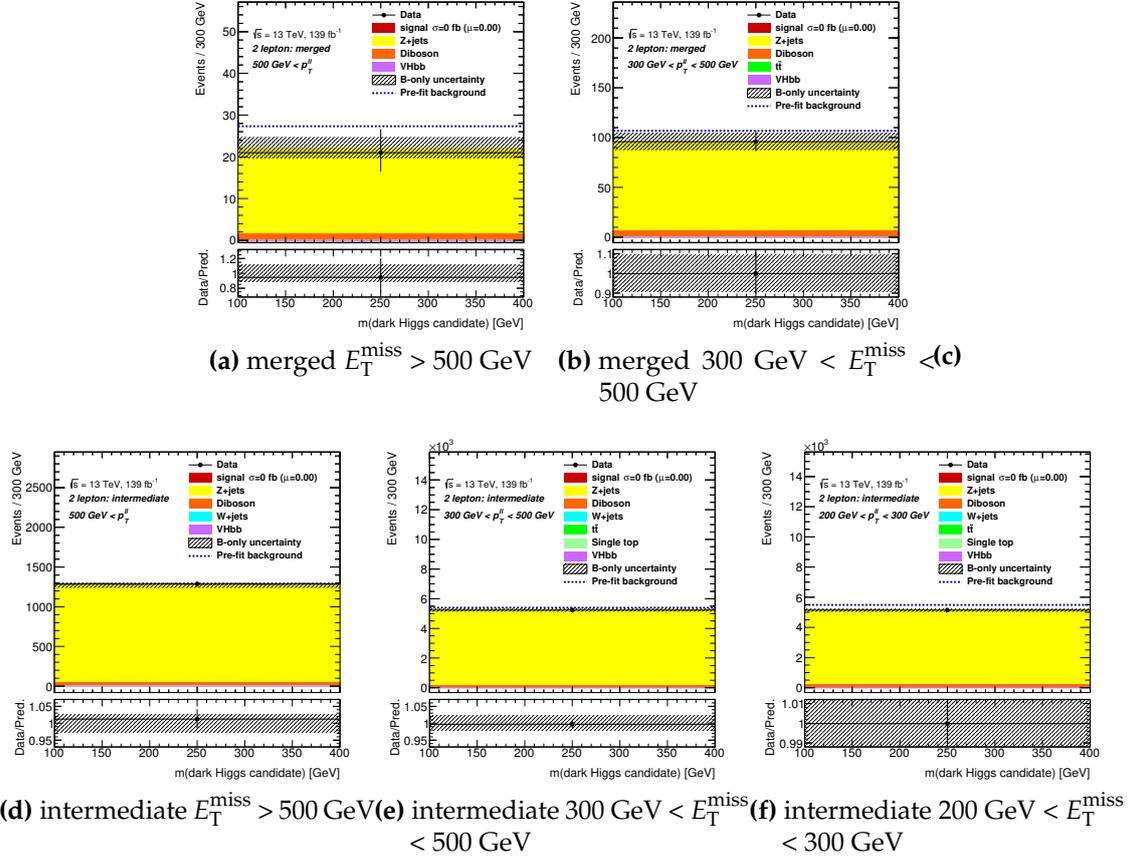


Figure 11.5: Distributions showing the background model after the conditional $\mu = 0$ fit and overlaid data in the 2 lepton control region (post-fit). The hatched band indicates the post-fit uncertainty on the total background prediction. The dotted-line indicates the theory prediction (pre-fit) for the total background.

The $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ sensitivity can be compared to the one estimated using the generic limits extrapolated from the $E_T^{\text{miss}} + H(b\bar{b})$ [97] and $E_T^{\text{miss}} + V(\text{had})$ [98] analyses, showing remarkable agreement. It was shown in the previous Sections in Figure 7.1.

The breakdown for the separate decay scenarios and for each of the analysis categories are shown in the Appendix A.9.

No evidence for New Physics in the Dark Higgs scenario is found.

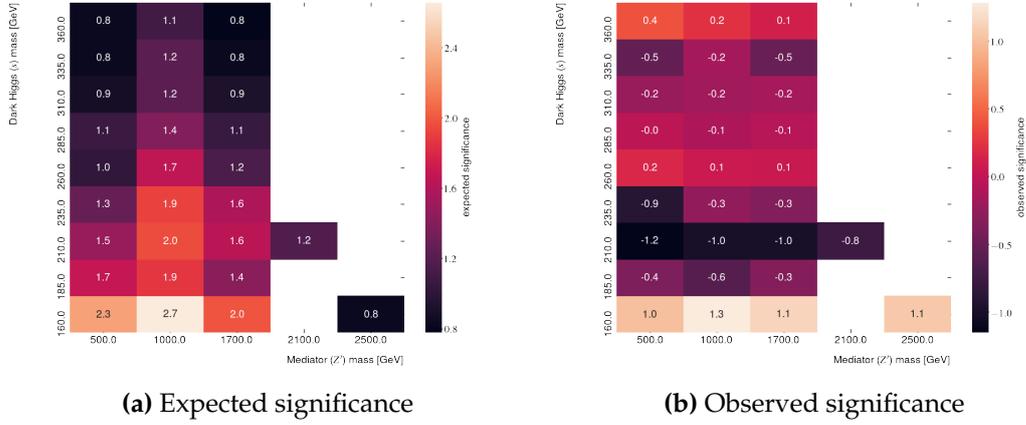


Figure 11.6: On the left, the expected significance (z-axis) on the signal strength μ for Dark Higgs signals in the combined VV decay modes with fixed $m_\chi = 200$ GeV is shown. On the right, the observed significance is displayed for the same decay mode and m_χ . On x-axis, the mass of the Z' and on the y-axis the mass of the Dark Higgs boson can be read.

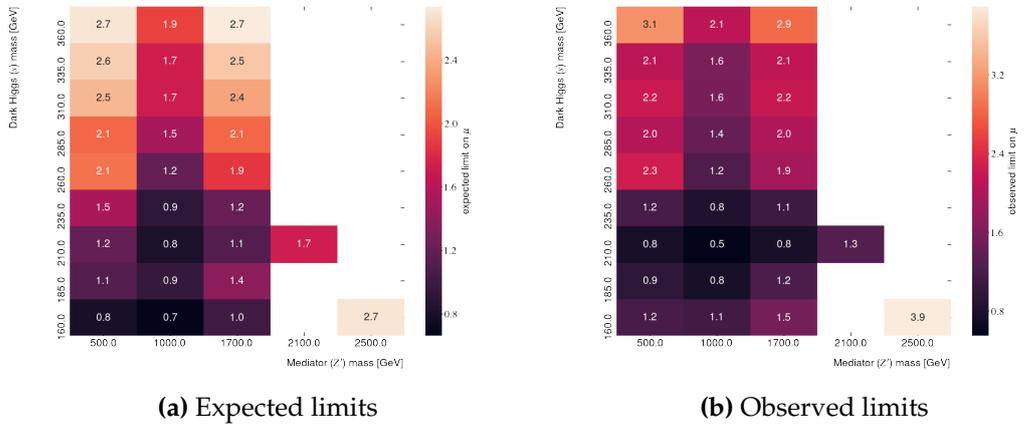


Figure 11.7: On the left, the expected limits (z-axis) on the signal strength μ for Dark Higgs signals in the combined VV decay modes with fixed $m_\chi = 200$ GeV are shown. On the right the, observed limits are displayed for the same decay mode and m_χ . On x-axis, the mass of the Z' and on the y-axis the mass of the Dark Higgs boson can be read.

11.3. Exclusion Limits on the Dark Higgs Model

Since no significant excess for the Dark Higgs model over the SM background is found, exclusion limits are set. Limits are placed at 95% confidence level (CL_s) on the Dark

Higgs model, as explained in Chapter 10. Expected and observed limits are shown on Figure 11.7. The limits are set in the parameters space and on the signal strength μ , as shown in Figure 11.8 as a function of m_s . Points are experimentally excluded if their limits on the signal strength are below one. The analysis is able to exclude points up to $m_s \lesssim 240$ GeV. The exclusion limits are also shown for the different $m_{Z'}$ hypothesis, $m_{Z'} = 1.7$ TeV and $m_{Z'} = 0.5$ TeV, in the Appendix A.10. The detailed tables of the exclusion limits for the different masses of the Dark Higgs scenario are shown as well in the Appendix A.10.1.

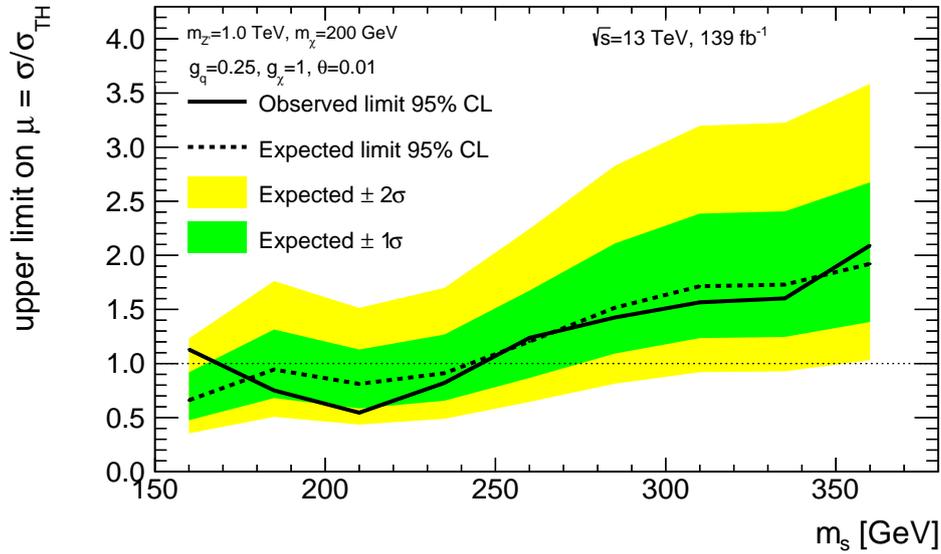


Figure 11.8.: 95% CL_s expected and observed limits on the signal strength μ as dashed and solid line, respectively, for the $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ analysis in the signal hypothesis with $m_{Z'} = 1$ TeV. The band around the expected limits shows the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties. The horizontal line denotes $\mu = 1$. Masses are excluded if the limits are below this line.

The limits shown can be compared to those already presented in Chapter 3.2 based on the reinterpretation of the $E_T^{\text{miss}} + H(b\bar{b})$ analysis. The limits are presented together in Figure 11.9. Since the branching ratio of $s \rightarrow b\bar{b}$ quickly decreased around 160 GeV, the exploration of the Dark Higgs model can be done only in the $s \rightarrow VV$ final state, as in this $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ analysis. For the first time the region $m_s > 160$ GeV is investigated, covering a large part of the available parameter-space.

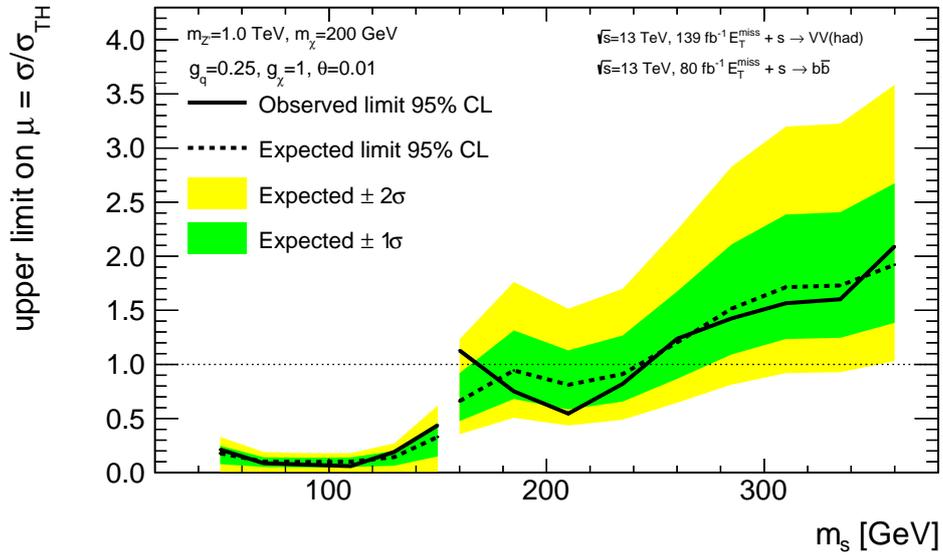


Figure 11.9.: 95% CL_s expected and observed limits on the signal strength μ as dashed and solid line, respectively. The $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ analysis in the signal hypothesis with $m_{Z'} = 1$ TeV is shown in the range (160-360) GeV. The $E_T^{\text{miss}} + s \rightarrow b\bar{b}$ as reinterpreted from the $E_T^{\text{miss}} + H(b\bar{b})$ analysis is shown in the range (50-150) GeV. The band around the expected limits shows the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties. The horizontal line denotes $\mu = 1$. Masses are excluded if the limits are below this line.

12. Conclusions and Outlook

The production at the LHC of states connected to the Dark Sector provides a complementary approach to the quest for Dark Matter. The highest-than-ever energies available, the large amounts of integrated luminosities collected by the ATLAS detector, and the more refined reconstruction techniques make possible to experimentally explore beyond the SM theories that include Dark Matter particles. The final states expected in such scenarios comprise highly energetic objects recoiling against a large momentum imbalance. The reconstruction challenge of energetic, hadronically decaying Lorentz boosted objects is presented. These challenges can be overcome combining the measurement of the tracker and calorimeter systems. TAR jets are a unique tool in the ATLAS collaboration, which were designed to improve the signal efficiency and background rejection. This can be achieved via its flexibility of the reclustering radius and the superior track-based jet substructure. They have been demonstrated to outperform the standard reconstruction techniques for energetic objects [88].

The Dark Higgs scenario introduces an Higgs-like particle, s , responsible for the generation of the masses in the Dark Sector [50]. The constraints on this model are relaxed if the s is lighter than Dark Matter. Conventional searches would remain insensitive, and the Dark Sector secluded [52]. The Dark Higgs can decay to different SM states: for $m_s \lesssim 160$ GeV the branching ratio is dominated by $s \rightarrow b\bar{b}$, while for $m_s \gtrsim 160$ GeV it is dominated by $s \rightarrow VV$. In the former case, the $E_T^{\text{miss}} + H(b\bar{b})$ analysis [97] was reinterpreted under the Dark Higgs hypothesis [4]. To address the latter case, a new analysis has been performed and presented in this thesis. The $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ targets the unexplored $s \rightarrow VV$ in the fully hadronic final state recoiling against Dark Matter, using 139 fb^{-1} of proton-proton collision at 13 TeV of centre-of-mass energy. TAR jets have been used in the most sensitive categories to reconstruct the Dark Higgs candidate, exploiting the peculiar substructure of VV system. No significant excess is found in the observed data over the Standard Model; the upper limits on the production cross-section of $E_T^{\text{miss}} + VV$ in the fully hadronic final state for the Dark Higgs scenario are set at 95% confidence level.

To achieve even stronger limits, a search in the semi-leptonic final state of the $s \rightarrow VV$ system is already at advanced stage in the ATLAS collaboration; the combination with the fully hadronic search is envisioned. The limits set by $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$

extend up to ~ 250 GeV in m_s . After this point, the cross-section decreases because of new decay channels that become kinematically allowed. The Dark Higgs model can be further explored at higher masses targeting its decay mode, sizeable starting from $m_s \gtrsim 250$ GeV: the $s \rightarrow HH \rightarrow b\bar{b}b\bar{b}$. This final state can achieve, despite the low signal cross-sections, strong SM background suppression that is guaranteed by the presence of four b quark jets. For $m_s \gtrsim 350$ GeV also the process $s \rightarrow t\bar{t}$ becomes relevant. Other searches for beyond the SM physics that comprise energetic hadronically decaying objects, are currently considering or adopting the TAR jet because of flexibility and the high sensitivity that TAR-based taggers offer.

Part IV.
Appendices

A. Analysis Appendix

A.1. Constraints on the Dark Higgs Model

In Figure A.1, a summary plot of the ATLAS analyses that have put limits on the coupling between the Z' boson and its coupling to quarks g_q .

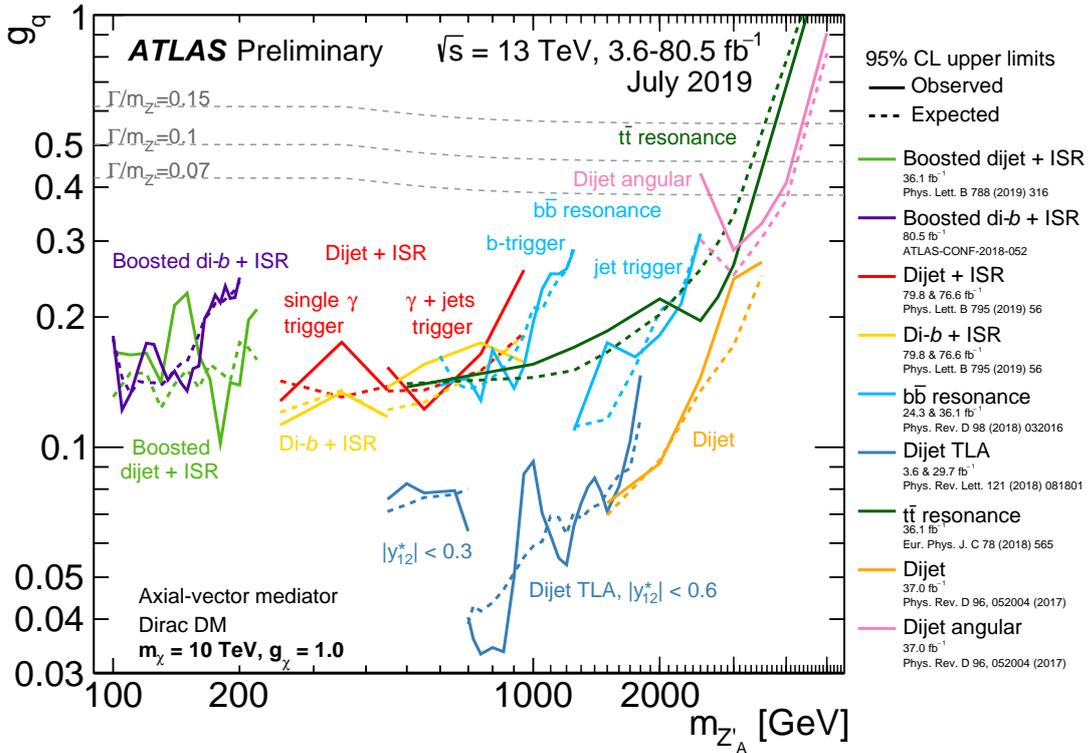


Figure A.1.: Summary of limits on the coupling between Z' boson and quarks for various Z' boson mass $m_{Z'}$ placed by ATLAS di-jet searches [120].

Associated production of Dark Matter and a hadronically decaying vector boson search [98] can also be interpreted in term of the Dark Higgs model [51].

As already stated in the body of the thesis, these constraints hold for fixed benchmark values for the coupling g_χ . For raising values of g_χ , however, the limits become less and less stringent as the process $qq \rightarrow Z' \rightarrow qq$ loses sensitivity with respect to $qq \rightarrow Z' \rightarrow X + DM$.

A.2. Triggers used

Table A.1 summarizes the details of the trigger configuration used for the $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ analysis.

A.3. Generic Limits Procedure

When requesting the MC signal points, it is important to make sure that the parameter choice (in this case the mass of the Dark Higgs and the mass of the Z' mediator) is covering the final exclusion contour, and the search is appropriately optimized for the signal's kinematic properties (which can vary depending on the model's parameters). If this is not the case, new MC request have to be requested, making the signal interpretation of the analysis unnecessary slow. There are two ways to make the estimate for the signal grid:

- Use particle-level signal and detector-level background MC.
- Use generic limits, provided the signature is already covered by an analysis that provided the limits.

Sensitivity estimate using generic limits (GL) is a technique that makes use of GL in order to establish the MC signal grid.

This is evaluated in three steps:

1. The cross-section for each parameter choice is calculated with Madgraph.
2. The cross-section is binned in E_T^{miss} (this is necessary since the GL are provided in bins of E_T^{miss}).
3. A bin-wise reconstruction-level sensitivity is constructed multiplying by $A * \epsilon / \sigma_{GL}$, where A is the acceptance, ϵ is the efficiency and σ_{GL} is the GL on the cross-section.
4. To get the full reconstruction-level sensitivity all the E_T^{miss} bins are then again summed together.

Table A.1.: Summary table of triggers used in 2015, 2016, 2017, and 2018 data.

Period	0 lepton	1 lepton	2 lepton
2015	HLT_xe70_MHT	HLT_xe70_MHT	HLT_E24_LHMEDIUM_L1EM20VH OR HLT_E60_LHMEDIUM OR HLT_E120_LHLOOSE OR HLT_MU20_ILOOSE_L1MU15 OR HLT_MU50
2016 (A)	HLT_xe90_MHT_L1XE50	HLT_xe90_MHT_L1XE50	HLT_E24_LHTIGHT_NOD0_IVARLOOSE OR HLT_E60_LHMEDIUM_NOD0 OR HLT_E60_MEDIUM OR HLT_E300_ETCUT OR HLT_E140_LHLOOSE_NOD0 OR HLT_MU24_ILOOSE OR HLT_MU24_IVARLOOSE OR HLT_MU40 OR HLT_MU50
2016 (B-D3)	HLT_xe90_MHT_L1XE50	HLT_xe90_MHT_L1XE50	HLT_E24_LHTIGHT_NOD0_IVARLOOSE OR HLT_E60_LHMEDIUM_NOD0 OR HLT_E60_MEDIUM OR HLT_E300_ETCUT OR HLT_E140_LHLOOSE_NOD0 OR HLT_MU24_IVARMEDIUM OR HLT_MU24_IMEDIUM OR HLT_MU50
2016 (D4-E3)	HLT_xe100_MHT_L1XE50	HLT_xe100_MHT_L1XE50	HLT_E26_LHTIGHT_NOD0_IVARLOOSE OR HLT_E60_LHMEDIUM_NOD0 OR HLT_E60_MEDIUM OR HLT_E300_ETCUT OR HLT_E140_LHLOOSE_NOD0 OR HLT_MU24_IVARMEDIUM OR HLT_MU24_IMEDIUM OR HLT_MU26_IVARMEDIUM OR HLT_MU26_IMEDIUM OR HLT_MU50
2016 (F1)	HLT_xe100_MHT_L1XE50	HLT_xe100_MHT_L1XE50	HLT_E26_LHTIGHT_NOD0_IVARLOOSE OR HLT_E60_LHMEDIUM_NOD0 OR HLT_E60_MEDIUM OR HLT_E300_ETCUT OR HLT_E140_LHLOOSE_NOD0 OR HLT_MU26_IVARMEDIUM OR HLT_MU26_IMEDIUM OR HLT_MU50
2016 (F2-)	HLT_xe110_MHT_L1XE50	HLT_xe110_MHT_L1XE50	HLT_E26_LHTIGHT_NOD0_IVARLOOSE OR HLT_E60_LHMEDIUM_NOD0 OR HLT_E60_MEDIUM OR HLT_E300_ETCUT OR HLT_E140_LHLOOSE_NOD0 OR HLT_MU26_IVARMEDIUM OR HLT_MU26_IMEDIUM OR HLT_MU50
2017 (B-D5)	HLT_xe110_PUFT_L1XE55	HLT_xe110_PUFT_L1XE55	HLT_E26_LHTIGHT_NOD0_IVARLOOSE OR HLT_E60_LHMEDIUM_NOD0 OR HLT_E140_LHLOOSE_NOD0 OR HLT_E300_ETCUT OR HLT_MU26_IVARMEDIUM OR HLT_MU50
2017 (D6-)	HLT_xe110_PUFT_L1XE50	HLT_xe110_PUFT_L1XE50	HLT_E26_LHTIGHT_NOD0_IVARLOOSE OR HLT_E60_LHMEDIUM_NOD0 OR HLT_E140_LHLOOSE_NOD0 OR HLT_E300_ETCUT OR HLT_MU26_IVARMEDIUM OR HLT_MU50
2018 (B-C5)	HLT_xe110_PUFT_70_L1XE55	HLT_xe110_PUFT_70_L1XE55	HLT_E26_LHTIGHT_NOD0_IVARLOOSE OR HLT_E60_LHMEDIUM_NOD0 OR HLT_E140_LHLOOSE_NOD0 OR HLT_E300_ETCUT OR HLT_MU26_IVARMEDIUM OR HLT_MU50
2018 (C5-)	HLT_xe110_PUFT_65_L1XE55	HLT_xe110_PUFT_65_L1XE55	HLT_E26_LHTIGHT_NOD0_IVARLOOSE OR HLT_E60_LHMEDIUM_NOD0 OR HLT_E140_LHLOOSE_NOD0 OR HLT_E300_ETCUT OR HLT_MU26_IVARMEDIUM OR HLT_MU50

The MC production request is then placed covering the parameter space up to the edge of the expected sensitivity of the search.

A.3.1. The $E_T^{\text{miss}} + H(b\bar{b})$ RECAST Example

To make an example, the procedure applied to the $E_T^{\text{miss}} + H(b\bar{b})$'s reinterpretation in the Dark Higgs scenario [4] is shown. The GL are taken from the $E_T^{\text{miss}} + H(b\bar{b})$ analysis [97].

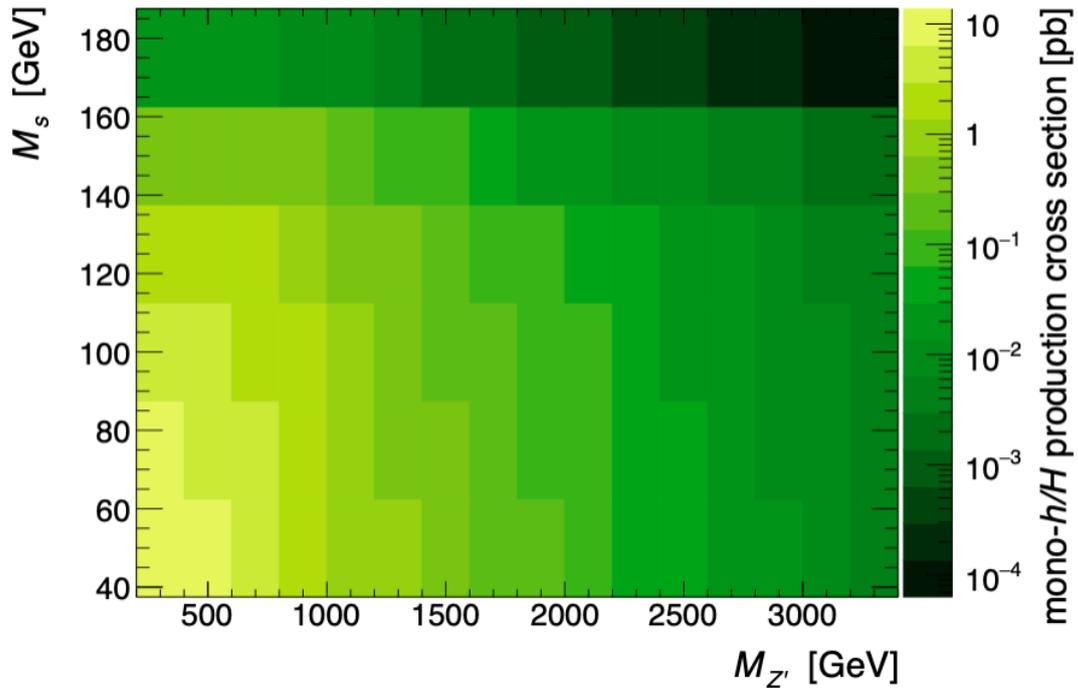


Figure A.2.: Madgraph cross-section computed for the signal model in the parameter space.

The procedure is illustrated from Figure A.2 to Figure A.5. First, Figure A.2, the signal cross-section is evaluate, then Figure A.3, binned in E_T^{miss} . This is multiplied by $A * \epsilon / \sigma_{GL}$ to evaluate the binned reconstruction-level sensitivity, Figure A.4, and finally, Figure A.5, all the E_T^{miss} bins are summed up together.

The expected and observed exclusion contour from RECAST reinterpretation of the Dark Higgs signal was shown in the body of the thesis, Figure 3.8; the GL estimated limits are in fact quantitatively consistent (cf. Figure A.5).

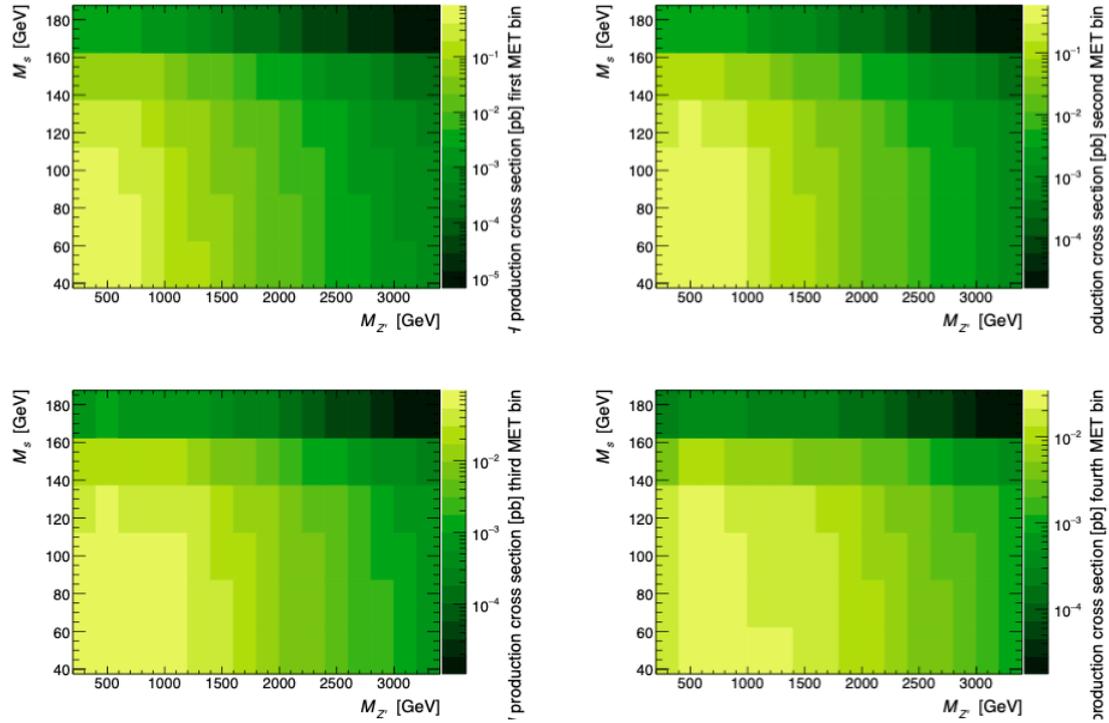


Figure A.3.: The Madgraph cross-section is then binned in E_T^{miss} .

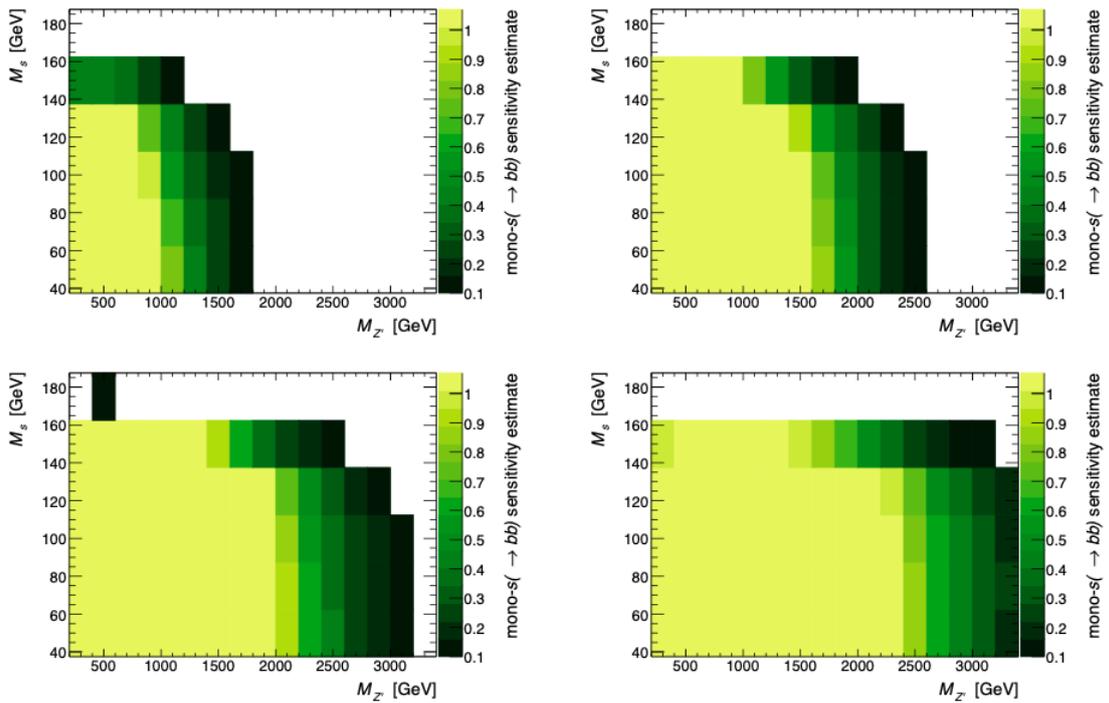


Figure A.4.: Bin-wise sensitivity multiplied by the $A * \epsilon / \sigma_{GL}$.

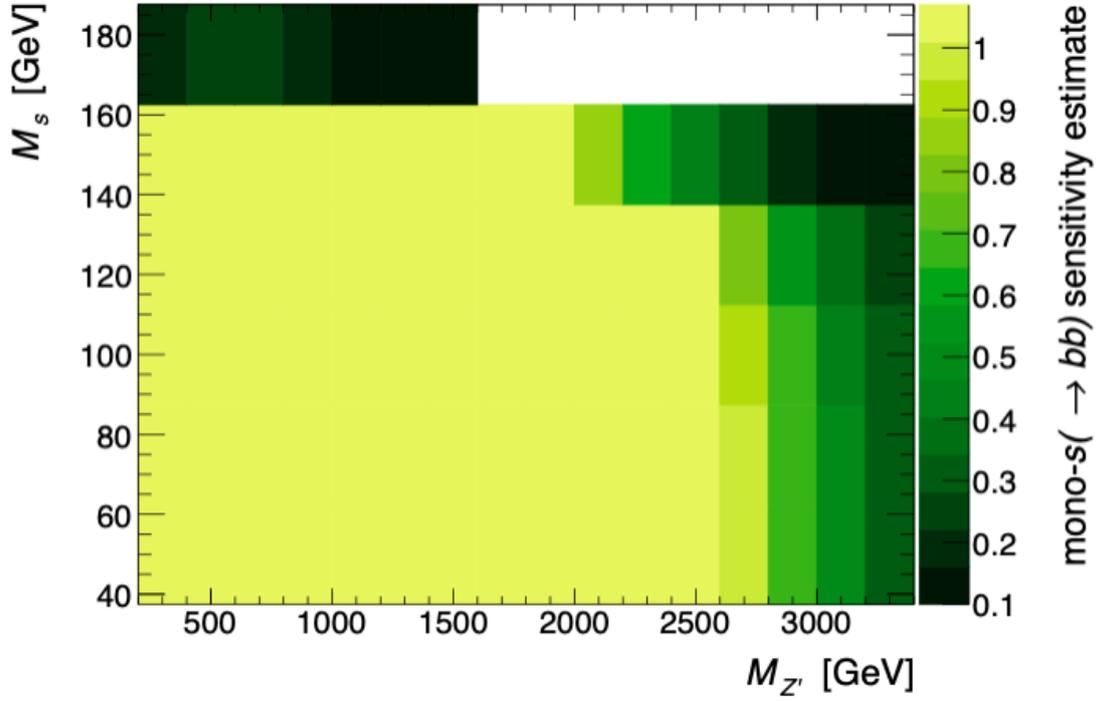


Figure A.5.: Final reco-level sensitivity.

A.3.2. The $E_T^{\text{miss}} + s \rightarrow W^+ W^- (\text{had})$ Sensitivity Estimate

The example above showed the application of the sensitivity estimate using GL for the $E_T^{\text{miss}} + H(b\bar{b})$'s RECAST to the signal model. The same cannot be done directly in $E_T^{\text{miss}} + s \rightarrow W^+ W^- (\text{had})$, since there are no available GL limits for the resonant $W^+ W^-$ final state. It can be, however, estimated with care from the $E_T^{\text{miss}} + V(\text{had})$ analysis, targeting hadronically decaying W and Z boson and E_T^{miss} [98], which do provide GL. These two analysis can be compared for a number of reasons:

- Same main background (Z+jets, W+jets, $t\bar{t}$),
- Similar fit setup with 2 lepton Control Region to constrain Z+jets, 1 lepton Control Region to constrain W+jets,
- Similar 2 prong decay of W/Z/H,
- Similar strategy for $E_T^{\text{miss}} + s \rightarrow W^+ W^- (\text{had})$.

In these setting, the GL are mostly a function of the background level in the Signal Region. This can be seen in Table A.3.2.

E_T^{miss} bin	GL $E_T^{\text{miss}} + H(b\bar{b})$, fb^{-1}	Bkg events in SR $E_T^{\text{miss}} + H(b\bar{b})$	GL $E_T^{\text{miss}} + V(\text{had})$ fb^{-1}	Bkg events in SR $E_T^{\text{miss}} + V(\text{had})$
150 - 200 GeV	19.1	2835	750	198500
200 - 350 GeV	13.1	1390	248	133350
350 - 500 GeV	2.4	98	26	11004
500 - GeV	1.7	11	9.8	1878

The table of the E_T^{miss} binned estimated background (from the post-fit) in the SR and the GL.

Figure A.6 moreover shows the linear dependency of the GL on the background.

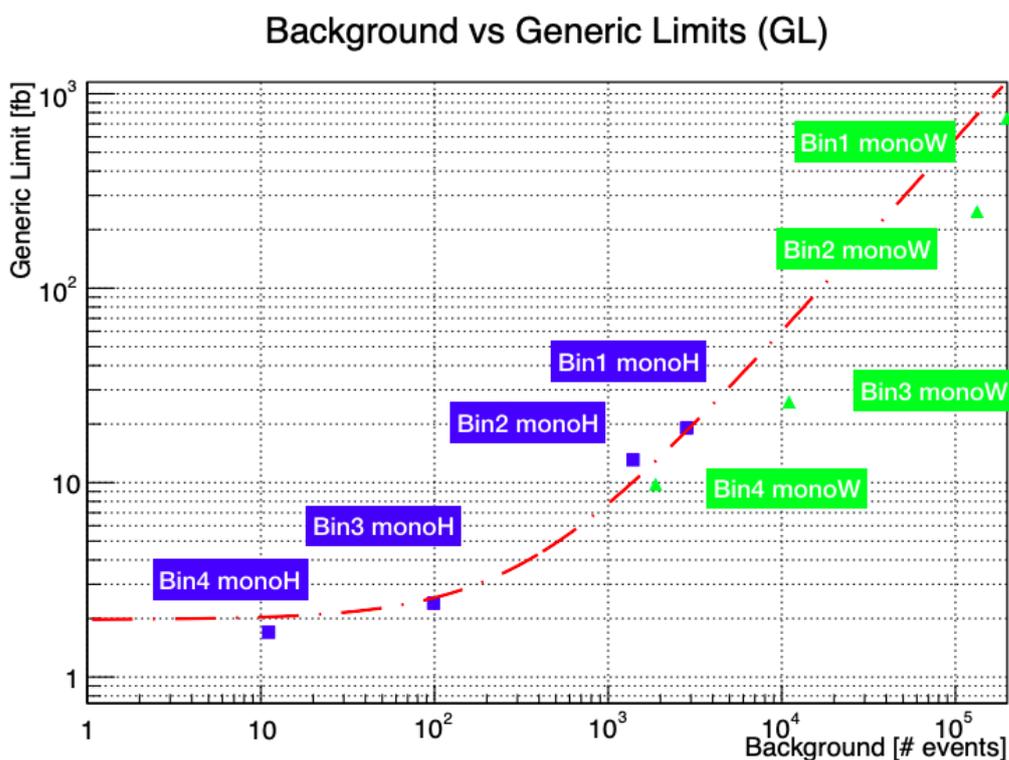


Figure A.6.: Dependency of the GL on the background in the Signal Region. The E_T^{miss} bins are highlighted in the plot and the a linear fit is shown in red line.

Assuming a similar behaviour of the GL as the one seen in these analyses and estimating the background level in the $E_T^{\text{miss}} + s \rightarrow W^+ W^- (\text{had})$ SR, the linear relation found to get the expected GL can be used.

This is done in two ways, an optimistic and a pessimistic one. This assures that the assumptions are not dominating the sensitivity evaluation.

In the pessimistic one, the background efficiency for $E_T^{\text{miss}} + s \rightarrow W^+W^- (\text{had})$, ϵ_{WW} are assumed to be ϵ_W^2 , where ϵ_W is taken from the $E_T^{\text{miss}} + V(\text{had})$ analysis.

In the optimistic case, some more background rejection criteria based on the likely event strategy of the analysis is hypothesized, such as:

- For $E_T^{\text{miss}} < 500$ GeV at least 4 jets are required. Additional $\epsilon=0.35$.
- The mass window of the Large-R jets as in $E_T^{\text{miss}} + V(\text{had})$ is moved to values at least bigger than 160 GeV (thus benefiting from the more than exponential decay of the background with the mass of the Large-R jet). Additional $\epsilon=0.05$.
- Enforce 2 Large-R jets for the intermediate case. This assumption was not really followed by the analysis, where the second Large-R jet is usually not reclustered. However, the dedicated algorithm develop could cure much of the mis-evaluation. Additional $\epsilon=0.05$.

The final estimates are shown in Figure A.7. This can be compared with the choice of the signal grid, Figure 7.1.

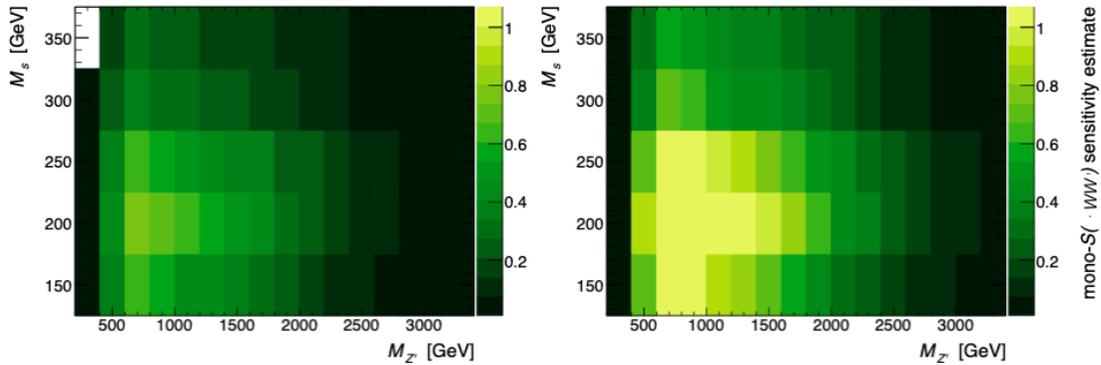


Figure A.7.: Final sensitivity for the two scenarios hypothesized for the background level in the Signal Region, the worst and the best case, left and right plot, respectively.

A.4. Overlap Removal

This section summarizes the working points used in the object definition for the removal of H_T overlap: this is a very important check to make sure that there is no double counting in the calorimeter deposits and tracks. The working points are:

Table A.2.: Overlap removal

Reject	Against	Criteria
Electron	Electron	shared track, $p_{T,1} < p_{T,2}$
Tau	Electron	$\Delta R < 0.2$
Tau	Muon	$\Delta R < 0.2$
Muon	Electron	is Calo-Muon and shared ID track
Electron	Muon	shared ID track
Photon	Electron	$\Delta R < 0.4$
Photon	Muon	$\Delta R < 0.4$
Jet	Electron	$\Delta R < 0.2$
Electron	Jet	$\Delta R < \min(0.4, 0.04 + 10 \text{ GeV}/p_T(e))$
Jet	Muon	NumTrack < 3 and (ghost-associated or $\Delta R < 0.2$)
Muon	Jet	$\Delta R < \min(0.4, 0.04 + 10 \text{ GeV}/p_T(\mu))$
Jet	Tau	$\Delta R < 0.2$
Photon	Jet	$\Delta R < 0.4$
Fat-jet	Electron	$\Delta R < 1.0$

(ΔR) is calculated using rapidity by default.

In addition to the procedure listed in Table A.2, after performing the steps in order as listed, an overlap removal is performed with the $R = 0.2$ jets vs. electrons, muons and taus, always in favour of the leptons, meaning a jet is removed if it overlaps within 0.2 with a lepton.

A.5. TAR+Comb

TAR+Comb algorithm is presented here in a pictorial way, Figure A.8. The algorithm can be constructed both upon R -scan or Small- R jets. It is used in the intermediate category.

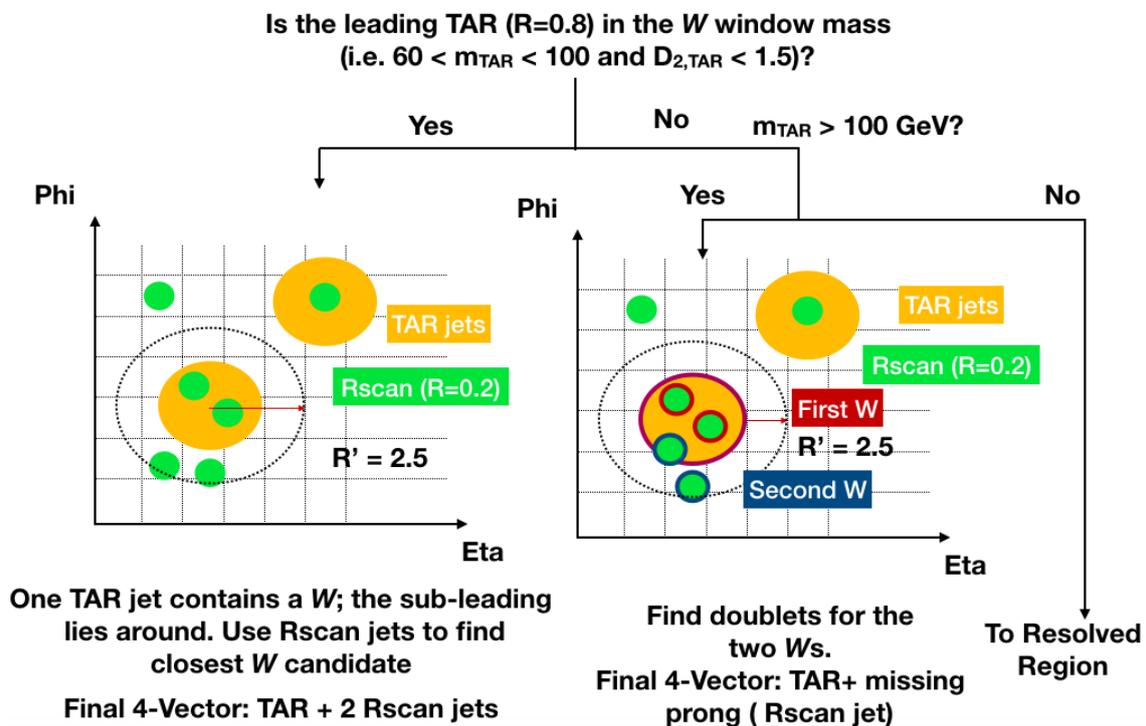


Figure A.8.: The representation of the flow of the TAR+Comb algorithm. First a TAR jet consistent with a V -jet is looked for; if one is found (left), the two prong from the V jets not contained in the TAR jet are added. If the TAR jet is above the V mass window (right), only one V prong is assumed to be missing; this is selected and added. If the TAR jet has a mass below the V window, the event is passed to the resolved category according to the priority based selection. In the η - ϕ maps, the yellow circle is the TAR jet, the green ones the R-scan jets. R' is radius in which the algorithm looks for the missing prong(s). This example is shown for the optima case of R-scan jets, while in the analysis Small- R jets are used instead.

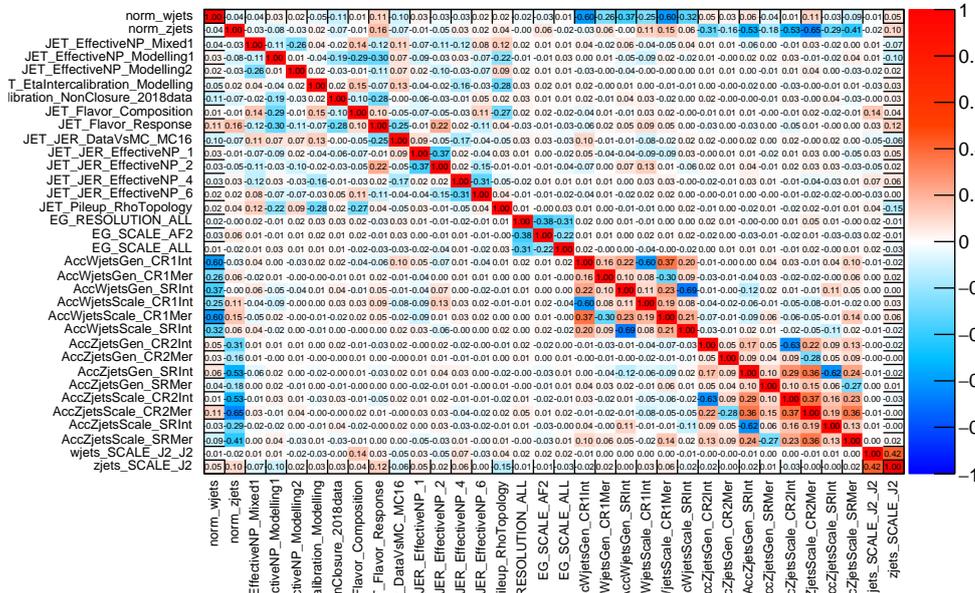


Figure A.10.: Correlation matrix showing the largest correlations of nuisance parameters (excluding nuisance parameters associated to Poisson terms in the likelihood function) for conditional $\mu = 0$ fit to a Asimov data set generated with $\mu = 1$.

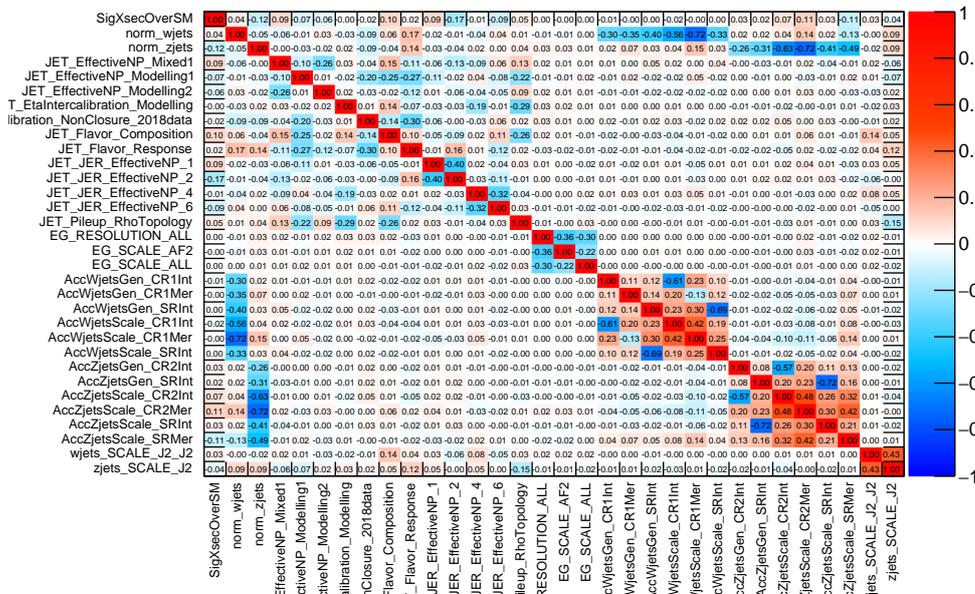


Figure A.11.: Correlation matrix showing the largest correlations of nuisance parameters (excluding nuisance parameters associated to Poisson terms in the likelihood function) for unconditional fit to a Asimov data set generated with $\mu = 0$.

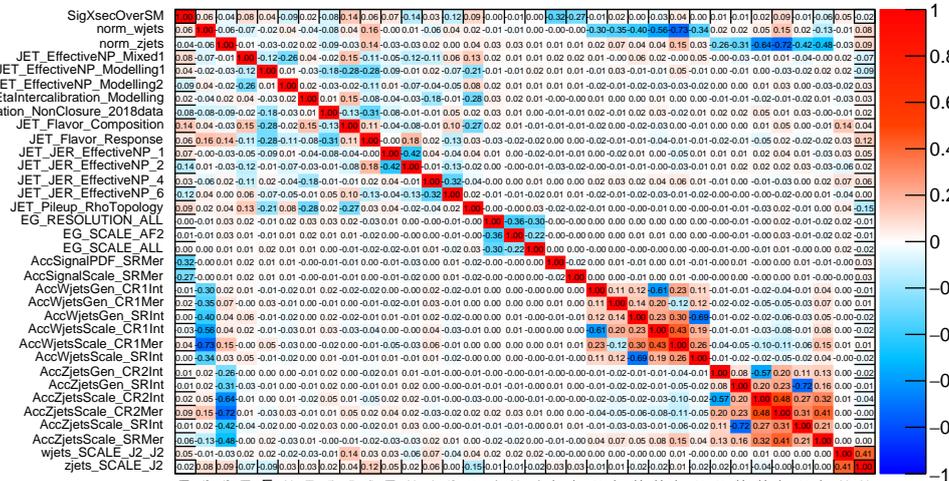


Figure A.12.: Correlation matrix showing the largest correlations of nuisance parameters (excluding nuisance parameters associated to Poisson terms in the likelihood function) for unconditional fit to a Asimov data set generated with $\mu = 1$.

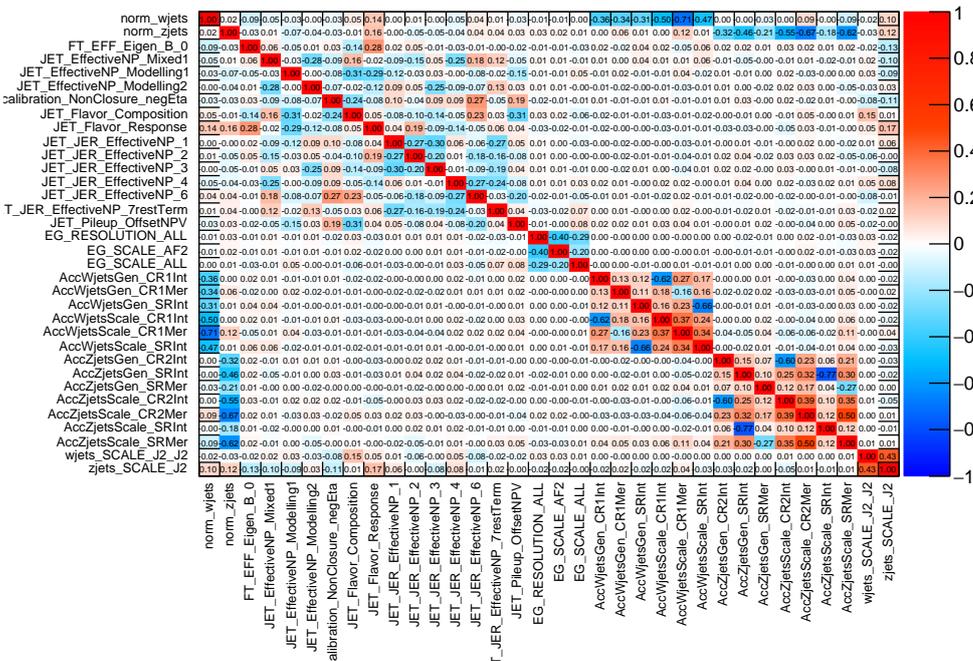


Figure A.13.: Correlation matrix showing the largest correlations of nuisance parameters (excluding nuisance parameters associated to Poisson terms in the likelihood function) for conditional $\mu = 0$ fit to the observed data.

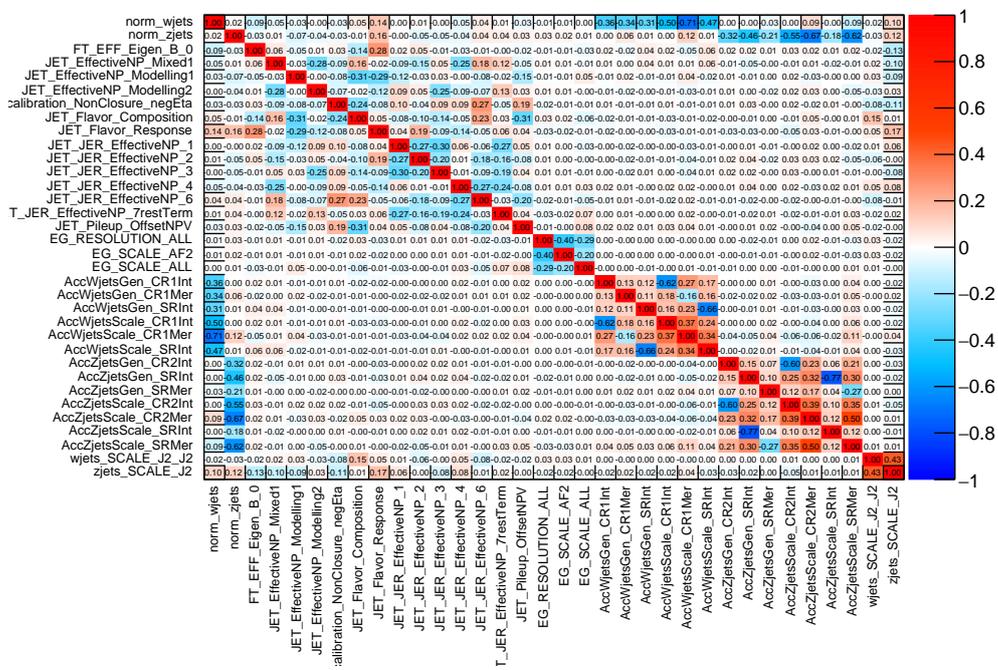


Figure A.14.: Correlation matrix showing the largest correlations of nuisance parameters (excluding nuisance parameters associated to Poisson terms in the likelihood function) to the observed data.

A.7. Resolved Category

A *resolved* category was studied but not included in the analysis. It targets the case when the Lorentz boost of the VV system is so low that none of them can be reconstructed with a single Large- R jet. In that case, Small- R jets are individually used to reconstruct the two V bosons. An ad-hoc algorithm was investigated for this purpose, selecting a set of four jets that are consistent with the masses of a pair of W bosons. The final state $s \rightarrow WW$ has in fact an higher cross-section than $s \rightarrow ZZ$, and the two boson are not distinguishable in the fully hadronic decay. However, since the resolved category was shown to have a contribution below 5% of the total significance of the search at the cost of substantial additional complexity, it was not included in the analysis. It is defined in this Appendix for completeness and since its use is considered in future analyses.

The resolved category addresses the kinematic region of the analysis where the Lorentz boost of the $s \rightarrow VV$ system is not sufficient to collimate the decay products into a Large- R jet. The four quarks in the final state are therefore reconstructed with four different Small- R jets. The resolved category helps gaining sensitivity in the region of the parameters space, such as in the case of high m_s , where the decay products are typically further apart with respect to lower m_s hypothesis. The most naive way to reconstruct the mass of the scalar is just to consider the invariant mass of the four leading p_T jets, m_{jjjj} . This choice, however, still does not take advantage of the fact that pairs of jets from the W/Z decay can be discriminated against backgrounds by means of their compatibility with the m_W . To enhance the sensitivity, a simple algorithm has been developed and optimized, the *2W algorithm*, which selects the four jets with the best combination of pairwise invariant masses. The algorithm is explained below:

- A pool is formed with all the pairwise combinations of two Small- R jets in the event.
- Two pairs, 1 and 2, which minimize $(m_1 - m_W)^2 + (m_2 - m_W)^2$, are picked from the pool.
- The pairs are checked for no overlapping constituents, e.g. the members of one pair cannot be member of other pairs.

While this algorithm has the potential to increase the significance, both signal and backgrounds are sculpted.

Table A.3.: A summary of the three different reconstruction algorithms used in each of the three prioritized signal categories, with the resolved category which is detailed in this Appendix. The final kinematic selection requirements of the reconstructed Dark Higgs candidate are also listed.

Merged	Intermediate	Resolved
$\geq 1 R = 0.8$ TAR jet	= 1 Dark Higgs candidate	≥ 4 small-R jets
$p_{\text{TAR}} > 300$ GeV	reconstructed with TAR+Comb	= 1 Dark Higgs candidate
$0. < \tau_{42} < 0.3$	$100 \text{ GeV} < m_{\text{TAR+Comb}} < 400 \text{ GeV}$	reconstructed with 2W algorithm
$0. < \tau_{43} < 0.6$		$E_{\text{T}}^{\text{miss}} * \exp(\Delta R_{1,2}) > 40$
$100 \text{ GeV} < m_{\text{TAR}} < 400 \text{ GeV}$		$100 \text{ GeV} < m_{2W} < 400 \text{ GeV}$

In order to take advantage of all possible kinematic correlations present in the signal and not in the background, the $E_{\text{T}}^{\text{miss}}$ and the opening angle between the leading and sub-leading p_{T} jets are combined into a single variable, $\text{METEXPDR12} = E_{\text{T}}^{\text{miss}} * \exp(-\Delta R_{1,2})$. As the background peaks at lower values (driven by higher $\Delta R_{1,2}$) and the signal peaks at higher values (driven by lower $\Delta R_{1,2}$ and higher $E_{\text{T}}^{\text{miss}}$). A cut for increased significance can be defined $\text{METEXPDR12} > 40$ GeV. The optimization was cross-checked via a BDT analysis in order to obtain the relative variable importance.

The requirements made on all the analysis categories are summarized together in the Table A.3.

The $E_{\text{T}}^{\text{miss}}$ binning is indicated on Table A.4 for all categories.

Table A.4.: The binning for the different regions 0LSR / 1LCR / 2LCR as a function of their $E_{\text{T}}^{\text{miss}} / E_{\text{T}}^{\text{miss,no}\mu} / p_{\text{T}}^{\ell\ell}$ bins for the merged, intermediate, and resolved categories. The same ranking is merged (M) \rightarrow intermediate (I) \rightarrow resolved (R).

	200-300 GeV	300-500 GeV	more than 500 GeV
0LSR ($E_{\text{T}}^{\text{miss}}$)	I,R	M,I,R	M,I,R
1LCR ($E_{\text{T}}^{\text{miss,no}\mu}$)	I,R	M,I,R	M,I,R
2LCR ($p_{\text{T}}^{\ell\ell}$)	I,R	M,I,R	M,I,R

A.8. Event Yield

A.8.1. Pre-fit Event Yield Tables

Pre-fit event yield tables are shown here for the 0LSR (Table A.5), 1LCR (Table A.6) and 2LCR (Table A.7).

Table A.5.: Expected and observed number of events in the 0 lepton signal region with 139 fb^{-1} data, shown separately in each topology category. The background yields and uncertainties are shown before the profile-likelihood fit to the data. The quoted background uncertainties include both the statistical and systematic contributions. The total uncertainties can be smaller than for individual components due to anti-correlations.

Process	merged	merged	intermediate	intermediate	intermediate
	300 - 500 GeV	> 500 GeV	200 - 300 GeV	300 - 500 GeV	> 500 GeV
Signal	65 ± 35	44 ± 40	170 ± 105	589 ± 250	211 ± 126
Z+jets	363 ± 158	108 ± 59	14881 ± 7511	25210 ± 11923	6167 ± 2890
W+jets	219 ± 139	34 ± 53	10935 ± 6626	13572 ± 7438	2275 ± 1296
$t\bar{t}$	47 ± 30	8 ± 14	1290 ± 427	1269 ± 390	87 ± 60
Single top-quark	9 ± 15	3 ± 6	312 ± 224	361 ± 222	47 ± 72
Diboson	47 ± 39	16 ± 27	1028 ± 442	1554 ± 529	367 ± 204
Bkg	685 ± 342	169 ± 147	28448 ± 11578	41971 ± 14837	8944 ± 3364
Data	647	178	29143	42653	8654

Table A.6.: Expected and observed number of events in the 1 lepton control region with 139 fb^{-1} data, shown separately in each topology category. The background yields and uncertainties are shown before the profile-likelihood fit to the data. The quoted background uncertainties include both the statistical and systematic contributions. The total uncertainties can be smaller than for individual components due to anti-correlations.

Process	merged	merged	intermediate	intermediate	intermediate
	300 - 500 GeV	> 500 GeV	200 - 300 GeV	300 - 500 GeV	> 500 GeV
Signal	2 ± 4	0 ± 0	6 ± 10	11 ± 21	3 ± 6
Z+jets	12 ± 15	2 ± 4	507 ± 365	668 ± 382	127 ± 87
W+jets	428 ± 175	135 ± 85	16439 ± 8005	28920 ± 13142	7599 ± 3502
$t\bar{t}$	127 ± 55	25 ± 32	2866 ± 781	3655 ± 842	381 ± 174
Single top-quark	44 ± 35	19 ± 39	861 ± 386	1389 ± 490	262 ± 208
Diboson	36 ± 36	15 ± 30	700 ± 352	1220 ± 452	341 ± 187
Bkg	647 ± 286	197 ± 182	21375 ± 8772	35854 ± 13608	8710 ± 3714
Data	595	143	20426	34579	8086

Table A.7.: Expected and observed number of events in the 2 lepton control region with 139 fb^{-1} data, shown separately in each topology category. The background yields and uncertainties are shown before the profile-likelihood fit to the data. The quoted background uncertainties include both the statistical and systematic contributions. The total uncertainties can be smaller than for individual components due to anti-correlations.

Process	merged	merged	intermediate	intermediate	intermediate
	300 - 500 GeV	> 500 GeV	200 - 300 GeV	300 - 500 GeV	> 500 GeV
Signal	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Z+jets	98 ± 55	24 ± 18	5243 ± 2748	5171 ± 2171	1236 ± 525
W+jets	0 ± 0	0 ± 0	4 ± 8	2 ± 2	0 ± 0
$t\bar{t}$	0 ± 0	0 ± 0	10 ± 18	2 ± 4	0 ± 0
Single top-quark	0 ± 0	0 ± 0	2 ± 6	0 ± 0	0 ± 0
Diboson	8 ± 8	3 ± 5	236 ± 115	213 ± 80	52 ± 33
Bkg	107 ± 61	27 ± 23	5496 ± 2825	5389 ± 2193	1288 ± 535
Data	96	21	5143	5230	1286

A.8.2. Post-fit Event Yield Tables

Post-fit event yield tables are shown here for the 0LSR (Table A.8), 1LCR (Table A.9) and 2LCR (Table A.10)

Table A.8.: Expected and observed number of events in the 0 lepton signal region with 139 fb^{-1} data, shown separately in each topology category. The background yields and uncertainties are shown after the profile-likelihood fit to the data. The quoted background uncertainties include both the statistical and systematic contributions. The total uncertainties can be smaller than for individual components due to anti-correlations.

Process	merged	merged	intermediate	intermediate	intermediate
	300 - 500 GeV	> 500 GeV	200 - 300 GeV	300 - 500 GeV	> 500 GeV
Signal	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Z+jets	414 ± 50	129 ± 15	13246 ± 943	22860 ± 1512	5562 ± 357
W+jets	161 ± 50	22 ± 9	13346 ± 1045	16621 ± 1404	2705 ± 252
$t\bar{t}$	44 ± 9	5 ± 3	1362 ± 169	1337 ± 177	79 ± 15
Single top-quark	6 ± 3	1 ± 1	279 ± 31	334 ± 41	30 ± 8
Diboson	36 ± 7	8 ± 2	926 ± 144	1444 ± 230	310 ± 53
Bkg	661 ± 24	165 ± 11	29160 ± 888	42601 ± 1308	8687 ± 291
Data	647	178	29143	42653	8654

Table A.9.: Expected and observed number of events in the 1 lepton control region with 139 fb^{-1} data, shown separately in each topology category. The background yields and uncertainties are shown after the profile-likelihood fit to the data. The quoted background uncertainties include both the statistical and systematic contributions. The total uncertainties can be smaller than for individual components due to anti-correlations.

Process	merged	merged	intermediate	intermediate	intermediate
	300 - 500 GeV	> 500 GeV	200 - 300 GeV	300 - 500 GeV	> 500 GeV
Signal	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Z+jets	9 ± 4	0 ± 0	539 ± 285	722 ± 384	132 ± 70
W+jets	391 ± 37	112 ± 10	15424 ± 513	27440 ± 802	7100 ± 183
$t\bar{t}$	129 ± 21	20 ± 6	3007 ± 368	3904 ± 509	391 ± 61
Single top-quark	39 ± 7	6 ± 3	852 ± 86	1400 ± 158	226 ± 43
Diboson	26 ± 7	5 ± 2	608 ± 120	1056 ± 212	281 ± 59
Bkg	594 ± 23	144 ± 9	20430 ± 221	34526 ± 353	8131 ± 113
Data	595	143	20426	34579	8086

Table A.10.: Expected and observed number of events in the 2 lepton control region with 139 fb^{-1} data, shown separately in each topology category. The background yields and uncertainties are shown after the profile-likelihood fit to the data. The quoted background uncertainties include both the statistical and systematic contributions. The total uncertainties can be smaller than for individual components due to anti-correlations.

Process	merged	merged	intermediate	intermediate	intermediate
	300 - 500 GeV	> 500 GeV	200 - 300 GeV	300 - 500 GeV	> 500 GeV
Signal	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
Z+jets	90 ± 9	21 ± 3	4920 ± 120	5040 ± 121	1224 ± 33
W+jets	0 ± 0	0 ± 0	2 ± 1	2 ± 1	0 ± 0
$t\bar{t}$	0 ± 0	0 ± 0	6 ± 2	1 ± 1	0 ± 0
Single top-quark	0 ± 0	0 ± 0	1 ± 1	1 ± 0	0 ± 0
Diboson	6 ± 2	2 ± 1	215 ± 40	201 ± 38	46 ± 10
Bkg	96 ± 9	22 ± 3	5144 ± 115	5245 ± 114	1270 ± 33
Data	96	21	5143	5230	1286

A.9. Expected Significance Breakdown

In this appendix, the breakdown of the expected significance for the search is presented. It is shown for different decay modes of the Dark Higgs in the merged and intermediate categories (which the fit setup presented in the main body of the thesis) on Figure A.15 for the W^+W^- (*had*) final state and on Figure A.16 for the ZZ (*had*) final state. The per-categories breakdown is shown on Figure A.17 for the merged category in the VV (*had*) final state. On Figure A.18 the intermediate category in the VV (*had*) final state. Finally, on Figure A.19 the resolved category in the VV (*had*). As it can be appreciated here, the resolved category contributes in a negligible way to the total significance, and it was subsequently dropped from the fit analysis setup. This had the advantage to make the fit more stable and sound. The expected limits were calculated with Asimov data only.

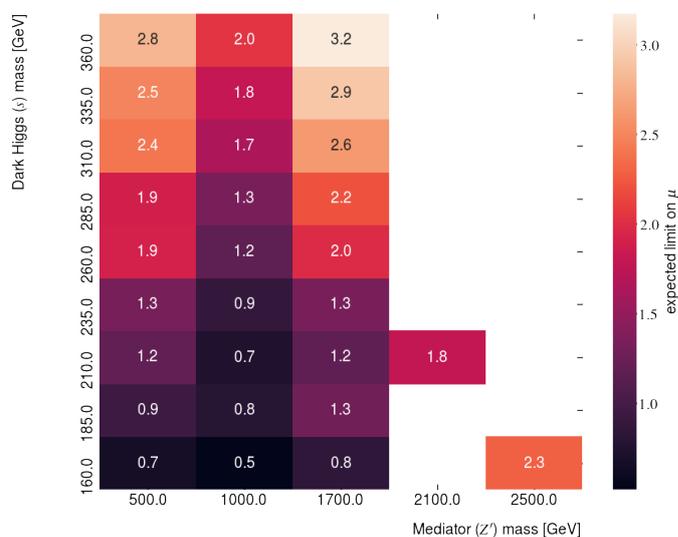


Figure A.15.: Expected limit on the signal strength μ for Dark Higgs signals in the WW decay mode with fixed $m_\chi = 200$ GeV, scanned in $m_{Z'}$ and m_s . The merged and intermediate categories are considered.

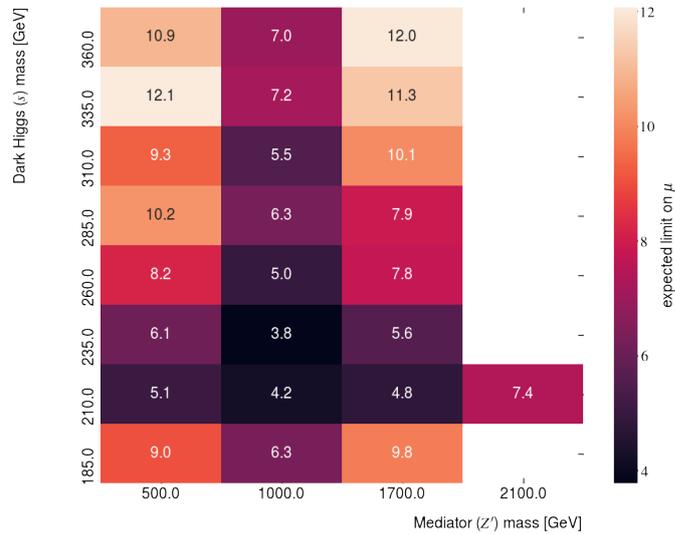


Figure A.16.: Expected limit on the signal strength μ for Dark Higgs signals in the ZZ decay mode with fixed $m_\chi = 200$ GeV, scanned in $m_{Z'}$ and m_s . The merged and intermediate categories are considered.

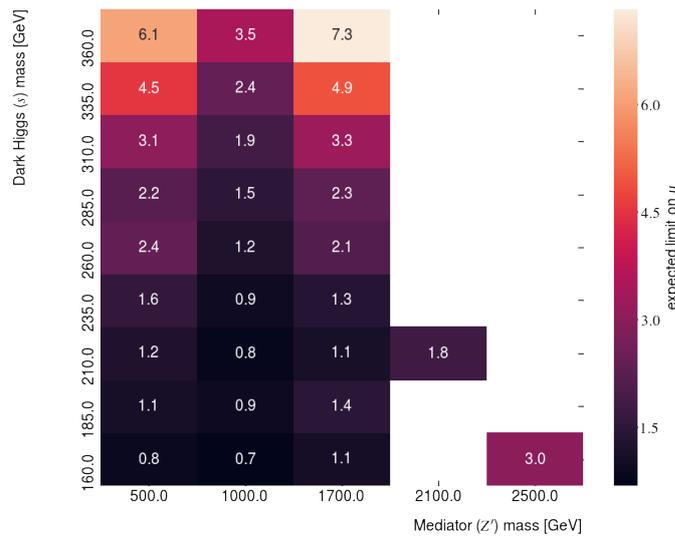


Figure A.17.: Expected limit on the signal strength μ for Dark Higgs signals in the combined WW and ZZ decay modes with fixed $m_\chi = 200$ GeV, scanned in $m_{Z'}$ and m_s . Only the merged category is considered.

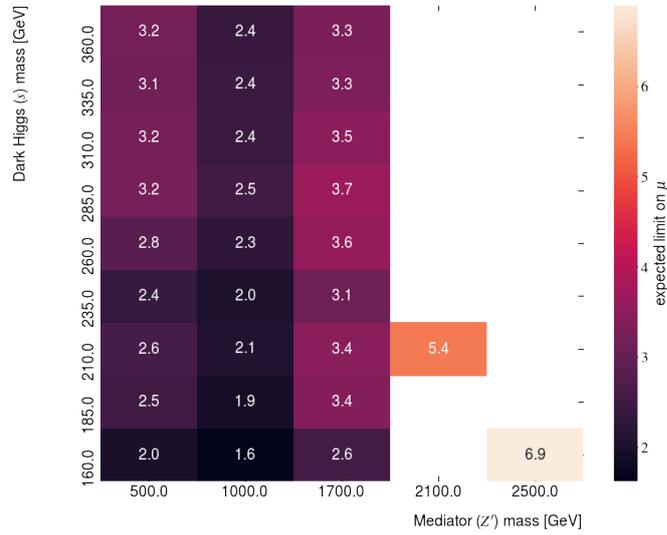


Figure A.18.: Expected limit on the signal strength μ for Dark Higgs signals in the combined WW and ZZ decay modes with fixed $m_\chi = 200$ GeV, scanned in $m_{Z'}$ and m_s . Only the intermediate category is considered.

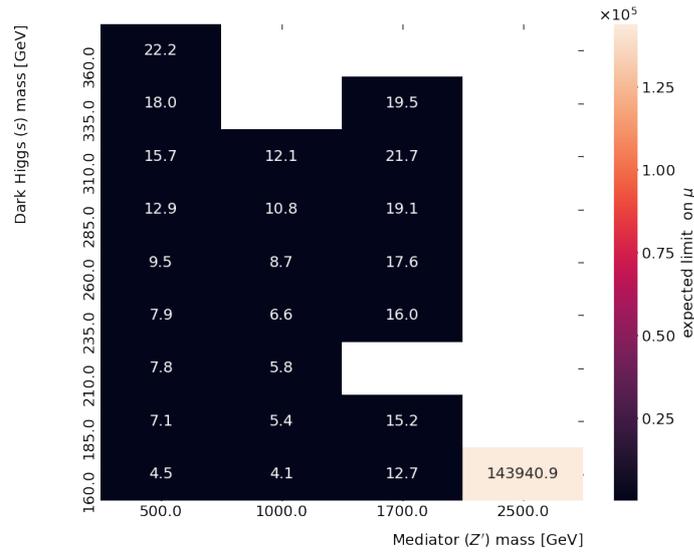


Figure A.19.: Expected limit on the signal strength μ for Dark Higgs signals in the combined WW and ZZ decay modes with fixed $m_\chi = 200$ GeV, scanned in $m_{Z'}$ and m_s . Only the resolved category is considered.

A.10. Exclusion Limits for Different Z' masses

In this Section, the exclusion limits of the $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ analysis are presented for different values of the $m_{Z'}$ than $m_{Z'} = 1$ TeV, which is shown in the body of the thesis 11.3. In Figure A.20 the 95% CL_s exclusion limits on the signal strength are shown for $m_{Z'} = 1.7$ TeV. In Figure A.21 they are shown instead for the mass hypothesis $m_{Z'} = 0.5$ TeV. Expected and observed limits are worse with respect to the ones on the Dark Higgs hypothesis with $m_{Z'} = 1$ TeV. This is because of a delicate balance of effects. For low masses of the Z' boson, the Lorentz boost generated on the Dark Higgs is consequently lower. This results in a migration of events from the most sensitive category of the search, the merged category, to the less sensitive intermediate category. For heavier masses of the Z' boson, instead the decrease of the production cross-section takes its toll on the signal yield, degrading the overall sensitivity. Even if the Dark Higgs signal is highly energetic, its yield is much more limited.

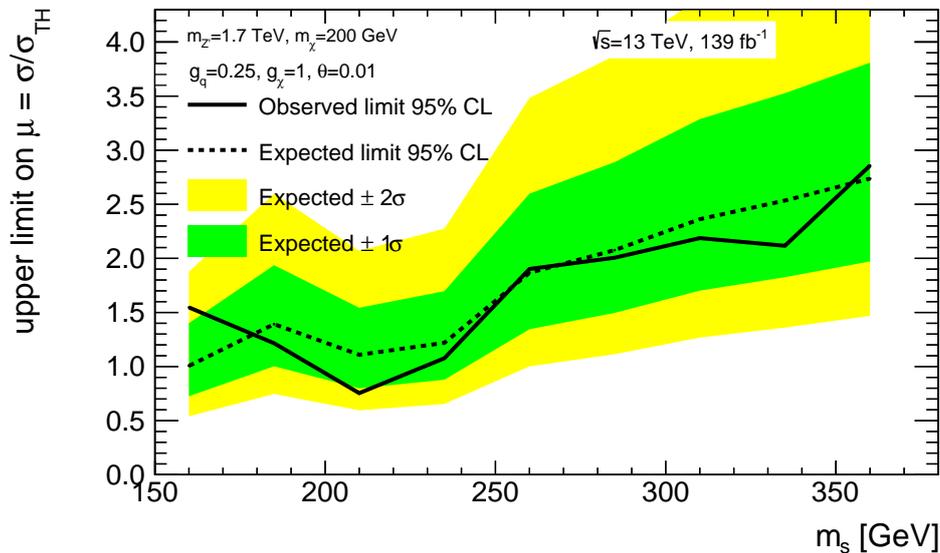


Figure A.20.: 95% CL_s expected and observed limits on the signal strength μ as dashed and solid line, respectively, for the $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ analysis in the signal hypothesis with $m_{Z'} = 1.7$ TeV. The band around the expected limits shows the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties.

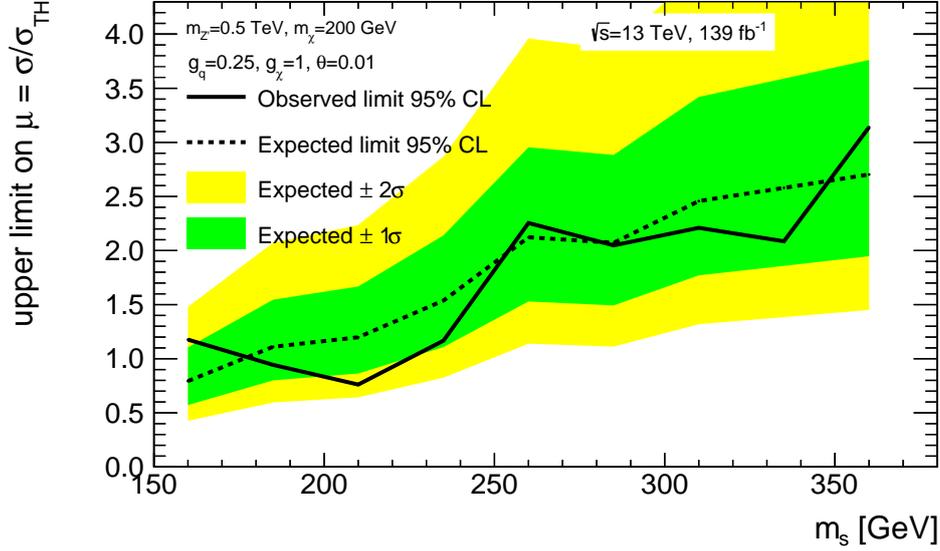


Figure A.21.: 95% CL_s expected and observed limits on the signal strength μ as dashed and solid line, respectively, for the $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ analysis in the signal hypothesis with $m_{Z'} = 0.5$ TeV. The band around the expected limits shows the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties.

A.10.1. Exclusion Limits Tables

Detailed tables of the observed and expected limits with their ± 1 and $\pm 2\sigma$ for different mass hypothesis of the Dark Higgs and Z' massive gauge boson. The Dark Matter mass is fixed at 200 GeV. Table A.11 shows the expected and observed limits for $m_{Z'} = 0.5$ TeV, Table A.12 shows the expected and observed limits for $m_{Z'} = 1$ TeV and Table A.13 shows the expected and observed limits for $m_{Z'} = 1.7$ TeV.

Table A.11.: Observed and expected limits on the Dark Higgs scenario with $m_{Z'} = 0.5$ TeV

m_s [GeV]	observed limit on μ	expected limit on μ	-2σ	-1σ	$+1\sigma$	$+2\sigma$
160	1.17	0.79	0.42	0.57	1.10	1.48
185	0.94	1.11	0.59	0.80	1.54	2.07
210	0.76	1.19	0.64	0.86	1.66	2.23
235	1.16	1.53	0.82	1.10	2.13	2.86
260	2.25	2.12	1.13	1.52	2.95	3.96
285	2.04	2.07	1.11	1.49	2.88	3.86
310	2.20	2.45	1.31	1.77	3.42	4.58
335	2.08	2.57	1.38	1.85	3.58	4.81
360	3.14	2.70	1.45	1.94	3.76	5.04

Table A.12.: Observed and expected limits on the Dark Higgs scenario with $m_{Z'} = 1$ TeV

m_s [GeV]	observed limit on μ	expected limit on μ	-2σ	-1σ	$+1\sigma$	$+2\sigma$
160	1.12	0.66	0.35	0.47	0.91	1.23
185	0.75	0.94	0.50	0.68	1.31	1.76
210	0.54	0.81	0.43	0.58	1.12	1.51
235	0.82	0.91	0.48	0.65	1.26	1.69
260	1.23	1.20	0.64	0.86	1.67	2.24
285	1.42	1.51	0.81	1.09	2.10	2.82
310	1.56	1.71	0.91	1.23	2.38	3.19
335	1.60	1.72	0.92	1.24	2.40	3.22
360	2.09	1.92	1.03	1.38	2.67	3.58

Table A.13.: Observed and expected limits on the Dark Higgs scenario with $m_{Z'} = 1.7$ TeV

m_s [GeV]	observed limit on μ	expected limit on μ	-2σ	-1σ	$+1\sigma$	$+2\sigma$
160	1.54	1.00	0.54	0.72	1.40	1.87
185	1.21	1.39	0.74	1.00	1.93	2.59
210	0.75	1.10	0.59	0.79	1.54	2.06
235	1.07	1.21	0.65	0.87	1.69	2.27
260	1.90	1.86	1.00	1.34	2.59	3.48
285	2.00	2.07	1.11	1.49	2.88	3.87
310	2.18	2.36	1.26	1.70	3.28	4.40
335	2.11	2.53	1.36	1.82	3.52	4.72
360	2.85	2.73	1.46	1.97	3.80	5.10

B. Full Data/MC Comparison

In this Appendix, the data/MC distributions are shown, with the full breakdown per each E_T^{miss} bin. In the body of the thesis, Section 8.3, the distributions integrated in E_T^{miss} are displayed.

B.0.1. Signal Region

Merged Category

The MC distributions for the 0 lepton signal region in the merged category are shown separately for $E_T^{\text{miss}} > 500\text{GeV}$ in Figure B.1 and for $300\text{GeV} < E_T^{\text{miss}} < 500\text{GeV}$ in Figure B.2 respectively. All requirements of the event selection are applied.

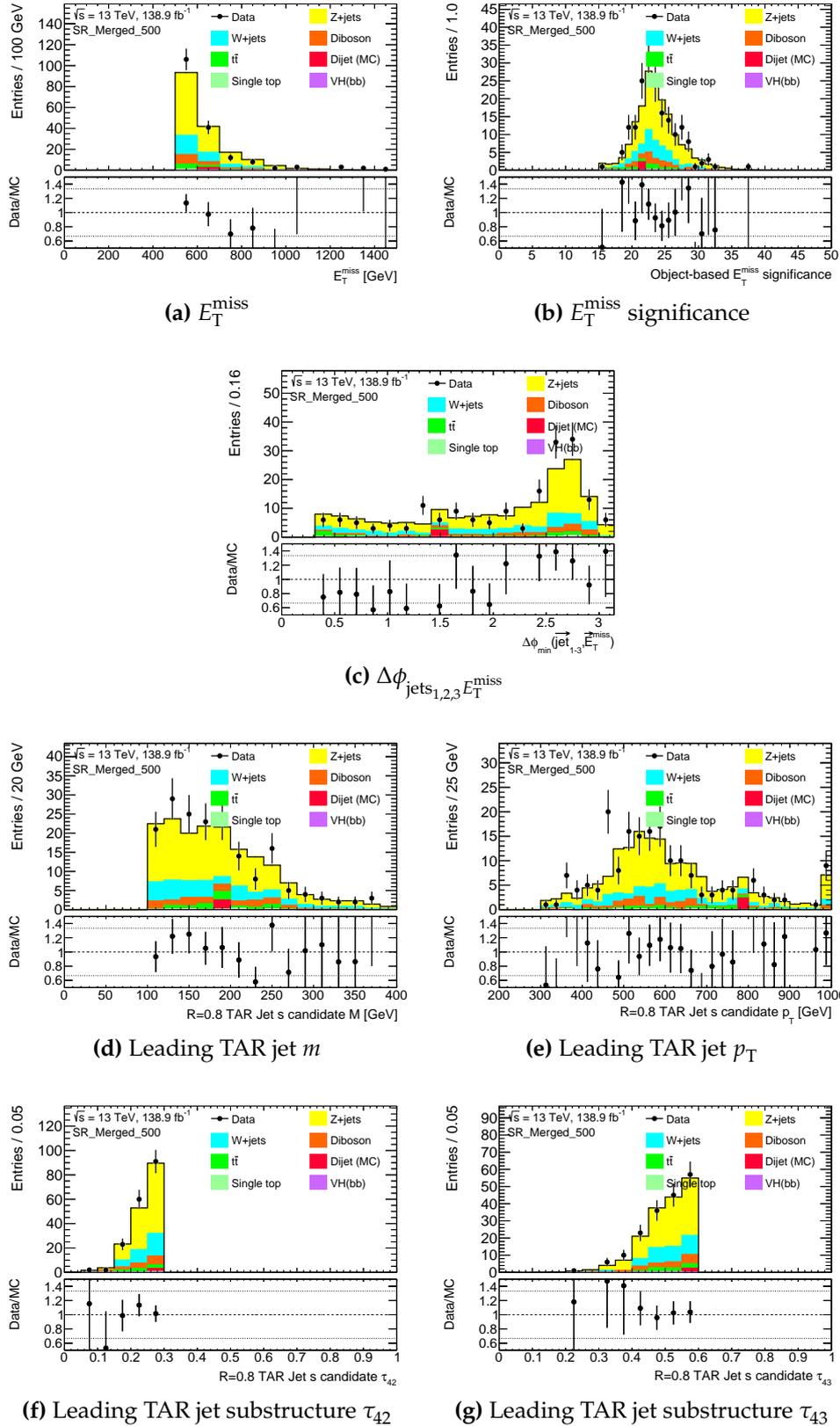


Figure B.1.: MC distributions in 0 lepton signal region for the merged category with $E_T^{\text{miss}} > 500 \text{ GeV}$.

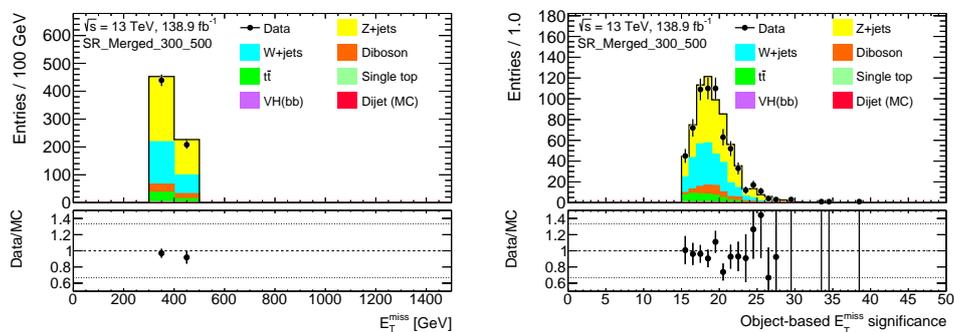
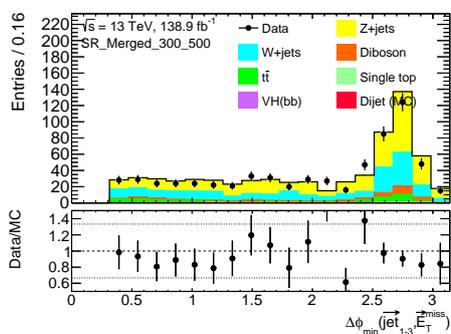
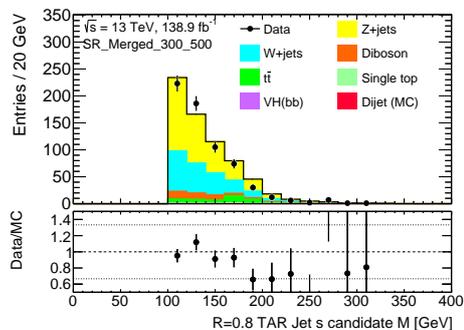
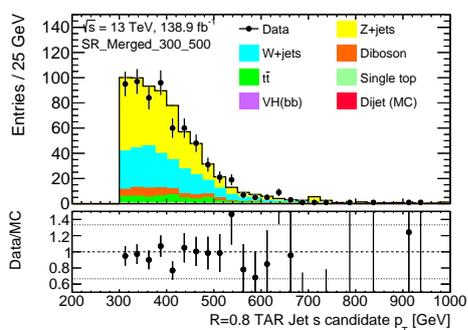
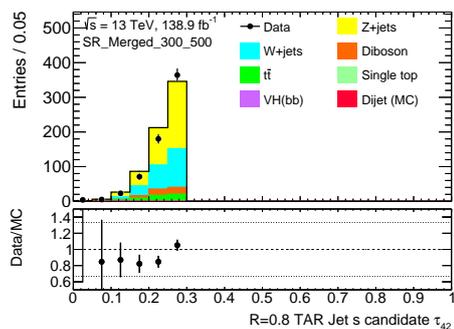
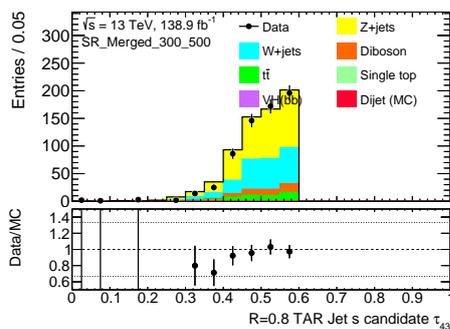
(a) E_T^{miss} (b) E_T^{miss} significance(c) $\Delta\phi_{\text{jets } 1,2,3} E_T^{\text{miss}}$ (d) Leading TAR jet m (e) Leading TAR jet p_T (f) Leading TAR jet substructure τ_{42} (g) Leading TAR jet substructure τ_{43}

Figure B.2.: MC distributions in 0 lepton signal region for the merged category with $300\text{ GeV} < E_T^{\text{miss}} < 500\text{ GeV}$.

Intermediate Category

The MC distributions for the 0 lepton signal region in the intermediate category are shown separately for $E_T^{\text{miss}} > 500\text{GeV}$ in Figure B.3, for $300\text{GeV} < E_T^{\text{miss}} < 500\text{GeV}$ in Figure B.4 and for $200\text{GeV} < E_T^{\text{miss}} < 300\text{GeV}$ in Figure B.5 respectively. All requirements of the event selection are applied.

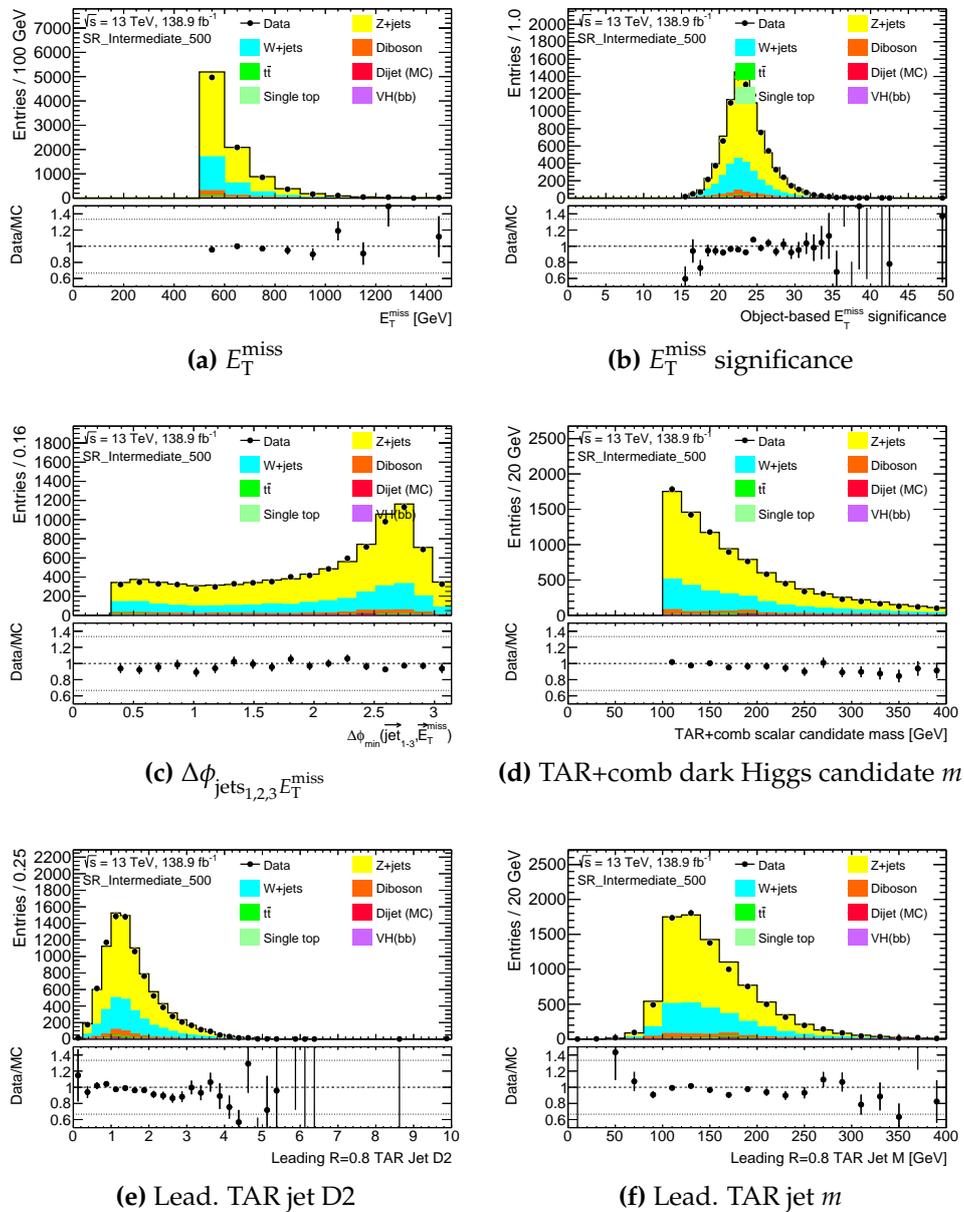


Figure B.3.: MC distributions in 0 lepton signal region for the intermediate category with $E_T^{\text{miss}} > 500\text{GeV}$.

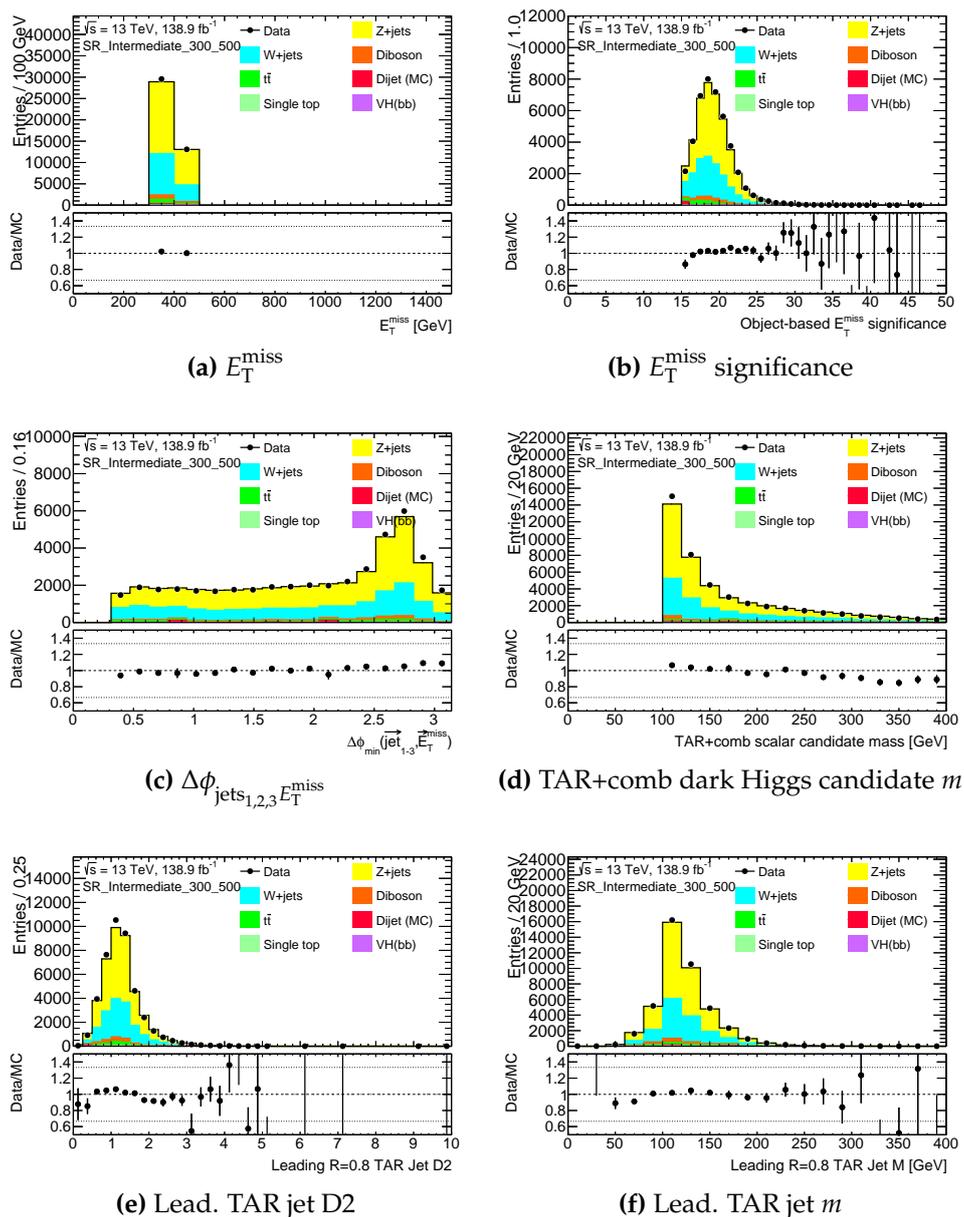


Figure B.4.: MC distributions in 0 lepton signal region for the intermediate category with $300\text{GeV} < E_T^{\text{miss}} < 500\text{GeV}$.

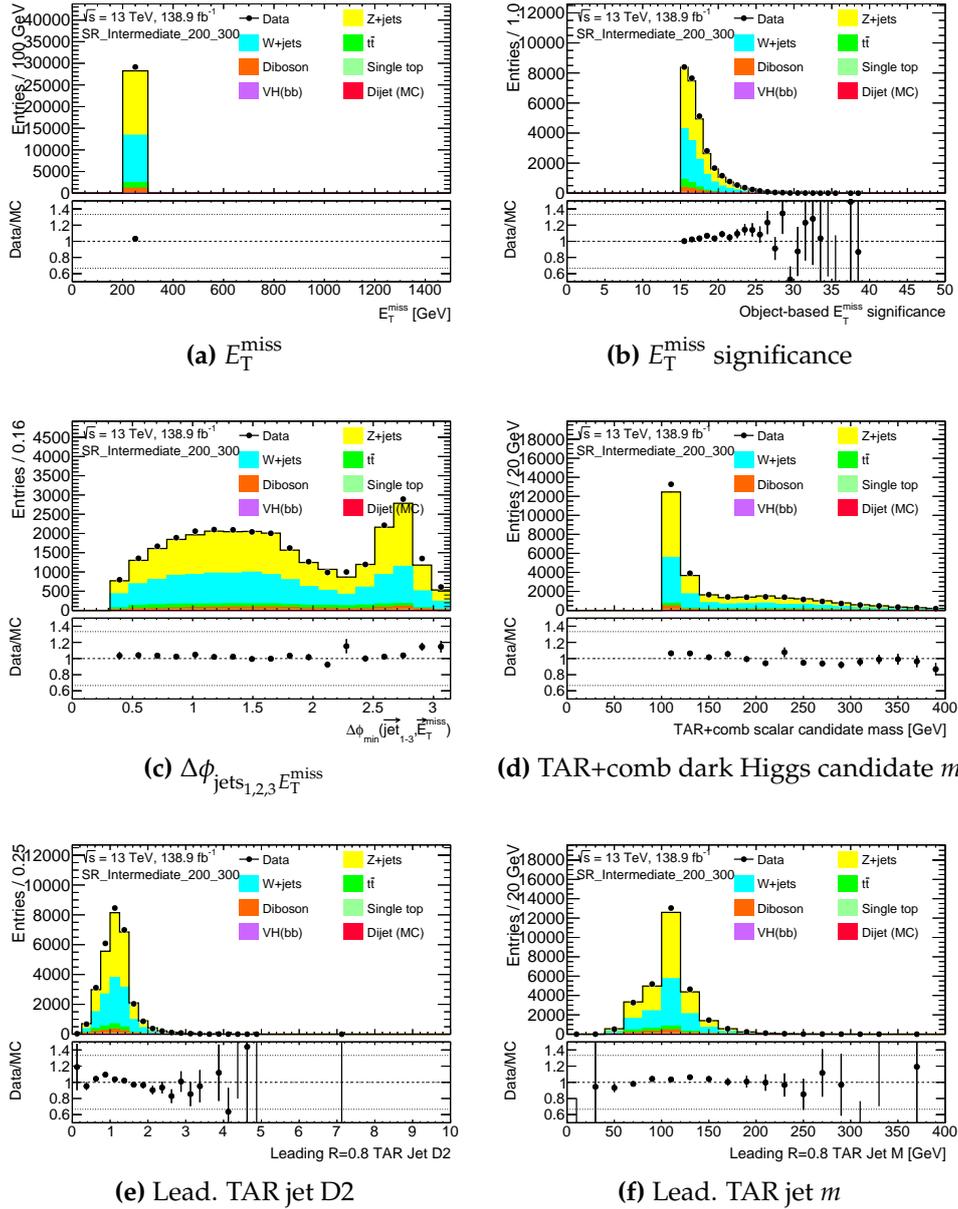


Figure B.5.: MC distributions in 0 lepton signal region for the intermediate category with $200\text{GeV} < E_T^{\text{miss}} < 300\text{GeV}$.

Resolved Category

The MC distributions for the 0 lepton signal region in the resolved category are shown separately for $E_T^{\text{miss}} > 500\text{GeV}$ in Figure B.6, for $300\text{GeV} < E_T^{\text{miss}} < 500\text{GeV}$ in Figure B.7 and for $200\text{GeV} < E_T^{\text{miss}} < 300\text{GeV}$ in Figure B.8 respectively. All requirements of the event selection are applied.

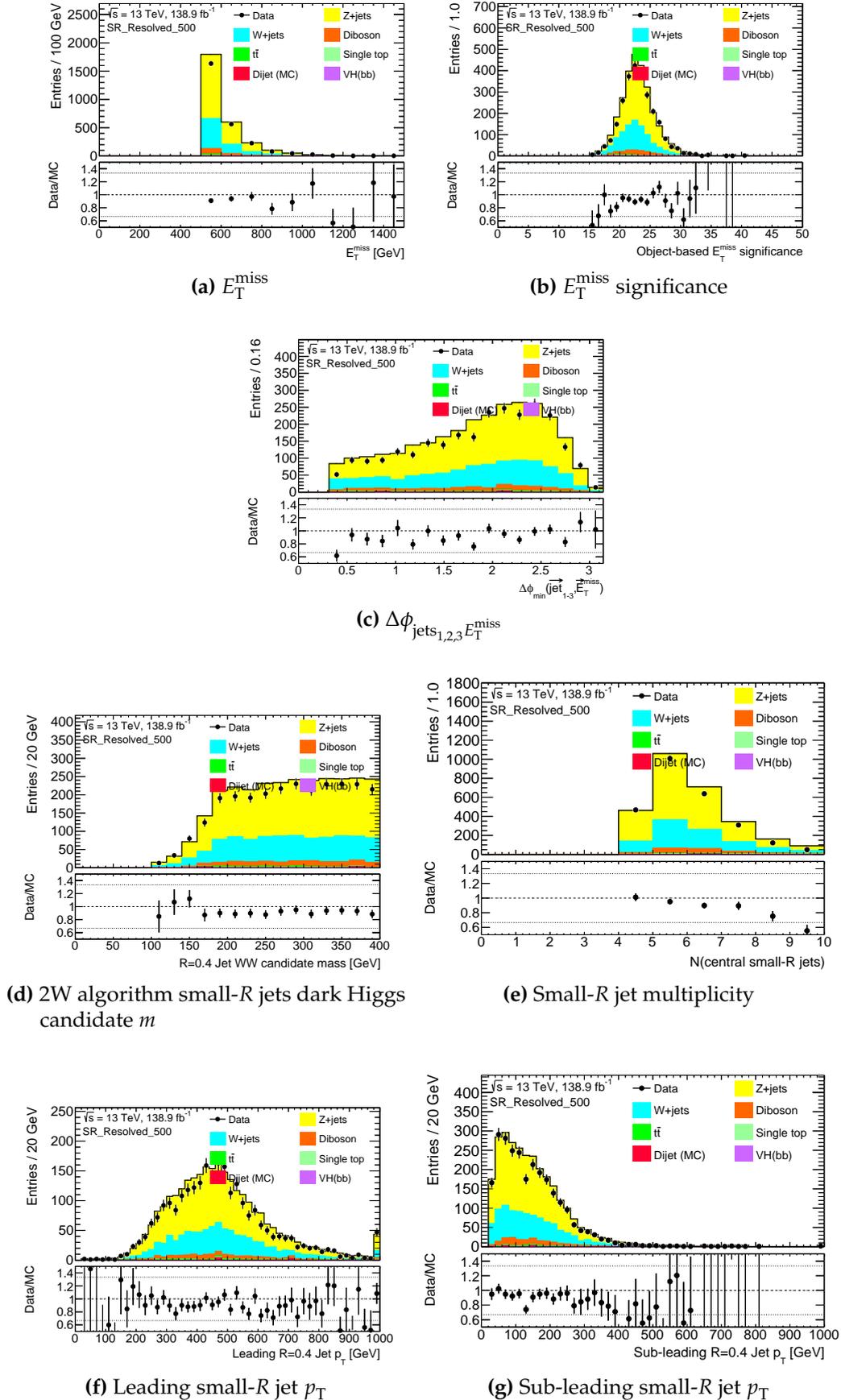


Figure B.6.: MC distributions in 0 lepton signal region for resolved category with $E_T^{\text{miss}} > 500\text{GeV}$.

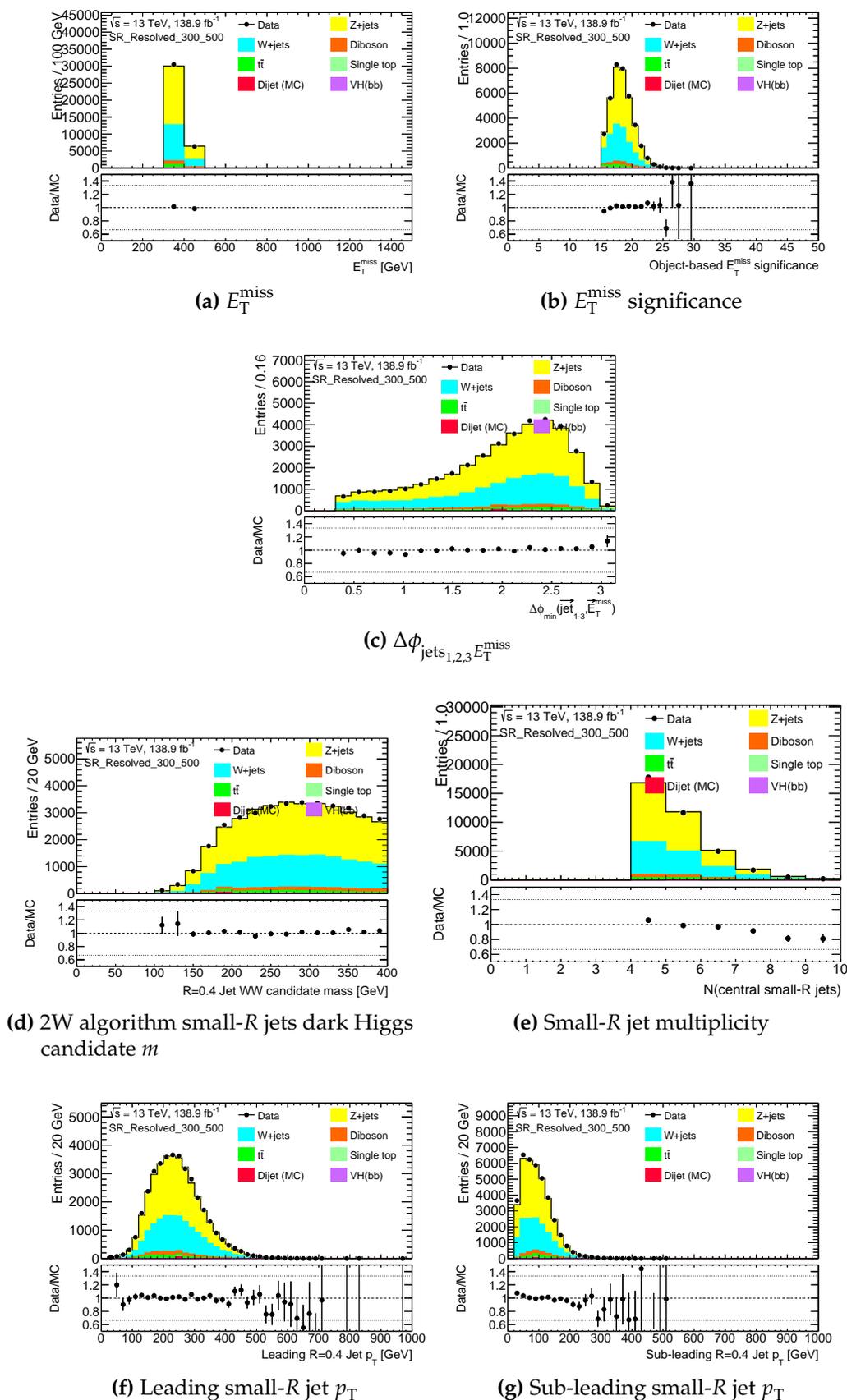


Figure B.7.: MC distributions in 0 lepton signal region for resolved category with $300\text{GeV} < E_T^{\text{miss}} < 500\text{GeV}$.

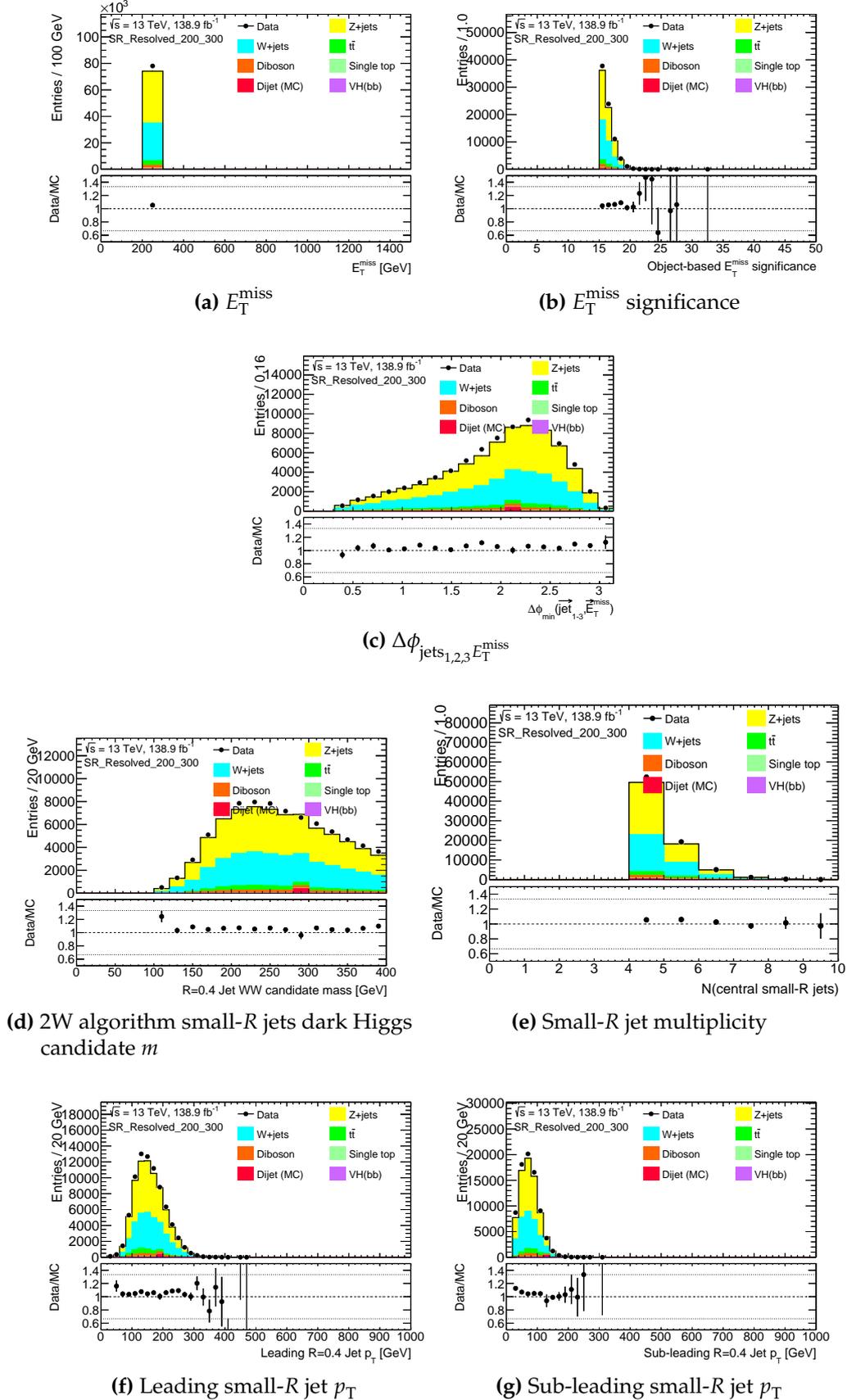


Figure B.8.: MC distributions in 0 lepton signal region for resolved category with $200\text{GeV} < E_T^{\text{miss}} < 300\text{GeV}$.

B.0.2. One μ Control Region

Merged Category

The data / MC comparisons for the 1 lepton control region in the merged category are shown separately for $E_T^{\text{miss,no}\mu} > 500\text{GeV}$ in Figure B.9 and for $300\text{GeV} < E_T^{\text{miss,no}\mu} < 500\text{GeV}$ in Figure B.10 respectively. All requirements of the event selection are applied.

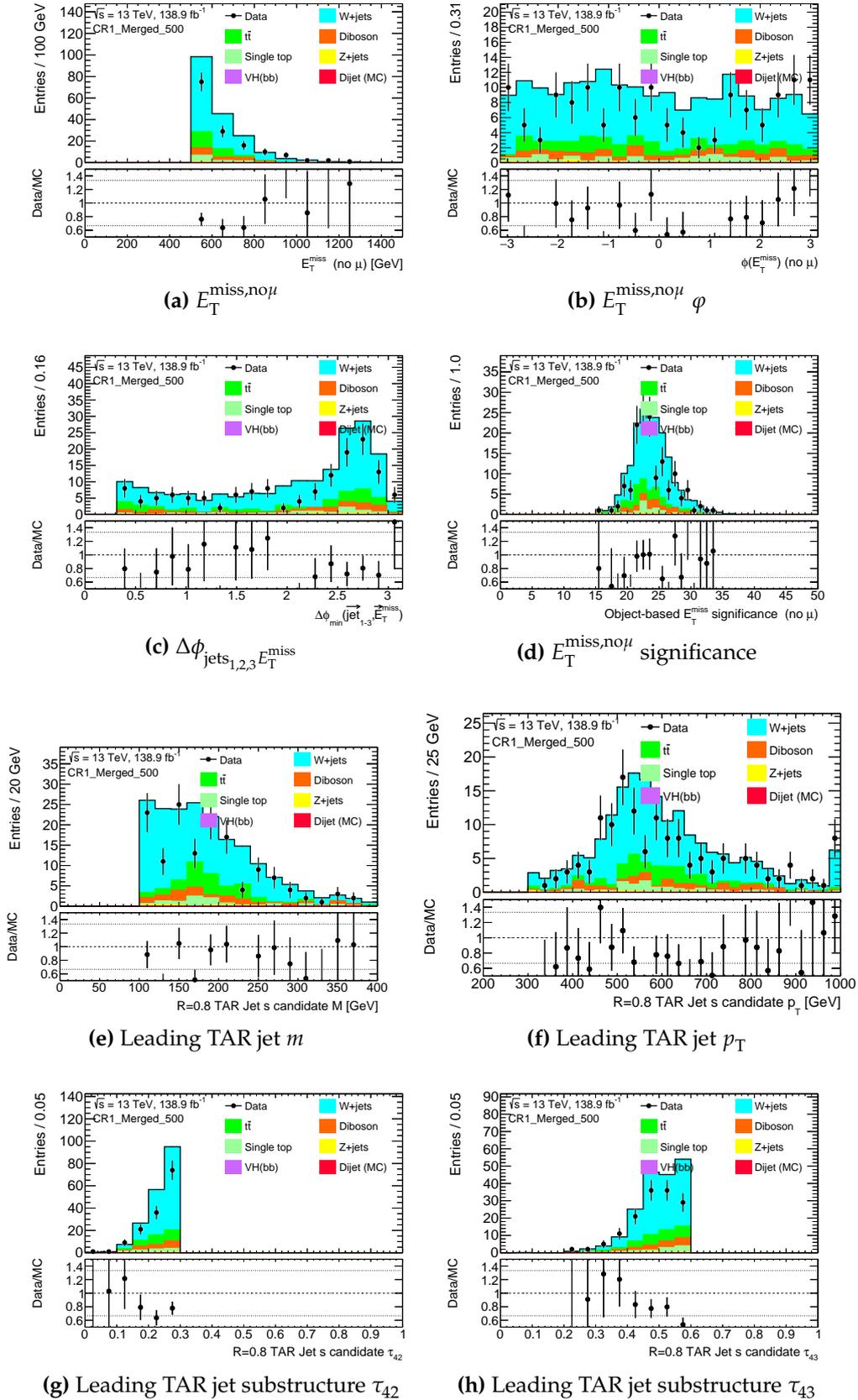


Figure B.9.: Data / MC comparison in 1μ control region for the merged category with $E_T^{\text{miss,no}\mu} > 500 \text{ GeV}$.

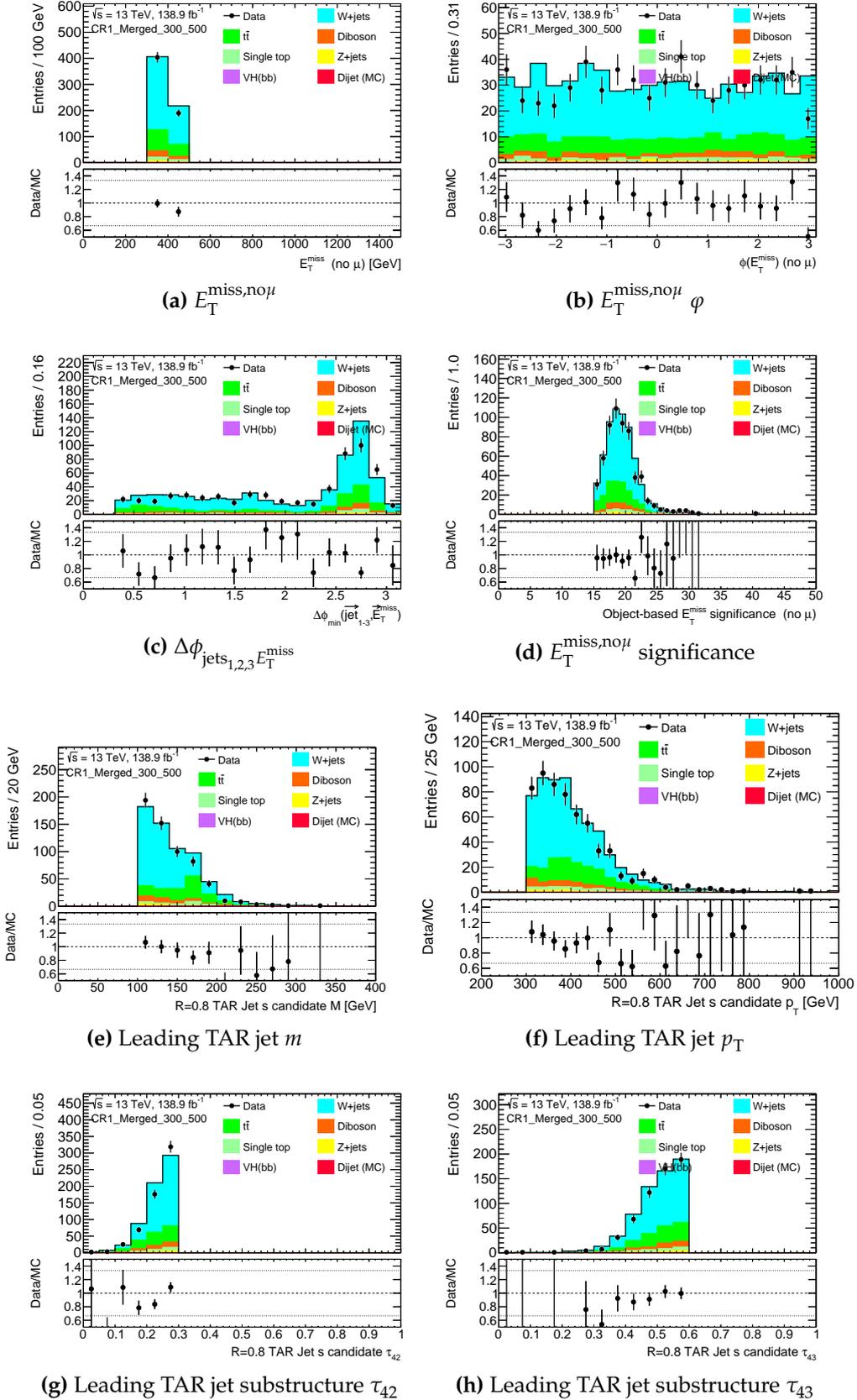


Figure B.10.: Data / MC comparison in 1μ control region for the merged category with $300 \text{ GeV} < E_T^{\text{miss,no}\mu} < 500 \text{ GeV}$.

Intermediate Category

The data / MC comparisons for the 1 lepton control region in the intermediate category are shown separately for $E_T^{\text{miss,no}\mu} > 500\text{GeV}$ in Figure B.11, for $300\text{GeV} < E_T^{\text{miss,no}\mu} < 500\text{GeV}$ in Figure B.12 and for $200\text{GeV} < E_T^{\text{miss,no}\mu} < 300\text{GeV}$ in Figure B.13 respectively. All requirements of the event selection are applied.

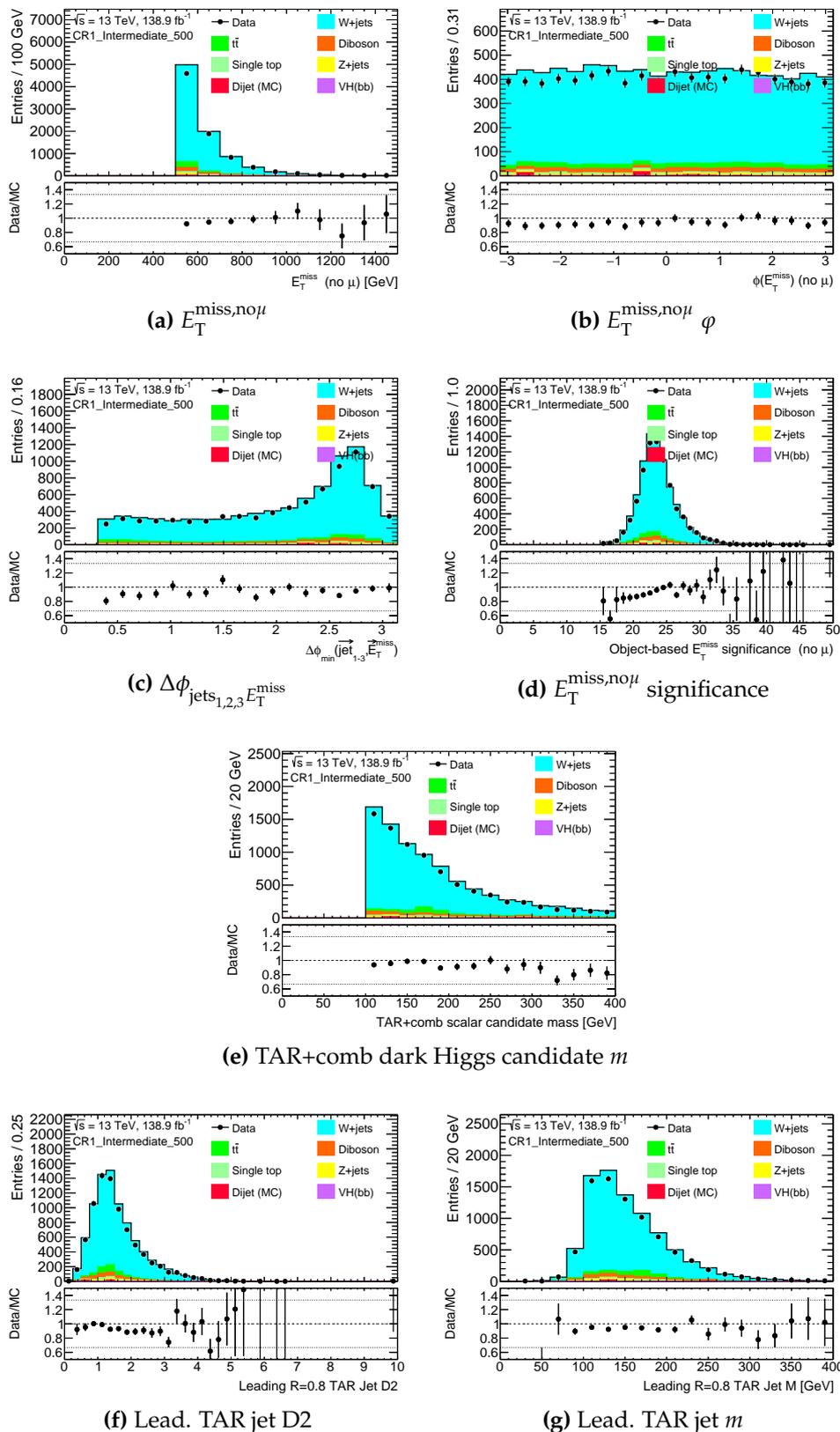


Figure B.11.: Data / MC comparison in 1μ control region for the intermediate category with $E_T^{\text{miss,no}\mu} > 500 \text{ GeV}$.

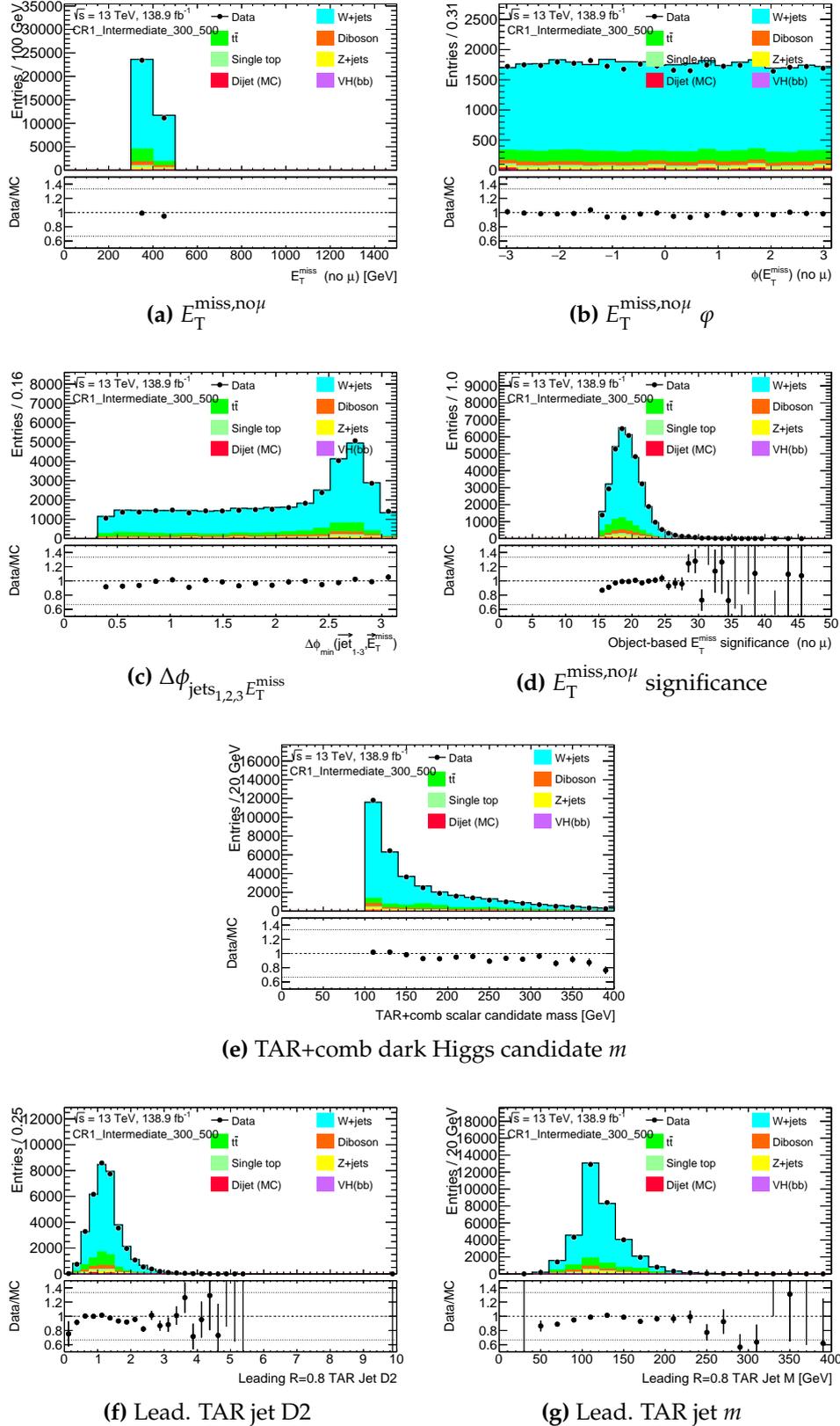


Figure B.12.: Data / MC comparison in 1μ control region for the intermediate category with $300 \text{ GeV} < E_T^{\text{miss,no}\mu} < 500 \text{ GeV}$.

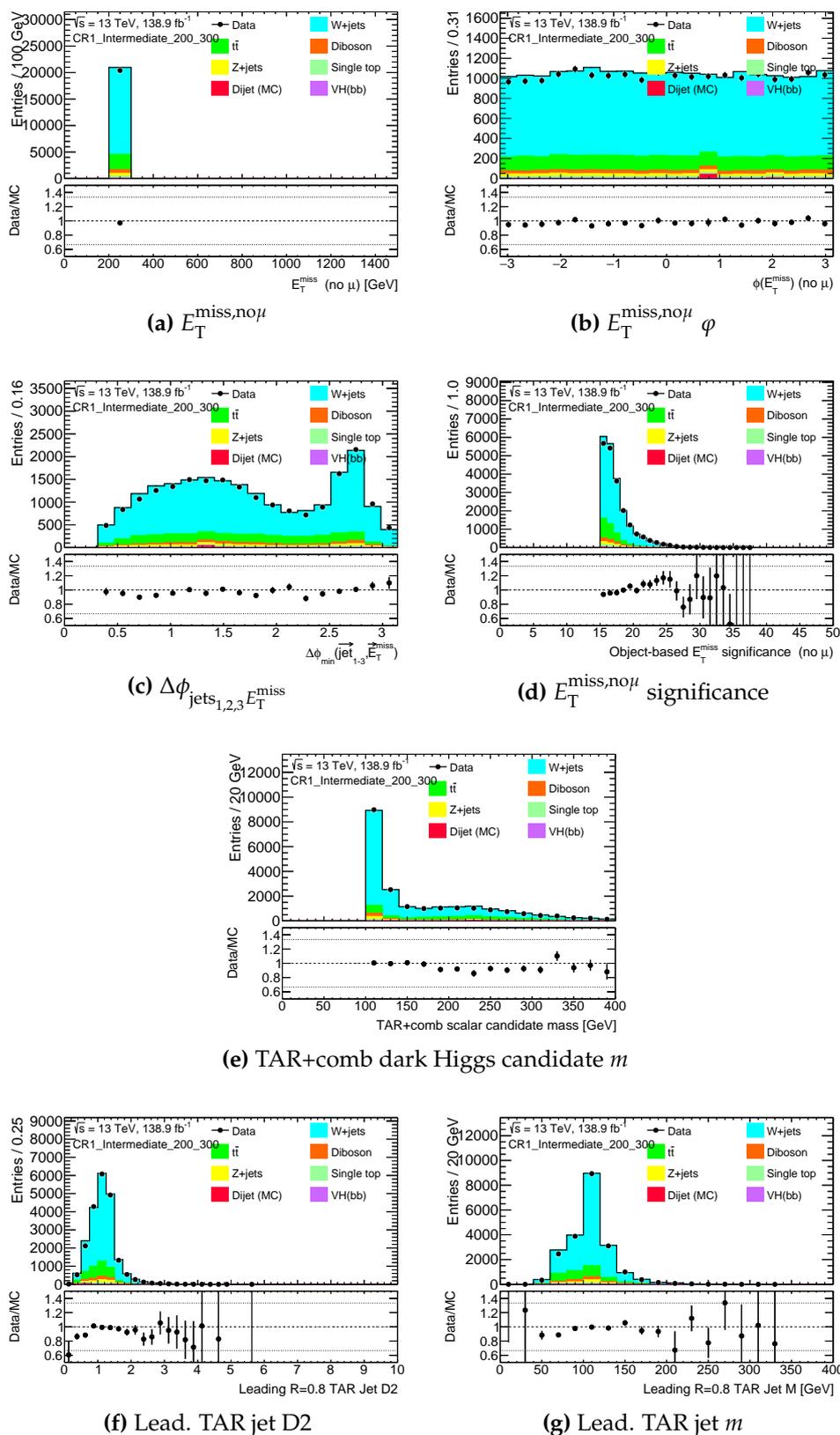


Figure B.13.: Data / MC comparison in 1μ control region for the intermediate category with $200\text{GeV} < E_T^{\text{miss,no}\mu} < 300\text{GeV}$.

Resolved Category

The data / MC comparisons for the 1 lepton control region in the resolved category are shown separately for $E_T^{\text{miss,no}\mu} > 500\text{GeV}$ in Figure B.14, for $300\text{GeV} < E_T^{\text{miss,no}\mu} < 500\text{GeV}$ in Figure B.15 and for $200\text{GeV} < E_T^{\text{miss,no}\mu} < 300\text{GeV}$ in Figure B.16 respectively. All requirements of the event selection are applied.

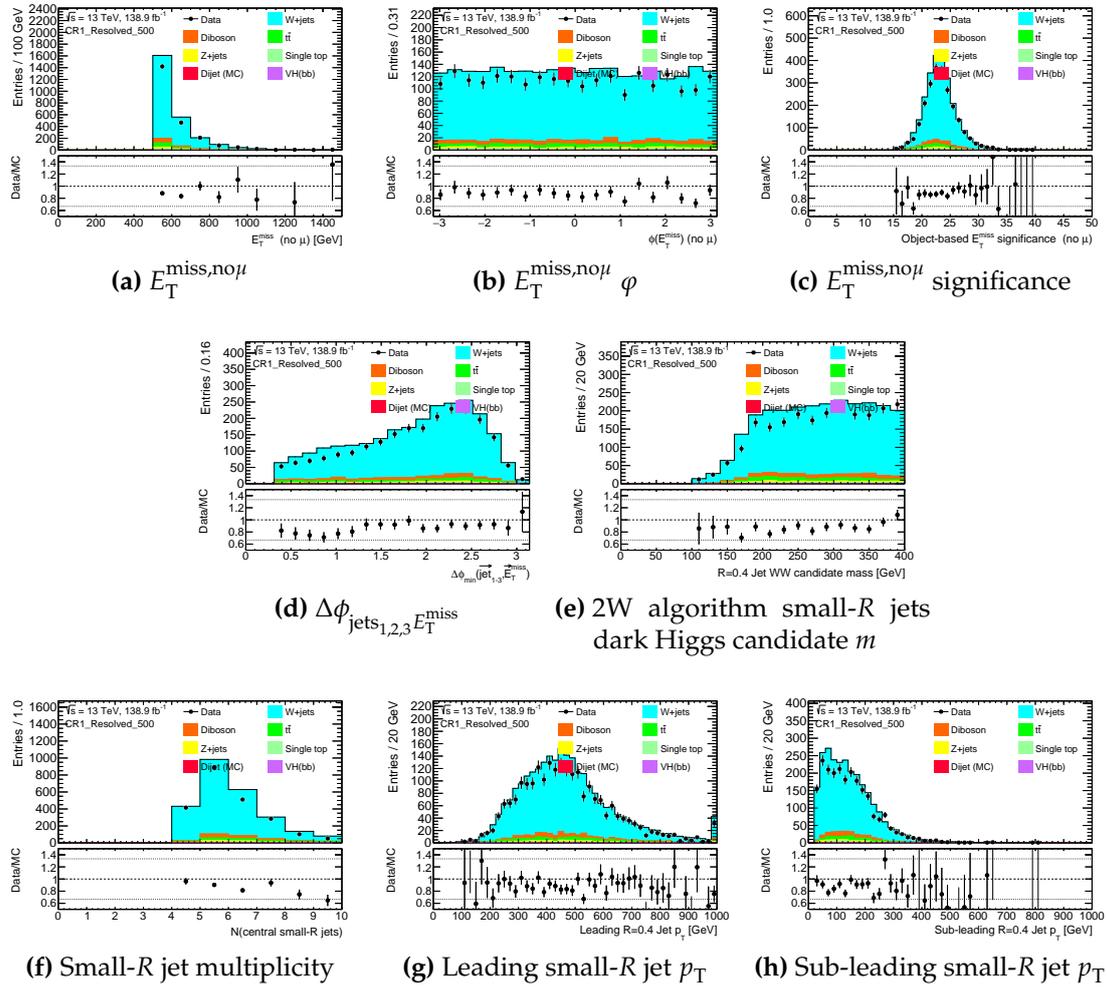


Figure B.14.: Data / MC comparison in 1 μ control region for resolved category with $E_T^{\text{miss,no}\mu} > 500\text{GeV}$.

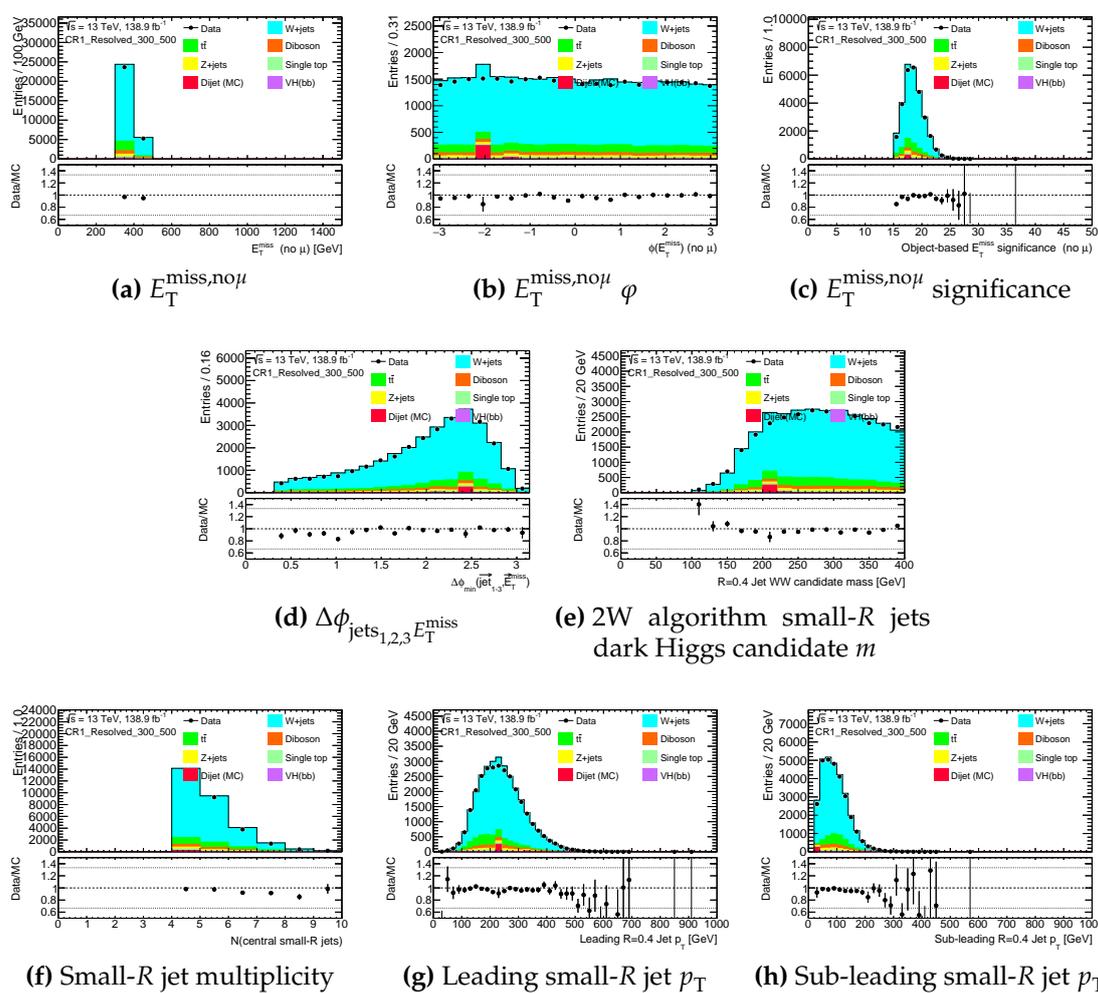


Figure B.15.: Data / MC comparison in 1μ control region for resolved category with $300 \text{ GeV} < E_T^{\text{miss,no}\mu} < 500 \text{ GeV}$.

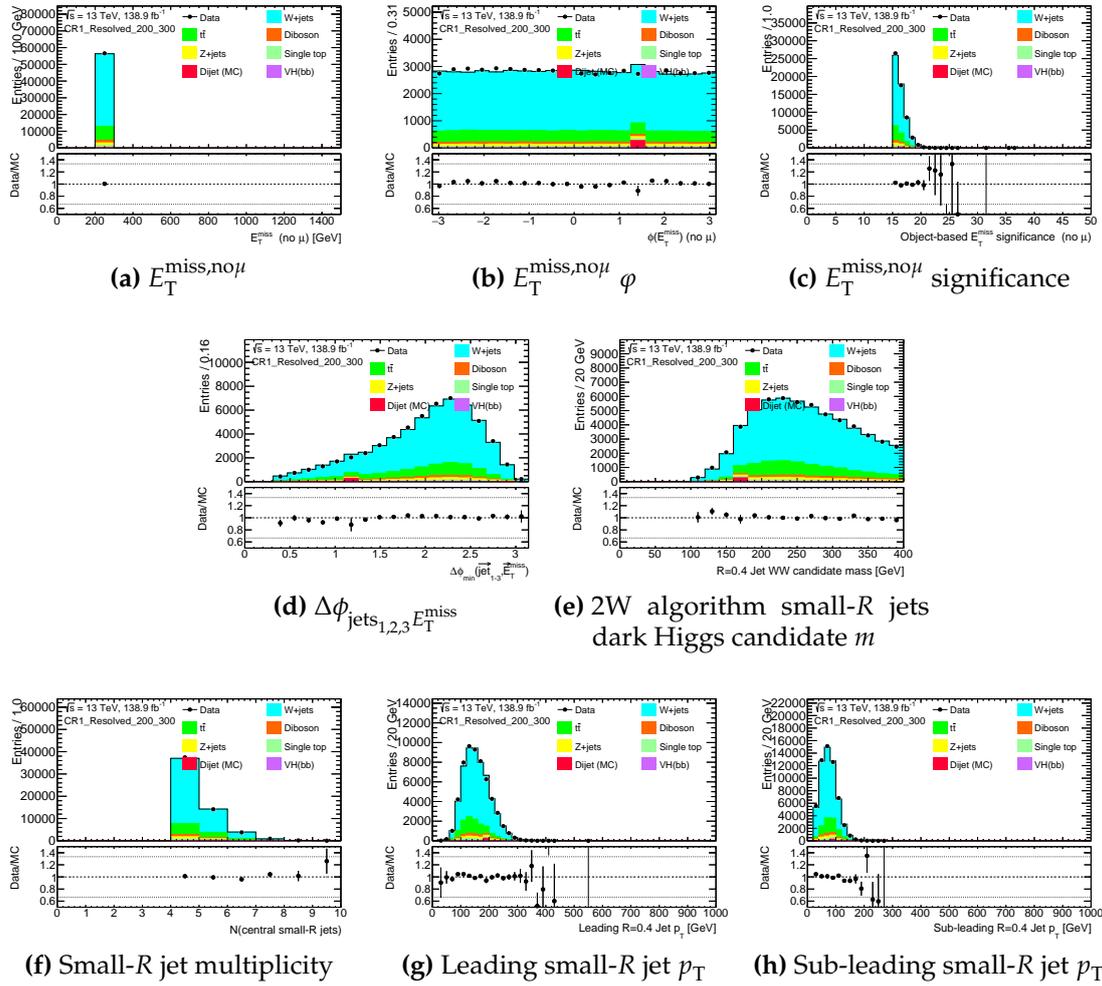
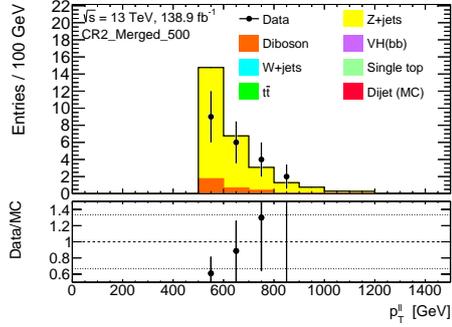
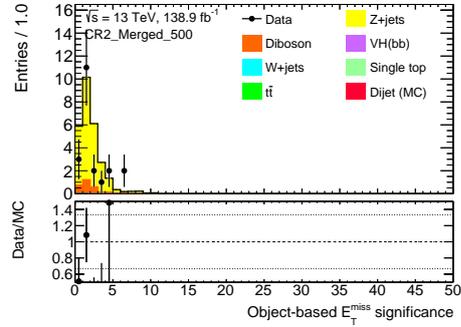
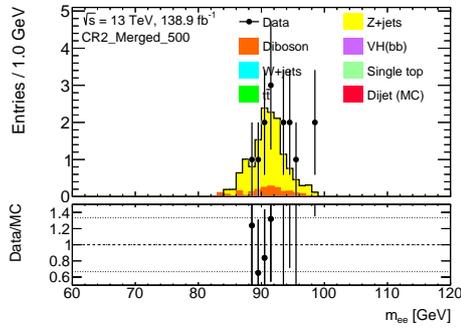


Figure B.16.: Data / MC comparison in 1μ control region for resolved category with $200 \text{ GeV} < E_T^{\text{miss,no}\mu} < 300 \text{ GeV}$.

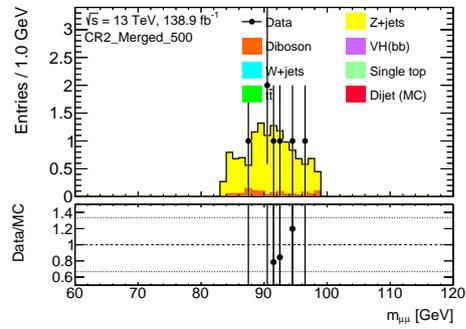
B.0.3. Two Leptons Control Region

Merged Category

The data / MC comparisons for the 2 leptons control region in the merged category are shown separately for $p_T^{\ell\ell} > 500\text{GeV}$ in Figure B.17 and for $300\text{GeV} < p_T^{\ell\ell} < 500\text{GeV}$ in Figure B.18 respectively. All requirements of the event selection are applied.

(a) Di-lepton $p_T^{\ell\ell}$ (b) E_T^{miss} significance

(c) di-electron system invariant mass



(d) di-muon system invariant mass

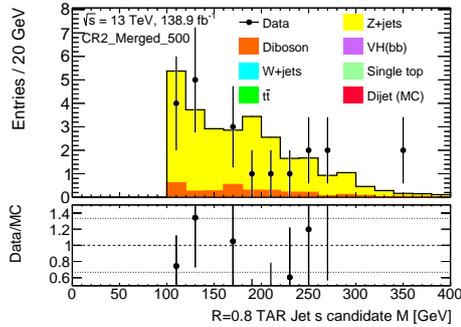
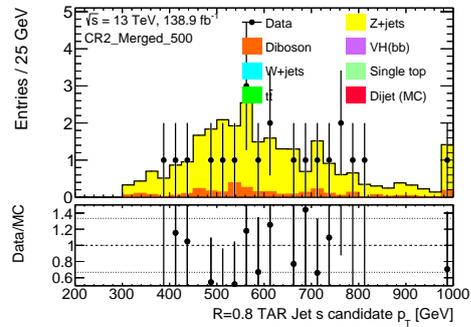
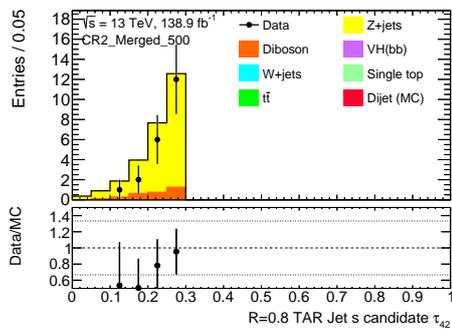
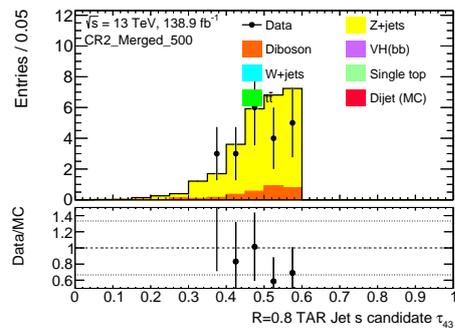
(e) Leading TAR jet m (f) Leading TAR jet p_T (g) Leading TAR jet substructure τ_{42} (h) Leading TAR jet substructure τ_{43}

Figure B.17.: Data / MC comparison in 2ℓ control region for the merged category with $p_T^{\ell\ell} > 500 \text{ GeV}$.

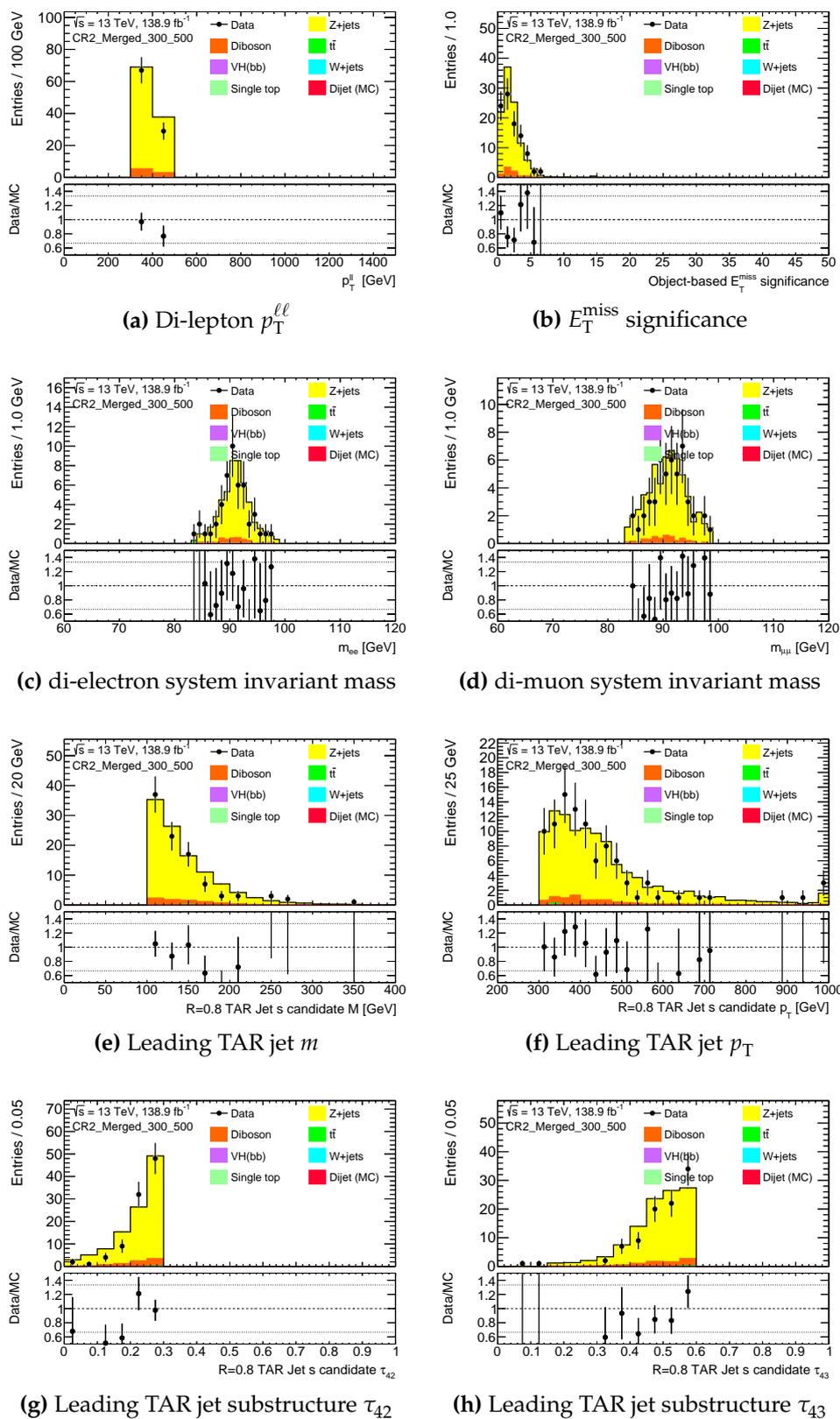
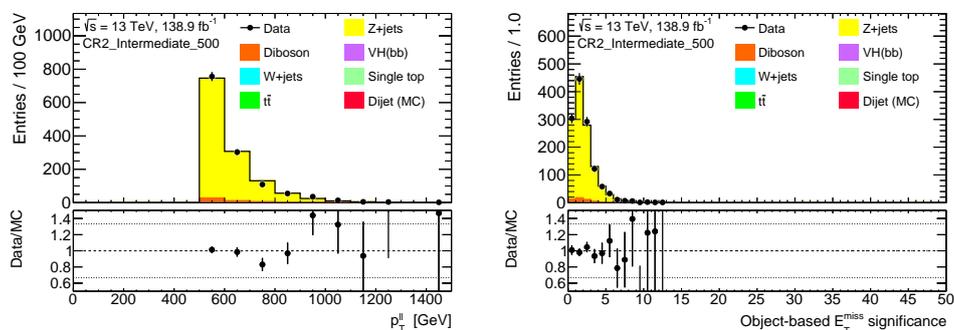


Figure B.18.: Data / MC comparison in 2ℓ control region for the merged category with $300\text{GeV} < p_T^{\ell\ell} < 500\text{GeV}$.

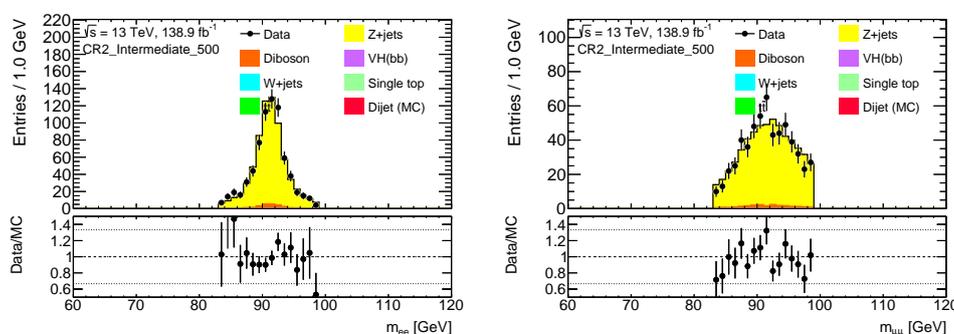
Intermediate Category

The data / MC comparisons for the 2 lepton control region in the intermediate category are shown separately for $p_T^{\ell\ell} > 500\text{GeV}$ in Figure B.19, for $300\text{GeV} < p_T^{\ell\ell} < 500\text{GeV}$ in Figure B.20, and for $200\text{GeV} < p_T^{\ell\ell} < 300\text{GeV}$ in Figure B.21 respectively. All requirements of the event selection are applied.



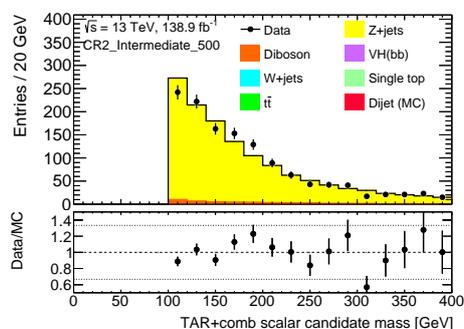
(a) Di-lepton $p_T^{\ell\ell}$

(b) E_T^{miss} significance

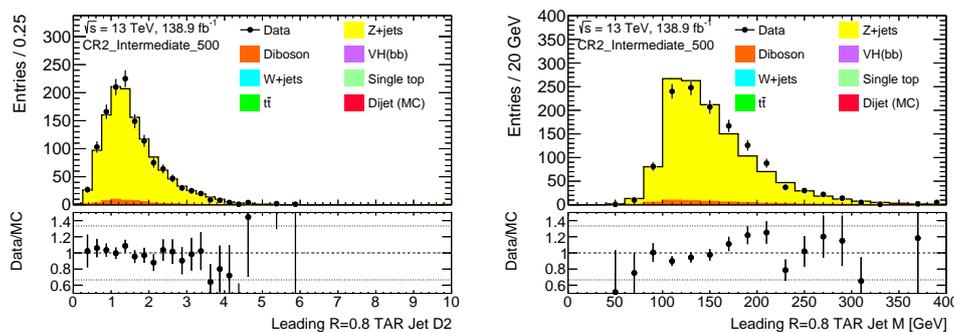


(c) di-electron system invariant mass

(d) di-muon system invariant mass



(e) TAR+comb dark Higgs candidate m



(f) Lead. TAR jet D2

(g) Lead. TAR jet m

Figure B.19.: Data / MC comparison in 2ℓ control region for the intermediate category with $p_T^{\ell\ell} > 500 \text{ GeV}$.

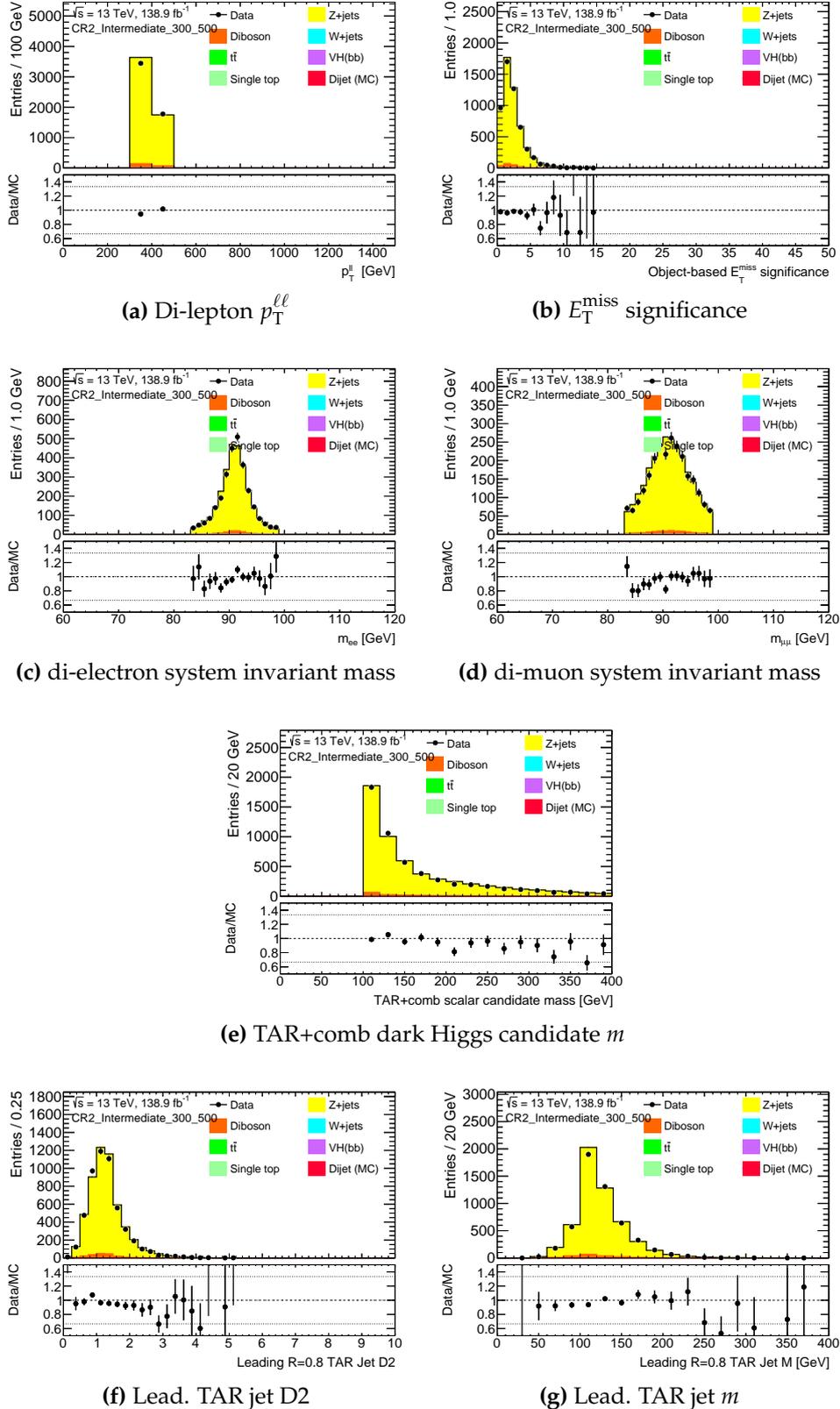
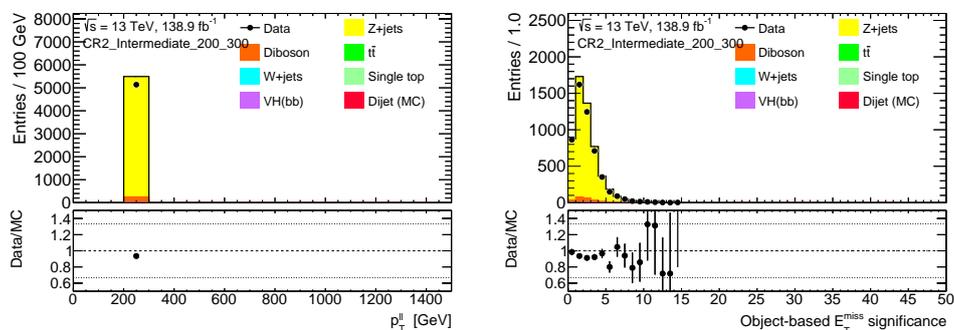
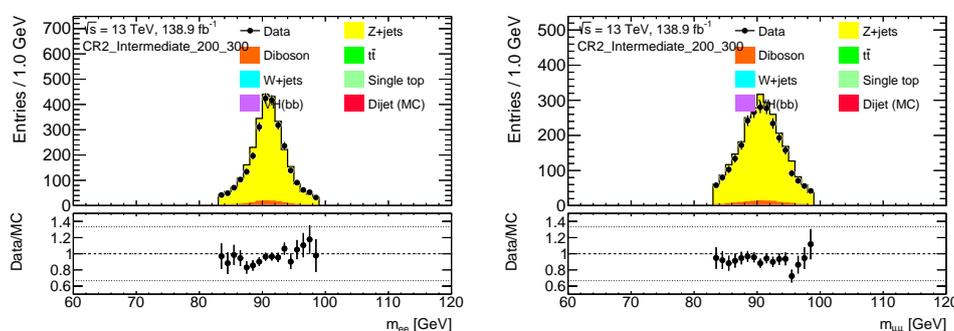


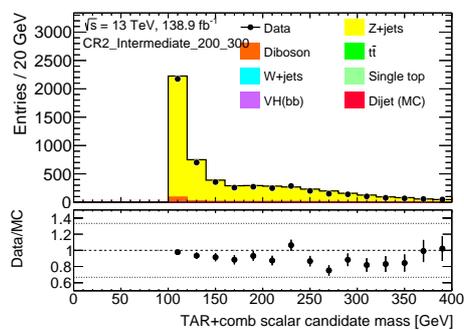
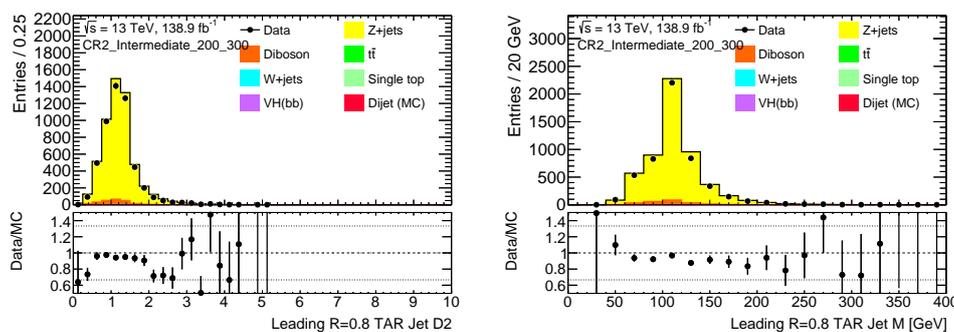
Figure B.20.: Data / MC comparison in 2ℓ control region for the intermediate category with $300\text{GeV} < p_T^{\ell\ell} < 500\text{GeV}$.


 (a) Di-lepton $p_T^{\ell\ell}$

 (b) E_T^{miss} significance


(c) di-electron system invariant mass

(d) di-muon system invariant mass


 (e) TAR+comb dark Higgs candidate m


(f) Lead. TAR jet D2

 (g) Lead. TAR jet m

Figure B.21.: Data / MC comparison in 2ℓ control region for the intermediate category with $200\text{GeV} < p_T^{\ell\ell} < 300\text{GeV}$.

Resolved Category

The data / MC comparisons for the 2 lepton control region in the resolved category are shown separately for $p_T^{\ell\ell} > 500\text{GeV}$ in Figure B.19, for $300\text{GeV} < p_T^{\ell\ell} < 500\text{GeV}$ in Figure B.20, and for $200\text{GeV} < p_T^{\ell\ell} < 300\text{GeV}$ in Figure B.21 respectively. All requirements of the event selection are applied.

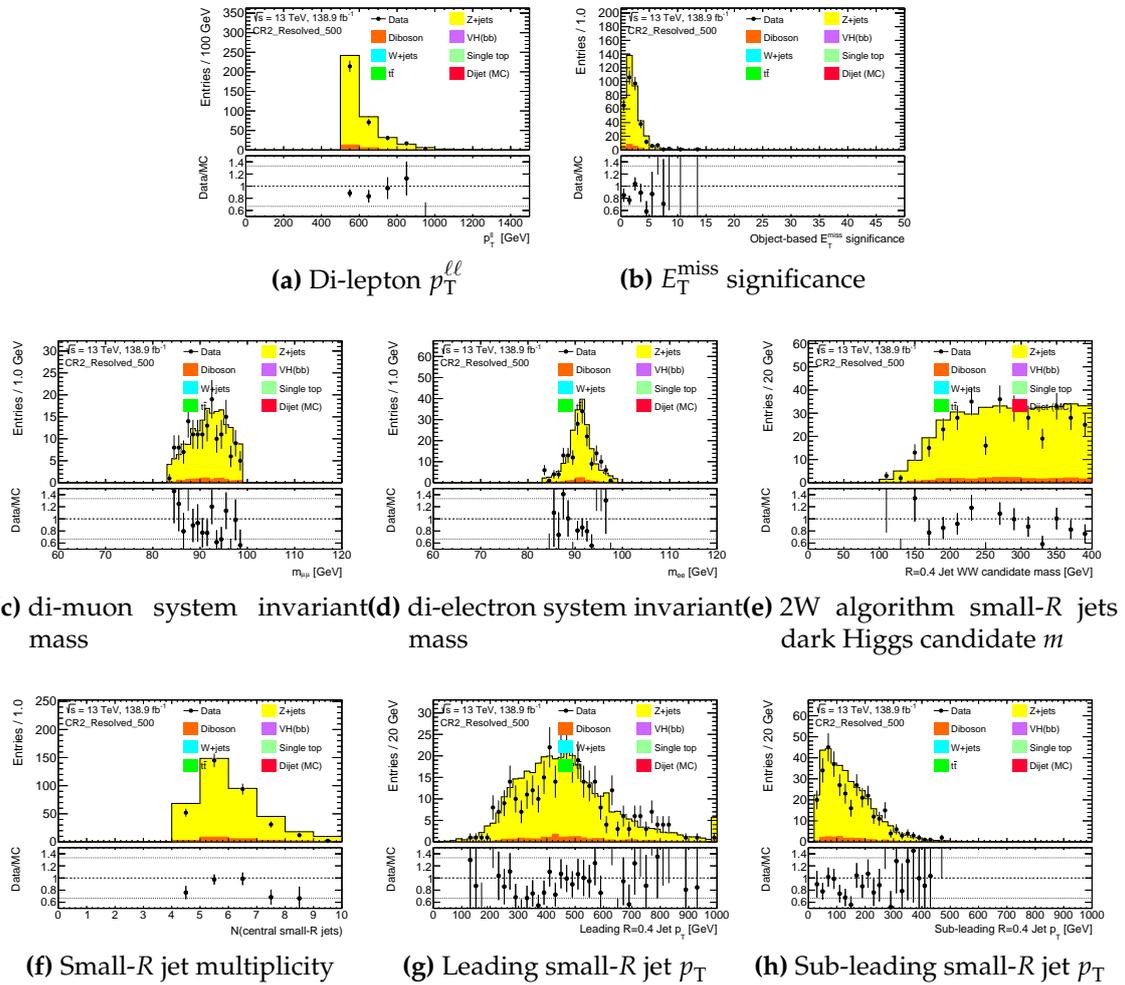


Figure B.22.: Data / MC comparison in 2 ℓ control region for resolved category with $p_T^{\ell\ell} > 500\text{GeV}$.

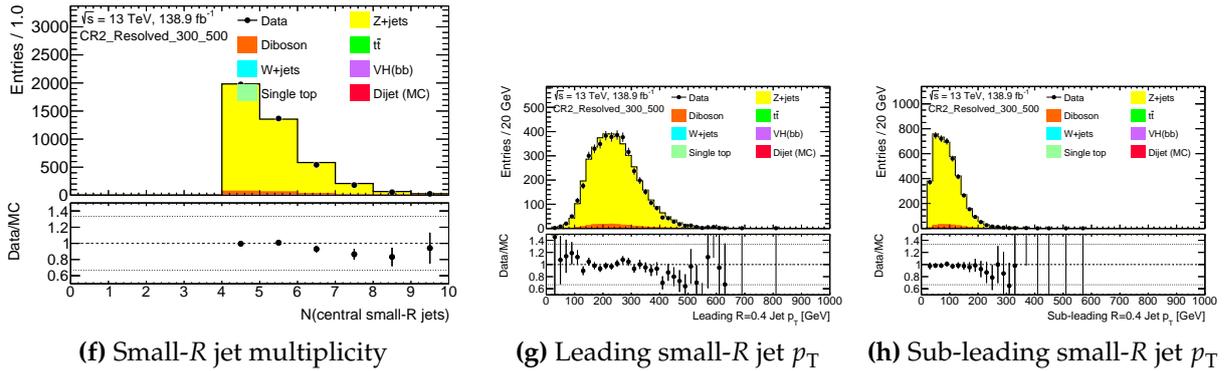
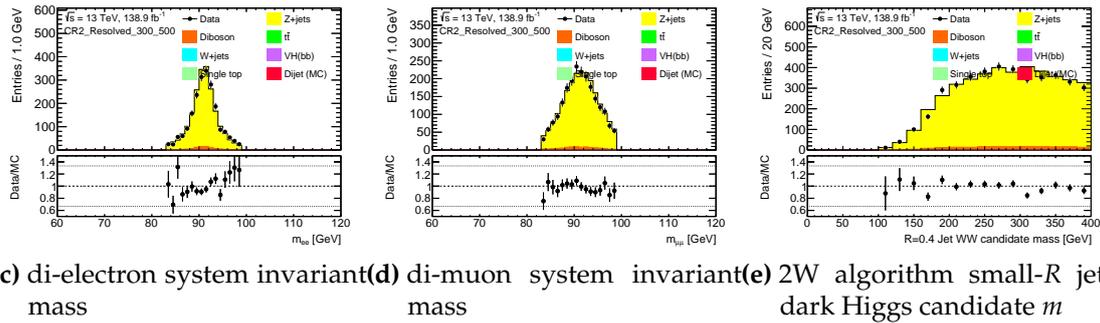
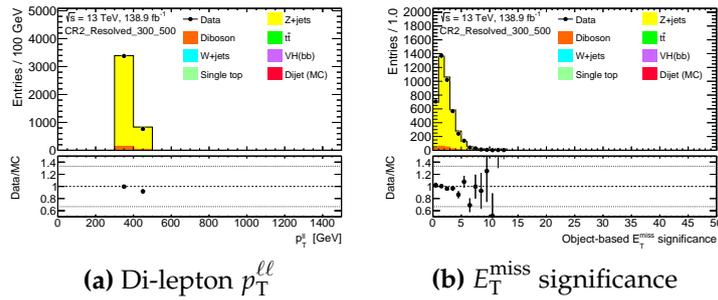


Figure B.23.: Data / MC comparison in 2ℓ control region for resolved category with $300\text{GeV} < p_T^{\ell\ell} < 500\text{GeV}$.

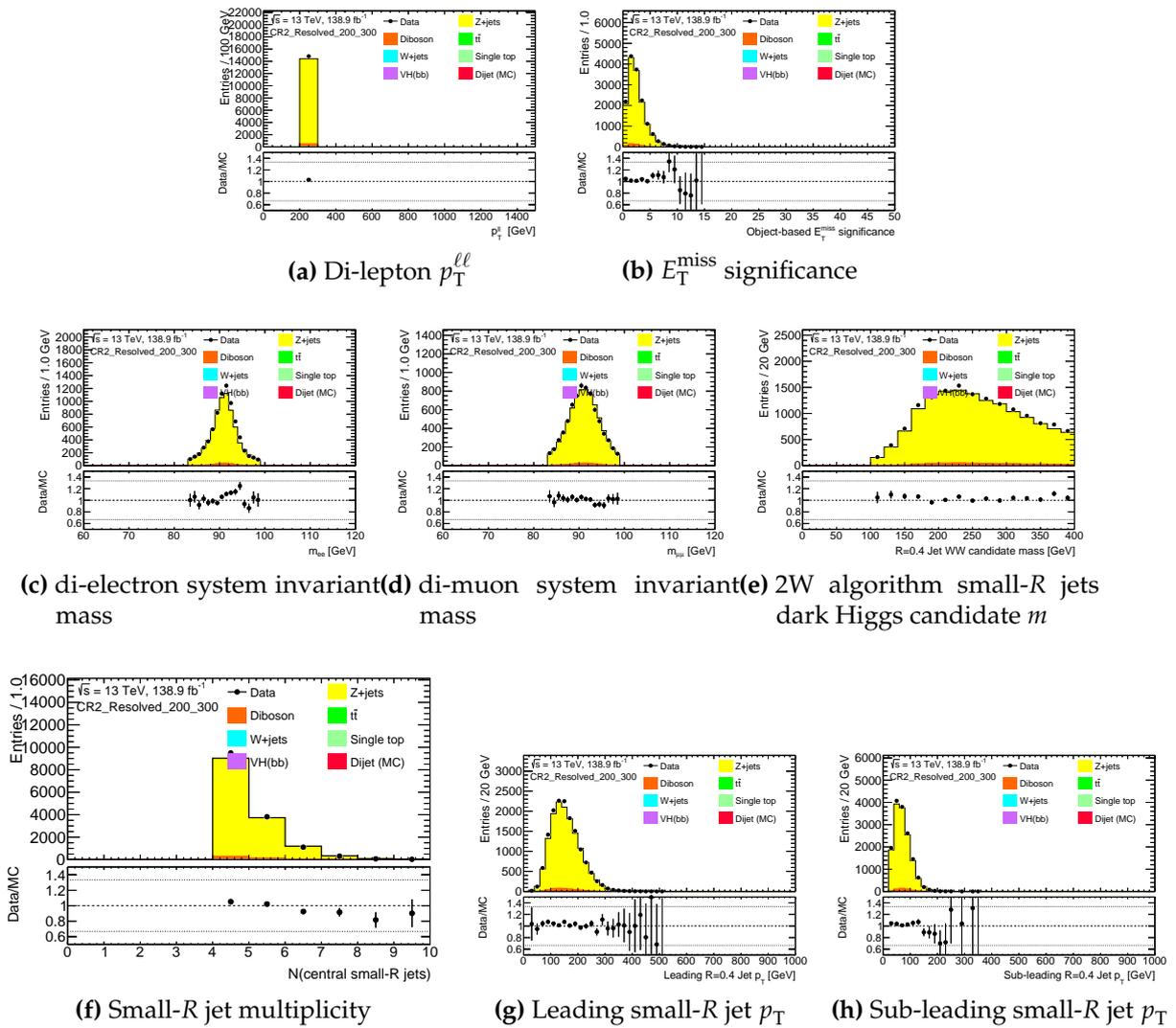


Figure B.24.: Data / MC comparison in 2ℓ control region for resolved category with $200\text{GeV} < p_T^{\ell\ell} < 300\text{GeV}$.

C. Systematics Uncertainties Additional Tables

C.0.1. Acceptance Uncertainties

Table C.1.: Generator-related acceptance uncertainties for the W -jets SM process in the signal region for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column.

	inclusive	200-300 GeV	300-500 GeV	more than 500GeV
SR merged	0.161	-	0.191	0.005
SR intermediate	0.420	0.470	0.032	0.322
SR resolved	0.103	0.085	0.184	0.247
CR1 merged	0.130	-	0.166	0.033
CR1 intermediate	0.335	0.390	0.110	0.044
CR1 resolved	0.095	0.066	0.175	0.266
CR2 merged	0.721	-	0.701	-
CR2 intermediate	0.347	0.347	-	-
CR2 resolved	0.018	0.027	0.146	-

Table C.2.: Generator-related acceptance uncertainties for the Z+jets SM process in the signal region for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column.

	inclusive	200-300 GeV	300-500 GeV	more than 500GeV
SR merged	0.127	-	0.158	0.033
SR intermediate	0.360	0.414	0.117	0.089
SR resolved	0.097	0.071	0.172	0.261
CR1 merged	0.125	-	0.160	0.012
CR1 intermediate	0.394	0.451	0.050	6.851
CR1 resolved	0.080	0.072	0.107	0.411
CR2 merged	0.096	-	0.106	0.046
CR2 intermediate	0.232	0.236	0.156	0.805
CR2 resolved	0.083	0.066	0.162	0.293

Table C.3.: PDF acceptance uncertainties for the W+jets SM process in the signal region for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column.

	inclusive	200-300 GeV	300-500 GeV	more than 500GeV
SR merged	0.013 / -0.001	-	0.010 / 0.000	0.026 / -0.006
SR intermediate	0.010 / 0.000	0.012 / 0.000	0.000 / -0.002	0.026 / 0.000
SR resolved	0.010 / 0.000	0.010 / 0.000	0.011 / 0.000	0.014 / 0.000
CR1 merged	0.027 / -0.007	-	0.025 / -0.005	0.034 / -0.010
CR1 intermediate	0.023 / -0.001	0.024 / -0.000	0.017 / -0.005	0.046 / -0.004
CR1 resolved	0.021 / 0.000	0.020 / 0.000	0.021 / -0.004	0.022 / -0.021
CR2 merged	0.008 / -0.043	-	0.003 / -0.050	0.089 / 0.000
CR2 intermediate	0.034 / 0.000	0.034 / 0.000	-	-
CR2 resolved	0.029 / -0.000	0.031 / -0.002	0.028 / -0.021	-

Table C.4.: PDF acceptance uncertainties for the Z+jets SM process in the signal region for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column.

	inclusive	200-300 GeV	300-500 GeV	more than 500GeV
SR merged	0.027 / 0.000	-	0.025 / 0.000	0.032 / 0.000
SR intermediate	0.022 / 0.000	0.023 / 0.000	0.019 / 0.000	0.006 / -0.013
SR resolved	0.021 / 0.000	0.022 / 0.000	0.019 / 0.000	0.018 / 0.000
CR1 merged	0.024 / 0.000	-	0.023 / 0.000	0.028 / -0.005
CR1 intermediate	0.021 / 0.000	0.024 / 0.000	0.000 / -0.008	0.000 / -0.188
CR1 resolved	0.022 / 0.000	0.024 / 0.000	0.015 / 0.000	0.021 / -0.006
CR2 merged	0.027 / 0.000	-	0.027 / 0.000	0.027 / 0.000
CR2 intermediate	0.025 / 0.000	0.025 / 0.000	0.019 / 0.000	0.021 / 0.000
CR2 resolved	0.020 / 0.000	0.020 / 0.000	0.021 / 0.000	0.019 / 0.000

Table C.5.: Scale acceptance uncertainties for the W+jets SM process in the signal region for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column.

	inclusive	200-300 GeV	300-500 GeV	more than 500GeV
SR merged	0.319 / -0.218	-	0.321 / -0.218	0.321 / -0.219
SR intermediate	0.381 / -0.239	0.388 / -0.241	0.332 / -0.223	0.482 / -0.275
SR resolved	0.404 / -0.250	0.413 / -0.252	0.363 / -0.242	0.328 / -0.230
CR1 merged	0.295 / -0.207	-	0.296 / -0.206	0.292 / -0.208
CR1 intermediate	0.346 / -0.228	0.358 / -0.231	0.296 / -0.211	0.425 / -0.266
CR1 resolved	0.386 / -0.246	0.400 / -0.249	0.349 / -0.238	0.317 / -0.222
CR2 merged	0.357 / -0.235	-	0.366 / -0.241	0.274 / -0.149
CR2 intermediate	0.433 / -0.264	0.433 / -0.264	-	-
CR2 resolved	0.269 / -0.266	0.256 / -0.264	0.505 / -0.295	-

Table C.6.: Scale acceptance uncertainties for the Z+jets SM process in the signal region for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column.

	inclusive	200-300 GeV	300-500 GeV	more than 500GeV
SR merged	0.297 / -0.207	-	0.300 / -0.206	0.288 / -0.209
SR intermediate	0.343 / -0.229	0.354 / -0.234	0.293 / -0.207	0.425 / -0.269
SR resolved	0.388 / -0.246	0.402 / -0.248	0.351 / -0.239	0.313 / -0.219
CR1 merged	0.310 / -0.217	-	0.314 / -0.218	0.297 / -0.211
CR1 intermediate	0.385 / -0.245	0.397 / -0.249	0.318 / -0.220	0.495 / -0.674
CR1 resolved	0.403 / -0.247	0.419 / -0.250	0.334 / -0.233	0.362 / -0.236
CR2 merged	0.338 / -0.225	-	0.343 / -0.227	0.312 / -0.220
CR2 intermediate	0.405 / -0.252	0.408 / -0.253	0.336 / -0.222	0.465 / -0.265
CR2 resolved	0.401 / -0.251	0.409 / -0.253	0.369 / -0.242	0.314 / -0.225

Table C.7.: PDF acceptance uncertainties for a mono-s(W W) signal with $m_{Z'} = 1$ TeV, $m_\chi = 200$ GeV, $m_s = 160$ GeV for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column.

	inclusive	200-300 GeV	300-500 GeV	more than 500GeV
SR merged	0.088 / -0.181	-	0.088 / -0.183	0.093 / -0.178
SR intermediate	0.092 / -0.187	0.087 / -0.222	0.098 / -0.158	0.084 / -0.262
SR resolved	0.074 / -0.197	0.076 / -0.193	0.071 / -0.207	0.078 / -0.158

Table C.8.: PDF acceptance uncertainties for a mono-s(Z Z) signal with $m_{Z'} = 1$ TeV, $m_\chi = 200$ GeV, $m_s = 210$ GeV for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column.

	inclusive	200-300 GeV	300-500 GeV	more than 500GeV
SR merged	0.102 / -0.174	-	0.094 / -0.157	0.112 / -0.210
SR intermediate	0.079 / -0.192	0.123 / -0.167	0.087 / -0.211	0.016 / -0.264
SR resolved	0.089 / -0.183	0.091 / -0.181	0.089 / -0.188	0.068 / -0.172

Table C.9.: Scale acceptance uncertainties for a mono-s(W W) signal with $m_{Z'} = 1$ TeV, $m_\chi = 200$ GeV, $m_s = 160$ GeV for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column.

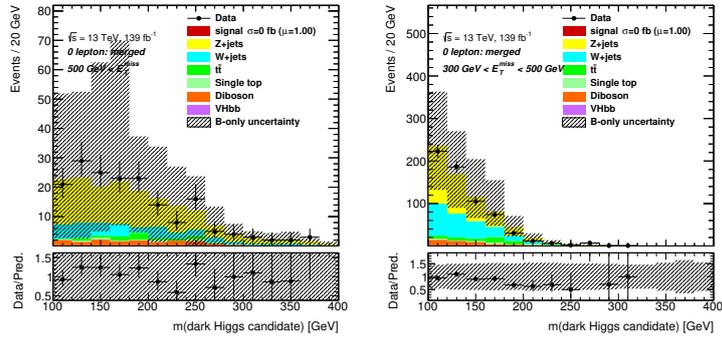
	inclusive	200-300 GeV	300-500 GeV	more than 500GeV
SR merged	0.084 / -0.147	-	0.077 / -0.125	0.097 / -0.184
SR intermediate	0.069 / -0.113	0.071 / -0.124	0.067 / -0.099	0.093 / -0.249
SR resolved	0.068 / -0.124	0.064 / -0.108	0.073 / -0.149	0.092 / -0.219

Table C.10.: Scale acceptance uncertainties for a mono-s(Z Z) signal with $m_{Z'} = 1$ TeV, $m_\chi = 200$ GeV, $m_s = 210$ GeV for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column.

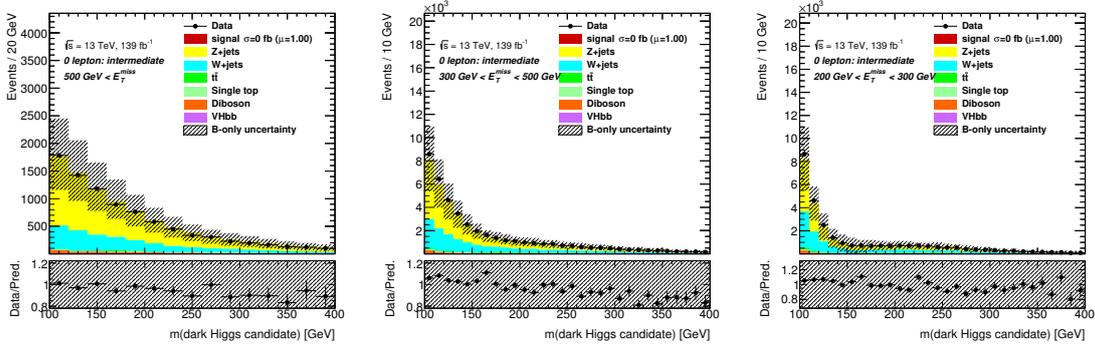
	inclusive	200-300 GeV	300-500 GeV	more than 500GeV
SR merged	0.088 / -0.148	-	0.079 / -0.126	0.099 / -0.175
SR intermediate	0.071 / -0.117	0.072 / -0.129	0.070 / -0.107	0.070 / -0.222
SR resolved	0.068 / -0.106	0.064 / -0.092	0.073 / -0.125	0.084 / -0.189

D. Pre-fit Distribution

In this Chapter, the pre-fit distributions used in the statistical treatment can be seen, for all the analysis regions and all the categories.



(a) merged $E_T^{\text{miss}} > 500 \text{ GeV}$ (b) merged $300 \text{ GeV} < E_T^{\text{miss}} < 500 \text{ GeV}$



(d) intermediate $E_T^{\text{miss}} > 500 \text{ GeV}$ (e) intermediate $300 \text{ GeV} < E_T^{\text{miss}} < 500 \text{ GeV}$ (f) intermediate $200 \text{ GeV} < E_T^{\text{miss}} < 300 \text{ GeV}$

Figure D.1.: MC distributions for the background model and overlaid data in the 0 lepton signal region for theory prediction (pre-fit). The bottom panel shows the relative uncertainty on the background prediction. The hatched band indicates the pre-fit uncertainty on the total background prediction.

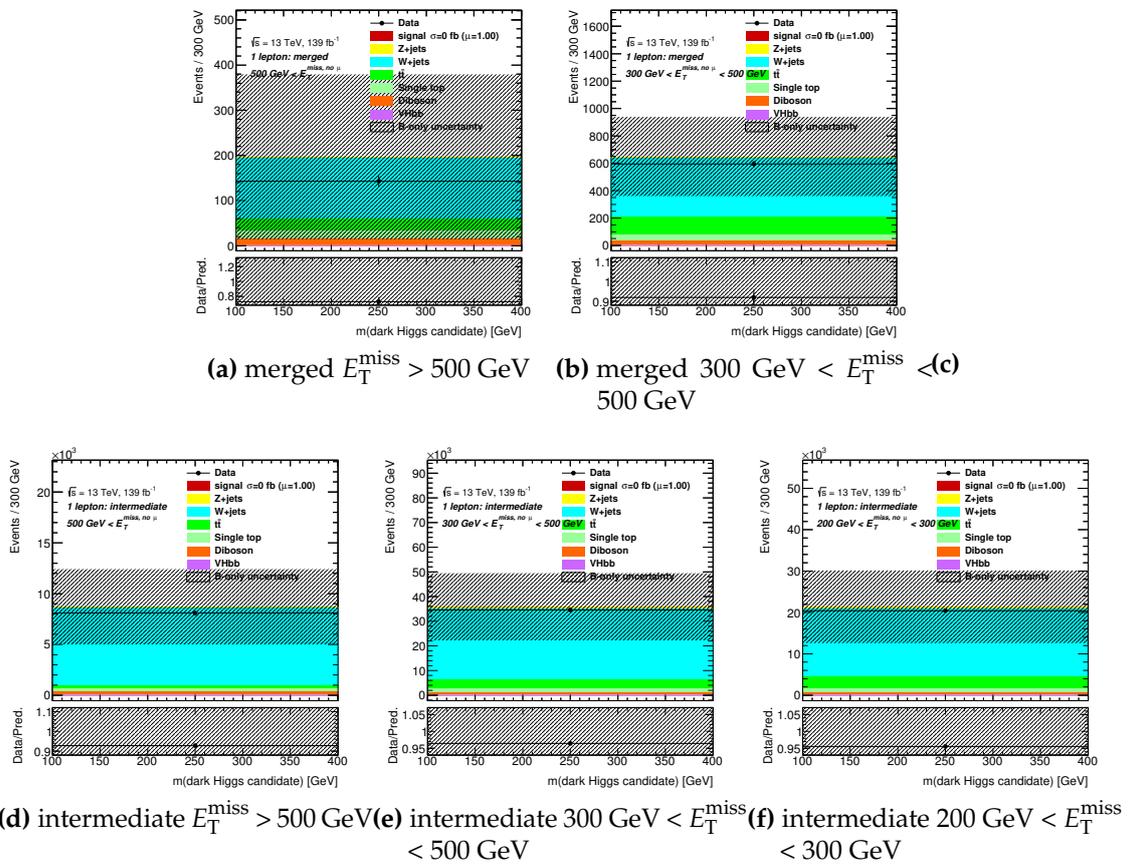
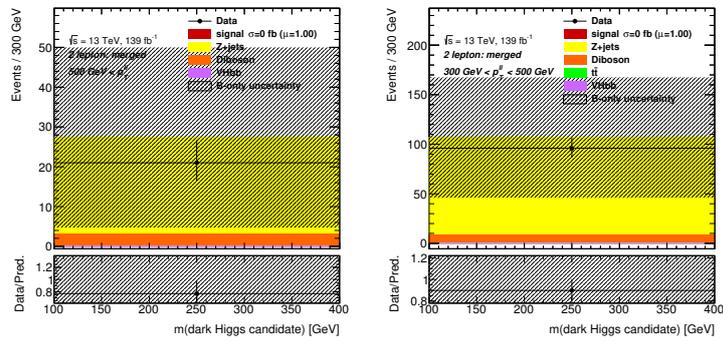
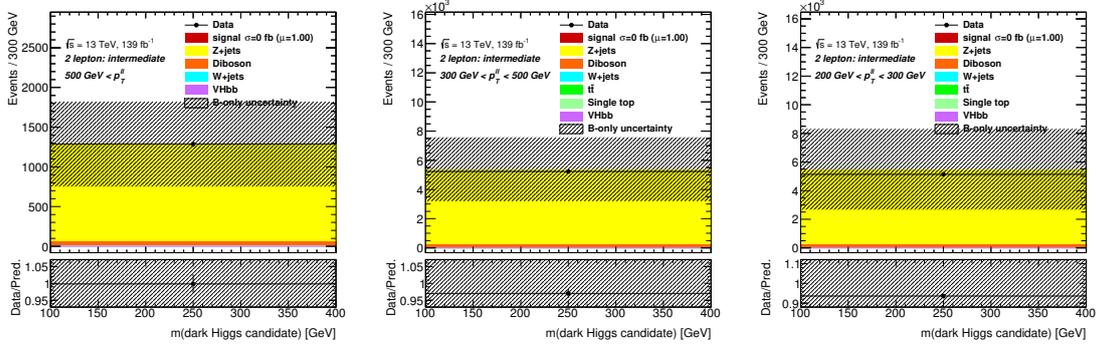


Figure D.2.: Distributions showing the background model after the conditional $\mu = 0$ fit and overlaid data in the 1 lepton control region (post-fit). The hatched band indicates the post-fit uncertainty on the total background prediction. The dotted-line indicates the theory prediction (pre-fit) for the total background.



(a) merged $E_T^{\text{miss}} > 500 \text{ GeV}$ (b) merged $300 \text{ GeV} < E_T^{\text{miss}} < 500 \text{ GeV}$



(d) intermediate $E_T^{\text{miss}} > 500 \text{ GeV}$ (e) intermediate $300 \text{ GeV} < E_T^{\text{miss}} < 500 \text{ GeV}$ (f) intermediate $200 \text{ GeV} < E_T^{\text{miss}} < 300 \text{ GeV}$

Figure D.3.: Distributions showing the background model after the conditional $\mu = 0$ fit and overlaid data in the 2 lepton control region (post-fit). The hatched band indicates the post-fit uncertainty on the total background prediction. The dotted-line indicates the theory prediction (pre-fit) for the total background.

E. Introduction to QFT Lagrangians

E.0.1. Introduction: QED

In the Lagrangian formulation, the fundamental law of motion is given by the Euler-Lagrange equation, which can be generalized for the fields as

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \right) = \frac{\partial \mathcal{L}}{\partial \phi_i}$$

If the Euler-Lagrange equation are applied to the Klein-Gordon Lagrangian for a spin=0 field, i.e.

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2}m^2 \phi^2$$

This Lagrangian is important since it is used in the description of the Higgs field. The well known Klein-Gordon equation can be derived from the previous equation as:

$$\partial_\mu \partial^\mu \phi + m^2 \phi = 0$$

On the other hand, applying the Euler-Lagrange equation (on the field $\bar{\psi}$) to the Dirac Lagrangian for spinor field ψ (and its adjoint $\bar{\psi}$):

$$\mathcal{L} = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi$$

From the relation $\frac{\partial \mathcal{L}}{\partial(\partial_\mu \bar{\psi})} = 0$ and the relation $\frac{\partial \mathcal{L}}{\partial \bar{\psi}} = i\gamma^\mu \partial_\mu \psi - m\psi$ the Dirac equation describing quantum fields of 1/2 spin particles with mass m can be derived:

$$i\gamma^\mu \partial_\mu \psi - m\psi = 0$$

Similarly, the Dirac equation for the antiparticle associated to the field $\bar{\psi}$ can be derived applying the Euler-Lagrange to the field ψ instead.

Finally, repeating the same procedure to the Proca Lagrangian for a vector field of spin 1, A^μ :

$$\mathcal{L} = -\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \frac{1}{8\pi}m^2A^\nu A_\nu$$

where for simplicity $F^{\mu\nu} = (\partial^\mu A^\nu - \partial^\nu A^\mu)$. Then, applying the Euler-Lagrange equations:

$$\partial_\mu F^{\mu\nu} + m^2 A^\nu = 0$$

which is the Proca equation describing a field associated to a particle of spin 1 and mass m . Setting $m = 0$, the Maxwell equations in the vacuum can be recovered.

The Dirac and Proca Lagrangians are 'automatically' combined enforcing a gauge symmetry. The Dirac Lagrangian is invariant under constant phase transformation: $\psi \rightarrow \psi' = e^{i\alpha}\psi$, $\bar{\psi} \rightarrow \bar{\psi}' = e^{-i\alpha}\bar{\psi}$. However, if the phase α is not constant any more but a function of the four momentum $\alpha(x)$, then it can be verified that the Dirac Lagrangian is not invariant any more. Therefore, a transformation D_μ such that $D_\mu\psi \rightarrow e^{i\alpha(x)}D_\mu\psi$ is looked for, hoping that this way the extra terms would be taken care of. This can be realized introducing a spin 1 gauge field A_μ , which transforms as $A_\mu \rightarrow A_\mu - \frac{1}{e}\partial_\mu\alpha(x)$ (Remembering that this transformation leaves the Maxwell equations unchanged; moreover, it makes the Proca Lagrangian $m = 0$ invariant). With this, the *covariant* derivative $D_\mu = (\partial_\mu + ieA_\mu)\psi$ can be defined. The rationale behind this is that this definition allows the Proca and Dirac Lagrangians together to be invariant, sometimes in fact the terms *compensating* is used instead of gauge, reflecting the fact that this field is compensating or restoring the symmetry. Plugging the covariant derivative into the Dirac Lagrangian, the following equation is found:

$$\begin{aligned}\mathcal{L} &= i\bar{\psi}\gamma^\mu D_\mu\psi - m\bar{\psi}\psi \\ &= i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - (e\bar{\psi}\gamma^\mu\psi)A_\mu\end{aligned}\tag{E.1}$$

Where the last term now describes the coupling of the fermions from the original Dirac Lagrangian with the spin 1 field introduced, with coupling constant e . Now the massless Proca Lagrangian can be added (the mass term would break the invariance). The Lagrangian then reads:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} - (e\bar{\psi}\gamma^\mu\psi)A_\mu$$

which is the well known Lagrangian of the quantum electrodynamics, simply obtained forcing the Dirac Lagrangian to be invariant under local transformation (and by local it is meant that $\alpha = \alpha(x)$). The transformations $\psi \rightarrow \psi' = e^{i\alpha}\psi$ can be in fact thought as $\psi \rightarrow \psi' = U\psi$, where U is an unitary matrix ($U^\dagger U = 1$). In other words, this is a $U(1)$ gauge transformation.

E.0.2. Electroweak Theory

The electroweak sector of the SM can be formalized following the Yang-Mills theories: in order to achieve the chiral V-A (vector minus axial vector) interactions that was needed to described processes where parity (P) was violated¹, another symmetry has to be used: the $SU(2)$.

Assuming to have two spin 1/2 fields, ψ_1 and ψ_2 , the free Lagrangian can be more compactly written using $\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$ ($\bar{\psi} = (\bar{\psi}_1 \bar{\psi}_2)$):

$$\begin{aligned} \mathcal{L} &= [i\bar{\psi}_1 \gamma^\mu D_\mu \psi_1 - m_1 \bar{\psi}_1 \psi_1] + [i\bar{\psi}_2 \gamma^\mu D_\mu \psi_2 - m_2 \bar{\psi}_2 \psi_2] \\ &= i\bar{\psi} \gamma^\mu D_\mu \psi - m \bar{\psi} \psi \end{aligned} \quad (\text{E.2})$$

assuming that $m_1 = m_2$, which looks exactly as the Dirac Lagrangian. However, this new Lagrangian admits new global and local symmetries. These can be shown using the transformation $\psi \rightarrow \tilde{U}\psi$ ($\bar{\psi} \rightarrow \bar{\psi}\tilde{U}^\dagger$), where \tilde{U} is a 2×2 unitary matrix ($\tilde{U}^\dagger \tilde{U} = 1$, $\tilde{U} \in U(2)$). \tilde{U} can be conveniently rewritten as $\tilde{U} = e^{iH}$ where H is Hermitian ($H = H^\dagger$), which can be expressed in the form $H = \theta 1 + \alpha^a \sigma^a$ where $a=1,2,3$; α^a and θ are real numbers and σ^a are the Pauli matrices. Therefore $\tilde{U} = e^{i\theta} e^{i\alpha^a \sigma^a}$. The first exponential can be recognized to be the same as in the QED case ($U(1)$), but now the second one, $e^{i\alpha^a \sigma^a}$ is responsible for a $SU(2)$ global transformation (effectively this factorizes the $U(2)$ in $U(1) \times SU(2)$).

Another key ingredient is the decomposition the fermion fields into left and right handed components. This can be achieved by means of the projection operator P_L and

¹This was a requirement since the experiment carried out by madame Wu in 1956 [121], finding that P was maximally violated in charged current processes and in the Goldhaber experiment 1957 [122] that proved that only left handed neutrinos participate to the weak interactions.

P_R :

$$\psi = \left(\frac{1 - \gamma_5}{2} + \frac{1 + \gamma_5}{2} \right) \psi = (P_L + P_R) \psi = \psi_L + \psi_R$$

The left and right handed component divide the Lagrangian for massless fermions:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu D_\mu \psi = i(\bar{\psi}_L + \bar{\psi}_R)\gamma^\mu D_\mu (\psi_L + \psi_R) = i\bar{\psi}_L\gamma^\mu D_\mu \psi_L + i\bar{\psi}_R\gamma^\mu D_\mu \psi_R$$

here using the anticommutation relation $\{\gamma^5, \gamma^\mu\} = 0$ and the fact that $P_L P_R \psi = 0$. This effectively decouples the Lagrangian into right handed and left handed component: $\mathcal{L} = \mathcal{L}_L + \mathcal{L}_R$ for the kinetic term. This is desirable, since the aim was to provide a description in the Lagrangian that is different for left and right handed fermions. Following this prescription, the left handed components can be set to be doublets (e.g. $\ell_L^i = \begin{pmatrix} e \\ \nu_e \end{pmatrix}_L$ for the leptons and $Q_L^i = \begin{pmatrix} u \\ d \end{pmatrix}_L$) and the right handed ones to be singlets (e.g. ℓ_R^i for leptons and u_R^i, d_R^i for quarks). The mass terms, however, look like $\bar{\psi}\psi = (\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L)$ and there is no way to take care of them, since they will break the local gauge invariance. This forces the Lagrangian up to this point to be massless; however, masses can be generated via the Higgs mechanism.

Again following the examples of the QED and repeating the same procedure of going to global to local symmetry $\alpha^a = \alpha^a(x)$, expanding the exponential keeping only the first-order terms, the Lagrangian is not invariant any more. This is fixed, again in analogy with the QED, introducing two gauge fields, B_μ and W_μ^a (or in short $W_\mu = W_\mu^a \sigma^a / 2$, again $a = 1..3$). Their free Lagrangian can be for B_μ the same as in QED, $\frac{1}{4} B_{\mu\nu} B^{\mu\nu}$ (with a slight abuse of notation $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$), for W_μ this is not allowed any more. In fact, W_μ transforms as $W_\mu \rightarrow \tilde{U} W_\mu \tilde{U}^\dagger - i\tilde{U} \partial_\mu \tilde{U}^\dagger$, forces the tensor to be

$$W_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu + ig_2 [W_\mu, W_\nu]$$

where $[W_\mu, W_\nu]$ can also be written in terms of the structure constants of the $SU(2)$ group (in this case they are the Levi-Civita tensors), and g_2 a constant. However, this time also $W_{\mu\nu} W^{\mu\nu}$ will not be adapt since not yet invariant; what can be done instead is considering $\text{Tr}(W_{\mu\nu} W^{\mu\nu})$, where Tr indicates the trace.

Bibliography

- [1] B. Malaescu, F. Napolitano, M. F. Daneri, G. Otero y Garzon, J. D. Bossio Sola, L. Henkelmann, M. Toscani, O. Brandt, R. Piegai, M. R. Devesa, Y. W. Y. Ng, and J. Veatch, “In situ calibrations for R-scan jets,” Tech. Rep. ATL-COM-PHYS-2018-407, CERN, Geneva, Apr 2018.
- [2] “Search for Dark Matter Produced in Association with a Higgs Boson decaying to $b\bar{b}$ at $\sqrt{s} = 13$ TeV with the ATLAS Detector using 79.8 fb^{-1} of proton-proton collision data,” Tech. Rep. ATLAS-CONF-2018-039, CERN, Geneva, Jul 2018.
- [3] ATLAS Collaboration, “Search for dark matter in events with a hadronically decaying vector boson and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector.” ATLAS-CONF-2018-005, 2018.
- [4] “RECAST framework reinterpretation of an ATLAS Dark Matter Search constraining a model of a dark Higgs boson decaying to two b -quarks,” Tech. Rep. ATL-PHYS-PUB-2019-032, CERN, Geneva, Aug 2019.
- [5] C. N. Yang and R. L. Mills, “Conservation of isotopic spin and isotopic gauge invariance,” *Phys. Rev.*, vol. 96, pp. 191–195, Oct 1954.
- [6] P. W. Higgs, “Broken symmetries and the masses of gauge bosons,” *Phys. Rev. Lett.*, vol. 13, pp. 508–509, Oct 1964.
- [7] F. Englert and R. Brout, “Broken symmetry and the mass of gauge vector mesons,” *Phys. Rev. Lett.*, vol. 13, pp. 321–323, Aug 1964.
- [8] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global conservation laws and massless particles,” *Phys. Rev. Lett.*, vol. 13, pp. 585–587, Nov 1964.
- [9] J. Goldstone, “Field theories with « superconductor » solutions,” *Il Nuovo Cimento (1955-1965)*, vol. 19, pp. 154–164, Jan 1961.
- [10] J. Goldstone, A. Salam, and S. Weinberg, “Broken symmetries,” *Phys. Rev.*, vol. 127, pp. 965–970, Aug 1962.
- [11] S. L. Glashow, “Partial-symmetries of weak interactions,” *Nuclear Physics*, vol. 22, pp. 579–588, Feb 1961.

- [12] S. Weinberg, "A model of leptons," *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, Nov 1967.
- [13] A. Salam, "Weak and Electromagnetic Interactions," *Conf. Proc.*, vol. C680519, pp. 367–377, 1968.
- [14] D. J. Gross and F. Wilczek, "Ultraviolet behavior of non-abelian gauge theories," *Phys. Rev. Lett.*, vol. 30, pp. 1343–1346, Jun 1973.
- [15] H. D. Politzer, "Reliable perturbative results for strong interactions?," *Phys. Rev. Lett.*, vol. 30, pp. 1346–1349, Jun 1973.
- [16] M. Gell-Mann, "A schematic model of baryons and mesons," *Physics Letters*, vol. 8, no. 3, pp. 214 – 215, 1964.
- [17] G. Zweig, "An SU_3 model for strong interaction symmetry and its breaking; Version 2," p. 80 p, Feb 1964. Version 1 is CERN preprint 8182/TH.401, Jan. 17, 1964.
- [18] M. Y. Han and Y. Nambu, "Three-triplet model with double $SU(3)$ symmetry," *Phys. Rev.*, vol. 139, pp. B1006–B1010, Aug 1965.
- [19] M. e. a. Tanabashi, "Review of particle physics," *Phys. Rev. D*, vol. 98, p. 030001, Aug 2018.
- [20] W. Commons, "File:standard model of elementary particles.svg — wikimedia commons, the free media repository." https://commons.wikimedia.org/w/index.php?title=File:Standard_Model_of_Elementary_Particles.svg&oldid=416441632, 2020. [Online; accessed 9-May-2020].
- [21] D. Griffiths, *Introduction to elementary particles*. 2008.
- [22] G. Costa and G. Fogli, "Symmetries and group theory in particle physics. An introduction to space-time and internal symmetries," *Lect. Notes Phys.*, vol. 823, pp. 1–288, 2012.
- [23] W. N. Cottingham and D. A. Greenwood, *An introduction to the standard model of particle physics*. Cambridge University Press, 2007.
- [24] J. Ellis, "Higgs Physics," pp. 117–168. 52 p, Dec 2013. 52 pages, 45 figures, Lectures presented at the ESHEP 2013 School of High-Energy Physics, to appear as part of the proceedings in a CERN Yellow Report.

- [25] G. Bertone and D. Hooper, "History of dark matter," *Reviews of Modern Physics*, vol. 90, Oct 2018.
- [26] F. Zwicky, "Die Rotverschiebung von extragalaktischen Nebeln," *Helvetica Physica Acta*, vol. 6, pp. 110–127, Jan 1933.
- [27] H. W. Babcock, "The rotation of the Andromeda Nebula," *Lick Observatory Bulletin*, vol. 498, pp. 41–51, Jan 1939.
- [28] Q. Guo, H. Hu, Z. Zheng, S. Liao, W. Du, S. Mao, L. Jiang, J. Wang, Y. Peng, L. Gao, J. Wang, and H. Wu, "Further evidence for a population of dark-matter-deficient dwarf galaxies," *Nature Astronomy*, 2019.
- [29] T. S. van Albada, J. N. Bahcall, K. Begeman, and R. Sancisi, "Distribution of dark matter in the spiral galaxy NGC 3198.," *apj*, vol. 295, pp. 305–313, Aug 1985.
- [30] X. ray: NASA/CXC/CfA/M.Markevitch et al.; Optical: NASA/STScI; Magellan/U.Arizona/D.Clowe et al.; Lensing Map: NASA/STScI; ESO WFI; Magellan/U.Arizona/D.Clowe et al.
- [31] G. Gamow, "The origin of elements and the separation of galaxies," *Phys. Rev.*, vol. 74, pp. 505–506, Aug 1948.
- [32] G. GAMOW, "The evolution of the universe," *Nature*, vol. 162, no. 4122, pp. 680–682, 1948.
- [33] R. A. Alpher and R. C. Herman, "On the relative abundance of the elements," *Phys. Rev.*, vol. 74, pp. 1737–1742, Dec 1948.
- [34] A. A. Penzias and R. W. Wilson, "A Measurement of Excess Antenna Temperature at 4080 Mc/s.," *apj*, vol. 142, pp. 419–421, Jul 1965.
- [35] Planck Collaboration and A. et al., "Planck 2015 results - xiii. cosmological parameters," *A&A*, vol. 594, p. A13, 2016.
- [36] D. Hooper, "Tasi 2008 lectures on dark matter," 2009.
- [37] B. Penning, "The pursuit of dark matter at colliders—an overview," *J. Phys.*, vol. G45, no. 6, p. 063001, 2018.
- [38] M. C. Smith, G. R. Ruchti, A. Helmi, R. F. G. Wyse, J. P. Fulbright, K. C. Freeman, J. F. Navarro, G. M. Seabroke, M. Steinmetz, M. Williams, O. Bienaymé, J. Binney, J. Bland-Hawthorn, W. Dehnen, B. K. Gibson, G. Gilmore, E. K. Grebel, U. Mu-

- nari, Q. A. Parker, R.-D. Scholz, A. Siebert, F. G. Watson, and T. Zwitter, "The RAVE survey: constraining the local Galactic escape speed," *Monthly Notices of the Royal Astronomical Society*, vol. 379, pp. 755–772, 07 2007.
- [39] R. Bernabei, P. Belli, A. Bussolotti, F. Cappella, R. Cerulli, C. Dai, A. d' Angelo, H. He, A. Incicchitti, H. Kuang, and et al., "The dama/libra apparatus," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 592, p. 297–315, Jul 2008.
- [40] C. E. Aalseth, P. S. Barbeau, J. Colaresi, J. I. Collar, J. D. Leon, J. E. Fast, N. E. Fields, T. W. Hossbach, A. Knecht, M. S. Kos, M. G. Marino, H. S. Miley, M. L. Miller, J. L. Orrell, and K. M. Yocum, "Search for an annual modulation in three years of cogent dark matter detector data," 2014.
- [41] S. K. Liu, Q. Yue, K. J. Kang, J. P. Cheng, H. T. Wong, Y. J. Li, S. T. Lin, J. P. Chang, N. Chen, Q. H. Chen, Y. H. Chen, Y. C. Chuang, Z. Deng, Q. Du, H. Gong, X. Q. Hao, H. J. He, Q. J. He, H. X. Huang, T. R. Huang, H. Jiang, H. B. Li, J. M. Li, J. Li, J. Li, X. Li, X. Q. Li, X. Y. Li, Y. L. Li, H. Y. Liao, F. K. Lin, L. C. Lü, H. Ma, S. J. Mao, J. Q. Qin, J. Ren, J. Ren, X. C. Ruan, M. B. Shen, L. Singh, M. K. Singh, A. K. Soma, J. Su, C. J. Tang, C. H. Tseng, J. M. Wang, L. Wang, Q. Wang, S. Y. Wu, Y. C. Wu, Y. C. Wu, Z. Z. Xianyu, R. Q. Xiao, H. Y. Xing, F. Z. Xu, Y. Xu, X. J. Xu, T. Xue, C. W. Yang, L. T. Yang, S. W. Yang, N. Yi, C. X. Yu, H. Yu, X. Z. Yu, X. H. Zeng, Z. Zeng, L. Zhang, Y. H. Zhang, M. G. Zhao, W. Zhao, Z. Y. Zhou, J. J. Zhu, W. B. Zhu, X. Z. Zhu, and Z. H. Zhu, "Limits on light wimps with a germanium detector at 177 eVee threshold at the china Jinping underground laboratory," *Phys. Rev. D*, vol. 90, p. 032003, Aug 2014.
- [42] R. Agnese, A. J. Anderson, T. Aramaki, M. Asai, W. Baker, D. Balakishiyeva, D. Barker, R. Basu Thakur, D. A. Bauer, J. Billard, A. Borgland, M. A. Bowles, P. L. Brink, R. Bunker, B. Cabrera, D. O. Caldwell, R. Calkins, D. G. Cerdeno, H. Chagani, Y. Chen, J. Cooley, B. Cornell, P. Cushman, M. Daal, P. C. F. Di Stefano, T. Doughty, L. Esteban, S. Fallows, E. Figueroa-Feliciano, M. Ghaith, G. L. Godfrey, S. R. Golwala, J. Hall, H. R. Harris, T. Hofer, D. Holmgren, L. Hsu, M. E. Huber, D. Jardin, A. Jastram, O. Kamaev, B. Kara, M. H. Kelsey, A. Kennedy, A. Leder, B. Loer, E. Lopez Asamar, P. Lukens, R. Mahapatra, V. Mandic, N. Mast, N. Mirabolfathi, R. A. Moffatt, J. D. Morales Mendoza, S. M. Oser, K. Page, W. A. Page, R. Partridge, M. Pepin, A. Phipps, K. Prasad, M. Pyle, H. Qiu, W. Rau, P. Redl, A. Reissetter, Y. Ricci, A. Roberts, H. E. Rogers, T. Saab, B. Sadoulet,

- J. Sander, K. Schneck, R. W. Schnee, S. Scorza, B. Serfass, B. Shank, D. Speller, D. Toback, R. Underwood, S. Upadhyayula, A. N. Villano, B. Welliver, J. S. Wilson, D. H. Wright, S. Yellin, J. J. Yen, B. A. Young, and J. Zhang, "New results from the search for low-mass weakly interacting massive particles with the cdms low ionization threshold experiment," *Phys. Rev. Lett.*, vol. 116, p. 071301, Feb 2016.
- [43] G. Angloher, A. Bento, C. Bucci, L. Canonica, A. Erb, F. von Feilitzsch, N. F. Iachellini, P. Gorla, A. Gütlein, D. Hauff, P. Huff, J. Jochum, M. Kiefer, C. Kister, H. Kluck, H. Kraus, J. C. Lanfranchi, J. Loebell, A. Münster, F. Petricca, W. Potzel, F. Pröbst, F. Reindl, S. Roth, K. Rottler, C. Sailer, K. Schäffner, J. Schieck, J. Schmalzer, S. Scholl, S. Schönert, W. Seidel, M. von Sivers, L. Stodolsky, C. Strandhagen, R. Strauss, A. Tanzke, M. Uffinger, A. Ulrich, I. Usherov, S. Wawoczny, M. Willers, M. Wüstrich, and A. Zöller, "Results on low mass wimps using an upgraded cress-ti detector," *The European Physical Journal C*, vol. 74, no. 12, p. 3184, 2014.
- [44] E. Armengaud, Q. Arnaud, C. Augier, A. Benoit, A. Benoit, L. Bergé, T. Bergmann, J. Bläzemer, A. Broniatowski, V. Brudanin, P. Camus, A. Cazes, B. Censier, M. Chapellier, F. Charlieux, F. Couã«do, P. Coulter, G. Cox, T. de Boissière, M. D. Jesus, Y. Dolgorouky, A. Drillien, L. Dumoulin, K. Eitel, D. Filosofov, N. Fourches, J. Gascon, G. Gerbier, M. Gros, L. Hehn, S. Henry, S. Hervé, G. Heuermann, N. Holtzer, V. Humbert, A. Juillard, C. Kéfélian, M. Kleifges, H. Kluck, V. Kozlov, H. Kraus, V. Kudryavtsev, H. L. Sueur, M. Mancuso, C. Marrache-Kikuchi, S. Marnieros, A. Menshikov, X.-F. Navick, C. Nones, E. Olivieri, P. Pari, B. Paul, M. Piro, O. Rigaut, M. Robinson, S. Rozov, V. Sanglard, B. Schmidt, B. Siebenborn, D. Tcherniakhovski, M. Tenconi, L. Vagneron, R. Walker, M. Weber, E. Yakushev, and X. Zhang, "Axion searches with the EDELWEISS-II experiment," *Journal of Cosmology and Astroparticle Physics*, vol. 2013, pp. 067–067, nov 2013.
- [45] E. Aprile, J. Aalbers, F. Agostini, M. Alfonsi, F. D. Amaro, M. Anthony, B. Antunes, F. Arneodo, M. Balata, and et al., "The xenon1t dark matter experiment," *The European Physical Journal C*, vol. 77, Dec 2017.
- [46] M. e. a. Ackermann, "Constraining dark matter models from a combined analysis of milky way satellites with the fermi large area telescope," *Phys. Rev. Lett.*, vol. 107, p. 241302, Dec 2011.

- [47] H. e. a. Abdallah, "Search for dark matter annihilations towards the inner galactic halo from 10 years of observations with h.e.s.s.," *Phys. Rev. Lett.*, vol. 117, p. 111301, Sep 2016.
- [48] "First result from the alpha magnetic spectrometer on the international space station: Precision measurement of the positron fraction in primary cosmic rays of 0.5–350 gev," *Phys. Rev. Lett.*, vol. 110, p. 141102, Apr 2013.
- [49] F. Halzen and S. R. Klein, "Invited review article: Icecube: An instrument for neutrino astronomy," *Review of Scientific Instruments*, vol. 81, no. 8, p. 081101, 2010.
- [50] M. Duerr, A. Grohsjean, F. Kahlhoefer, B. Penning, K. Schmidt-Hoberg, and C. Schwanenberger, "Hunting the dark higgs," *Journal of High Energy Physics*, vol. 2017, Apr 2017.
- [51] M. Autran, K. Bauer, T. Lin, and D. Whiteson, "Searches for dark matter in events with a resonance and missing transverse energy," *Phys. Rev. D*, vol. 92, p. 035007, Aug 2015.
- [52] M. Pospelov, A. Ritz, and M. Voloshin, "Secluded wimp dark matter," *Physics Letters B*, vol. 662, no. 1, pp. 53 – 61, 2008.
- [53] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *Journal of High Energy Physics*, vol. 2014, no. 7, p. 79, 2014.
- [54] O. S. Bruning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*. CERN Yellow Reports: Monographs, Geneva: CERN, 2004.
- [55] E. Mobs, "The CERN accelerator complex. Complexe des accélérateurs du CERN," Jul 2016. General Photo <https://cds.cern.ch/record/2197559>.
- [56] C. T. Department, *The LHC Underground Facilities*. Geneva: CERN, 2015. <http://te-epc-lpc.web.cern.ch/te-epc-lpc/machines/lhc/pagesources/LHC-Underground-Layout.png> consulted on the 11.09.2019.
- [57] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider," *JINST*, vol. 3, p. S08003, 2008.

- [58] J. Pequenaó, "Computer generated image of the whole ATLAS detector." Mar 2008.
- [59] J. Pequenaó, "Computer generated image of the ATLAS inner detector." Mar 2008.
- [60] J. Pequenaó, "Computer Generated image of the ATLAS calorimeter." Mar 2008.
- [61] M. Aharrouché *et al.*, "Energy linearity and resolution of the ATLAS electromagnetic barrel calorimeter in an electron test-beam," *Nucl. Instrum. Meth.*, vol. A568, pp. 601–623, 2006.
- [62] W. Buttinger, "The ATLAS Level-1 Trigger System," Tech. Rep. ATL-DAQ-PROC-2012-024, CERN, Geneva, Jun 2012.
- [63] R. Achenbach *et al.*, "The ATLAS level-1 calorimeter trigger," tech. rep., 2008.
- [64] G. Aad *et al.*, "Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s} = 8$ TeV," *Eur. Phys. J.*, vol. C75, p. 120, 2015.
- [65] A. Aukerman and T. M. Hong, "Commissioning and validation of the ATLAS Level-1 topological trigger," in *Proceedings, 16th International Conference on Accelerator and Large Experimental Physics Control Systems (ICALEPCS 2017): Barcelona, Spain, October 8-13, 2017*, p. TUPHA070, 2018.
- [66] M. Aaboud *et al.*, "Performance of the ATLAS Trigger System in 2015," *Eur. Phys. J.*, vol. C77, no. 5, p. 317, 2017.
- [67] ATLAS Collaboration, "Performance of the ATLAS trigger system in 2015," *Eur. Phys. J. C*, vol. 77, p. 317, 2017.
- [68] "Presentation at the l1calo joint meeting [internal]." https://indico.cern.ch/event/638444/contributions/2639269/attachments/1490807/2317180/Timing_update_2017.pdf. Accessed: 2020-05-04.
- [69] "Performance of the upgraded PreProcessor of the ATLAS Level-1 Calorimeter Trigger," Tech. Rep. arXiv:2005.04179, CERN, Geneva, May 2020. 49 pages in total, author list starting page 33, 19 figures, 1 table, submitted to JINST. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/TDAQ-2019-01>.

- [70] S. Schätzel and M. Spannowsky, “Tagging highly boosted top quarks,” *Physical Review D*, vol. 89, Jan 2014.
- [71] Z. Nagy and D. E. Soper, “What is a parton shower?,” *Physical Review D*, vol. 98, Jul 2018.
- [72] A. Ali and G. Kramer, “Jets and qcd: a historical review of the discovery of the quark and gluon jets and its impact on qcd,” *The European Physical Journal H*, vol. 36, p. 245–326, Sep 2011.
- [73] ATLAS Collaboration, “Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1,” *Eur. Phys. J. C*, vol. 77, p. 490, 2017.
- [74] G. P. Salam, “Towards jetography,” *The European Physical Journal C*, vol. 67, p. 637–686, May 2010.
- [75] G. Sterman and S. Weinberg, “Jets from quantum chromodynamics,” *Phys. Rev. Lett.*, vol. 39, pp. 1436–1439, Dec 1977.
- [76] J. E. Huth *et al.*, “Toward a standardization of jet definitions,” in *1990 DPF Summer Study on High-energy Physics: Research Directions for the Decade (Snowmass 90) Snowmass, Colorado, June 25-July 13, 1990*, pp. 0134–136, 1990.
- [77] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm,” *JHEP*, vol. 04, p. 063, 2008.
- [78] ATLAS Collaboration, “Jet reconstruction and performance using particle flow with the ATLAS Detector,” *Eur. Phys. J. C*, vol. 77, p. 466, 2017.
- [79] ATLAS Collaboration, “Improving jet substructure performance in ATLAS using Track-CaloClusters.” ATL-PHYS-PUB-2017-015, 2017.
- [80] ATLAS Collaboration, “Performance of jet substructure techniques for large- R jets in proton–proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector,” *JHEP*, vol. 09, p. 076, 2013.
- [81] J. Thaler and K. Van Tilburg, “Identifying boosted objects with n-subjettiness,” *Journal of High Energy Physics*, vol. 2011, Mar 2011.
- [82] R. Kogler, B. Nachman, A. Schmidt, L. Asquith, E. Winkels, M. Campanelli, C. Delitzsch, P. Harris, A. Hinemann, D. Kar, and et al., “Jet substructure at the large hadron collider,” *Reviews of Modern Physics*, vol. 91, Dec 2019.

- [83] A. J. Larkoski, I. Moult, and B. Nachman, “Jet substructure at the large hadron collider: A review of recent advances in theory and machine learning,” *Physics Reports*, Nov 2019.
- [84] A. J. Larkoski, G. P. Salam, and J. Thaler, “Energy correlation functions for jet substructure,” *Journal of High Energy Physics*, vol. 2013, Jun 2013.
- [85] ATLAS Collaboration, “Jet mass reconstruction with the ATLAS Detector in early Run 2 data.” ATLAS-CONF-2016-035, 2016.
- [86] M. Son, C. Spethmann, and B. Tweedie, “Diboson-jets and the search for resonant zh production,” *Journal of High Energy Physics*, vol. 2012, Aug 2012.
- [87] A. Katz, M. Son, and B. Tweedie, “Jet substructure and the search for neutral spin-one resonances in electroweak boson channels,” *Journal of High Energy Physics*, vol. 2011, Mar 2011.
- [88] ATLAS Collaboration, “Track assisted techniques for jet substructure.” ATL-PHYS-PUB-2018-012, 2018.
- [89] B. Nachman, P. Nef, A. Schwartzman, M. Swiatlowski, and C. Wanotayaroj, “Jets from jets: re-clustering as a tool for large radius jet reconstruction and grooming at the lh,” *Journal of High Energy Physics*, vol. 2015, Feb 2015.
- [90] W. Lukas, “Fast Simulation for ATLAS: Atlfast-II and ISF,” Tech. Rep. ATL-SOFT-PROC-2012-065, CERN, Geneva, Jun 2012.
- [91] ATLAS Collaboration, “The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim.” ATL-PHYS-PUB-2010-013, 2010.
- [92] ATLAS Collaboration, “Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,” *Phys. Rev. D*, vol. 96, p. 072002, 2017.
- [93] ATLAS Collaboration, “Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV.” ATL-PHYS-PUB-2015-015, 2015.
- [94] A. Collaboration, “Jet energy resolution in 2017 data and simulation, 5 september 2018.” https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2018-005/fig_04.png. Accessed: 2020-02-20.

- [95] “Presentation at the atlas in-situ jet calibration meeting [internal].” <https://indico.cern.ch/event/874082/contributions/3791034/attachments/2006942/3352020/In-situ-20-03.pdf>. Accessed: 2020-05-04.
- [96] ATLAS Collaboration, “The ATLAS Simulation Infrastructure,” *Eur. Phys. J. C*, vol. 70, p. 823, 2010.
- [97] ATLAS Collaboration, “Search for Dark Matter Produced in Association with a Higgs Boson Decaying to $b\bar{b}$ using 36 fb^{-1} of pp collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS Detector,” *Phys. Rev. Lett.*, vol. 119, p. 181804, 2017.
- [98] ATLAS Collaboration, “Search for dark matter in events with a hadronically decaying vector boson and missing transverse momentum in pp collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector,” *JHEP*, vol. 10, p. 180, 2018.
- [99] ATLAS Collaboration, “Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach.” ATL-PHYS-PUB-2017-022, 2017.
- [100] ATLAS Collaboration, “Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13 \text{ TeV}$,” *Eur. Phys. J.*, 2019.
- [101] ATLAS Collaboration, “Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton–proton collision data,” *JINST*, vol. 14, p. P03017, 2019.
- [102] ATLAS Collaboration, “Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13 \text{ TeV}$,” *Eur. Phys. J. C*, vol. 76, p. 292, 2016.
- [103] ATLAS Collaboration, “Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13 \text{ TeV}$.” ATL-CONF-2017-029, 2017.
- [104] ATLAS Collaboration, “The Optimization of ATLAS Track Reconstruction in Dense Environments.” ATL-PHYS-PUB-2015-006, 2015.
- [105] ATLAS Collaboration, “Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2,” *Eur. Phys. J. C*, vol. 77, p. 673, 2017.

- [106] ATLAS Collaboration, "Forward Jet Vertex Tagging: A new technique for the identification and rejection of forward pileup jets." ATL-PHYS-PUB-2015-034, 2015.
- [107] ATLAS Collaboration, "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV," *Eur. Phys. J. C*, vol. 78, p. 903, 2018.
- [108] ATLAS Collaboration, "Object-based missing transverse momentum significance in the ATLAS Detector." ATLAS-CONF-2018-038, 2018.
- [109] D. Krohn, J. Thaler, and L.-T. Wang, "Jets with variable r ," *Journal of High Energy Physics*, vol. 2009, p. 059–059, Jun 2009.
- [110] ATLAS Collaboration, "Variable Radius, Exclusive- k_T , and Center-of-Mass Subjet Reconstruction for Higgs($\rightarrow b\bar{b}$) Tagging in ATLAS." ATL-PHYS-PUB-2017-010, 2017.
- [111] ATLAS Collaboration, "Measurements of b -jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV," *JHEP*, vol. 08, p. 089, 2018.
- [112] A. Collaboration, "ATLAS b -jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV," *Eur. Phys. J. C*, vol. 79, p. 970. 36 p, Jul 2019. 52 pages in total, author list starting page 36, 12 figures, 6 tables, submitted to EPJC. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/FTAG-2018-01/>.
- [113] ATLAS Collaboration, "Optimisation of the ATLAS b -tagging performance for the 2016 LHC Run." ATL-PHYS-PUB-2016-012, 2016.
- [114] G. Cowan, *Statistical data analysis*. Oxford University Press, USA, 1998.
- [115] L. Lyons, *Statistics for Nuclear and Particle Physicists*. Cambridge University Press, 1986.
- [116] W. Eadie, D. Drijard, F. James, M. Roos, and B. Sadoulet, "Statistical methods in experimental physics," *Journal of the American Statistical Association*, 07 2013.
- [117] S. S. Wilks, "The large-sample distribution of the likelihood ratio for testing composite hypotheses," *Ann. Math. Statist.*, vol. 9, pp. 60–62, 03 1938.

- [118] A. Wald, "Tests of statistical hypotheses concerning several parameters when the number of observations is large," *Transactions of the American Mathematical Society*, vol. 54, no. 3, pp. 426–482, 1943.
- [119] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics," *The European Physical Journal C*, vol. 71, Feb 2011.
- [120] "Summary plots from the atlas exotic physics group." <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/EXOTICS/>, 2020. [Online; accessed 9-May-2020].
- [121] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, "Experimental test of parity conservation in beta decay," *Phys. Rev.*, vol. 105, pp. 1413–1415, Feb 1957.
- [122] M. Goldhaber, L. Grodzins, and A. W. Sunyar, "Helicity of neutrinos," *Phys. Rev.*, vol. 109, pp. 1015–1017, Feb 1958.

List of figures

2.1.	Representation of all the known particle and anti-particles, with the fermions organized in their respective generations, and the bosons associated to the fundamental forces plus the Higgs boson are shown. Each particle has the mass, charge and spin indicated in the top-left. From [20].	9
2.2.	Representation of the Higgs potential, for the parameter choice $\lambda > 0$ and $\mu < 0$. From [24].	13
3.1.	The galactic rotation curve of NGC 3198, from the 1985 measurement [29]. In separate components, the contribution to the observed values from the galactic disk and the Dark Matter halo are shown.	20
3.2.	The <i>bullet cluster</i> as shown in different superimposed components: the pink clumps in the image are the hot gas distribution in the as measured by the Chandra-X in the X-rays spectrum. The one on the right shows the bullet shape, which passed through the left one. The optical image is taken from the Magellan and Hubble Space Telescope. The blue areas represent the mass of the two systems as measured from micro-lensing. The hot gas was slowed down during the impact by drag force, surpassed by the Dark Matter component [30].	22
3.3.	The temperature power spectrum of spherical harmonics as measured by <i>Planck 2015</i> [35].	24
3.4.	Schematic of the evolution of the number density that undergoes the process of the Freeze-out [36]. As the temperature decreases (or m/T increases) due to expansion of the universe, the density is frozen at a certain value according to the interaction rate. Higher interaction rates make DM annihilation more likely, decreasing its density.	27

3.5.	The three main experimental lines to address the experimental search for Dark Matter [37]. From top to bottom, the DM state scatters off SM state that is then detected: this is the direct detection. From right to left, two DM states annihilate into SM states, which are then detected, in what is called the indirect detection. Finally, from left to right, the production at colliders, where two SM states annihilate into DM. What is looked for in this case is missing momentum in the event, as DM states fly away undetected right after production.	28
3.6.	Feynman diagram of the Dark Higgs-Strahlung. s is the Dark Higgs, Z' and χ are the gauge boson and the Dark Matter particle, respectively.	32
3.7.	Branching ratio of the Dark Higgs into SM states: b-quarks, W^+W^- , ZZ and HH as calculated from MADGRAPH5_aMC@NLO 2.6.6 at leading order [53], from [4].	33
3.8.	The exclusion contour for the Dark Higgs model with benchmark parameters $m_\chi = 200$ GeV, $g_q = 0.25$ and $g_\chi = 1.0$, as a function of $m_{Z'}$ (x-axis) and m_s (y-axis). The solid line represents the observed limit. The dashed line represents the expected limit; green and yellow band show the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty on the expected limit respectively. The pink dotted line indicates the parameter points, for which the observed relic density is reproduced [4].	34
3.9.	On the left, the most contributing Feynman diagram for the W^+W^- final state: the Dark Higgs from the Z' that subsequently decays to Dark Matter. On the right, the second most contributing diagram: the Dark Higgs is radiated off a Dark Matter particle.	34
4.1.	Schematic of CERN's accelerator complex, with the experiments served. Taken from [55].	40
4.2.	The LHC underground layout [56].	42
4.3.	Overview of the ATLAS detector, with its sub-components indicated on the picture. The detector is 25 m high and 44 m long, weighting approximately 7000 tonnes. From [58].	44
4.4.	Computer generated image of the ATLAS inner detector [59].	46

4.5.	Computer generated image of the ATLAS calorimeters [60].	48
4.6.	Structure of the EMB calorimeter with the granularities in η - ϕ dimensions shown [57].	50
4.7.	Tiles, fibres and photomultipliers composing the optical readout of a tile calorimeter module [57].	52
4.8.	The ATLAS Trigger and Data Acquisition (TDAQ) system in Run 2 [67]. The Fast Tracker was decommissioned and will not be part of the TDAQ in future.	55
5.1.	Illustration of the infrared and collinear safety that a jet clustering algorithm should have. The jets are shown as cones and their components as arrows. In the top left diagram, two different jets were identified. In case of non infrared safety, the algorithm merges the two jets if a soft component is added between the two (top right diagram). On the bottom left, a jet is identified from three components. In case of non collinear safety of the jet algorithm, a collinear splitting of the component in the middle causes the jet axis to shift and skip the left most component (bottom right).	60
5.2.	The anti- k_t (left) and k_t (right) jet clustering algorithms, shown in the rapidity azimuth plane with p_T on the z-axis for the same event. The anti- k_t exhibits circular shapes of its active area. In case of overlap between a hard and a softer jet, the harder one will extend its border into the softer one as much as R . If two neighbouring jets are at the same p_T scale, none of them will prevail and the border will be in-between. The k_t algorithm clusters with preference to softer components, and results in asymmetrical shapes. Hard jets could be clustered with a smaller area with respect to softer ones. From [77].	62

- 5.3. Both pictures show the angular separation between their decay products as a function of the p_T of the decaying particle, in a simulated sample with energetic di-tops [80]. On the left (a) the ΔR between the W boson and the b quark from the top quark decay; the hyperbolic behaviour follows the $\sim 1/p_T$ law. While a Small- R jet is not sufficient to contain both W and b , a Large- R with $R = 1.0$ assures the containment for $p_T^{top} > 300$ GeV. On the right (b), the ΔR between the quarks from the W bosons decay. If $p_T^W > 300$ GeV, the ability to resolve the quarks individually with Small- R jets starts degrading quickly, but on the other hand a Large- R jet can contain both of them. 65
- 5.4. The flow of the trimming algorithm [80]. 66
- 5.5. The mass distribution of Large- R jets in high p_T regimes ($600 \text{ GeV} < p_T < 800 \text{ GeV}$) before trimming (dashed) and after trimming (continuous line), from boosted Z (red) and QCD di-jet background (black) [80]. The trimmed jets from boosted Z loose a small fraction of their mass, but they see an increase in the resolution. Trimming removes the majority of the energy from the QCD jet, achieving a strong suppression of the background at the higher end of the mass spectrum. 67
- 5.6. $\eta - \phi$ map of Large- R jets constituents originated from QCD (a), W bosons (b) and Top quarks decays (c) [81]. Constituents of the same subjet are indicated by the same colour, the open square indicates the total jet direction, the open circles indicate the two subjet directions, and the crosses indicate the three subjet's directions. 68
- 5.7. The performance of the reconstruction of W/Z jets using the m^{TA} , m^{calo} and their combination m^{comb} (the standard in ATLAS) is compared. On x-axis, the p_T of the MC truth jet and on y-axis the interquartile range (IQnR) divided by twice the median of the response distribution is shown. The IQnR is a robust way to access the resolution. It can be seen that the Track-Assisted mass achieves better resolutions at higher p_T 's and that the combination m^{comb} achieves an even superior performance. From [85]. 72

5.8. Description of the TAR jet algorithm. The $R=0.2$ jets are clustered to form a Large- R jet that is then trimmed. Tracks are subsequently associated (both with ghost association and ΔR matching) to the $R=0.2$ jets surviving the trimming procedure. The momentum of the tracks is then rescaled per $R=0.2$ jet in the so-called 'Track-Assisted procedure'. The Small- R jets can be as well used instead of $R=0.2$ jets, although they do not resolve the subjets as well. 75

5.9. The resolution of the mass observable for W jets (left) and for the $D_2^{\beta=1}$ substructure observable (right) as a function of the MC truth p_T . The TAR algorithm is shown in beige. The m^{comb} is shown on the left in magenta and an alternative (and older) definition of TAR, TAS (Track-Assisted Subjet) is shown in green. On the right, the substructure observable D2 for TAR and TAS is compared against calorimeter-based substructure only, since there exists no combined version as for the mass. The Figure is based on the author's work in [88]. 76

5.10. ROC curves comparing the separation power of a 68% efficient mass cut plus a substructure observable for the identification of four-prong decays, τ_{42} , against QCD jets for $250 \text{ GeV} < p_T^{jet} < 600 \text{ GeV}$. On the x-axis the efficiency of the identification of the pair of W from Higgs decay, and on the y-axis the efficiency of the background rejection. Again TAR (TAS) in beige (green) outperforms standard techniques, in blue. The Figure is based on the author's work in [88]. 77

5.11. The calibration sequence for jets at the EM scale [92]. 78

5.12. An example of the absolute MC-based calibration on the left, showing the jet energy response as a function of the jet η for different values of the jet p_T in different markers, before the correction is applied. On the right, a stage of the Global Sequential Calibration using the number of tracks n_{trk} (x-axis) and the response (y-axis), for three different p_T bins in different markers, before the correction is applied. The bottom plot shows the relative fraction per p_T bin [92]. 79

- 5.13. Overview of the composition of the JES systematic uncertainty in early Run 2, as a function of p_T . On the left, the total JES uncertainty as calculated from the γ +jets balance, where the main components are the MC generator, which accounts for potential mis-modelling, and out-of-cone effects. In the central plot it is shown the overview of the collective impact of the in-situ (γ +jets measurement depicted with cyan pluses, which dominates at higher p_T 's), pile-up and flavour uncertainties, which dominates at the low end of the spectrum. Finally, on the right, the plot showing the combined effect of the in-situ uncertainties, together with the flavour composition and response, pile-up uncertainties and punch-through [92]. 80
- 5.14. The breakdown of the systematic uncertainties components on the fractional jet energy resolution as a function of p_T [94]. The black line indicates the total uncertainty. The green line shows the impact of the noise term component, which is dominant at the lower end of the spectrum. The blue component is derived from the in-situ di-jet direct balance method. 82
- 7.1. The expected sensitivity S , as calculated with the GL procedure from the $E_T^{\text{miss}} + H(b\bar{b})$ [97] and $E_T^{\text{miss}} + V(\text{had})$ [98] analyses (left), is depicted on the $(m_{Z'}, m_s)$ plane on the x-y axis and the S on the z-axis. The points with $S > 1$ are expected to be excluded in the parameter space. The grid of produced signal samples in the plane $(m_{Z'}, m_s)$, chosen on the basis of the GL estimate (right). Blue points are generated with hadronic-only decays of the W boson, which is the target of this analysis. Red points are generated instead with the full decay, for preliminary studies of the semi-leptonic analysis. 93
- 8.1. The E_T^{miss} (left) and TAR jet mass (right) for selected signal W^+W^- final state. Different mass hypotheses are shown, magnified as displayed in the label to be visible in the same y-axis range. 103

- 8.2. Ratio of significances between TAR jets ($R = 0.8$ constructed on R -scan jets) and standard large- R jet in the merged category for $m_{Z'} = 1$ TeV. The mass of the Dark Higgs is shown on the x -axis, and on y -axis the E_T^{miss} lower boundary. The significance is computed after the requirement that the candidate mass is within 50 GeV window centred around the Dark Higgs mass, and after the $\tau_{42}^{\text{TAR}} < 0.3$ requirement. Significance is defined as $S = \sqrt{2((s+b)\log(1+s/b) - s)}$ 105
- 8.3. Data and MC distributions in the inclusive E_T^{miss} selection for the Dark Higgs candidate mass as reconstructed in the merged and in the intermediate category, for the 0LSR, 1LCR and 2LCR. 110
- 11.1. Nuisance parameters pulls and constraints for the conditional $\mu = 0$ fit to Asimov data (top) and for the unconditional (μ free to float) fit to observed data using the full 139 fb^{-1} Run 2 dataset (bottom). The x -axis shows the different NP as defined in Chapter 9, and the first two items on the left are the free floating normalization of the W +jets and Z +jets processes. 126
- 11.2. NPs ranking plot based on the Asimov data set generated with $\mu = 1$ for different signal mass hypothesis (in caption). The filled boxes represent the $\sigma_{\hat{\mu}}/\hat{\mu}$ on pre-fit and post fit (scale on top x -axis). The pull of each NP is also overlaid (scale on bottom x -axis). The ranking is ordered based on the post-fit impact on the fitted value of μ , and are displayed from top (most important) downwards. 127
- 11.3. MC distributions for the background model and overlaid data in the 0 lepton signal region for theory prediction (post-fit). The bottom panel shows the relative uncertainty on the background prediction. The hatched band indicates the pre-fit uncertainty on the total background prediction. 128
- 11.4. Distributions showing the background model after the conditional $\mu = 0$ fit and overlaid data in the 1 lepton control region (post-fit). The hatched band indicates the post-fit uncertainty on the total background prediction. The dotted-line indicates the theory prediction (pre-fit) for the total background. 129

- 11.5. Distributions showing the background model after the conditional $\mu = 0$ fit and overlaid data in the 2 lepton control region (post-fit). The hatched band indicates the post-fit uncertainty on the total background prediction. The dotted-line indicates the theory prediction (pre-fit) for the total background. 130
- 11.6. On the left, the expected significance (z-axis) on the signal strength μ for Dark Higgs signals in the combined VV decay modes with fixed $m_\chi = 200$ GeV is shown. On the right, the observed significance is displayed for the same decay mode and m_χ . On x-axis, the mass of the Z' and on the y-axis the mass of the Dark Higgs boson can be read. 131
- 11.7. On the left, the expected limits (z-axis) on the signal strength μ for Dark Higgs signals in the combined VV decay modes with fixed $m_\chi = 200$ GeV are shown. On the right the, observed limits are displayed for the same decay mode and m_χ . On x-axis, the mass of the Z' and on the y-axis the mass of the Dark Higgs boson can be read. 131
- 11.8. 95% CL_s expected and observed limits on the signal strength μ as dashed and solid line, respectively, for the $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ analysis in the signal hypothesis with $m_{Z'} = 1$ TeV. The band around the expected limits shows the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties. The horizontal line denotes $\mu = 1$. Masses are excluded if the limits are below this line. 132
- 11.9. 95% CL_s expected and observed limits on the signal strength μ as dashed and solid line, respectively. The $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ analysis in the signal hypothesis with $m_{Z'} = 1$ TeV is shown in the range (160-360) GeV. The $E_T^{\text{miss}} + s \rightarrow b\bar{b}$ as reinterpreted from the $E_T^{\text{miss}} + H(b\bar{b})$ analysis is shown in the range (50-150) GeV. The band around the expected limits shows the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties. The horizontal line denotes $\mu = 1$. Masses are excluded if the limits are below this line. 133
- A.1. Summary of limits on the coupling between Z' boson and quarks for various Z' boson mass $m_{Z'}$ placed by ATLAS di-jet searches [120]. . . . 139
- A.2. Madgraph cross-section computed for the signal model in the parameter space. 142
- A.3. The Madgraph cross-section is then binned in E_T^{miss} 143

A.4. Bin-wise sensitivity multiplied by the $A * \epsilon / \sigma_{GL}$	143
A.5. Final reco-level sensitivity.	144
A.6. Dependency of the GL on the background in the Signal Region. The E_T^{miss} bins are highlighted in the plot and the a linear fit is shown in red line.	145
A.7. Final sensitivity for the two scenarios hypothesized for the background level in the Signal Region, the worst and the best case, left and right plot, respectively.	146
A.8. The representation of the flow of the TAR+Comb algorithm. First a TAR jet consistent with a V -jet is looked for; if one is found (left), the two prong from the V jets not contained in the TAR jet are added. If the TAR jet is above the V mass window (right), only one V prong is assumed to be missing; this is selected and added. If the TAR jet has a mass below the V window, the event is passed to the resolved category according to the priority based selection. In the η - ϕ maps, the yellow circle is the TAR jet, the green ones the R-scan jets. R' is radius in which the algorithm looks for the missing prong(s). This example is shown for the optima case of R -scan jets, while in the analysis Small- R jets are used instead.	148
A.9. Correlation matrix showing the largest correlations of nuisance parameters (excluding nuisance parameters associated to Poisson terms in the likelihood function) for conditional $\mu = 0$ fit to a Asimov data set generated with $\mu = 0$	149
A.10. Correlation matrix showing the largest correlations of nuisance parameters (excluding nuisance parameters associated to Poisson terms in the likelihood function) for conditional $\mu = 0$ fit to a Asimov data set generated with $\mu = 1$	150
A.11. Correlation matrix showing the largest correlations of nuisance parameters (excluding nuisance parameters associated to Poisson terms in the likelihood function) for unconditional fit to a Asimov data set generated with $\mu = 0$	150

- A.12. Correlation matrix showing the largest correlations of nuisance parameters (excluding nuisance parameters associated to Poisson terms in the likelihood function) for unconditional fit to a Asimov data set generated with $\mu = 1$ 151
- A.13. Correlation matrix showing the largest correlations of nuisance parameters (excluding nuisance parameters associated to Poisson terms in the likelihood function) for conditional $\mu = 0$ fit to the observed data. . . . 151
- A.14. Correlation matrix showing the largest correlations of nuisance parameters (excluding nuisance parameters associated to Poisson terms in the likelihood function) for unconditional fit to the observed data. 152
- A.15. Expected limit on the signal strength μ for Dark Higgs signals in the WW decay mode with fixed $m_\chi = 200$ GeV, scanned in $m_{Z'}$ and m_s . The merged and intermediate categories are considered. 159
- A.16. Expected limit on the signal strength μ for Dark Higgs signals in the ZZ decay mode with fixed $m_\chi = 200$ GeV, scanned in $m_{Z'}$ and m_s . The merged and intermediate categories are considered. 160
- A.17. Expected limit on the signal strength μ for Dark Higgs signals in the combined WW and ZZ decay modes with fixed $m_\chi = 200$ GeV, scanned in $m_{Z'}$ and m_s . Only the merged category is considered. 160
- A.18. Expected limit on the signal strength μ for Dark Higgs signals in the combined WW and ZZ decay modes with fixed $m_\chi = 200$ GeV, scanned in $m_{Z'}$ and m_s . Only the intermediate category is considered. 161
- A.19. Expected limit on the signal strength μ for Dark Higgs signals in the combined WW and ZZ decay modes with fixed $m_\chi = 200$ GeV, scanned in $m_{Z'}$ and m_s . Only the resolved category is considered. 161
- A.20. 95% CL_s expected and observed limits on the signal strength μ as dashed and solid line, respectively, for the $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ analysis in the signal hypothesis with $m_{Z'} = 1.7$ TeV. The band around the expected limits shows the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties. 162

A.21. 95% CL_s expected and observed limits on the signal strength μ as dashed and solid line, respectively, for the $E_T^{\text{miss}} + s \rightarrow VV(\text{had})$ analysis in the signal hypothesis with $m_{Z'} = 0.5$ TeV. The band around the expected limits shows the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties.	163
B.1. MC distributions in 0 lepton signal region for the merged category with $E_T^{\text{miss}} > 500\text{GeV}$	166
B.2. MC distributions in 0 lepton signal region for the merged category with $300\text{GeV} < E_T^{\text{miss}} < 500\text{GeV}$	167
B.3. MC distributions in 0 lepton signal region for the intermediate category with $E_T^{\text{miss}} > 500\text{GeV}$	168
B.4. MC distributions in 0 lepton signal region for the intermediate category with $300\text{GeV} < E_T^{\text{miss}} < 500\text{GeV}$	169
B.5. MC distributions in 0 lepton signal region for the intermediate category with $200\text{GeV} < E_T^{\text{miss}} < 300\text{GeV}$	170
B.6. MC distributions in 0 lepton signal region for resolved category with $E_T^{\text{miss}} > 500\text{GeV}$	172
B.7. MC distributions in 0 lepton signal region for resolved category with $300\text{GeV} < E_T^{\text{miss}} < 500\text{GeV}$	173
B.8. MC distributions in 0 lepton signal region for resolved category with $200\text{GeV} < E_T^{\text{miss}} < 300\text{GeV}$	174
B.9. Data / MC comparison in 1 μ control region for the merged category with $E_T^{\text{miss,no}\mu} > 500\text{GeV}$	176
B.10. Data / MC comparison in 1 μ control region for the merged category with $300\text{GeV} < E_T^{\text{miss,no}\mu} < 500\text{GeV}$	177
B.11. Data / MC comparison in 1 μ control region for the intermediate category with $E_T^{\text{miss,no}\mu} > 500\text{GeV}$	179
B.12. Data / MC comparison in 1 μ control region for the intermediate category with $300\text{GeV} < E_T^{\text{miss,no}\mu} < 500\text{GeV}$	180

B.13. Data / MC comparison in 1 μ control region for the intermediate category with $200\text{GeV} < E_T^{\text{miss,no}\mu} < 300\text{GeV}$	181
B.14. Data / MC comparison in 1 μ control region for resolved category with $E_T^{\text{miss,no}\mu} > 500\text{GeV}$	182
B.15. Data / MC comparison in 1 μ control region for resolved category with $300\text{GeV} < E_T^{\text{miss,no}\mu} < 500\text{GeV}$	183
B.16. Data / MC comparison in 1 μ control region for resolved category with $200\text{GeV} < E_T^{\text{miss,no}\mu} < 300\text{GeV}$	184
B.17. Data / MC comparison in 2 ℓ control region for the merged category with $p_T^{\ell\ell} > 500\text{GeV}$	186
B.18. Data / MC comparison in 2 ℓ control region for the merged category with $300\text{GeV} < p_T^{\ell\ell} < 500\text{GeV}$	187
B.19. Data / MC comparison in 2 ℓ control region for the intermediate category with $p_T^{\ell\ell} > 500\text{GeV}$	189
B.20. Data / MC comparison in 2 ℓ control region for the intermediate category with $300\text{GeV} < p_T^{\ell\ell} < 500\text{GeV}$	190
B.21. Data / MC comparison in 2 ℓ control region for the intermediate category with $200\text{GeV} < p_T^{\ell\ell} < 300\text{GeV}$	191
B.22. Data / MC comparison in 2 ℓ control region for resolved category with $p_T^{\ell\ell} > 500\text{GeV}$	192
B.23. Data / MC comparison in 2 ℓ control region for resolved category with $300\text{GeV} < p_T^{\ell\ell} < 500\text{GeV}$	193
B.24. Data / MC comparison in 2 ℓ control region for resolved category with $200\text{GeV} < p_T^{\ell\ell} < 300\text{GeV}$	194
D.1. MC distributions for the background model and overlaid data in the 0 lepton signal region for theory prediction (pre-fit). The bottom panel shows the relative uncertainty on the background prediction. The hatched band indicates the pre-fit uncertainty on the total background prediction.	201

-
- D.2. Distributions showing the background model after the conditional $\mu = 0$ fit and overlaid data in the 1 lepton control region (post-fit). The hatched band indicates the post-fit uncertainty on the total background prediction. The dotted-line indicates the theory prediction (pre-fit) for the total background. 202
- D.3. Distributions showing the background model after the conditional $\mu = 0$ fit and overlaid data in the 2 lepton control region (post-fit). The hatched band indicates the post-fit uncertainty on the total background prediction. The dotted-line indicates the theory prediction (pre-fit) for the total background. 203

List of tables

7.1.	Summary of the simulated background processes. The V +jets cross-section refers to the NNLO normalization. For the Z +jets process this is calculated for $66 \text{ GeV} < m_{ll} < 116 \text{ GeV}$. The alternative W/Z +jets sample are used to evaluate the theoretical modelling uncertainties. The $t\bar{t}$ cross-section is normalized at NNLO+NNLL, using a top quark mass of $m_t=172.5 \text{ GeV}$. The di-boson process is treated differently in the leptonic, semi-leptonic and loop-induced, in the accuracy of the additional partons. Similarly, single-top is generated and treated differently in tW associated production, single top s- and t-channel production.	90
8.1.	The baseline selection that is required in all the signal and control regions.	102
8.2.	A summary of the different reconstruction algorithms used in each of the prioritized signal categories, explained in detail in this Section. The final kinematic selection requirements of the reconstructed Dark Higgs candidate are also listed.	108
8.3.	The $E_T^{\text{miss}} / E_T^{\text{miss,no}\mu} / p_T^{\ell\ell}$ bins for the different regions 0LSR / 1LCR / 2LCR for the merged and for the intermediate categories are shown. The ranking is merged (M) \rightarrow intermediate (I).	109
9.1.	Qualitative summary of the experimental systematic uncertainties that are considered in this analysis. The name of the systematic uncertainties and their description is shown. The smoothing is used to mitigate the impact that limited MC statistics has on the fit.	113
9.2.	Summary of the evaluation strategy for the theoretical uncertainties. The strategy is shown for W/Z +jets, other background processes and for the signal. The shape uncertainties are derived at reconstruction level, the acceptance with the simplified analysis using truth level Monte Carlo.	116

9.3. Qualitative summary of the theoretical systematic uncertainties that are considered in this analysis. The name of the systematic uncertainties and their description is shown. The smoothing is used to mitigate the impact that limited MC statistics has on the fit.	119
A.1. Summary table of triggers used in 2015, 2016, 2017, and 2018 data. . . .	141
A.2. Overlap removal	147
A.3. A summary of the three different reconstruction algorithms used in each of the three prioritized signal categories, with the resolved category which is detailed in this Appendix. The final kinematic selection requirements of the reconstructed Dark Higgs candidate are also listed.	154
A.4. The binning for the different regions 0LSR / 1LCR / 2LCR as a function of their $E_T^{\text{miss}} / E_T^{\text{miss,no}\mu} / p_T^{\ell\ell}$ bins for the merged, intermediate, and resolved categories. The same ranking is merged (M) \rightarrow intermediate (I) \rightarrow resolved (R).	154
A.5. Expected and observed number of events in the 0 lepton signal region with 139 fb^{-1} data, shown separately in each topology category. The background yields and uncertainties are shown before the profile-likelihood fit to the data. The quoted background uncertainties include both the statistical and systematic contributions. The total uncertainties can be smaller than for individual components due to anti-correlations.	155
A.6. Expected and observed number of events in the 1 lepton control region with 139 fb^{-1} data, shown separately in each topology category. The background yields and uncertainties are shown before the profile-likelihood fit to the data. The quoted background uncertainties include both the statistical and systematic contributions. The total uncertainties can be smaller than for individual components due to anti-correlations.	156
A.7. Expected and observed number of events in the 2 lepton control region with 139 fb^{-1} data, shown separately in each topology category. The background yields and uncertainties are shown before the profile-likelihood fit to the data. The quoted background uncertainties include both the statistical and systematic contributions. The total uncertainties can be smaller than for individual components due to anti-correlations.	156

- A.8. Expected and observed number of events in the 0 lepton signal region with 139 fb^{-1} data, shown separately in each topology category. The background yields and uncertainties are shown after the profile-likelihood fit to the data. The quoted background uncertainties include both the statistical and systematic contributions. The total uncertainties can be smaller than for individual components due to anti-correlations. 157
- A.9. Expected and observed number of events in the 1 lepton control region with 139 fb^{-1} data, shown separately in each topology category. The background yields and uncertainties are shown after the profile-likelihood fit to the data. The quoted background uncertainties include both the statistical and systematic contributions. The total uncertainties can be smaller than for individual components due to anti-correlations. 158
- A.10. Expected and observed number of events in the 2 lepton control region with 139 fb^{-1} data, shown separately in each topology category. The background yields and uncertainties are shown after the profile-likelihood fit to the data. The quoted background uncertainties include both the statistical and systematic contributions. The total uncertainties can be smaller than for individual components due to anti-correlations. 158
- A.11. Observed and expected limits on the Dark Higgs scenario with $m_{Z'} = 0.5 \text{ TeV}$ 164
- A.12. Observed and expected limits on the Dark Higgs scenario with $m_{Z'} = 1 \text{ TeV}$ 164
- A.13. Observed and expected limits on the Dark Higgs scenario with $m_{Z'} = 1.7 \text{ TeV}$ 164
- C.1. Generator-related acceptance uncertainties for the W +jets SM process in the signal region for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column. . 195
- C.2. Generator-related acceptance uncertainties for the Z +jets SM process in the signal region for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column. 196

- C.3. PDF acceptance uncertainties for the W +jets SM process in the signal region for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column. 196
- C.4. PDF acceptance uncertainties for the Z +jets SM process in the signal region for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column. 197
- C.5. Scale acceptance uncertainties for the W +jets SM process in the signal region for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column. 197
- C.6. Scale acceptance uncertainties for the Z +jets SM process in the signal region for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column. 198
- C.7. PDF acceptance uncertainties for a mono- $s(W W)$ signal with $m_{Z'} = 1$ TeV, $m_\chi = 200$ GeV, $m_s = 160$ GeV for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column. 198
- C.8. PDF acceptance uncertainties for a mono- $s(Z Z)$ signal with $m_{Z'} = 1$ TeV, $m_\chi = 200$ GeV, $m_s = 210$ GeV for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column. 198
- C.9. Scale acceptance uncertainties for a mono- $s(W W)$ signal with $m_{Z'} = 1$ TeV, $m_\chi = 200$ GeV, $m_s = 160$ GeV for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column. 199
- C.10. Scale acceptance uncertainties for a mono- $s(Z Z)$ signal with $m_{Z'} = 1$ TeV, $m_\chi = 200$ GeV, $m_s = 210$ GeV for inclusive selection in E_T^{miss} in the first column, and for individual E_T^{miss} bins in the second, third and fourth column. 199